Title

How to Describe What? Towards a Theory of Modality Utilization

Permalink

Journal

Authors

Arens, Yigal
Hovy, Eduard

Publication Date

1990

Peer reviewed

How to Describe What?
Towards a Theory of Modality Utilization

Yigal Arens and Eduard Hovy*
USC/ISI
ARENS@ISI.EDU, HOVY@ISI.EDU

In this paper we outline the first steps of an investigation of the nature of representations of information, an investigation that uses as a starting point the various ways in which people tend to communicate different kinds of information. Our hope is that by identifying the regularities of presentation, in particular by finding out when people decide to switch presentation modalities and what they then tend to do, we will be able to shed light on the nature of the underlying representations and processes of communication between people.

1 Introduction: The Use of Multiple Modalities

In extended discussions of a technical nature, there invariably comes a point when someone reaches for a pen and draws a diagram or figure. Why? When and why does language, which is after all the most powerful means of communication available to humankind, fall short in expressive power? What additional features do other modalities of communication have? How do our cognitive abilities manage all the disparate kinds of information, splitting them apart during communication and allocating them to various modalities, and then integrating them again? Why is translating to another "visualization" not a simple process — for example, why do some problems seem unsolvable when presented in one way (say, in language) and straightforward when presented in another (say, diagrammatically)? What does it mean for some people to be more "visual" than others?

Questions such as these are interesting to the cognitive scientist because they may shed light both on our internal representations and on our manipulations of them, as reflected by the natures of the modalities we have developed to communicate ideas. This is a deep issue: humans *need* multiple media when they communicate. You cannot speak normally if restricted in hand gesture and facial expression. It is a rare nonfiction book that does not contain photographs, illustrations, or charts. Speakers in most workshops would consider themselves severely handicapped if denied use of overhead or slide projectors. And so forth.

The problem of display design has not yet been given a thorough computational analysis elsewhere. There does exist a general theory of graphical presentation, aimed at the human practitioner ([Bretin 83]), as well as computational treatments of certain subclasses of presentations ([Mackinlay 86, Feiner 88]).

While we do not pretend to have a theory to explain the phenomena, we believe that a careful study of the types of modalities people use, and the types of information they typically utilize them for, will single out characteristics of the underlying cognitive representations and shed light on people's communicative processes. With these issues in mind, initiating a study of the characteristics of representation as expressed through communication, we decided to examine first two aspects:

- communication-related characteristics of information
- modes of human-human and human-computer communication

We decided to take into account modes that are used in interactions with computers as well, in order eventually to test the rules we develop and implement on a computer against the display decisions made by people.

When identifying characteristics salient to the display of information, the vocabulary should:
- describe all features of the information that are salient for presentation purposes,
- describe all features of presentation modalities that can be utilized to convey information,
- be general enough to allow comparisons and specific enough to differentiate between different modalities and information.

We first define some useful terms, and then provide characterizations of media and information. The paper ends with an example.

2 Characterization of Modalities

2.1 Definition of Terms

The following terms are used to describe presentation-related concepts. We take the point of view of the communicator (indicating where the consumer's subjective experience may differ).

1. Consumer: A person interpreting a communication.

2. Modality: A single mechanism by which to express information. Examples: spoken and written natural language, diagrams, sketches, graphs, tables, pictures.

3. Exhibit: A complex exhibit is a collection, or composition, of several simple exhibits. A simple exhibit is what is produced by one invocation of one modality. Examples of simple exhibits are a paragraph of text, a diagram, a computer beep. Simple exhibits involve the placement of one or more **Information Carriers** on a background **Substrate**.

4. Carried Item: That piece of information represented by the carrier; the 'denotation' of the carrier.

For purposes of rigor, it is important to note that a substrate is simply one or more information carrier(s) superimposed. This is because the substrate carries information as well[1]. In addition, in many cases the substrate provides an internal system of semantics which may be utilized by the carrier to convey information. Thus, despite its name, not all information is transmitted by the carrier itself alone; its positioning (temporal or spatial) in relation to the substrate may encode information as well. This is discussed further below.

5. Channel: An independent dimension of variation of a particular information carrier in a particular substrate. The total number of channels gives the total number of independent pieces of information the carrier can convey. For example, a single mark or icon can convey information by its shape, color, and position and orientation in relation to a background map. The number and nature of the channels depend on the type of the carrier and on the exhibit's substrate.

2.2 Internal Semantic Systems

Some information carriers exhibit an internal structure that can be assigned a 'real-world' denotation, enabling them subsequently to be used as substrates against which other carriers can acquire

---

[1]Note that from the information consumer's point of view, Carrier and Substrate are subjective terms; two people looking at the same exhibit can interpret its components as carrier and substrate in different ways, depending on what they already know.

information by virtue of being interpreted within the substrate. For example, a map used to describe a region of the world possesses an internal structure — points on it correspond to points in the region it charts. When used as a background for a ship icon, one may indicate the location of the ship in the world by placing its icon in the corresponding location on the map substrate. Examples of such carriers and their internal semantic systems are:

Carrier	Internal Semantic System
Picture	'real-world' spatial location based on picture denotation
NL sentence	'real-world' sentence denotation
Table	categorization according to row and column
Graph	coordinate values on graph axes
Map	'real-world' spatial location based on map denotation
Ordered list	ordinal sequentiality

Other information carriers exhibit no internal structure. Examples: icon, computer beep, and unordered list.

