UNIVERSITY OF CALIFORNIA
SANTA CRUZ

TELLTALE HEARTS:
ENCODING CINEMATIC CHOICE-BASED ADVENTURE GAMES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER SCIENCE

by

John T. Murray

September 2018

The Dissertation of John T. Murray
is approved:

Professor Noah Wardrip-Fruin, Chair

Professor Michael Mateas

Professor Katherine Isbister

Lori Kletzer
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Telltale Hearts:

Encoding Cinematic Choice-based Adventure Games

by

John T. Murray

Dramatic choices are the core of the content in contemporary adventure games like those produced by Telltale Games and Dontnod Entertainment, developers of the /Life is Strange/ series. These consist of emotionally-charged and dramatically contextualized player choices rather than strategic or skill-based challenges. Current approaches to understanding and assessing player experience in these games include surveys, interviews, think-aloud and various biometric measurements. These measures, taken alone, are useful for understanding traditional gameplay challenges such as platformers or puzzles but fall short of assessing how emotional content influences player experience or leads to variations in player responses. I collaboratively conducted a study of six players playing the first episode of *The Wolf Among Us* (*TWAU*) by Telltale Games, recording gameplay and facial expressions along with heart rate and skin conductivity. To compare and analyze player responses and classify content features, I developed an encoding, the Interactive Cinematic Experience Surface schemata, and used it to annotate player traversals with feature locations and player choices. These align player response data to content for comparison and analysis. Researchers can use the model to compare multiple paths through the game that can span hours of gameplay. In order to understand the dataset better, I developed a web-based visual data mining tool, SHERLOCK, for annotating and analyzing player experience datasets using the schemata and assessed it with a group of game researchers. I compare the

encoding method to two other encodings using content from $TWAU$: Elson's Story Intention Graphs and Mawhorter's choice poetics. Analyzing the first half of the first episode revealed a spectrum of variations in responses to content, supporting the presence and effectiveness of emotional content anchors and the characteristics of the Telltale effect. Thematic content that depended on choices differed in the values portrayed from content that was always included. The approach of using multiple player responses, especially facial expression analytics, is essential to a medium read of experiences and is a promising direction for interactive narrative research using existing commercial storygame content.

For My Parents

Acknowledgments

There is not enough room even in a document of this size to express my gratitude for the people that have made this journey possible. It has been a long and rewarding one mostly due to the people I have me along the way.

I am humbled by the mentoring, guidance and support that my advisors, Noah Wardrip-Fruin and Michael Mateas have provided me. Their work has set the agenda for the field and I'm very grateful to have had the opportunity to be a part of it. I am particularly grateful for Noah Wardrip-Fruin's relentless insights and maddening ability to see directly through obfuscation and misdirection to the heart of any issue. I will miss leaving his office feeling ready to take on the world. I thank Katherine Isbister, who made time and inspired me to do the same for others. I am grateful for the other members of the faculty whom I have worked with and learned from over the years at UCSC, including Arnav Jhala, Marilyn Walker, James Davis, and many others. I am particularly grateful for the chance to have worked with Jim Whitehead and for the many opportunities he provided to teach and conduct research. I'm grateful also David Elson for his guidance as a member of my advancement committee, and for his inspirational research in the area of story modeling.

One of the greatest opportunities of being a graduate student was to be part of the Expressive Intelligence Studio, a group whose work and alumni have made a mark on the field. While a member there, I met both lifelong friends as well as collaborators who I hope to continue knowing for the rest of my life, including Jacob Garbe, Aaron Reed, Kate Compton, Johnathan Pagnutti, Stacey Mason, James Ryan, Peter Mawhorter, Anne Sullivan, Josh McCoy, Jo Mazeika, Gillian Smith, Joe Osborn, Ben Samuel, Chris Lewis, Adam Smith, Chaim Gingold, April Grow, Eric Kaltman, Mike Treanor, James Skorupski, and many others. I also thank my colleagues at Seebright for supporting my goals of pursuing both a startup and a graduate degree simultaneously, including Dirk Kanngiesser, Bill Keating, Doug Michels, Robert Rossi, Ezra Goldman, and Ron Lachman, as well as the various contributors to the company over the years and its goal of creating an affordable augmented reality headset.

I have the privilege of collaborating with a number of talented scholars and artists who have shaped my work. Among these, I am most grateful to my partner and collaborator, Anastasia Salter. I also am grateful to Raquel Robinson and Joellyn Rock for their patience and willingness to take chances as my collaborators. I am particularly grateful for the support and friendship of Mark Marino, whose work in critical code studies, netprov and electronic literature continually inspire me. I am honored by the support and friendship of other members of the electronic literature community, including my former advisor Matthew Kirschenbaum, as well as Nick Montfort, Caitlin Fisher, Dene Grigar, Deena Larsen, Mark Sample, Rob Wittig, Leonardo Flores, Bill Bly, Stuart Moulthrop, Stephanie Strickland, Alex Mitchell, Hartmut Koenitz, Marjorie Luesebrink, Stephanie Boluk, Fox Harrell, Jason Nelson and many others.

I also thank people who played an important role in my life, but are not as much in it these past few years, including Heather Marusiak, Cristin Moore, David Cardigan, Jack O'Sullivan, Maya Encarnación, Daniel Notley, Florence Cadigan, and others. I cannot thank enough my close friends Chris Salter and Maraget Salter for their support and attempts at project management during the toughest times of this dissertation.

Finally, I thank my family, Wayne Murray, Mary and Mark Moranville, Jennifer and Marc Bialek. Their support over the years and unwavering faith that I'd find the right path gave me the time to pursue my dreams.

Chapter 3 contains substantially the same content as my master's thesis, and is included for context and its relationship to the other contributions.

Chapter 1

Introduction

My original goal was to create a novel, generative, and interactive digital storytelling system that leveraged computational reasoning and knowledge representations. I have now had the chance to play with and observe several story-focused systems, each emphasizing different aspects of the total experience. Some, like parser-based interactive fiction (IF), emphasize writing content that relates to an underlying model of a world, while others focus on the generation of specific content that invokes responses such as surprise or suspense. I learned that authors of computational media must balance these two forces: the possibility space of content that any given individual (player, reader) experiences and the desired experience from having engaged with the material. These "target" experiences are numerous but include an increased understanding of the human condition, an emotional experience relating to a particular character, or a realization of the nature of some system. These are just a few of the possible goals, of course, but each relies on a translation step that authors make in etching an experience into a given media: they anticipate a reception and craft the content to elicit that response. This realization led me to begin investigating how to reveal this translation's presence in the existing content of contemporary storygames and map it onto the variety of emotional responses elicited. This dissertation directly builds on the observation that, embedded in existing storygames, there are a variety of creative, imaginative, and effective solutions for authoring procedural content to elicit emotional responses. Understanding and mapping these patterns provides a path to both improve the theory used to evaluate storygame content and create a basis for future computational support for authors.

In the last two decades, there has been a renaissance in research into the human experience using computational methods, in part due to the rapid increase in the capabilities of computers, the development of efficient algorithms, and the availability of sensors [1]. Approaches taken by researchers range from directly measuring and modeling players [2] to the creation of tools and instruments that measure games and record gameplay (also called play traces or traversals, with the latter term used in this dissertation) [3], [4], [5]. Storygames present a unique challenge in this respect. Their **content** is dissimilar enough to other games that many of the methods used to evaluate platformers and action games, such as the measures of difficulty and pacing employed in Tanagra [6], do not address the layer of meaning encoded and expressed in language and gesture performances. Concepts such as *fun*, *challenge*, and *flow* are often used to distinguish games from other types of narrative media and so do not naturally map onto moments that aim to evoke sadness or remorse in players. The game research community recognizes this challenge. Reed, for instance, describes the dynamics at the core of the tension between stories and games in his dissertation [7], and my thesis benefits from his insights into the overall tensions in the genre. More generally, there have been many efforts to formalize and consolidate into a coherent field the study of the combination of story and gaming. Computational linguistics has provided an example that computational narratology has followed [8], while others have borrowed techniques such as telemetry data from other genres of digital games [9]. While the fields have yet to coalesce under a single name, **Interactive Digital Narrative** (**IDN**) is one candidate label that has been used to describe both practice and theory, and it is the term used in the dissertation. The full field includes traditions ranging from the **interactive digital storytelling** (**IDS**) community to communities that prioritize generation [10] and modeling, either for authoring new stories [11] or for understanding player interactions and responses [12]. The field label of **IDN** reflects a focus on the narratives as artifacts themselves, while the term **IDS** emphasizes the process and action of conveying a story. The approach taken in this dissertation focuses more on the relationship of player responses to the storygame content structure itself. This focus recognizes that contemporary storygame content is designed to elicit responses, which in turn are conditioned and shaped by the content as well as individual variations in personality, preferences, background experiences, and emotional states.

Content in the choice-based cinematic adventure game (CCAG) subgenre is distinct from other types of game content. Specifically, the content found in this subgenre (currently) requires human-level interpretation in order to both experience (e.g., play) and author; therefore, the subgenre provides an opportunity to develop and test new models that describe the current limits of computational modeling. I call the subgenre **cinematic choice-based adventure games** not because it is the common name, but rather because the label highlights the dynamic of cinematically presented choices. The subgenre belongs to a broader category of storygames called branching narratives, which includes work created using the hypertext platform Twine [13] and interactive fiction platforms, both of which rely more heavily on text than cinematic conventions. However, Telltale Games "produces" its works in a way that has more in common with television than other genres, including a strong emphasis on scriptwriting, voice acting, and animation. The content lends itself to multiple types of processing and is well-suited for corpus-based approaches. The game displays player selections for choices alongside available options and presents content segments with consistent timing. Player responses can also be mapped to specific content without requiring the eye-tracking needed in more text-heavy forms of content. The limited variation is also a benefit over exploratory or emergent narratives, as it provides greater alignment between traversals of similar content. To study player responses to this subgenre, I collaborated with Raquel Robinson and collected a set of traversal records of a CCAG by Telltale Games, *The Wolf Among Us* (*TWAU*). I describe the dataset and initial annotations further in Chapter 3. An initial analysis of these traversals and the player response data associated with them suggests that a content analysis that incorporates a formal model provides more details than the simple tagging of the relevant content. As a result, I develop formal schemata (described in Chapter 4) to manage the complexities involved in the player and traversal relationships and to assist in identifying patterns in content and choice design. The schemata and the annotations of game traversals made with it may, in turn, advance the field of interactive digital narratives.

1.1 Interactive Digital Narratives

Games have incorporated some aspect of narratives, and many computational narrative experiences (such as hypertext) have incorporated game-like elements. This dissertation is concerned primarily with how the combination of games and stories in computational media can evoke emotions.

Developers have combined games and stories in a wide variety of ways over the years. It is useful to think of narrative as a cognitive capability in addition to the usual abilities associated with games, such as an understanding of physics, reaction times, memory, or logical reasoning. Narrative thinking can reflect an understanding of others' motivations and emotions; it can include making inferences based on context; and it involves our ability to remember past events, find meaning, and imagine future outcomes. The study of narrative has shifted from an approach that emphasizes narratives' status as an artifact (as "stories" or "text" that can be studied apart from their reception) to an approach that instead welcomes insights into the mental capabilities needed to both create and understand narratives. This shift allows for a more interdisciplinary approach to storygame modeling, incorporating theories from psychology and cognitive science in addition to traditional concepts from game studies and classical and contemporary narratology. I review this shift in more detail in Section 2.2, but the two perspectives, artifact and mental phenomena, provide a useful lens into the treatment of IDN content by various storygame systems and approaches.

One can classify IDNs by how content is manipulated and selected[1]. Figure 1.1 classifies several prominent works in IDNs, including the social simulation platform *Versu* [15], the social physics puzzle game *Prom Week* [16], *Façade* [17], and the augmented reality IF game *Ice-Bound Concordance* [18]. It also includes the variety of games typical of several game producers: Failbetter Games, creators of the browser narrative game *Fallen London* [19]; Lucasarts, creator of classic adventure games; Telltale Games, creators of *The Wolf Among Us* [20]; and Sierra, a prolific developer of adventure games in the 1980s [21]. Given the often academic origins of works from the first two categories, there are many published details about how their architectures organize and present content.

When I use the term *model* in this context, I mean specifically a mapping re-

[1]The concept of Content Selection draws from conversations and correspondence with Michael Mateas, and the particular strategy and architecture used in *Façade* is one example of the types of models described [14].

Content Manipulation in IDN

Model Type	Mechanic/Method	Example
Explicit Model	Simulationist ↕ Logical	Versu Prom Week Facade
Implicit Model	Symbolic/ Combinatorial	Failbetter Games Ice-Bound Concordance
Mental Model	Decision Tree	Telltale Games LucasArts Sierra

Figure 1.1: Comparison of IDN content model strategies

lating the semantic meaning of the content and its computational representation, where such a model is used by both authors and systems to represent relationships present in a *target domain.* The target domain varies but could be real-world systems, such as social dynamics based on modeling relationships (*Prom Week* [22], Versu [23]) or story theories such as dramatic arcs and character affinities (Façade [14]). An **explicit model** is one where the underlying variables are mapped directly to a target variable, such as tension (in *Façade*), relationship strength (*Prom Week*), or utility (used to select actions based on desires in *Versu*). An **implicit model** is one where a computational algorithm or process is *analogous* to a desired expressive process but not necessarily a simulation of the values themselves. One example is a combination of symbols that produce meanings outside of the algorithm, such as in procedural poetry. Reed and Garbe's *The Ice-Bound Concordance* [18] invites players to manipulate glowing symbols to decide whether to reveal specific content and manage the variations in the ongoing story. One way to think of it is to think of this class as having desirable features for telling stories, but they are not directly representing some aspect of the simulation or of the presentation through a model that is defines as such.

For **mental model** storygames, the minds of the authors and players are responsible for interpreting the primary meaning of the content. This interpretation occurs in addition to rules such as collision detection, simulation conditions, or game logic. The meaning-producing assets (such as sound files and animations) in this category are not manipulated by a computational process except concerning ordering and rendering (e.g., the order of the content follows a graph structure). There are general aspects for how the mechanics and logics are related to their representation. These are better understood through the lens of operational logics (see [24], [25], and [26]), but they play a smaller role compared to the instantial content (hand-authored assets) [27] concerning the story itself. Game designers often use simple flags and states to represent the branching logic (the variations in content triggered by player input). The process does not consider the nature of the content being branched: A choice based on a life-or-death situation is the same as a minor choice around a conversational reply. A vital item in an adventure game is often represented in the same manner as an unimportant one, except that the vital item is a condition for progression. The system does not model the implications of the social and personal storyworlds of the events and states. In other words, the game makes selections or modifications to the presentation of the content (usually in the form of 3D models, animations, and audio files) without a value representing the various system dynamics observed by a player. I describe many of these interactive digital narrative efforts and their relationship to my work in more detail in Section 2.2, but in this introduction, I focus on the significance of this last class of storygames for descriptive modeling approaches.

Games in the CCAG subgenre are, as members of the **mental model** category, particularly good candidates for applying a descriptive model. The hand-selected content means that there are fewer, more controlled variations. The game presents the story content in a mostly fixed order.[2] This tight control of sequencing and the hand-authored content lead to the probability of simulation-introduced "bugs" in the story being low to non-existent. Wardrip-Fruin observes that games that attempt to use a more free-form structure, such as computer role-playing games, may make mistakes and reveal that the underlying logic is in fact based on flags and decision trees[3] [29]. The content flaws that are present in CCAGs do not result from some computational model that failed to model a real-world system accurately or perform some human-level task such as language understanding. These features are why I selected a specific Telltale Games production as the target for developing a descriptive model.

Telltale Games has set its work apart from other contemporary adventure game production studios through its use of film and television conventions such as shot composition, voice acting, and animation. A human must consider each alternative path and variation following a choice point in a Telltale Game for its impact and interpretation. This does not mean that the combination of possible traversal paths is not large, only that the variations are designed to make the experience consistent regardless of any particular set of choices in a traversal. Like traditional stories, the content is crafted to manipulate player responses, giving us the opportunity to identify structures, particularly in the combination

[2]Although chunks of content may have different orderings based on player choices at times, there are usually not more than a handful of different sequences. *TWAU* has two orders for the player visiting two scenes.

[3]Sullivan developed a solution that addressed the inflexibility of quests in her system, the Grail Framework [28].

of traditional content and choices, which is present in the (mental) models used by the author(s).

1.1.1 Telltale-Style Adventure Games

Adventure games hold a special place in the history of storygames and in the hearts of many game researchers. Salter chronicled the power of their content to retain strong communities willing to curate and recreate the games [21]. These games evoke player emotions associated with overcoming challenges, loving characters, and otherwise spending time with the stories and worlds. Aarseth describes ergodic media as consisting of challenges or gaps in interpretation (aporia) that are overcome through epiphany [30]. The challenges, however, are often hard. These take the form of puzzles or riddles [31]. Reed labels one design pattern found in adventure games as an *eurekon* [32] that requires the reader to connect their knowledge of objects and relationships from their background to the possibilities of how objects and ideas relate within a game world puzzle. This property inspired Aarseth to label such works as “ergodic literature” [33] or “ergodic discourse” [30].

Compared with previous adventure games, those produced by Telltale Games are far less challenging, resulting in the successful broadening of adventure game audiences. Besides making the challenges easier, Telltale Games also chooses not to incorporate advances in AI such as drama management into their core game engine or production process. Former Telltale CEO Kevin Bruner acknowledges in a keynote at the International Conference on Interactive Digital Storytelling (ICIDS) in 2016 that not every possible story is a good story and states that until the success rate of AI storytelling systems surpasses a hand-crafted approach, the studio will continue to produce their games using a more traditional production process. Their approach to combining interactivity and drama is one solution to the authorial burden challenges identified for the authoring highly dynamic AI storytelling systems [34], specifically intended to ensure a high probability of both satisfying traversals and good stories. Bruner also mentions that part of the inspiration behind the studio's approach is the experience of discussing episodes with coworkers on a regular basis. The episodic structure enables a similar mode of community engagement, where speculation and discussion around shared experiences are a key part of the value proposition. As part of this, the consistency of the content is important, even if the branching content will necessarily vary from player to player. With most players faced with the same decisions, the decisions themselves can become a valuable topic for community discussions. The founders of Telltale Games are acutely aware of the commercial viability of the genre of adventure games, and their efforts that led to the subgenre reflect an awareness of needing to appeal to a broader audience. This results in the key features of the subgenre, including the prominence of dramatically contextualized timed choices and the cinematic mode of story content, as well as an overall reduction of the challenges present in the games.

The "Telltale-style adventure game" [35] is very cinematic, to the point that the style has even been called interactive cinema. But not everyone sees Telltale's style in a positive light. Bycer, for instance, blames the genre for removing player agency: "Good player agency is about having the player make choices both large and small so while yes, not every choice should be a matter of life and death, they should still mean something to the character and the plot" [36]. Bycer's perspective has come up in many other reviews that focus in on the level of impact player choices have on the plot of the game and often the number of possible endings, another measure sometimes used to evaluate the complexity or variability of the content. Gerardi collects some discussions on the popular website

The AV Club, noting how choice and results play a role in an experience based on expectations [37].

Telltale Games has had broad critical acclaim for their productions, as described by Maltbie in a post defending Telltale-style games from critics who would debate whether they were games or not:

> My mind stayed engaged, I took responsibility for what happened, and I was absorbed by the idea that I need to **pay attention to details if I want the story to play out a certain way**. And that's an incredibly personal experience when you get down to it. And at the end of the chapter in the Telltale games, a menu shows how your choices differ from other players who are going through the game for the first time. If the experience is malleable like this, how can it not be a game? [38] (emphasis added).

The degree to which the "experience" is malleable is a topic of some disagreement, as many expect this changeability to include more variations in the plot. Reviewers and critics have accused Telltale Games of only providing stories that are in the genre of games that treat story content as "beads on a string," even if some recognize the value of the experience. Frauschauer, a German journalist, describes this as an illusion of choice:

> But one major complaint comes up again and again: Fast-paced choices and moral quandaries are the most prominent feature of *The Walking Dead*, but they seem to have no real consequences. Why did this game receive approximately all of the awards in 2012 if its main selling point is used so inconsequentially? I suggest that the most important aspect about this game is not the many difficult choices it offers the player––it's the illusion of choice the game constructs [. . .] The amount of different content that needs to be produced is limited to a minimum while, ideally, still giving you a feeling of agency [39]

On the contrary, one of the core factors in a player's experience is not the sense of meaningful control often associated with the term *agency* but, as I will argue in this dissertation, is instead that the choices they can make are designed to integrate player responses with the dramatic tensions. Ultimately, these choices give players an avenue for expressing their feelings towards events and characters rather than acting to change them decisively. This expression, in turn, can evoke a sense of emotional agency and even responsibility, a feeling that is one consequence of the content design patterns that give rise to what I propose as the **Telltale effect**, described more in Section 8.1. The effect is not merely an illusion of agency or the result of meaningless choices but rather an experience of emotional engagement with the story based on expressions and choices that reflect player identity and values.

It is useful to compare the effect to a different approach to IDNs, one that strives to provide players with a high degree of agency and that is a stated goal of many of the works described in Section 2.2. Ben Samuel, for instance, proposes the term "shared authorship" for the collaborative creation of narrative artifacts, emphasizing that it is important that "players have an immensely ontological impact" and that this excludes works such as CCAGs from falling under the moniker of "shared authorship". In this dissertation, I consider the ways in which player experience is structured both by surrounding content and the limited set of choices provided to players. My goal is to understand, highlight, and explicate the way that choices are integrated into dramatic structures through the model described in Chapter 4, as this represents the crucible where CCAG content and player experience are formed.

Molly Maloney, one of the few Telltale narrative designers who has spoken publicly about the process, says this about her role as a narrative designer:

> The biggest part of my job as a narrative designer is actually trying to anticipate what a player will care about before they play the game and building my choices around those things. This requires a lot of testing and working with players, a lot of iterating content, figuring out what they care about and tying those ideas to our choices. [40]

These three elements form the core of branching narrative design in the "Telltale style": anticipating what players will care about, building choices around that care, and integrating those choices into a coherent story. I encoded the ways these elements intersect in part through the "Story Value Charges" concept in the formal schemata proposed in this dissertation (discussed in detail in Chapter 4 and further in Chapter 8). The concept assigns numerical values and positive and negative valences to particular content pieces that relate to what players "care about" and further identifies common themes between different parts of the story. The structural unit of a "beat" is useful for encoding the relationship between character behavior and the dramatic flow (see Section 2.1.5 and 3.4.1), enabling a method for understanding how choices are integrated.

Telltale leverages its strengths of strong storytelling and the acquisition of familiar franchise licenses, and it has even been compared to a production studio. Former Head of Creative Communications Job Stauffer says in an interview with Christopher Dring, "Our DNA is to be more like a TV studio. We are an interactive HBO or an interactive Netflix" [41]. Telltale Games highlights the emotionally significant choices after each game along with the percentages of other players' choices in a way that is similar to the Netflix strategy of using audience genre preferences to market its catalog.

By making players aware of their own choices, Telltale Games replaces player agency with a sense of identity and responsibility for decisions made. As noted before, by limiting the sequences of events that surround and structure choices, episodic games provide players with the ability to discuss the game without resorting to complicated reconstructions of previous events as they might in a more freeform emergent storytelling genre, another primary motivation for the format, as cited by Bruner in his ICIDS keynote.

1.1.2 The Wolf Among Us

The Wolf Among Us (*TWAU*) [20] was released by Telltale Games first for PC and subsequently for most other platforms. The game received numerous positive critical reviews, becoming an exemplar of the CCAG subgenre. Vertigo Comics has even translated the game into a comic book, the original form of the *Fables* series on which the game is based [42]. In it, players act as the protagonist, Bigby. Bigby is the Sheriff of a town of immigrant fairy tale characters who are struggling with issues of class and race. The player "tailors" (in the words of Telltale Games founders) the story through the selection of options throughout gameplay and by using player controls to perform characters' actions during timed sequences known as quicktime events.

Figure 1.2: Example choice space in Episode 1

For example, a significant choice is depicted in Figure 1.2, a selection of possible responses to a tense encounter in a hallway with one of the episode's antagonists. Telltale narrative designers describe these points in the storygame as **choice spaces**, an apt term that captures the various dimensions that go beyond the prompt of **display texts** indicating player options. I will go further into formal descriptions of the various components of choice spaces and their relationships in Chapter 4.

The Wolf Among Us is episodic: future episodes must account for previous player decisions. Most games in this genre, even if not explicitly released in episodic installments, limit their dependencies in order to control the space of possible traversals. An example of another episodic storygame in the same subgenre is the *Life is Strange* series [43]. In contrast to standalone games, episodic storygames build on previous decisions made by players and have limited replay value due to the amount of content that is necessarily constant between playthroughs, though this is not true of all games. *Heavy Rain* [44] and *Detroit: Become Human*[45] are closely related in their use of dialogue interfaces and could be suitable for analysis using the tools and methodology presented in this dissertation. One of the principle differences with respect to content is that *Life is Strange* emphasizes exploration and exploratory content and provides a mechanism of time-traveling which significantly alters the effects of choice contexts. Instead, the work relies on an accumulation of meaning and considered choices with the constant reminder that certain choices will have consequences (much in the way that Telltale Games often present a notice that characters will remember decisions). Quantic Dream's style is closer to traditional interactive drama in that it emphasizes branching narratives and presents a smaller percentage of the total content in any given playthrough. Likewise, the distribution of possible stories is greater and relationships between content segments are explicitly indicated through both the flowchart feature as well as the lock and unlock icons in the dialogue interface. The effects described in this dissertation principally focus on the dramatic content surrounding and supporting emotional choices and apply to various aspects of other games in the subgenre, but the exact distribution of mechanics and other factors such as branching are subjects of future work.

The initial schemata were designed around the technique that CCAGs use of conservation by considering how features before and after a choice are related to both the dramatic context and the player response. The number of possible relationships between content and choices determines whether this technique is applicable. The larger the combinatorial space of possible sequences, the greater the variability in context effects and the less valuable a particular set of player traversals will be for direct comparison. Simultaneously examining multiple player experiences of traversals that share more in common brings us closer to a practice of engaging with works through the lens of multiple experiences, which I call **medium reading**, described more in Chapter 5. This dissertation focuses primarily on the first traversals of a work as its subject, and so does not address effects that might result from revisiting content with the outcome known.

1.1.3 *TWAU* Summary

The following is a summary of the *The Wolf Among Us*, provided to give context to references throughout this dissertation. I recommend playing the game or viewing related YouTube videos.

Chapter 1. Disturbance

The story starts with Sheriff Bigby Wolf, known in fairy tales as the Big Bad Wolf, investigating a disturbance at Mr. Toad's house. Toad does not look human, which is a problem because non-human fables (fairy tale characters) are required to procure glamours (spells that make them look human) so as not to arouse suspicion from the Mundie (human) community. Upon going upstairs, Bigby finds the Woodsman slapping a woman. After Bigby gets no answers from him, the Woodsman insults the woman, picks a fight, and provokes Bigby into tackling him through the wall of the apartment, landing on Toad's car. Another struggle ensues, revealing that Bigby has werewolf attributes when provoked. The woman stops the Woodsman by swinging an ax into his head. After an unsuccessful interview (which I describe and encode in Chapter 5), Bigby invites the woman for further questioning back to the Woodlands where he lives, and she agrees to meet him there.

Chapter 2. The Woodlands

Bigby encounters Beauty, who, suspiciously, is hiding from her husband Beast. After agreeing or not to keep her whereabouts secret, Bigby runs into Beast in the elevator on his way up to his apartment. There, he encounters one of the three little pigs, Collin, who uses Bigby's guilt about his past deeds to procure a sip of liquor. After a short nap, Bigby is awoken by Snow White, the assistant to the deputy mayor, who brings him down to the front of the building. There, a decapitated head sits, and Bigby identifies it as the head of the fable that he earlier rescued. The careful placement was clearly a message. However, in order to investigate the crime, Bigby needs to identify her name and story.

Chapter 3. Mirror, Mirror

At the mayor's office, Grendel, in human form and standing in line waiting to be seen, accosts Bigby. After Bigby blows him off, Grendel is seen staring at Bigby through the door as it slams shut. Bigby encounters Mayor Ichabod Crane berating Snow White for the handling of the murder and has a choice of whether and how to intervene. After Crane's departure, Bigby, Snow and a flying monkey named Bufkin use the Magic Mirror and several books to investigate the identity of the dead woman. She is revealed to be Faith, also known as Donkeyskin. The Magic Mirror provides the ability to see people if their name is known and so shows Faith's husband, Prince Lawrence, to be nearby and potentially wounded with a knife. At this point, Toad calls and reports an intruder in his house.

Chapter 4. Choices

In this short chapter, Bigby decides whether to investigate Prince Lawrence or Toad's call.

Chapter 5. Consequences Pt 1

Based on the choice in Chapter 4, these events follow one of two paths. The first path involves interrogating Toad about covering up an intruder's presence in his house, only to learn that the intruder intimidated him into keeping silent and threatened to kill his son, coloring Toad's behavior in a very different light.

Chapter 5. Consequences Pt 2

In the alternative path, Bigby arrives to (potentially) save Prince Lawrence from committing suicide. If Prince Lawrence is alive (and visited as the first in the sequence), Bigby finds out that Prince Lawrence does not know Faith's whereabouts, and a suspicious character appears and flees. Bigby chases Tweedle Dee only to be knocked unconscious by his brother just after being told the location of the Woodsman.

Chapter 6. Trip Trap

After a car ride in which Snow communicates misgivings about the current relationships between the rich and poor of Fabletown, Bigby enters the Trip Trap Bar to find that its residents, including Grendel and Holly, are not very welcoming. Instead, it is clear they are covering up for the Woodsman, who joins them at the bar. At this point, Bigby questions him only to find out that the Woodsman is not aware of Faith's death. Grendel then starts a fight with Bigby in which they both transform, and at the end of the fight, Bigby has the option of ripping off Grendel's arm. This action has consequences: If the player chooses to remove the arm, they must enact the actual dismemberment through a series of button presses and witness the horror of the other characters in the room.

At this point, Tweedle Dee arrives looking for the Woodsman. Bigby must choose which suspect to pursue. Regardless of the choice made, Bigby returns to the Woodlands only to find another decapitated head: Snow White's. This scene concludes Episode 1 and is followed by a set of previews for the next episode.

This summary glossed over many of the important events that I will analyze in this thesis. *TWAU* is a critically acclaimed storygame that manages to balance suspense and interest with strategically located quicktime sequences and a variety of moods and locations. The story paints a picture of an immigrant community with deep fissures and a player character with a compelling inner conflict between usefulness to his community and the pursuit of justice by whatever means necessary. The game incorporates genre conventions from its inspirations of film noir and the *Fables* comics by Bill Willingham. The game has a clear narrative arc, compact scene design with integrated choices, and a consistent theme throughout. These features, along with its positive critical reception, make it a good candidate for studying the intersection of player response and storygame content through the lenses of emotions and values.

1.1.4 Computational Modeling

CCAGs are good candidates for descriptive modeling due to their careful attention to crafting the player experience and the fact that their content is visible on the surface of the work. Computational models of narrative (described further in Section 2.1.2) attempt to define relationships between elements of narrative such as actions, events, plot points, characters, settings, and even the more abstract elements of goals, values, and concepts, in such a way that they can be processed.

There are two models that are particularly important with respect to IDNs that I build on throughout this dissertation. The first is choice poetics, developed by Peter Mawhorter along with a generative system for creating choices using it, *Dunyazad* [46]. Choice poetics includes a process for encoding a choice using goal-based choice analysis of an individual decision. The process considers the outcomes and various factors that would shape the perception of the choice in terms of player goals. Telltale Games take full advantage of the expressiveness of choice, although not every choice requires a careful decision based on expected outcomes. The theory and method, however, are useful for considering which features to encode.

David Elson developed the Story Intention Graph [47], a set of schemata that represent key narrative relationships. The SIG was primarily developed using textual stories. It enables the collection of a set of annotations made by different annotators that describe the relationships between actions and characters' plans and then maps them back to the original textual story.

Both formal models capture different aspects of choice and narrative. Choice poetics describes the relationships between outcomes and framing, while the Story Intention Graph models a theory-of-mind view of the story structure from the perspective of a set of annotators. Telltale games involve a series of dramatically presented actions that change the player's perceptions while integrating choices into a dramatic structure. The importance of both the dramatic embedding and the player response are the focus of my schemata, described in Chapter 4.

In the next section, I review the specific research questions this thesis addresses, followed by a description of its research contributions and overall methodology.

1.2 Research Questions

Why do we find storygames such as the kind produced by Telltale Games compelling? What are the structures that govern the relationship between choice design, our experience of emotional engagement, and the overall narrative? What elements of the surface content can be encoded and compared, and how do we leverage those models and the data collected in interpreting the work itself? Recall that the acronym CCAG refers to the subgenre of interest, choice-based cinematic adventure games. These questions are aimed primarily at the modes of presentation and interactivity in this subgenre, although the results may be useful to the study of interactive digital narratives more broadly.

This line of questioning leads to the specific research questions that this dissertation addresses:

1. If we collect **biometric data** during player traversals, can we use these to **identify commonalities and differences** in the experiences of players?

2. Can we design a **descriptive encoding schemata** that captures the **surface content structures** in CCAGs? Does an encoding of **story craft** concepts such as **story values** and **beats** reveal new insights into how Telltale Games integrates narrative structure and choice structure in **existing works**?

3. Can **simultaneously** considering **multiple player traversals** of CCAGs, using an **encoding schemata** of the games, reveal **patterns** in how **surface content structure** is used to shape **audience response**?

4. Can we create an **analytical tool** that could help a target audience of **game scholars and researchers** apply this model and compare player traversals more efficiently than other available annotation platforms or systems?

I will break these research questions down for precision.

1. If we collect **biometric data** during player traversals, can this be used to **identify commonalities and differences** in the experience of players?

For this dissertation, I define **biometric data** as information obtained by either electronic sensors (average heart rate, skin conductivity) or processing of videos of player faces using facial expression analysis libraries (see Section 3.1.2). Facial expression data is output as values indicating engagement, valence, and identified combinations of action units associated with classes of recognizable emotions. These values, taken in their role in interpreting player reactions to content, can be considered as affective signals. **Commonalities and differences** refers to the patterns of player responses and actions observed in relation to similar content while playing a storygame, differences that are associated with specific segments of content. This reflects a desire to identify both the larger structural patterns present in the work as well as reveal individual player actions and reactions, not just the sequence of choices or the dependent content that can be identified through telemetry data.

2. Can we design a **descriptive encoding schemata** that does a better job of capturing the **surface content structures** in CCAGs? Does an encoding of **story craft** concepts such as **story values** and **beats** reveal new insights into how Telltale Games integrates narrative structure and choice structure in **existing works**?

By **descriptive encoding schemata**, I mean formal computational schemata that encodes relationships among elements. **Surface content structures** are the visible performances, choice displays, choice selections, and dramatic elements that are identifiable from a gameplay video. They are contrasted with **deep content structures** such as the underlying goal networks, player intentions, and other possible relationships that are present in IDNs. The patterns that these structures form can be useful in classifying specific instances of either the content itself or the combination of content structures and player experience. The type of schemata I am interested in should apply to **existing works** and not just storygames that can be modified through access to the source code.

I use the term **story craft** to designate concepts from the practice of screenwriting and use them in a content coding scheme. In particular, I use two terms, **beats** and **story values**, from Robert McKee's work [48] to describe the subjective role that certain types of narrative elements have in player engagement and interest. McKee identifies the **beat** as a unit of scene design that involves an action and a reaction between characters or a character and the environment. **Story values** are a categorization schema for the stakes and meanings involved. McKee uses pairs to describe the different valences between which the content shifts during particular moments in a drama. These can be used to characterize the dynamics of the meaning through the course of a short dramatic scene. Examples include justice/injustice and duty/preference. These story values relate to the topics and ideas with which the players may find resonance or by which characters are motivated to pursue their goals.

3. Can **simultaneously** considering **multiple player traversals** of CCAGs, using a **encoding schemata** of the games, reveal **patterns** in how **surface content structure** is used to shape **audience response**?

Simultaneously refers to the assessment of multiple traversals at once in an encoding. **Encoding schemata** refer to the output, a set of related annotations that classify content types and relates values based on logical relationships among content elements. **Audience response** is a broader and more subjective term than the **affective signal** term used previously; it incorporates verbal and other means of understanding how the content affects an audience. By **patterns**, I mean recurring configurations of relationships that can be generalized across multiple works or multiple instances within the same traversal.

> 4. Can we create an **analytical tool** that could help a target audience of **game scholars and researchers** encode content and compare player traversals more efficiently than other available annotation platforms or systems?

The **analytical tool**, in this dissertation, is a piece of software that provides functionality that aids in the various stages of analysis of digital artifacts, specifically the records of a player's experience with a game in the form of a time series and video recordings. The tool relies on a formal encoding scheme, such as that described in the Text Encoding Initiative (TEI) guidelines. These models can be used to organize content and as a format for encoding software tools to apply to various artifacts. The schemata approach permits identifying and counting patterns and occurrences of specific elements in a set of works and is particularly useful for addressing large sets of data. The **analytical tool**, like those developed to incorporate or produce TEI encodings, can provide assistance to analysts in the encoding process itself by searching, sorting, and collating content in a corpus that has been previously encoded. It also provides assistance in navigating to specific content within a corpus. These features provide a means for analysts to find new patterns through the model or explore the distribution of features in the dataset.

1.3 Contributions

There are many barriers to performing analyses on storygames. These range from the inability to cite specific states [3], to the lack of easy acquisition of telemetry data from players experiencing interactive narratives, to the need to incorporate a player-specific series of experiences and actions. While PlaySpecs addresses some of these issues [49], it is clear that we cannot fully characterize the responses of games such as *The Wolf Among Us* through telemetry data alone. Instead, I argue that we must employ an approach of telemetry data combined with examination of the work's affective content using story values and depicted representations of character behaviors, a player response, and the related choice structures.

The main contribution of this dissertation is a method and software tool for labeling and analyzing interactive narrative content using a formal schemata alongside player response data.

The primary contributions are:

1. A **schemata, the Interactive Cinematic Experience** (ICE) ontology, a Web Ontology Language format (OWL 2.1) ontology that relates the protostory, player traversal record segments, player data, and poetics into schemata for analysis.

2. A **process for encoding multiple traversals** and an **encoding** of the first three chapters of a player study using *TWAU* using the ICE schemata. I use an example scene to compare the encoding process and features to Elson's theory-of-mind model of story, the Story Intention Graph [47], and Mawhorter's goal-based choice analysis [46].

3. A **multi-channel corpus of data and annotations of *TWAU* Episode**

1 player traversals using ICE and including both self-reported and extracted player experience measures.

4. A **web-based annotation and visual data mining platform**, SHERLOCK, designed to support researchers studying multi-channel player experience data through filters, queries, and visualization using annotations based on formal schemata such as ICE. I evaluate the tool using a preliminary user study of seven game researchers as well as through using it to construct an encoding.

To investigate the relationship between players, game, and narrative, I conducted a player study with my collaborator, Raquel Robinson. We co-developed the study design and procedure and collected video records and response data for *The Wolf Among Us* Episode 1 using Robinson's tool, *All The Feels* (ATF), that renders biometric data on each frame in the form of several graphs and emoticons that indicate the presence of detected emotional expressions [50]. I describe my initial annotation schema used to analyze the dataset and initial results in Chapter 3. The successes and failures of the annotation schema in Chapter 3 led to the development of ICE to sharpen the relationships between content and the details of player traversals. The schemata itself and several patterns that can be described using it can be found in Chapter 4. The process and an example encoding can be found in Chapter 5, while an overview of annotation and an analysis of the encoded dataset can be found in Chapter 7.

The ICE model relates protostory elements to concrete instances of player traversals through a parent-child relationship. Instances of content are associated with a single protostory element, which in turn connects traversal to one another. I draw terminology and structure from Koenitz's concept of the protostory [51] as well as McKee's definition of the beat and story value [48] (see Section 2.1.5).

Beats are used successfully in Mateas and Stern's interactive drama [52] as a method for both mixing and sequencing content. I adopt a formal component-based method of analyzing story values based on Mawhorter's goal-based choice analysis [46], and Elson's success using a formal model inspires the idea of using a graph-based representation [53] (see Section 2.1.3.2). These theoretical apparatuses result in a granularity level of individual elements that can be associated with emotional content in the form of components, each of which has a story value, a magnitude, and a valence that represents the subjective interpretation of the content by an interpreter. I argue that the variations in story-value charge within the context of choices and emotional anchors are useful for understanding how content structure shapes player experience in the subgenre, as demonstrated in the analysis of the encoding performed in Chapter 7 and the analytics conducted in 7. I further argue that these patterns play a role in producing the Telltale effect, a player response to the constrained emotional situations discussed further in Chapter 8.

1.4 Overall Methodology

My method adopts a descriptive and data-focused approach to understanding IDNs at the intersection of interactive narrative, game studies, and affective computing. I am interested in understanding the diversity and causes of player experiences. While the field of computational narratology has a vested interest in this topic, it has, with several notable exceptions, not conducted detailed and openly available corpus projects in the context of existing storygames, instead focusing on evaluating specific effects of research systems on particular dimensions. The work detailed here shares much in common with the emerging field of affective computing, where narrative is only beginning to become an object of interest [54], although the methods of my dissertation are firmly situated within the critical-technical practice of systems building and modeling through both the formal model and the platform used to work with it.

The methodology follows other projects at the intersection of artificial intelligence and expressive processing. One of the principle differences is the focus on the data available that represents player experiences of existing works as well as the emphasis on the role that surface features play in storygame content. This results from the observation that complex narrative artifacts rely on player responses to "instantiate" their content, and so bringing both content and response into a relationship is critical to investigating the relationship between choices and their intended effects.

Previous research in this tradition has followed a process that roughly follows these three steps, which I also adopt:

1. Identify an object worthy of study.

2. Develop a system (and, consequently, a model) that reveals more about it.

3. Learn about the system, the object, and the process, then iterate.

The object worthy of study in my case is the relationship between dramatic "content" and player responses to that content in storygames. My interest is not in hypothetical or possible responses but actual responses based on a human experiencing the storygame itself. The system (and model) are designed to reveal relationships that may be hidden, either masked by our own responses as critics and researchers or by the multitude of responses that the work has elicited, both in research settings and in the various settings where the game is played. The two contributions of ICE and Sherlock represent a strategy for approaching the challenge of documenting and understanding how players respond to interactive narratives. Through the course of the work described in this dissertation, I have come to further appreciate the roles that come together to create such works: authors, performers, developers, animators, and, most importantly, players. It is not possible to put oneself truly in another's shoes, as the saying goes, but that is the role that authors such as Molly Malone must perform. In craft writing advice, it is often a rule of thumb to have respect for one's readers, in the sense that a writer should give them the room to process and make their own judgments even as a writer must necessarily lead them through the experience of the narrative. This dissertation builds on this rule of thumb, developing the insights gleaned from a close examination of player responses alongside the content representing authorial decisions as a source for expanding our knowledge of interactive storytelling.

The following sections discuss several lenses that reflect different approaches taken by research systems, particularly as motivations behind taking a systems-building approach.

1.4.1 Systems as Hypotheses

The field of artificial intelligence has long used artificial systems to understand natural ones. Sometimes, these systems are used to obtain similar results, as is the case in genetic algorithms through the use of generations and selection. Other systems, such as neural networks, describe a version of collective pattern-matching and adaptation believed to be present in our own nervous systems. Sherlock benefits from several technologies incubated in the field of artificial intelligence. These include representing schemata as ontologies (formal representations of knowledge that mirror our own understanding of concepts and their relationships) and the use of pattern-matching to identify and classify phenomena (particularly the use of facial expression classification libraries). While Sherlock itself does not represent a novel AI technology, it does represent a new approach to leveraging existing work. It represents a hypothesis of how AI systems can be used to assist in human interpretation rather than to mimic or automate it.

Sherlock functions as an analytical tool to support comparison of player traversals in a way that traditional video editors and annotation tools currently do not. The tool produces structured representations of an experience that combine top-down interpretations from human annotations with bottom-up observations from various sensors and detection libraries. Hypertext systems provide a similar capacity for navigating among lexia [55], or chunks of text, that offers certain advantages over print media, whereas current video players offer fewer options for structuring the display and analysis of video content. For instance, a researcher might need to compare points in time in different traversals in order to understand how a strategy or content choice is related and to develop new theories. The use of actual player experiences enables a degree of rapid testing of hypotheses of potential relationships and patterns that may not be possible with other, more speculative approaches. The system design embodies the hypothesis that researchers require the original content form as well as a computational representation to improve detection of related patterns and describe them in nonlinear storygames. The system adopts the assumption more generally held in visual analytics software that, by making fewer assumptions and visualizing as many features as possible, the space of possible relationships is greater than that of a single purpose visualization of one or two variables or even of a visualization that includes an assumed relationship.

Several works take an approach that could be considered hypothetical. Elson's *Scheherazade* seeks to represent the most salient aspects of a story through a theory of mind interpretation by annotators. Elson uses it to discover and test analogies in a corpus of Aesop's fables. Another example of a system that uses visualization is Nowviskie's Temporal Modeling Project, which explores the effects of visualizing the subjective interpretation of time through metadata and how this can inform interpretation and understanding [56]. Moreover, Mawhorter's *Dunyazad* and the theory of choice poetics that was co-developed with it serves as an example of the value of detailed analysis in a computable fashion [46].

One of the key issues with taking a computational approach is the danger of falling into relying solely on computationally "simple" methods of evaluation. Emotion recognition provides an example of this: By seeking to classify emotions, the task becomes an easy problem with similarities to other classification tasks. We can trace this trap to the labels assigned by the various platforms that suggest they are objective representations of emotions but that in fact amount to pattern-matching based on machine learning models trained primarily on static expressions. The idea that classes of emotions are sufficient to describe experience has already been subject to scrutiny, and alternative models that incorporate scales are increasingly gaining popularity. I return to and review affective computing and its treatment of emotions in Section 2.3 and Section 3.1.1.

My primary hypothesis for SHERLOCK is that organizing player traversal data and visualizing it alongside the original content will reveal patterns of interest to researchers of interactive narratives.

1.4.2 Systems as Platforms

Flash: Building the Interactive Web, by myself and my coauthor, Raquel Robinson, explores software platform relationships, specifically the complicated relationship between the systems themselves, the companies that create them, their artifacts, and their creators [57]. Research systems are usually far more modest in scope than the Flash Platform, but their goals of easing certain types of computational and conceptual work remain the same. Most platforms in academia aspire to be as general as possible. For instance, the *Story Workbench*, designed explicitly to enable a variety of different models to be annotated [58], follows the software and philosophical footsteps of the Eclipse platform. *Scheherazade*, the annotation tool for the Story Intention Graph schema, also includes many features that intended to make the system more generalized, including an API and a narrative rules engine separate from the schema itself [53]. Elson forms hypotheses about the use of schemata to derive facts about encodings and further tests whether interface features such as a textual feedback mechanism ease the process for annotation. Elson examines whether the model can identify analogies between parallel encodings of Aesop's fables.

Platforms such as the Flash Platform embed facts and assumptions that go beyond their technological capabilities and limitations. In the case of Finlayson's *Story Workbench* and Elson's *Scheherazade*, a necessary privileging of linguistic modes of representation make these platforms much more accessible for textual annotation projects, necessitating a new approach for video annotation in SHERLOCK.

Games User Research, Computational Narratology, and IDNs are three related areas that have seen immense growth in recent years, as we will see in the next chapter. There is no clear single framework or approach that has proven more valuable than others, and so there remains a diverse set of competing theories, models, and approaches. For this reason, one system goal is to be capable of using different theoretical models for comparison and in order to remain useful for longer. This dissertation contains one such model, but others can be translated into an ontological representation and used for annotation or analysis. Another important topic in the data science and AI communities is the ability to reproduce research results [59]. Thus, another goal for the system is to retain as much of the original provenance (relationship of derived results to the original source data) as possible and to record any explicit representations of theoretical and conceptual commitments, both within the tool and within the annotations. The modularity is made possible through the flexibility of the OWL 2 ontology language, which enables the coexistence of multiple representations. Future work would include creating translations of other models, such as the Story Intention Graph, and comparing or combining them. The semantic web technologies described in Section 4.2 enable combining or extending the approach with other models. Through this, I hope to make the work more accessible to future researchers.

1.4.3 Observation and Interpretation

There remains a divided view of approaches in the academic community, which separates work into quantitative and qualitative or empirical and interpretive.

The first approach, of quantitative and qualitative, relies on operationalizing and distilling theories into a formal system such as mathematics or an implemented computational system, followed by evaluating them by automated means, usually through formulating them as well-defined "problems." This mode has proven very productive for both scientific endeavors and many aspects of human behavior, but it still makes an uneasy fit with many of the values and insights sought in the humanities. One of the most common approaches in computational linguistics, now becoming increasingly popular among digital humanists, is what Franco Moretti has termed *distant reading* [60], [61]. In his book, *Distant Reading*, Moretti applies a version of network theory as a tool to analyze the play *Hamlet* [60]. This approach has grown into a family of methods that take a model and abstract various aspects of a work or collection of works to tabulate and compare properties. This term serves as the source for Chapter 5's title "Toward Medium Reading," moving the emphasis back from the collection of artifacts to a collection of player experiences of an interactive artifact as a lens into the work itself. Moretti's work has provided an example for quantitative methods to be applied to a variety of traditionally humanistic topics [62], although it has also been subject to criticism within the digital humanities [63]. Another example of the quantitative approach inspired by Moretti is David Elson's approach to a literary hypothesis using extracted dialogue networks in a corpus of novels.

One further quantitative approach adopted by the machine learning community involves using a well-specified benchmark to evaluate the performance of a system, typically a machine learning model, on a dataset. This approach has been used to develop systems for recognizing emotions [64]. Some of the systems used in Sherlock benefited from the availability of the datasets and libraries that result from this approach. Other examples of data-intensive, quantitative "corpus" approaches include Google N-Gram, TimeML [65], The Penn Discourse Treebank [66], and Elson's DramaBank [67]. The two major drawbacks of this approach are the necessity for knowledge engineering and the loss of individual features present in members of the corpus. The knowledge engineering necessarily needs to take place before encoding the works and may not reflect all of the desired features. This usually involves a schema of annotations, although only a few are based on a formal model that can be described as an ontology. During the development of ICE, I retained the original, simpler dataset annotations (found in Chapter 3) and used Sherlock to import and translate them into various possible encodings while providing appropriate features for adding new relationships as one possible solution to this challenge. This commitment to one layer of representation makes experimentation with the connections simpler and can support future evolution of the model from the same dataset.

One qualitative approach taken in game studies and the humanities relies on critical theory and interpretation to guide understanding of what may be beyond current numerical abstractions. This method almost exclusively uses selected pre-existing works and synthesizes from them approaches that may not be specific or complete enough to construct a computational system. Theoretical results include proceduralist readings of works [68], craft knowledge, and recommendations [69], much of narrative theory (see Section 2.1.1), as well as critical and close-playing methods.

Combining the two methods of qualitative and quantitative seems like a promising direction and one that goes by the appropriate name of **mixed methods**. Lieberoth and Roepstorff describe some of the benefits of approaching games from this perspective by comparing how different lenses can only identify layers within the totality of the game-playing practice. These four layers, including three based on Stevens et al. [70], include in-game, in-room, in-world, and in-body. These layers represent the various influences that context plays in the reception of a piece and also provide a basis for the commitment to combining content and context for interpretation. Another example of a mixture of qualitative and quantitative is Mawhorter's approach to choice poetics. It involves co-designing a system alongside the theory, incorporating insights from one into the other [46]. My development methodology also features co-developing the theory represented in ICE alongside a tool that incorporates it: SHERLOCK. McCoy's evaluations of *Prom Week* [16] use both methods to evaluate different aspects of their performance.

These different approaches reflect different perspectives on knowledge creation and values, indicating a prioritization of observations, theory, or various combinations of the two. SHERLOCK provides one approach that keeps the various elements separate: Observations, theory, and interpretations are modeled as distinct elements, and I strive to retain their independence as much as possible. By retaining as much of the original content and responses as possible and documenting the methods and assumptions of the approach, future additions or replacements, refinements, and alternatives can be made without sacrificing the original dataset's value. My ultimate goal is to create a platform for applying various tests that draw from new models while still supporting and encouraging traditional humanistic interpretation of the fundamentally human processes represented within it. This balance of valuing qualitative insights while making available quantitative data in the form of descriptive statistics that could provide insights for future quantitative studies is a primary contribution of this particular mixed methods approach.

This dissertation offers another method for bringing these two primary approaches, humanistic and scientific, into a conversation without necessarily merging them or valuing one over the other. The attentiveness to others' interpretations and the incorporation of models to assist in a broader view of interpretation can be considered a form of *medium reading* and is described in more detail in Chapter 6. The challenge is not so much defining a methodology for others to adopt as it is providing scholars of one tradition the tools and resources to experience insights and findings often taken for granted in the other. Many of the quantitative tools, such as those incorporated into SHERLOCK, are inaccessible or unavailable to scholars without a computer science background, while many of the theoretical insights that could provide value in the hunt for meaningful interactions with conversational agents and computational narratives are scattered across various programming languages, systems, and reports or are described but not implemented (as, for instance, the system described by Marie-Laure Ryan, the "recursive graph model" [71]).

1.4.4 Storygames as Intersection

Storygames represent an ideal nexus for such work, as they contain content best assessed through interpretive and humanistic lenses while presenting opportunities to evaluate efforts to computationally model aspects of their organization, presentation, and reception.

Stories encode complex experiential information in a way that not only allows for communication but also functions as a form of creating an experience itself [72]. They combine engaging situations with characters that are worth getting to know. They capture the types of emotions that we experience at times yet do not ordinarily encounter in everyday life. Writing a good story requires making observations about human life and conveying them in a form that is both engaging and moving. Humans do not always have all the information necessary, much less the time, to make ideal decisions, and our decision-making process is complex and multifaceted: We are emotionally attached to outcomes, express values, and empathize with others. The time-based choices found in the cinematic choice-based adventure game genre involve value judgments and represent many of the reasons we enjoy and write stories. These storygames recognize and exercise our ability to judge social cues and make quick decisions in a form that is both artistic in its cinematic familiarity and mimetic in its closeness to our everyday experience of conversations and values.

The author's view of a story is one of a sequence of creative choices: What would best serve the characters and the effect at this given juncture? Traversals, on the other hand, reflect the path and decisions made by players as they pursue their goals. These two views provide a dual lens with which to understand a story design and the audience that receives it. Nick Montfort popularized the term "traversal" [31] to designate the possible paths through multicursal artifacts such as interactive fiction or hypertexts. The term includes variants that indicate a successful or, as Reed prefers, satisfying traversal. Koenitz, on the other hand, addresses the **multicursal** (multiple paths) nature of storygames by proposing the term **protostories** [51] to describe what amounts to the possibility space of narrative elements encountered in an interactive story. We take advantage of this idea of abstraction in our representations, particularly through adapting the prefix "proto-" in various protostory layer elements described in Section 4.3.

Managing these choice structures without revealing the artifice is no easy task. David Cage describes this act of carefully managing choice branching and merging as "bending stories," providing the right hints, feedback, and sets of reactions to ensure that most of the authoring effort is conserved through most traversals [73].

Analyzing storygames, their content, and specific player traversals provides an opportunity to explore how art can shape our feelings, why certain choices are more engaging than others, and, most importantly, how we can understand others' viewpoints and what influences their decisions.

1.5 Dissertation Outline

In the next chapter, I will review the three primary fields that form the foundation for the approach to studying player experience of interactive narratives: affective computing, computational narratology, and interactive digital narratives. In Chapter 3, I describe the motivations, setup, and initial results of a player study executed in collaboration with Raquel Robinson. Chapter 4 describes the ICE schemata, which relates surface features to player actions and responses. Chapter 5 demonstrates an encoding of a scene from *The Wolf Among Us* using ICE and compares it to two other computational models. Sherlock, the annotation, visualization, and analysis platform, is presented in Chapter 6, which supports researchers in exploring player traversals and in particular the dataset described in Chapter 3. I use Sherlock to analyze an encoding of the first three chapters of *TWAU* in 7. Finally, I summarize and discuss general implications for the research and its contributions in Chapter 8, along with a description of the Telltale effect, a response to the shaping of player emotional agency through dramatically embedded choices.

Chapter 2

Related Works

The authoring and study of interactive digital narratives (IDN) incorporate many disciplines and fields. My goal is twofold: to incorporate insights from affective computing and computational narratology to explain how a particular subgenre of IDN structures content to give rise to player affect, and to create a tool that can help scholars compare player traversals. SHERLOCK (described in Chapter 6) is designed to help humanists examine parallel traversals of IDN video records alongside related datasets, while being modular enough to help researchers create formal models and investigate interactive digital narratives and artificial intelligence. Using both formal models and biometric and extracted data, this chapter sketches the motivations behind research on player affect.

Three threads weave through this dissertation and this chapter. The first connects narratology to computational models and their uses, including primarily modeling and understanding using computational methods. The second discusses how computational media systems have combined interactivity and narratives. The third thread surveys player experience itself, making use of insights from affective computing and games user research.

Player experience in interactive digital narratives is difficult to measure or study, even using instruments such as surveys or player studies. Issues include the specificity of the questions as well as the presence of various post-experience effects that may change a player's perception of their experience from the actual moment-to-moment sensations and feelings. Combining formal representations with collected measurements of player response will provide a basis for developing and assessing encodings for interactive digital narratives and how their structures relate to player responses. This dissertation contributes encoding schemata in the form of an OWL ontology and demonstrates an encoding method that annotates storygame content with effects relating to its reception. These are in turn related to player experience datasets for finding and describing patterns in choice-based cinematic adventure games, leading to the observation of a set of features that produce a distinct effect for players, described in Chapter 8. To facilitate both annotation and analysis, I developed a tool that can guide researchers using the data in their interpretation of the patterns that emerge and that can help them relate features back to the original game content and player videos.

Sherlock and the Interactive Cinematic Experience (ICE) schemata are designed to further the study of player experience; they therefore target game researchers. These researchers are likely to address prominent questions in the field of IDN, such as the following: How can interactive content be modeled? What properties does it have? What elements or relationships cause players to engage or disengage? How can you computationally manipulate annotated content? How do the effects (and affect) of narrative content change when combined with interaction? How can IDN structure be measured or assessed? Do some player types respond differently to narrative content? Is it possible to predict or assess player attributes, such as personality or values, that might influence players' experience? Does the genre or emotional nature of a story dictate the types of structural design that work for a storygame?

These questions are far from being answered, and their pursuit motivates my work. All the disciplinary perspectives that examine narrative, interaction, and player experience provide useful tools for addressing these questions. Telltale Games has found a formula in their subgenre that reflects a great deal of innovation in creating engaging player experiences of story, and I believe a close examination of their work can reveal how their unique fusion of story and game can provide insights for researchers in the IDN community and beyond.

2.1 Computers and Narrative

Stories provide a form with which to communicate experiences, and their mode of communication benefits from a variety of advances in knowledge. Shakespeare forever changed theater, itself an ancient method for storytelling, by developing stories that are now valued across generations and in many translations, depicting and evoking core aspects of the human experience. The close study of Shakespeare's works is taught in classrooms across the United States, illustrating storytelling and providing context for other works that reference them. In *Hamlet on the Holodeck*, Janet Murray considers the implications of such attention to future storytelling media, using a thought experiment to consider its possibilities and consequences [74]. Future technologies for storytelling are often born out of previous explorations into the possibilities and their promises, and the innovations that Telltale Games has pioneered in the subgenre fall into this category.

Stories and storytelling have increasingly been the object of multidisciplinary approaches. Keith Oatley's *Such Stuff as Dreams* [72] attempts to reconcile Shakespeare's attitude and use of storytelling with psychological theories of mind, while Herman's *Storytelling and the Sciences of the Mind* [75] proposes that cognitive science researchers would benefit from studying narratology and that narratologists would benefit from a knowledge of cognitive science. In what follows, I outline a few of the many approaches that are relevant to understanding the relationship between technology, specifically computation, and narrative. Efforts to formalize and classify elements of narrative are also attempts to clarify the art form and its various elemental relationships, much as visual artists study anatomy and color theory. Visual artists who use the human figure in their works often need to understand bones and musculature to render the surface form in a way that achieves their expressive ends, even if those ends include exaggeration. The audiences of both art forms can detect certain kinds of flaws, especially if they interfere with the desired effect.

2.1.1 Narratology

As anatomy is to visual artists seeking to render the figure, so narratology is to writers seeking to convey a story. This section reviews the field of research dedicated to the study of narratives, paying particular attention to the structural approaches because they provide features useful for annotating narrative content.

The modern field of narratology blossomed during a renewed interest in structuralism led by Roland Barthes [55], Todorov [76], and Genette [77], among others. The Russian formalists popularized narratology, especially Vladimir Propp's study of the similarities in the morphology of Russian folktales [78]. The narratology approach was distinguished from other approaches to literature by focusing on underlying patterns rather than specific meanings and interpretations of a single work. These formalist and structuralist roots can still be found in the computational narratology and linguistic approaches that classify, generalize, and abstract, as I observe later in this chapter. Other influences on the growth and popularity of narratology as an academic discipline include its recognition within linguistics as a concern distinct from semantics, syntax, and pragmatics [77].

Narratology benefits from distinct classes and their relationships, a feature that any structured representation requires. Mieke Bal consolidates the views of many previous theories of narratology in her three editions of *Narratology: Introduction to the Theory of Narrative* [79]. Bal, like others, separates narrative into layers that correspond to different classes of meaning and materials: *text*, *story*, and *fabula*. The *fabula* is the set of events and their relationships that make up the storyworld, whereas the story corresponds to the artful presentation of some elements from the *fabula* to achieve a particular effect. Distinguishing between story and *fabula* is a common tool among narratologists: Chatman uses the alternate terms *discourse* and *story* [80], whereas Todorov uses *discours* and *histoire* [76].

ICE focuses almost exclusively on the *text* and *discourse* layers, including aspects of the *story* layer only as necessary. The visual representations of the cinematic game form a complex and multilayered text.

2.1.2 Computational Models of Narrative

My goal of studying interactive narrative using schemata follows the computational narratology practice of annotating artifacts to gain greater insight. Recent approaches to the automation of narrative understanding and generation seek to mimic human processes, either to support creative processing or to understand the processes of creativity or understanding itself. While a comprehensive review of the research on narrative generation is beyond the scope of this dissertation, narrative generation shares an interest in representing narrative concepts for processing. My approach is inspired by a desire to improve the theoretical and representational knowledge of how humans interpret and respond to stories and choices, based in part on the success of several interactive narrative systems described in this dissertation.

Computational narratology emerged from several approaches to formalizing narrative in the existing disciplinary traditions [8] while having a parallel existence within the AI community. Inderjeet Mani describes computational narratology as "the study of narrative from the point of view of computation and information processing" [81].

Several conference series have created a community of researchers who have explored different perspectives on computation and narratives. The ACM Workshops on Story Representation, Mechanism, and Context, which were held in 2004, 2006, and 2011, investigated both the creation and analysis of stories and related technologies in a variety of media, but mainly in nonfictional scenarios. The Intelligent Narrative Technologies (INT) workshop series began in 2007 alongside the AAAI Conference on Artificial Intelligence, sponsored by the Association for the Advancement of Artificial Intelligence Conference. The Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE), which has been running since 2005, also focuses on digital games, and AI has been a significant venue for research in the area. The International Conference on Interactive Digital Storytelling (ICIDS) conference series provides a venue for theoretical and creative works in interactive storytelling as well as discussions of how to evaluate such works. These conferences reflect the past decade's growing research interest in combining artificial intelligence, storytelling, and digital games.

2.1.3 Annotation-based Approaches

This dissertation contributes an annotation and collation of a collaboratively collected dataset of player traversals as well as a method for annotating video data with an encoding based on a formal model. What other disciplines adopt corpus-based methods, and how can these be adapted for understanding player experiences of interactive digital narratives? This section addresses work in both the digital humanities and the social sciences.

Stories can be considered artifacts, both as interactive programs that when run produce a story and as linear stories that reside in media such as printed books, film, and graphic novels. The distinction between the content contained in the artifact and the experience of it represent two distinct classes of records that are useful to consider in a player experience corpus. The annotation of meaning and features using controlled schemata is one method employed by the Text Encoding Initiative [1] to share research and schemas. Assembling annotated corpora for use in analysis and classification is also one of many approaches in computational linguistics, and their use provides the basis for current consumer speech recognition products. The collection and sharing of annotated corpora inspired the Penn discourse treebank to provide a set of example parses for sentences [66].

Another corpus-based approach is to document and label distribution of content. Schmierbach [82] surveys research on content distributions in games, noting that previous studies have focused on violence and gender and have adopted a variety of methodologies. Krippendorf [83] details six components for content analysis: unitizing, sampling, recording/coding, reducing data, drawing inferences, and narrating the result. The approach taken in an ICE encoding adopts some of the interpretation and coding methods from this style of content analysis, with its

[1] `http://www.tei-c.org/index.xml`

focus on both the player experience and the relationship of content *structure* to player input.

One open-access game corpus research project that shares many goals is the Platformer Experience Dataset ([84]). This corpus is based on a collection of recorded player actions, demographics, and generation parameters, and it uses a simple comma-separated values format for metadata. However, the goal of ICE is to represent relationships between player choices and dependent content, requiring a more structured approach to retain the dependencies present in the choice structures. The project's results suggest that fusing a multimodal database of both player and gameplay is a promising undertaking, and the goal of reducing the future costs of research using available datasets is laudable.

Despite the connections between interactive digital storytelling and computational narratology, there are no open-access corpora of player traversals of storygames. Researchers have often created novel systems and conducted user studies to collect quantitative insights into interactive digital narratives or have used more traditional user study methods that target a particular game. The dataset collected and described in the next chapter would be the first openly available corpus of correlated player response, game content, and player choices of a storygame, and we intend to release the videos that correspond to the annotations and extracted features provided by this dissertation. The corpus represents not only a record of player interactions but also a computable representation of the structure of the content and of the nature and timing of player responses.

2.1.3.1 Story Workbench Tool

Mark Finlayson is an early pioneer in creating open-source and openly available tools for narrative annotation [85]. Finlayson has also been actively involved in the computational models of narrative community, acting as organizer of the workshop series that was active from 2010 to 2017 [86]. In one review of the development of the software, Finlayson describes three loops of the annotation process: individual annotation, double annotation, and annotation development [58]. The dataset that constitutes one of the contributions of this dissertation currently uses my annotations and annotation development, though future work will include the collection of multiple sets of annotations from different annotators using the web-based tool and evaluating inter-rater agreement.

2.1.3.2 Story Intention Graphs

David Elson developed a set of discourse relations comprising schemata that are collectively entitled "Story Intention Graph." The semantic graph is defined by a set of labeled nodes and arcs and a set of logical relationships that determine "legal" relationships among them. In this section, I briefly review the key features of the SIG schemata that are relevant to my encoding, and I conduct a further comparison of applying it and ICE in Section 5.3.1.

Elson demonstrated that a corpus of encodings could be used for automated analysis using the Story Intention Graph (or SIG). The SIG built on insights from previous formalisms from narrative discourse and narratology, including Wendy Lehnert's Plot Units [87]. Elson describes Story Intention Graphs as a formalism whose goal is to be **robust**, "emphasizing the key elements of narrative rather than attempting to model the entire semantic world" [47]. He also describes it as **expressive and computable**, such that it is "formal enough to allow us to find analogies, identify patterns, and design summaries of narrative content" and **accessible** and so able to be used by human annotators with minimal training [47]. He used the representation to automate recognition of well-defined "patterns" that composed a library of story patterns much like Lehnert's plot units. This approach of both creating a library of possible patterns as a measure of expressiveness and the use of the structure to detect patterns that exist are two methods of use for evaluating ICE encodings. The importance of being accessible is also shared by ICE, in the sense that the model should be usable by game researchers without formal training in annotation or programming. This is part of the motivation behind SHERLOCK and a basis for the user study described in Chapter 6.

2.1.3.3 Drammar

Drammar is both an ontology and a theory of drama, described in [88]. The notion of breaking up dramatic content into units and then mapping these onto values is one taken in this dissertation, though the focus is on a subjective evaluation of the thematic impact of content rather than an evaluation of the goals with respect to fulfillment or frustration. Lombardo et al. also used the web ontology language to represent the formal encoding [89]. In the Drammar ontology, they specify a taxonomy of classes and a set of Semantic Web Rules Language (SWRL) rules that function similarly to the Story Intention Graph in modeling key relationships and features. Lombardo et al. use the ontology and a process for encoding as a teaching tool in drama classes, and have even used it as the basis for visualizing *Hamlet*. Drammar and the success of creating such an ontology are of particular relevance to my work, because it is an annotation of dramatic media and includes entailment rules that incorporate character emotions in dramatic media, though it does not go so far as to consider issues of reception or audience emotions. A more detailed comparison of the two models is beyond the scope of the current study, however, because ICE focuses more on the surface representation and moment-to-moment dramatic reception of content, whereas Drammar focuses on structural elements relating to the larger-scale conflict.

2.1.3.4 Video Game Ontology

Another video game ontology project that shares some of the goals with the current work is the Video Game Ontology (VGO)[2] by Parkkila, Radulovic, Garijo, Poveda-Villalón, and Ikonen [90]. The ontology provides "a model for enabling interoperability among video games and enhancing data analysis of gameplay information." Another related ontology is the Game Character Ontology (GCO)[3] by Sacco, Liapis and Yannakakis [91]. The ontology was created with the vision of extracting character information dynamically from online sources. This kind of encyclopedic content is provides an example of a taxonomic approach, one which puts an emphasis on naming and identifying game-relevant properties such as armor, actions and relationships. The approach of this dissertation is to identify the results of lower level event descriptions such as content time indices and biometric and response measurements. This approach presumes that labels could be determined in an emergent fashion.

2.1.4 Communication and Film

The discipline of media studies connects the notion of content to the measures of emotion that I review in section 2.3.1, including the effects of particular choices on dimensions of emotional response in audiences [92]. Detenber and Lang present an argument for the role that non-content attributes of media play in creating emotions, such as cutting, color, and other elements. It is also interesting to note that the *Routledge Handbook of Emotions and Mass Media* contains a chapter on the automated analysis of emotions through facial expressions, confirming

[2] `http://purl.org/net/VideoGameOntology`
[3] `http://autosemanticgame.eu/ontologies/gco`

that computational methods are becoming a part of this burgeoning field [93]. However, the chapter provides a set of applications and methods for collecting and classifying expression that is closer in tone and content to work done in affective computing than to that done in media studies, and these applications and methods are consequently not utilized or leveraged by many of the other authors in the volume.

2.1.5 Storywriting Craft Advice

A nonacademic source of insight for creating an encoding process is how writers and authors view their own craft, and how books teach it. One such author, Robert McKee, wrote the 1997 book *Story: Substance, Structure, Style and The Principles of Screenwriting*, which has inspired many projects and approaches in computational narratology in addition to being widely familiar to screenwriters. I have adapted several of the book's concepts and terms, primarily that of *story value* and *beats*, and it is worth considering their origin.

McKee has been criticized for presenting his ideas as truth and for his accessible writing style, which is sometimes more entertaining than it is formal or precise [94]. Along with similar books by other published authors, his works have been valued for their representation of the craft of storytelling as a worthwhile endeavor and as an example of a method for analysis. McKee's approach utilizes examples from film to build a theory of genre, expectation, and form that, while not precise, captures some aspects of the process of writing that are important to both writers and media theorists.

I am primarily interested in the insights structural analysis can provide, and two of McKee's terms, "beats" and "story values," provide a temporal and behavioral structure and a means of classifying content, respectively. Other definitions of "dramatic beat" have similar meanings, including those offered by Judith Weston [95] and Nicholas Proferes [96]. The ultimate connection of these terms relies more on the underlying narrative elements and their relationship to their presentation, as described above. McKee's distinction between action and reaction provides a straightforward method for breaking up the content. Other methods might include action decomposition or linguistic analysis, but the concept integrates the sense of intentionality often found in dramatically presented content.

A unresolved challenge in both academia and the commercial game industry is *how* to combine narrative with opportunities for player interaction. That is the topic of the next section.

2.2 Interactive Digital Narratives

This section focuses on the more general category of works to which *The Wolf Among Us* (*TWAU*) belongs and situates the contribution in that category's historical and disciplinary context.

Interactive digital narrative (IDN) genres include interactive fiction (IF), adventure games, and hypertext works including classical works such as Michael Joyce's *afternoon: a story* [97] and Stewart Malthrop's *Victory Garden* [98]. IDN also includes modern genres such as works created using the HTML-based hypertext environment *Twine* [13], including Porpentine's *Howling Dogs* [99]. The subgenre of interest, choice-based cinematic adventure games, is a relatively new subgenre of IDN, following experiments in adventure games that were closer to the model of the classic adventure game.

2.2.1 Interactive Drama

The dream of players directly interacting with virtual characters in real time while ideally retaining the emotional power of dramatic works captures the concept of interactive drama. Many scholars regard Laurel's dissertation [100] as one of the first studies to describe interactive drama as a grand challenge, because it lists a set of requirements that involve a computer system shaping a narrative in real time for an audience. In the book *Hamlet on the Holodeck*, Janet Murray further imagines and theorizes the dream of interactive drama [74]. Its first recognized instance was the interactive drama Façade [52], which implemented Mateas's theory of a neo-Aristotelian model that emphasizes agency and story. However, few subsequent works have achieved the same level of immediacy and freedom as *Façade/*, especially in view of Laurel's definition. Telltale Games can be considered a very restricted form of interactive drama, one in which the player's agency is constrained to a few possible options for progressing the story. An advantage of such restrictions is that they make the works simpler to author given current production workflows and authoring strategies. These restrictions provide the necessary structure, though, to assess how narrative and drama can be integrated, and they are the subject of this dissertation's descriptive encoding.

2.2.2 Simulation

To understand the decision Telltale Games made in developing their particular approach, it is worth looking at interactive digital narratives that incorporate simulation.

The most famous example of direct simulation is that of *Dwarf Fortress*[4] [101], which uses simulation to create an entire history of the world as well as languages

[4] *Slaves to Armok: God of Blood Chapter II: Dwarf Fortress*

and cultures. The technique of using generation and simulation to provide back-story and context for a narrative is also employed in Ryan, Samuel, and Summerville's *Bad News*. [102], though the story itself emerges as a result of the situation rather than any specific drama management.

McCoy et al.'s *Prom Week* [103], [104] ,[22] simulates social relationships using integers to represent the strength of relationship components labeled "friendship," "romance," and "buddy." These values combine with various effects and statuses to form a puzzle minigame structure whose goal is to motivate play that emulates the types of relationship changes and dynamics found in high school. *Prom Week* positions social relationships as a strategic problem involving a series of steps that rely on a player's intuition about how to achieve a certain goal state. While this gameplay does result in emergent stories and represents a substantial possibility space, it also relinquishes control over the pacing and emotions a player might experience in favor of a more analytical mode of play.

2.2.3 Craft Knowledge for IDN

It is also important to recognize the contributions of the creators and developers of Telltale games. Molly Malone, for instance, gave a talk at Konsole 2017 [40] in which she detailed many of the idioms and insights gleaned from design choices. I incorporate her insight into how the creators conceptualized the design space, and I consider her concept of the "choice space" a particularly compelling metaphor that encompasses not only the components Mawhorter identifies but also the psychological aspects of the dramatic presentation and context. Other insights relevant to modeling are the importance of identifying content that players would care about. I introduce the concept of a "story value charges" and extend McKee's characterization of dramatic movement to encompass a discrete labeling system that captures magnitude, valence and theme for emotional content events as subjectively identified by an annotator. My term specifies the valence and magnitude contribution of individual acts or lines in a beat that enact the change that McKee considers critical to the turning of a scene either from positive to double positive or from negative to positive. Malone also cites two Telltale terms that relate directly to my representation, namely "choice text," which corresponds to option labels, and "feeder lines," which Malone describes as the speaker line or situation immediately preceding a choice prompt. The importance of feeder lines for narrative designers suggests that content has a momentum that occurs before a choice. Compared to alternative models, separating actions within the immediate context from the general framing of choice is one advantage of a finer granularity in categorizing the content and context surrounding dramatically represented choices.

2.2.4 Narrative Research Systems

Peter Mawhorter's *Dunyazad*, as discussed previously, focuses on the effects conveyed by choices themselves. Mawhorter addresses the context of choices and their embedding within narratives, because these features both motivate players to engage in certain modes and generate a variety of interesting choices. Sarah Harmon developed the *Rensa* framework [105] as an intermediary human-friendly language for representing narrative data and processes. Harmon compares this framework's representational capacity directly to that of Story Intention Graphs, and intends for the system to improve access to storytelling technologies. Implemented in Python, *Rensa* focuses on the creation of narrative tools rather than on a system for annotating existing narratives, which is the primary goal of SHERLOCK, described further in Chapter 6. Like *Dunyazad*, *Rensa* focuses on textual narratives and has been successfully adapted for use in several projects. These narrative research systems apply computational methods to represent and evoke our understanding of linear narratives and choices. Both focus on creating generative systems, though *Dunyazad* was developed alongside an encoding that shares several features with the encoding presented in my dissertation. I compare Mawhorter's encoding for goal-based choice analysis to ICE and the SIG in Chapter 5, noting the complementary emphasis of each approach.

2.2.5 Revisiting Cinematic Choice-Based Adventure Games

The subgenre of cinematic choice-based adventure games' was selected because it relies heavily on interpretation instead of simulation or combination, while still incorporating alternate performances and even ontological changes in the story world based on player choice. Such games do not offer a sophisticated simulation of social interactions, complex procedural rhetorics [106], or the enactment of drama managers. Instead, the subgenre implements a straightforward translation of certain narrative effects into an interactive space, thus providing a fairly stable environment for testing the relationship between narrative and interaction.

TWAU has earned critical acclaim for its storytelling. Both game critics and scholars interested in applying alternative methodologies such as microethnography have regarded Telltale's brand of graphical adventure games as noteworthy [107]. The close relationship between story content and interactivity makes this particular game a prime subject for a study using computational narratology. The subgenre's avoidance of emergent narrative and simulation makes it particularly appropriate for the approach described below, that is, developing an encoding for the surface content contained in video recordings of traversals.

In the next section, I will review the formal study and computational representations of emotion. Although the connection between narrative and emotion should be clear, their respective fields of research exhibit only a few clear overlaps.

2.3 Affect

Everyone has an intuitive sense of what emotions are. If you do something to make someone mad, you can expect them to feel angry, and likewise, if something terrible happens to them, you can interpret their solemnity as an appropriate expression of sadness by relating it to the feelings you might have had in a similar situation. Emotions play a significant role in daily life and are consequently an important dimension of artworks, but their relative value has been the subject of debate over the centuries.

A broader term than “emotion,” “affect” incorporates the physiological sensations typically associated with emotions – the flush of neurotransmitters such as norepinephrine or the contentment associated with oxytocin. Emotions play a significant role in decision-making and are a core part of why humans act the way they do. A play, for instance, reverses the emotional implications of a story, because it requires actors to assist the audience in experiencing the necessary affective states.

In this section, I survey theories of affect and their application to computational methods of understanding and measuring emotions, in the fields of both affective computing and games user research. I provide a more detailed review of the specific measures incorporated in my study in Section 3.1.1, including the role that measurable values such as heart rate and skin conductivity play in indicating emotional states.

2.3.1 Emotion

There are two primary classifications of emotion within the affective science research community. The first is a categorical measure, where a person's emotional state can be mapped to a discrete set of states such as fear, anger, happy or sad. The second is a set of continuous dimensions that represent arousal and valence. In addition to classifying the response, the term *elicitor* labels an event or condition that acts as a stimulus and gives rise to emotion. There are several competing models for how emotion arises, of which I will describe the most popular, the appraisal model and relate it to the goal of connecting encoding of interactive digital narrative content and player response.

2.3.1.1 Appraisal Model

One set of theories that have gained widespread support is that of emotional appraisal [108]. The theory focuses attention on structural features of the situation and the agent experiencing emotion. It differs from other theories by emphasizing the dynamic interpretation of the situation of an agent over static associations between emotions and elicitors. It addresses many issues raised by previous theories such as stimulus-response. "A common pattern of appraisal is found in all the situations that evoke the same emotion" [108]. For example, losing a laptop may give rise to either joy or sadness based on the knowledge of whether it was insured, which can further classify the event as a positive or negative valenced outcome.

In *TWAU*, the main game content consists of a sequence of emotionally charged situations, in which the player has severely constrained agency to shape the outcome of events. The player's moment to moment reactions is visible through facial expressions and various other biometric signals (Heart rate, Skin Conductivity) that are indicators of physiological responses. The content is also visible through the gameplay video, providing the context for these responses. The appraisal model suggests that an encoding method that incorporates features from the content (as described in Chapter 5), when compared with the various measurable signs of player affect, could be useful in understanding types of associations and variations between players.

2.3.1.2 Categorical Models of Emotion

These "basic emotions" are families of related states, each of which may include some temporal duration or even a mixture of different emotions. Paul Ekman conducted a seminal study documenting facial expressions from numerous cultures [109]. The emotion categories are used as labels for particular expressions classified by the Affdex Software Developer Kit (SDK), a library produced by Affectiva described further in Section 3.1.2 [110].

Table 2.1: Ekman's 6 basic emotions

Emotion		
Joy	Fear	Sadness
Anger	Surprise	Disgust

2.3.1.3 Dimensional Models of Emotions

As opposed to the categorical model, the dimensional model asserts that there are many continuums on which a particular emotion could fall. The most popular is the valence-arousal circumspect model [111], a two-dimension model of emotion where valence can be either positive or negative. The Affdex SDK outputs several values for each category based on the detection of distinct action units, including valence which I use in this dissertation, and engagement which can be an indicator of arousal along with other biometric signals. Other methods include measuring the activation of facial muscles, though the technique is more invasive and likely inappropriate for gameplay sessions.

The detection and rating of valence are useful as they share features in common with story value charges and represent the perception that dramatically portrayed events and actions have a positive, negative or mixed valence with respect to some stake that represents why they were chosen for inclusion by the authors.

2.3.1.4 Emotions in Literature and Film

It is no surprise that emotions play a major role in the study and appreciation of media. In fiction, for instance, it is often desirable to experience emotions that are not encountered in everyday life and to explore scenarios that are plausible without the associated risks. From an evolutionary psychology perspective, emotions play an instantaneous appraisal role that assists and regulates social behaviors, in particular ensuring that the concerns and goals of a group are sufficiently aligned and providing a means of recognizing when they are not. This explanation extends to media, where the exercise of emotional appraisal in a fictional environment can aid in understanding and responding in real life. In this way stories, regardless of media, provide a form of emotional training for life, as well as a means of experiencing the emotional system's effects [72].

In the next section, I will describe the current intersection of emotional research and computing, with the goal of assessing the player's experience during an interactive digital narrative.

2.3.2 Player Experience

Player Experience (PX) has been positioned as a game-focused alternative to User Experience (UX).

In *How Games Move Us*, Katherine Isbister describes a variety of interactive games to demonstrate the breadth and types of emotions that games can evoke [112]. Isbister's observations about the reality of interactions with NPCs is relevant to my goal of charting player feelings about actions: "Interactions with NPCs move players beyond 'para-social' feelings into consequential social experiences with accompanying social emotions and behaviors" [112].

Scholars have used several terms to describe aspects of the player experience. *Flow* [113] describes the feeling of getting lost in a task, when players often lose track of time. *Immersion* [114] describes the sense of actually being in a situation, which is distinct from the sense of *presence* [115] that is often used to distinguish VR from non-VR experiences. Murray's definition of *agency* [74] is the satisfying sense of being able to take actions afforded by the environment. Researchers have examined the emotions of *surprise* ([116] and [117]), and *suspense* ([118], [119], and [120]) in terms of their relationship to both platformers and narrative gameplay. While narrative games certainly evoke a sense of flow and immersion, as well as surprise, it is feelings relating to responsibility and social interactions that often distinguish interactive digital narratives from player emotional experiences in other types of games, such as pride or satisfaction in solving a puzzle or defeating an enemy. In Chapter 5, I identify several key points and their related actions that lead to a player's feeling of regret and uncertainty about whether giving Faith money led to her demise.

2.3.3 Games User Research

As the economic value of computer games has increased over the last decade, specialized approaches and methodologies have been used to improve their quality and the efficiency of their production. Many of these techniques originated in human-computer-interaction research, which focused initially on improving usability and accessibility. These topics are also relevant to games, because ease of use is necessary in many aspects of games, such as menu systems, but the primary goal is not to be productive in a traditional sense. With the creation of a specific CHI conference (CHI-PLAY), the CHI community has expanded to incorporate the distinct goals associated with game design. The overall approach of measuring and assessing the experience component is of particular importance to my goal.

The thread that distinguishes games user approaches from the more humanistic perspective on player experience is the former's use of the scientific method to verify results as well as its use of biometrics and telemetry. Ravaja et al. investigates the association of specific gameplay events with physiological measures [121] [122]. The increasing availability of affordable biometric sensors has supported the increased use of computational methods to measure and classify emotional states [123]. Nacke's dissertation [124] is closely related to mine, though Nacke focuses on the relationship between subjective measures such as questionnaires and objective psychophysiological measures, whereas I focus on the relationships between narrative-specific content features and psychophysiological and facial expressions. Nacke calls the intersection of game studies and psychophysiological measures "affective ludology" [125].

2.4 Summary

Documenting, modeling, and encoding a set of player traversals requires insights from many fields. This chapter reviewed studies of narrative modeling, interactive digital storytelling, and games user research. This review provides the background necessary for developing a model of how content structures can be related to player response, and it contextualizes the necessity of a tool for visualizing and annotating player response data.

The approach described in the rest of this thesis addresses challenges identified in the fields mentioned above. In the previous chapter, I argued for the significance of the emergence of a subgenre with properties that made it attractive for descriptive modeling, as well as a source of insight into player response. In this chapter, I described efforts to advance our understanding of narrative, interactive narratives, and the player experience through models and systems. The success of corpus-based approaches suggests that a corpus and model approach can be applied to the content and responses present in the subgenre. Finally, the availability of player traversals in the form of streaming video records motivates the need for additional tools with which researchers can develop and evaluate new theories that build on player behavior and player response data. The availability of libraries that can extract data relating to the player experience makes the approach useful to a broader range of scholars. In the next chapter, I will describe the initial player study, which collected a corpus of traversals and data that was essential to the model development.

Chapter 3

Player Experience Study

Player experience is one of the main reasons for playing a game and determines its success. The field of games user research is increasingly relying on data collection and analysis to assess and guide design decisions. This chapter describes a player experience study that involved collecting player traversal videos from seven participants who played *The Wolf Among Us* (*TWAU*) by Telltale Games. The study design and the collection of the data was conducted collaboratively, as well as some of the initial annotations. I developed an annotation schema that is based on narrative features suggested by story craft and narratology and used it to annotate portions of the dataset.

3.1 Summary and Overview

The study design and the collection of the data were conducted collaboratively, as well as some of the initial annotations. In this chapter, I describe an annotation schema based on narrative features suggested by story craft and narratology, and I use it to annotate portions of the dataset. The study was designed to obtain a dataset that could be used for a variety of purposes. The collection and initial results were promising but revealed several shortcomings in the available tools and the initial annotation schema.

The approach provided evidence supporting the schema's usefulness in detecting content preferences among the players, and it further provided a means of understanding the overall emotional flow of the work by aggregating the different measures. By manually inspecting the videos, including reviewing the think-aloud and watching the gameplay itself, I confirmed that the features indicated were significant. A vast number of potential interactions exist among the content, context, and player physiology, and these could cause variations in the observed signals.

However, the process of annotating the data was labor intensive. Once tabulated, the annotations were challenging to associate back to specific features in the original footage, as the process used to calculate the results (Google Sheets) was not connected with a video player that could directly move to the relevant locations. These challenges motivated the formal encoding schemata described in the next chapter (the Interactive Cinematic Experience ontology) described in Chapter 4. I also developed a web-based visualization tool that addresses the identified challenges in annotation and analyzing related traversals for non-experts in 6.

My collaborator and I (hereafter, we) were interested not only in studying the relationship between narrative contexts and player responses but also in comparing the evaluation methods used [126].

3.1.1 Measuring Affect

Affective computing has provided methods to assess emotions through various sensors and processing of facial videos. In this chapter, I review several key measurements used in the player study and that are present in the dataset, as well as the tools and methods associated with them.

3.1.2 Facial Action Units

One of the primary ways by which computing systems can understand and classify facial expressions is through an existing taxonomy known as the facial action coding system (FACS). A Swedish anatomist, Hjortsjö [127], developed the original taxonomy, which Ekman and Friesen adapted and published in 1978 [128]. The system was updated in 2002 and has served as a gold standard for research on automated facial recognition and emotion classification [129]. The reasoning behind using visual images and movies to classify emotions is that humans are capable of recognizing emotions based on pictures alone and that these expressions are the result of combinations of common facial muscles. These combinations are called action units (AUs).

Two systems are available to analyze player facial expressions, and both are incorporated into the dataset. The Affdex SDK is a toolkit that incorporates trained models that have been created using open datasets of images of faces with various features labeled, such as eye position and facial feature positions. The second library, OpenFace by Baltrusaitis et al. [130], provides a set of algorithms for training a classifier, as well as a set of pre-trained models based on openly available facial expression datasets. Both systems can identify a subset of AUs, described in Table 3.1. One of the advantages of the Affdex SDK is that it provides a ready set of classifications for the traditional categories of emotions, and can be readily incorporated into other software through a library.

Raquel Robinson et al. incorporated the openly available Affdex SDK system [110] into All The Feels (ATF) used during the study [131]. The set of action units used to classify each emotion is shown in Table 3.2, from the iMotions website [132], which uses Affdex to process emotional recognition. The classification method is based on EMFACS, the emotional facial action coding system developed

Table 3.1: Facial Action Units Recognized by OpenFace and Affdex SDK

AU	Description	OpenFace	Affdex
1	Inner brow raiser	X	X
2	Outer brow raiser	X	X
4	Brow lower	X	X
5	Upper lid raiser	X	
6	Cheek raiser	X	
7	Lid tightener	X	
9	Nose wrinkler	X	X
10	Upper lip raiser	X	X
12	Lip corner puller	X	X
14	Dimpler	X	
15	Lip corner Depressor	X	X
17	Chin raiser	X	X
18	Lip pucker		X
20	Lip press	X	X
23	Lip tightener	X	
25	Mouth open	X	X
26	Jaw drop	X	
28	Lip suck	X	X
43	Eyes Closed		X
45	Blink	X	
	Smirk		X

Table 3.2: Categorical Emotion Values provided by Affdex SDK

Emotion	Action Units
Joy	6+12
Sadness	1+4+15
Surprise	1+2+5+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+16
Contempt	12+14(one side)

by Friesen and Ekman [133].

3.1.3 Measuring Body Responses

The results over the past decade clearly show that the relationship between mental processes and physically measurable responses is not straightforward. Nacke describes the relationships between mental processes and bodily responses as being either one-one, one-many, many-one, many-many, or null [123]. These classifications characterize how one system may influence or be indicative of either a single or many mental or physiological processes, such as those described in the following sections. Nacke further recommends starting the process of selecting measures with electro-dermal activity (skin conductivity [SC]), as it has a direct relationship with arousal and can represent responses in short time windows.

3.1.4 Heart Rate

Some physiological response measures relate to the central nervous system. Many hospital dramas rely on omnipresent monitors of vital signs to set up the scene of a patient near death. Beyond simply indicating physiological arousal, heart rate (HR) can be a valuable indicator of mental workloads. Heart rate variability, a measure of the variation in the beat-to-beat interval, can indicate increased mental workloads and stress [134]. Heart rate can also be measured through changes in skin tone, which may indicate a means of measuring the heart rate of players who were not previously attached to sensors, such as streamers [135]. Heart rate measures, however, can be affected by other activities, such as think-aloud [123], and so possible sources of variability must be accounted for.

3.1.5 Skin Conductivity

Skin conductance level , also shortened as skin conductivity or SC, is a physiological measure of electric resistance across the skin associated with the polygraph (lie detector) test. It can be measured simply and affordably with two electrodes, often attached to the fingers. The idea is that certain microscopic variations in SC result from the responses of the body's sweat produced by the *eccrine* glands as a result of changes in neurotransmitter levels associated with psychological arousal [136]. These variations can be measured for certain stimuli, in which case it is called galvanic skin response, or can be measured over time, in which case it is called electro-dermal activity (EDA). In this thesis, these are generally referred to as SC, though the usage is more an indication of EDA. The delay of a response to a stimulus is approximately five seconds. The measure tends to drift over time as sweat accumulates, which suggests that a shorter time window average might be of use ([123], [137], [54], and [138]). Examples of events that may cause detectable changes include the nervousness one feels when on a first date or when telling a lie with high stakes. In using EDA to obtain features, [139] find local minima as points of interest, as a decrease in the SC is considered a measure of an increase in sudation, "which usually occurs when one is experiencing emotions such as stress or surprise." Its use in game design evaluation has been primarily in high temporal resolution detection of events. An example of the use of GSR in an application is an automated indexing system for an audio stream [140]. For a comprehensive summary of the brain and its relationship to research in play, see [141].

3.2 Study Design

To understand how player responses were related to the content, my collaborator and I organized a study. This study consisted of seven uncompensated graduate students whose age ranged from 20 to 35. Each participant was instructed to play through the first episode of *TWAU*, wearing an Empatica E4 wristband sensor, and to speak any thoughts aloud. We collected the data through the output ATF, described further in Section 3.1.2, which uses the sensor data from the wristband to visually present the state of a streamer in the form of bars and emoticons. We also encouraged the participants to use the SEI [142], [143], designed for nonverbal self-report, shown in Figure 3.1. Because of some technical issues, only six player traversals are included in this thesis (noted as P2 to P7).

3.3 Sensual Evaluation Instruments

We decided to collect a variety of measures of player emotions through traditional measures , such as think-aloud and surveys, as well as lesser-used methods, such as the SEI. The SEI functions as a non-verbal, cross-cultural self-report method for augmenting physiological and qualitative measures by allowing participants to express their feelings using a set of sculptures developed by Isbister et al. See [142] for an example of its use in a narrative game context. The SEI consists of eight sculptures that are designed to be touched, held, or gestured with (see Figure 3.1 for the shapes and names of each). Players interact with these sculptures to indicate some internal feeling was taking place, and these moments serve both as a self-report of their experience, as well as anchors for discussion during a post-game interview.

We also hypothesized that the SEI would provide insights into the more subtle

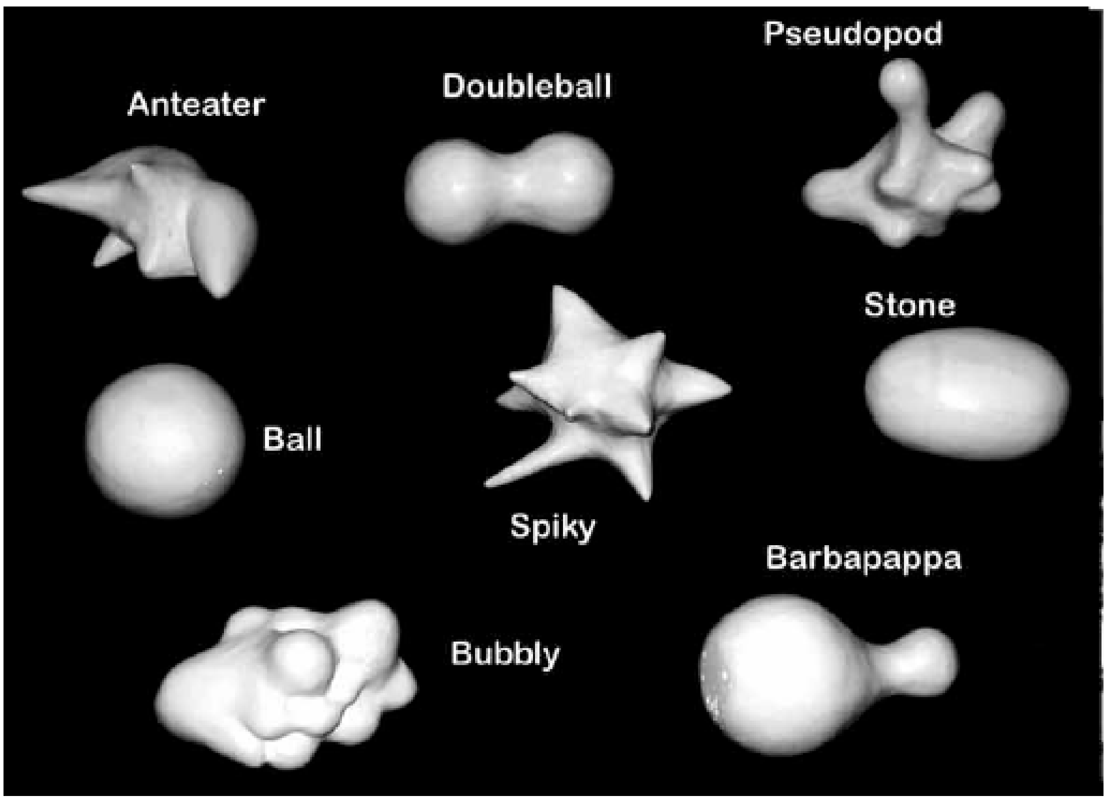


Figure 3.1: SEI Models

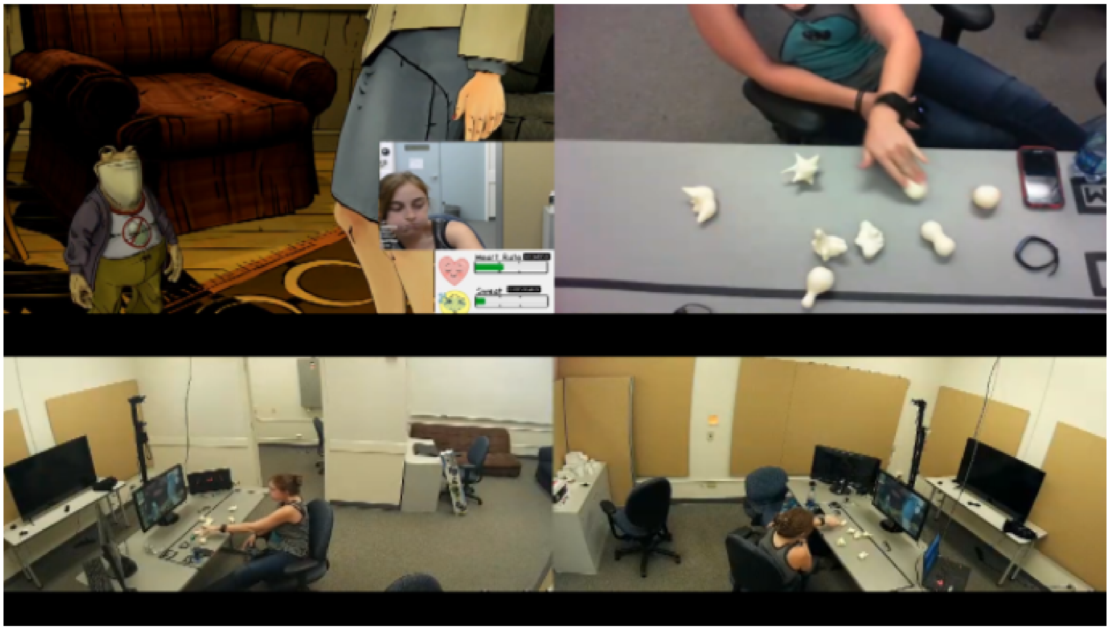


Figure 3.2: Video Recording Setup. The top left is the gameplay video along with the ATF display, the top right is an overhead view of the usage of the SEI, and the bottom left and bottom right are the other views of the space and game screen.

emotional experiences that were prevalent in this type of narrative game experience, indicating points of interest that may not be otherwise evident from the biometric signals alone. For this analysis, I treated the SEI as an indicator of the presence of some player affect, without differentiating between each instrument. As I will demonstrate, this is useful in identifying locations of emotion through frequency measures, although multiple uses of the SEI within a short period may indicate ambiguity of player feelings rather than increased intensity. For this chapter, any SEI usage is considered an affect event (which, together with other events, can be combined into an affect signal) and is undifferentiated by the particular instrument used. Attending to patterns that relate to specific instruments or relating instruments through their usage patterns is a topic of future work.

3.3.1 Protocol

Each session was conducted as follows. The players first took a pre-game questionnaire, with questions covering familiarity with the *Fables* comic series and whether they have played other titles by Telltale Games. As a prerequisite to participate in the study, we required that the participants should not have played the game before so that their reactions would be initial traversals of the game content. However, we learned afterward that one participant was familiar with the comic series on which it was based, coloring this participant's knowledge of the final scene. We then calibrated the SEI by showing the users a series of 10 photos from the International Affective Picture System [144] and having participants indicate with the SEI which association came to their mind for that particular instrument. This was used for familiarizing the users with the process of reporting affect using the instruments instead of recording specific reactions to particular stimuli. We then instructed the participants to play the entire first episode while expressing aloud any thoughts that emerged, as well as encouraged them to use the SEI as much as possible. We found that some early participants (before P2) were engrossed in the game and forgot to gesture with the SEI, so the usage varied. After noting this, we set a timer to go off periodically and serve as a reminder, although this was not as much of an issue with the other participants. We left the room for the duration of the gameplay session, only returning in case of technical difficulty.

The facial recognition software, Affdex, from ATF was also running and is displayed along with SC and HR readings in the lower right-hand corner of the video. At the end of the session, we conducted a retrospective think-aloud with the players about how they felt at certain peak moments of the game, how they felt about various characters, and why they used specific SEI objects during the game.

They also took questions from the Immersive Evaluation Questionnaire [145]) and combined them with questions relating to emotion. These provided broad strokes of the players' retrospective assessment of their experience. However, these are not included in this thesis, as I instead focus on moment-to-moment measurements and annotations.

Once we finished collecting data, I processed the gameplay screen capture and the external cameras' videos into a single video and adjusted the timing to synchronize them by using the sounds and movements shown in the videos. I used Adobe Premiere Pro to create a single video for each participant containing the gameplay video (with ATF data as a picture-in-a-picture), the top-down view, and left and right room cameras (see Figure 3.2). We both used the marker feature of Adobe Premiere Pro to hand-annotate the features described in the next section because of the sheer amount of footage, totaling over 13 hours. These annotations were exported as a comma-separated value file and tabulated with Google Sheets using a sheet for each data table.

The annotations consisted of features, including beats, choices, and affect events. My collaborator identified peaks in HR and SC through visual inspection. Choice prompts are locations in the game where a set of choices associated with controller buttons is displayed, often only for a brief time. I created a text annotation consisting of the selected choice, the choice texts, and marked each point when a player either pressed a button (indicated by a white highlight of the corresponding option) or when the timer expired.

At this point, I used the location of the annotations of choices as a basis for a set of time windows of eight seconds before or after each choice. These were used to select the presence of the affect annotations using a formula. Next, I charted various statistical measures of the choices, paying attention to the distributions

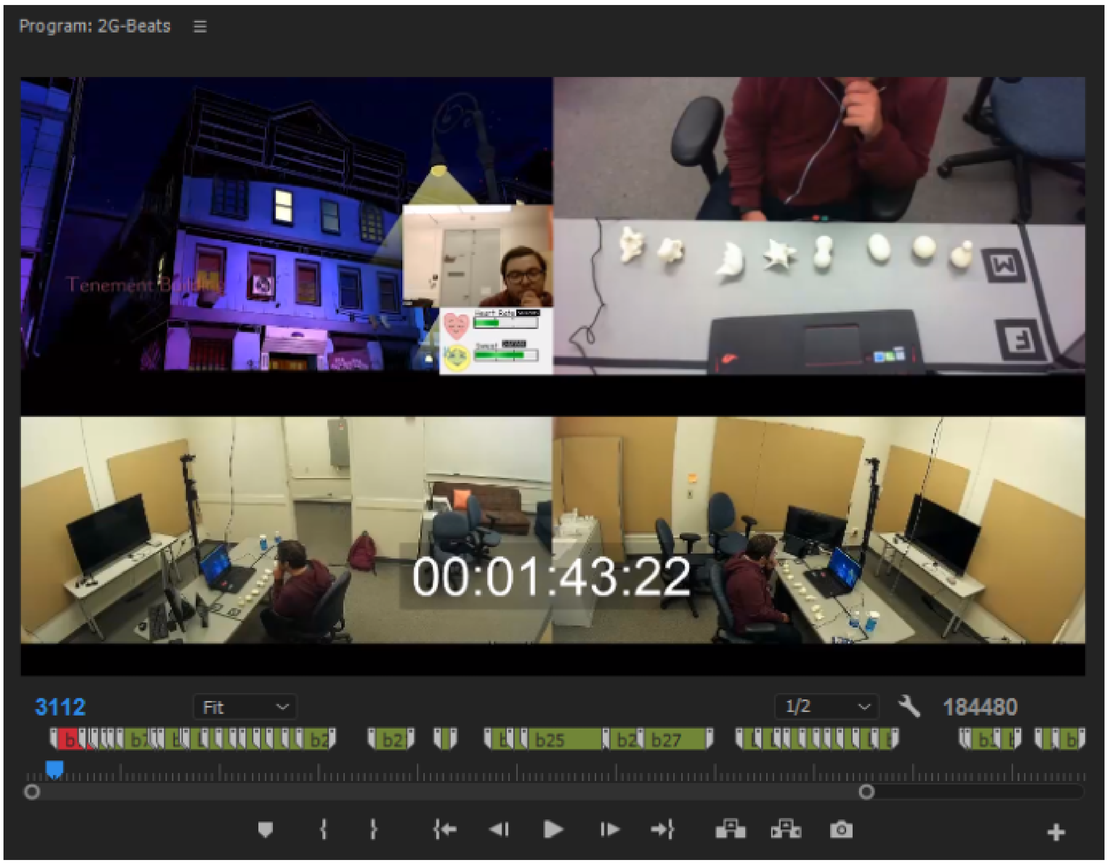


Figure 3.3: Beat span annotations in Premiere Pro

of the choices themselves. The results of these are detailed in the Section 3.6.4.2.

Using pairs of videos in a side-by-side configuration, I identified and annotated segments of each traversal as containing specific dramatic beats, described further in Section 3.4.1. The annotation was completed by identifying the particular shot or frame associated with the start and end of the beat and matching it. The next section presents and discusses the initial features used to annotate the narrative elements of the playthroughs.

The multi-channel nature of the data necessitated that the videos be synchronized to attribute response measures to particular elicitation events or to determine that they were a part of a non-game related activity, such as interruptions. These data can be divided into player response data and game content data. The narrative elements selected for analysis include the dramatic beat, story values, and character. Refer to Section 1.1.3 for a summary of the plot and story.

3.3.2 Expected Results

TWAU has many emotionally charged events in which players might react, as well as ambiguous outcomes for many of the choices provided. The game takes advantage of the fact that the characters have pre-existing relationships with the player character and his role both as a sheriff and as the eponymous Big Bad Wolf.

I anticipated that the content segments would contain the most instances of affect (excluding events that require carefully timed player input, also called QTEs, which, through their nature, evoke stress and an increased heart rate). I also predicted that there would be common patterns among all players in their emotional involvement over the sequencing of beats, as the story is relatively fixed in its composition of events.

One of the more surprising results was the amount of variation in the expressiveness of each of the players, which necessitated the normalization process to account for the different quantities of expressive events while facilitating comparison of like content. Some of the variations in SEI usage were more prominent than physiological responses were. While some events consistently caused players to report feelings, many of the spikes in usage observed occurred during different segments. The approach used in this chapter allows the consideration of these variations with more than a pair of player videos by aligning the data to the content.

In the next chapter, I will describe the annotation schema, followed by the results and the discussion.

3.4 Narratologically-inspired Annotation Schema

TWAU has three types of gameplay. The first is choosing between narrative options. The second is quicktime events (QTE), which are button-pressing sequences that require fast reaction times, usually corresponding to a violent or risky event and an on-screen prompt. The third is free-roam exploration coupled with hotspots, most akin to classical adventure games, where the player can move the player character about a region of interest and engage with elements of the environment. I decided to base my analysis of the narrative content on the story beat and focus primarily on the segments including the narrative choice type of gameplay. The story beat is also used as an organizing "content piece" in the interactive drama *Façade* [146]. The beat functions as a means of segmenting character interactions using behaviors and values. Although McKee's book is not grounded in a scholarly tradition, his work has inspired a number of approaches to narrative and writing in the interactive digital narrative community. I also selected a set of key story events that I believed to have the potential for player response and which I believed to be directly related to the primary plot. These provided spans of time that I used to focus comparisons of each player-decided narrative traversal and that served to align the player data. I also annotated interruptions and excluded those periods from the analysis, although the periods of interruption may have an effect that isn't accounted for in terms of the duration of the interruption and the return of a participant's arousal levels to a neutral state. The following sections describe each annotation feature in detail.

3.4.1 Beats

McKee defines the "beat" as a unit of action [48] that involves one character initiating a behavior and another character or environment reacting. The classic example is one character using a strategy to gain something that they desire while another character either deflects or concedes. In *TWAU*, one tense scene involves a character (Ichabod Crane, deputy mayor) savaging the player character with blame (beat #31). This beat ends when the first character turns their attention to another subject and ceases expressing blame. Each beat is further annotated with a story value, which is a label that describes the overall values at stake in the story. It can be considered an empathetic label, as it is primarily based on the reception of the story material by an audience, rather than purely the concerns of the characters involved, and so is distinct from systems which look purely from an agent's perspective. The story value I assigned to this particular beat as a binary-opposition is ego/community, as Crane is attempting to save face at the expense of the community by laying blame on others. During that beat, the receivers of his blame react: in the case of the beat, this reaction is in part determined by player choice. Bigby can avoid blame by staying quiet, assume blame from Snow, or try to redirect the conversation.

I analyzed the first three chapters of *TWAU* for the location of individual beats. I identified a total of 42 beats, listed in Table 3.3. Other annotators may select different segments for beats, but these beats served the desired purpose of segmenting content and aligning player data on characters and story values. The criteria used for designating a dramatic exchange as a beat was whether a single behavior was being pursued by a character and was directed toward another character. A new beat begins when either the agent of the action or the behavior itself changes.

Each beat was further tagged with the characters involved, including the character taking action and the responding character, as well as the story value(s) at stake. I found, as expected, there were recurring themes of truth/lies, justice/injustice, and community/ego. The idea of a story value pair was introduced by McKee. His idea was that during each unit of action, some movement between a positive and negative valence of a story value would occur, such as from negative to positive, or from positive to double positive. The story value of justice, for instance, is invoked initially when Bigby visits Toad's apartment, as Toad is getting justice for the disturbance upstairs. The value swings directly to negative, however, when he realizes that he is breaking the currently unjust rules of Fabletown by not being in glamour. This variation from one side of value to another is often imprecise and dependent on context, but it generally recognizes the emotional nature and thematic unity of actions in a scene. The dynamics caused by value shifts keep the reader's interest, and the cause of the change is often the conflict itself.

In this initial annotation, each story value pair is used to describe the main values that are at stake in a particular beat. This classification approach doesn't

Table 3.3: Dramatic Beats and Cast in Chapters 1, 2 and 3

ID	Description	Cast
1	Toad discovered not being in Glamour	Bigby;Toad
2	Toad asks Bigby to deal with the Woodsman	Bigby;Toad
3	Toad yells at TJ to get back in his room	Toad;TJ
4	Woodsman asks Faith who he is, beating on her.	Woodsman;Faith
5	Faith spits blood on Woodsman	Woodsman;Bigby;Faith
6	Bigby asks Woodsman why he did it	Woodsman;Bigby
7	Woodsman attacks Bigby, who defeats him.	Woodsman;Bigby
8	Bigby asks Faith to leave	Woodsman;Bigby;Faith
9	Woodsman threatens Faith and refuses to pay	Faith;Bigby;Woodsman
10	Woodsman insults Faith	Bigby;Faith;Woodsman
11	Bigby destroys Toad's car.	Toad;Woodsman;Bigby
12	Toad laments car and Bigby's help	Toad;Bigby
13	Woodsman attacks Bigby	Woodsman;Bigby
14	Faith loots Woodsman, Bigby's transforming	Bigby;Faith;Woodsman
15	Faith kicks Woodsman then pushes ax in	Bigby;Faith;Woodsman
16	Bigby gently questions Faith (Cigarette)	Bigby;Faith
17	Bigby Asks about Faith's employer	Bigby;Faith
18	Faith stops Bigby from pursuing Woodsman	Bigby;Faith
19	Bigby asks about obligation; Faith demures	Bigby;Faith
20	Faith touches Bigby's wounds; plans meeting.	Bigby;Faith
21	Beauty asks Bigby to cover for her	Bigby;Beauty
22	Beast asks about Beauty's whereabouts	Bigby;Beast
23	Bigby asks Colin to get out of chair	Colin;Bigby
24	Bigby provides Colin a smoke	Colin;Bigby
25	Colin entreats Bigby to give him a drink	Colin;Bigby
26	Snow retrieves Bigby	Snow;Bigby
27	Snow asks Bigby how he knows Faith	Snow;Bigby;Faith
28	Snow wants to know Bigby's early impressions	Snow;Bigby;Faith
29	Snow wants to tell Crane	Snow;Bigby;Crane
30	Grendel is upset that Bigby isn't more helpful	Bigby;Grendel
31	Crane is attempting to assign blame	Crane;Snow;Bigby
32	Crane blames Bigby and Snow; asks for leads	Crane;Bigby;Snow
33	Asks for massage and wine before storming out	Crane;Snow;Bigby
34	Snow stops Bigby from provoking Crane	Snow;Bigby
35	Bufkin says hi	Bufkin;Bigby
36	Snow asks for help in researching identity	Snow;Bigby;Bufkin
37	Mirror accuses Bigby of being Cruel	Mirror;Bigby
38	Bigby can't read it, but Bufkin offers to help	Bufkin;Snow;Bigby
39	Bufkin reads story of Faith	Bufkin;Snow;Bigby;Faith
40	Declare Husband person of interest	Bigby;Snow;Lwrnc.
41	Finds out Husband is in trouble	Bigby;Mirror;Snow;Lwrnc
42	Toad calls about intruder	Bigby;Snow;Toad

capture finer details of the how values manifest in a scene. For instance, the portrayal of values in a scene is often characterized as moving from one “charge” to another, either from a positive to a negative. The annotation schema avoids this level of annotation and instead focuses on which binary pair is associated with a particular beat.

The annotation process was applied for sequences involving dramatic choices. Some of the quicktime segments were integrated into a dramatic beat, such as the first fight sequence (Bigby and the Woodsman fighting), which was consolidated into a single beat with the Woodsman as the agent initiating the action (starting the fight). There are segments of content that are dependent on players selecting hotspot while exploring areas. Sometimes, upon selecting an item or option, a character will interact with the player character based on the object selected, as is the case in the business office scene and the final scene with Grendel. These beats are significant in that they are not always required to progress, which means they may require a different strategy than the simple annotation schema described in this chapter. One possibility is a graph-based approach, which accounts for the dependent relationship between beats and various dependencies. These beats are significant in that they are not always required to progress, and this optional nature is revisited at the end of this chapter and in the schemata described in the next chapter (Chapter 4).

In this initial annotation, each story value pair is used to describe the main values that are at stake in a particular beat. This classification approach avoids some of the finer details of the usage of values. For instance, the portrayal of values in a scene is often characterized as moving from one “charge” to another, either from a positive to a negative. In the next iteration described in the next chapter, I propose a finer-grained annotation to story values, “story value charges,” and associate them with individual content segments corresponding to actions that take place on screen. These actions could take the form of speech acts, physical gestures or actions, or even spectacles with no clear agent. For more details, see Section 4.3.2.4.

3.5 Choices

Dramatically presented choices are the central game mechanic in this kind of storygame and consist of two or more options shown to the player for selection. In timed prompts, the default option is chosen if the timer expires. It is often silence but can also be some form of inaction. For instance, when Bigby is presented a choice of whether to give Faith money after noting she would be short, the timer expiring indicates that Bigby chose not to give her any.

I developed a simple annotation schema that records the location and content of "choice prompts" in the game as integers representing the frame location from the beginning of the file. I assigned each choice prompt an ID (noted as "ChID" or "ChoiceID" in the database records) based on having the same choice text (the labels displayed to the player as part of a prompt). In non-dramatic segments, I encode these with a separate annotation type, indicating that they are menu options, as they play a different role than dramatic, timed choices. They are usually part of some process to obtain exposition (plot points delivered outside of a dramatic context), such as interactions with the Magic Mirror that revealed Prince Lawrence's being in trouble.

I develop this idea of annotating choices, options and structural relationships between content further in the next chapter in the ICE schemata, associating specific story value components with more detailed representations of their relative impact and associating them with specific content events such as dialogue acts or speech acts. See Chapter 5 for an example encoding.

3.5.1 Story Events

Table 3.4: Story Events

ID	**Event Description**
1	Bigby Catches Toad Out of Glamour
2	The Woodsman Hits Faith
3	Faith Rescues Bigby
4	Bigby Responds to Beast
5	Bigby (doesn't) Give Colin a Drink
6	Faith's Head is Discovered
7	Faith's Identity is Revealed
8	Prince Lawrence is Shown in Mirror
9	Toad Calls for Help
10	Prince Lawrence is Informed/Discovered
11	Bigby (Fails to) Catch Toad in a Lie
12	Snow White Confides in Bigby
13	Grendel Insults Snow
14	Grendel Fights Bigby
15	Bigby (Doesn't) Remove(s) Grendel's Arm
16	Snow's Head is Revealed

Another concept which McKee uses but which has a more general meaning in narratology is the story event, also known as a plot point, where an irreversible change that propels the story forward occurs. The story event can include information revealed, significant actions taken, and revelations about other characters or situations. Some of these events invite player participation in their outcome, but the presence and content of the events are mostly fixed. I selected the events listed in Table 3.4 that describe the plot of the first episode. Each propels Bigby forward in his story. These usually are the key point of a scene, and so beats usually precede a story event or build up to one. For instance, the first scene between Bigby and Toad reveals the nature of glamours as a means for non-human residents to pass as humans and that it is against the rules to be out without having one. This sets up one of several reasons behind the class resentment that leads to Grendel provoking and attacking Bigby in the climax.

3.5.2 Lenses

Combining objective and subjective measures into a single analysis requires elements that relate them to one another. I developed a set of low-level features that enable parts of the player experience to be analyzed. Interpretive lenses can be used to guide the application of these feature sets. I applied these lenses to analyze the first three chapters of the game, "Disturbance," "Woodlands," and "Mirror Mirror." The fourth chapter provides two variants based on player choice at the end of the third chapter, which presents a challenge for linearly analyzing beats. The content is largely the same between playthroughs for the first three chapters. The following sections go into more detail into how these features will be used to reveal the relationship between players and the story, as well as defining the objectives for the various charts and calculations that I employ.

3.5.2.1 Story Values

Table 3.5: Story Values

Value	**Associated Beats**
Community/Ego	4,5,7,9,11,13,15,31,33,34
Survival/Bravery	8,14,19,
Truth/Lies	6,21,22,26,27,29,36,
Justice/Injustice	1,2,10,12,16,17,25,30,41,42
Duty/Hedonism	32
Other/None	28,40

The first perspective looks at a story as a set of values expressed through the discourse. According to McKee, scenes "turn" on a value change. Each value is expressed as a binary opposition, where each beat revolves around the dominance of one or the other. The principle value-pairs in the first three chapters of *The Wolf Among Us* are listed in Table 3.5. These represent important motivations for both the characters in taking actions and the audience in caring about them and are present in virtually every dramatically portrayed scene that includes conflict. A player's affinity for particular types of content can be approximated by comparing their response during beats associated with story value pairs.

3.5.2.2 Character

Table 3.6: Cast Involvement in First Three Chapters

Cast Name	Beats Present
Bigby	All but 3
Toad	1,2,3,11,12,42
TJ	3
Woodsman	4,5,6,7,8,9,10,11,13,14,15
Faith	4,5,8,9,10,14,15,16,17,18,19,20
Beauty	21
Beast	22
Colin	23,24,25
Snow	26,27,28,29,31,32,33,34,36,38,39,40,41,42
Crane	29,31,32,33
Grendel	30
Bufkin	35,
Mirror	37,41
Lawrence	40,41

Some characters may be more compelling or resonate with players more than others. Toad's plight as a disadvantaged member of the community is overshadowed by his acrimonious attitude toward Bigby, leading to the suspicion that Bigby (and the player) expresses in some scenes. Grendel's valid grievances are hidden behind a wall of rage and disrespect. Faith and Snow are both sympathetic and conflicted. I investigated ways to estimate individual players' feelings during scenes in which characters are involved (if not physically present). In this phase, I did not include either the valence or the intensity beyond measuring the frequency of occurrences, although other measures of emotion may shed additional light.

3.5.2.3 Choices

There are two ways I analyzed the choices and the selections players made. First, I analyzed the decisions players made for agreement (e.g., whether many or all players made the same decision when presented with the same choice prompt). This provided a rough way of determining if there were perceptions of different outcomes and also was a way of finding segments of gameplay where player experience diverged due to either preference or the experience itself.

These three perspectives represented different layers of narrative engagement:

1. **Story value**: specific labels associated with the stakes for which a character is fighting and that make players relate to said characters.

2. **Characters affinity**: how some characters may be more or less engaging to the player than others.

3. **Decisions and interactivity**: where the recorded player actions may or may not line up with other players, including choices and decisions.

In the next section, we will review the results of these perspectives, while a full encoding of a sample scene is demonstrated in Chapter 5.

3.6 Results

In this chapter, I describe and discuss the results of applying the annotation schema to the first three chapters and present a preliminary analysis.

The annotations, which we recorded through analysis of individual videos, were imported through the CSV files into a Google Sheets document. Each table corresponds to a sheet. The calculations were performed by a combination of table formulas and references, often using selections based on values stored in temporary tables that recorded preliminary data. For instance, one such table summed the number of events in a given category along columns for each of a set of segments based on the start and stop point for a given player's traversal.

The analysis process starts by assessing the relationships of each player using the lenses to filter the player affect events into bins based on either time or content segmentation, which are then displayed as charts. These charts fall into two categories: segment-oriented charts and content-oriented charts. The segment-oriented charts provide a means to interpret player responses associated with specific segments, in this case beats, across multiple response measures. The content-oriented charts take a particular content annotation, such as values or characters, and represent evidence for a player's affinity for engaging or responding during content segments that feature that element (such as a particular character, value pair) over others. Although a direct causal relationship cannot be established from these measures alone, they provide a means of assessing the large number of annotations with respect to different aspects of the narrative, and serve to focus attention on areas of interest in the larger space of the content recorded. These charts also facilitate comparisons of larger trends among players over the entirety of the dataset.

3.6.1 Normalization

Each player has a different amount of expressiveness for a given type of affect event. This makes it difficult to compare players on an even footing, so I developed a formula to represent a "normalized" value that was scaled according to the rate found in the dataset. Since the normalization is novel, I will use two examples to describe the process and the rationale. The first uses sample data, whereas the second uses actual data. The first is shown in Figure 3.4, where there are two players. The diamonds represent an expressive event, whereas the rectangles indicate segments of interest. For Player A (PA) , three events occur within the segments of interest (labeled 5 min and 5 min), whereas Player B (PB) exhibits 9 events during the same labeled segment. The difference is both in the amount of time each set of segments occupies as well as the total number of events. For PA, this calculation results in a ratio of 1/3, whereas for PB this is 1/2. This indicates that PB had a higher concentration of events in the period (1/2) to PA (1/3). The comparison of these ratios indicates relative concentration of events in the corresponding segments of time, even if the segments differ in length.

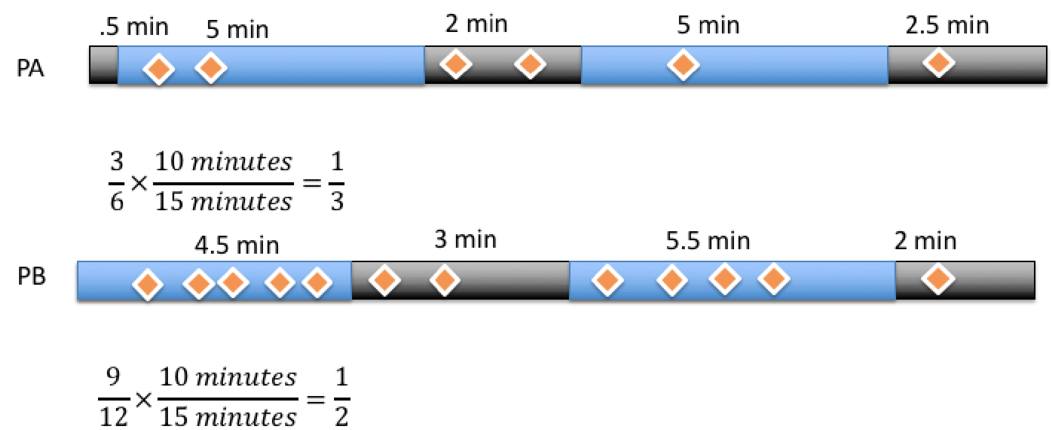


Figure 3.4: Normalized ratio of affect events to time

For the second example, I will use actual data. Player 2 (P2) triggered 60 facial expressions from ATF (labeled as valence events) during beats in which Faith was present. These beats occupied 23,831 frames, or 13 minutes and 14 seconds. The total time measured, however, was 96,134 frames (the total number of frames from the start of the video until the end of beat 42). Beats in which Faith was involved occupied 24.79% of the time measured. During the total period, P2 had

177 expressions recorded. This means that P2 presented 34% of their expressions during a beat where Faith was involved, whereas the beats only occupied 24.79% of game time. Multiplying these two percentages produces the dimensionless scaled ratio of the number of events that occur during the time windows to the total number of events, scaled by the proportion of time the event was measured by the total amount of time considered. The time considered in the denominator excludes interruptions for players that have them, as it only looks at time in which the player is engaged in the game.

The formula for the ratio is:

$$\frac{events(Faith)}{totalEvents} \times \frac{time(Faith)}{totalTime}$$

The ratios can then be compared to one another as they represent relative preferences compared to the total number of expressions per unit of time, rather than absolute counts that may be skewed by either the total number of events or the total amount of time that the given content occupied.

An example of a short scene with an important impact on player response is the introduction of the antagonist, Grendel, which may fall by the wayside without this step, as the beat occupies a relatively small amount of time compared to the total time measured.

Table 3.7: Sample Database Records

ID	PID	Type	ID	ID1	ID2	In
1-2	3	1	1-2	1	2	5857
1-3	3	1	1-3	1	3	6372
1-4	3	1	1-4	1	4	6369

The quantitative data consists of a database that relates each of the inter-player content types to individual playthrough video timelines. Annotations are identified by type, with types including the content types such as the chapter and beat types, as well as affect reports (think-aloud, SEI usage, facial expressions,

Table 3.8: Sample Database Records (Continued)

ID	Ch	A	B	X	Y
1-2	0	I don't make the rules	Not my problem	[...]	Get it fixed.
1-2	0	Do it yourself.	I'm heading up.	[...]	Why's he so pissed?
1-3	2	So what have I walked into?	[Head Upstairs]	[...]	What do you want me to do?

and HR/SC peaks). Each annotation represents this type as an integer (1-12) and an ID if present. These are separated into parts by the hyphen for choices, with the first part indicating the scene and the second indicating the choice sequence within the scene. For choices, it was important to correlate choices with the same set of options even if a choice was present in some traversals and not others, and so these are assigned the same id by identifying shared choice options. This enabled identifying shared choices between traversals. Each row contains associated data, such as the text of the choice-prompt and the decision the player made (See Fig. 3.7). Several columns were only used by some of the types, such as the `out` column, which was useful for beats but was not used for choices. Choices were associated with the time at which the player made a selection or when the choice timer expired. Secondary tables collect the independent items, including the complete set of beats, choices, characters, and values. These are used to select related events according to the time windows for each narrative content type in a relational-database fashion.

For this initial study, I focused on the first three chapters in the first episode of *TWAU*, which amounts to about the first hour of gameplay and is summarized in section 1.1.3. The three chapters provide a classic introduction to a mystery, with the player character being called to the scene of a disturbance only to meet and become attached to the victim. Bigby, the sheriff of Fabletown, is confronted with his past as a violent character and is affected by the violent death of the just-introduced character named Faith. The third chapter describes the community context of the leadership of Fabletown and the core relationships that define Bigby, particular those with Snow and Colin.

3.6.2 Participants

Due to technical difficulties, for this analysis, I excluded P1's data, as the physiological measures and camera data were not sufficiently complete for this analysis during the first part of the session. The remaining six participants are discussed in the results below. Of the six remaining participants, five previously played a Telltale Games title (four played Telltale's *The Walking Dead* series [147], one played *Sam and Max* [148], and one played numerous other titles). There were three men (P2, P4, and P5) and three women (P3, P6, and P7).

3.6.3 Emotional Signals

For this chapter, I use the term **emotional signal** to refer to a time series of measurable events that indicate the presence of affect. As discussed before, these include heart rate, skin conductivity (sweat), facial expression recognition by ATF (at the peak of an expression detection as indicated by an emoticon), player self-reports by way of one of the eight SEI, and think-aloud. These signals occur at different times due to the nature of interactive digital narratives, and thus one of the challenges that I address is that of aligning player's emotional signals to each other. While I was mostly concerned with narrative content and raw affect, it would be worthwhile to collect additional measures for insight into motivations and personality in the future. For instance, players had varying attitudes toward flirtation and romance elements hinted at during the first chapter, and this may have elicited other responses throughout the game.

3.6.4 Story Elements

The following sections use the lenses described in Section 3.5.2. These represent the transformation of objective observations into narrative-focused metrics for understanding both the players' relationship to the content and the variations in player experience.

3.6.4.1 Character Affinity Analysis

The goal of a character-oriented analysis is to differentiate between players based on which characters engaged them the most. Since each sensory measure is a measure of engagement, the combination, when associated with characters, can be used to differentiate player attitudes towards characters. Using the annotations described in Section 3.4, I calculate players' affinity for each character using the normalization procedure from 3.6.1.

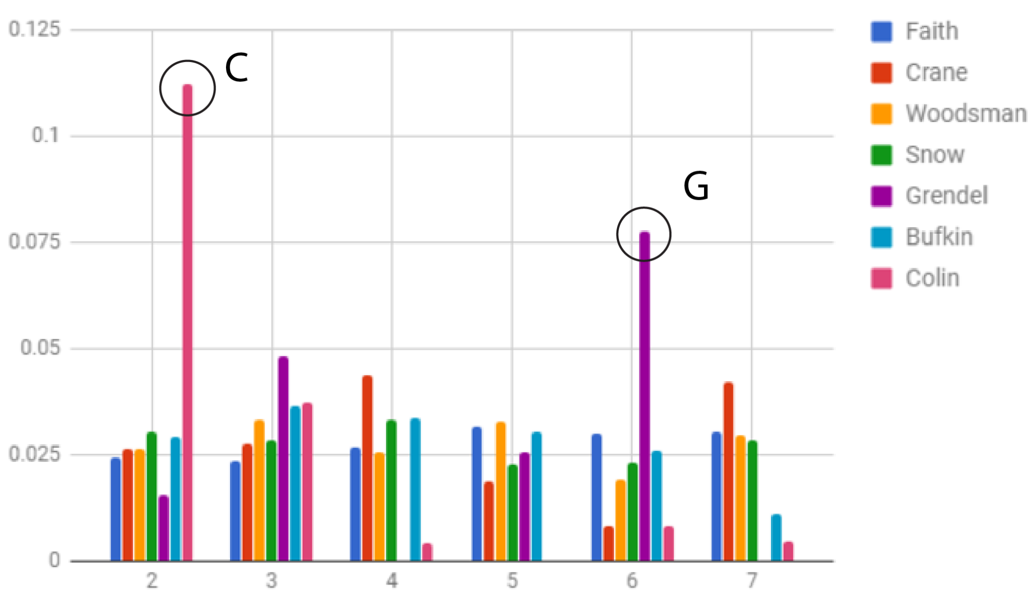


Figure 3.5: Normalized mapping of players to characters according to affect measures

The most prominent feature is P2's response to Colin (indicated as annotation C). During the think aloud, the player reacted strongly to the character and especially to the story on which he was based ("The Three Little Pigs"). In Figure 3.5, the two character-oriented reactions that stand out are P6's response during scenes where Grendel is present (indicated by Annotation G) and P7 and P4's responses to scenes involving Crane (the red bar that is higher than the rest). Future work will involve charting these affective responses through player decisions. For instance, during the climax scene, P6 and P3 decide to dismember Grendel's arm, which may have been indicated by previous responses and decisions. The unique feeling that a player has toward each character in a narrative defines their reactions and their enjoyment. This measure of player-to-character relationships may be preliminary and more associative than causal, but the concept of creating a fingerprint of how a player feels about a character indicates that storygame researchers and even the systems themselves may be able to focus more on measuring and adapting to players with appropriate modeling. Such assessment would benefit from a finer-grained approach to assessing responses, possibly at the level of individual lines of dialogue or character actions.

3.6.4.2 Emotional Choice Analysis

I annotated choices for the entire episode, noting the player's choice as well as the options presented. Some choice sets had different options based on previous decisions, in which case they were assigned a further variant as a lowercase letter in the id. In Premiere Pro, I labeled in each player video every choice prompt occurrence with a semicolon-delimited list of elements including the ID, the choice the player made, and the choice options, in a predetermined order. I analyzed the choices in three ways: first, through determining the location of choices within their respective beats, and second through analyzing their proximity to emotional events, and finally for agreement.

For the first three chapters, there were 50 unique choice sets, including variants. Of these, 10 were shown only to some players, while the other 40 were shown to all players in each traversal. 47 of these choices occurred during a beat. 34 of these choices occurred within the middle of a beat, defined by the timestamp falling between the middle 60% of the beat based on time. 6 occurred during the final 20% of the beat (2-5a, 7-2, 7-7, 12-1 and 12-2), and 8 took place during the first 20% of the beat (2-4, 2-5b, 2-5d, 2-7, 5-1, 7-3, 7-4, 8-2).

Table 3.9: Choices with emotion proximity. ChId represents Choice ID, Ch_[Button] represents the button used to select a particular choice label.

ChId	Count	Ch_X	Ch_Y	Ch_A	Ch_B
2-2	44	This is your last warning	[threaten him]	[...]	You're drunk
3-1	35	Sorry about the car.	How's your insurance?	[...]	Get off the street
7-5	35	My job	Don't need advice	[...]	Not my fault
2-7	33	HEY!	Will you excuse me a moment	[...]	[Throw him out]
19-16	31	Keep it	[Open it]	[...]	It belongs to Lawrence.
12-1	30	I'm fine	Fuck off	[...]	I'm not great.

In this analysis, I use the choices from the entire episode. In Table 3.9, I show the top six choices based on their proximity to emotional annotations. The number of players that chose the same response is listed along with the most popular choice. These reveal some problems with the straightforward approach of mapping affect annotations onto content: the scene where the top two options occur are artificially boosted due to the combat sequence that follows both. While using a time window may help, a type-based approach to dealing with player affect may be more effective than a single blanket time window.

Table 3.10: Unanimously Agreed Choice Prompts. DID represents Decision ID, ChID represents Choice ID, while Ch_[Button] represents button associated with choice option

ChID	DID	Ch_X	Ch_Y	Ch_A	Ch_B
13-1	2	Glamour	**What do you want?**	[...]	The car.
7-1	2	Yeah. Get out	**C'mon, I'm tired.**	[...]	There's only the one.
7-7	1	[**Give Colin a Drink**]	[Take Drink]		
19-9	4	What toy could have made this mark?	He left the toy for that long?	[...]	**The broken lamp was** here.
19-16	2	Keep it	[**Open it**]	[...]	It belongs to Lawrence.

For the final method, 3.10 tabulates a set of choices in which all 7 participants selected the same option. All players chose to give the character Colin their drink, and all characters responded with the most polite response for `Choice ID 13-1` and `7-1`. The `Decision ID` represents the most popular choice as an index of the column of the selected choice. These agreed choices provide some insight into the overall nature of choices where the answer is obvious, as in the case where all participants decided to be nicer to Toad (rather than bring up the two previous incidents) as well as giving Colin the drink he desired. More interesting is that none of the players decided to follow the high road and not read the final note from Faith to her husband, choice 19-16, and instead opted to read it. Correlating player agreement and responses is another area worth investigating further.

3.6.4.3 Beat Analysis

This section describes the benefits of using beats as an organizing content element and considers the trade-offs between self-report (such as SEI usage) and psychophysiological measurements.

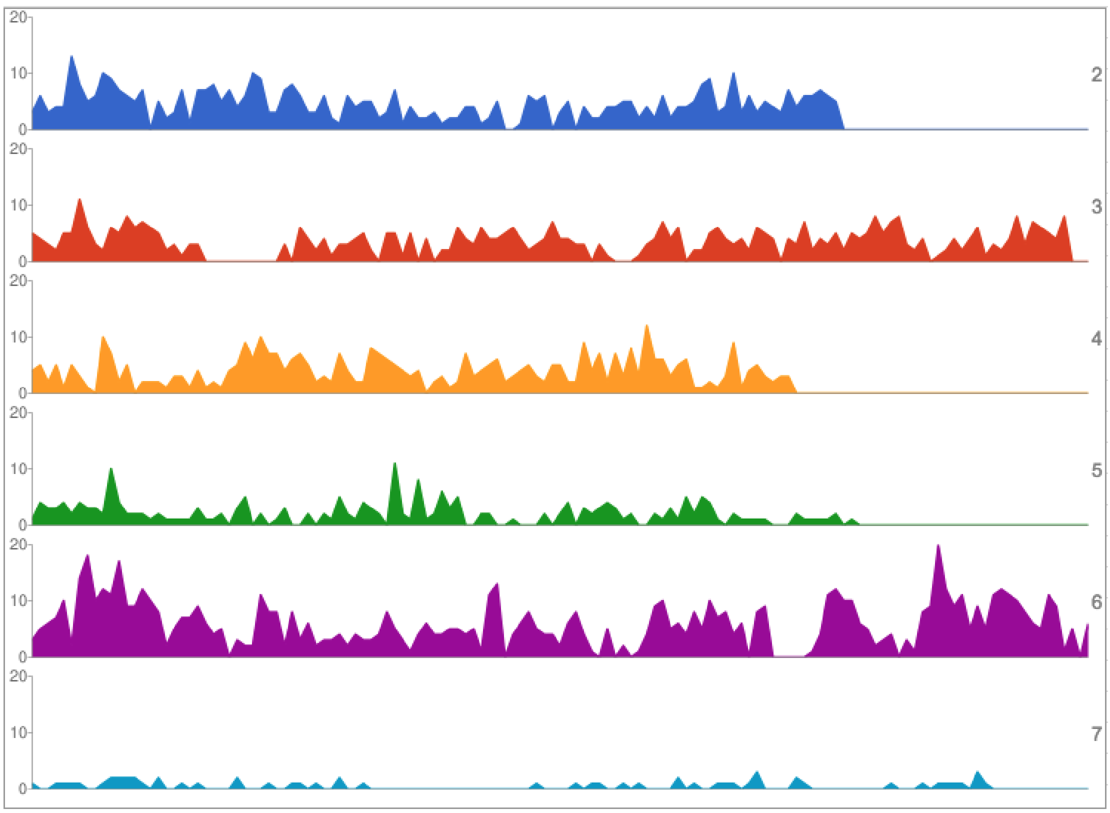


Figure 3.6: Sum Methods per Time Window

Each player spends a different amount of time in each content segment due to the nature of gameplay. One method of analyzing player affective signals is to view each segment of time uniformly, as a sort of binning approach. This is shown in Figure 3.6, where any expressive event during each period of 60 seconds is summed in a time series. The chart shows both the variations in timing as well as the variations in expressiveness between the players. Where there are no data, players experienced a technical difficulty during the testing session, as indicated by the flat line on P3's graph. The variance in the timing made it clear that another approach to collating the data for analysis was required.

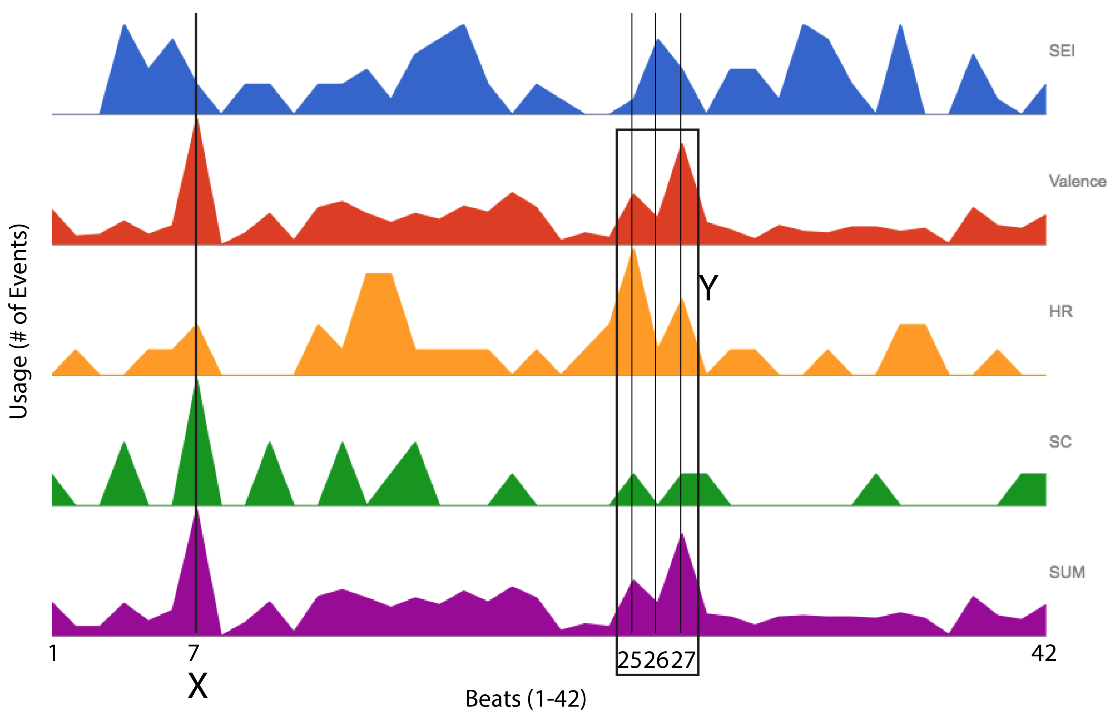


Figure 3.7: Sum of Each Method Per Beat

The next two analyses are based on segmenting the first three chapters into 42 beats, as described previously. Each beat was used to select events that occurred between the start and end timecodes. Each expressive event category is summed across all players and shown in Figure 3.7. The time window for annotations was shifted 8 seconds later from the start and end of the content annotations to account for delays in response to the content, although further fine-tuning of this shift may be worthwhile. The x-axis represents an ordered sequence of beats from 1-42, whereas the y-axis represents the number of affect annotations in the time window for each beat. In this chart, beat 7 (noted as X) indicates when the Woodsman attacks Bigby in the game's first quicktime sequence. The encounter requires quick reflexes and exhibits a spike across the various physiological measures that reflects the tension involved. Beats 25-27, indicated as annotation Y, starts with the beat where Bigby argues with Colin, followed by the one where Snow retrieves Bigby while looking distressed, followed by the revelation that Faith was killed. These also showed clear physiological responses up to the revelation from players across all measures, with a dip in response measures during the suspenseful beat where Snow retrieves Bigby (beat 26). Discovering Faith's death is a key moment in the game and was carefully set up by the previous chapter, and this moment elicited sympathy from the players and foreboding at her fate.

The next chart breaks out each player's individual SEI usage 3.8. It revealed a spike (annotation A) for P3 when Snow retrieves Bigby. Unlike the physiological measures, a higher number of SEI usages usually indicates an increase in ambiguity, as the player is selecting multiple objects to indicate their feelings rather than using the same instrument multiple times as an indicator of intensity. P3 used three different SEIs in that beat, indicating a shift in feeling in one of the more suspenseful moments. P4 and P7 have spikes in their respective signals for SEI usage at beats 32 and 33 (annotation B and annotation C), during a tense scene with Mayor Crane. Another peak that occurs in the SEI signal is that of beat 35, where P5 and P3 both select multiple SEIs during a beat where the assistant flying monkey, Bufkin, is introduced. This scene follows the intense scene with

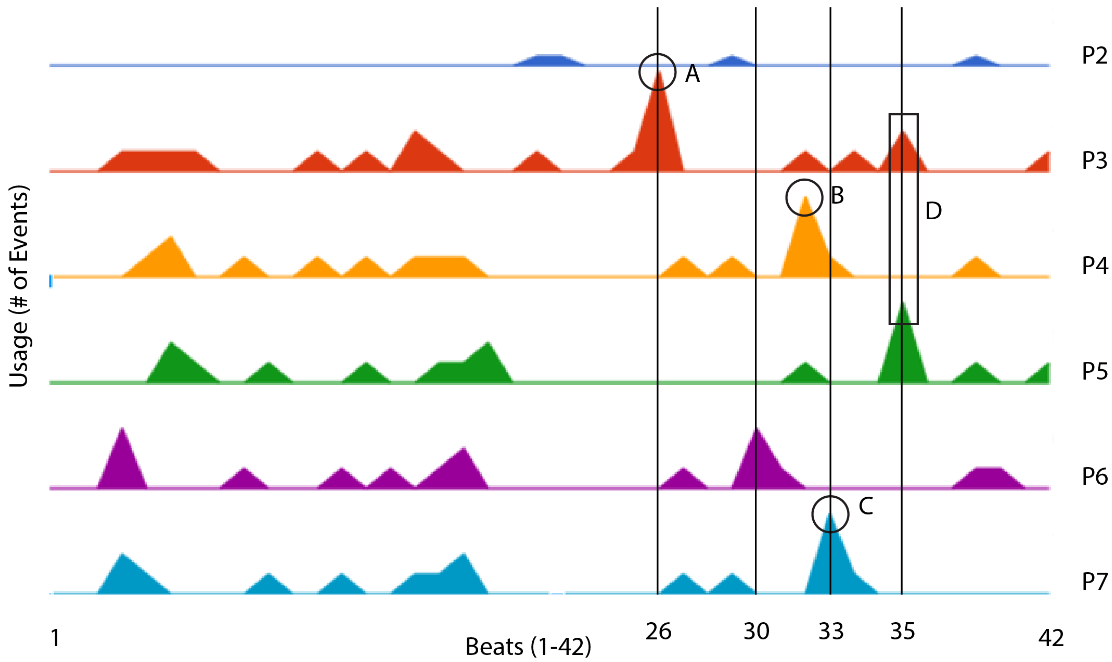


Figure 3.8: SEI Usage Per Beat

Crane and, like the introduction of Colin and Toad, drives home the fantastical nature of the community and its inhabitants.

Future versions of this analysis should further discriminate between the different instruments used, although each player may end up assigning different meanings to different instruments. Additionally, the duration and nature of the SEI usage may be significant, although here we have only noted the location of the initial use. The nature of usage can further be understood using other statistical measures. I used each in a summative signal to achieve an overview of the space of potential affective events across multiple players, although breaking each signal into individual players assisted in locating the potential causes of the various features described. A more finely grained view of content can further distinguish between instruments used in response to one part of a beat or another.

I also used the timing information from the beats to analyze the position of the choice prompts in relation to their containing beats. Of the 48 choice prompts annotated that occurred within a story beat, six occurred in the first 20% of a containing beat, and eight occurred in the final 20% of a containing beat. The remaining 34 took place somewhere in between. This makes sense, as the characters (including the player character) have more of a chance to frame the decision, whereas an initial choice-prompt would put a lot of pressure on the narrative designers to shape the rest of the content of the beat based on the initial player choice.

3.6.4.4 Value analysis

The final analysis involves examining how players reacted to the main themes in the story. These are embedded throughout, and so simple ordered sequences of beats may not reveal a preference toward one value over another. I came up with the following list of binary-oppositions to describe the underlying force in each beat: Justice/Injustice, Ego/Community, Truth/Lies, Nice/Bad, and Duty/Hedonism. I could not classify two beats. These beats did not play a dramatic role in the sense of an active conflict involving a story value. One involves a character asking Bigby about his early guesses as to who the culprit might be, which functions to further solidify Snow's relationship and to pique the player's interest. The second is likewise an opportunity for the player to engage in speculation regarding the husband of the victim.

I labeled each beat with the story values and applied the same normalization as was described in Section 3.6.1. These are shown in Figure 3.9.

It is worthwhile to note that P5 stands out from the rest, with a strong reaction to beats 8, 14, and 19, all of which concerned survival or bravery. These beats involve Faith's pursuit of her self-interest in the face of the Woodsman and the

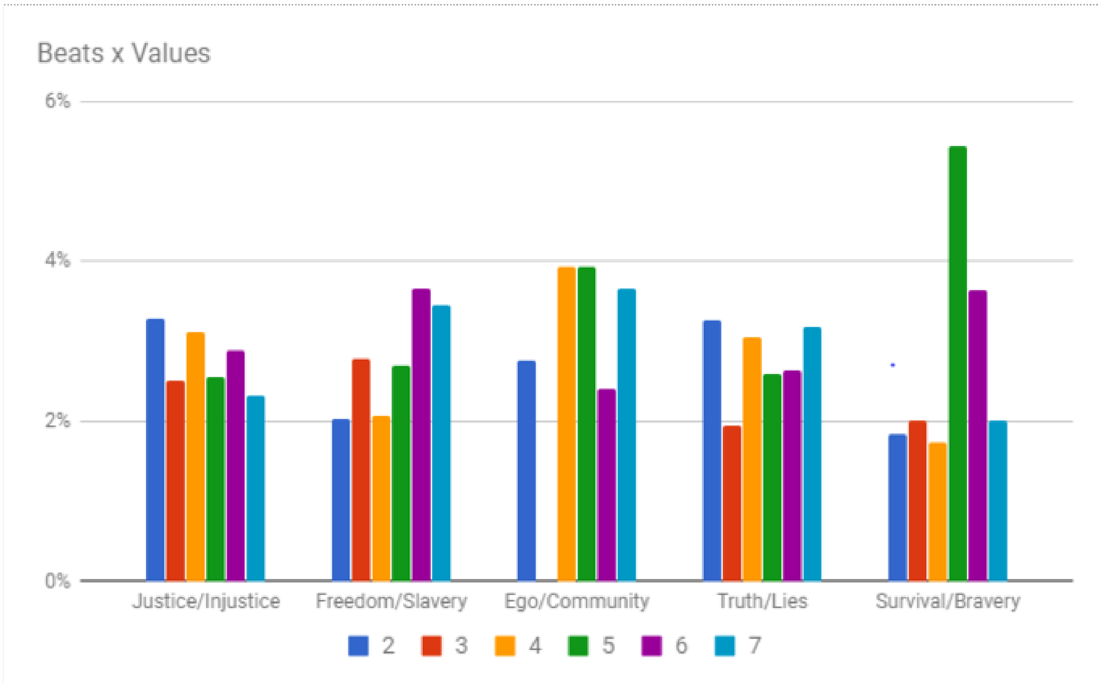


Figure 3.9: Value-oriented sorting. Each player's relative affective response to each value is shown.

world at large.

In Figure 3.10, the distribution of story values is not normalized, and so the absolute relationship between the prominent themes of Truth/Lies and Justice/Injustice are contrasted with the less common beats that focus on the story values of Freedom/Slavery and Survival/Bravery. These story values are further developed later in the episode and the series.

The analysis yielded patterns that were not immediately apparent from watching the videos themselves, especially given the total length of the gameplay videos. It was difficult to remember each of the player's responses to particular scenes, and analyzing the response data required a closer inspection of the relevant scenes relating to abnormal spikes, such as the response to Colin. The other challenge was the difficulty of ensuring annotations corresponded across traversals and retained the relevant structural features. While I was interested in how themes were

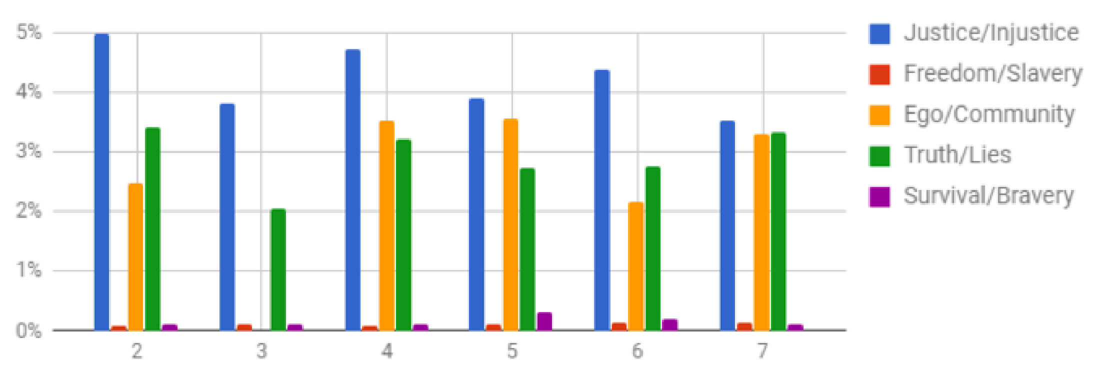


Figure 3.10: Player-sorted Value Beat Emotional Signals

related to player responses to content, I did not create a label in the annotation schema for how choice texts or outcomes were related to the story values. Both of these features are included in the ICE schemata in Chapter 4 and the encoding in Chapter 5.

3.6.4.5 Story Analysis

I recorded in each video key events and selected a time window of 25 seconds after to record the number of affective events correlated with each key event. The results are shown in a stacked bar chart in 3.11. This chart, unlike the others, is *not* normalized according to the player's event frequency and therefore shows the contrast in the expressiveness of P6 relative to P7.

3.7 Summary

The results presented in this chapter provide evidence that using various metrics in conjunction with a content-oriented schema can help guide not only assessment of traditional responses associated with games but also assist in interpreting the preferences that players exhibit with respect to certain classes of narrative content. While the granularity at this point is fairly coarse, this methodology

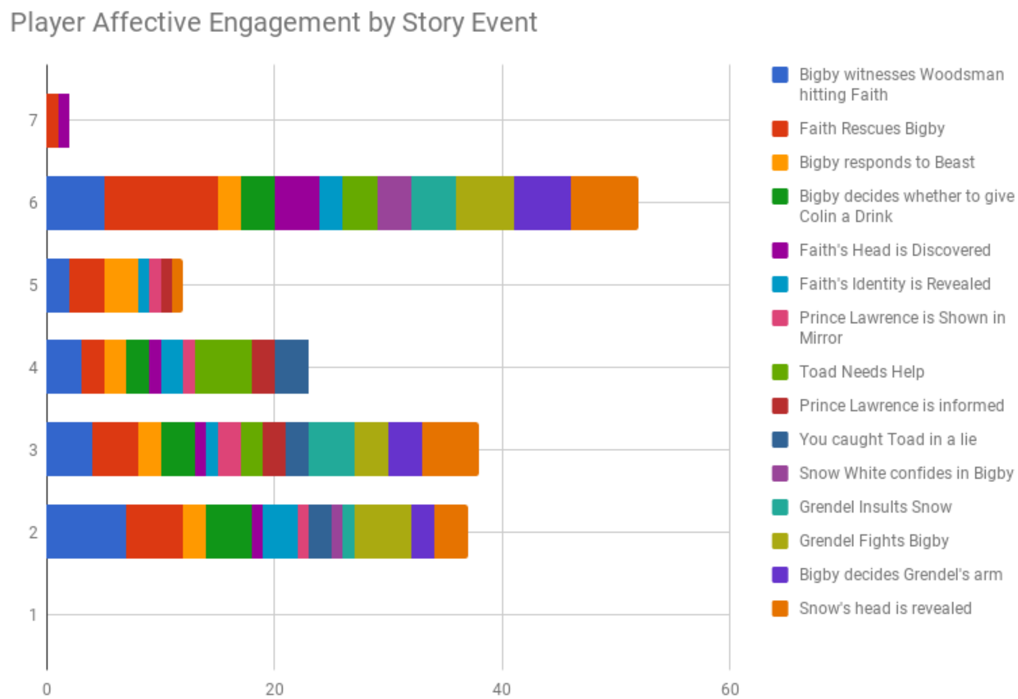


Figure 3.11: Player affect by event

does provide a means of comparing multiple players at once in a way that would be difficult to do using traditional analysis of player traversal videos, as it highlights both points of interest (through the points where SEI differed) as well as common patterns among all players (where the narrative was designed to elicit similar responses).

3.8 Conclusions

This chapter describes evidence that combining measurable aspects of player response can highlight properties of the larger content structure, as well as indicate individual player preferences. The method of using content annotations to sort and compare player traversals can be seen as a promising direction to combine qualitative and quantitative measures on longer storygame experiences and is the basis for the model developed in Chapter 4 and the visualization and annotation tool described in Chapter 6.

The study revealed several challenges for future work in relating content and player experience, and these were as follows:

1. The relationships between measures of player affect and content were implicit and difficult to disentangle with the current set of labels.

2. The summative method of analyzing the affective signal was coarse: it did not identify particular choices that were related to story values or their valence (positive or negative charges), nor did it allow for discriminating positively and negatively valenced emotional responses.

3. Questions that depended on patterns of content and player response were difficult to pose in the database representation.

4. The role that story value played in the mixed reception of content was observed but not in fine-enough granularity to identify potential causes.

Long-form narrative works are challenging to describe using a simple schema. The annotation schema described in this chapter provides one example of a set of criteria for breaking down content into segments and assigning labels; however, the exact labels and beat locations may be interpreted differently by another annotator, which could result in different segments and story values being selected. We are far from achieving a solution for the automated recognition of beats within the context of a modern adventure game. My annotations demonstrate evidence that even a small number of player traversals can be used to obtain insights when annotated. The model described in the next chapter provides some of these features, for instance, the ability to assess the consistency of choice values of individual players. They also provide a way of evaluating finer variation in content through the concept of story value charges associated with individual dialogue lines and actions. An encoding process that describes these features is demonstrated in Chapter 5.

The analysis inspired several new research questions, which provide a basis for future work: Were players consistent in their decisions concerning story values? Did the variations in content play a role in their level of engagement? Were specific content design patterns present that caused the effects observed? To what extent can such an analysis be automated, and what are the ethical implications of using such a method? Beginning to address these questions would require a fully annotated dataset and an annotation schema that can associate appropriate features of the content with points of interest in the traversals.

This chapter operationalizes a set of labels and concepts to classify content and associate it with an affective signal derived from annotated player responses.

Another set of binary-opposition pairs may have also suited, and thus certain aspects of this analysis must be necessarily qualitative and interpretative. Likewise, the exact specification of what individual practitioners and researchers consider to be a beat in drama and narrative design is fluid and ill-defined and could use additional formalization in this context. The usage of beats and story values provided a set of constraints with which to categorize the data. By recording the experiences, I was able to extract additional values from the dataset, as detailed in Appendix B. These values provide an additional level of granularity and open up the process to finer-grained analysis of player expressions, whereas the labeling method described in this chapter focused on features that could be hand-annotated.

In the next chapter, I develop schemata that addresses this initial annotation schema's limitations, including supporting content that encodes variable paths based on choices and states, pattern detection and a more detailed set of relations between choices, story values and individual acts depicted in the content.

Chapter 4

Relating Content to Experience

What is storygame content, particularly the kind found in choice-based cinematic adventure games (CCAG), and how can we encode it?

This question is the focus of this chapter and builds on the previous analysis described in Chapter 3. I contribute schemata for annotating interactive cinematic experiences and relating them to time series records of player response. I argue that this encoding strategy can represent important features of this subgenre of adventure games, thereby creating a descriptive framework to describe and analyze design patterns and player experience, which is suitable for building further analysis tools and methodologies. In the next chapter, I demonstrate an encoding of a scene contained in the traversal dataset obtained from the study described in Section 3.2.

Why do we need additional schemata? Existing approaches that could be applied to CCAG either highlight salient features of linear, textual works that have been read in their entirety (as Elson's Story Intention Graph) or they focus on discrete, goal-oriented choices (Mawhorter's choice poetics). These approaches are both powerful and provide examples of how to address their respective content; however, the unique character of the content in CCAG provides both a challenge and an opportunity to understand the moment-to-moment experience of narrative choice-based content, particularly when used with a rich source of player response data and gameplay video.

This chapter makes the following contributions:

1. I contribute a **descriptive encoding schemata**, the *Interactive Cinematic Experience* (ICE) schemata, with an example representation in the form of an OWL 2 ontology. The name draws from the key features of the content that it is designed to address, specifically computational media artifacts that rely on video to produce experiences and which are primarily about surface content. It represents a set of relationships between choice-based dramatic content, poetics features of content, player traversal data and player response data.

2. **Extending Koenitz's theory of Interactive Digital Narratives (IDN)**, specifically his concept of *protostory*, and providing a taxonomy of relationships for use in describing various content elements in CCAG as they relate to concrete traversals of interactive choice-based stories.

3. Example **patterns** that can be detected using ICE that the expressiveness of the schemata for detecting phenomena and describing relationships.

ICE is designed to integrate interpretive annotations of content for use in analyzing concrete player traversal data comprising both raw measurements and extracted features. It leaves deeper, goal-oriented or causal analysis of the story world and its relationships to other schemata and models; however, these can be integrated in the future.

In CCAG, choices and their context are designed to evoke particular feelings in players through selection and through several features innovated by Telltale

Games. I call the response to this choice and content strategy the Telltale effect, after their innovations, and I describe it further in Chapter 8. It describes which features give rise to player responses as a result of the interplay between content and choice, and how these can direct player feelings about particular subjects and characters. Some of the principles described in practice-oriented talks, such as that given by Molly Maloney [40], provide insights into features that play a large role, such as the feeder line and role-playing "rails." The integration of choice-based content, key content callbacks and payoffs, and the expression of story values are all key to the effectiveness of the Telltale effect. The approach outlined in the next chapter of encoding player traversals using the formal schemata described in this chapter was useful in identifying the relationships between the structural patterns and player responses, which was further described in the analyses on the encoded data set described in Chapter 7.

I encountered several challenges in the analysis in Chapter 3 in terms of labeling content. For example, the order of choices in certain beats varied, and content segmentation at the level of beat and choice position alone did not capture how players responded to the particular content from a choice result. However, the database representation was good at selecting events based on content spans. The specific patterns used to detect events relied heavily on table organization and field names rather than the underlying structure.

The remainder of this chapter is organized in the following manner: First, a definition of storygame content is developed and used to motivate the specific goals for ICE. Then, I review and contextualize the semantic web technologies selected for the schemata. The four layers of the schemata are then described along with my rationale for their structure, representing **storygame content** (protostory layer), **traversal details** (instantial layer), **data sets and processing** (data layer) and **intended effects** (interactive poetics layer). Next, I demonstrate potential expressiveness of ICE through example IDN patterns and provide sample queries that represent them. I conclude by discussing the limitations and potential extensions for future work.

4.1 Content and Storygames

"Content" is a slippery term. However, to formalize its representation, it is important to define it in such a manner to separate it from other elements in a storygame that are often seen as interchangeable to some degree, such as the game engine, the computational platform the game is run on, and the particular controls used (e.g., a handheld game controller or a mouse or keyboard). I propose one candidate definition: "Data that encode a valuable experience for presentation to some audience." This definition agrees with our understanding of media such as periodicals, books, and film as "content": the content represents the outcome of an audience seeing, reading, or otherwise experiencing the information that informs, entertains, or otherwise provides value.

Defining content for interactive works is more challenging than such fixed media. Wardrip-Fruin decomposes expressive media artifacts into three components: surface, data, and process [149]. Roughly, these correspond to the final rendered experience, the underlying data structures and instantial data, and the manipulations and algorithms that operate to transform one into the other. Many computational media artifacts that incorporate narrative vary in their emphasis on the surface content as a function of processes. The works that ICE can describe in this chapter have a stronger focus on surface content than on the procedural manipulation of that content.

Yannakakis and Togelius provide the following definition of content within the context of experience-driven procedural content generation:

> Game content refers to all aspects of a game that affect game play but are not nonplayer character behavior or the game engine itself. This definition includes such elements as terrain, maps, levels, stories, dialogue, quests, characters, rule sets, camera profiles, dynamics, music, and weapons. The definition explicitly excludes the most common application of learning and search techniques in academic games research, namely, NPC artificial intelligence. [54]

This definition of content includes the entirety of what is authored for a typical Telltale Games production, as none of the agents or gameplay, to date in its main episodic titles, are controlled by artificial intelligence (AI). In this sense, the game, apart from the engine itself, is content. Content that is authored and produced by writers, directors, and performers. Authoring such content could be supported or augmented in the future by generative methods, either at the plot level at authoring time or through a process of content generation at run-time. For the moment, though, it is handcrafted and represents a design structure that is intended to elicit player responses.

4.1.1 Storygame Content

Although the story content is lumped in with other instantial assets, its overall impact on players and the experience *feels* like a different category of content. The creative production of compelling stories remains mysterious, despite the extensive research over the years, and many generative systems, tools, and research projects continue to explore what makes narratives interesting [1].

[1] Façade [150], discussed further in Section 4.1.2, blurs the boundary between storygame content and AI. In it, the authored content (in the form of beats, or sequences of programmed behavior of the various characters, and the handwritten lines of dialogue) and the presentation (in the form of the drama manager) do not explicitly model the impact of the content on the reader, but instead approximate it through a heuristic of the plot arc. The plot arc models the increase or decrease of tension over time toward a resolution. However, this conceptualization of content and model as a single intertwined form has proven challenging from an authorial burden perspective, as the potential impact on the player of each piece of content is not independent.

A single consistent theory of interactive narrative would simplify the design and study of IDN. Hartmut Koenitz proposes one promising framework for IDN content, which proposes a set of terms including "protostory" that begins to articulate the level of abstraction the schemata hopes to capture:

> Protostory denotes the concrete content of an IDN system as a space of potential narratives. Any realised narrative experience is related to the respective protostory through a process of instantiation. Protostory can most easily be understood as a pre story containing the necessary ingredients for any given walkthrough. [51]

In Koenitz's terminology, the segments within player traversals can be seen as "instances" of their respective protostories elements. Koenitz noted the need for further taxonomies with particular attention to segmentation and primitives, which ICE constitutes.

Schemata such as Elson's SIG support testing theories on a set of encodings. The Story Intention Graph is a set of schemata that served as an inspiration for ICE. Generative models can provide sufficient detail to create new artifacts, as demonstrated by Mawhorter's *Dunyazad*, but most fall short in producing complex or high-quality works compared to those created by humans. ICE represents a set of schemata for relating content and response data; it is designed to annotate player traversals of existing games to learn more about how their content structure relates to the effects they produce in players.

I developed ICE after observing the promising results of the previous chapter. The cost for changing the annotations and revisiting the content was high, as the representation was closely tied to the questions themselves. The consolidation of the observed relationships into formal representation instantiates a theory of how the subgenre operates. It also demonstrates a method for using semantic web ontologies to organize and relate objective measures with interpretive annotations in the context of media artifacts.

4.1.2 Revisiting *Façade*

I'd like to return for a moment to the interactive drama *Façade* [150]. *Façade*'s internal representation is documented at length in Mateas' dissertation, but several features of his choices for content selection and modeling are worth reviewing.

Façade includes a Drama Manager process that selects and populates beats based on two story values: tension and affinity. The drama manager selects beats based on preconditions, which include whether a previous beat has been played or whether a topic has been asserted. Importantly, changes to story values are represented as effects of a beat, representing to the system the content's overall effect on the perceived experience. These changes can be absolute or relative. ICE only includes an absolute numerical representation of story values, although it is worth exploring other systems in the future to capture the subjective interpretation of context, which I revisit in Chapter 7.

There are two insights that I adopt from this approach: the numerical modeling of story values and the association of beats with distinct story values. Tension and affinity can be considered as a single story value that exists independently but combines to characterize the experience.

Façade's drama manager selects content templates and chunks based on an author-specified trajectory of player experience rather than attempting to model agent behaviors in a more simulationist manner. The system calculates the value of the content at any given point based on labels associated with the content and uses the content pieces to redirect attention from mistakes and to order and present content that both makes sense and matches the desired progression. ICE also focuses on player experience but considers content that has not been labeled or categorized by its authors as a potential resource for understanding the forces at work. It can be used to annotate commercially available games and traversals using some of the same insights that *Façade* had with regard to the importance of the impact of modeling content features on the player experience.

Discourse acts, the primary means of player input in *Façade*, were used to control the selection criteria and preconditions for beats. Players had a responsibility in maintaining the role-play, and Mateas discusses how the design of the content used the interaction of the two non-player characters with one another to address potential mishaps with either understanding or player statements. Telltale Games handles this issue for scenes with two characters by providing not only a selection of fictionally viable responses but by having player characters initiate actions before players are provided options to shape the choice space to their intended experiential goals.

Both Telltale Games productions and *Façade* use hand-authored content in the form of voice acting and performances in a dramatic mode and both require careful attention to timing and pacing to achieve their respective effects. Whereas *Façade* has successfully achieved a level of player agency with the variety of strategies to handle player input and a powerful system for mixing behaviors and dramatic beats, CCAGs take an opposite approach to evoke emotions from players by carefully constraining agency and carefully managing each beat's variations to maximize the impact on player perceptions. Fully comparing the two approaches is beyond the scope of this thesis, but the wealth of research available on both the design and evaluation of *Façade* provides a promising foundation for describing how content is interpreted and how to annotate it using schemata such as ICE.

4.1.3 Challenges for Analysis

The Wolf Among Us (*TWAU*) contains a complex mix of choice poetics and standard dramatic devices that draw players into the story. In the previous chapter, I introduced three perspectives from which the content could be understood between player traversals, using concepts inspired from narratology. I also evaluated whether these could be combined with content segments based on time windows of player response data and whether player choices could be used as indexes. The results provided sufficient evidence that the use of affect events as useful signals to classify player response could be used to reveal other types of patterns, but they also suggested that further granularity was needed to understand how specific elements in the content and player choices were related.

As noted at the end of Chapter 3, I identified the following issues during our initial analysis:

1. The relationships between measures of player affect and content were implicit and difficult to disentangle.

2. The summative method of analyzing the affective signal was coarse: it did not identify particular emotions or discriminate between positive and negative valences either in responses or content.

3. Questions that depended on patterns of content and player response were difficult to pose in the database representation.

4. The role that story-value played in the varied reception of content was apparent but not sufficiently distinct to isolate potential relationships.

The first challenge represents a growing concern within the game research and AI community: that of reproducibility and transparency of research results. Peng distinguishes reproducible research from replication [59]. Ideally, the same experiment would be conducted with new data to verify the results. In the absence of the resources to do so, the availability of the data itself and the code used to verify it is useful in evaluating the claims. I plan to make the tool and data set openly available as a resource for additional analysis. There are many opportunities for further analyzing the data set, particularly in the form of testing machine learning models and other descriptive computational representations.

4.2 Semantic Web Technologies

The term "Semantic Web" was coined by Tim Berners Lee [151] for an interconnected web of meaning that would be understood by machines. Many of the underlying infrastructure and resulting technologies are employed in cloud-based applications and data centers.

The Semantic Web architecture supports a distributed method of representing relationships, supporting the combination of multiple data sources and providing mechanisms and reasoners to infer new facts. Popular reasoners include Pellet [152] and HermiT [153]. In the web ontology world, reasoning is the execution of a set of rules that results in asserted triples representing inferred facts. One example is that of the grandfather relationship in a representation of a family tree. Instead of encoding every possible relationship between family members, a rule can be used to detect and represent particular structures as a relationship.

Description Logics (DL) strike a balance between the expressiveness of first-order logic and the tractability of propositional logic. OWL 2 is in the family of description logic languages (DL); thus, many of the ontologies created with it can be reasoned over efficiently.

While a review of all the details of description logics and their formal expressiveness are beyond the scope of this dissertation, there are some relevant features of OWL that are relevant to the ontology's structure and function.

The core technologies of the Semantic Web build on top of existing ideas such as the internationalized resource identifier (IRI), an extension of URI schema [154]), and include the Resource Description Framework (RDF) [155]. RDF relies on the fact that labeled graphs can be represented as a set of triples, corresponding to subject, predicate, and object resources. RDF has multiple serializations, including XML, but the turtle syntax is used in this thesis and in the attached files due to its conciseness. The term "ontology" is used in this thesis to refer to both the computational representation of an encoding and the schemata itself, as both are represented as a collection of triples that have fixed semantics, and not the philosophical sense. Instead, when a specific term is needed, either ICE is used to refer to the schemata or an "encoding" to refer to a particular set of annotations based on the ICE schemata. Ontologies provide a graph representation for modeling relationships that would be laborious to maintain and query using relational database semantics. Examples include the use of graph patterns that incorporate property paths and the ability to use inferences in the query itself, such as the class of a segment (Beat vs. Scene) based on its relationship to another. These are all features that make the development and implementation of patterns that are described later in this chapter more straightforward using an ontology, although they also do not preclude implementing semantics in other languages, such as Javascript. The first version of SHERLOCK implements many of the semantics of ICE using traditional object-oriented programming techniques; however, ideally, these would be directly inferred from an ontology to support a broader audience of ontology authors, with some standardization of the relevant concepts.

The following is an example of an RDF triple in Turtle format:

```
ice:StoryEvent a           owl:Class .
<subject>        <predicate> <object>
```

The prefix "ice" is a namespace representing a shortening of the full ontology prefix represented as an IRI: "http://purl.org/ontology/ice". Note that although the format is identical to the IRI representing web sites on the internet, they do not always resolve to a specific network location. The full `StoryEvent` "resource" (used interchangeably with the term class) would then be "`http://purl.org/ontology/ice#StoryEvent`". The `a` used in the place of the predicate is syntactic sugar for `rdf:type`, a relationship which asserts that the subject is an `owl:Class`. There are three separate namespaces: `ice:` represents the primary ontology classes and relations; `twau:` and `ucsc:` for the elements in the protostory and instantial layers, respectively. UCSC is the prefix for the data relating to the specific dataset, as the study was conducted at the University of California, Santa Cruz. I also use the term "instance" to describe an individual member of a class, even though the semantics are different from the usage in programming.

Multiple triples can be associated with a single subject, with the comma used to repeat the subject and object and the semicolon used to repeat the subject for subsequent triples. For example:

```
ucsc:Episode3_0_238231
        a                  ice:Episode , owl:namedIndividual ;
        ice:instanceOf    twau:ProtoEpisode1 ;
        time:hasBeginning  ucsc:Frame_3_0 ;
        time:hasEnd        ucsc:Frame_3_238231 .
```

In this set of triples, the subject is a specific instance of the episode, whereas it is both an `ice:Episode` and an `owl:namedIndividual`. It has three properties: a `ice:instanceOf` relation to a protostory element, a beginning `Frame`, and an ending `Frame`.

Authoring OWL constructs using this format is tedious; therefore, I used the

Protege ontology editor [156]. It is widely used in the biomedical community and is even used in other narrative ontology projects such as in the construction of the Drammar ontology. I used it to do most of the work of identifying classes and authoring inference rules, and the figures below that present class relationships that are produced from its screen captures.

Object properties describe how individuals are related to one another. Object properties can indicate a directed relation, and are named accordingly: `time:-hasBeginning` and `time:hasEnd`, for example, suggest a relationship between some temporal entity and a specific time instant. **Data properties** contain raw information in the triple itself, often in the form of an integer or string.

Reasoning over ontologies is supported directly through the use of libraries that implement specific inference rules. While some inferences are represented in the core OWL semantics, more expressive rules need to be written using rules in a format such as the Semantic Web Rules Language (SWRL) [157]. In SWRL, for example, one can write a statement stating that a relationship exists when several facts are also present in the ontology. For example, this can be used to apply a label based on a set of relationships, such as assigning a label to a `Beat` based on whether a positive valence player response event (or whether certain `Story-ValueChargeComponent` instances have positive valences) occurred within it. In the following section, these and other ICE elements are described along with the overall goals of the schemata.

4.3 Interactive Cinematic Experience Schemata

The core structure of CCAG is a fusion of choice and dramatic narrative. Both the content and the records need to be related logically and legibly to reveal patterns for interpretation in large sets of player traversal data. These observations led to the development of ICE. The CCAG subgenre employs emotionally charged contexts that vary their story value charge based on player decisions. My approach draws from several insights that are in turn drawn from the architecture and design decisions underlying the interactive drama *Façade*, which used beats as the "story pieces" to represent the content driving joint AI behaviors selected by a drama manager. Modern CCAGs are one solution in the industry to deliver interactive narrative content that has some of the pleasures of drama, while they avoid the issue of the authorial burden associated with producing sufficient content to provide true player agency and instead focus on authoring choice spaces that evoke player emotional responses.

I draw inspiration from McKee's definitions of story writing concepts (whose use is discussed further in Section 2.1.5) [48] and formalize them to increase precision and identify how interactivity combines with narrative elements in these games. The dramatic **beat** can be formally defined as a sequence of content events (represented by children of the `StoryContentEvent` class in ICE) that are associated with either the behavior of an agent pursuing an action or reacting to the action. Each `StoryContentEvent` instance, whether it is a physical act such as a gesture or movement, or a discourse act (dialogue line), can be associated with an agent and classified using a **story values.** Story values represent an annotator's subjective evaluation of the reception of a piece of content by an audience, and is further described in Section 4.3.2.4. **Scenes** are contiguous segments of content that may contain ordered sequences of beats, but are not defined by them. A story value-based encoding provides a different perspective of story content than that of a player-goal oriented or a character-goal oriented encoding. It increases the granularity and addresses subtle relationships in content between option, framing, and outcomes.

ICE treats content segments that include beats and scenes as spans of time within which both player and content events can be classified and incorporated into queries and patterns. Relationships between spans are represented through both the **protostory layer**, which includes the element instances that are common among multiple traversals with a "proto-" prefix, such as `ProtoBeat`. ICE does not represent causality or intentionality. Instead, it focuses on associations and organizations that may play a role in triggering player responses while representing the explicit game structures observed in concrete traversals. The `StoryState` is used to manage choices with effects that change content selection within a set of traversals.

The approach adopted here does not claim to represent all aspects of an IDN's effects. Instead, it focuses on content relationships that can be associated with specific player traversal data. Insofar as the variety and number of traversals are high, the total coverage of the potential stories contained in an IDN improves. The approach also addresses the fact that not all possible traversals are equally likely, and that specific traversal patterns can reflect properties of either player differences or the nature of content experienced in a particular order that could affect the likelihood of player choices. For example, a player who makes a value-charged decision to choose justice earlier in the game may be less likely to make a choice that goes against that value later in the game.

4.3.1 ICE Overview

ICE comprises a set of classes and relationships for features on four layers of representation that pertain to IDN. It can be combined with other taxonomies of cinematography, such as that created by Winer et al. [158]. A simplified version of the schemata that emphasizes the progression from data to interpretation is

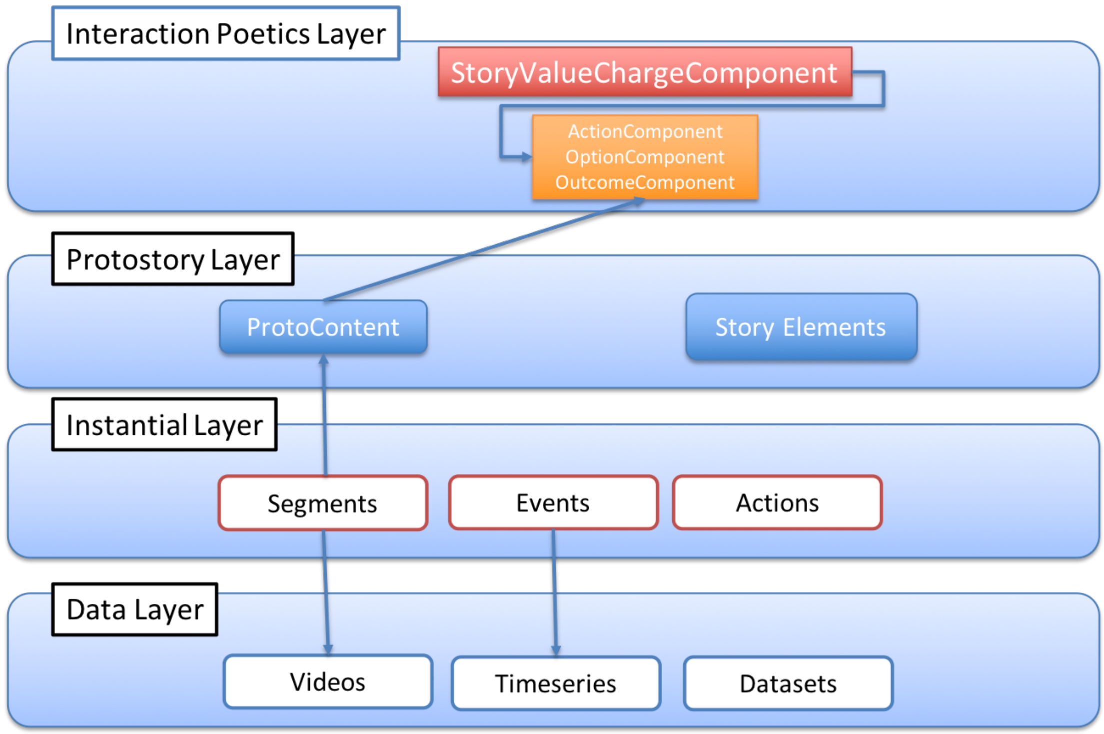


Figure 4.1: ICE Layers

presented in Figure 4.1.

In Figure 4.2, all the layers of the schemata are represented. The data layer represents the raw traversal records in the form of files, both video (1B.mp4 and 2B.mp4, representing videos of two player traversals) and biometric data sets that could be stored in a database. The instantial layer shows a single beat in which several `StoryContentEvents` are annotated. There is also a single choice, represented by the diamond that both players encounter. The choice has two options. These elements are related through the protostory layer, which captures their shared identity. In the protostory layer, other relevant content elements such as state (whether a player chooses an option in which the player character loses their money) and value (justice is a value at stake in the decision). Finally, the interpretation of two aspects (called components in the schemata) of one of the events of the story content is represented by two `StoryValueChargeComponent`

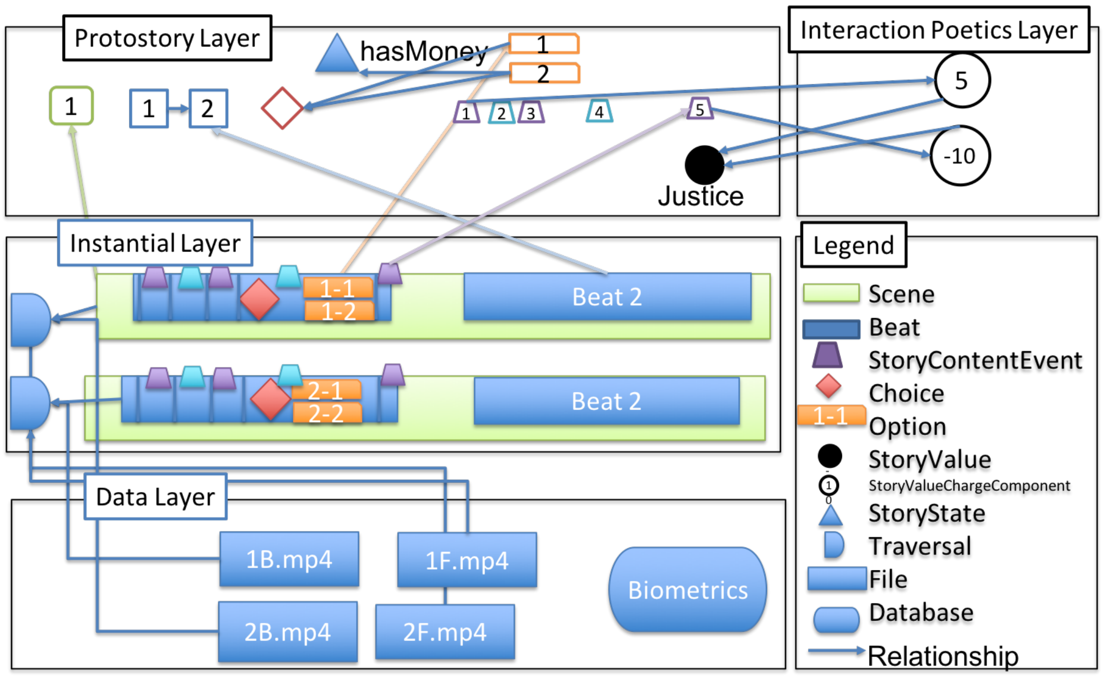


Figure 4.2: Simplified ICE Example

instances in the Interaction Poetics layer, which in turn connects the value to a specific Protostory version of the `StoryContentEvents`, or the specific lines of dialogue, actions, or spectacles that can be interpreted as being emotionally connected to a story value. These detailed annotations of time points can be used to analyze player response events, as I will demonstrate in Chapter 7.

The following sections describe each layer in detail.

4.3.2 Layers

The **protostory layer** represents the *potential* story content, including the traditional narrative elements such as characters and settings, but also the set of choices and options that can be presented to the player. It operationalizes an extension of Koenitz's theory of IDN [51]. These elements are prefixed with **Proto** to distinguish them from instances in the subsequent layer.

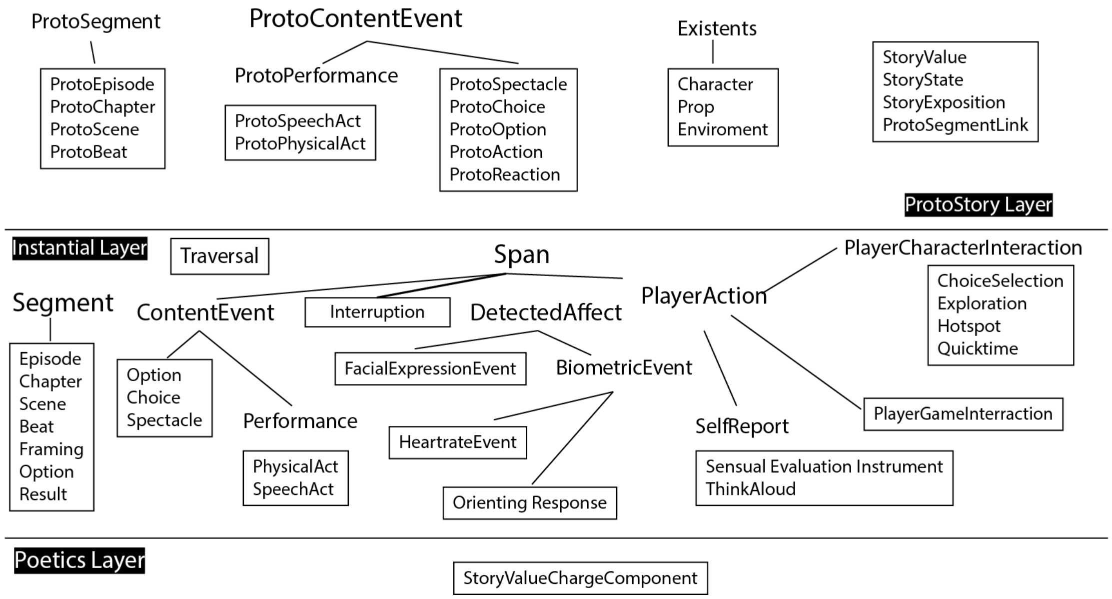


Figure 4.3: ICE Protostory, Instantial and Interactive Poetics Layer Overview

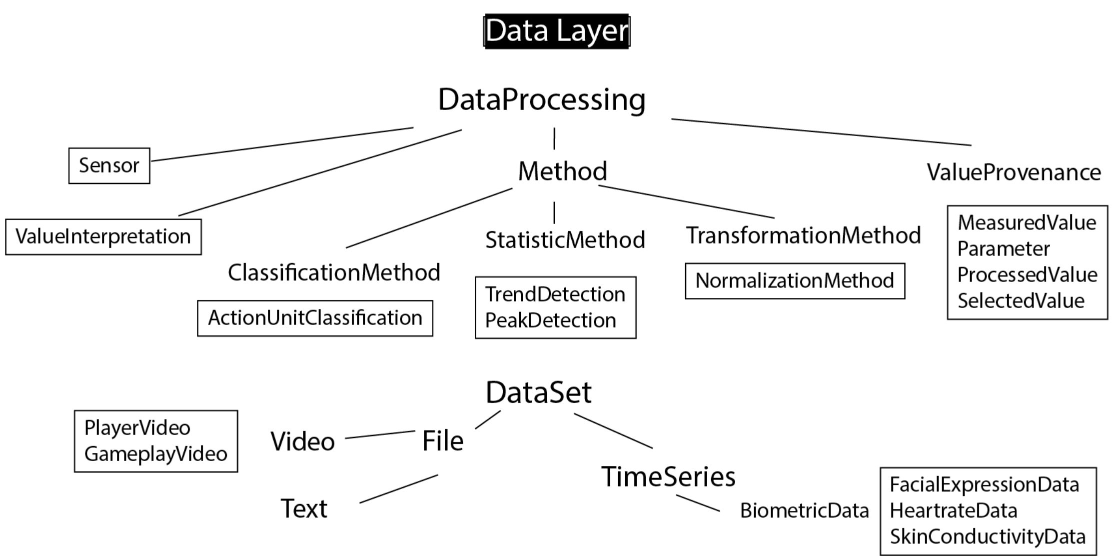


Figure 4.4: ICE Data Layer

The **instantial layer** includes labeled time spans relating content elements from the protostory layer to each player's traversal. Segments are annotated with their start and end locations and are related to their corresponding protostory elements.

The **poetics** layer captures the different authorial intentions concerning the organization of content. This extends the idea of traditional poetics and includes elements that contribute to poetic effects. It includes a class that represents how content relates to value-oriented choices using the concept of **story value charges**, described in Section 4.3.2.4, and can be extended to incorporate aspects of Mawhorter's choice poetics theory [46]. I describe particular charge components as story value charge components (SVCC) throughout the dissertation.

The **data layer** represents time series data sets and video data of player traversals. This includes the videos as well as measurements and extracted data.

The following sections describe each layer in greater detail as well as motivate their role within ICE compared to other similar approaches.

4.3.2.1 Instantial Layer

All elements that are subclasses of `InstantialElement` are events that can be linked to a specific point in time in a particular traversal. Elements can either be instantaneous (represented by the `time:hasTime`) or by a span (of `time:hasBeginning` and `time:hasEnd`). Frame indexes are recorded in the data property `time:hasNumericPosition` on `Frame` instances, while the traversal it is associated with is noted by `ice:inTraversal`. Subclasses of `Segment` correspond to non-overlapping classes of content while subclasses of `Span` may have overlapping intervals. The `Player` class is used to track which player is associated with which `Traversal`, whereas the `Traversal` instance is used to connect related elements and has a single time frame.

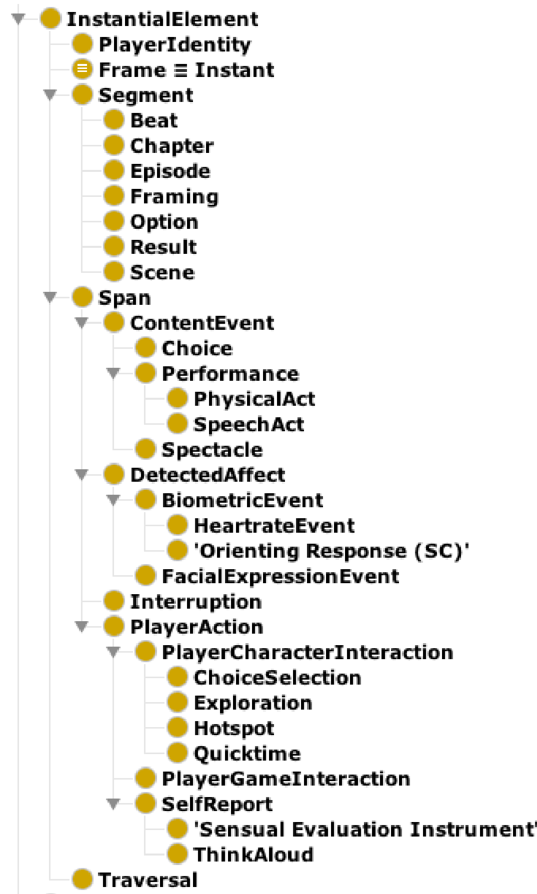


Figure 4.5: Instantial Hierarchy

These temporal annotations are used to group events in a similar manner to the previous annotation schema. The organization of content by a fixed timeline implies that any changes to a particular classification can be added without having to modify existing elements. This provides a more robust strategy than encoding hierarchical or containment relationships as properties that would need to be managed each time an annotation changed or a different representation was applied.

Detected Affect.

`PlayerAffect` instances are extracted features from some data source related to player response. They comprise spans associated with segments where some statistic is calculated (maximum, minimum, mean) or points where a peak is recorded within a time series of affect data.

`FacialExpressionEvent` includes features obtained from the time series generated by the Affdex SDK, including “Engagement,” “Valence,” “Joy,” “Sadness,” “Fear,” “Disgust,” “Contempt,” and “Anger.” These values are associated with their source through the `DataSet` class.

By representing player affect as entities in time in the ontology, a single query can relate the presence or absence of these features into a particular result set. For example, one can find all "beats" that have an instance of joy contained in them that follow another content element with an instance of "fear" in them. The syntax for representing this relationship is much simpler than the equivalent set of database queries and logic required to access secondary tables.

Player Action.

This class encompasses any event caused by a player during a traversal, including self-reports and actions taken within the game. An example of a player action is an instance of `ChoiceSelection`, which indicates that a player has pushed a button associated with a choice. These are hand-annotated in the present data set, although computer vision techniques could lead to automated extraction of navigation commands or selection of hotspots and choices. Player participation can be used to evaluate features other than simply the chosen path; explicit player actions can be used to evaluate the success or failure of a player or to understand particular patterns of command usage. This is distinct from the observation of player response, either through GSR or self-report methods (such as SEI or think-aloud).

Content Segments.

Table 4.1: Narrative Segments

Segment	Description
Episode	A single game episode, usually lasting 1.5-2hours
Chapter	A segmentation from the game
Scene	A contiguous set of actions in time and space
Beat	Exchange of behavior or presented spectacle involving value change

A segment is an element that does not overlap with other elements of its class in a single traversal, and which has a corresponding proto-element in the protostory layer. Each segment has a start and end location that is used for identifying related elements. The elements are summarized in Table 4.1.

Beats are a significant part of the player experience of Telltale Games. They follow cinematic conventions and incorporate scenes where characters interact and exchange lines with conventional cuts and camera actions. They are fundamental both to the uniformity of the experience as well as the variation.

4.3.2.2 Choices

I incorporate the primary mechanic of the genre, choices, with a `Choice` segment and the `Option` class. Options represent concrete labeled options presented on screen, while choices are instances of a specific set of options (`ProtoChoice`). Content events that affect the story value charge are classified by the association with either an option or a content event that is dependent on an option. These are related through the `StoryValueChargeComponent` instance, described in Section 4.3.2.4. Story value estimation changes on the basis of the target element: evaluation of choice text (a `ProtoOption`) can be considered an evaluation of the option presented, while when it is associated with content that is dependent on that `ProtoOption`, it depicts a measure of outcome. This approach was inspired by Mawhorter's choice poetics, [46], but departs in the sense that they are independent of a player goal and maintain a uniform representation of the estimate comprising a story value classification, a valence, and a magnitude.

Telltale narrative designer Moley Maloney describes the label Telltale designers give to the last line before the choice itself as the "feeder line" [40], although often the actions of the player-character will shape the nature of the option just as much as the responding character. An example of this is when in the Trip Trap Bar, Bigby shatters his glass and asks for a second one. This context, the violence, and demand for another glass, provides sufficient motivation for the player to select

Figure 4.6: Pay for Drink Choice

the "[Pay for drink]" option in Figure 4.6.

A further note is that the option "[Pay for drink]" itself is a device for shaping players' impressions and can be considered part of the "framing," as it implies that Bigby can pay. The state of whether or not Bigby has the money depends on a previous choice with Faith and is, therefore, associated with that `StoryState` instance.

4.3.2.3 Protostory Layer

In the **protostory layer**, represented by the `ProtoElement` class, there are "existant" classes (using terminology from Chatman [80]) representing the characters, props, and locations. `Character` elements are related to particular `StoryContentEvent` elements. `StoryValue` and `StoryState` serve as anchors for relations involving the interaction poetics elements (which are described further in the next section, Section 4.3.2.4). The `StoryValue` are the set of classes representing stakes that are evoked or related through `StoryContentEvent` instances. In addition, `StoryExposition` represents key aspects where facts relevant to the story are delivered, and is often associated with key spectacles, actions, or dialogues.

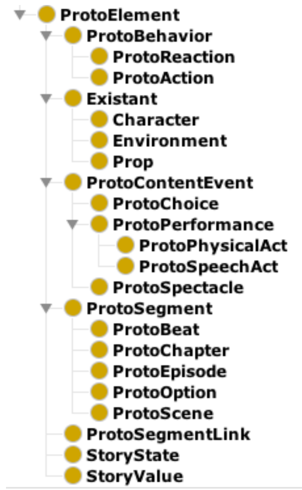


Figure 4.7: ProtoElement Hierarchy

A `ProtoBeat` is described by how its content is received, and these individual actions are represented by `ProtoContentEvent` instances: in particular, the ProtoStory versions of `Choice` and `StoryContentEvent`. The `ProtoContentEvent` is associated through a set of `StoryValueChargeComponent` instances to a story value, a valence, and a magnitude. A `ProtoContentEvent` can further be related through a `isResultOf` data property to show a dependent relationship with a `ProtoOption` that triggered it. This can support the selection of options whose results involve particular story values, or where options differ in their value charges.

Choices and the resulting content vary in CCAGs and can be dependent on previous choices. The `ProtoChoice` represents the set of potential choices that share context, options, and effects. The reason that the exact number or proportion is used is the tendency for choices to substitute alternative options based on states. In rare cases, a `ProtoChoice` will not include a fourth option at all; however, in most cases either an alternative is present or an option based on the state is offered. One example of this is when the player picks up a matchbook between the first two scenes. This changes a future `ProtoChoice` to include the option of offering Faith a matchbook. The change in the single option does not change the `ProtoChoice` that an individual `Choice` is associated with. The "Default" option is often the silent option (represented as "[...]" in choice text and selected when a timer runs out). A set of `ProtoOption` instances may not include every possible `ProtoOption` associated with a `ProtoChoice`. One method for identifying `ProtoChoice` in different traversals is the feeder line described by Malone: a ProtoChoice is defined by a sharing similar contexts and having some `ProtoOptions` in common. An example where a choice appears twice might, in fact, represent two distinct `ProtoOption` that happen to share the same option text, as is the case in the initial encounter with Faith. This is not always the case, as Bigby has multiple opportunities to ask Toad similar questions during the investigation sequence, where the questions represent distinct instances of the same `ProtoChoice` and `ProtoOption`, which depend on Bigby making observations in the environment.

A `ProtoOption` represents the idea that an option may also occur in multiple locations, but is usually associated with similar framing and possible outcomes. Most times, but not always, all instances of a `ProtoOption` share the same text. `ProtoOption` contains the option text itself, along with any associated precondition states, and any associated post-condition states. Aspects particular to a traversal, such as the selection, the display or the button offered, are associated with the instantial layer element, `Option`.

The `ProtoSegmentLink` describes traversal paths taken by players between content encountered in a traversal. Each `ProtoSegment` contains a data property, `nextSegment`, which indicates the next `ProtoSegment` in a linked list fashion. In cases where a `ProtoSegment` can be followed by several potential protostory elements, more than one `ProtoSegmentLink` element can be used with preconditions that describe how selection occurs. An example is Snow's response to Bigby after the argument with Crane: her reaction depends on whether Bigby agreed or disagreed with her. Each link has a `nextSegment` relation and a `hasCondition` that relates to the `StoryState` that is associated. If a previous option has an outcome that triggers the state, then a property path can be found for certain graph patterns. To the best of my knowledge, CCAGs produced by Telltale Games, up to this point in time, are deterministic in the outcomes of dramatic choices. In the future, if randomness needs to be integrated, the distribution of traversals could be described instead. The `ProtoOption` class supports dependency relationship to a `StoryState` through two relations: a `requiresState` relation as well as a `setsState` relation.

These elements enable graph-based queries such as those made possible by SPARQL and OWL to identify specific traversal patterns. For example, one selection criteria might be player traversals that involve a path that includes an option that sets the `hasMatchbook` state to true.

4.3.2.4 Poetics Layer

This layer provides a separate space for annotations of other elements that describe aesthetic effects and specifically content-oriented interpretation. I extend and adapt Mawhorter's choice poetics method [159] of decomposing choices and developed story value charges to represent content patterns as the sole element of an interactive poetics class, represented by `InteractivePoeticsElement`. Other poetic effects and theories could be included in this layer for combining ICE with other approaches, such as choice poetics.

Story Value Charges I adopt and formalize the concept of story value from craft knowledge as a **story value charge** with three facets: a categorical value, a valence, and a magnitude. The use of practice-oriented concepts is also discussed in Section 2.1.5. The category acts as a classification of content based on a valenced emotional concept, such as Justice or Intimacy, that is enacted in character behavior. The valence represents the positive or negative feeling that is associated with the treatment of the concept, while the magnitude represents the degree to which the value is challenged or upheld or the stakes for a character who holds the value. The role that story value charges play in the schemata can be compared to the concept of Koenitz's "narrative vectors" [51]. They represent the tendency for the content to influence the current scene's interpretation or overall response of a player through the actions or words. They address the core questions at the heart of any story: Why are these characters taking these actions? Why is it important to show these actions? What are the stakes that make these actions worthwhile to these characters? What should the player be thinking or feeling about at any given moment?

In the first scene between Bigby and Toad, the initial charge of the scene is moderately positive for **Justice**, as Bigby has arrived to settle a domestic disturbance and deliver justice for Toad's grievance. However, the **Community** value shifts to a negative value due to Toad's appearance and Toad's anticipation that it would cause problems. Because Toad is a non-human Fable, he is required by the community to possess glamour to disguise himself. Regardless of the player choice, Bigby admonishes Toad for his lack of responsibility toward the community. On the other hand, the story value charge for `Justice` shifts to the negative as Toad complains that it is not fair for nonhuman residents to bear the financial burden for the good of the community. This dynamic is often used to maintain interest in dramatic craft advice: a value moves from one valence to another, representing conflict and the dynamics of an exchange.

Story value charges do not depend on a goal-oriented breakdown of the characters' motives. Toad has a goal of not being sent to The Farm, which can be interpreted as relating to Toad's freedom (using the value hierarchy concept from Elson's encoding). However, this value is underneath the immediate danger and dynamic in Bigby and Toad's exchange around justice and community. The beat plays several roles for the storygame: it introduces Toad as a rule-breaker, the concept of glamour, and the social tension between the "haves" and the "have-nots" in the community. Despite these many roles, the content (actions and lines of characters, in this case) can be classified according to how they portray the conflict between the story values **Justice** and **Community** and, therefore, drive the scene. The specific location of content associated with a story value charge makes this representation useful for representing the player's immediate responses to the dramatic content. Other formal representations can capture other roles, such as exposition.

The `StoryValueChargeComponent` class relates actions to particular story value charges. While Mawhorter focuses on the player as an agent pursuing goals, and Elson focuses on the affectual impact on specific characters of goal-oriented actions taken by other characters, my schemata focuses on identifying locations of authorial control of story values and how they interact with other elements in order to evoke player response. I demonstrate the process of encoding a scene using multiple traversals in Section 5.2.

Encoding only a subsection of a storygame may yield interesting results due to the independence and locality of the representational unit. The Story Intention Graph structure requires a complete encoding of the story to account for causal relationships, whereas an ICE encoding does not require a goal for each action and agent. This is particularly important for ambiguous or mysterious events, such as when the player discovers Faith's head. The agent responsible and their ultimate goal remains mysterious even after the conclusion of Episode 1. It also enables us to explore the design patterns underlying choice inclusion within beat sequences and how events function as affect anchors for the overall player experience. The story value can be assigned as an ambiguous combination of independent value components without assigning a clear intentional goal, which would be necessary for a valid SIG encoding. By focusing on the surface impression of the content, the distribution of these impressions can be analyzed with respect to their successful (or unsuccessful) effect on players.

The charge is represented as an integer data value associated with the `StoryValueChargeComponent`. The method for estimating them is described in Section 5.2.5, and their application is further analyzed in Chapter 7 for the first three chapters.

Not every beat or every choice plays the same role, and this is a key aspect when attempting to encode content using a uniform representation. Many of the choices are better described using story-value decomposition, as they represent the subtle moment-to-moment value judgments involved in conversational choices. However, certain decisions are better understood through a full choice poetics analysis, particularly when a significant branch of the story is involved.

4.3.3 Data Layer

While the other layers represented relationships between content and traversal, this layer provides elements that locate and identify sources of data outside of the schemata. The primary data sources are video files and databases located outside of the primary triple store. In order to reference content features, it employs a standoff index to relate elements from other layers with specific player response data spans and features. Key features are entered as elements into the ontology, such as average values, peaks, and other derived statistics. These may be used in a query or pattern without directly querying the database. This caching function makes it so the client does not have to store the entirety of the dataset, which becomes more critical as the number and length of traversals and data points increase.

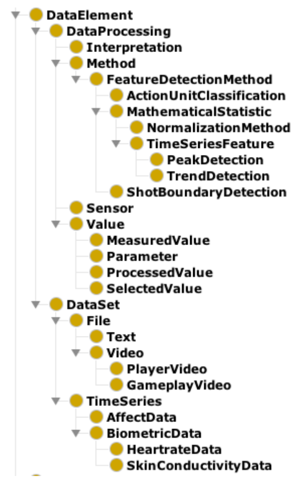


Figure 4.8: DataElement Hierarchy

SHERLOCK provides an interface for these relations, presenting disparate modalities such as video and graphs of the time series data to describe the dynamics of player affect alongside annotations representing spans of content and their relationships through the protostory layer. The visualization and interface are described in detail in Chapter 6. This section describes how a single representation framework would be useful for both interpretive and computational analyses of player traversal data sets.

4.3.3.1 DataSet

The `DataSet` class represents assets used for analysis. The extracted time series from a particular player traversal are recorded as instances of the `TimeSeries` class, including the values extracted from the data set detailed in Appendix B.

Video files have many technical variations, although these often do not affect the content or presentation beyond limiting the players that can be used to display the file. Videos are identified by their role (`GameplayVideo`, `PlayerVideo` and `OverheadVideo` are example roles designated as subclasses of `VideoContentRole`) and associated with specific player identities (`PlayerIdentity`). These enable processing operations to locate the files, reformat them, or serve them, as is the case in the Sherlock architecture.

The data set comprises biometric sensor readings for each player on two dimensions (`HeartrateData` and `SkinConductivityData`) for each frame of recorded data. Further, I added nine additional values extracted from the facial expression of players using the output from the Affdex SDK, thereby resulting in over 11 million data points spanning over a million video frames. In addition to these continuous dimensions, my collaborator provided annotated points where players reported affect via the Sensual Evaluation Instrument [143], which provided an additional 317 points representing self-reported player affect.

To support various filtering methods, I include classes that describe the methods employed. This provides researchers with some preliminary indication that additional study might reveal patterns, including traditional studies with a larger number of participants. The individual sensor readings are not included as triples in the ontology, as the individual values place a burden on any reasoning. Instead, labels are associated with segments based on the presence or absence of features or spans, which can be used to further designate classes of player response.

Peaks (calculated using the methods described in Appendix B) are inserted into the ontology directly from R[2]. An example encoding of a peak can be found in Figure 4.3.3.1.

[2] The R algorithm can be found here `https://github.com/stas-g/findPeaks`. I used a value of `M=200`, representing minima within a six-second window, and passed in the negative of the data set to find local minima instead of maxima.

```
ucsc:PeakReading7_54713
        a                   ice:or ;
        ice:hasFloatValue  "0.02687887"^^xsd:float ;
        ice:isProcessedBy  ucsc:RRollingWindowPeak7 ;
        time:hasTime        ucsc:Frame_7_54713 .

ucsc:RRollingWindowPeak7
        a               ice:Procedure ;
        ice:hasMethod  ucsc:RollApplyNorm1 ,
                        ucsc:RFilteredPeakDetection1 ;
        ice:firstMethod ucsc:RollApplyNorm .

ucsc:RollApplyNorm1  a      ucsc:NormalizationMethod ;
        ice:hasParameter  ucsc:SelectedInteger210 .
        ice:nextMethod ucsc:RFilteredPeakDetection .

ucsc:SelectedInteger25
        a                     ice:SelectedInteger ;
        ice:hasIntegerValue  25 ;
        ucsc:parameterName    "filter" .

ucsc:RFilteredPeakDetection1 ;
        a                  ice:PeakDetection ;
        ice:hasCode       "R Code Here" ;
        ice:hasParameter  ucsc:SelectedInteger25 ,
                            ucsc:SelectedInteger200 ;
        ice:hasSource     "https://stackoverflow.com/..." .

ucsc:SelectedInteger200
        a                     ice:SelectedInteger ;
        ice:hasIntegerValue  200 ;
        ucsc:parameterName    "m_value" .
```

This representation captures the nature of the parameters and the algorithm and records it in the ontology alongside the derived values. It also distinguishes calculated values (such as the mean) from researcher-selected values. Importantly, it documents the source code of the algorithm, which could be used to generate or share functions of use alongside the data set. This may be particularly important when deploying code with collaborators who may not be familiar with source control or package management systems.

4.4 ICE Patterns

Like the Story Intention Graph, ICE encodings can be treated as labeled graphs. Graph patterns are topographical relationships that describe the manner in which sets of elements can be connected along paths. These can represent situations that are encoded across multiple elements which are of interest, within or between any of the four layers. These patterns can be formally described by SPARQL queries that are used to detect them.

The `spanID` data property is used to identify specific instances of `ProtoSegment` or `Segment`, which is important when a search is conducted using a particular content segment or protostory element.[3] For elements in the protostory layer, these comprise the element name's first and last letters and an integer. For example, the protostory element `twau:ProtoBeat_bt41` has a `spanID` of `bt41`. Another use of `spanID` is to distinguish elements as being unique, given the open world assumption. OWL provides a mechanism to use a string or data attribute as a key that establishes that the element is unique and not necessarily an alias of another element. Otherwise, `twau:ProtoBeat_bt40` could refer to the same element as `twau:ProtoBeat_bt41`, given the semantics of OWL. Rules require this distinction to recognize two elements as unique.

The ICE schemata are capable of expressing the relationships underlying concepts such as tempo or dynamics, thereby reflecting an increased emphasis on the reception aspects of the content and parallel features with other temporal art.

[3]Recall that ProtoSegment instances represent the element in the ProtoStory, and are uniquely identified, whereas each player traversal has an instance of a `Segment`, which is associated with potentially the same `ProtoSegment`.

These are particularly important given the temporal elements of drama, where the action should take place in a single session to maximize the emotional impact of the content.

ICE patterns can be classified by the principle layers involved: those primarily describing the content organization or protostory layer, and those primarily describing the player response and patterns or instantial layer. The following sections describe the terms relevant in each category that are "matched," to use the terminology common in regular expressions, by the patterns themselves. Note that some patterns take parameters that describe the scope or sensitivity of the algorithm or graph pattern used to detect the pattern. They may describe a particular dynamic range using a selected value or values, for example, when a particular "large amount" needs to be specified. These parameters are selected through trial and error, although in the future they may be determined through various optimization or clustering algorithms.

4.4.1 Protostory Patterns

Protostory patterns involve the content and its potential set of relationships based on a set of player traversals. They do not directly draw from concrete instances to describe specific traversal paths. These patterns are useful for considering the potential relationships of content elements without considering how players made choices using them. A list of several protostory patterns can be found in Table 4.2.

An **expected choice** is an option that is strongly suggested through the associated story value as the preferred choice by being annotated with a strongly positive valenced option. An example is the "Beautiful" response to Faith's question of whether Bigby likes her ribbon. While that is not Faith's desired response, it is considered an expected choice by the relatively positive valence compared to the other options. This pattern reveals some of the pitfalls when it comes to using valenced annotation to identify patterns in content: occasionally, the meaning is necessary to understand certain types of interpretations. This limitation is inherent in any annotation schema that seeks to reduce complex interpretive elements into a classification and numerical representation. However, the goal behind the patterns is to guide an interpretation of the original content, and not necessarily to serve as an "objective" interpretation. Therefore, the labels are necessary but potentially misleading in certain patterns.

One of the ideas from McKee's work is that values should shift back and forth from negative to positive. The **alternation** pattern describes this sort of movement. Like other patterns, it requires a bound in order to find matches. The `span_length` parameter determines the maximum number of beats allowed before an opposite valenced beat must occur. With a `span_length` of 1, the pattern would only match beats that alternated valence every beat, for example, one beat was negative, one was positive, and so forth. The `method` describes the means by which beats are classified as positive or negative, and can be either max, last, or average. These are calculated from the `StoryContentValue` valence and magnitude values associated with the beat.

The **Long-Term** pattern captures the idea of long-term payoffs. The pattern would match a content set where one choice references a prior choice that changed a state after a set number of beats. This pattern takes advantage of the ability to specify path length in SPARQL queries using graph patterns, along with requiring a common edge with a `StoryState` element with the necessary object property.

Table 4.2: ICE Protostory Patterns

Label	Parameters	Description
Alternation	`Min` , `Max` `spn_len`, `method`	The same value shifts from positive to negative to positive in a sequence of [`min` , `max`] beats.
Long-Term	`Time`	A value is present in two beats separated by more than a specified number of beats (`Time`).
Minority Choice	`Bias`	A player chooses an option which is contrary to the **expected choice** (`Bias`)
Dynamic Range	`Amount`	A specific value charge changes between two beats by an amount (`Amount`)
Value Repetition	`Gaps` `Min` `Max`	A value is incorporated into multiple choices in sequence to emphasize its importance
Value Reflection	None	A value which was initially presented to the player as a trigger for response becomes an option for the player to decide as the agent
Sneaky Exposition	`Proximity`	important exposition is located close to an highly charged beat or `StoryContentEvent`
Choice Determination	None	A set of choices representing divergent values identified more than 1 positive values present
Distant Consequence	None	An outcome whose result is dependent on a previous choice in a previous scene through a `StoryState`
Value Inception	None	Encouraging adoption of a value through many instances of `StoryContentEvent` instances in a scene and many choices
Fast Pacing	`Cutoff` `Window`	A high or low (`cutoff`) frequency of value charge oscillations over a period of time (`Window`)
Scene Control	None	The ability for a player to make a key choice relating to a scene's turning point.
Big Branch	None	A choice whose options include a state which connects to a `ProtoStoryLink`
Consequential	None	A choice whose options include a state change

4.4.2 Instantial Patterns

These patterns reflect various potential relationships among content drawn from specific player traversal data. A list of sample patterns can be found in Table 4.3.

In these patterns, **Scope** describes the type of `Segment`, such as `Beat` or `Scene`. It is used in patterns to define what type of content segments are of interest. SHERLOCK uses the term in much the same manner.

The **Sustained** pattern indicates that a player has continued to be engaged with the content through their expression over a number of beats. The `Cutoff` parameter provides a minimum cutoff for the number of beats in order for this pattern to match, although like other pattern matching languages, greedy and non-greedy versions could be implemented.

For example, the **Singular Response** and **Singular Choice** patterns can provide a method for finding choices where one player is distinct in either the option selected or their response to a given `scope`.

Instantial patterns focus on actual player data, and so can be used both to compare players to one another as well as to find player response patterns that match them in a corpus.

These patterns are just a sample of the full set of possible combinations. Other patterns can include whether a player is consistent in their choices with respect to a particular `StoryValue`.

4.5 Conclusion

The ICE schemata encode a unified representation of content and player response data that is unique to this genre of storygames. The OWL version of

Table 4.3: ICE Instantial Patterns

Label	Parameters	Description
Sustained Engagement	`Cutoff`	A player has at least `cutoff` engagement in each beat during a scene
Alternating Valence	`Min` `Max`	A player expresses positive and negative valence during a sequence
Mixed Valence	`Scope` `Cutoff`	A player expresses both positive and negative valence during a segment
Re-Engagement	`Range`	A player experiences low engagement for one segment followed by a high engagement
Singular Response	`Threshold`	A player expresses a highly different valence and engagement than other players during the same content segment
Singular Choice		A player chooses an option that none of the other players select

Affective Choice	`v_cutoff`
	`e_cutoff`
	`Window`
		A player expresses themselves in a time window following a choice

Peak Beat	`Scope`	A beat whose expression and valence is the highest in the `Scope` for an individual player
Nadir Beat	`Scope`	A beat whose valence and expression are lowest in the `Scope`

the ontology can be found at `sherlock.jtm.io` or included with this dissertation. Chapter 7 describes the results of analyzing an encoding of the first three chapters of the player experience study dataset described in Chapter 3,

ICE demonstrates how researchers can combine top-down theory with descriptive statistical measures to analyze content. The approach lays the groundwork for more focused research studies in the future. For example, one study could compare the role of interactivity by measuring player responses from players who were playing a game, to players who only watched the player gameplay video, to those who watched both the player gameplay video and the player responses. Story value charge can be assessed for its validity as a measure of content by examining inter-annotator agreement on segments of content drawn from the present data set. Finally, player experience measurement may be a source for predictions, as there may be relationships between content selection, response, and various other measures of experience such as post-experience surveys. It would be interesting if the method could reveal indicators that an experience would be memorable, particularly as these are sources for community discourse and a significant value proposition for the genre.

Developing ICE also raised many new questions about the value of descriptive models and interpretation in studying IDN. For example, how would one compare different versions of the same model? Is there a method for evaluating a model for its efficiency, legibility, and completeness? Can an ontology incorporate multiple different phenomena, such as the Story Intention Graph and the Interactive Cinematic Experience schemata, or are their respective encoding processes incompatible? While I am able to encode the same content in the next chapter, the results confirm that each approach attends to different aspects of the artifact: the reasoned interpretation of the storyworld, the moment-to-moment responses and feelings toward the content, and the consideration and understanding of how options relate to outcomes and likelihoods with respect to a set of goals.

There remains a lot of work in cataloging the exact relationships between lower-level choice poetics core patterns, but the schemata described in this chapter provide evidence of the types of relationships and patterns that can be revealed. The complete cataloging of beat-level content patterns that integrate choices in modern storygames has only begun. Additional work is also needed to understand the role that the genre of the story, its potential cultural resonance, and specific level of familiarity with individual players and the types of potential patterns created as a result.

Story value charges are just one possible coding system for the affective content role. The underlying contribution of ICE is to demonstrate how to relate content structures and their potential effects on an audience. Other theories could be applied using different coding strategies that follow the same approach. In Chapter 5, I compare the advantages of the story value charge encoding process with two other encodings.

A comparative analysis of *TWAU* to other games in the subgenre may reveal interesting patterns associated with the narrative genre expectations. Each genre often highlights different emotional ranges and produces different effects, given player's knowledge about the characters and franchise. These are features that are often highlighted as being a part of what makes a Game of Thrones episode feel like it belongs within the franchise, and getting this sense of flavor right is often cited as a reason that Telltale Games is selected for creating games for different intellectual properties.

One potential future direction would be to collect bottom-up classification labels for content using clustering or machine learning methods. These might classify options, content, or other elements using both the detected affect present in the game as well as player responses at points before, during, and after content events and decisions. This type of approach would be possible given a classification schema that consists of a set of interpreted labels. At this point, heuristics and human insight are still required to extract the type of information needed to interpret the empirical data relating to complex content such as storygames. I predict that various stages of the encoding process and the elements it identifies will be automated, such as identifying key actions and estimating value charges based on sentiment and affect analysis.

Finally, the process does not account for the types of experiences that result from revisiting or replaying, where a single player's traversals are encoded as two sets of instantial layer elements. For example, Alex Mitchell has investigated the effects of rereading in the context of interactive narratives [160], and his approach might be fruitfully modeled in the context of ICE as well. One of the stated goals of interactive drama and IDNs is to provide different experiences upon replaying the same work, and this is one of the principal advantages of the generative techniques such as those employed by *Façade* or *Prom Week* [103].

This chapter contributed an integrated, explicit schemata for interactive narrative content and player response data that can be used to identify and express patterns in IDN content structure. The extensions to Mawhorter's theory of choice poetics address the challenges of smaller, subjective, impressionistic choice sequences that are not always goal-oriented, but which affect players in ways that are unique to this form of presentation within traditional dramatic structures of the scene and beat. I further evaluate these ideas in the next chapter, where I demonstrate an encoding of a scene from *TWAU* using the set of player traversals.

Chapter 5

Toward Medium Reading

Choice-based cinematic adventure games (CCAGs) are similar to traditional media, and yet vary significantly based on the choices players make. This variability requires considering a different approach to encoding and studying than stories or games would on their own.

This chapter contributes an encoding process that accounts for the timing and relationship of specific character actions with respect to story values. It provides a set of annotations that can be correlated with player response, based on the observation that player response is often directly related to the perceived valences and emotional content of the scenes. It encodes different aspects of the structure of story game content when compared with goal-based choice analysis and story intention graph (SIG) encodings. I demonstrate the process by encoding a scene from *The Wolf Among Us* (*TWAU*), using a set of traversals with the ICE schemata described in 4, including identifying character actions and options that relate to thematic values and estimating their level of impact on the scene. The results of encoding multiple traversals can provide the materials for conducting a type of hybrid analysis that I term “medium reading" and describe further in

Section 5.4.

5.1 Questioning Faith

The main plot is set up in the first chapter of *The Wolf Among Us*. A summary of the entire episode can be found in Section 1.1.2. The first scene involves Bigby's chastisement of Toad for being out of glamour. It accomplishes many authorial goals simultaneously, as it must, since it is the introduction. It establishes Bigby's role as an enforcer and authority figure; moreover, it demonstrates an uneasy relationship between Bigby and Toad, which mixes duty and sympathy and further sets up the larger conflict between the classes in Fabletown.

Bigby investigates a disturbance in Toad's tenement upstairs just in time to see the Woodsman assaulting Faith, a woman whose identity is unclear. The second set of scenes pick up the pace by introducing two new characters and their respective desires. Bigby drives the scene through his desire to discover the truth of the situation by questioning Faith and the cause of the violence. The Woodsman's ego drives him not only to hit Faith when she does not recognize him but to turn his ire toward Bigby when he is interrupted. Faith's desires are more complicated because she is restrained from speaking about her purpose by the ribbon around her neck (saying instead that "my lips are sealed"); in fact, apart from an assumption about her occupation, her goals are unknown to the player even at the end of Episode 1. The exact motivations and identity of Faith that is present in the first chapter is, in fact, the subject of debate in community forums, as Telltale Games has left open the interpretation of whether or not she was another character in Glamour, with evidence supporting either side. Stories that incorporate ambiguity and uncertainty are appealing, and so it is worth developing analytical methods that can address them and their effect on the player experience.

After Bigby and the Woodsman struggle, Bigby attempts for a single beat and several choices to obtain some information from Faith, which fails. The defeated Woodsman insults Faith once more, triggering Bigby (regardless of player choice) to tackle the Woodsman through the side of the building and having Bigby land on Toad's car.

At this point, Toad is shown stunned next to his destroyed car, participating in a single beat that reverses the prior power dynamic between the two. Bigby has, instead of achieving the scene goal of justice, further hurt Toad's finances by destroying his vehicle and part of his business. The Woodsman is revealed to be still alive and struggles with Bigby, including a QuickTime event where the player rapidly taps a button to represent Bigby's struggle to avoid being strangled, causing Bigby to begin to transform into his Big Bad Wolf form. Faith rescues him by swinging an axe into the Woodsman's head. The transcript of the next scene is presented in the following pages, and I provide a detailed example of its encoding in the following section.

5.1.1 Notes on the Transcript

First, I provide a few comments on the transcript notation and relationships. The transcript is simply a tool to describe the protostory encoding process. It is useful for describing the content in a textual form, although the use of a video browser is preferred. The `StoryValueChargeComponents` are assigned to key events, but not every event, line, or action has sufficient emotional charge to be annotated with one. The transcript is formatted in a manner that is standard for scripts with the addition of protobeat and protochoice headers, indicating points where the `ProtoBeat` and `ProtoChoice` elements would be segmented. `ProtoChoices` are described further in Section 4.3.2.3 and are indicated by labeled headings, while `ProtoOption` instances are arranged according to the selected `Option` by at least one player, which means that some options do not have the content that results included. Instances (elements on the instantial layer from Chapter 4) are not indicated in the transcript but are annotated during each stage of the encoding for each traversal. `ProtoChoices` are referenced starting at "PC1," and `ProtoOptions` are referenced starting at "PO1". Each column in the transcript following a ProtoChoice represents dependent content that is present after selecting that option.

The transcript includes marginal labels for `StoryContentEvent` instances from ICE, described further in Section 4.3. These consist of **SpeechAct**, **PhysicalAct**, and **Spectacle**. For Faith, these would be indicated as FS1, FA1, whereas if any spectacles were present, they would begin with "S1" as they are not attributed to an agent.

SpeechAct represents a line of dialogue that contains a single meaning. These can be divided if two distinct purposes exist, but are most useful when they share an interpretation of intent and reception. They are identified by an integer and referenced in this chapter using a two-letter abbreviation "XS" in the right margin and in the encoding description, with X being the first letter of the character's name. For instance, Bigby's first line is noted as BS1. Faith's begin at FS1, while the Woodsman's begin with WS1.

PhysicalActs are any depicted performances that may be significant, such as movement, gesturing, or performing an expression, and are represented in the model as `PhysicalAct`, and each instance in each traversal is connected through a single `ProtoPhysicalAct` parent. These are noted as XA, using the same format as speech acts. Since the scene is presented cinematically, some actions are completed or started off-camera or where the player is unaware, such as Faith's retrieval of the Woodman's ax and her swing, but inferred actions are omitted. This is the case for the Woodsman getting up and leaving.

Spectacles are content events where a person, object, or environment are shown to the player and have an affective impact but no immediate agent. Some spectacles are subsequently learned to be the action of an agent, but the encoding focuses on the ordering and sequence of visible events. An example of a spectacle in Episode 1 is Faith's head. The head is presented to the player (and Bigby) even though the significance is not based on the act of revealing the head or even the implied act of some Fable leaving the head (as that action takes place off-screen). The head's owner is related, but not the agent of the spectacle.

Action*/*Reaction are behaviors that can take place over multiple **PhysicalActs** and/or **SpeechActs.** Each `ProtoBeat` has a pair of `ProtoAction` and `ProtoReaction` elements associated with it. These have corresponding `Action` and `Reaction` instances that represent the core behavior and response that takes place over the course of the beat. Each `ContentEvent` that enacts those behaviors is linked to the `Action` or `Reaction`.

All `ProtoContentEvent spanID` have been renumbered from 1 for ease of reference. They normally start at the beginning of each traversal. It is worth noting that modeling each action and line of dialogue is *not* necessary for a complete encoding. That being said, it can be useful to simply encode each line for completeness and for determining which actions and speech acts are attributable to specific story values. For a valid ICE encoding, only actions that directly relate to a story value charge must be included.

This is an encoding of merely the protostory layer of the content based on a transcript and the set of videos. The estimation of SVCCs comes in a later phase, and an analysis of player response based on an encoding of the first three chapters can be found in Chapter 7.

5.1.2 Faith and Bigby Transcript

`ProtoChoice`

ProtoBeat 1

FAITH looks concerned a moment then turns an angry look to WOODSMAN **FA1**

Woodsman crawls a few steps and then collapses **WA1**

FAITH crouches and begins to search the WOODSMAN **FA2**

ProtoChoice 1

PO1 *Thanks...*	PO2 *What are you doing?*	PO3 *Leave him alone.*
BIGBY **BS1** Thanks. FAITH pauses, looking back **FA3** FAITH **FS1** Don't mention it.	BIGBY **BS2** What are you doing?	BIGBY **BS3** You can't be doing that. PO4 ...

FAITH **FS2**

I'm just getting what he owes me.

FAITH **FS3**

You alright back there? I mean... Your eyes. and the teeth... you're not really supposed to do that, are you?

BIGBY **BS4**

Not if I can avoid it.

FAITH finds a few coins and throws them down in frustration. **FA4**

FAITH **FS4**

Great.

ProtoBeat 2

FAITH stands and begins to kick the WOODMAN's unconscious body **FA5**

BIGBY **BS5**

The guy's got an axe in his brain. He's not feeling that.

FAITH **FS5**

It's more for me. He'll be fine.

WOODSMAN **WS1**

Aisllle killl yew. Yew fuckin' bidch

Woodsman begins flailing arms trying to get at the axe. **WA2**

FAITH **FS6**

Let me help you with that.

FAITH steps on the axe, looking determined **FA6**

ProtoChoice 2

PO5 *Let her*

BIGBY **BS6**

Guy's having a bad day.

FAITH **FS7**

That makes two of us.

FAITH continues to push down with her foot on the axe head. **FA7**

FAITH **FS8**

I guess it's a good thing Fables are hard to kill.

BIGBY **BS7**

Suppose it is.

PO6 *Stop her*

BIGBY takes FAITH by the arm and pulls her back. **BA1**

BIGBY **BS8**

He's had enough

FAITH **FS9**

He's a Fable... sonuvabitch can take plenty!

ProtoBeat 3

FAITH walks over to the door and looks in purse, dissapointed. **FA8**

BIGBY begins pulling axe out **BA2**

She picks up her purse before starting to walk away. **FA9**

BIGBY finishes pulling axe out and tosses it down, looking after her. **BA3**

FAITH tries to light cigarette with a lighter. **FA10**

FAITH **FS10**

Shit... just... come on...

BIGBY follows and pauses as she continues to try to light **BA4**

ProtoChoice 3

PO7 *[Give her the matchbook]*

FAITH **FS11**

Thanks.

FAITH strikes a match back and leans back and puffs. **FA11**

PO8 *[Light her cigarette...]*

BIGBY **BS9**

Here.

BIGBY walks up to her and offer her light **BA5**

FAITH leans down and accepts. **FA12**

FAITH **FS12**

Thanks.

FAITH leans back and puffs. **FA13**

PO11 *Got an extra?*

BIGBY **BS10**

Can I bum one of those?

FAITH **FS13**

Sorry... Last one...

FAITH gets her lighter to work and lights her cigarette. **FA14**

PO9 *Make a joke...*

PO10 *...*

ProtoBeat 4

BIGBY **BS11**

Who do you work for, really?

FAITH **FS14**

These lips are sealed... sorry...

FAITH **FS15**

Hey, you like my ribbon?

FAITH gestures to her neck, her expression stressed. **FA15**

ProtoChoice 4

PO12 *I'm trying to help you.*

BIGBY **BS12**

If you don't answer my questions, I can't help you.

FAITH **FS16**

I'm answering them, the best that I can.

PO13 *Stop changing the subject.*

BIGBY **BS13**

Stop changing the subject.

BIGBY **BS14**

If you don't answer my questions, I can't help you.

CT1: She'll remember that.

FAITH **FS17**

I'm answering them, the best that I can.

PO14 *...*

BIGBY looks on silently.
BA6

FAITH **FS18**

FAITH: What's the matter Bigby, tempted to take a bite?

FAITH **FS19**

FAITH: I guess those days are behind you.

CT2: You avoided her question.

PO15 *Beautiful...*

BIGBY **BS15**

I feel like we've met before...

FAITH **FS20**

We probably have.

FAITH **FS21**

We all sort of knew each other at one point or another. But things change I guess.

BIGBY **BS16**

I guess.

BIGBY goes over to see that WOODSMAN is gone. **BA7**

Faith walks to stand beside him. **FA16**

——— ProtoBeat 5 ———

BIGBY turns, intending to pursue him. **BA8**

FAITH grabs BIGBY's arm **FA17**

FAITH **FS22**

Stop. We don't have to make any more of a thing out of it than it already is.

ProtoChoice 5

PO16 *Are you sure?*

BIGBY **BS17**

BIGBY: Are you sure?

FAITH **FS23**

Eh, he's had enough. For tonight at least. Wouldn't be surprised if he has an axe to grind though.

BIGBY gives her a look. **BA9**

FAITH **FS24**

That wasn't intentional... sorry.

PO17 *This is about Fabletown.*

BIGBY **BS18**

This isn't just about you. I can't have him running around the city in the state he's in.

FAITH **FS25**

Good. Cause I was just starting to worry it was all about me. (hurt).

BIGBY **BS19**

That's not what I mean.

PO18 *He hit you.*

BIGBY **BS20**

He hit you. He needs to pay for that.

FAITH **FS26**

He's got nothin' to give, Sheriff. I checked.

FAITH **FS27**

Besides, I just swung an axe into the guy's head. I'd say we're even.

BIGBY **BS21**

Maybe I should be arresting you.

FAITH **FS28**

I'm not going to be doing this for much longer.

FAITH **FS29**

So now's your chance.

PO19 *...*

FAITH **FS30**

I'll be fine. Really. Don't worry about me.

ProtoBeat 6

FAITH removes her hand from BIGBY's arm, turning away from him. **FA18**

FAITH looks worried and scared toward the camera. **FA19**

BIGBY **BS22**

How much was it he owed you?

FAITH **FS31**

A hundred...

BIGBY **BS23**

...and I'm guessing it'd be bad for you to show up empty handed.

FA21 FAITH turns toward him and smiles wanly. **FA20**

FAITH **FS32**

I'll be fine.

ProtoChoice 6

PO20 *[give her some money]*

BIGBY **BS24**

Let's see... I've got uh... It's twenty, forty... seven– eight. Fifty eight. It's all I got– it'll help, right? Take it.

FAITH looks at the money and up at BIGBY. **FA21**

FAITH **FS33**

It's okay. I'll be fine. You've done enough. Really.

BIGBY **BS25**

Just take the money, okay?

CT3: You chose to give her money.

PO21 *Wish I could help*

BIGBY **BS26**

I really wish I could help.

FAITH **FS34**

It's okay. I'll be fine. You've done enough. Really.

CT4: You chose not to give her any money.

FAITH **FS35**

You got me out of a bad situation back there. Thanks.

5.2 ICE Encoding

This section describes the encoding in the context of a specific demonstration, along with a set of traversals. Because of the interpretive nature of the encoding, many of the final values and choices may differ from analyst to analyst. What is important is not whether the labels are correct, but whether the relative values capture the dynamics at work.

This process takes advantage of multiple actual player traversals, not only supporting the goal of correlating player responses to the content but also accurately reflecting the traversals that involve real player choices and the resulting content, as it only encodes content that some player has selected. The more information associated with each traversal, including biometrics and think-aloud, the more an analyst can discriminate between instances of mistaken choices and intended decisions. The encoding process highlights authorial decisions regarding choice spaces and specifies a regular process for classifying segments of content experienced during a player traversal. Labeling content for its interpreted affect can reveal how story values relate to individual options without using propositional logic to represent meaning. Without a structure to interpret them, player choices are difficult to compare between traversals, or even to compare among the choices in the same traversal. Finally, labeling content in a set of traversal timelines reveals timing relationships and hidden dependencies that may vary from traversal to traversal.

Such an encoding is made easier by a tool such as Sherlock, which displays player traversals side by side and supports annotating directly within the timelines. Indeed, Sherlock's annotation interface was co-developed alongside the encoding process with the goal of highlighting variations in the content and making it simpler to review segments of the content for transcription and evaluation. The more the traversals that were incorporated, the better the potential for representing not only the structure of the content but the possibility space for players who navigate it.

This approach identifies the forces at work within a scene of drama to contextualize the choices that players are offered and select. However, instead of focusing on modeling the propositional content of actions or words, I focus on the impact of various actions and reactions to the overall value charges, which may or may not directly relate causally to individual agents' affect. This distinguishes the ICE encoding process from the proposition-network approach of Story Intention Graphs, where each interpreted proposition must relate to an affective impact on an agent as well as the goal-based choice analysis described in Mawhorter's dissertation.

5.2.1 ICE Encoding Process

For any given scene, the encoding process can be divided into the following six steps:

1. Context Analysis
2. Action Decomposition
3. Beat Segmentation
4. Story Value Assignment
5. Story Value Charge Estimation

These steps progress from one to the next, although it is useful to occasionally revisit a prior step. The steps and a general description follow.

Context Analysis.

1. What is the overall situation thus far? What are the genre conventions referenced? What recent events occurred within the story? What have players learned about the characters involved? What are their character's scene goals?

2. This step breaks down the overall structure and accounts for the overall forces leading up to this point, as well as where the story may be headed. Knowing the complete story is useful but not necessary.

3. After this step, all major `ProtoChoice` elements and `ProtoOption` elements should be associated with individual `Choice` and `Option` instances with start and end points in each traversal. Each `Option` includes the selection made by a player, as well as whether or not it was displayed, whereas the `ProtoOption` contains the choice text and any changes to `StoryStates`.

Action Decomposition.

1. Which actions play a role in shifting the value perception in the scene? This does not yet assign specific story values to actions but rather selects actions that feel particularly significant, including all types of actions: dialogue, gesture, tasks, often including attention to paralanguage, which is the manner and style of an expression that could substitute or enhance speech. For example, when Bigby lights Faith's cigarette, the distance between them closes both physically and socially in the sense of *proxemics*.[1] In addition, the affective intent can also be in the choice of shot distance (close, medium, long) or duration. A longer shot of a character's expression may indicate

[1] Film studies scholars often refer to Hall's four patterns of proxemics [161]: **Intimate** (skin contact to 18 inches), **Personal** or 18 inches to 4 ft, **Social** 4ft-12ft and **public**, or 12-25 ft.

its significance, such as when Faith responds positively to Bigby's "Thanks" (PO1, BS1, and FA3).[2]

2. These actions and their start and end locations should be encoded as instances of the appropriate `ContentEvent` subclass for each player, ideally while simultaneously stepping through the same scene in each traversal.

Beat Segmentation.

1. Now that the significant actions have been identified and encoded, how can the scene be decomposed into action/reaction pairs, usually with a gerund (-ing) as a description? How do these actions communicate a character's pursuit of a desire?

2. This is useful for spotting repeated actions, as is the case in Bigby questioning Faith in the first sequence in the Woodsman's apartment. It also helps with distinguishing which character originated the action, and whether or not that action was part of a choice. If any actions are relevant to the beat at this point, they can be encoded as per the previous step.

3. The results of this step are `ProtoBeat` and `Beat` instances, as well as `ProtoAction`, `ProtoReaction` and `Action` and `Reaction` instances corresponding to respective behaviors. These are assigned to the `ProtoBeat` through the `ofBeat` relation, and each `ContentEvent` that pertains to an `Action` or `Reaction` element is associated through the `demonstratesBehavior` relation.

Story Value Assignment.

[2]PO1: ProtoOption 1, the first ProtoOption. BS1: Bigby SpeechAct 1, and FA3: Faith PhysicalAct 3.

1. Using the context and action decomposition analyses results, I come up with a list of story values (often in the form of pairs) that could potentially describe the actions taken. What do these characters desire? What are the reasons they could or did evoke sympathy from the audience? Why are these characters taking the actions they do? What are the stakes, what does a character stand to gain or lose by pursuing a course of action? Unlike the representation in SIG, the full goal as a series of anticipated or projected steps is not necessary. Here, values that are responded to negatively by the players are just as worth recording as values that are attractive to players. For example, the intimacy value is considered off-putting by P6, whereas P2 embraces the role-playing opportunity of lighting Faith's cigarette in PC3.[3] Thus, the valence of a value does not determine the player response; it is often the opposite.

2. With a list of values and actions, the goal then becomes to narrow it to two to four story values that play the largest role. A story value is encoded by associating an instance of `StoryValue` using `StoryValueChargeComponent` with the `ProtoStoryContent` instance. Select values that change across the scene. If the value does not change, then the action it is associated with may serve some other purpose such as exposition or characterization, or it may be put in as a distractor. PO11[4] (in PC3, "Got an Extra?") is a good example of a distractor: it does not further any story values and, in fact, represents a mistake in a social sense that is commented on by the player who chose it.

Story Value Charge Estimation.

[3]PC3: ProtoChoice 3
[4]PO11: ProtoOption 11

1. At this point, the goal is to represent the relative value changes that each action represents. This can be small (-1/+1), moderate (-10/+10) or large (-50/+50), depending on the overall change in perceived value. The dynamic range represents the most significant actions or beats in the game, which can be assigned +/- 100 and these can be compared to the least, which correspond to +/- 1.

2. These values are assigned to the respective `StoryValueChargeComponent` through the data property `hasValue`. Although these charges share the concept of valence with player and character expression valence, they are distinct concepts. The valence represented here is the *content* valence, as opposed to the *experienced* or *performed* valence.

This subjective analysis results in a graph of traversals of the scene that relates story value charge, distribution, and relationship to a sequence of beats and choices. Each alternative traversal can be evaluated for its relative impact on the resultant actions on the beats they occur in, including peaks, differences in subsequent actions and the last value, all of which can be worth considering in how a choice may have an effect on the overall beat. I discuss some of these possible patterns in Section 4.4. Examining player affect and videos of player responses can help in estimating the potential impact, in particular examining engagement and valence following a choice.

The following sections break down and demonstrate each stage of the encoding process for the scene above.

5.2.2 Context Analysis

The context analysis phase of the encoding process begins with the introduction of Section 5.1. Faith has just been assaulted. She rescued Bigby from exposing his wolf form, which would result in negative consequences for the community if witnessed. Bigby has just hurt Toad, who called for his help, by damaging his tenement and his car. Bigby wants to help and has been motivated to protect Faith by defending her in the Woodman's apartment, though he was goaded by the Woodsman into attacking after the clear danger had passed.

Molly Maloney, in her talk at Konsoll 2017, describes "role-play rails" as one of the guiding principles for designing choices. This concept involves incorporating character "pillars" as options. Character pillars, in Maloney's definition, are core aspects of the character's nature or personality. Bigby has several core pillars that are evident from his actions. He has a strong sense of justice. He has a desire for intimacy, perhaps owing to his lone wolf existence. Third, he has a strong sense of loyalty to Fabletown, as it has given him a role and a place within it as sheriff. Two of the rails are represented in the value pairs `Justice/Injustice` and `Duty/Preference` (loyalty), the third by the `Intimacy/Isolation` pair.

Examining the relationship is also worthwhile. Bigby and the Woodsman have a history stretching back to the homelands, one which involves the Woodsman besting Bigby in the events that are told as the story of Little Red Riding Hood. Bigby witnessed Faith being attacked, which resulted in a protective response, triggered again when Faith was insulted.

Finally, what are the emotional and player goals associated with this scene? In this case, the emotional goal is for the player to feel attached to Faith. The player goal is to find out what might help Faith or Fabletown (by investigating the original cause for the disturbance and Faith and the Woodsman's roles in it). Player goals are often synonymous with player-character scene goals, particularly in the role-play mode.

5.2.3 Action Decomposition

The action decomposition is represented in the transcript in section 5.1, although each action would also be associated with a subclass of `StoryContentEvent` and located in the player timeline.

Several actions that were part of a sequence were consolidated into a single action (such as pulling the ax out of the Woodsman's head and throwing it down), as the individual parts were not significant in themselves. I included actions in the transcript for completeness, but only those that explicitly change the value charge are necessary. These are described alongside their estimation subsequently in the chapter in Table 5.2. I describe story value assignment in Section 5.2.5. If encoded, they can be used for generating transcripts or other types of analyses.

The primary criteria for identifying whether a segment of content is necessary to encode as a `ContentEvent` is its pragmatic role within the sequence. Does it participate in an action through a character or a response? Is it distinct from adjacent actions or speech acts? Does it communicate a unit of meaning to an audience with respect to the feelings and behaviors of the characters involved?

5.2.4 Beat Segmentation

This phase of the encoding involves identifying action/reaction pairs in the form of gerunds (-ing words), as well as linking the set of actions that perform them to the respective beats. A single beat, and thus a single `ProtoBeat,` can incorporate repeated actions, if they are in fact different expressions of the same underlying behavior, and each of these actions is in turn associated with the `ProtoAction` element. Minor actions can also be present in a beat, as long as it does not shift the orientation toward a different intention on the part of one or the other agents involved.

Silence can have a powerful effect and Malone describes it as a requirement for Telltale narrative designers to include it in most of the choice spaces in Telltale Games [40]. The interpretation of the player character's silence often depends on the expectations of the other character and the implications of the feeder line (discussed in Section 2.2.3).

The decomposition of the scene into beats that is indicated in the script is shown in Table 5.1.

Table 5.1: Example Beat Decomposition

Bt	**Action**	**Reaction**
1	Faith: Looting	Bigby: Variable (PC1)
2	Faith: Assaulting	Bigby: Variable (PC2)
3	Faith: Leaving	Bigby: Pursuing
4	Bigby: Questioning	Faith: Answering
5	Bigby: Pursuing	Faith: Restraining
6	Faith: Despairing	Bigby: Empathizing

There are 6 `ProtoChoice` (PC) instances taking place within 6 `ProtoBeat` (PB) instances. Each choice fits into its respective `ProtoBeat` in different ways, so let us consider each individually.

Protobeat 1: Looting/Responding. Faith, having just rescued Bigby, moves to her scene intention, which is to fulfill her obligations to her employer and to connect with Bigby. Bigby's response is determined by PC1, which includes one option that increases intimacy (gratitude, PO1[5]) and three that increase the story

[5]PO1: ProtoOption 1

value of duty (PO2-PO4). The silent option is implied to be a judgment of her action, given his role, though neither this nor the default option[6] was chosen by any of the players.

Protobeat 2: Assaulting/Responding. After finding nothing, Faith proceeds first to kick, then torment the Woodsman by pressing down on the ax lodged in his head. Bigby did not move to stop Faith from kicking him, and the player is only offered an option once the violence has escalated. Faith's revenge represents a sort of justice for his having beaten her earlier, and foreshadows the later decision Bigby makes regarding Grendel's arm to deliberately torment a fallen opponent. Bigby's choice reflects a conflict between empathizing with her violent desire for justice (PO5), and his duty to protect Fabletown residents (PO6), and the choice players make here is one of many regarding violence and justice. Players were evenly split.

Protobeat 3: Leaving/Pursuing. Faith starts walking away, trying to light a cigarette. Bigby has yet to complete his scene intention, which is to find out why the disturbance took place, so he pursues (BA6).[7] The player is offered an opportunity to pick how Bigby begins a new conversation with Faith (PC3)–one of the first actions that the player chooses that is an initiating action (as opposed to a reaction). Option PO7 depends on whether Bigby picked up a matchbook earlier, but it is a popular choice as everyone who has the option takes it. Player 3 selected option PO11, but said aloud "That's not what I wanted to say." This indicates an assumption of the role, and a sense of dissatisfaction with the loss of control (a mistaken choice), and the negative response from Faith to the resultant

[6]Default options are the option selected when the time runs out in a timed choice. This is most often the silent option, but in the case of giving Faith money, it represents not deciding to give her money.

[7]BA6: Bigby PhysicalAct 6 from the Transcript

action confirms that it was not a desirable option.

Protobeat 4: Questioning/Answering. Bigby asks about her employer (BS7), which Faith does not answer. Instead, she asks Bigby if he likes her ribbon. The player is given an opportunity to interpret the response (PC4), but not initiate a new action. Faith's response can be considered a continuation of the questioning beat despite the apparent change in subject, given that she is attempting to answer his question.

Protobeat 5: Pursuing/Restraining. Here, Faith uses the opportunity of Bigby's attempting to pursue the Woodsman to initiate extended contact with him for one of the more value-oriented PCs. In it, the player must choose between valuing intimacy with Faith (PO18) or an assertion of community and duty (PO17). Deciding to pursue community results in a hurt response from Faith, while expressing intimacy inspires Faith to confide an essential piece of information (FS27).

Protobeat 6: Despairing/Empathizing. Faith turns the response into an action that positions Bigby to feel empathy for her position as she turns away and let go of his arm (FA17).[8]

Faith consciously removes her hand from holding Bigby's arm and turns away, her expression suggesting despair and self-consciousness. This action evokes a response from Bigby empathizing with her situation (BS13). The player's choice factors into whether this response includes an act of charity (PC6), or whether it is purely a recognition and expression of sympathy, but it does not negate the empathetic gesture, and Faith recognizes that by returning to the value of intimacy with gratitude (FS30). Interestingly, the money in the player's inventory serves as a reminder of this choice for the rest of the game; as Player 4 noted, it may have

[8]FA19: Faith PhysicalAct 19

played a role in the character's death, and that alone serves as a potent source of regret. Bigby's recognition of her situation forms the turning point of the scene, specifically with BS23, FA19, and FS31. They both understand that the situation is dire, that there is no real help for what's happening that put Faith into this position. This moment of facing away from one another represents the pinnacle of intimacy and the nadir of duty, as Bigby's job is not to fix her situation, regardless of whether he offers her his money or not. Each action is assigned to a beat by associating it either directly to the `ProtoBeat` or through a `ProtoReaction` or `ProtoAction`, thereby indicating the role of the action.

Some lines and actions are not associated with beats as part of an action/reaction pair. In particular, lines BS15, FS20, FS21, and BS16 are out of place in the sequence but are used in the final montage of the final sequence to suggest that Faith and Nerissa may, in fact, be the same person. This theory is discussed in section 5.3.1, but reflects the ambiguous ending to the series where Nerissa says the same lines as Faith during this scene and the general theory is that one was as glamoured as the other. While the notion of beats as a means of organizing content is useful, there is a need to experiment with other content labels to account for the diverse roles that content can play in shaping the player response.

This view of beats represents actions as a series of negotiations between the player and the character with regard to particular values. This scene is critical, as it sets up Bigby's realization of his powerlessness to stop the death that drives him through the rest of his quest to find justice.

5.2.5 Story Value Assignment

There are two steps here: developing a list of story values and narrowing it down to two to three, with as many as four.

I identify four forces at work behind each of the beats: **intimacy/isolation**, **community/ego**, **justice/injustice,** and **duty/preference**. These four values are encoded in ICE as instances of `StoryValue`, as they are involved in the protostory layer.[9] For intimacy/isolation, these physical acts or dialogue lines can include flirtation, a form of romantic intimacy, or times when a character is confiding in another. Flirting could be judged by actions involving proximity (whispering in Bigby's ear, lighting Faith's cigarette, holding on to Bigby's arm for a longer period), or by the nature of the dialogue, such as playful teasing. For example, Faith's response to the silent option in PC4 represents an attempt by Faith to re-establish a playful intimacy with Bigby (FS17: "What's the matter Bigby, tempted to take a bite?").

The core values of the story are Bigby's desire for justice and community and these are often at odds with opportunities for intimacy and duty. Far from a simple distinction between a good wolf and a bad wolf, Bigby represents a complex character with a negative reputation seeking a place within his community.

Table 5.2 depicts a set of assignments for actions that represent story values in the scene, as well as the general direction the actions move in or set the value to. It is clear that the majority of actions from this stage of the encoding process revolve around intimacy and isolation, but not every action influences the scene to the same extent, which is addressed in the next phase. Each is further annotated by whether the action has a negative or positive valence.

[9] In an editor, these may be assigned symbols such as `StoryValue1`, `StoryValue2`, etc, using the "twau" namespace.

Table 5.2: Story Value Assignments

Story Value	Actions
Duty/Preference	-BS6,+BA1,-PO5,+PO6,+PO14,+BA8,-FS21
Community/Self	+FA6,-BS10,-PO11,-PO9,-PO10,+PO12,-FS20, +PO17,+BS18,-BS26,+BS25
Justice/Injustice	+FA2,+BS2,+BS3,+FA5,+FA7,
Intimacy/Isolation	+BS1,+BS7,+FS3,+PO5,-FA10,+PO7,+PO8,+FA15 -PO14,+PO15,+BS15,+FS17,-FS16,+BS15 +BS15,+BS16,+PO16,-PO17,+PO18,-FS29,-FS24 +BS20,-BS21,-FS27,-FS29,-FA18,-FA19,+BS22 +BS23,-FS31,+BS24,-BS26,+FS33,+FS32

5.2.6 Story Value Charge Estimation

Story value charges are represented in ICE as instances of the `StoryValue-ChargeComponent` class, with `OptionComponent` and `OutcomeComponent` corresponding to the ideas of the prospective impression of an option concerning a story value and the actual events that occur as a result of an option, as used in Mawhorter's choice poetics. The primary difference is that these are calculated individually with respect to a story value rather than with respect to a player goal. Moreover, these values are interpreted at the surface level of the reception of the content, and there is no estimate of the likelihood of a given outcome, as is the case in choice poetics. This focuses on the immediate evaluation of the impression that accounts for the timed nature of many of the options and the dramatic presentation. Each `StoryValueChargeComponent` instance associates an integer value that represents the change in charge due to a specific act, ranging from -100 to 100 with a specific `StoryValue` instance from the previous step. These can be separated into a valence (+1/-1) and a magnitude (1-100).

The numerical scale is fundamentally subjective. However, it does mean that different assignments of values to each story value component could be possible, and this subjective range is worth future investigation in the form of studies determining whether participants agree on assignments to action. Other coding systems are also worth exploring, as story value is simply one of many possible labeling schemas that capture the dynamics and meaning involved in the reception of content. For this encoding, I went through each action and assigned either a positive or negative valence based on assessing the act's weight and valence. For example, f, there is a small magnitude positive story value of Bigby realizing that the Woodsman is gone. The small magnitude reflects the fact that he is easily stopped by Faith and the act takes place quickly compared to some of the other content segments.

On the other hand, Faith's response (FS34) to Bigby's choice of not giving money (PO21) functions as a significant value charge. She defuses the potential tension that could result from the player's choice not to give money and sets up the following beats to have an opportunity to attain the ultimate goal of increasing the player's affinity for Faith.

A full assignment of relative and absolute story value charges for the scene can be found in Appendix A; the final two beats are presented in Table 5.3. The ProtoChoice (PC) represents the associated ProtoChoice for each ProtoOption, while the Parent ProtoOption (PPO) column indicates which content is dependent on which `ProtoOption`. In addition to valence and magnitude, the players that experienced the content are also listed.

In an encoding, each instance of a subclass of `ProtoContentEvent` or `Proto-Option` is assigned the charge using a `StoryValueChargeComponent` subclass based on the relationship of the event to a choice. If it is an option, then it is associated using an `OptionComponent`, if it is a result, an `OutcomeComponent`, and if neither, then an `ActionComponent`.

Table 5.3: Story Value Charge Estimates for ProtoBeats 5 & 6

X	ID	V	M	StoryValue	PC	PPC	PB	Players
26	BA8	1	5	Duty/Preference			5	All
27	PO16	1	1	Intimacy/Isolation	5		5	4
27	PO17	1	10	Community/Self	5		5	2,5,7
27	PO17	-1	10	Intimacy/Isolation	5		5	2,5,7
27	PO18	1	10	Intimacy/Isolation	5		5	3
28	FS24	-1	1	Intimacy/Isolation		5	5	4
28	BS18	1	10	Community/Self		5	5	2,5,7
28	BS20	1	10	Intimacy/Isolation		5	5	3
29	BS21	1	5	Intimacy/Isolation		5	5	3
30	FS28	-1	10	Intimacy/Isolation		5	5	3
28	FS30	1	5	Intimacy/Isolation		5	5	6
31	FA18	-1	10	Intimacy/Isolation			6	All
32	FA19	-1	5	Intimacy/Isolation			6	All
33	BS22	1	5	Intimacy/Isolation			6	All
34	FS31	-1	10	Intimacy/Isolation			6	All
35	BS23	1	20	Intimacy/Isolation			6	All
36	FS32	-1	10	Intimacy/Isolation			6	All
37	PO20	1	20	Intimacy/Isolation	6		6	2,3,5,6,7
37	PO21	-1	20	Intimacy/Isolation	6		6	4
38	BS24	1	20	Intimacy/Isolation		6	6	2,3,5,6,7
39	FS33	1	10	Intimacy/Isolation		6	6	2,3,5,6,7
38	BS26	-1	20	Intimacy/Isolation		6	6	4
38	BS26	-1	20	Community/Self		6	6	4
39	FS34	1	10	Duty/Preference		6	6	4
40	FS35	-1	10	Intimacy/Isolation			6	All

The scene's turning point resides in the sequence BS22, FS31, BS23, FA20, and FS32 just before PC6. This is the moment where Bigby discovers the true depths of the situation, where Faith finally admits how much trouble she is in and Bigby expresses empathy for her. This is important, as it functions as a fulcrum for both the player goal and the emotional goal of the scene. The actual choice (whether or not to give Faith money) is just a piece of Bigby's response to this critical revelation, even though it may inspire feelings of regret later. These high stakes set the high point for this scene at a magnitude of 20.

Time steps (X-axis) are represented by integers. Some actions have multiple story value charges associated with them. The orderings are determined from the relative start and end points of the respective subclass of `StoryContentEvent` or the `StoryOption`.

The magnitude of a story value charge associated with a `StoryContentEvent` can be difficult to get right, but a scale of 1-5-10-20-50-100 appears to capture the relative scale of values, although this scene does not achieve the same magnitudes as others.

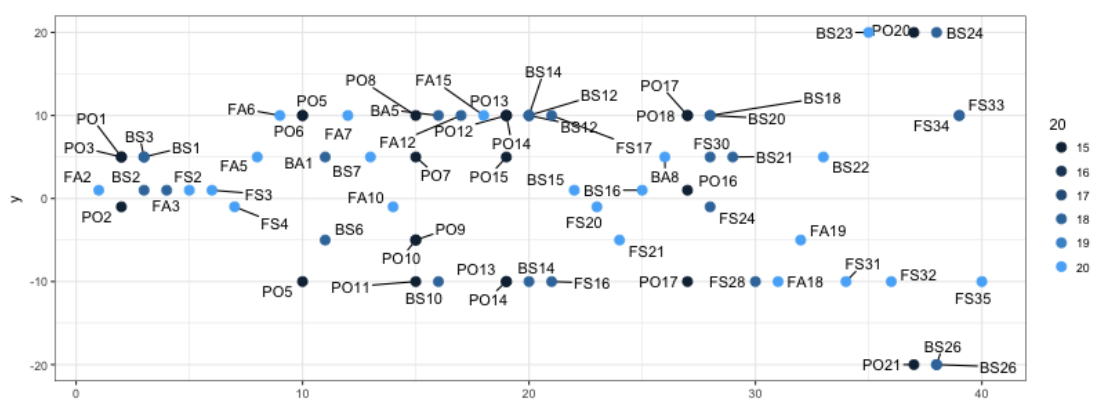


Figure 5.1: All Story Value Charge Assignments. Note lines indicate the association of labels to points in the scatterplot.

After this step, there are a set of annotations that relate individual traversals with both choices and protostory elements. The combination of all value charge assignments for each of the values can be seen in Figure 5.2.

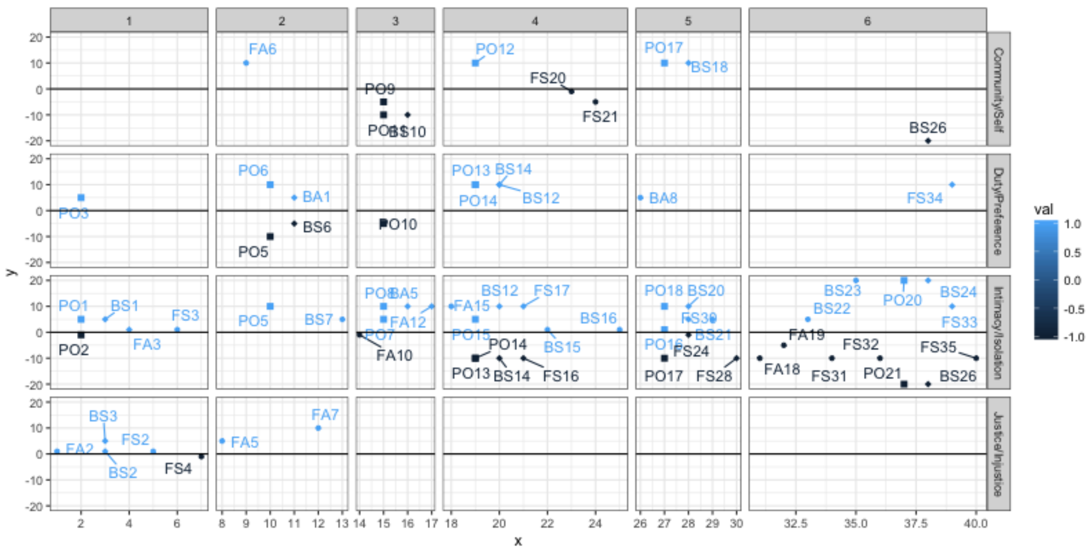


Figure 5.2: Value Charge Assignment By Story Value and Beat

A few things become apparent at this point. The distribution of story values may hint at shifting focuses within the scene. For example, the Justice/Injustice value is primarily in the early beats but is largely absent from subsequent beats. Some values may have higher stakes; these are core values for this particular scene. Sorting and evaluating acts and their relative values is one way of ensuring an accurate judgment of their impacts. Again, these are subjective measures of fundamentally immeasurable features of the content, but the more the detail that becomes available, the greater the number of patterns that can be identified and the greater the number of comparisons that can be made between the experience of particular traversals.

5.2.7 Discussion

The types of choices that this encoding lends itself to are those that are more value-focused and less strategically oriented, and the insights it reveals are based on a combination of interpretation of the individual player responses and the distribution of the different traversals in terms of story value charges. This provides several insights into the nature of choices in Telltale Games, in particular, how the choices players make are distributed and what the choices are really about. In this subgenre, player agency is severely constrained. It is replaced by a means for players to express their feelings and responses, which provides a different type of satisfaction than the satisfaction of taking meaningful actions to change the story.

These choices are reflected in the value-charged nature of the content, where the outcomes to decisions are not strategically significant, but where the possible value implications are made legible through the presentation of the choices themselves and of the expressions of the player character. Bigby leads many of the beats and sets up the player with expectations regarding the possible responses should the player decide to select a negative valenced option.

In coding the various actions, one common trend was that there were no neutral story value charges before a choice space (called a feeder line or feeder situation). This resulted in a value trend that positioned the player's selection of a choice for the player character into a responding role that asked how they felt about the previous action. FS22 does not have a story value charge, and so it is one of the most balanced actions, although its option selection has two ProtoOptions in the ProtoChoice that follows it that have an intimacy value associated with them. Here, also, the magnitude comes into play. "He hit you" is a much more intimate statement, whereas "Are you sure?" is far more ambiguous and potentially negative one.

Finally, the coding of the silent option is a challenge, as it is often uncertain how the other character will respond. It can even function a bit as a wildcard. I delve into the details of the implications for what this encoding reveals in Section 8.1, consolidating many of these observations as part of the Telltale effect.

In the next section, I use subsets of this scene to compare the encoding to two other models that map aspects of narrative and choice poetics and discuss the aspects that each attempts to model.

5.3 Comparison to Other Models

Two models that served as inspiration and as starting points also provide a means of encoding interpretive content, although they prioritize different goals in their representations. In this section, I compare the ICE encoding process shown above with two formal models of story and choice, the Story Intention Graph by David Elson [47] and the goal-based choice analysis approach by Peter Mawhorter.

5.3.1 Story Intention Graph Encoding

One challenge of encoding a single scene using the Story Intention Graph is that it requires a complete reading of a story to encode it. For example, when modeling the two character's actions as part of larger plans, it is unclear (at this point) precisely what their goals are or what actions they are taking to achieve them. I incorporate the overall narrative information from the entire episode series for this encoding. The encoding presented in Figure 5.3 uses labels referring to key actions in the script, although it only accounts for one potential traversal, as SIG does not currently support multiple possible stories.

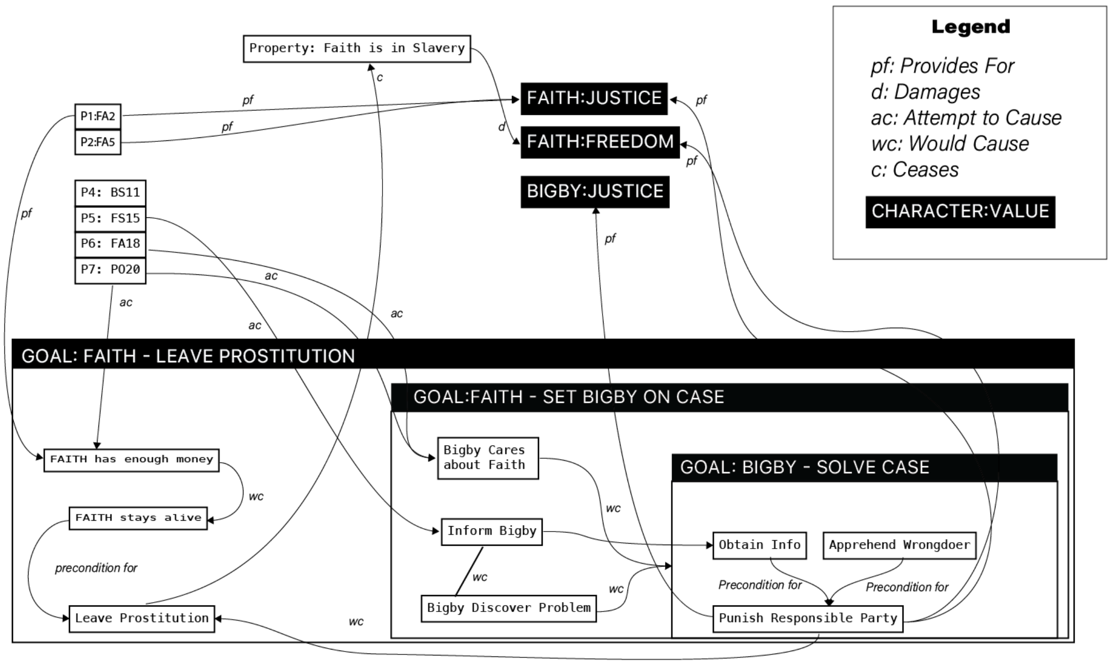


Figure 5.3: SIG Encoding for Scene

In the scene, Bigby appears to be pursuing Justice, represented by an affect node in the SIG schemata. In the initial scene, Bigby admonished Toad for breaking laws and his role as sheriff indicates that he would be motivated to achieve that value.

The scene can be encoded as involving two conflicting Goal Frames: those of Bigby and those of Faith's. Faith's goals are to escape her slavery (by pursuing the Affect node `Freedom`, FS28 seems to suggest as much, and we learn more from Nerissa later in the series). She also appears to want `Justice` through attempting to get Bigby to become interested in the case. This involves taking a risk in informing Bigby and potentially causing her magical agreement to trigger, thus introducing the need for subtlety. Bigby's goal is to achieve `Justice`, which motivates many of his actions. Elson provides a means of distinguishing individual character values, and so I separate Faith's justice from Bigby's, as they may imply different things.

Some questions are difficult to answer and are valuable for a complete SIG encoding. Was the incident at the Woodsman's apartment intentional? Was Faith's plan to pique Bigby's interest by initiating the entire situation in the first place motivating her desire to keep Bigby from pursuing the Woodsman? There are two main theories regarding the true identity of Faith and Nerissa, another woman who is caught in the ribbon agreement. The ribbon is a magical artifact that guarantees the wearer's silence through magic and a threat of death. The exact properties of that artifact are unknown during the first episode, but it is later revealed that Faith's decapitation is a direct result of the ribbon's removal, whereas her statement about her "lips being sealed," along with the Mirror's similar wording, suggests that her ability to speak on certain topics is magically constrained. The theories are premised on the capacity of the glamours and their use in the initial sequence. One involves Faith glamouring herself as Nerissa for the remainder of the storyline, whereas the other involves Nerissa taking on Faith's identity in order to lead Bigby onto the trail. The exact details and implications are beyond of the scope of this discussion, but the power that the lingering questions had on players after the completion of the game is clearly an important element of the story's resonance. There are fewer ambiguities in Aesop's fables, which Elson uses as the source of narratives in DramaBank, but they are a common a source of narrative interest in Telltale Games productions.

For this scene, we can assume, given what the player knows, that Faith was intending to flee and her main goal was to avoid interference. This leads to the affect node of `Faith:Freedom`. One of the inferences a player could draw from the ending is that the flirtation Faith employed to inspire Bigby to pursue the case in the first place was calculated and not in fact out of a desire to achieve intimacy or a closer relationship, as might be suspected from a first impression. This fits with the femme fatale trope in the neo-noir genre *TWAU* occupies, and the deception is certainly possible to model within SIG. The encoding would involve Faith planning for Bigby to adopt a plan to desire intimacy with Faith. This type of planning for a belief structure is well-supported in SIG and demonstrated in Aesop's Fables. Faith would have a separate plan that would involve Bigby accepting and successfully solving the case that pertains to her slavery.

Bigby's goals can be considered to either achieve justice or an attempt to achieve intimacy. The action of lighting Faith's cigarette can be interpreted as either a plan to get Faith to open up or an attempt to get Faith to like him. The same options that have the potential to achieve an affect goal of intimacy can also be seen as a means of obtaining information needed to do his duty. These conflict with Faith's goal of staying alive, which requires her to not reveal any information. Faith does take an action that can be interpreted as her having a goal of informing Bigby about the overall situation through highlighting her ribbon (FS15), though she is hindered by a magic spell.

While this encoding provides clear insights into the network of relationships among the story elements, clarifying the types of conflicts that occur at the goal level, the encoding does not provide indications of where players might respond. Further, the encoding focuses on the overall structure of the story as it exists in the mind of an annotator after the story has been read, and not necessarily as the story is being experienced. Finally, the role that choices play in either revealing motivations or playing a role in the traversal is unclear, as a choice text is a direct signal to the player and functions more to condition a player's attitude as well as provide an opportunity for expression. The SIG directly addresses insights into the character that are based on inferences and anticipated plans and goals, and so these elements are not as simple to adapt.

One of the primary motivations behind the development of ICE was to address the challenge of articulating the "experiencing" of the story. This includes the encoding method of identifying and coding moment-to-moment valenced impressions that are most likely to be related to moment-to-moment measurable evidence of player responses.

In the next section, I explore how Mawhorter's goal-based choice analysis can be used to encode a choice from the sequence.

5.3.2 Choice Poetics Encoding

Applying choice poetics to the choices in *The Wolf Among Us* revealed a number of interesting features regarding the nature of reasoning under duress and uncertainty. The types of decisions that are common in CCAG are not ones with clearly distinct outcomes, nor are they decisions where time is allowed to weigh the various options. Many of these decisions are in fact impulsive and based on an intuitive assessment of the values at stake. However, there are several key decisions where choice poetics provides valuable insights. One of the border cases, PC5, was chosen as it provides both a set of clearly distinct options as well as demonstrates some of the ambiguity that is common in choices in CCAG.

Mawhorter emphasizes the importance of understanding and accounting for player motivation, asserting that it is more important than player preferences as player motivations "establish the contexts in which they can be expressed" [46]. Mawhorter proposes that assessing the modes of engagement, or motivations, is the first step of analyzing the choice poetics of a work.

TWAU primarily rewards two types of engagement, identified by Yee [162] in a study of motivations in massively multiplayer online RPGs and further noted by Mawhorter: *role-play* and *avatar play*. Role-play involves the player assuming the perspective of the sheriff, Bigby Wolf, as a character. Avatar play encourages the player to express more of their personality. While both are reasonable, we focus on the role-playing mode for this, as Bigby's personality is a strong influence on the set of options. In Mawhorter's view, motivations are the desires that brought players to engage with the experience in the first place, and which underlie a decision-making policy. These might be to role-play as the character and to be faithful to that character. They might also be to win, in whatever sense winning is for the game. These motivations are distinct from preferences, which are often based on choices that do not map onto the underlying motivations. An example is a preference of one color over another in a choice–both options provide an opportunity to express the same motivation for aesthetic control, but one relies on a player's whims in the moment. In fact, many choices in *The Wolf Among Us* are based on preferences, usually in the form of selecting a value to express or identify with, and few actually diverge in terms of player motivations.

Mawhorter details a process that can be used to analyze a particular choice, which involves the following steps:

1. Goal Analysis
2. Likelihood Analysis
3. Option Analysis
4. Outcome Component Analysis
5. Relative Option Analysis
6. Full Outcome Analysis

7. Retrospective Analysis

I apply the first five steps as a method of encoding choice PC5, comparing the process and encoding to the ICE encoding process. PC5 takes place when Bigby begins to leave, and Faith stops him, saying that "we don't have to make any more of a thing out of it than it already is."

Goal Analysis

Observing the players provided some insight into their goals. Three goals (with listed priority) are plausible given the scene:

- [high] Discover the reason for the disturbance (justice)
- [medium] Improve the relationship with Faith (intimacy)
- [medium] Perform his role as sheriff (duty)

These are justified by the various options at each point and Bigby's stated objectives.

5.3.2.1 Likelihood Analysis

What are the outcomes of this particular choice? Here, we run into some issues regarding the lack of clear outcomes in the storyworld and the murky nature of conversational choices. PC6 has clearer outcomes than PC5: Bigby may need the money and not have it, but not giving it may result in negative consequences. PC5 is more of a "What do you think/How do you feel" choice, one which has many possible social outcomes given what we know about Faith, Bigby, and the situation. I chose to analyze PC5 because it is closer to a hybrid of the two types of choices in CCAG: it provides distinct options for players to consider (and does not merely provide a set of similar valenced options), and yet it contains ambiguity over the outcomes and their consequences that are also common to choices in the game.

I divide the outcomes based on the scene goals:

1. Improve relationship with Faith (also called improving intimacy in the previous section). (`improve-relationship`)

2. Find out what happened (perform duty). (`obtain-information`)

3. Achieve justice. (`pursue-justice`)

What do these mean in this particular situation? If Faith were to reveal more information, that would advance the "find out what happened" goal. However, Faith's question is posed directly to stop Bigby from pursuing what he thinks is justice by apprehending the Woodsman. This suggests that the option, questioning Faith's motives for stopping him, may achieve the goal of pursuing the Woodsman. Finally, it is possible that selecting the option "he hit you" (PO28) could improve Faith's view of Bigby. With these observations, I assigned likelihoods to each of the possible outcomes, shown in Table 5.4.

Table 5.4: Likelihood analysis for PC6.

Outcome Component	**Are you sure?**	**This is about Fabletown**	**He hit you**	**[...]**
Damage-relationship	[likely]	[unlikely]	[unlikely]	[unlikely]
Improve-relationship	[unlikely]	[neutral]	[likely]	[unlikely]
Pursue-justice	[likely]	[neutral]	[neutral]	[neutral]
Stop-justice	[unlikely]	[neutral]	[neutral]	[neutral]
Obtain-information	[unlikely]	[likely]	[likely]	[unlikely]
Not-obtain-info	[likely]	[unlikely]	[unlikely]	[unlikely]

5.3.2.2 Option analysis

The next step assigns labels to each option based on likely and unlikely components. The outcome components of `pursue-justice` and `obtain-information` can be considered to relate to the justice goal. Note that the nature of conversational outcomes may differ from material outcomes that are experienced in PC6 when giving the money represents a state change. These are shown in Table 5.5. Further, there are two possible label sets that can be assigned to each goal based on the encoding procedure: The **threatens** label is applied if any of the negative outcomes are *possible*. In this case, each outcome component is possible for each option, thereby reflecting some of the ambiguity of the situation. Similarly, if any positive outcomes are possible, then the **enables** label is applied. The **advances*/*hinders** labels are similar, except only applied if an outcome is marked as likely. Improve-relationship is hindered by "Are you sure" as the `damage-relationship` outcome component is likely, whereas Find-out-more is advanced by both "This is about Fabletown" and "He hit you", due to the fact that these are likely to elicit a response from Faith that may shed additional light.

Table 5.5: Option analysis for PC6.

Goal	**Are you Sure?**	**This is about Fabletown**	**He hit you**	[. . .]
Improve-relationship	threatens enables hinders	threatens enables	threatens enables advances	threatens enables
Find-out-more	threatens enables hinders	threatens enables advances	threatens enables advances	threatens enables
Achieve-justice	threatens enables advances	threatens enables	threatens enables	threatens enables

5.3.2.3 Relative Option Analysis

This step involves assigning a prospective impression label based on the prior steps, such as **depressing** when all choices hinder some goals or **obvious** when a single option is better than the others. It is evident that there is a variant of an apparent **dilemma** pattern between "Are you Sure?" and "He Hit You." "Are you Sure?" has a higher potential to provide more information than a more roundabout statement, at least from a prospective analysis, but it can damage the relationship (doubting Faith). Instead, choosing "He hit you" will likely improve the relationship, which in turn might improve the chance of Faith providing more information. The silent option is challenging to interpret, and I believe it was challenging for the narrative designers as well, as it is unclear how Faith will respond or interpret a lack of an answer to that line. The outcome (FS30) involves Faith reassuring Bigby about her being fine, which may or may not have been the player's intent in remaining silent (indicating caring about Faith).

5.3.2.4 Outcome Component Analysis

Next, we encode the actual results of the two most distinct options, "He hit you" and "This is about Fabletown."

The two outcomes of "He hit you" are a positive impact on `improve-relationship` and a positive impact on `obtain-information`. There is no real impact on `pursue-justice`, as the exchange primarily involves achieving a degree of intimacy and empathy, but it does result in the insight that Faith is getting out of the business (FS28, FS29), which is not in other responses.

The outcomes of "This is about Fabletown" are a negative impact on the improve-relationship goal and a major negative impact on the discover-reason-for-disturbance goal, although the player is not made aware of the information they missed. This content further foreshadows the events of the next chapter. Bigby's actual line, saying that it is not just about Faith, not only expresses his love for his community but does so in a manner that was not indicated in the choice text. The statement is met with a careful riposte by Faith that redirects Bigby's attention to the fact that he is not concerned about her and expresses her disappointment in that, thereby attempting to recover intimacy.

The two final steps of the goal-based choice analysis, the *Full Outcome Analysis* and the *Retrospective Analysis*, are not part of the encoding. They represent an analysis step. Since we are interested in comparing the encodings, we can compare the encoding details up to this point for all three models.

5.3.3 Discussion

These three approaches (ICE, SIG, and Choice Poetics) each encode different aspects of experience that are important to storygames. The network of goals and character actions that the SIG schemata encode are essential features of narrative plots, but are mostly static in choice-based games and do not provide a simple means of associating moment-to-moment responses with content properties. The network defines the overall structure and shape of a story as it relates inside a person's head after having understood it. The act of experiencing a story is more of a process, with new information causing different effects based on previous information. While the timeline provides a means of modeling the temporal relationships, the poetic potential of temporal relationships is not modeled. Similarly, it is unclear whether players will understand the ultimate goals that a character pursues in any given scene at the time that they are depicted. It encodes information regarding the relationships between characters and the ultimate structure of their plans, but many of the subtle moments that hint at a desire for intimacy within the scene are not mapped onto specific goals. This information is critical for capturing the deeper elements and motivations of story structure but does not provide a means of organizing and labeling the *discourse level* with elements that would influence the *reception* of the story.

A *goal-based choice analysis*, representing Mawhorter's choice poetics, is useful when the player has more information regarding potential outcomes and can make an informed choice. However, each encoding provides insights. The goal-based choice analysis revealed that the player was making a tradeoff choice between intimacy and justice, which provides one facet of the ICE analysis. However, the additional power of estimating likelihoods and accounting for different outcomes did not necessarily clarify the decision further, partially due to the ambiguity of the decision context and the mapping of outcomes to positive and negative expectations. The option outcomes were uncertain because the other character's motives and behavior were also unknown. The labels involving *risk and rewards*, that is, the threatens (risk) and enables (rewards), would be much more valuable in key decisions that players make towards a particular goal, and these decisions are found throughout the episode. Further, in a generation context, these values might be of more use, as ambiguity without a clear structure may backfire, whereas ambiguity in this situation is a result of player knowledge being incomplete.

The choice timer also plays an important role with respect to the amount of time players have to consider and weigh the options. Most of the choices in Telltale Games emphasize split-second decisions rather than considered ones. This is not always true, as many of the larger branches have an unlimited time to decide, such as which location to visit after Chapter 1. Such decisions may be made on the basis of impressions that may be identified through story-value charge estimations of the different options relative to the previous actions taken.

Second, the immediate relation of options to outcomes is unclear for many of the choices concerning player goals. Life or death situations are often dealt with in quicktime events, although there are cases when a deliberate decision relates to a character's fate. Player goals may be role-playing, but as previously noted, a strategy for the narrative design of Telltale Games is to ensure that there are several role-playing "rails" available for players. However, it may be a mistake to map one of these suggested role-playing rails onto a player goal based on a history of choices, as it may obscure other aspects of the content that players may be responding to, such as simply liking or disliking the character they are interacting with.

The cumulative effect of choices, which often occur in a sequence or a narrative context, are also considered by Mawhorter, although not through a detailed example. Mawhorter's example of "Papers, Please" provides an excellent example of where a core choice is repeated and its possibility space becomes a major part of the game's rhetorical effect. The choices in *TWAU* each differ in their own way according to the dramatic context. Mawhorter focused his explication on choices with a minimal context that related to character abilities and clear success criteria to isolate elemental components relating to specific postulated (or real) player goals, and so was able to surface several interesting properties of choices that may be present in CCAG decisions, albeit in more subtle ways.

Molly Maloney distinguishes between two types of choices and their relationship with one another in the eyes of narrative designers, supporting the idea of an accumulation of meaning and investment over time rather than the significance of each choice and outcome:

> Small branches support big branches. You need both. Small branches are your day to day choices. I'd say only 5 percent of choices in a game are really big ones that are going to branch the content in a super obvious way. All of those incidental choices are where you get your player investment. That's where players get excited about characters and ideas that help inform the big choices that we offer later. [40]

The integration and annotation of timing for the payoff (outcome related content events) and the cumulative effects of choices in a dramatic context are two advantages that ICE provides for analyzing content in games such as *TWAU*. Further, the story value charge provides a means of assigning values estimating the reception of related pieces of content that can, in turn, analyze individual player choices and responses.

5.4 Toward Medium Reading

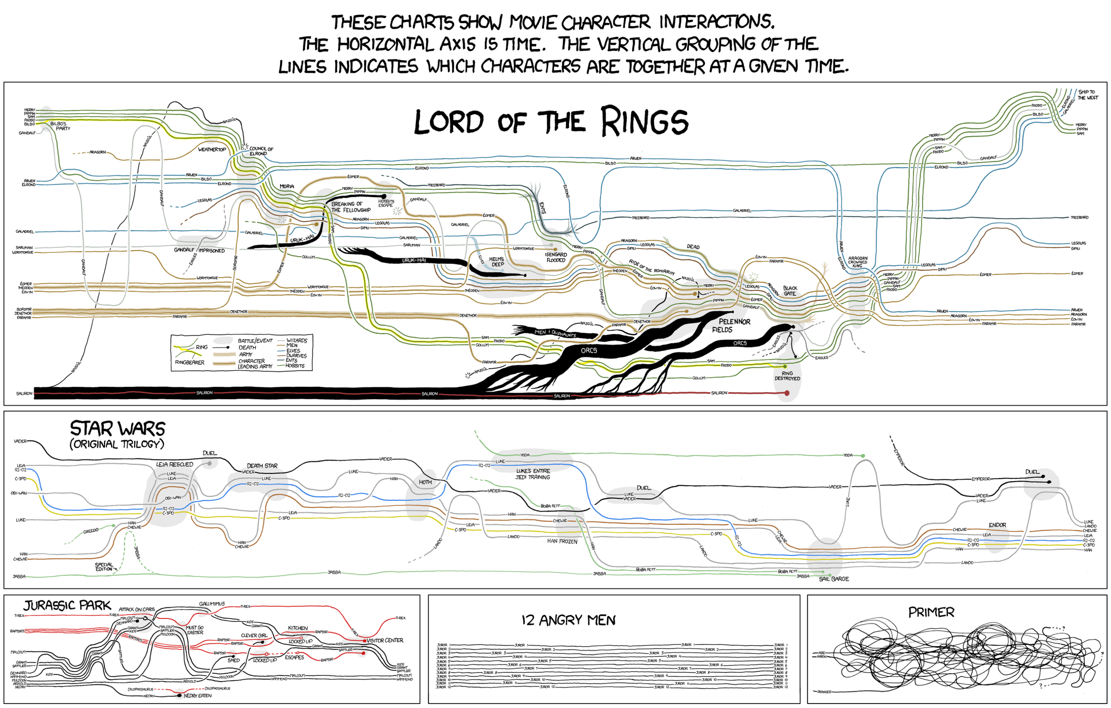


Figure 5.4: Movie Narrative Charts

In this section, I sketch an outline of *medium reading*, an analytical approach that is related to both *close playing* [163] and *distant reading*, the technique whose term was popularized by Franco Moretti [60]. Another term that could apply even better is *faceted reading*, which better describes the multiple viewpoints that are incorporated. Moretti considered the use of quantitative techniques in studying literature to be a means of surveying and understanding the underlying forces that shaped it, particularly economic and social forces. He did not see it as a superior approach to studying literature but rather as one that acknowledged the limitations in studying closely the sum of literature produced. One of the examples from his book of the same name includes a series of diagrams representing the co-occurrence of characters in Hamlet, analyzing the implications for how characters appear together. This type of relationship mapping is also employed in Story Charts, popularized in XKCD's Movie Narrative Charts,[10] shown in Figure 5.4, and subsequently automated to a certain extent [164]. Moretti's approach has inspired several works, including Elson's [47], while essentially giving a name to an approach that was already becoming popular within the digital humanities, also called humanities computing.

Close playing is to games what *close reading* is to analyzing literature or poetry: A subjective analysis based on close attention to elements of style, presentation or figures. Some suggested strategies for close playing include pair playing, or watching another person play, or using screen captures. ScholarsPlay[11] is a Twitch channel where the authors provide a running commentary while performing a close play of a video game, often explicating the mechanics and dynamics in real time. Other styles of analysis include *content analysis* [82], which involves

[10] https://xkcd.com/657/.

[11] https://www.twitch.tv/scholarsplay

coming up with a set of labels and assigning them to either individual games or segments of games. Popular topics for this style of analysis include violence and gender depictions. Closely related to what I would consider medium reading is *micro-ethnography*, an example of which is demonstrated by Taylor [107]. That analysis goes into substantial depth regarding a particular set of players, employing interviews, surveys, and other instruments to develop a complete picture of the nature of their relationship with the game. In contrast, a medium reading may include some information about the background of the subjects but focuses instead on their encounter itself.

Why, then, would there be a need for a new mode, *medium reading,* and what distinguishes it from these analytical approaches? A close play faces issues of how to compare features, how to cite particular game states (many of which are addressed by tools such as GISST [3]). For a successful method, we need to clearly define the materials for analysis and a transparent process for approaching them.

While reviewing player experience data, I realized that the approach of *simultaneously* analyzing several player's experiences was providing me insights into the different ways the same content could be interpreted. I appreciated subtleties in how players' attitudes toward Faith in options such as the moment where Faith stops Bigby (PC5) provide players who want Bigby to open up to Faith with payoffs including additional insights into Faith's motives. These discoveries of related rewards and content structures were only clear after comparing related content segments alongside the player's choices and responses, as reviewing a single playthrough and then watching another makes distinguishing contingent content difficult, as can juggling more than two or three video players to track the content common to them.

Four elements are necessary for what I propose as *medium reading*, an analysis method for approaching artifacts based on the moment-to-moment recorded responses of others:

1. A **media artifact** that is created to be interpreted and responded to.

2. A **formal model** that can be used to create an encoding of content, of which ICE, SIG, and Choice Poetics are all candidates. The model provides a means of categorizing content features that relate to an interpretation of intent and/or reception and identifies patterns.

3. Records of multiple humans **interpreting** the same content in such a manner as to be related to a content model and the original **media artifact**, possibly through a video record. These can be player traversal videos, as in this thesis, but may also be records of readers encountering traditional media in the form of movies, text, or comics.

4. A record of the moment-to-moment **response** to the content that can be connected to the part of the content model being responded to.

The above mentioned four elements provide one definition of medium reading as a method that includes both linear and non-linear media as potential subjects. Elements 1-3 can be found in the method Elson employed to collect annotations for stories using *Sheherazade* and the SIG formalism [67], although Elson used a computational approach to interpret the resulting records for patterns. Element 4 is one of the contributions described in this thesis as well as facilitated through SHERLOCK, as it captures the moment-to-moment interpretation of players. CCAGs have unique properties relating to player choice and are designed to elicit emotional responses, whereas other media artifacts may not require processing but could still benefit from a medium reading approach to understanding their reception.

Two important departures from other methods are worth noting. First, a simple gameplay video that does not include a player's responses or a set of telemetry data is not sufficient to provide enough detail to facilitate a medium reading, as they are missing the emotional and interpretative responses of the audience. Telemetry data often does not contain player response in any form, and many professional streamers realize that their performance of responding to the content is of great interest and is part of the entertainment they provide. Second, the approach, like close playing, is designed to be interpretive rather than scientific. While medium reading may reveal potential effects or processes that can be isolated and confirmed through experimentation, the underlying methodology proposed here does not include a numerical *evaluation* metric, as one does not seem to exist for the many possible meanings and interpretations possible. Instead, medium reading is a method of approaching media through the eyes and minds of others using computationally assisted observation and using those observations to better understand and articulate the symbiotic relationship between content and reception.

This approach is particularly attractive for computational media artifacts based on the availability of potential source materials. There are many player traversals of games on popular video sharing websites, and players can obtain their own using standard streaming technology without necessarily setting up a formal user study. Many of these display the face of the player interacting with the source materials, thereby providing the necessary details of a response and interpretation. Further, additional analyses can be conducted on a corpus without necessarily assembling a new corpus of traversals. The availability and ease of applicability are two reasons why both the tool that can support medium reading and the corpus of traversals are planned to be released openly.

Medium reading, when applied to textual or static image content, may need additional insights into where a reader's attention is focused using tools such as eye-tracking hardware or software. This could provide researchers a method for investigating the different effects of composition or reading patterns with respect to the content as well as visual layout or design. When applied to cinematic media, it has the potential to leverage additional advances in computational classification, such as scene detection[12], scene diarization (identifying speakers and generating transcripts), expression annotation, and others. These descriptive layers can aid in organizing and making comparisons, in the same manner that they are used in Sherlock. They can also form the basis for generative theories if the desire is to create systems that can emulate the experience of structural patterns, such as the experience of the Telltale effect, particularly with existing works. This same goal motivated Mawhorter in his theory of choice poetics: The utility of computationally modeling how choice, risk, and evaluation are related and how they can be used to both create a generative system and a theory of how they relate to creating poetic effects.

5.5 Conclusion

Encoding interactive digital narrative content using the process described in this chapter can provide the materials for insights not only into the content but also into the various potential meanings and effects that are contained within it. The results are summarized in Table 5.6. The encoding process accounted for a larger set of features specific to the subgenre than by two other formal models.

[12] PySceneDetect is one such software library, `https://github.com/Breakthrough/PySceneDetect`

Table 5.6: Summary of the three encoding method advantages

	ICE	**Choice Poetics**	**SIG**
Scope	Acts/Lines	Situation	Whole Story
Approach	Thematic	Analytical	Causal
Emphasis	Emotion	Reason	Understanding
Target	Responses	Perceptions	Analogies
Goal	Reception	Generation	Comparison

While the SIG can capture a variety of patterns of goals and structures, it does not handle smaller subsections of a story well, and its emphasis on the theory of mind relationships of storyworld entities limits its ability to account for reception effects and the ambiguities present in *TWAU*. Choice Poetics does very well in choices where there are clear outcomes that can be related to player goals, but it is less useful in short, ambiguous choices that are the bread and butter of many sequences in Cinematic Choice-Based Adventure Games.

ICE provides a model that captures choices and subtle variations in content, aligning these to player time series data. We are still far from having a full taxonomy or ontology of the potential story strategies that incorporate dramatic portrayals, but we do have many published successful storygames produced by Telltale Games and others. The method of annotating fine-grained shifts in value addresses the need to locate subtleties of how characters' actions and events are related to player choices and broader themes. These small inflections may not be detected in a goal-based choice analysis that focuses on discrete choice contexts and which have clear risks and outcomes. It also accounts for features that do not require a complete interpretation of goals and plans that may be ambiguous and are often only suggested in dramatic works. n ICE and the encoding method presented in this chapter provide a complementary set of instruments with which to assess interactive narrative traversals alongside player response data. It can be used to encode games that are visually presented and provides a means of distinguishing fine-grained differences in content between traversals, particularly when applied using supporting software such as Sherlock, which is presented in the next chapter.

Chapter 6

Sherlock

This chapter describes SHERLOCK, a multimodal, multichannel annotation and visual analytics tool that incorporates the structured representation documented in Chapter 4. Comparing a player's traversal data in real time is an ad hoc process that requires either nonlinear video editing tools or video annotation software designed for analyzing linear content. In this chapter, I describe the initial prototype and a user study to provide evidence of the validity of my approach. Combining structured annotations based on a formal model, player response visualizations, and features that support querying, while comparing player response data, represents a valuable approach for interactive digital narrative researchers.

Curation and display are essential steps in any data-focused analysis, and working with multimodal data has its own challenges. Data that fit into tables are relatively straightforward to work with, as is data that can be shown in a time series. Combining time series data with video and a dependency graph, however, requires custom visualization. One of the challenges was the number of dimensions and the size of the dataset. Another critical analytical step was aligning the content between traversals and mapping response data back to the original context, especially for artifacts and missing data. My collaborator and I used Adobe Premiere Pro, a commercially available nonlinear video editor, to create the original annotations, but the approach did not scale well to more complex schemata.

As a platform design, SHERLOCK makes several contributions to the field of games research. It supports both experts and non-experts in three phases of the analysis process: parallel annotation using a formal model, point of interest discovery, and pattern matching. The interface associates time series data with gameplay videos and provides a query interface. It integrates a novel visualization for representing the measured player expression values of valence and engagement alongside the biometric measures of heart rate (HR) and skin conductivity (SC). These measures are combined with interactive annotations representing modeled content segments that can be used to align videos for comparison. The overall user interface is depicted in Figure 6.1. After encountering challenges in annotating and analyzing the original dataset described in 3, I developed the tool alongside the Interactive Cinematic Experience (ICE) schemata described in 4. Conducting user studies is time-consuming and laborious. According to Fisher and Sanderson (1996), "The analysis time to data sequence time ratio (AT:ST) typically ranges from 5:1 to 100:1" [138]. The analysis time (AT) refers to the time taken to code data. In the dataset described in 3, the data sequence time (ST) refers to the length of the video data being coded, which totals to more than 11 hours. According to that metric, it would take between 55 and 1,100 hours to code, depending on the granularity of the coding scheme. The coding time includes not only the media itself but also the number of times it must be reviewed. Entering annotation data also takes time, including typing labels and selecting values. The time can be reduced through hotkeys and reducing the number of values in menus.

The interface screen captures presented in this chapter reflect two iterations of the tool, one that was completed before the user study and one that incorporated feedback from the user study. These are denoted as v1 and v2 and include improvements to the timeline visualizations. I added additional annotation support for the full ICE model, which is documented in Chapter 7. The black rectangles on the lower right-hand side of the gameplay videos represent the results of extracting a player's facial data into a separate channel. Masking the video allowed me to apply the facial expression detection library, the Affdex SDK, to the gameplay video (described further in Section 6.4.2).

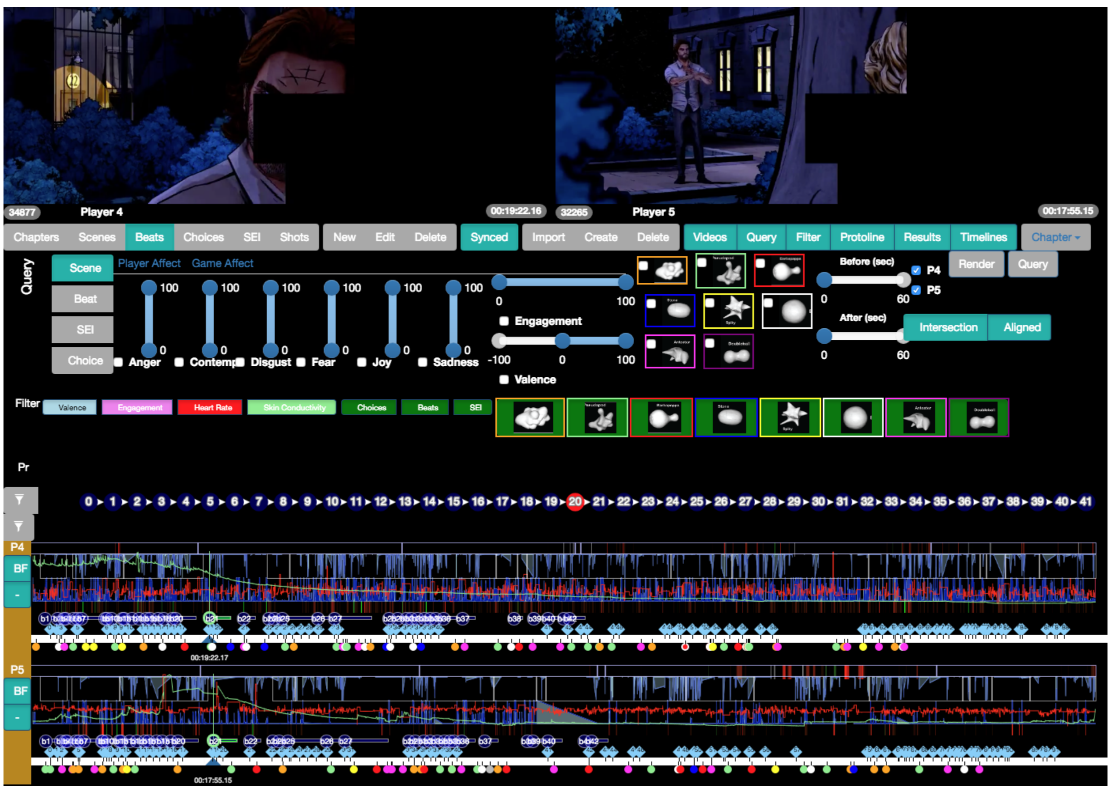


Figure 6.1: SHERLOCK user interface, v2

This chapter describes the design process of and motivations behind the system and presents its evaluation through an informal user study that involved seven self-identified game researchers and practitioners with expertise in narrative games.

While the application's goals was evident from the positive feedback recorded on the survey, several issues emerged:

1. The latency of rendering visualizations was too long.
2. The visual appearance and selection of annotations could be improved.
3. The color scheme could be confusing or difficult to read for color-blind individuals, and the colors might be unbalanced.

The responses were positive overall and indicated the need for a tool that addresses these features in the research community. One participant, game researcher James Ryan, described the system as "an exciting tool that enables fundamentally new kinds of computational analyses, both quantitative and qualitative, of video games and interactive-narrative experiences."

6.1 Positioning Sherlock

Player response and player action data contain valuable records but are complex. Sherlock is designed to analyze records of player traversals in storygames, providing a visualization of the players' raw affect data alongside annotations based on a descriptive and queriable model of the content. Interactive artifact records, also known as playtraces, are sets of correlated data that correspond to both the actions a player takes as well as the responses they might experience as they interact with a game or computational media artifact. As a result, it is distinct from systems that focus on annotation. It is a web application that can be accessed without installation and with minimal client capabilities. The server component can be installed either on the same device or separately, enabling flexibility in the size of the dataset and the availability of resources.

I chose the name "Sherlock" for the annotation and research system for several reasons. Its most direct inspiration was the annotation tool developed by Elson for the Story Intention Graph, *Scheherazade*. The name evokes a talented detective who understands how to identify, organize and relate evidence to human motivations and experiences. The tool is intended to grow in capability over time, with the capacity to incorporate different schemata in order to understanding game experience and its relationship to both content and player responses and actions.

The following sections review the existing systems and their relation to SHERLOCK.

6.1.1 Tools and Visualizations

One of my motivations for this research project was to make the conceptual and software tools of affective computing accessible to researchers in computational media who are interested in modeling and analyzing digital artifacts. The use of visualization to address complex and large datasets is a standard technique that is applied by a variety of tools.

6.1.1.1 Game and Interactive Scholarship Toolkit (GISST)

The GISST is "a custom suite of tools that manage the reference and creation of executable game data, game performance video recordings, and game executable state from emulated programs" [3]. The GISST can reference and play emulated versions of certain games within a browser. The ability to load gameplay data and perform research activities using a web application makes the tool valuable as it can be used in classrooms and on different devices to reference and discuss computational media works; this also informed my decision to use web technologies for development.

The authors identify four categories of references: objective bibliographic references, references to some performative event during specific gameplay, references to a specific game state (such as an image or copy of the system data and processes at a particular time), and ontological references. The authors describe the last category as the most difficult to accomplish due to its dependency on the other categories. Hence, the approach described in this chapter is twofold: to design a solution for a narrow genre of digital games, choice-based cinematic adventure games (CCAG)and to focus on the surface of a computational media artifact alongside player response records.

6.1.1.2 Façade Visualizations

One project evaluated the expressive capacity for dynamic content in the form of interactive drama beats in an interactive drama called Façade [165]. Sali and Mateas took advantage of treemaps to chart the relative frequency of beats, but one limitation of using telemetry data is the inability to trace the record directly back to the original experience of gameplay and the context that caused it.

Sali conducted an in-depth analysis of the behavior of Mateas and Stern's Façade using an instrumented version of the game [165] in which special functions were installed to record the internal workings. His research focused on the frequency and distribution of the occurrence of beats, and the tool itself consisted primarily of a means of recording and visualizing the highly dynamic behavior of the system. This analysis, which used dramatic elements and statistics, mirrors my own, though *The Wolf Among Us* (*TWAU*) is almost linear in its ordering of beats. Instead, I focus on the responses to beats, so the tool highlights these features.

6.1.1.3 Vixen

Another recent tool that focuses on 3D games, Vixen [5], was designed primarily for movement-based games designed in Unity. The tool uses paths, heatmaps, and custom event specifications to explore sets of user data while navigating the corresponding level geometry. This approach is particularly useful for level-based games where issues can arise as a result of physical obstacles. Movement paths of characters, especially in first-person shooters, can indicate places where players had difficulty in progressing or where a challenge was too easily overcome. In narrative games, where movement is often a secondary concern, these path-based approaches are less useful, but it is useful to have a visualization of the choice paths and the players who take these paths.

6.1.1.4 ANVIL

My collaborator and I initially began annotating the dataset described in 3 using *ANVIL* [166], a freely available video annotation research tool developed by Michael Kipp. The tool handles multimodal annotation and provides a track-based annotation interface that can handle hierarchical decomposition of various annotations as well as display visualizations, such as waveforms, to aid in locating features of interest. It can be used to annotate gestures, spatial patterns, and speech features.

ANVIL is complementary to the features of SHERLOCK in its focus, as SHERLOCK focuses on aspects particular to player traversals of interactive digital games. *ANVIL* has the capacity to handle 3D pose and motion capture data that are of particular interest for future work in analyzing performance data. The interface supports rapid annotation through hotkeys and uses a shared schema in the form of an Extensible Markup Lanugage (XML) file that describes how the data is laid out with respect to the tracks that it uses. Each track can contain multiple child types that are subsumed within an annotation, a strategy called subdividing and spanning. The annotations range from words to gestures, and include anything that can be observed in a video.

ANVIL sets a strong example for interoperability with other tools through its intermediary data formats and export and import options. The thorough documentation and usability of this tool are models for SHERLOCK.

6.1.1.5 Bio Storyboard (BioSt)

The design of Bio Storyboard (*BioSt*) inspired several of the features described in this chapter, and both tools provide a visualization of biometric data, which are mapped to player traversal data. The purpose of *BioSt*, however, is for evaluating intended design experiences and matching these with actual player experiences [4]. The tool also focuses on non-narrative games, ignoring the task of associating content with narrative structures. I was unable to compare the tools directly, as BioSt was not available online at the time of writing. Moreover, the experiences documented were shorter in duration and complexity than the player traversals recorded in *TWAU*, which motivated the approach of using a client-server architecture to store the video and dataset and to stream the data to the client using the most recent version of the Hypertext Markup language (HTML5) specification which includes video player elements that can be controlled using a Javascript interface.

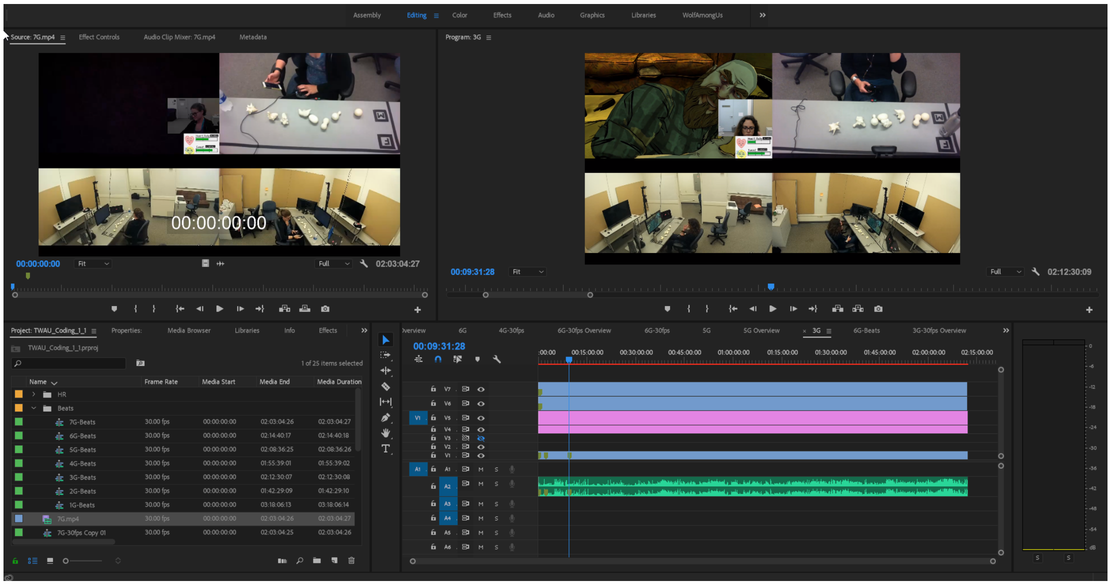


Figure 6.2: Adobe Premiere Pro interface

6.1.1.6 Adobe Premiere Pro

While not a research tool, Adobe Premiere Pro (shown in Figure 6.2) supports annotation. It is a commercial software tool designed for video production. It is limited to the creation and exporting of markers. Markers can be associated with a string and have a beginning and end location that can be recorded as either timecodes in a format standardized by the Society of Motion Picture and Television Engineers (SMTPE) or as frame counts in a comma-separated values (CSV) file.

The user interface for Adobe allowed the initial annotation to examine multiple videos, though it did not support synchronization at different time points in each video. I synced the various video channels for each traversal. Combining videos into a single video does not allow the user to see related segments of content or change the distribution of screen real estate dynamically based on the videos of interest. Both of these features are implemented in Sherlock, as each panel can be hidden or revealed as needed. The diagrammatic view of segments enables the user to see all related content in other traversals, as I will describe further in Section 6.5.4.

6.1.1.7 Microsoft TRUE

Several systems were designed to aid in the analysis of player experience data. Two of these are Microsoft Tracking Real-time User Experience (TRUE) [167], and Data Cracker [168], a tool for online user experience evaluation built alongside the game *Deadspace 2* and with the input of the development team. Microsoft TRUE combined data analytics with specific measurement protocols built in, such as surveying users during gameplay. Microsoft TRUE's subjective analysis was accomplished by pausing the game every three minutes and prompting the players to answer a survey question. The system was deployed in a large-scale case study in a released game, providing evidence that the technique could provide valuable insights into aggregate datasets. Both Microsoft TRUE and Data Cracker focus on spatial and logical components that are key to the experience of many genres, whereas the present work focuses on the story content found in storygames. The present study also looks for smaller sets of players, and the method of annotating player response data and actions can be used with already-released games without instrumenting the system itself, making it useful in research contexts outside of a particular game development context.

6.1.2 Visualization in the Digital Humanities

Another approach of interest in the digital humanities is the visualization of large sets of cultural data. Lev Manovich coined the term "cultural analytics" to describe the need for computational methods to address the quickly growing amount of cultural data being produced by mobile devices and other tools [169]. Manovich was involved in a project at the University of California, San Diego that utilized fan-made scans, or "scanlations," to visualize and organize one million manga pages [170]. The project established a precedent for the approach of addressing the media usage "in the wild," as there are many openly available streams of players playing games, including those by Telltale. SHERLOCK currently works with video files, but I have planned a future version that will analyze and annotate online streaming videos from popular services.

Another researcher whose work is significant within the digital humanities is Johanna Drucker. Drucker has argued for a careful approach to the use of visualization tools and quantitative methods within the humanities [171], seeing them as a "kind of intellectual Trojan horse" that can cause researchers to mistake data for facts. Drucker's concern is worth considering given the target audience of researchers for SHERLOCK. The goal of SHERLOCK is not to provide tools that ingest a player's traversals and produce an interpretation of what is important or what experience is taking place in them; rather, the various visualizations and instruments are tools to *facilitate* interpretation and theorize about potential relationships.

SHERLOCK deals with ambiguity and uncertainty on several levels, which are essential for an in-depth analysis of expressions, performances, and interpretations when working with narratives.Drucke (2011) notes the following: "[. . .]the task of representing ambiguity and uncertainty has to be distinguished from a second task – that of using ambiguity and uncertainty as the basis on which a representation is constructed" [171]. Encoding story value charges is a way of articulating a subjective interpretation of content that does not assign specific meaning within a goal or agent-based framework. The model includes ambiguity in the way that multiple story charge components can be attached to a given story event. This uncertainty of referent is present in both the content (what is specifically causing the interpretation) and the interpretation (what does the interpretation *mean* specifically). Both of these apparent of precision are actually necessary to account for the wide range of potential causes and the nonlinear effects of both surrounding content and previous content at this time. An elicitor may be some hidden memory or state that is not apparent on the surface, while the work itself may be unclear regarding the meaning of a particular expression.

Drucker's experience with combining visualizations with humanistic modes of inquiry is important to the present study as these aspects correspond to some of the assumptions made in the design decisions to separate and highlight the provenance (source) of data from the annotations. Drucker also distinguishes *capta* from *data* in that *capta* represents information that has been recorded and must be interpreted, whereas *data* are often treated as a "given." As we develop automated methods for analysis and apply tools to various traces outside of a lab setting, it is worth considering how to capture and retain as much of the original the uncertainty and noise as possible within the dataset on which Sherlock is intended to operate.

6.2 Design Process

6.2.1 User Profiles

Sherlock is designed for game developers, media scholars, and game researchers interested in studying records of player interactions. This section describes several user groups that guided the development of the requirements and planned features.

6.2.1.1 Game Studies Researchers

For the purposes of the design, a game design researcher is assumed to have the goal of finding insights into a game by examining the playthrough records, either by charting game mechanics or by combining a model of the gameplay with a set of player responses. This researcher would require access to the original gameplay video as well as the ability to annotate features related to the game. SHERLOCK was designed with a focus on storygames, but other genres and gameplay content could be annotated with additions or modifications to the schemata.

6.2.1.2 Game Developers

Existing games are increasingly incorporating features from competitors, with game mechanics established by one studio being adopted by others. This practice makes existing games an important resource in the game development process. The cost of producing an experimental game is high compared to assessing the reception and success of a pre-existing work that may share some features in common. Game developers may also assess new versions of a game, quickly locating points of interest for further analysis and directing their attention to larger traversal sets.

6.2.1.3 Humanities Researchers

Researchers studying traditional linear media or fan cultures are another potential user group. Fan performances represent one such resource for researchers investigating fan culture and engagement. SHERLOCK presents an opportunity to examine publicly released streams on platforms such as Twitch.tv and YouTube.

These videos, coupled with fan discourse on community websites, forums, and chat rooms, provides another layer of data for assessing a game's reception and impact.

6.2.2 Design Goals

The goals for *Sherlock* were to take the lessons from previous analyses and modeling efforts and reduce the effort needed to both annotate and analyze player traversal data. The desire to appeal to a broad game research community and to not constrain the relationships found by the tool resulted in several design goals:

1. Zero installation requirements (web distribution)
2. Parallel, synchronized videos
3. One-screen visualization of raw data
4. Connection of protostory and instantial layer content (see Chapter 4)
5. Real-time feedback for queries and filters

6.3 Architecture and Data Structures

Openly released and available research projects have played a major role in the architecture and development of Sherlock, both as components that contribute to the functionality as well as models for the design, development, and distribution.

The systems that have play a role in the architecture and development include the semantic web triple datastore, *Apache Jena* [172], and the the ontology editor, *Protege* [156], produced by the Stanford Center for Biomedical Informatics Research. The availability of the purpose-built narrative annotation system *Scheherazade* [53] initially inspired my approach of using encoding to analyze narrative media. The facial expression data extracted were produced using the openly available Affdex SDK library [110], detailed in Appendix B. Finally, *R* [173], a statistical processing software library, played a vital role in the real-time client-server visualizations for the tool as well as in handling queries.

These technologies provided a technological stack that enabled me to address the challenge of relating and displaying story-game content and player affect. They also demonstrated the value of open research systems and datasets in furthering research in the field.

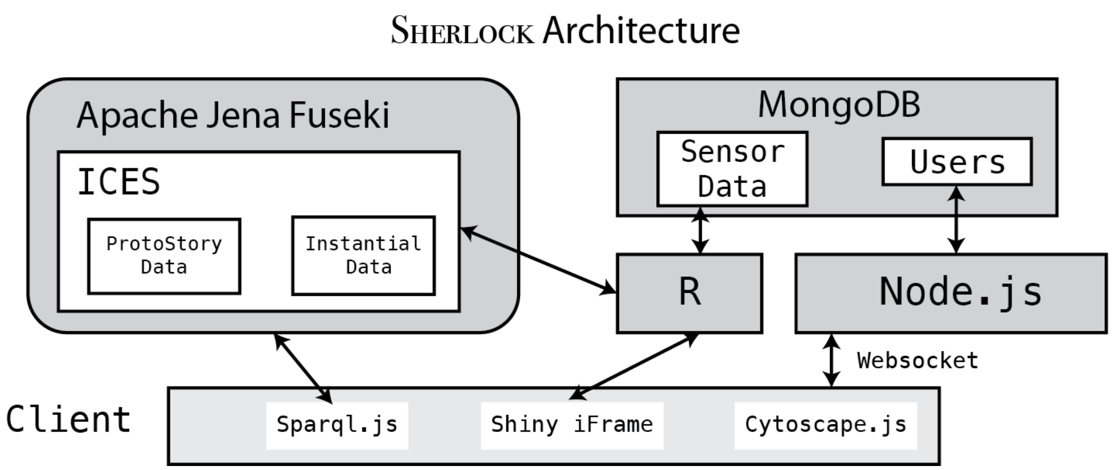


Figure 6.3: Sherlock architecture

The full architecture of Sherlock is shown in Figure 6.3. The tool incorporates the data and processes from several open source frameworks to support its goal of combining visualization, annotation, and analysis. The annotation data are stored on an Apache Fuseki Jena server [172] as RDF triples, which are described further is Chapter 4. RDF triples are essentially a ternary data structure used for describing graph data structures, consiting of a subject, object and predicate that can be any symbol but is usually constrained by a semantic interpretation such as the ontology web language (OWL). The sensor data are stored in MongoDB tables. The biometrics table stores the following: floating point values, user ID, and data type (e.g., HR or SC). The affect-related measurements are stored as objects with fields for each value extracted from each frame, as described in Section 6.4.2. The server-side rendering of the visualizations and querying is handled using R [173], a statistics platform. The server library for R is Shiny [174], used for interactive sessions with an R instance. that maintains a connection between a web page and the server. For patterns, it uses the SPARQL, an extension of SQL for making queries on triplestores, client for R to query the Apache Fuseki Jena server [175]. Cytoscape.js handles the graph visualization of the protostory layer [176]. For a given video, the time series measurements far outnumber the number of pixels, so the server renders and processes a dataset before sending it to the client [173]. Annotations and the interface are implemented with a mixture of Bootstrap [177], JQuery[1], and TypeScript.

6.4 Visualizing Player Response

The sheer scale of the multimodal data in this study made it clear that traditional graphs would not be sufficient. The term visual analytics (VA) describes the use of interactive information visualization tools to assist in analyzing complex and dynamic datasets. In 2006, Thomas and Cook summarized the visual analytics genre in a report from a panel that was assembled to address challenges facing the United States after the September 11 attacks and Hurricane Katrina:

> The panel defined [VA] as the science of analytical reasoning facilitated by interactive visual interfaces. People use [VA] tools and techniques

[1] `https://jquery.com/` Version 3.2.1

> to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data; detect the expected and discover the unexpected; provide timely, defensible, and understandable assessments; and communicate assessment effectively for action. [178]

The report includes a section on the use of data transformation, which is a process that changes data into a new (potentially more useful) form. One example is rendering a graph, and another is applying a facial expression library to obtain values for various emotion labels that are associated with combinations of facial action units.

The VA tool addresses the complexity of identifying relationships between the various values over time associated with player response. The use of a visual search avoids making a commitment about searching for a particular relationships. This emphasis on data enables researchers to formulate and address a variety of research questions, which makes the tool useful for other genres of games. This commitment to neutrality with respect to patterns of data that can be mapped onto either the content or the response to it led to three design decisions:

1. Associate data consistently with a uniform time axis across multiple traversals. (This simplified the comparisons but required that segments start at locations that could be aligned on demand.)

2. Detailed data views should be rendered on demand based on the region of interest.

3. As many values as possible should be visible at once so that they can be filtered selectively.

These decisions represent best practices for information visualization, known as the visual-information-seeking mantra, as introduced by Ben Schneiderman:

overview first, followed by zoom and filter, and then details on demand [179].

As described in Section 3.1.1, several values are available to measure player response over time. These include HR, SC, and aggregate values obtained by analysis of facial expression, engagement, and valence. Other values exist, which include specific emotions detected by a facial expression processing library, but these have too many dimensions for an overview (though peaks may be included in a future version of the research). They are available for selection based on queries and filters.

6.4.1 Visualizing Biometrics

Time series have the potential to contain both local and global features. The raw sensor data provided a useful means of comparing players to one another. These data also showed general trends in a player's physiological state. The processed time series, described in Appendix B, required experimentation to select the right set of features and exhibited a great deal of noise and artifacts due to breathing, talking, and moving. The visualization incorporated colored line graphs for the two biometric measures, HR and SC. Figure 6.4 shows Player 2's (P2's) SC and HR layered on top of one another in an initial version of the color scheme. The color representing skin conductivity was later changed from blue to green so that it would stand out better from the light blue engagement data in the visualization.

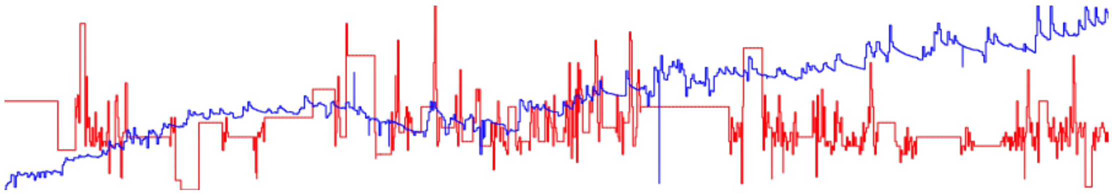


Figure 6.4: Player 2, Chapter 1: Heart Rate & Skin Conductivity

6.4.2 Visualizing Facial Measures of Affect

In developing a visualization for facial affect, I settled on the two main dimensions provided by the Affdex SDK: valence and engagement. Valence is the measure of whether a perceived emotion is positive or negative. Valence can be understood as either positive or negative, so a two-color scheme that highlighted extreme values seemed appropriate. The exact details of the data points were less important than the patterns, so I settled on a linear graph because color variation is complementary to the high-frequency visual details of line graphs.

An example of this visualization is P2's affect during the first chapter, where valence is combined with engagement in Figure 6.5. While it is possible to map a value onto a hue or color directly, a filled uniform color forms positive shapes from the outline ; in this context, the positive/negative space worked well next to the color-only valence measure, as it relied on distinct visual features. HR and SC often showed patterns of global movement and variation that may have been lost in a filled-in line chart, which would disguise small variations and take up space for unchanging periods and may have further interfered with the facial engagement visualization.

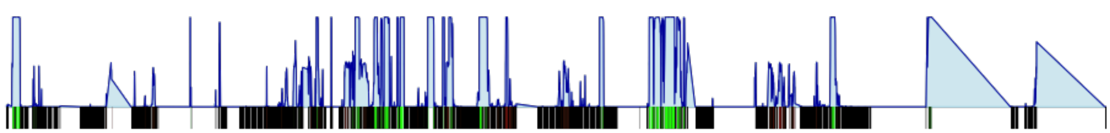


Figure 6.5: Player 2, Chapter 1: engagement vs. valence

The negative valence segments are indicated by faint vertical lines in red along the bottom, while positive segments are in green. This color scheme has several issues, though, as green is far more visible (as a higher value) than red, and the middle point, represented by a neutral valence, is black, which is also the color of the background. This obscures locations where there are no data and points where small magnitude variations occur. The black background in the interface also hides areas where there are gaps in the data, which should be visible but not distracting. The combination of features enables a user to survey the timeline for points where the two measurements interact (valence, engagement) juxtaposed with content annotation time boundaries.

In v2, each visualization occupies vertical space in the total timeline (the two graphs occupy the center of the topmost layer, while the valence is positioned along the top and bottom). Separating the annotation tracks did address the overlapping information, but the missing data is still an issue. One solution is to use a neutral gray to represent the middle value and reserve black for missing values. This strategy could be expanded to a light gray background for the engagement segments with missing values; however, the v2 line graph connects and suggests values where there are none, resulting in artifacts. Regardless, the color scheme should be adjustable for different user populations that may have visual impairments.

6.4.3 Comparing Game and Player Data

While the tool's focus is on player data, I realized that another source of emotional data was already present in the dataset: the facial expressions of the characters themselves. With this insight, I set out to combine both the content and the player visualizations in a way that would keep the two distinct to enable visual comparisons along the same time segments, with each vertical line representing a slice of time.

Using the method for charting valence and engagement described in the previous section, I created a mirror image (top down for engagement with valence along the top edge) and populated it with data extracted from running the same library in the game itself. Many scenes or characters in the game are dynamic or have unusual poses, which provides a set of potentially valuable features that indicate emotional signals present in the game content. These signals can be aligned with common points between multiple traversals, although the same segments of content do not always yield the same values. This strategy worked surprisingly well, especially for the human fables, such as Snow's negatively valenced facial expressions in the scene leading up to the discovery of Faith's decapitated head. This additional dataset of content-derived features that mirror a player's expression may be a source for additional analytical insights, such as automating the detection of emotional events performed in the content, but it is not used in the current thesis beyond its inclusion in the visualization.

6.4.4 Combining Visualizations

One of the advantages of interactive visualization is the ability to overlay and filter information. In *Sherlock*, users are encouraged to make these comparisons between the different measures that are juxtaposed, such as content annotations and heart rate changes, by visually filtering information.

As discussed in the previous section, the first version of the timeline visualization did not constrain each graph to a specific vertical extent, so the line graph overlapped with the valence lines. This version of the timeline visualization is shown in Figure 6.6, where boxes represent beats or scenes.

In the second version, shown in Figure 6.7, the annotations' vertical size was reduced, and a background shape was used for the ID labels. Each time series value

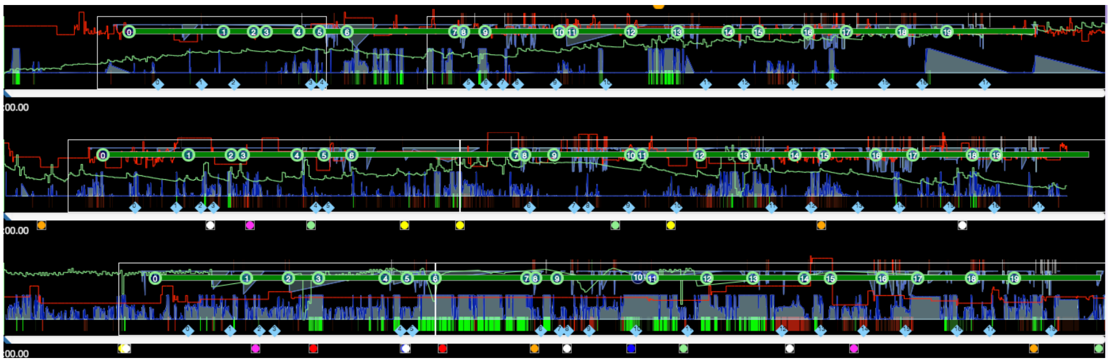


Figure 6.6: Player 4, Episode 1, v1: all content types are present in the overview mode of the timeline for the first episode

and segment annotation was constrained to occupy a specific vertical segment or horizontal "track." Separating features into tracks enabled the annotations and information to remain close enough for visual comparison without overlapping or excessive interference. Users could still filter or hide annotations using the filter panel. These include Sensual Evaluation Instrument (SEI) annotations, choice locations, beat locations, and the multivariate time series visualization of six continuous variables (valence + engagement for the game and players; HR + SC).

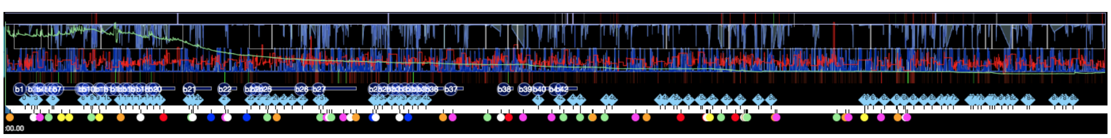


Figure 6.7: Player 4, Episode 1, v2: annotations have been moved to a distinct "track," while scenes are rectangles that are highlighted when the mouse hovers over them

This visualization has several points of interest. The raw GSR visualization enables response events and locating features in addition to relative peak sizes.

Although the interface may seem dense when compared to other types of high-information-content datasets, such as visualizing Waveform Audio Format (WAV) files or stock data (Figure 6.9 and Figure 6.8 [2]), the complexity is comparable. The ability to scan for complex and potentially nonlinear relationships is useful when dealing with longer time series in which variables may be highly interdependent.

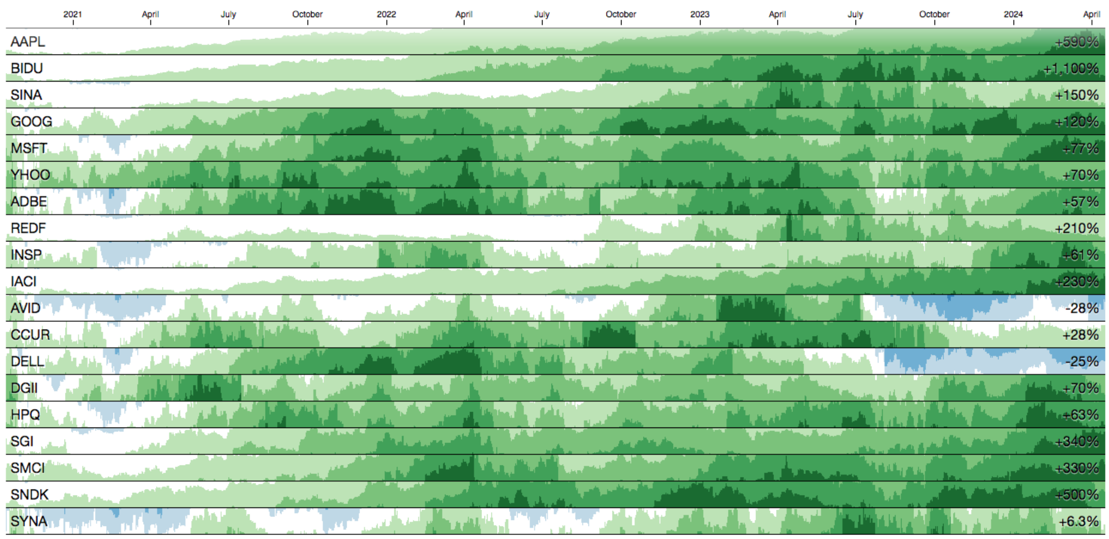


Figure 6.8: Stock data used in the cubism demo

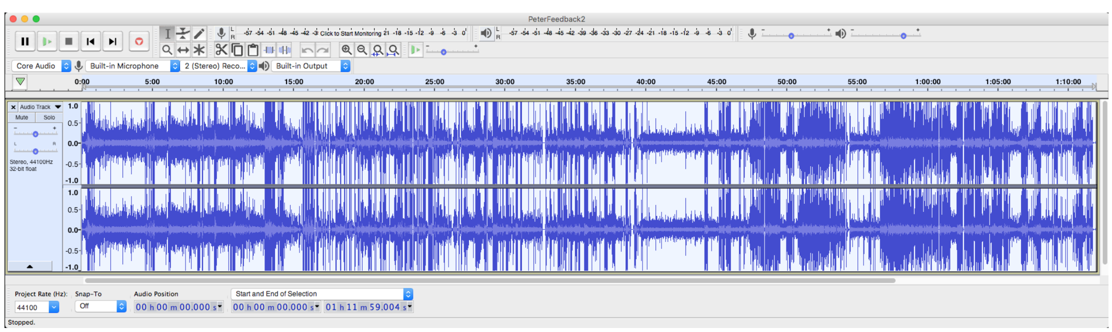


Figure 6.9: Audio in Audacity for 1 hour and 10 minutes

6.5 User Interface

The interface implements several principles that resulted from studying the dataset and conducting the analysis, as described in Chapter 3:

[2] `https://bost.ocks.org/mike/cubism/intro/#0`, Mark Bostock, Square, Inc. Cubism.js talk SF Metrics Meetup. May 2, 2012

1. Relevant expression values are combined for each x position.

2. One-screen visualization is used.

3. Navigation on a timeline assumes a desire to compare similar content.

4. Users can filter information based on the segment type and matching criteria.

5. Data should be aligned and scaled uniformly (a single pixel is equivalent to the same amount of time).

6. All connected elements should react to user interaction, indicating their corresponding elements. Hovering over one element highlights related annotations.

6.5.1 Videos

The player and game video elements use standard HTML5 video players with the addition of frame, time, and ID labels. All videos are arranged in a grid at the top of the screen. Individual video channels can be toggled on or off on each player's timeline.

Displaying multiple videos at once in order to compare their contents is one of the requirements. Videos are arranged as a grid along the top of the frame alongside their current timecode (hour:minute:second format) and frame count. I found, however, that only about six videos could be played effectively at any given time in Chrome at the time of writing. This limitation can potentially be addressed through server-side combinations of video channels.

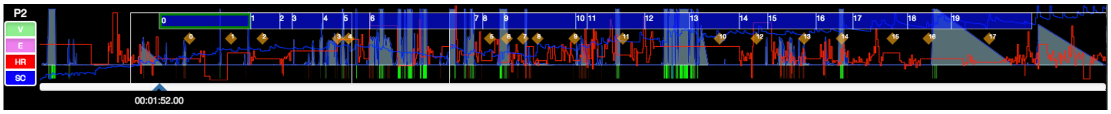


Figure 6.10: Player 2 timeline: overview (v1)

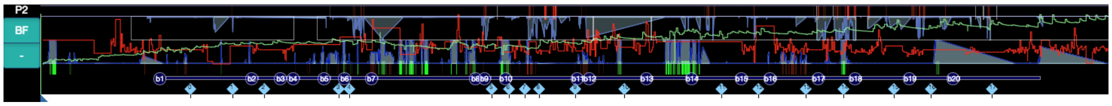


Figure 6.11: Player 2 timeline: overview (v2)

6.5.2 Timelines

Time is an important dimension of player experience. Events occur in response to other events, emotions evolve, and most of the action takes place in time, as opposed to space, in dramatic works. A horizontal slider represents the current frame. The slider track is aligned to the annotations on the x coordinates. By clicking on an annotation, the cursor moves to the start of the related content in each traversal. Currently, the timelines are sized to fit six traversals at most on a typical screen, but in the minimized mode, as many as 12 can be fit.

The timeline has two x-axis “scopes”: overview, which is comprised of the outer scope segment type and a specified ID, and detail, which is a zoomed-in version. The outer scope refers to the level of the segment that is used to contain the rest at the level of the browser window. Ordinarily it is a Chapter, but it can also be a Scene, Beat or Episode. Users can activate the detailed view by holding the Ctrl (or command) key while clicking on a segment of interest; this causes the overview timeline to switch to a minimized state. The timeline can also be minimized, which reduces the height by half and changes several of the

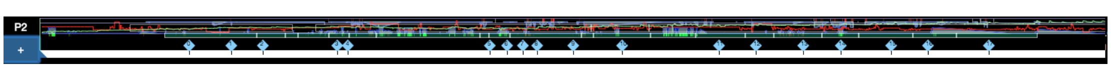


Figure 6.12: Player 2 timeline: minimized overview (v2)

annotations to more compact versions (Figure 6.12). The timeline allows users to select segments or annotations, and currently scene 0 is selected in Figure 6.12).

The buttons on the left in this version reflect a set of value filters that can be toggled on or off to display different biometric data. These toggles are consolidated into a single panel in version 2.

6.5.3 Synchronizing Videos

As we have seen, comparing data across playthroughs requires the corresponding content to be aligned. This comparison is performed by playing videos based on the active content segment. When a content segment is selected, that segment's protostory element is used to locate and set each player's timeline to the start of their corresponding segment. For instance, if the first beat in P2's timeline is selected, then the player positions in all timelines are set to their respective locations for the first beat. This is particularly useful for comparing the facial reactions of players to similar content and is important in connecting the affective data to the source materials. Synchronization of videos occurs when a timeline is activated based on the most specific content in the "synced" mode. The frame for each timeline is calculated by finding the most specific containing segments of the selected timeline and using the offset to position the other current frames relative to their respective segments. The offset is calculated by taking the difference of the start locations of the corresponding segments on the other timelines whenever a new timeline is selected.

In addition to the slider and clicking on elements, users can navigate elements using hotkeys. The "n" and "p" keys select the next and previous elements of the current type, respectively. The forward and backward arrow keys increment the frame index, and the control or command key can be used to jump one second forward or backward (30 frames). These shortcuts are useful as they reduce the time needed to navigate, which is especially important when comparing frames or deciding on subjective evaluations of content variations. Users can change the selected player by simultaneously hitting the Ctrl or command key and the "n" or "p".

6.5.4 Protoline

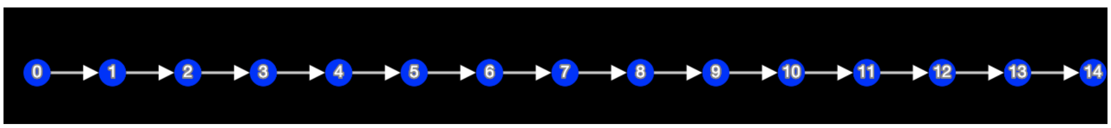


Figure 6.13: Protoline (scenes)

Sherlock primarily supports navigating, organizing, and querying multicursal traversal records alongside content. Since content segments vary in length and location, I introduced a panel that renders each element as a node in a graph using Cytoscape.js. For segments, the display lays out each node in a breadth-first traversal (left to right) using the links that designate the preceding and following content.

There are two modes for displaying choices: one that displays an option text as an edge and another that shows each option text (also known as a `ProtoOption` in ICE) as a node. Each option node is connected to the choice that occurs next, creating a graph that shows each possible path a user could take through the set of possible options. Clicking on a choice node, a `ProtoChoice` in ICE, selects each traversal that has an instance of that `ProtoChoice`, shown in Figure 6.14.

The protoline panel, or a diagram that shows the protostory relationships between available elements, can be used to highlight selected content segments

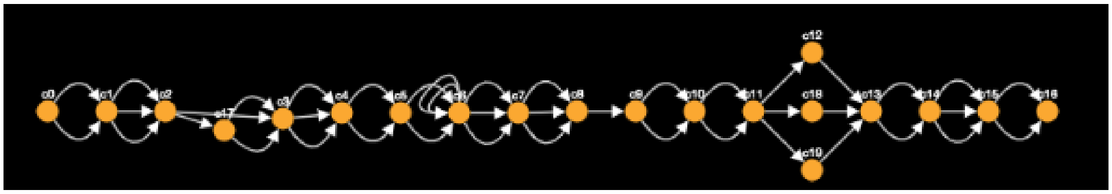


Figure 6.14: Protoline (choices). Each edge represents a path taken by players, each node a distinct ProtoChoice.

from a query (each of these appear in a different color) and to manually select content segments for comparison using the render command (i.e., adding to a selection set by holding the Shift key). The protoline panel renders a graph diagram that provides a uniform, clickable annotation for each protostory element, making short segments easier to select.

6.5.5 Querying and Filtering

Figure 6.15: Sherlock filter panel

As aforementioned, one of the primary features of Sherlock is the combination of visualization, annotation, and querying in a single interface designed for exploring multichannel player data. However, the information presented on the screen may be overwhelming. The *filter panel*, shown in Figure 6.15 includes a set of toggles that turn the respective features' visibility on and off. These include individual SEI instruments, choice text, and the various graphs of time series data.

The query interface tests subsets of the data for the presence of a value in a specified range. Subsets are either segments, specifically scenes and beats, or points, SEI usage, and choices, with time windows before and after. The query interface, shown in Figure 6.16, displays a query of Chapter 1 matching beats that contain engagement and valence values between 80 and 100. Like the ICE encoding, valence values provided by Affdex SDK from a set of input frames vary between -100 and 100. Originally, each filter selection had a single handle, but it quickly became apparent that the relationships among the values were not clear and that a second button was needed to express “greater than” or “less than” relations. Therefore, a range selection mode was added for the sliders, which enabled any contiguous set of values to be selected. Although the current implementation (v2) does not have a method of excluding segments that contain missing values, the user may want to take this into account or return an annotation indicating this for some patterns, as was the case with the visualizations. Note in Figure 6.16 that Player 7 (P7) did not match any segments, while Player 6 (P6) matched multiple beats.

The design of the query interface emphasized real-time feedback and making the underlying dataset available. While both SPARQL and SQL are difficult for novices, their accessibility can be improved through the interface design. By presenting the results in the context of a navigable set of annotations that link them directly to the original context, an analyst can focus on the underlying relationships of content to both time, player response and to other aspects modeled.

Players can be filtered for inclusion. The **Scope** is the segment type matched by the query. The idea is that at each stage, the set of candidate elements is filtered. Each **Value** is represented by a range slider. Indexes are used to select a segment if a value falls within the maximum and minimum values. Both **Choices and SEI** have a single frame location. Consequently, sliders for before and after (in seconds) define the time windows for selecting values relative to the points.

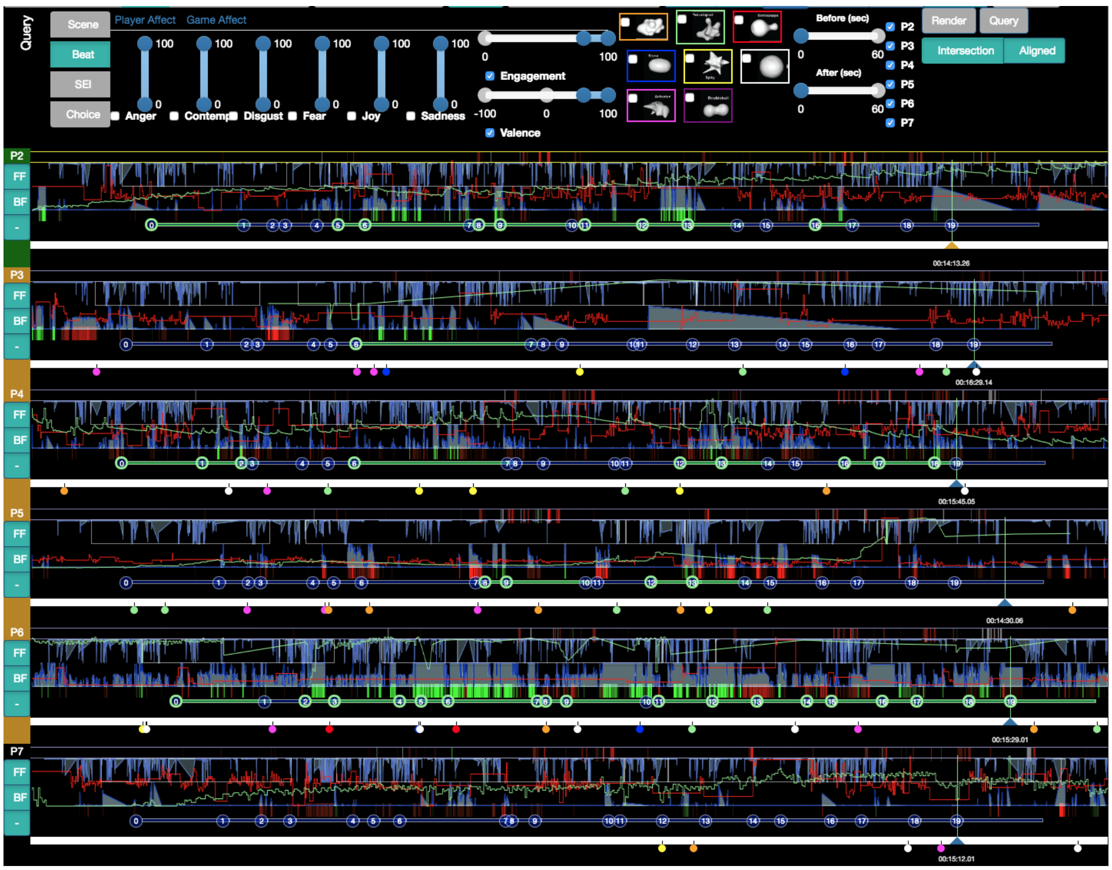


Figure 6.16: Query panel

6.5.6 Rendering Results

Since the videos depict players encountering the same content, Sherlock supports aligning subsections related to content segments in a way that previous tools could not. The “render” button uses the currently selected segments and aligns them based on common protostory elements. This allows the user to compare the timings of annotations and player affects.

The sets of beats shown in Figure 6.17 and Figure 6.18 are those analyzed in the last chapter (Chapter 5). Version A (Figure 6.17) includes option text and selections for each choice, whereas Version B (Figure 6.18) has choices hidden by a filter to show the data.

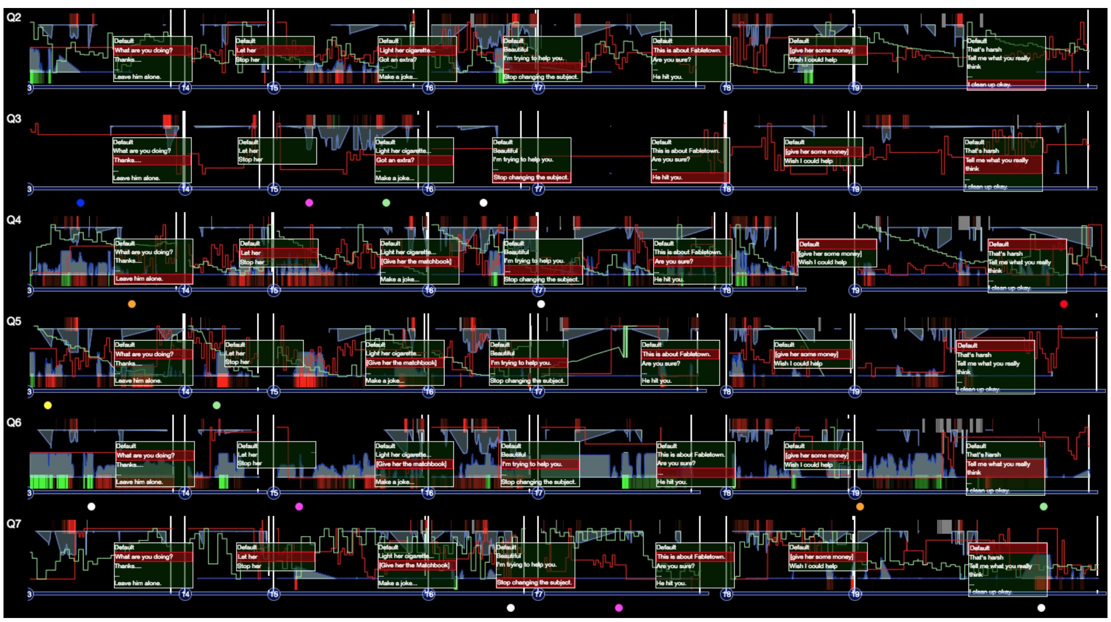


Figure 6.17: Rendering results of beats 14–19 (A): choices shown

Each annotation is rendered by a specific instance of Shiny (running in its own memory space) on the server with the same annotations and visualizations as those of the main timelines. I was able to shorten the latency of rendering results considerably using a subsection of the existing image as a proxy, and a higher

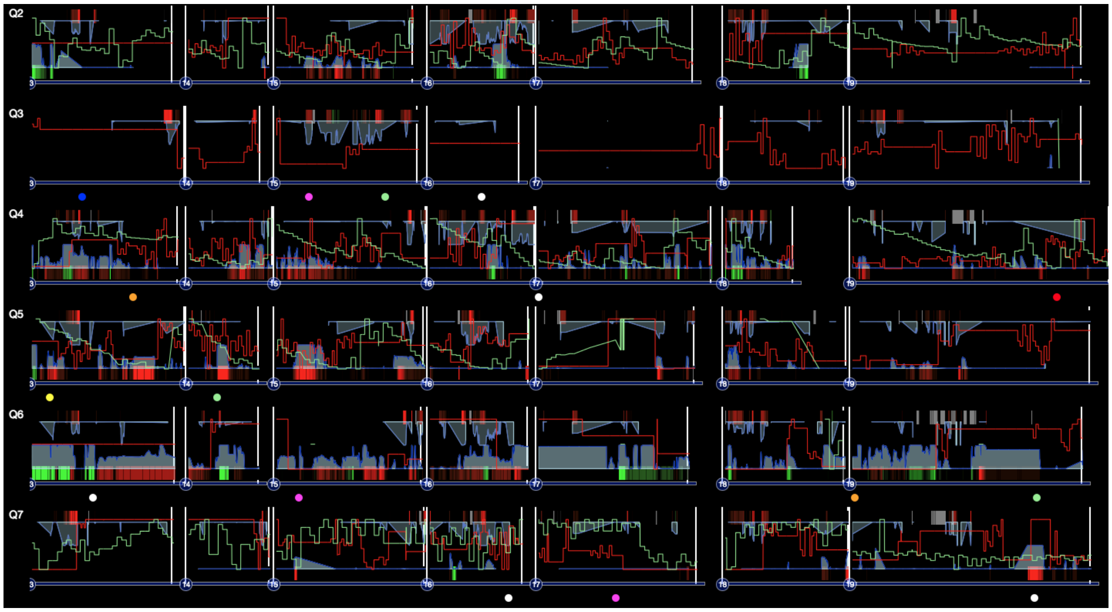


Figure 6.18: Rendering results of beats 14–19 (B): choices hidden

resolution version was rendered on the server.

6.6 Sherlock User Study

To evaluate the tool, I conducted an informal user study that asked game researchers to use and assess the tool, its purpose, and the interface. The primary goals of the study were to validate the hypothesis that there was a desire for this type of VA tool in the community and that the combination of features was of interest for research and teaching. I also wanted to gather feedback before the next iteration of the features.

6.6.1 Study Setup

I guided the participants through the features of the interface and asked them to complete a survey consisting of a set of Likert-style items that captured the

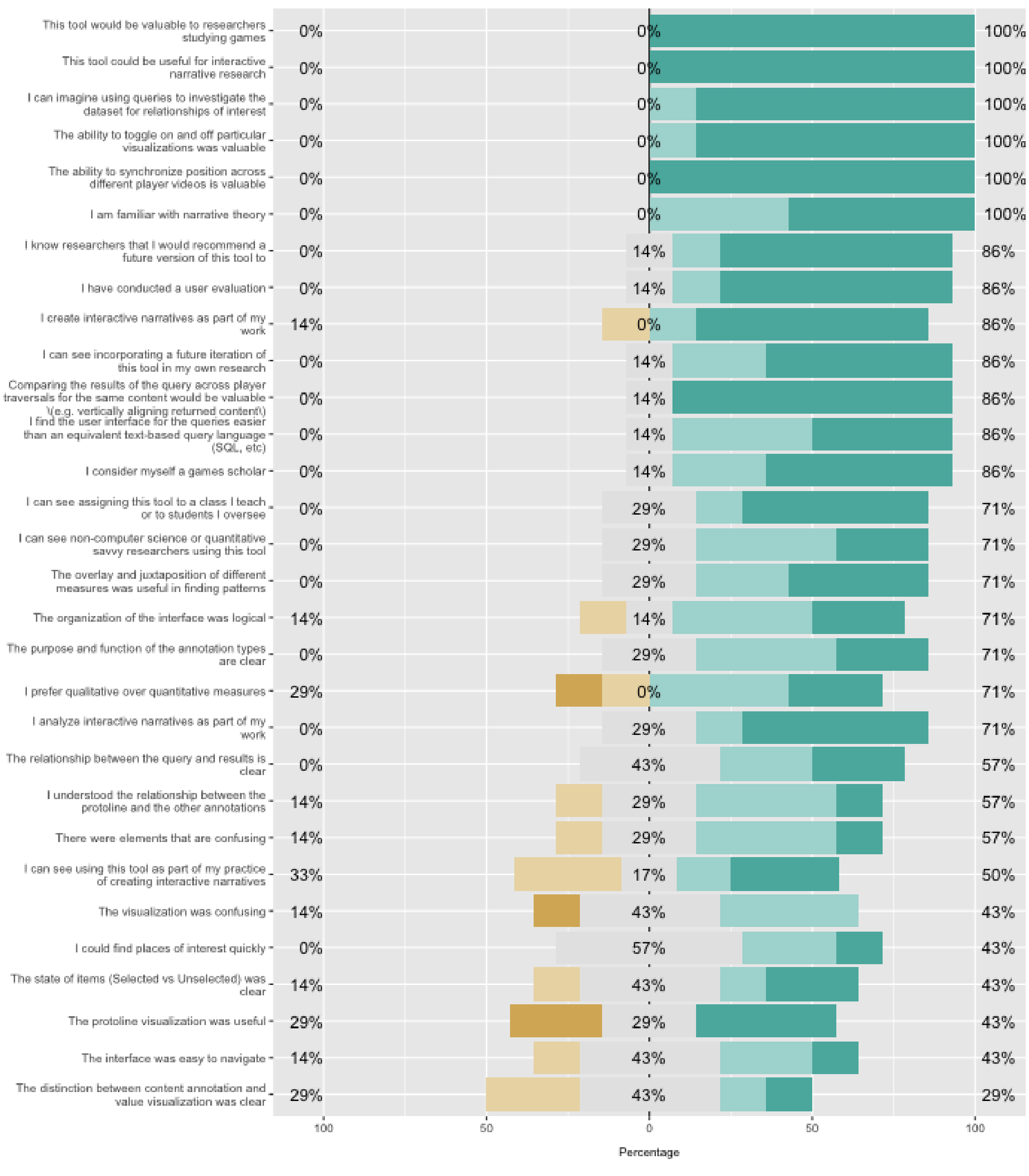


Figure 6.19: Sorted survey results: 1 = strongly disagree, 5 = strongly agree

opinions of the researchers on the visualizations and features. I was not measuring any underlying variables, so the questions were not repeated with reversed valences or different wording. During the walkthrough, I encouraged participants to experiment with various features and controls. Two participants were unable to load the application, so I asked them to assess it as demonstrated in the video chat screen share. The participants completed an anonymous survey that consisted of 30 questions in 5 ordinal categories with options on a scale between "strongly disagree" and "strongly agree." The participants provided feedback in a freeform text box at the end of the survey, from which quotations were drawn. The survey included screenshots of the core elements to help the participants recall the interface and visualizations, which are identical to those in the v1 version presented in the previous sections of this chapter.

The full set of questions and the responses can be found in Figure 6.19, though I will present the relevant subset for each category here. The survey asked several self-identification questions about the participants' expertise and field (Table 6.1). I designed these questions to assess the degree to which the participants identified as one of the core sets of users described in Section 6.2.1.

Table 6.1: Participant self-identification

I consider myself a games scholar.
I am familiar with narrative theory.
I create interactive narratives as part of my work.
I analyze interactive narratives as part of my work.
I have conducted a user evaluation.
I prefer qualitative over quantitative measures.

6.6.2 General Feedback

The surveyed researchers were enthusiastic about the key features, and several indicated they would use or recommend a future iteration of the tool. One of the respondents was troubled at the level of detail available from the tool, musing on what it represented in terms of the potential objectification of an art form. Another researcher mentioned that this sort of analysis could raise ethical questions about the types of knowledge it could reveal. The topic is important, and I will revisit it in Chapter 8, although a full discussion of the implications is outside of the scope of this thesis.

6.6.3 Visualizations

One of SHERLOCK's key contributions is the visualization of player affect data alongside the content model and original video of player traversals. The questions provided to participants are listed in Table 6.2.

Table 6.2: Visualization questions

The state of items (selected vs. unselected) was clear.
The ability to toggle particular visualizations on and off was valuable.
The overlay and juxtaposition of different measures were useful in finding patterns.
I could find places of interest quickly.
The visualization was confusing.
The distinction between content annotations and value visualization was clear.

This section of the survey showed some disagreement among the participants. The variations may have something to do with the comfort level of participants with objective measures of player response and could be addressed through appropriate through a tutorial of the interface built into the software Another possibility is that additional signifiers related to the content are needed to orient users. The current version of the interface moves annotations into separate tracks and positions point annotations next to the timeline (Figure 6.20).

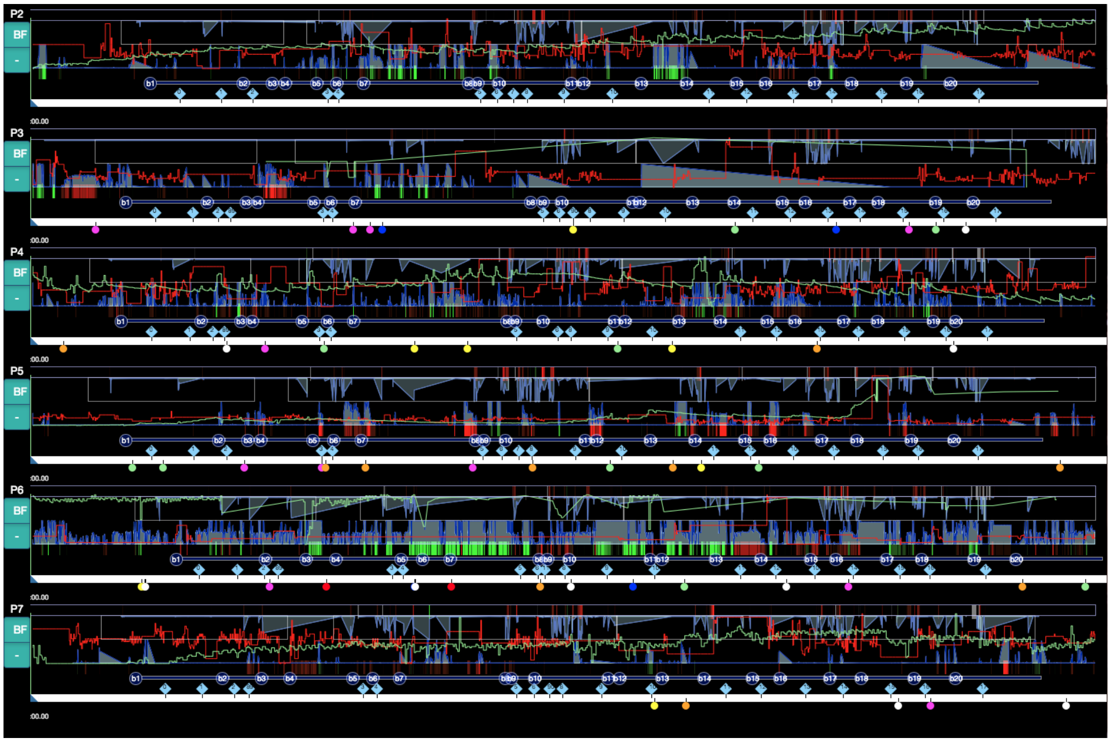


Figure 6.20: Current timelines

6.6.4 Core Interface

Table 6.3: Interface questions

The purpose and function of the annotation types were clear.
The interface was easy to navigate.
There were elements that were confusing.
The organization of the interface was logical.
The ability to synchronize positions across different player videos was valuable.

The survey results provide some evidence in support of the hypothesis that game researchers would find the alignment of content and synchronization of traversal videos useful. The questions can be found in Table 6.3. One issue was the nature of the sliders and the values they indicated. The labels and functions of various interface elements needed to be explained. It is likely that a tutorial will need to be developed to accompany the interface, as suggested in one of the comments. However, the amount of screen real estate for labels is limited.

6.6.5 Bugs and Issues

Several bugs occurred in the tool during a few sessions, especially involving the demonstration of the query feature. These did not interfere, however, with the study's general purpose of assessing the concepts and interface features. Another issue that was brought up was the complexity of the interface. Tutorials that introduce each of the features and visualizations can partially address this, as suggested by a participant.

The protoline assessments (described in Section 6.5.4) were divided, as shown in Table 6.4. This feature captured how content related, and although the use of the protoline to select content across multiple traversals was clear according to the participants during the demonstration, its use and representation could be clarified via additional user trials and labels. As one participant pointed out, the variation in protoline assessments may have been due to the reversed order of strongly agree/strongly disagree on the survey in comparison with the Likert scale. Protoline is the least common of the features shown, so users could benefit from additional guidance in its use and in understanding the relationship between the protostory and the timelines.

Table 6.4: Protoline questions: 1 = strongly agree, 5 = strongly disagree

Question	**1**	**2**	**3**	**4**	**5**
The protoline visualization was useful.	2		2		3
I understood the relationship between the protoline and other annotations.	0	1	2	3	1

6.6.6 Selected Quotations

Below is a selection of quotes provided by two participants in the general feedback section that captures many of the sentiments expressed by the participants during the sessions:

> Its use in game making also had its troubling side to me as an artist. It made me wonder if we do really want to see the responses of our readers at this level.

> It's a quite useful tool, with the video synchronization and choice structure detection being perhaps the most useful functions relative to performing an equivalent analysis manually with, for example, raw YouTube videos.
>
> I think that the interface still needs some work, and the issues with the queries during my trial meant that in its current state, the tool would be pretty frustrating to use, but these issues are easily addressed.
>
> I think that the other big issue is tutorialization: the interface is quite complicated and was only comprehensible because I had the designer in live chat explaining things to me.

6.6.7 Discussion

Users quickly moved from talking about the interface itself to beginning to assess the underlying dataset it represented. This was encouraging, as it indicated the appeal of using raw data visualizations as a design element. Instead of having to formulate a single question, the information contains several potential patterns that can provoke new questions. This means the tool can be used for a variety of purposes, not just for the subgenre addressed in this thesis. The use of the Affdex

SDK means that any video portraying facial expressions can be visualized using the engagement/valence visualization. Many streamers make a practice of including their faces in gameplay streams, in part due to the value found in watching their expression while playing. This suggests that such videos are valuable for game design researchers analyzing player response in a variety of different game genres.

6.7 Feedback and Iterations

The version of Sherlock described here served as a valuable tool to open up a dialogue with game researchers about what sorts of insights and capabilities might be useful in utilizing a visual data mining tool to study player response. I implemented a few enhancements based on the feedback obtained from the user study. First, I solved the latency issues by pre-rendering the timeline visualizations in a standard resolution and replacing them when the more accurate image data is available from the server. This represents a tradeoff between the server load and latency, and it was clear that the data needed to become available quickly. In a similar vein, I further optimized the usage of the server R package, Shiny, so that two instances could operate simultaneously, as I used the community version that operates in a single thread. This doubled the memory requirements, however. I also improved the visualization of the timeline to avoid having annotations obscure the engagement and biometric data.

6.8 Summary

Sherlock represents a new type of hybrid visualization/annotation interface that combines annotations of video traversal data with a visual data mining approach to player response data. In its current form, it enables researchers to explore a set of related traversals and add annotations and query datasets using a user interface. The participants in the user study indicated that the tool and its related features would be of use to the target user community of game and narrative game researchers, especially when coupled with additional tools for ingesting data.

The latest version of Sherlock, along with the source code, will be made available at `http://sherlock.jtm.io`, with the goal of making computational methods of analysis available to both game researchers and game designers.

In the next chapter, I will describe a set of analyses are completed using a portion of the study dataset annotated with Sherlock. Precise timing information across multiple traversals enables a more fine-grained analysis of individual content events and their resulting responses.

Chapter 7

Analyzing Content and Experience Encodings

Emotionally compelling content is at the core of storygames. I defined Storygame content in Section 4.1 as "data that encode a valuable experience for presentation to some audience." In the last chapter, I presented a tool, Sherlock, that assists the interpretive task of *decoding* a record of experience and *encoding* both properties and location of the content and values representing its reception into a format that is suitable for analysis. The format, the interactive cinematic experience (ICE) schemata, associates content in individual traversals to response recordings. The model and the tool work together to produce datasets that can be used to analyze both individual and group player experiences. In this chapter, I describe an analytical approach using the encoded data from the first three chapters of *The Wolf Among Us* (*TWAU*) to assess records of player experience in the context of specific types of storygame content and to evaluate whether ICE reveals additional insights as compared to the encoding approach described in Chapter 3. By connecting the subjective evaluation of the desirable qualities in the game to measurable qualities of longer player-traversal records, we can evaluate larger datasets and discover content strategies and response patterns in CCAGs and other storygame genres.

Storygame content in CCAGs focuses on emotional decisions, events, and themes, and a successful game is one that organizes content for maximum effect. As discussed in Section 4.1, this content consists primarily of authored performances (animations and voice-acting) and text (dialogue, scripts) and is produced in a similar fashion to television content, incorporating work by writers, actors, and animators. Because of the cost of such content, it is important to maximize its effectiveness by reusing lines of dialogue and animations as much as possible while delivering the experience. *TWAU* demonstrates an efficient and effective use of content that pulls players into the story, as evidenced by both the uniformity and variation in user responses to different content segments as required by the desired. It does this in part by leveraging combinations of positive and negative values in emotional content events and focuses on the story values of community and intimacy. The analysis described in this chapter demonstrates that some values are more consistent in their presentation than others and that variability in the emotional *expression* of themes is where CCAGs excel. Emotionally charged events are positioned to anchor player emotions and to ground players after sequences where the emotional responses vary the most while providing rewards in the form of contingent content that is emotionally tied to the key themes players have control over, particularly justice, community, and intimacy. The analysis also shows that the content strategy used in *TWAU* successfully elicits a variety of player responses by showing a range of variance in engagement and valence across the same content events. Drawing from the previous analysis and the practical experience from performing an ICE encoding of the content of the first three chapters of *TWAU*, I address the following questions in this chapter:

- How can the emotional content of storygame content be analyzed using subjective annotations?
- How can responses to storygame content be measured and described using multiple player-traversal records?

This chapter is organized as follows. First, I describe the method used to encode the first three chapters of *TWAU* using the ICE schemata described in Chapter 4. An ICE encoding consists of annotations of content according to type and their interpretation as emotional content events, and includes both a value for magnitude (or strength) and valence (positive or negative) of the content's effect. I then analyze the encoded data from three perspective: the distribution of story charges across values, annotated story value charges in scenes, and player response data. These results are compared to those from the previous encoding approach, providing evidence that the new schemata gives additional insight into how *TWAU* provokes a varying range of variance in emotional responses while using the same content. Finally, I discuss the implications of this analytical approach and some resulting theories, particularly the use and charaterization of anchor events on the content structure.

ICE schemata provides finer granularity segmentation and labeling compared to the previous approach by incorporating individual "content events" and supporting multiple attributes that describe them. The expansion of the analytical scope to include valence, as discussed in Section 4.1.3, supports representing the often ambiguous treatment and individual facets of major content themes. Valence, as discussed in Section 2.3.1, is a measure of the positive or negative dimensions of emotion, and can be estimated from the presence of certain facial expressions.

Through the process of annotating the content and reviewing player response data, I observed details about the role emotional events played in the game content. The same events were responsible for eliciting a wide variety of emotionally intense experiences. This observation motivated the investigation into the causes of differing user responses to similar content events. After comparing respones and considering the needs of the developer, the role and value of emotional anchors in shaping player responses emerged.

These anchor content events appear designed to align player emotions regardless of the emotional variations in the previous content. The encoding and its correlations to the player response data are a starting point for future investigations using more refined labels, more machine-learning support, or an improved interface for assigning and shifting values and annotating the reasoning behind their selection. Story values were a useful tool for understanding how choices not only provide opportunities to influence actions in the storyworld but also color players' perceptions of the lifeworld of the characters and their communities.

7.1 Encoding Details

For the encoding of the first three chapters of *TWAU*, the procedure detailed in Chapter 5 was used with additional support for specifying label data from Sherlock to locate and label the same content segments across different player traversal records. For details on the interface and its organization, see Appendix C. The additional granularity over the original schema included identifying dependency structures (e.g., content that depends on player choice), specific dialogue and physical acts, and both classifying content events by themes and assigning estimates of valence and magnitude to individual actions using the concept of

SVCC (described further in Section 4.3.2.4). The player response data used in this chapter were obtained using the method described in Appendix B, and only the valence and engagement values extracted from videos of the players using the Affdex SDK [110] were used in this chapter.

The relevant data subsets were extracted from SHERLOCK as a Javascript Object Notation (JSON) file that contained the full set of content segment frame locations, associated IDs, text content, and dependency relationships (if an instance of a `StoryPerformanceEvent`). This dataset was imported into R [173], where the full dataset of player expression values was also imported and both were indexed by a common reference frame count. Then, the data for each lens were filtered and processed, and the results were tabulated as described in the following sections.

7.1.1 Scenes

The first three chapters of *TWAU* were annotated with 19 scenes that contained the original 42 beats, with the vast majority present in the first chapter. The second and third chapters involve more scenes and more free-roaming content between the investigation of the courtyard and the movement around the business office. The scenes are primarily organized around character entrances or exits, as these change the nature of the dynamics, although Scene 13 includes Crane's exit. When Crane leaves the business office, the entire atmosphere changes from one of antagonism to one of camaraderie, with the contrast between the two emphasizing the effect. Not only do scenes provide another scope for assessing how beats are organized (allowing aggregation of story value occurrences and overall player responses), but they also accelerate annotation by limiting the possible character

Table 7.1: Scenes in Chapters 1–3 in *TWAU*

Scene	Description	Beats
1	Toad admonished by Bigby	1,2
2	TJ scolded by Toad	3
3	Faith beaten by Woodsman	4,5,6,7,8,9,10,11
4	Toad's car destroyed	12
5	Struggle with Woodsman	13
6	Faith interrogation	14,15,16,17,18,19,20
7	Beauty asking for promise	21
8	Beast asking for Beauty	22
9	Collin asks for a drink	23,24,25
10	Snow asks for Bigby	26
11	After Faith's head	27,28,29
12	Grendel	30
13	Crane and Snow	31
14	Snow and Bigby	32, 33
15	Bufkin	34,35,36
16	Mirror	37
17	Books	38,39,40
18	Mirror 2	41
19	Toad	42

labels and relative start locations of content events.

7.1.2 Story Value Definitions

One of the principal components of creating an ICE encoding is the subjective evaluation of emotional content based on the annotator's impressions of the content, both in terms of which value it pertains to and its valence and magnitude. The story value labels, or values for short, selected for the annotation of the first three chapters are listed in Table 7.2. Each value describes a perceived influence of an act or content event on player feelings about a theme. For more details on story values and SVCC, see Section 5.2.6. Each value label can further be annotated as possessing either a positive and negative valence. Value changes from one valence to another that describe player or character actions provide insight into character relationships and overall themes. For instance, a player's options to respond to choices involving Faith are constrained by how the relationship has progressed: Bigby's responses that are player-controlled, even when relatively negative, reflect Bigby's positive regard toward Faith. The value charge as represented by the label's meaning, though, captures the varying perception of the situation, such as acknowledging the situation Faith is in (negative) compared with the actions that Bigby might take to improve it (positive).

7.1.3 Facial Expression Measurements

Two metrics for measuring player response relative to annotated content were examined: the facial expression metrics of valence, which range from -100 (negative) to 100 (positive), and engagement, which ranges from 0 to 100. Heart rate and skin conductivity visualization graphs were available from SHERLOCK's

Table 7.2: Values in Chapters 1–3 in *TWAU*

Value	Positive	Negative
Community	Living in harmony	Living in conflict
Ego	Confidence	Asshole
Intimacy	Achieving closeness	Isolation
Survival	Achieving life	Losing life/health
Justice	Achieving justice	Injustice
Bravery	Addressing conflict	Bravado
Duty	Responsibility	Obligation
Truth	The truth is worthwhile	(Necessary) Lies

interface when reviewing particular content segments but were not incorporated into the numerical analysis presented in this chapter. The reason is that the specificity of heart rate and skin conductivity to particular event types is still unclear at the level of granularity of individual content events. GSR was excluded from this analysis primarily due to the presence of artifacts caused by participant movement, breathing, and talking and loss of sensor data during segments of the traversals. While these limitations prevented a more global analysis of skin conductivity events, future studies that control for these factors should include GSR as a feature, including the rate and frequency of both GSR peaks that can be attributed to events and nonspecific peaks. Heart rate variability and other measures could also be incorporated into an analysis in the same way that engagement and valence are used in this chapter.

7.2 Results

As in the previous analysis chapter, results are organized based on various aspects of the content and player responses. The first section presents as analysis of the protostory layer encoding for variations in content event charges and themes as well as content dependencies between option text and contingent content events.

Next, I discuss the player responses, first with a labeling method based on player response and, second, with a method for measuring player response variations and confirming that the content in *TWAU* elicits a range of responses, with some exhibiting more variations than others. The full encoding used in these analyses containing annotations of the locations and values of the dataset can be found as a supplementary file.

7.2.1 Story Values

Story values are a fundamentally subjective measure, and provide an index of perceived tension, theme and intended emotional reception as judged by an annotator. In this dataset, I performed the annotations, though in the future it would be worthwhile to compare different annotators labels for agreement. The sensitivity of selecting events for inclusion, such as `SpeechAct`, `PhysicalAct`, or `Spectacle`, depends on a subjective judgement on whether it was significant enough in its perceived influence and not on any automated or numerical measure.

In Figure 7.1, each scene is associated with the sum of the SVCC magnitudes annotated. Each `ContentEvent` may contain multiple charges, each of which is included in the value. This can explain, in part, the bias in this graph's sums to events that contain multiple charge components and thus ambiguity. The expectation is that scenes that are more emotionally charged will stand out from the accumulation of smaller estimations, and indeed the first half of the three chapters appear to contains more ambiguity and value-charged moments than the second half.

In Figure 7.2, each scene is further broken down by the valence of the estimates. Each magnitude is shown side-by-side, using the absolute value of the charge. The

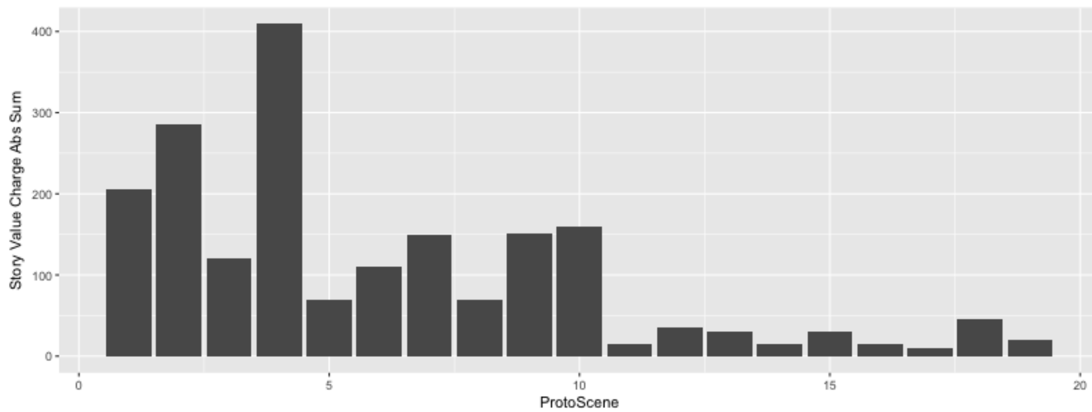


Figure 7.1: SVCC sums for each scene

general content skew is negative, as would be expected by the overall tone of the film noir genre of the game, which incorporates dark and mature themes. The distribution indicates a bias due to the increased density of emotional events and charges at the beginning of the three-chapter sequence. While the revelation of Faith's death is powerful, many content events in the first chapter appear to be intended to persuade the player to feel sympathetic toward Faith before her death.

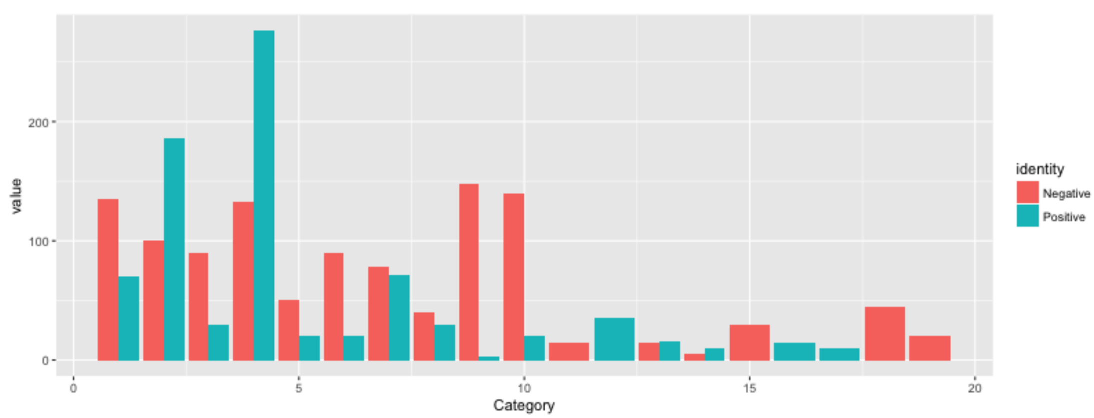


Figure 7.2: SVCC valence breakdown for each scene. When only one event charge valence is present (negative or positive), the bar expands to fill the entire width of the scene.

The relative value shift from the start of the first chapter to the end of the third chapter reflects several possible factors. First, the emotional intensity of the fight scene and the tension between Faith and Bigby were drawn out over multiple beats, allowing story value charges to accumulate. On the other hand, the encounter with Crane was brief but explosive; it only lasted for a single beat. Second, the primary positive value events all seem intended to establish a relationship between the player and Faith. Finally, the second half of the sequence contained more free-roaming periods and a much lower density of emotional interactions.

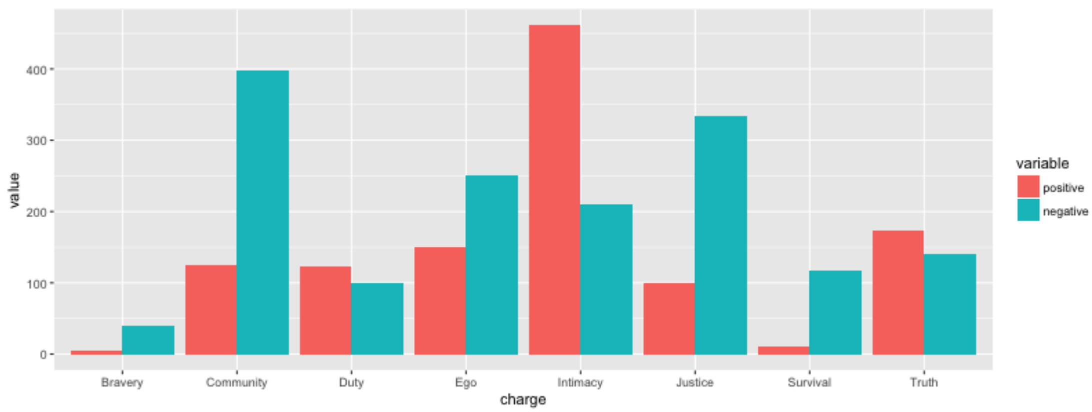


Figure 7.3: StoryValueChargeComponent sums for each value category

Figure 7.3 shows the distribution of story charges across each value category. For most values, there is a clear emphasis on presenting both positive and negative treatments, with the exception of survival and bravery, which may be biased toward negative and positive interpretations, respectively. This view also demonstrates the overall tone of the work as a function of which themes are more prevalent in emotionally charged content.

7.2.1.1 Options Vs. Content

SVCCs can be classified by associated content type. Figure 7.5 shows a comparison of different value categories for contingent content (content that is played only in a traversal where a player selects a relevant option) while Figure 7.4

shows the same distribution for content events that do not depend on particular choices. These are again the sums of respective absolute magnitudes. It is clear that there are many opportunities for Bigby to pursue isolation through negative valence events, although it is also interesting to note the number of positive content events that reinforce the story value of community. This reflects the identification of Bigby as a lone wolf detective stereotype while confirming emphasis placed on possibilities for positive changes through actions within the player (and player-character's) control. It is worth noting that these events also reflect the distributions of content selected by the players during their traversals and not the overall distribution of content in all possible traversals; thus, this analysis reflects the overall tendencies of the group of players towards the content and not the total possibility space of possible traversals. The contrast and distribution of values in each set of option texts demonstrates some of these action tendencies. In other words, a negative option is perceived as even more negative when in a set containing positive options.

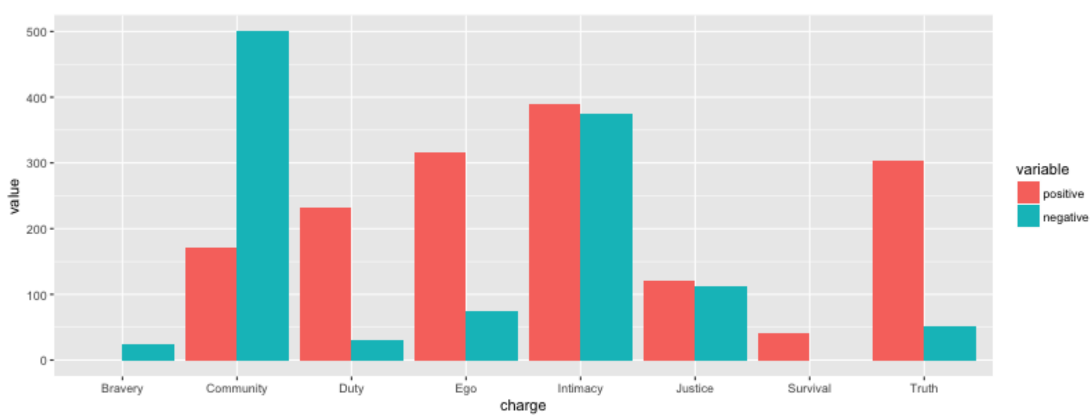


Figure 7.4: SVCC independent option selection

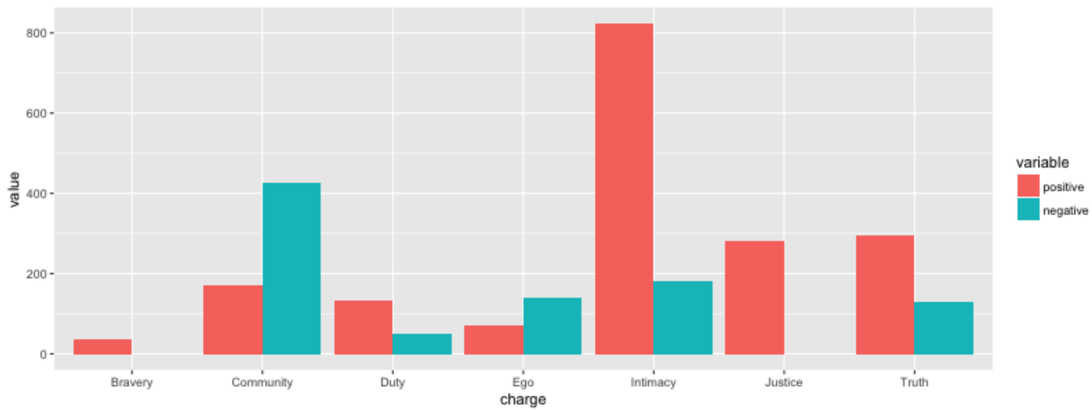


Figure 7.5: SVCC contingent on a particular option

7.2.2 Player Responses

This analysis looks at two features: choices and content events. As opposed to the previous analytical approach of aggregating all affect events that occur during each beat, ICE annotations support isolating responses to specific events by looking for responses during or after the event in a time window. The response can be analyzed either within a single player traversal record or by comparing responses to the same event across different player records. Each event can be considered an elicitor, although other sources can be the cause of the response. By collecting responses to similar events, this provides a means of collating player responses in the same way that using relative density and time windows of event occurences did in Chapter 3.

Individual content events can be classified according to the presence or absence of certain types of player responses. In this section, I use event timestamps to provide bounds for classify player responses to each event using the detected engagement and valence values. I use the minimum or maximum value from each time window, which results in three labels, shown along with their thresholds in Table 7.3, while the counts for each class are tallied in 7.6. For this analysis, valid

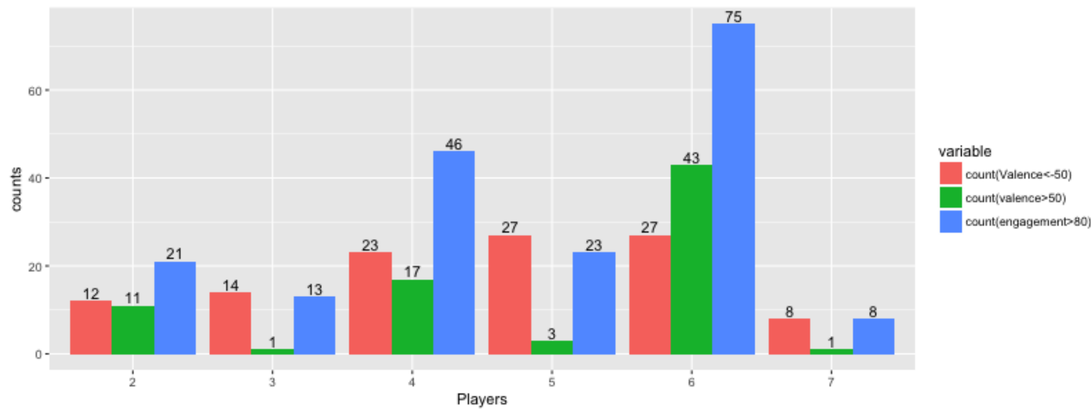


Figure 7.6: Story Event Responses

ranges for each label were set based on observations, but in the future, individual player's emotional distributions may be a better approach.

Table 7.3: Content labels based on player experience measurements

Label	**Value**	**Calculation**	**Threshold**
Engaged	Engagement	Max	80
Negative	Valence	Min	-50
Positive	Valence	Max > 50	50
Mixed	Valence	Max > 50 and Min < -50	-50,50

Player response variation was calculated by examining the responses by each player for the annotated `StoryContentEvents` that did not depend on an option or state. Variance was calculated based on the squared distribution of differences between each individual and the population mean for each time window's maximum and minimum valence, and engagement. This, along with the labeling approach described in Section 5.2.5, provided a solid foundation to examine which content may have elicited which types of responses. The following tables are top 10 events after sorted by each of the specified fields (in either ascending or descending order). The associated protostory element ID and text were included to identify content, and most lines of dialogue and actions that were unique and memorable were recorded during the annotation process to assist in estimating magnitude and assigning story values.

7.2.2.1 Choices

I selected an eight-second time window following each decision point, and set the engagement threshold to 70. This means that if a player's engagement value exceeded 70 during any frame in the window, the choice would be tallied. During the three chapters included in the encoding, there were 51 protochoices, although not every player encountered every choice. The results are shown in Table 7.4.

Table 7.4: Count of Player Response Types to Choices

Player	Positive Valence	Negative Valence	Mixed Valence	Engagement
2	7	10	2	8
3	3	8	1	8
4	5	14	2	15
5	3	15	0	7
6	21	20	9	35
7	2	8	1	2

While the overall counts reflect the expressivity of the other measures, this particular view shows some interesting balances and bias toward negative effects following choices, when I assumed that there would be a higher percentage of choices that would be labeled with mixed or ambiguous responses. There were only a few choices that triggered both positive and negative effects. After totaling each choice using the method above, all choices that had resulted in at least one player response in each of the three categories (positive, negative, and mixed) were considered and are listed in the following three tables. Table 7.5 presents all choices for which at least one player had experienced both negative and positive affective readings within the eight second time window after the choice. Tables 7.6

and 7.7 show choices where at least one player responded positively or negatively, showing that there were far more negative responses than positive responses. This agrees with the overall charges anticipated by the assigned labels in the previous section. The ChoiceIDs are detailed in the dataset stored with this dissertation.

Table 7.5: Mixed Choices

ChoiceID	Positive	negative	mixed	eng
8	0	1	1	2
17	0	0	1	1
18	0	0	1	1
26	0	1	1	2
38	1	1	1	2
39	0	0	1	1
46	0	0	1	1

Table 7.6: All Positive Labeled Choices

ChoiceID	Positive	negative	mixed	eng
3	2	1	0	3
4	2	0	0	2
5	2	0	0	3
14	2	1	0	3
1	2	0	0	2
2	1	0	0	2
6	1	1	0	2
9	1	0	0	1
11	1	2	0	1
16	1	0	0	2
21	1	0	0	2
23	1	0	0	1
27	1	1	0	2
28	1	0	0	1
36	1	0	0	1
38	1	1	1	2
25	1	0	0	1

While these results were promising for dividing player responses into groups, a method for comparing player responses was needed to determine whether content

Table 7.7: All Negative Labeled Choices

ChoiceID	Positive	negative	mixed	eng
40	0	4	0	3
0	0	2	0	1
11	1	2	0	1
12	0	2	0	3
15	0	2	0	2
32	0	2	0	2
44	0	2	0	2
3	2	1	0	3
6	1	1	0	2
7	0	1	0	2
8	0	1	1	2
10	0	1	0	1
14	2	1	0	3
19	0	1	0	1
22	0	1	0	1
24	0	1	0	0
26	0	1	1	2
27	1	1	0	2
33	0	1	0	0
34	0	1	0	1
35	0	1	0	1
37	0	1	0	1
38	1	1	1	2
41	0	1	0	2
42	0	1	0	1
47	0	1	0	1
48	0	1	0	1
49	0	1	0	1

produced the variable responses that I anticipated, and so the focus of the next section is on the variance of the responses.

7.2.2.2 Story Content Performance Events

In order to focus on how the same content produced different results, this section's analysis ranks each content event according to the *variance* of the maximum and minimum values of each player's response. In 7.8, each choice is listed along with the calculated variance of all players' maximum engagement during the time window. The choices considered were selected from 64 choice prompts that were present in every player traversal.

Table 7.8: Highest **Engagement** Variance Content Events
FA: Faith Action, CS: Collin Speech, BuS: Beauty Speech, BA: Bigby Action, MS: Mirror Speech, WS: Woodsman Speech, FS: Faith Speech

EventID	Associated Text	Variance
FA10	FAITH steps on an axe, looking determined.	2992.21
FA18	FAITH removes her hand from Bigby's arm, turning away from him.	2907.40
FA7	Faith crouches and begins to search the WOODSMAN.	2804.27
CS120	People are scared of you.	2751.96
BuS70	Please, Bigby[...] promise me you won't tell 2689.72 Beast you saw me.	
BA16	Bigby turns, intending to pursue him.	2648.50
MS119	Through powerful magic, her whereabouts are concealed. Unfortunately for you, "These lips are sealed."	2580.45
WS131	I'm the woodsman, you whore.	2511.29
BS53	And I'm guessing it'd be bad for you to show up empty-handed.	2510.43
FS52	A hundred...	2496.33

Table 7.9 lists events whose engagement maximums exhibited the lowest variance. This means that most players had a similar engagement value following the

event. These events seem to function within their respective scenes as emotional anchors, providing a clear interpretation that signals players as to how to feel. For example, the mirror dialogue line is distinct in that it is clearly ambiguous about Faith's fate, and players all respond to the different voice and intriguing information, which thus elicits interest.

Table 7.9: Lowest **Engagement** Variance Content Events.
TS: Toad Speech, BfA: Bufkin Action. SS: Snow Speech, FA: Faith Physical Action, BS: Bigby Speech, Cr: Crane Speech

EventID	Associated Text	Variance
TS2	Just cut me a break.	910.66
BfA	Bufkin marks Faith as deceased in the Book of Fables.	1338.95
GS89	Must be nice being the Sheriff. Do whatever the 1353.89 fuck you like.	
SS105	Bigby do you have any idea what's going? on? How did this happen? Why her?	1389.58
FA19	Faith furrows her brows in anticipation.	1413.93
BS123	You don't want me catching you out of glamour again.	1461.81
Cr102	Any shred of evidence you two know what the hell you're doing?	1522.30
W30	Woodsman strangles Bigby.	1612.23
FS54	I'll be fine.	1619.13
FA20	Faith kisses Bigby on cheek.	1682.78

Table 7.10 lists events where the variance in player's maximum detected valence in the time window was highest. The response to these events depends heavily on other interactions and context, which may account for the variance.

These values suggest that categorizing and ranking content based on player responses can provide further insight into whether content is successful in providing both personal and "tailored" experiences while using the same content. In the following section, the results of the analysis of these values are assessed in close association to the content labeled during the annotation process.

Table 7.10: Highest **Valence** Variance Content Events.
TS: Toad Speech, BfA: Bufkin Action. SS: Snow Speech, FA: Faith Physical Action, BS: Bigby Speech, Cr: Crane Speech,

	EventID	**Associated Text**	**Variance**
59	Cr103	Call Vivian right this minute and let her know I'm coming in early for my... massage.	2925.79
60	Cr104	And where is the bottle of wine you were to purchase?	2877.63
9	TS134	Furry pricked gobshite. Tell me how to spend my money. <Spit>	2610.97
10	TS136	You want the big bad wolf to take you away? Then get the fuck back inside.	2610.97
6	TS122	Do you have any idea how much it costs to have an entire family in glamour?	2563.28
31	BA16	Bigby turns, intending to pursue him.	2540.79
11	TJS135	The lights are shaking again.	2537.51
45	Co81	Just gimme a drink, please!	2514.47
16	WS125	Aisle fuhged puhd yew in de ground.	2424.48
18	WS131	I'm the woodsman, you whore.	2393.37

7.3 Discussion

By encoding the content, I observed that certain patterns are only apparent when viewed without the attending effects and especially when viewed across a longer experience. The encoding process and results support the goal of identifying content variations, comparing context effects, and highlighted the role that each event played by comparing actual player responses.

The Telltale effect describes the variance and strength in the intended responses and is described in more detail in Section 8.1. It describes the often intense feelings a player has as a result of making primarily affective choices in response to dramatically presented content. There are three aspects of the phenomena that this chapter addresses:

1. The variations in the density of emotionally significant content

2. The ambiguity of content with respect to player goals and content themes

3. The variability of the reception of content based on the emotional context

These three content strategies are employed in CCAG to elicit emotions and support the production of interactive narratives using higher quality and more expensive content (including fully performed voice acting) without incorporating major branches of content that may not be experienced by all players, as is found in games such as *Detroit: Become Human* [45].

7.3.1 Content Variations

The content itself can vary in multiple ways due to the modality. In classical adventure games, content variation might include alternate content or changes is either the text displayed or the images involved, but contemporary games include both voice acting and cinematic presentations. Varying how a particular segment of content within the engine is one way CCAGs reuse content around and as a result of player choices. While the audio of the line for Faith accepting the offered matchbook is the same as the line for accepting the light, the cinematography is markedly different. Other lines, though similar or the same in their text and context, can be performed differently. One example of this is Bigby's reply to Bufkin's asking how he is doing. The text for this response is "Thanks for asking." The intonation of that part of the line, though, is different based on whether the player chooses to open up more or less by selecting one of the options, "I'm fine" and "I'm not great." This choice is also interesting due to the sheer contrast between the reply "Fuck off" and the previous content. While it could be considered a "Bad Wolf" role-playing rail, it seems inconsistent with the lines Bigby has so far delivered in the scene, especially given the presence of

Snow. Although several characters have insulted Bigby throughout the episode (e.g., Toad, Woodsman, Grendel, and Beast in some traversals), Bigby himself rarely employed expletives or cursed other Fables, and it may seem out of character to respond severely to a well-meaning character without similar responses available to other more antagonistic characters.

7.3.2 Context Effects

A temptation when performing an annotation task at the same time as developing and refining an annotation schema is to account for effects by introducing new features to address effects that were not originally anticipated. This was the case in the original conception for the SVCC. While estimating content values, the surrounding content made it difficult to isolate communicated emotional content without accounting for context. This suggests that minor variations may have a larger effect on surrounding content than the variation itself as an independent event. A line of dialogue can have a massive effect on the interpretation of a subsequent interaction. By using a fairly mechanistic procedure to annotate subjective ratings, the method served to defamiliarize the researcher enough to record units of emotional content in a more independent fashion. In the end, these context effects may need to be addressed through different types of annotations describing a function of the story event on preceding or following content. For example, the SVCC of the option "I'll tell you what I told Toad" is completely dependent on the previous decision made by the player, and so represents a challenge for a uniform annotation approach.

While a single dialogue line has independent charge, other lines serve to modify the perceived effects. One example of this is in the first scene. In each traversal, Bigby delivers the line, "You don't want me to catch you out of glamour again." The line is ambiguous in the sense that it can be interpreted as either a sympathetic warning or a veiled threat. The degree to which this line is received, either by Toad or by the player, is determined by the previous lines, which in turn are dependent on the options. This is accounted for by assigning a neutral charge to this line, though ideally it may function as an intensifier of the previous content's effects.

Combining a formal model of content with a tool to explore content variations has yielded a number of insights and results. As in other applications of content analysis, the use of a regular encoding format enables quantitative comparisons of content distributions and frequencies. The ability to evaluate not only content and corresponding responses but also patterns of dependencies among content provides evidence for specific strategies employed in TWAU and a method for evaluating their effectiveness using a collected dataset.

The encoding process discussed in this chapter provide a starting point for investigations into subjectivity. Some of the content design, such as the variability of content based on context and the roles that certain events that players responded to similarly played in resetting player expectations, echo the findings of Mateas and Stern on their design of beats in the interactive drama *Façade* [150]. They found that certain content would reset tension levels to an absolute value, whereas other content could build or decrease values in a relative fashion. In the current study, the SVCC was designed to be under-specified to allow for variations in interpretation and analysis rather than forcing interpretations of charges as having an absolute or relative effect. The fact that certain content events functioned as anchors was highlighted by both the rating process and the location and distribution of content events, as shown in the previous section. These actions occur in all traversals and serve to establish emotional context and important sequences, while serving as independent events that carried large enough charges to override previous event charges when scenes went against the valence trend.

SVCC estimation requires assigning numerical quantities to subjective phenomena, and there are a number of potential areas for improvement. First, the classification of content into story values benefits from analyzing the entire work for character values, thematic unity, and other elements before assigning specific labels or values. There are points where a character's text and subtext are at odds, and this provides much of the interest in a scene for players. For instance, Collin both strongly desires a drink and is using Bigby's guilt over his past misdeeds as a means to an end, but it is also clear that he cares about Bigby. Value classification does not provide much insight into the design of free-roam areas and the sub-beat dialogue interactions that accompany selected hot spots; however, it does provide insight into the design of choice content and the use of varying content to achieve desired effects. Finally, the ability to indicate that a particular story value charge is dependent or builds on previous story value expressions may be necessary for a complete picture of the particular story values of a given traversal.

7.3.3 Emotional Content Anchors

As I have found in my analyses, emotional content anchors are a content events that resets the current scene intention and emotional state. In *TWAU*, Toad's first line is "You going to just stand there?" This line conveys multiple facts and feelings simultaneously and has a strong emotional charge: it reminds the player of their scene intention (helping Toad with a disturbance, which may not have been clear from the introduction) while resetting their relationship to an antagonistic one in spite of any previous positive exchanges of sympathy. The accusatory subtext is unmistakable and is important for when Toad is cast in a suspicious light during a scene in which Bigby is investigating Toad's lies.

I discuss how minor content events can modify one another in terms of emotional values and the Telltale effect further in Section 8.1, but often, story value charges that are located at the end or beginning of a scene function either function as anchors or are strongly dependent on player choices. Several dependent lines occur at the end of scenes to reinforce relationships, such as Beast's parting shot at Bigby when the player chooses the silent option and lets the elevator door close in his face. Beauty's warm thanks when Bigby promises to not tell Beast about her also reflects a strong reinforcement at the end of the beat for the player's (verbal) support. These emotional cues are strategically located either before or after a key choice in a scene or at the beginning or ending of a scene for maximum effect, reminding players of their decision to aid or hurt another character, and they often provide closing scenes memorability with powerful emotional notes.

Emotional content anchors are present in all traversals and function as plot points that provide the backbone of the emotional trajectory of the scenes. They usually occur either at the beginning or the end of a beat. Examples include when Faith removes her arm from Bigby's after the concern for the Woodsman getting away, and the entire sequence of kissing Bigby on the cheek at the end of Chapter 1 (serving to leave the player with the impression that Faith desires intimacy as well). The player character can also participate in such anchor events through actions. For example, Bigby's statement, "And I'm guessing it'd be bad for you to show up empty-handed," reminds players of the stakes of the situation and sets up the next decision. Most players presented a negative valence expression during this speech act, likely responding to its implications as intended. The expression

Table 7.11: StoryContentEvents that function as emotional anchors in Chapters 1–3. IDs correspond to the ProtoContentEvent ID (e.g., BS53 refers to the ProtoSpeechAct53 associated with Bigby). Higher ID numbers indicate the time of creation, not the location within the traversal. Characters with the same initial have a second character that is the first letter different: Bs = Beast, Bu = Beauty, Bf=Bufkin. Comm.=Community,Brv=Bravery, Dty=Duty, Int.=Intimacy

ID	Value	Charges	Key Phrase/Action
TS12	Duty	-20	"Do something Bigby."
TS136	Justice	-15	"Then get the fuck back inside!"
WS131	Ego	-20	"I'm the Woodsman."
FA16	Intimacy	+5	Hand on arm.
BS53	Intimacy	+20	"Guessing it'd be bad for[...]."
FS61	Intimacy	+10	"You got me out of a bad situation[...]."
FA20	Intimacy	+10	"You're not as bad[...]."
Bu78	Brv.,Int.,Tru.	-10,+10,-10	"Please, Bigby[...], promise me."
BsS71	Truth,Comm.	-10,-10	"Have you seen my wife?"
CoS81	Community	-10	"Just gimme a drink, please!"
SnS84	Bravery	-10	"Bigby!"
S1	Surv.,Comm.,Dty	-100,-50,-50	Faith's head
Cr101	Comm.,Duty	-15,-20	"You are the one charged[...]."
Bf111	Comm.	+5	"How are you today, Mr. Bigby?"
Bu116	Justice	-20	"His daughter, Faith."

itself is one of intimacy, whereby Bigby empathizes with Faith's situation, but the response indicates that players empathize with Faith's negative situation more. Bigby's action guides the player into the interpretation and the desired emotional take on the situation. This further provides an indication that the annotation schemata may benefit from clarification of the believed intentional feeling, and an enacted feeling should be made in such events, especially since in humorous events there is an opposite response (e.g., Toad's increasing frustration leads to laughter or amusement).

I developed a candidate list of anchors prior to doing the analysis in this chapter, which are listed in Table 7.11. Many correspond to key plot points in the episode. Most anchors portray a single value, unlike other content events, which can often be ambiguous and offer multiple conflicting values. Finally, each major character has at least one emotional anchor content line that expresses a core value or interaction with the story, which makes sense as these are both memorable and important to establishing pre-existing relationships and setting impressions for players. It should also be noted that the words or action do not entirely contain the meaning. For instance, the anchor involving Faith's identity (Bu116) depends on the story it is embedded in. Others contain more obvious emotional words or actions.

During annotation, I encountered a number of edge cases that did not fit into the model of SVCC. In particular, the suspense created in the scene where Snow retrieved Bigby did not revolve around the portrayal of story values. This mode departed from other scenes in the episode in that it built itself around a series of inferences about Bigby and Snow's relationship and Snow's emotional state. The entire scene consisted of Snow looking worried and encountering another resident in the woodlands. Other content models may be necessary to capture this kind of effect.

7.3.4 Future Work

7.3.4.1 Valence Comparisons between Content and Response

Content can be both negative and positive, as can individual responses. This is most readily apparent in facial expressions. Players may furrow their brows or appear sad when a character they care about dies, in the same way that characters express their emotional responses based on their respective successes and failures. The relationship of the content valence to player response (including its own valence) is complex. While certain types of situations encourage players to adopt the emotions expressed by characters in a scene, several others evoke an opposite effect. This can be seen by analyzing the perceived content valence versus the response valence in the same way as the other analyses in this chapter. Another dimension to consider would be the estimate of the emotions expressed by characters in a scene. This is especially promising given the success of the Affdex SDK in identifying facial expressions of characters, which is discussed further in Appendix B.

7.3.4.2 Dependent Content Analysis

Choice text is a powerful and frequent location for emotionally charged content, as it provides a perspective into the character's view of a situation and a set of options for a player to change it. When annotating options in this research, many options were complex in the sense that they can be interpreted as conveying multiple values. Some choice sets have a clear choice that differs in value charge from others, while some choices have a uniform value expression.

One advantage of the ICE schemata is the ability to compare dependent content. Of the 206 `ProtoContentEvents` that I annotated, 57 depended on an option. This amounts to more than a quarter of the content events. This suggests that emotional content in this game was concentrated around player choices in the feeder line, the choice text associated with an option, and the varied content.

7.4 Comparison to the Previous Method

The ICE (described in Chapter 4) was built based on lessons learned from the first schemata described and evaluated in Chapter 3. The goal was to reveal how emotionally charged storygame content could be related to emotional responses using measurable features and to author formal schemata that support querying and analysis of patterns of content and responses across multiple player-traversal records.

Two key drawbacks of the simplistic annotation schemata were the lack of granularity and the inability to represent dependent content, such as the result of choices. The ICE schemata included relationships that capture the variability of content through links and states. Finally, SHERLOCK provides an interface for collating content and player-traversal data through visualizations, syncing video between content annotations, and aligning selected subsections of timelines based on element start positions.

In terms of the dataset itself, many factors may negatively impact the ability to extract features and to automate the identification of peaks. Some of these may have been present in the previous analysis. Longer player-traversal records, such as those in the dataset examined here, include artifacts in sensor readings, player mistakes (e.g., choosing the wrong option), and interference from factors outside the game content, such as interruptions.

ICE `StoryContentEvent` annotations contained the locations and content of emotionally charged events in each traversal that may at some point be automatically detected. They indicate where emotional content is located; thus, their distribution is an indication of the structure of the work. Story content events can be considered stimuli or elicitors in the sense that they are designed to provoke an emotional response in conjunction with other content, although it is difficult to link a particular event to a particular response using only the biometric indicators considered here. Future work should consider using these to automate surveys or queries to disambiguate the stimuli that players associate with while reviewing their own responses. The previous method, described in Chapter 3, used a classification system and a set of time windows at the granularity of dramatic beats to group and tabulate events by label for analysis, whereas this chapter focused on response windows at the granularity of individual actions, called story content events, for classifying responses. The content associated with these events uses these annotations to allow review of play through content, which, along with listening to players think-aloud and reviewing the various biometric measures over time, provide a fuller picture of the situation and context for the data.

7.5 Summary

This chapter described promising results from the analysis of player response data using an ICE schemata. While this study focused on the dramatic elements that are most evocative, future work should focus on other aspects of the experience, such as intentional actions or other types of associations or labels. Ultimately, the approach of using a schemata to organize annotations allows each aspect of the experience to build upon one another. Not every part of the annotated content can be directly tied to the responses recorded, as player attention and interest vary, and further work will be needed to refine the methodology and selection of features that incorporate passive and active cooperation on the part of players during study scenarios. The current approach relies on numerical representations extracted from videos of player pfaces, making it applicable to a much broader range of content than a method that requires skin conductivity or heart rate data.

One important topic for future work is to automate aspects of the analysis and to improve the interface for entering and evaluating annotations. For instance, locating potential content that could be emotionally significant could be aided by looking at both the performed emotions of characters and the textual content of the lines themselves, using automated dialogue extraction. Since games often have closed captions, the use of optical character recognition to extract dialogue lines could work with existing speech-to-text to improve the quality of diarization of the dialogue content for further analysis. Having a set of candidate events and their locations could further enable a human annotator to make assignments or to indicate whether or not an event is emotionally significant.

The annotation process itself can be a subject of study, especially when comparing annotations from different annotators. The additiion of an annotator ID could be used to reveal insights into ambiguity and disagreement; for instance, it may be that the story value labels themselves differ between two annotators, or that they include different events or the values ascribed to the same event. Each of these disagreements has the potential to further refine our understanding of intepretation and reception, both in terms of documenting discussion of content in a precise and computationally tractable representation while evaluating it against the raw response metrics. This could also be used to compare annotation schemata and corresponding theories based on how well they explain the relationships between content and response. I discuss how schemata and models can assist in comparing theories further in Section 8.2 along with the implications for the content design revealed through annotation and the study itself.

In the next and final chapter, I summarize the contributions of this dissertation and discuss the implications for the subgenre in the context of the Telltale effect.

Chapter 8

Conclusions

This dissertation makes several contributions to the field of interactive digital narratives, building upon insights and technologies from previous contributions in annotation and modeling as well as the related fields of affective computing and game studies. One of the approaches in affective computing is to focus on emotion as either a classification problem or as an optimization problem, treating an emotional state of a user/player as a response to an elicitor. In this thesis, I developed an encoding method that incorporated sets of interpretations and observations and integrated them into a model that describes both intent and context. I applied this encoding model to a set of player traversal records, gaining new insight into the role of emotional content within the work and providing additional insights into which events players may respond to differently. The model and approach of developing a corpus of both content and player responses are based on the observation that there are many effective patterns of combining story and games and many players that value those experiences, and we can build on a better understanding of those properties and relationships in our theory and future systems. I chose *The Wolf Among Us* as a case study based on the success of the game and the work by Telltale Games.

Telltale Games developed a subgenre that combines choices and cinematic presentation that departs from the traditional puzzle mechanics of the adventure game genre, introducing several innovations that are designed to increase the emotional impact of the works. These innovations include the core timed dramatic choice and the focus on cinematic presentations (including animated performances and voice acting) as fundamental to their method of storytelling. The types of emotional experiences that result are different from the kinds of feelings that many games involving strategic choices attempt to elicit, such as feelings of pride and accomplishment. Instead, Telltale Games aspires to draw forth feelings of responsibility and identification, which build on similar feelings from more traditional forms of storytelling. It does so by carefully constructing choices embedded in dramatic contexts with the goal of shaping players feelings toward characters, values, and their own decisions.

The description of this phenomena as an "effect" follows Wardrip-Fruin's descripton of other computational media effects in *Expressive Processing* [149]. These include a more refined and nuanced analysis of the more generally used term, the *Eliza* effect, to include authorial constraint and the nature of the breakdown in the illusion demonstrated by the famous program, *Eliza-Doctor*. Wardrip-Fruin contrasted this effect with the effects created by two other systems that represented different trade-offs between the simulation and legibility of their underlying processes: the *Tale-Spin* effect, after the generative system by Mehan that simulates plans of the various characters in detail, and the *Sim City* effect, which captures the internalization of the expressive process by the audience and the feeling of mastery that results. These three effects reflected a relationship between an underlying expressive process and the experience of a person interacting with it and thus were inspirational when considering how the Telltale Games approach to creating dynamic interactive narratives was distinct from other types of interactive narratives.

Choices are a significant element of narratives, both interactive and noninteractive. Murray defines *agency* as an aesthetic pleasure that results from the availability and exploitation of the properties of digital environments in her book *Hamlet on the Holodeck* [74]. Mateas extended this definition and clarified it as the "balance between the material and formal constraints," where the material constraints refer to the elements of narrative "available from the levels of spectacle, pattern, language and thought" and the formal constraints consist of the dramatic probabilities [180]. As discussed throughout this thesis, CCAGs do not seek to give players a high degree of agency. Instead, they focus the player's attention on dramatic scenes and sequences that maximize players' emotional responses to choices as well as their perceived and actual consequences. Payoffs highlight choice-based content through both immediate and longer-term consequences related to the story values involved. These features contribute to an emotional response that shapes the player experience in CCAG: the *Telltale effect*.

8.1 The Telltale Effect

The Telltale effect is the emotional response elicited from players by the integration of value-based choices in a charged cinematic context without the presence of significant agency within the story.

This effect is not merely an illusion of agency, nor is it merely an extension of the familiar narrative experiences created by cinema; rather, it is an effect that is based on the affordances of computational media to render dramatic performances and integrate content with constrained player choices. The choices and, significantly, the value-oriented content before and as a consequence of specific options selected, provoke players to have an affective response toward a topic, character, or previous choice. The exact feeling varies, but the context is designed to force players to express their feelings and responsibility for them even through inaction. This structure is present in *The Wolf Among Us* and was made evident through encoding a scene of the work as well as developing a model to account for player responses in a corpus of player traversals. The narrative structure persuades the player to adopt values as their own for the desired effect. For example, to make players feel invested in Faith's demise, the scene's objective from the point of view of the designers that was encoded in Chapter 5 is to persuade players to feel sympathetic toward Faith. These strategies result in several features that work well together to achieve the effect.

Three primary features of storygame content play a role in producing the Telltale effect in *The Wolf Among Us*: the cinematic presentation, the choices designed with social outcomes in mind, and finally the efficient integration of dramatic content and choices with payoffs both near choices and the inclusion of references to previous choices later in the games. I will describe each in further detail.

First, the story is told cinematically. Although variations in textual representation can be powerful and are used to great effect in several interactive fiction and hypertext forms, the unique capacity to significantly alter a given interpretation by changing timing and shot sequences is unique to cinematic presentation. Additionally, the ability to time options fits with the overall feeling of characters existing within a storyworld and prevents the type of situations where dramatic tension is dissolved as a result of an indeterminate time for a player to choose.

The cinematic presentation enables payoffs to be concentrated both in small ways following a choice selection and through more overt references to previous choices while using many of the same lines in each possible traversal. For example, as we saw, the initial response to Faith in Scene 4 following Bigby's rescue is based on the length and angle of the shot and her expression but is otherwise not reflected in any state changes or in many changes in dialogue. SHERLOCK supports making these comparisons between player traversals.

Second, player options for each choice are often designed with telegraphed social consequences. This means that for many choices, the potential response from the other character(s) is legible enough that players can imagine the emotional consequences of their choices. This is not always the case, as we saw in the silent response in the scene between Faith and Bigby. The imagined intention for short choice text is necessarily constrained, especially for certain ambiguous option labels. The option text itself is significant, such as when several options expressing sympathy with the other character can induce such sympathy in the player and highlight the significance of the choice that does not sympathize. This was the case in the minority response in the scene encoded and discussed in Chapter 5. The option text itself can, therefore, be associated with story value charges associated with potential consequences, and the distribution of options can reflect strategies to influence player perception. The social component incorporates the meaning of the value charge, as it represents the broad accessibility of the subgenre as well as what makes it a good candidate for encoding alongside player responses.

Third, alternative versions of content reinforce player selections, and each set of choices is usually believable within the context of a dramatic scene. This conceptualization of choices as integrated within a beat structure provides a different lens through which to view the decision-making process than that of Mawhorter's choice poetics. It moves the emphasis from player choices analyzed as strategic decisions, which either hinder or enable player goals, to content segments that depict story values that can be seen as opportunities for players to respond or choose according to their personal values or current feelings within the confines of the dramatic context.

While players can perceive some choices within *TWAU* and other Telltale Games as strategically motivated toward some larger goal, these are by far in the minority. The design goal is to provide a good story experience regardless of the success or failure of the player's individual dramatic choices (excluding actions relating to life-or-death situations of the player-character, which must be replayed in order to progress). Many choices are thus difficult to assess with the choice poetics analysis and further suggest that story value charge components can be a valuable tool for understanding how content structure design produces certain responses.

The Telltale effect results from the pattern of story value choices and character actions that set up specific opportunities and presents a solution to the challenges inherent in an engaging interactive drama that can be hand-authored using current technology and with reliable quality. The variety and strength of player responses that are represented both in the player response data set and in the reviews and fan attention show that this pattern is effective in managing the balance between authorability and player experience. Future interactive drama and interactive digital narrative systems may benefit from a closer look at the Telltale effect, and in particular, ways in which portions of it can be automated and operationalized into new systems. I believe that the rich set of authored decisions in this area in the form of released Telltale Games and the available player traversals can be a source of both new insights into how content and interaction can be understood as well as potential sources of training for future generative systems.

8.2 A Preliminary Agenda

Interdisciplinary work always must find a place within the context of multiple communities at once. This dissertation addresses the game studies community primarily, though the focus on interpretation and digital artifacts shares much in common with the digital humanities. The concerns of experience and assessment are also part of the larger concerns of the human-computer interaction field, which has recently focused more on player experience.

This section describes some of the ongoing conversations and where they might lead. One topic of interest to computational media scholars is how computational models can be used in productive discourse. In spite of the necessity of investigations around model development, the work of developing a model, such as ICE, in a form for broader adoption is often secondary to experimental results. AIIDE this year took the step of accepting artifact submissions, recognizing that working software is a significant yet distinct contribution from a research paper describing it. This trend is likely to continue, as both the practice of using and sharing models and datasets that implement them become more common.

This dissertation provides one example of schemata that not only functions to inform analysis but also serves to improve our ability to discuss features of the works itself. The ability to discuss interactive narratives where models can be tested against the same set of player traversals represents one possible way this sort of common ground could occur. Another would be to have a more formal invitation at a conference to use a single technological framework and compare models directly. I believe the approach of using a ground truth of experience records provides a means of escaping the difficult challenge of evaluating models on their own merits. A model, then, could be judged on whether it articulates some property that has not yet been described, or when its precision can facilitate a discussion about sets of features (such as the collection of story content events) that would be impossible otherwise. Currently, semantic web technologies such as those used in this dissertation have not won over a critical mass of the community, and so any effort to compare the results of computational models necessitates a translation and standardization of the models into the same framework. The goals, however, of making models compatible and comparable are worth pursuing, even if the specific technology may not have enough adoption.

In the digital humanities, it is likely that these experience records will become as important as the actual artifacts themselves. Already the importance of the interaction is recognized in works such as Malthrop, Grigar and Tabbi's *Traversals* [181] as a means of preserving electronic literature works. Ultimately, the goal should not just be to preserve works for a future population to understand but to make such experiences available to artists, scholars, and researchers to improve their craft and theories. The contributed dataset and platform that are described in this thesis hope to open some of the way to these sort of collaborations, especially focusing on the goal of bringing the toolset usability to the level that most scholars would feel comfortable using it to engage in research without coding or technical support needed at all.

Computational models offer this promise of being able to assess and communicate complex understandings, but like the theories behind interactive digital narratives, it will require more experimentation and growth to find the right combination of standardization and expressiveness.

8.3 Summary of Contributions

This thesis makes several contributions to the knowledge regarding interactive digital narratives. I collaboratively collected a dataset of player experience traversals of a commercially available choice-based cinematic adventure game and annotated it using narrative features in Chapter 3. The successes and failures of these annotations motivated the creation of a formal model of surface content, the ICE model, which associates player response features with content structures, adapting concepts from choice poetics and operationalizing Koenitz's term of the protostory in Chapter 4.

I developed a formal, descriptive model and used it to encode segments of traversals of an existing commercial storygame. This approach recognizes that storygame authors have an advanced understanding of the craft of shaping player experiences using choice, including the selection of dramatic incidents as well as the options that make sense during each choice point. The model includes a theory that surface-level features of story content can be used to interpret and catalog player responses to choices and content segments using story value charges and beats. It diverges from previous models that focus on either character goals or player goals, instead focusing on the dynamics and the consequences of surface content's interaction with player choices to shape story and player values. I provide an example encoding for a scene from *TWAU* in Chapter 5 and compare it to two other formal models relevant to storygames. I completed a full ICE encoding of the first three chapters and analyze the dataset using the new granularity in Chapter 7. I build on findings from both the analyses and method and develop a web-based annotation and visual data-mining tool that supports finding patterns in player responses across related player traversals, Sherlock in Chapter 6. In

7. I connect player response variations to aspects of the content in 8, using a encoding of the first three chapters of *TWAU*. The results provide the basis for a theory of how the subgenre of choice-based cinematic adventure games operates that is based on observations from encoding a scene of the dataset using the formal model as well as from working with the dataset in multiple capacities. The resulting player experience pattern is named the Telltale effect, which is described in this chapter.

8.4 Concluding Thoughts

Storytelling remains a frontier of computational media. Innovations in the modeling, generation, and theory related to interactive narratives in the field are exciting and ongoing. The complexity and diversity inherent in narrative discourse, such as the various genres and the many modes that storygames can occupy, reflects our own complexities and differences in terms of what we find valuable and how we respond to it. Emotions, both actual and represented, play a vital role in the experience of content, one distinct and yet integrated with other layers of meaning and structure. The ability to articulate these relationships and to understand them is particularly important as more and more human-level capabilities such as image recognition and emotion recognition are automated. The relationships between interactive experience and audience responses are important not just for their role in future generative systems, or even in assisting in the authoring of powerful stories, but for their capacity to increase our knowledge of how each of us values the experience itself.

Appendix A

Sample Encoding

This is the story value charge assignments for the protostory layer used in Chapter 5. In this table, "Par" represent the parent ID, the `ProtoChoice` which it is dependent on `ProtoChoice` for dependent content, whereas "PB" refers to the `ProtoBeat` it is contained in.

Table A.1: Story Value Assignments

X	ID	Valence	M	StoryValue	PC	Par	PB	Players
1	FA2	1	1	Justice/Injustice			1	all
2	PO1	1	5	Intimacy/Isolation	1		1	P3
2	PO2	-1	1	Intimacy/Isolation	1		1	P2,P5,P6
2	PO3	1	5	Duty/Preference	1		1	P4
3	BS1	1	5	Intimacy/Isolation		1	1	P3,P5,P6,P7
3	BS2	1	1	Justice/Injustice		1	1	P2,P5,P6,P7
3	BS3	1	5	Justice/Injustice		1	1	P4,P5,P6,P7
4	FA3	1	1	Intimacy/Isolation		1	1	P3,P5,P6,P7
5	FS2	1	1	Justice/Injustice			1	all
6	FS3	1	1	Intimacy/Isolation			1	all
7	FS4	-1	1	Justice/Injustice			1	all
8	FA5	1	5	Justice/Injustice			2	all
9	FA6	1	10	Community/Self			2	all
10	PO5	-1	10	Duty/Preference	2		2	P2,P4,P7
10	PO5	1	10	Intimacy/Isolation	2		2	P2,P4
10	PO6	1	10	Duty/Preference	2		2	P3,P5,P6,P7
11	BS6	-1	5	Duty/Preference		2	2	P2,P4
11	BA1	1	5	Duty/Preference		2	2	P3,P5,P6,P7
12	FA7	1	10	Justice/Injustice			2	all
13	BS7	1	5	Intimacy/Isolation			2	all
14	FA10	-1	1	Intimacy/Isolation			3	all
15	PO7	1	5	Intimacy/Isolation	3		3	
15	PO8	1	10	Intimacy/Isolation	3		3	P2
15	PO9	-1	5	Community/Self	3		3	all
15	PO10	-1	5	Duty/Preference	3		3	all
15	PO11	-1	10	Community/Self	3		3	all
16	BA5	1	10	Intimacy/Isolation		3	3	P2
17	FS11	1	5	Intimacy/Isolation		3	3	P3,P4,P5,P6,P7
18	FA12	1	10	Intimacy/Isolation		3	3	P2
16	BS10	-1	10	Community/Self		3	3	P3
19	FA15	1	10	Intimacy/Isolation			4	all
20	PO12	1	10	Community/Self	4		4	all
20	PO13	-1	10	Intimacy/Isolation	4		4	all
20	PO13	1	10	Duty/Preference	4		4	all
20	PO14	1	10	Duty/Preference	4		4	all
20	PO14	-1	10	Intimacy/Isolation	4		4	all
20	PO15	1	5	Intimacy/Isolation	4		4	all

Table A.2: Story Value Assignments (continued)

X	ID	Valence	M	StoryValue	PC	Par	PB	Players
21	BS12	1	10	Duty/Preference		4	4	P5,P6
21	BS12	1	10	Intimacy/Isolation		4	4	P5,P6
22	FS16	-1	10	Intimacy/Isolation		4	4	P5,P6
21	BS14	1	10	Duty/Preference		4	4	P3,P7
21	BS14	-1	10	Intimacy/Isolation		4	4	P3,P7
22	FS17	1	10	Intimacy/Isolation		4	4	P3,P7
23	BS15	1	1	Intimacy/Isolation			4	all
24	FS20	-1	1	Community/Self			4	all
25	FS21	-1	5	Community/Self			4	all
26	BS16	1	1	Intimacy/Isolation			4	all
27	BA8	1	5	Duty/Preference			5	all
28	PO16	1	1	Intimacy/Isolation	5		5	all
28	PO17	1	10	Community/Self	5		5	all
28	PO17	-1	10	Intimacy/Isolation	5		5	all
28	PO18	1	10	Intimacy/Isolation	5		5	all
29	BS17	1	1	Justice/Injustice		5	5	P4
30	FS24	-1	1	Intimacy/Isolation		5	5	P4
29	BS18	1	10	Community/Self		5	5	P2,P5,P7
29	BS20	1	10	Intimacy/Isolation		5	5	P3
30	BS21	1	5	Intimacy/Isolation		5	5	P3
31	FS28	-1	10	Intimacy/Isolation		5	5	P3
29	FS30	1	5	Intimacy/Isolation		5	5	P6
32	FA18	-1	10	Intimacy/Isolation			6	all
33	FA19	-1	5	Intimacy/Isolation			6	all
34	BS22	1	5	Intimacy/Isolation			6	all
35	FS31	-1	10	Intimacy/Isolation			6	all
36	BS23	1	20	Intimacy/Isolation			6	all
37	FS32	-1	10	Intimacy/Isolation			6	all
38	PO20	1	20	Intimacy/Isolation	6		6	all
38	PO21	-1	20	Intimacy/Isolation	6		6	all
39	BS24	1	20	Intimacy/Isolation		6	6	P2,P3,P5,P6,P7
40	FS33	1	10	Intimacy/Isolation		6	6	P2,P3,P5,P6,P7
39	BS26	-1	20	Intimacy/Isolation		6	6	P4
39	BS26	-1	20	Community/Self		6	6	P4
40	FS34	1	10	Duty/Preference		6	6	P4
41	FS35	-1	10	Intimacy/Isolation			6	all

Appendix B

Additional Dataset Processing

B.1 Dataset Preparation

The original analysis in Chapter 3 incorporated almost exclusively hand-annotated features, including annotating heart rate peaks, skin conductivity changes and various self-reports based on reviewing the video data and visual indications. These annotations were synthesized into a single expressive signal which provided a lens into player responses both as a group and as distinctive individual experiences. While I incorporate the Affdex SDK data into the analysis in Chapter 7, I found that the extracted features that are detailed in this chapter to be subject to issues relating to the resolution of the captured data and the likelihood that artifacts such as breathing, movement and talking interfered with the usefulness of the data for more global analysis, though the extracted and aligned data for each of these measures is included in the associated dataset for completeness.

The dataset consists of biometric sensor readings for each player on two dimensions, (`HeartrateData` and `SkinConductivityData`), for each frame of recorded data. I further extract nine additional values from the facial expression of players using the output from the Affdex SDK, resulting in more than 11 million data points over more than a million frames of video. In addition to these continuous dimensions, my collaborator indicated points where players reported affect via the Sensual Evaluation Instrument [143], which provided an additional 317 annotations of self-reported player affect.

One of the first steps was to remove some of the burden of hand-annotation of player response data. I obtained the original values for each frame of the heart rate and skin conductivity readings from the video by applying optical character recognition to each frame, and further applied the Affdex SDK to the masked video to obtain values facial expression values for each frame.

The general procedure was to collate the data into tables in a database based on the number of fields, using the frame index as a means of relating data to a common reference time. For the Affdex SDK dataset, this involved populating a database table with 32 fields, of which 6 were "basic" emotions whose value fell within the range [0:100]: joy, sadness, anger, contempt, fear, surprise. Two fields represented the aggregate measures of engagement [0:100] and valence [-100:100].

For the heart rate and skin conductivity readings, the optical character recognition method resulted in errors that had to be hand corrected. This was done by filtering readings that were outside of the expected range (for heart rate, this included values below 40 and above 150). For any value found, I inspected the original frame and corrected the value in the database. I used the openly available Tesseract OCR library found at `https://opensource.google.com/projects/tesseract`.

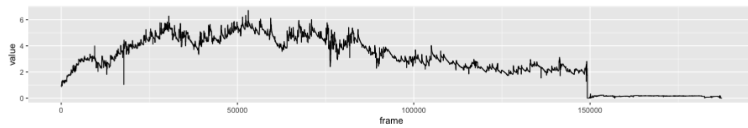
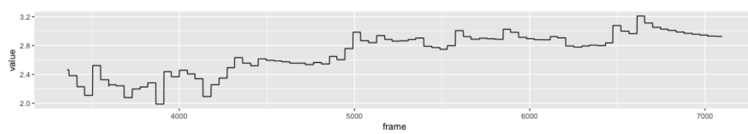
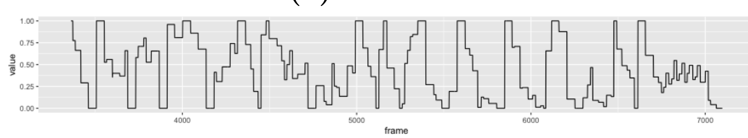


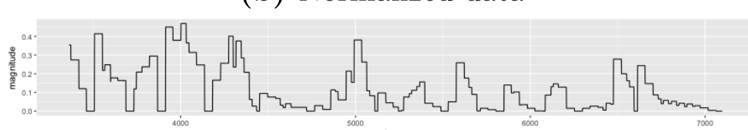
Figure B.1: Full SC for Player 2



(a) Raw data



(b) Normalized data



(c) Magnitude added

Figure B.2: GSR Data for Player 2 during Beats 1, 2 and 3

B.2 Feature detection and normalization

In order to extract features, one of the normal steps for processing skin conductivity is to normalize it. The values exhibited wandering on longer time scales (see Figure B.1), which is due to accumulation of oils. The formula for this is simply scaling each data point according to the max and min in the original, using a 7-second time window obtained from [182]. This amounts to 210 frames. We store these normalized datapoints into the database under a new collection in the database, resulting in the values depicted in B & C of Figure B.2. The magnitude is useful for further filtering of obtained features.

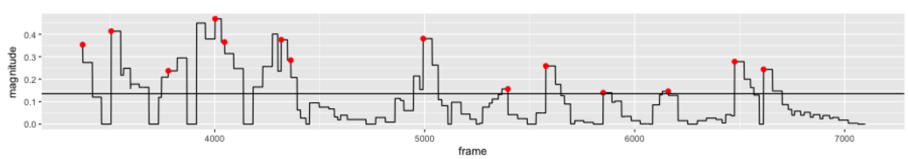


Figure B.3: GSR Peaks detected for Player 2 during Beats 1, 2 and 3

I next apply a peak detection algorithm, resulting in the peaks shown in Figure B.3. This required several steps, due to aliasing. The aliasing is the sharp lines resulting from the lower sample rate of the SC sensor than the framerate. I applied a simple peak detection algorithm, followed by removal of duplicates and a filter based on the mean magnitude of all peaks detected in the dataset. This results in 425 peaks detected in User 2, shown in figure B.4

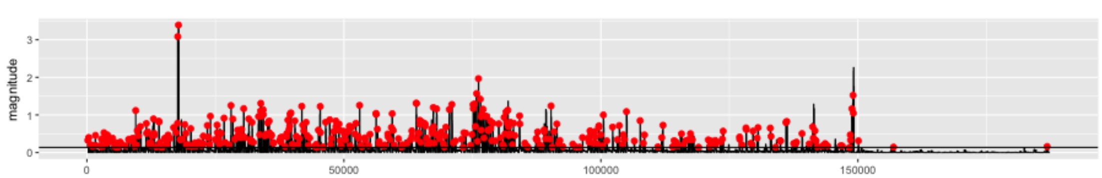
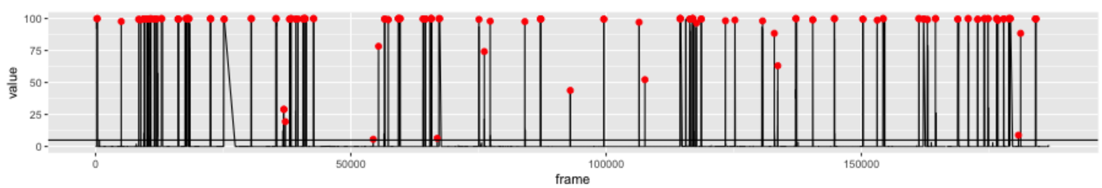


Figure B.4: GSR Peaks detected for Player 2 during Beats 1, 2 and 3

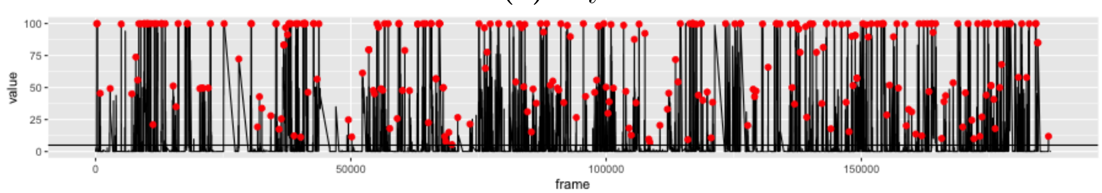
Another method used by [138] are time windows. I employ both feature detection and time windows as the two distinct methods, with features inserted as new entities in the ontology and time windows associated with their respective element from the ICE ontology.

The peak detection method resulted in the roughly one peak every 14.6 seconds for Player 2, with varying magnitudes. Also, note that the scaling of the SC values decreases rapidly around frame 150000. In this case, a stepwise function can be used to determine a new mean, dividing the dataset into frames before and after that point. Player 3 also did not have a sufficiently complete dataset, and so I excluded Player 3 from comparisons involving SC.

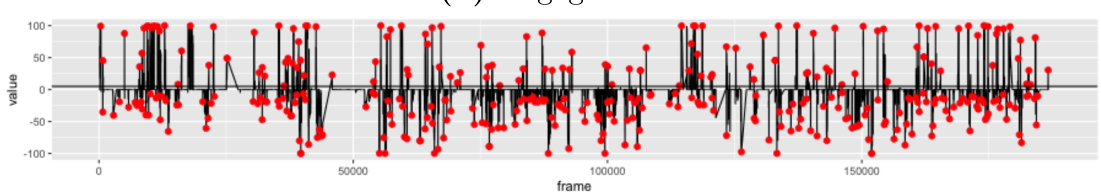
For the emotional values that are provided by the Affdex SDK, I apply a similar method of peak detection, except I filter out peaks with value less than 5 for all but valence, which I filter out [-5,5]. Charts for Joy, Engagement and Valence for User 2 are shown in Figure B.5. For the full set of visualizations, refer to the Appendix or the dataset itself. I obtained these by processing the video files using an objective C program and inserting the resulting values directly into the database.



(a) Joy



(b) Engagement



(c) Valence

Figure B.5: “Joy”, “Engagement” and “Valence” Peaks detected for Player 2 during Episode 1

Appendix C

Sherlock Annotation Interface

The SHERLOCK annotation interface supports identifying, labeling and analyzing content in interactive narratives. This section describes the relevant features and the process used to encode the first three chapters of *The Wolf Among Us* included as part of this dissertation and described in the analysis in Chapter 7.

The annotation process involves selecting a content type corresponding to the respective instantial layer ICE element (`Chapter`, `Scene`, `Beat`, `Choice`, `PhysicalAct`, `SpeechAct` and `Spectacle`) and setting "in" and "out" pointers (representing the start and end locations of a span of frames) on the timeline(s). Each timeline retains a reference to the last known pointer locations, and the length of the span can be used to set pointers on other timelines, as is common when the same content is present. The waveform data is rendered for each segment on the server using BBC's waveform tool[1]. These aid in annotating and aligning content by providing a visualization of the sound amplitudes over time. This is especially useful for annotating corresponding speech acts in several traversals, and it is incorporated in other annotation tools.

[1] Audio Waveform Image Generator. `https://github.com/bbc/audiowaveform`

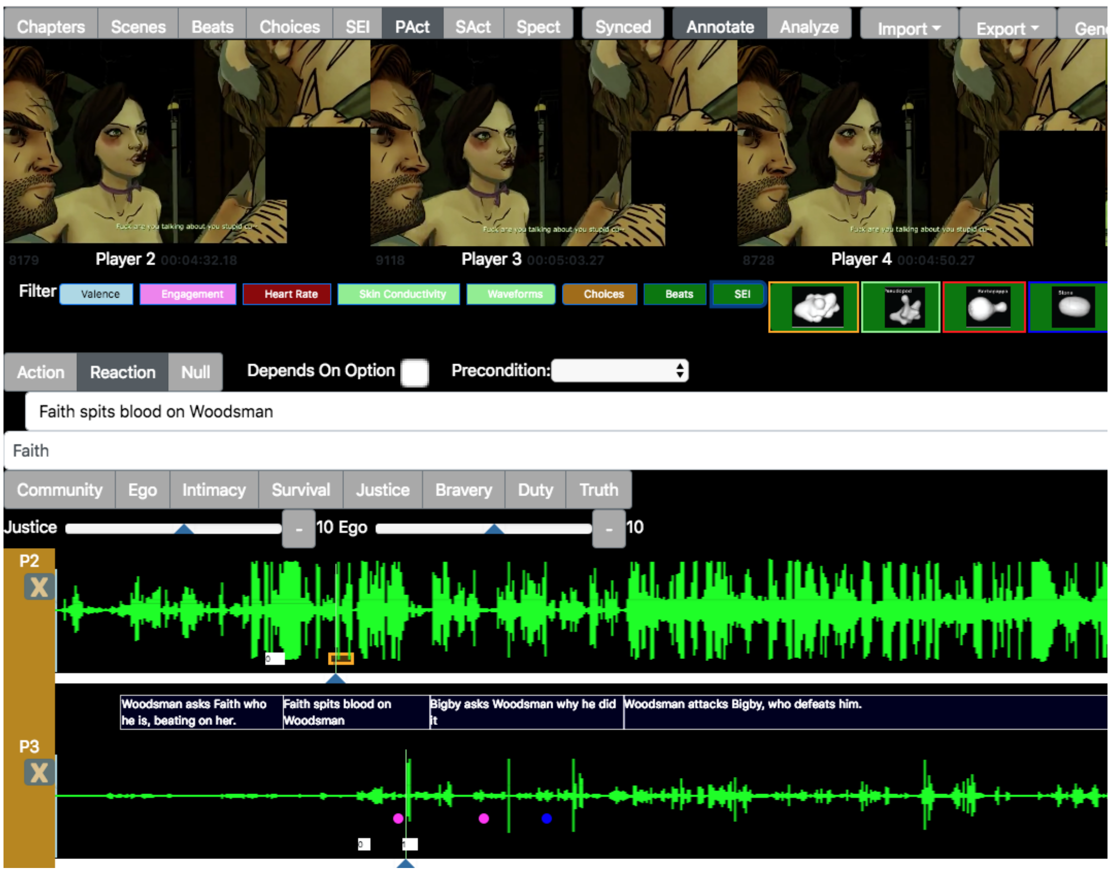


Figure C.1: Sherlock Annotation Panel

I used SHERLOCK to annotate each segment type in a top down fashion, starting with the chapters and finishing with individual `ContentEvents`. For a description of each element and its role in the overall schemata, see Chapter 4. Figure C.1 shows each major component of the annotation interface for a `StoryContentEvent`. While there is an important distinction in the schemata for the relationship between protostory and instantial elements, the interface combines the two concepts for convenience, storing the relevant information in the correct place.

Each segment includes a set of characters, such as Bigby or Snow, that are made available in segments contained within it to speed annotation. While I imported the previously labeled beat and choice locations and information, I annotated scenes to provide additional context and to better understand how beats were related. Many scenes only contained a single beat, such as scenes that involved Beauty and Beast, but most beat sequences built on one another, particularly the sequence of beats with Faith and Collin. These are listed in Table 7.1.

The annotation mode of SHERLOCK builds on several assumptions that make translating the data into the various representations simpler. These include the scope of various elements. For instance, when assigning a `StoryContentEvent` to a behavior (representing the ongoing action or reactions characterizing the beat that contains it), the relevant behaviors can be populated by finding the containing beat, and the corresponding agent is automatically selected based on the behavior's associated agent (or reaction agent). Likewise, for each segment, the set of potentially involved characters is a subset of the containing segment, reducing the number of options an annotator must choose from.

For choices, each option can have zero or more SVCCs assigned (see Section

5.2.6). The interface provides a dropdown box to select the current value, while each option has buttons for adding or deleting SVCCs and specifying a charge value. Future versions should render the assigned value on the timeline itself as well as enabling tweaking of values in a script form, as the balancing of value assignments requires relative evaluations rather than absolute evaluations, and currently having to enter edit mode makes those comparisons difficult.

The purple background of the top timeline label indicates that it is both the current timeline and is marked for insertions, whereas the second timeline (P3) is yellow and is also marked. The process for inserting a new annotation starts by selecting a toggle at the top of the screen, and then choosing the start and stop locations on the timeline (pressing "i" and "o" to set the in and out locations, respectively), which in turn alter the displayed frame of the corresponding video. `StoryValueChargeComponents`, the element in the ICE schemata that captures subjective assessment of value charged actions and events, are added by clicking one of the values, listed underneath the textbox. These create a new instance of a slider representing the charge, which can vary from -100 to 100. These can be deleted through the "-" button next to them. There is also a checkbox which indicates a particular content event is dependent on a story option. As mentioned in Chapter 6, there are two modes for playing and navigating videos: synced and separate. Playing in synced mode can assist in rapidly aligning content based on its position relative to a previously annotated element, such as a beat or chapter. Finally, each `StoryContentEvent` has an option to associate it with the previous `ProtoOption` through the "Depends on Option" checkbox. Behind the scenes, this dynamically associates the correct option on export and during any edits, in case the option is deleted or edited. In general, the architecture follows this strategy – dynamically associating relevant parts of the schemata during execution rather than maintaining the links through associations with object references. This has the advantage of simplifying export and import operations, as once the relevant objects have been imported, they can be discovered through the same search mechanisms when necessary. Finally, the interface supports exporting and importing from several sources, including the Fuseki server mentioned in Chapter 6 and a JSON file format for both the encoding as well as datasets for further analysis, used in this chapter, which includes the relevant data and segment locations in each record.

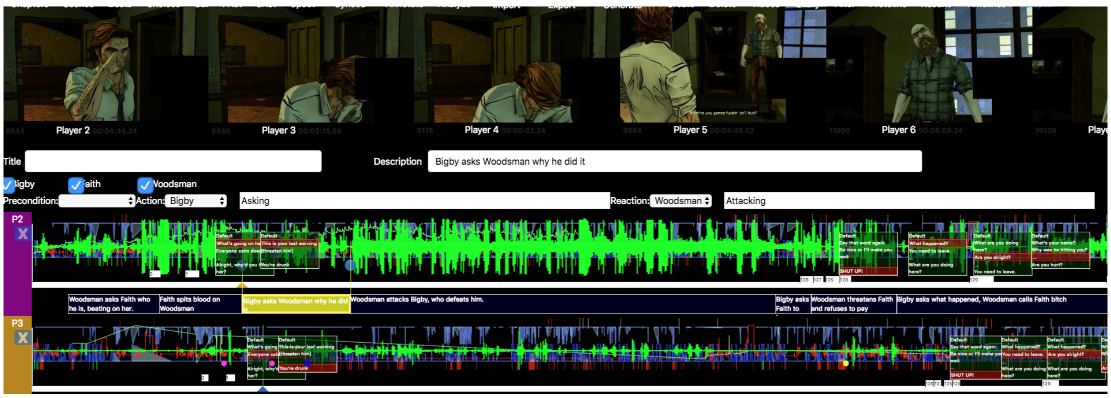


Figure C.2: Sherlock Beat Annotation Interface

There are two text inputs for every segment: a title and a description. A `StoryContentEvent` instance (`DialogueAct` or `PhysicalAct`) also has a toggle that selects whether it is a part of the action, reaction or neither, and by selecting it the relevant character is populated in the agent dropdown menu. For beats, the agent associated with the action and reaction are also available, along with a set of checkboxes to indicate characters are present. See Figure C.2 for the beat editing version of the annotation interface. In addition to creating and editing spans, the ability to add segments to an existing Protostory element relationship is important, and a feature I used heavily. If an existing element is already selected and the create key is pressed ("C"), any new segments will be associated with the same Protostory parent element. If all instantial segments have been deleted, the Protostory element is also deleted. Any selected spans can be edited by pressing the "e" key, which allows changing the start and end locations using the handles or the keyboard shortcuts, <alt> "i" and <alt> "o".

Bibliography

[1] S. Poria, E. Cambria, R. Bajpai, and A. Hussain. “A review of affective computing: From unimodal analysis to multimodal fusion”. In: *An international journal on information fusion* 37 (Sept. 2017), pp. 98–125.

[2] G. N. Yannakakis, P. Spronck, D. Loiacono, and E. André. “Player modeling”. In: *Dagstuhl Follow-Ups*. Vol. 6. 2013.

[3] E. Kaltman, J. Osborn, N. Wardrip-Fruin, and M. Mateas. “Getting the GISST: A Toolkit for the Creation, Analysis and Reference of Game Studies Resources”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. FDG ’17. New York, NY, USA: ACM, 2017, 16:1–16:10.

[4] P. Mirza-Babaei. “Biometric storyboards: a games user research approach for improving qualitative evaluations of player experience”. PhD thesis. University of Sussex, 2014.

[5] B. Drenikow and P. Mirza-Babaei. “Vixen: interactive visualization of gameplay experiences”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. ACM, Aug. 2017, p. 3.

[6] G. Smith, J. Whitehead, and M. Mateas. “Tanagra: A Mixed-Initiative Level Design Tool”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. 2010.

[7] A. A. Reed. “Changeful Tales: Design-Driven Approaches Toward More Expressive Storygames”. PhD thesis. University of California, Santa Cruz, June 2017.

[8] M. Cavazza and D. Pizzi. “Narratology for Interactive Storytelling: A Critical Introduction”. In: *Proceedings of the Third International Conference on Technologies for Interactive Digital Storytelling and Entertainment*. Ed. by S. Göbel, R. Malkewitz, and I. Iurgel. TIDSE’06. Springer Berlin Heidelberg, 2006, pp. 72–83.

[9] N. Szilas and I. Ilea. “Objective Metrics for Interactive Narrative”. In: *ICIDS*. Ed. by A. Mitchell, C. Fernández-Vara, and D. Thue. 2014, pp. 91–102.

[10] B. Tearse, N. Wardrip-Fruin, and M. Mateas. “Minstrel Remixed : Procedurally Generating Stories”. In: *Artificial intelligence* (2010), pp. 192–197.

[11] C. Martens. *Ceptre: A Language for Modeling Generative Interactive Systems*. 2015.

[12] D. L. Roberts, H. Narayanan, and C. L. Isbell. “Learning to Influence Emotional Responses for Interactive Storytelling”. In: *AAAI Spring Symposium:* (2009).

[13] C. Klimas. *Twine*. 2009. `http://http://twinery.org/`.

[14] M. Mateas and A. Stern. “Structuring content in the Facade interactive drama architecture”. In: *Proceedings of Artificial Intelligence and Interactive Digital Entertainment* (2005), pp. 93–98.

[15] R. Evans and E. Short. “Versu—A Simulationist Storytelling System”. In: *IEEE Transactions on Computational Intelligence in AI and Games* 6.2 (June 2014), pp. 113–130.

[16] J. McCoy, M. Treanor, B. Samuel, A. A. Reed, N. Wardrip-Fruin, and M. Mateas. “Prom week”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. New York, New York, USA: ACM Press, May 2012, p. 235.

[17] M. Mateas and A. Stern. “Façade: An experiment in building a fully-realized interactive drama”. In: *Game Developers Conference, Game Design track*. Citeseer, 2003.

[18] A. Reed, J. Garbe, N. Wardrip-Fruin, and M. Mateas. “Ice-Bound: Combining Richly-Realized Story With Expressive Gameplay”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. 2014.

[19] Failbetter Games. *Fallen London*. 2009. `https://www.failbettergames.com/fallen-london/`.

[20] Telltale Games. *The Wolf Among Us*. 2013.

[21] A. Salter. *What Is Your Quest?: From Adventure Games to Interactive Books*. University Of Iowa Press, 2014, p. 202.

[22] J. McCoy, M. Treanor, and B. Samuel. “Prom Week: social physics as gameplay”. In: *Proceedings of the International Conference on Foundations of Digital Games* (2011).

[23] R. P. Evans and E. Short. *The AI Architecture of Versu*. 2015.

[24] M. Mateas and N. Wardrip-Fruin. “Defining Operational Logics”. In: *Proceedings of the Digital Games Research Association* (2009), pp. 1–8.

[25] J. Osborn. “Operational Logics At Work”. PhD thesis. 2015.

[26] J. C. Osborn, N. Wardrip-Fruin, and M. Mateas. "Refining operational logics". In: *Proceedings of the International Conference on the Foundations of Digital Games*. ACM, Aug. 2017, p. 27.

[27] C. Crawford. *The Art of Computer Game Design by Chris Crawford*. Washington State University, 1982, p. 81.

[28] A. M. Sullivan. "The Grail Framework: Making Stories Playable on Three Levels in CRPGs". PhD thesis. UC Santa Cruz, 2012.

[29] N. Wardrip-Fruin. *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. The MIT Press, Sept. 2009, p. 441.

[30] E. Aarseth. "Aporia and Epiphany in "Doom" and "The Speaking Clock": The Temporality of Ergodic Art". In: *Cyberspace Textuality: Computer Technology and Literary Theory*. Ed. by M. Ryan. Indiana University Press, 1999.

[31] N. Montfort. *Twisty Little Passages: An Approach to Interactive Fiction*. The MIT Press, Dec. 2003.

[32] A. Reed, N. Wardrip-Fruin, and M. Mateas. "The Eurekon Design Pattern in Expressive Storygames". In: *Intelligent Narrative Technologies 7*. 2014, pp. 52–55.

[33] E. J. Aarseth. *Cybertext: Perspectives on Ergodic Literature*. The Johns Hopkins University Press, Aug. 1997.

[34] A. Grow, S. Gaudl, P. Gomes, M. Mateas, and N. Wardrip-Fruin. "A Methodology for Requirements Analysis of AI Architecture Authoring Tools". In: *International Conference on the Foundations of Digital Games*. 2014.

[35] K. Hamilton. *What Telltale Could Learn From Life Is Strange*. `http://kotaku.com/what-telltale-could-learn-from-life-is-strange-1683794466`. 2015.

[36] J. Bycer. *Player Agency: How Game Design Affects Narrative - Game Wisdom*. `http://game-wisdom.com/critical/player-agency-game-design-narrative`. Accessed: 2018-4-2. July 2014.

[37] M. Gerardi. *Readers debate the importance (or irrelevance) of choice in Telltale games*. `https://games.avclub.com/readers-debate-the-importance-or-irrelevance-of-choic-1798274854`. Accessed: 2018-4-2. Dec. 2014.

[38] B. Maltbie. *Stop Saying Telltale Games Aren't Games*. `http://dispatches.cheatcc.com/2038`. 2016.

[39] A. Froschauer. *Clementine Will Remember All of That*. `http://ontologicalgeek.com/clementine-will-remember-all-of-that/`. 2014.

[40] M. Maloney. *Out on a Limb: Practical Approaches to Branching Story.* 2017. https://www.youtube.com/watch?v=4-MGT6SZWWo.

[41] C. Dring. *Telltale: "We have yet to tackle a romantic comedy".* https://www.gamesindustry.biz/articles/2017-10-12-telltale-we-have-yet-to-tackle-a-romantic-comedy. Accessed: 2018-4-29. 2017.

[42] M. Sturges and D. Justus. *The Wolf Among Us.* 2014.

[43] *Life is Strange.* Dontnod Entertainment, 2015.

[44] *Heavy Rain.* Quantic Dream, 2010.

[45] *Detroit: Become Human.* Quantic Dream, 2018.

[46] P. A. Mawhorter. "Artificial Intelligence as a Tool for Understanding Narrative Choices". PhD thesis. University of California, Santa Cruz, 2016.

[47] D. K. Elson. "Modeling Narrative Discourse". PhD thesis. Columbia University, 2012.

[48] R. Mckee. *Story: Substance, Structure, Style and The Principles of Screenwriting.* Kindle. It Books, Dec. 1997.

[49] J. C. Osborn, B. Samuel, M. Mateas, and N. Wardrip-Fruin. "Playspecs: Regular Expressions for Game Play Traces". In: *Proceedings of the AIIDE.* Nov. 2015.

[50] R. Robinson, Z. Rubin, E. M. Segura, and K. Isbister. "All the feels: designing a tool that reveals streamers' biometrics to spectators". In: *Proceedings of the International Conference on the Foundations of Digital Games.* ACM, Aug. 2017, p. 36.

[51] H. Koenitz. "Towards a specific theory of interactive digital narrative". In: *Interactive Digital Narrative: History, Theory and Practice.* Ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, and T. İ. Sezen. Google Boo. Routledge, 2015. Chap. 6, p. 302.

[52] M. Mateas and A. Stern. "Integrating plot, character and natural language processing in the interactive drama Facade". In: 2003.

[53] D. K. Elson and K. R. Mckeown. "Extending and Evaluating a Platform for Story Understanding". In: *Intelligent Narrative Technologies II.* 2009, pp. 32–35.

[54] G. N. Yannakakis and J. Togelius. "Experience-Driven Procedural Content Generation". English. In: *IEEE Transactions on Affective Computing* 2.3 (July 2011), pp. 147–161.

[55] R. Barthes. *s/Z. trans. by Miller, Richard.* Blackwell Publishing, 2002.

[56] J. Drucker and B. Nowviskie. *Temporal Modelling: Conceptualization and Visualization of Temporal Relations for Humanities Scholarship.* Tech. rep. ACH/ALLC, 2003.

[57] A. Salter and J. Murray. *Flash: Building the Interactive Web*. MIT Press, 2014, p. 192.

[58] M. A. Finlayson. “The Story Workbench: An Extensible Semi-Automatic Text Annotation Tool”. In: *Proceedings of the 4th Workshop on Intelligent Narrative Technologies (INT4)* (2011), pp. 21–24.

[59] R. D. Peng. “Reproducible research in computational science”. en. In: *Science* 334.6060 (Dec. 2011), pp. 1226–1227.

[60] F. Moretti. *Distant Reading*. en. 1 edition. Verso, June 2013.

[61] F. Moretti. *Graphs, Maps, Trees: Abstract Models for a Literary History*. 2007th ed. Verso, 2007.

[62] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann. “On close and distant reading in digital humanities: A survey and future challenges”. In: *Eurographics Conference on Visualization (EuroVis)-STARs. The Eurographics Association*. 2015.

[63] L. Klein. *Distant Reading after Moretti*. `http://lklein.com/2018/01/distant-reading-after-moretti/`. Accessed: 2018-5-19. Jan. 2018.

[64] M. Wegrzyn, M. Vogt, B. Kireclioglu, J. Schneider, and J. Kissler. “Mapping the emotional face. How individual face parts contribute to successful emotion recognition”. en. In: *PloS one* 12.5 (May 2017), e0177239.

[65] R. Ingria, R. Sauri, J. Pustejovsky, R. Gaizauskas, A. Setzer, G. Katz, D. Radev, and J. Castano. “TimeML: Robust Specification of Event and Temporal Expressions in Text”. In: *New directions in question answering* 3 (2003), pp. 28–34.

[66] R. Prasad, N. Dinesh, and A. Lee. “The penn discourse treebank 2.0”. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).* (2008).

[67] D. K. Elson. “DramaBank: Annotating Agency in Narrative Discourse”. In: *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12)* (2012), pp. 2813–2819.

[68] M. Treanor and M. Mateas. “An Account of Proceduralist Meaning”. In: *DiGRA ’13 - Proceedings of the 2013 DiGRA International Conference: DeFragging Game Studies*. Vol. 7. Aug. 2014.

[69] B. Mears and J. Zhu. “Design patterns for silent player characters in narrative-driven games”. In: *Proceedings of the International Conference on the Foundations of Digital Games*. ACM, Aug. 2017, p. 59.

[70] R. Stevens, T. Satwicz, and L. McCarthy. “In-game, in-room, in-world: Reconnecting video game play to the rest of kids’ lives”. In: *The ecology of games* (2008).

[71] M.-L. Ryan. “Diagramming narrative”. In: *Semiotica* 165.1-4 (2007), pp. 11–40.

[72] K. Oatley. *Such stuff as dreams : the psychology of fiction.* eng. Chichester, West Sussex, U.K.: Wiley-Blackwell, 2011.

[73] D. Cage. *Postmortem: Indigo Prophecy.* 2006.

[74] J. H. Murray. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace.* The Free Press, 1997, p. 324.

[75] D. Herman. *Storytelling and the Sciences of Mind.* MIT Press, 2013, p. 428.

[76] T. Tzvetan and W. Arnold. “Structural Analysis of Narrative”. In: *NOVEL: A Forum on Fiction* 3.1 (1969), pp. 70–76.

[77] G. Genette. *Narrative Discourse: An Essay in Method.* Cornell University Press, 1983, p. 285.

[78] V. Propp. “Morphology of the Folk Tale”. In: *Folklore* (1928).

[79] M. Bal. *Narratology: Introduction to the Theory of Narrative.* University of Toronto Press, 1997, p. 264.

[80] S. B. Chatman. *Story and Discourse: Narrative Structure in Fiction and Film.* Cornell University Press, June 1980.

[81] I. Mani. “Computational Narratology”. In: *the living handbook of narratology.* Ed. by P. Hühn, J. C. Meister, J. Pier, and W. Schmid. Accessed: 2017-4-2. Hamburg: Hamburg University. URL: `http://www.lhn.uni-hamburg.de/article/computational-narratology`.

[82] M. Schmierbach. “Content Analysis of Video Games: Challenges and Potential Solutions”. In: *Communication methods and measures* 3.3 (Aug. 2009), pp. 147–172.

[83] K. Krippendorf. *Content Analysis An Introduction to Its Methodology Third Edition.* SAGE Publications, Inc, 2013, p. 456.

[84] K. Karpouzis, G. N. Yannakakis, N. Shaker, and S. Asteriadis. “The platformer experience dataset”. In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII).* 2015, pp. 712–718.

[85] M. A. Finlayson. “Collecting Semantics in the Wild: The Story Workbench”. In: *Proceedings of the AAAI Fall Symposium on Naturally Inspired Artificial Intelligence (published as Technical Report FS-08-06, Papers from the AAAI Fall Symposium)* 1 (2008), pp. 46–53.

[86] M. A. Finlayson, W. Richards, and P. H. Winston. “Computational Models of Narrative: Review of the Workshop”. In: *AI Magazine* 31.2 (2010), pp. 97–100.

[87] W. G. Lehnert. “Plot Units and Narrative Summarization*”. en. In: *Cognitive science* 5.4 (Oct. 1981), pp. 293–331.

[88] R. Damiano, V. Lombardo, and A. Pizzo. “Formal Encoding of Drama Ontology”. en. In: *Virtual Storytelling. Using Virtual Reality Technologies for Storytelling*. Springer, Berlin, Heidelberg, Nov. 2005, pp. 95–104.

[89] V. Lombardo, C. Battaglino, A. Pizzo, R. Damiano, and A. Lieto. “Coupling conceptual modeling and rules for the annotation of dramatic media”. In: *Semantic Web* 6.5 (Jan. 2015), pp. 503–534.

[90] J. Parkkila, F. Radulovic, D. Garijo, M. Poveda-Villalón, J. Ikonen, J. Porras, and A. Gómez-Pérez. “An ontology for videogame interoperability”. In: *Multimedia tools and applications* 76.4 (Feb. 2017), pp. 4981–5000.

[91] O. Sacco, A. Liapis, and G. N. Yannakakis. “Game Character Ontology (GCO) A Vocabulary for Extracting and Describing Game Character Information from Web Content”. In: *Proceedings of the 13th International Conference on Semantic Systems*. ACM, Sept. 2017, pp. 9–16.

[92] A. L. Benjamin H. Detenber. “The influence of form and presentation attributes of media on emotion”. In: *The Routledge Handbook of Emotions and Mass Media*. Routledge, 2010. Chap. 16.

[93] S. J. Ahn, J. Bailenson, J. Fox, and M. Jabon. “Using automated facial expression analysis for emotion and behavior prediction”. In: *The Routledge Handbook of Emotions and Mass Media*. Routledge, 2010. Chap. 20.

[94] J. Zinoman. *November 2009: Jason Zinoman on Robert Mckee*. `https://www.vanityfair.com/news/2009/11/robert-mckee-200911`. Accessed: 2018-5-2. Nov. 2009.

[95] J. Weston. *The Film Director’s Intuition: Script Analysis and Rehearsal Techniques*. en. Kindle. Michael Wiese Productions, Sept. 2003.

[96] N. T. Proferes. *Film Directing Fundamentals: From Script to Screen*. Woburn, MA: Focal Press, 2001.

[97] M. Joyce. *afternoon: a story*. 1987.

[98] S. Moulthrop. *Victory Garden*. {Eastgate Systems Inc}, Jan. 1992.

[99] P. C. Heartscape. *Howling Dogs*. `http://slimedaughter.com/games/twine/howlingdogs/`. Accessed: 2018-4-10. 2012.

[100] B. Laurel. “Toward the design of a computer-based interactive fantasy system”. PhD thesis. Ohio State University, 1986.

[101] T. Adams and Z. Adams. *Slaves to Armok: God of Blood Chapter 2: Dwarf Fortress*. 2006.

[102] B. Samuel, J. Ryan, A. J. Summerville, M. Mateas, and N. Wardrip-fruin. "Bad News : An Experiment in Computationally Assisted Performance". In: 2016.

[103] J. McCoy, M. Treanor, B. Samuel, and A. A. Reed. "Prom Week: Designing past the game/story dilemma". In: *FDG* (2013).

[104] J. McCoy, M. Treanor, B. Samuel, A. A. Reed, M. Mateas, and N. Wardrip-Fruin. "Social Story Worlds With Comme il Faut". In: *IEEE Transactions on Computational Intelligence in AI and Games* 6.2 (June 2014), pp. 97–112.

[105] S. Harmon. "Narrative Encoding for Computational Reasoning and Adaptation". PhD thesis. University of California, Santa Cruz, 2017.

[106] M. Treanor, B. Schweizer, I. Bogost, and M. Mateas. "The Micro-Rhetorics of Game-O-Matic". In: *Proceedings of the International Conference on the Foundations of Digital Games*. ACM, 2012.

[107] N. Taylor and N. Taylor. "Me and Lee : Identification and the Play of Attraction in The Walking Dead". In: *Game Studies* 15.2005 (2015).

[108] K. R. Scherer, A. Schorr, and T. Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research*. en. Oxford University Press, May 2001.

[109] P. Ekman. "An argument for basic emotions". In: *Cognition and Emotion* 6.3-4 (May 1992), pp. 169–200.

[110] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, and R. e. Kaliouby. "AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit". In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. New York, NY, USA: ACM, 2016, pp. 3723–3726.

[111] J. A. Russel. "A Circumplex Model of Affect". In: *Journal of Personality and Social Psychology* 29.6 (1980), pp. 1161–1178.

[112] K. Isbister. *How Games Move Us: Emotion by Design (Playful Thinking)*. en. Reprint edition. The MIT Press, Mar. 2016.

[113] B. Cowley, D. Charles, M. Black, and R. Hickey. "Toward an Understanding of Flow in Video Games". In: *Computer Entertainment* 6.2 (July 2008), 20:1–20:27.

[114] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. "Measuring and defining the experience of immersion in games". In: *International journal of human-computer studies* 66.9 (Sept. 2008), pp. 641–661.

[115] M. Slater and S. Wilbur. “A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments”. In: *Presence: Teleoperators and Virtual Environments* 6.6 (Dec. 1997), pp. 603–616.

[116] B.-C. Bae and R. M. Young. “A Use of Flashback and Foreshadowing for Surprise Arousal in Narrative Using a Plan-Based Approach”. In: *ICIDS*. Ed. by U. Spierling and N. Szilas. Vol. 5334. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, Nov. 2008, pp. 156–167.

[117] B.-C. Bae and R. M. Young. “A Computational Model of Narrative Generation for Surprise Arousal”. In: *IEEE Transactions on Computational Intelligence in AI and Games* 6.2 (June 2014), pp. 131–143.

[118] S. Giannatos, Y. G. Cheong, M. J. Nelson, and G. N. Yannakakis. “Generating narrative action schemas for suspense”. In: *AAAI Workshop - Technical Report* WS-12-14 (2012), pp. 8–13.

[119] Y. G. Cheong. “A Computational Model of Narrative Generation for Suspense”. PhD thesis. Raleigh, North Carolina, 2007.

[120] B. O’Neill. “Toward a computational model of affective responses to stories for augmenting narrative generation”. In: *ACII’11 Proceedings of the 4th international conference on Affective computing and intelligent interaction - Volume Part II*. Springer-Verlag, Oct. 2011, pp. 256–263.

[121] N. Ravaja, M. Turpeinen, T. Saari, S. Puttonen, and L. Keltikangas-Järvinen. “The psychophysiology of James Bond: phasic emotional responses to violent video game events”. en. In: *Emotion* 8.1 (Feb. 2008), pp. 114–120.

[122] N. Ravaja, T. Saari, J. Laarni, K. Kallinen, M. Salminen, J. Holopainen, and A. Järvinen. “The Psychophysiology of Video Gaming: Phasic Emotional Responses to Game Events”. In: *Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play*. 2005, pp. 1–13.

[123] L. E. Nacke. “An Introduction to Physiological Player Metrics for Evaluating Games”. In: *Game Analytics: Maximizing the Value of Player Data*. Ed. by M. Seif El-Nasr, A. Drachen, and A. Canossa. London: Springer London, 2013, pp. 585–619.

[124] L. Nacke. “Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment”. PhD thesis. Blekinge Institute of Technology, Karlskrona, 2009.

[125] L. Nacke and C. A. Lindley. “Affective Ludology, Flow and Immersion in a First-Person Shooter: Measurement of Player Experience”. In: *The Journal of the Canadian Game Studies Association* (2009).

[126] R. Robinson, J. Murray, and K. Isbister. ““You’Re Giving Me Mixed Signals!”: A Comparative Analysis of Methods That Capture Players’ Emotional Response to Games”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. New York, NY, USA: ACM, 2018, LBW567:1–LBW567:6.

[127] C.-H. Hjortsjö. *Man’s Face and Mimic Language*. en. Studen litteratur, 1969.

[128] E. Friesen and P. Ekman. “Facial action coding system: a technique for the measurement of facial movement”. In: (1978).

[129] B. Martinez, M. F. Valstar, B. Jiang, and M. Pantic. “Automatic Analysis of Facial Actions: A Survey”. In: *IEEE Transactions on Affective Computing* PP.99 (2017), p. 1.

[130] T. Baltrušaitis, P. Robinson, and L. P. Morency. “OpenFace: An open source facial behavior analysis toolkit”. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Mar. 2016, pp. 1–10.

[131] R. Robinson, K. Isbister, and Z. Rubin. “All the Feels: Introducing Biometric Data to Online Gameplay Streams”. In: *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY Companion ’16. New York, NY, USA: ACM, 2016, pp. 261–267.

[132] B. Farnsworth. *Facial Action Coding System (FACS) - A Visual Guidebook*. https://imotions.com/blog/facial-action-coding-system/. Accessed: 2018-3-26. Dec. 2016.

[133] W. V. Friesen and P. Ekman. “EMFACS-7: Emotional facial action coding system”. In: *Unpublished manuscript, University of California at San Francisco* 2.36 (1983), p. 1.

[134] F. Bousefsaf, C. Maaoui, and A. Pruski. “Remote Assessment of the Heart Rate Variability to Detect Mental Stress”. In: *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*. PervasiveHealth ’13. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013, pp. 348–351.

[135] K. Alghoul, S. Alharthi, H. A. Osman, and A. E. Saddik. “Heart Rate Variability Extraction From Videos Signals: ICA vs. EVM Comparison”. In: *IEEE Access* 5 (2017), pp. 4711–4719.

[136] P. Mirza-Babaei, L. E. Nacke, J. Gregory, N. Collins, and G. Fitzpatrick. “How Does It Play Better?: Exploring User Testing and Biometric Storyboards in Games User Research”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’13. New York, NY, USA: ACM, 2013, pp. 1499–1508.

[137] L. E. Nacke. “Games User Research and Physiological Game Evaluation”. In: *Game User Experience Evaluation*. Ed. by R. Bernhaupt. Cham: Springer International Publishing, 2015, pp. 63–86.

[138] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert. “Using psychophysiological techniques to measure user experience with entertainment technologies”. In: *Behaviour & information technology* 25.2 (2006), pp. 141–158.

[139] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun. “Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty”. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 41.6 (Nov. 2011), pp. 1052–1063.

[140] M. K. X. J. Pan, G. J.-S. Chang, G. H. Himmetoglu, A. Moon, T. W. Hazelton, K. E. MacLean, and E. A. Croft. “Galvanic Skin Response-derived Bookmarking of an Audio Stream”. In: *CHI ’11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’11. New York, NY, USA: ACM, 2011, pp. 1135–1140.

[141] C. Bateman and L. E. Nacke. “The neurobiology of play”. In: *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. ACM, May 2010, pp. 1–8.

[142] J. Laaksolahti, K. Isbister, and K. Höök. “Using the *Sensual Evaluation Instrument*”. In: *Digital Creativity* 20.3 (2009), pp. 165–175.

[143] K. Isbister, K. Höök, M. Sharp, and L. Jarmo. “The Sensual Evaluation Instrument: Developing an Affective Evaluation Tool”. In: *CHI 2006 Proceedings*. 2006, pp. 1163–1172.

[144] LANG and P. J. “International affective picture system (IAPS) : Digitized photographs, instruction manual and affective ratings”. In: *Technical report: NAVTRADEVCEN. Naval Training Device Center* (2005).

[145] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. “Measuring and Defining the Experience of Immersion in Games”. In: *International Journal Human-Computer Studies* 66.9 (Sept. 2008), pp. 641–661.

[146] M. Mateas and A. Stern. “Writing facade: A case study in procedural authorship”. In: *Second Person: Role-Playing and Story in Games and Playable Media* (2007), pp. 183–208.

[147] *The Walking Dead series*. Telltale Games, 2012.

[148] *Sam and Max Save the World*. Telltale Games, 2006.

[149] N. Wardrip-Fruin. "Reading Digital Literature: Surface, Data, Interaction, and Expressive Processing". In: *A Companion to Digital Literary Studies* (2013).

[150] M. Mateas and A. Stern. "Procedural authorship: A case-study of the interactive drama facade". In: *DIGITAL ARTS AND CULTURE*. 2005.

[151] T. Berners-Lee, J. Hendler, and O. Lassila. "The Semantic Web". In: *Scientific American*. Lecture Notes in Computer Science 284.5 (2001). Ed. by K. Aberer, K.-S. Choi, N. Noy, D. Allemang, K.-I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, and P. Cudré-Mauroux, pp. 34–43.

[152] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, and Y. Katz. "Pellet: A practical OWL-DL reasoner". In: *Web Semantics: Science, Services and Agents on the World Wide Web* 5.2 (June 2007), pp. 51–53.

[153] R. Shearer, B. Motik, and I. Horrocks. "HermiT: A Highly-Efficient OWL Reasoner". In: *OWLED*. Vol. 432. 2008, p. 91.

[154] A. Gangemi and V. Presutti. "The bourne identity of a web resource". In: *Proceedings of Identity Reference and the Web Workshop (IRW) at the WWW Conference* (2006).

[155] O. Lassila, R. R. Swick, W. Wide, and Web Consortium. *Resource Description Framework (RDF) Model and Syntax Specification*. 1998. `https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/`.

[156] M. A. Musen. "The Protégé project: A look back and a look forward". In: *AI Matters*. 1.4 (June 2015).

[157] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. "SWRL: A semantic web rule language combining OWL and RuleML". In: *W3C member submission* 21 (2003).

[158] D. R. Winer, J. P. Magliano, J. A. Clinton, A. Osterby, T. Ackerman, and R. Michael Young. "A Specialized Corpus for Film Understanding". en. In: *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*. Sept. 2017.

[159] P. Mawhorter, M. Mateas, N. Wardrip-Fruin, and A. Jhala. "Towards a Theory of Choice Poetics". In: *Proceedings of the International Conference on Foundations of Digital Games*. 2014.

[160] A. Mitchell. "Reflective Rereading and the SimCity Effect in Interactive Stories". In: 9445 (2015), pp. 27–39.

[161] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, J. A. Richard Diebold, M. Durbin, M. S. Edmonson, J. L. Fischer, D. Hymes, S. T. Kimball, W. L. Barre, S. J. Frank Lynch, J. E. McClellan, D. S. Marshall, G. B. Milner, H. B. Sarles, G. L. Trager, and A. P. Vayda. "Proxemics [and Comments and Replies]". In: *Current anthropology* 9.2/3 (1968), pp. 83–108. eprint: `https://doi.org/10.1086/200975`.

[162] N. Yee. "Motivations for play in online games". en. In: *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society* 9.6 (Dec. 2006), pp. 772–775.

[163] E. Chang. *Close Playing, a Meditation on Teaching (with) Video Games*. `https://www.hastac.org/blogs/changed/2010/11/12/close-playing-meditation-teaching-video-games`. Accessed: 2018-4-30. Nov. 2010.

[164] Y. Tanahashi and K. L. Ma. "Design considerations for optimizing storyline visualizations". In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2679–2688.

[165] S. Sali and M. Mateas. "Using Information Visualization to Understand Interactive Narrative: A Case Study on Façade". In: *ICIDS*. Springer, 2011, pp. 284–289.

[166] H. Bunt, M. Kipp, and V. Petukhova. "Using DiAML and ANVIL for multimodal dialogue annotation". In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. 2012.

[167] J. H. Kim, D. V. Gunn, E. Schuh, B. Phillips, R. J. Pagulayan, and D. Wixon. "Tracking Real-time User Experience (TRUE): A Comprehensive Instrumentation Solution for Complex Systems". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. New York, NY, USA: ACM, 2008, pp. 443–452.

[168] B. Medler, M. John, and J. Lane. "Data cracker: developing a visual game analytic tool for analyzing online gameplay". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. dl.acm.org, 2011, pp. 2365–2374.

[169] L. Manovich. "Cultural analytics: visualising cultural patterns in the era of 'more media'". In: *DOMUS* (2009).

[170] L. Manovich, J. Douglass, and T. Zepel. "How to Compare One Million Images?" In: *Understanding Digital Humanities*. Ed. by D. M. Berry. London: Palgrave Macmillan UK, 2012, pp. 249–278.

[171] J. Drucker. "Humanities approaches to graphical display". In: *Digital Humanities Quarterly* 5.1 (2011), pp. 1–21.

[172] A. Jena. "Apache Jena Fuseki". In: *The Apache Software Foundation* (2014).

[173] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2017. URL: https://www.R-project.org/.

[174] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *shiny: Web Application Framework for R*. R package version 1.0.5. 2017. https://CRAN.R-project.org/package=shiny.

[175] Willem Robert van Hage, T. Kauppinen, B. Graeler, C. Davis, J. Hoeksema, A. Ruttenberg, and D. Bahls. *SPARQL: SPARQL client*. R package version 1.16. 2013. https://CRAN.R-project.org/package=SPARQL.

[176] M. Franz, C. T. Lopes, G. Huck, Y. Dong, O. Sumer, and G. D. Bader. "Cytoscape.js: a graph theory library for visualisation and analysis". en. In: *Bioinformatics* 32.2 (Jan. 2016), pp. 309–311.

[177] T. Jacob and M. Otto. *Bootstrap*. June 2016.

[178] J. J. Thomas and K. A. Cook. "A visual analytics agenda". en. In: *IEEE computer graphics and applications* 26.1 (Jan. 2006), pp. 10–13.

[179] B. Shneiderman and C. Plaisant. *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Pearson Addison Wesley, 2004.

[180] M. Mateas. "Interactive Drama, Art and Artificial Intelligence". PhD thesis. Carnegie Mellon University, 2002.

[181] S. Moulthrop, D. Grigar, and J. Tabbi. *Traversals: The Use of Preservation for Early Electronic Writing*. The MIT Press, 2017.

[182] P. Mirza-Babaei, G. Wallner, G. McAllister, and L. E. Nacke. "Unified Visualization of Quantitative and Qualitative Playtesting Data". In: *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. CHI EA '14. New York, NY, USA: ACM, 2014, pp. 1363–1368.