An internal semantic system of the type described is always intrinsic to the item carried.

2.3 Characteristics of Modalities

In addition to the internal semantics listed above, modalities differ in a number of other ways which can be exploited by a presenter to communicate effectively and efficiently. The values of these characteristics for various modalities are shown in Table 1.

Carrier Dimension: Values: *0D, 1D, 2D*. A measure of the number of dimensions usually required to exhibit the information presented by the modality.

Internal Semantic Dimension: Values: *0D, 1D, 2D, >2D, 3D, #D, ∞D*. The number of dimensions present in the internal semantic system of the carrier or substrate.

Temporal Endurance: Values: *permanent, transient*. An indication whether the created exhibit varies during the lifetime of the presentation.

Granularity: Values: *continuous, discrete*. An indication of whether arbitrarily small variations along any dimension of presentation have meaning in the denotation or not.

Medium Type: Values: *aural, visual*. What type of medium is necessary for presenting the created exhibit.

Default Detectability: Values: *low, medlow, medhigh, high*. A default measure of how intrusive to the consumer the exhibit created by the modality will be.

Baggage: Values: *low, high*. A gross measure of the amount of extra information a consumer must process in order to become familiar enough with the substrate to correctly interpret a carrier on it.

2.4 How Carriers Convey Information

As part of an exhibit, a carrier can convey information along one or more **channels**. For example, with an icon carrier, one may convey information by the icon's shape, color, and possibly through its position in relation to a background map. The number and nature of the channels depends on the type of carrier and the substrate.

The semantics of a channel may be *derived* from the carrier's spatial or temporal relation to a substrate which possesses an internal semantic structure; e.g., placement on a map of a carrier representing an object which exists in the charted area. Otherwise we say the channels is *free*.

Generic Modality	Carrier Dimension	Int. Semantic Dim.	Temporal Endurance	Granularity	Medium Type	Default Detectability	Baggage
Beep	0D		transient	N/A	aural	high	
Icon	0D		permanent	N/A	visual	low	
Map	2D	2D	permanent	continuous	visual	low	high
Picture	2D	3D	permanent	continuous	visual	low	high
Table	2D	2D	permanent	discrete	visual	low	high
Form	2D	>2D	permanent	discrete	visual	low	high
Graph	2D	2D	permanent	continuous	visual	low	high
Ordered list	1D	#D	permanent	discrete	visual	low	low
Unordered list	0D	#D	permanent	N/A	visual	low	low
Written sentence	1D	∞D	permanent	discrete	visual	low	low
Spoken sentence	1D	∞D	transient	discrete	aural	medhigh	low
Animated material	2D	3D	transient	continuous	visual	high	high
Music	1D	?	transient	continuous	aural	med	low

Table 1: Modality characteristics.

Among *free* channels we distinguish between those whose interpretation is *independent* of the carried item (e.g., color, if the carrier does not represent an object for which color is relevant); and those whose interpretation is *dependent* on the carried item (e.g., shape, if the carrier represents an object which has some shape).

Most of the carrier channels can be made to vary their presented value in time. Time variation can be seen as an additional channel which provides yet another degree of freedom of presentation to most of the other channels. The most basic variation is the alternation between two states, in other words, a flip-flop, because this guarantees the continued (though intermittent) presentation of the original basic channel value.

3   Characterization of Information and Its Presentation

In this section we develop a vocabulary of presentation-related characteristics of information.

Broadly speaking, as shown in Table 2, three subcases must be considered when choosing a presentation for an item of information: intrinsic properties of the specific item; properties associated with the class to which the item belongs; and properties of the collection of items that will eventually be presented, and of which the current item is a member. These characteristics are explained in the remainder of this section.

Dimensionality: Some single items of information, such as a data base record, can be decomposed as a vector of simple components; others, such as a photograph, have a complex internal structure which is not decomposable. We define the *dimensionality* of the latter as *complex*, and of the former as the dimension of the vector.

Type	Characteristic	Values
Intrinsic Property	Dimensionality	0D, 1D, 2D, >2D, ∞D
	Transience	live, dead
	Urgency	urgent, routine
Class Property	Order	ordered, nominal
	Density	dense, discrete, N/A
Set Property	Volume	singular, little, much

Table 2: Information characteristics by type.

Since all the information must be represented in some fashion, the following must hold (where *simple* dimensionality has a value of 0, *single* the value 1, and so on, and *complex* the value ∞):

The Basic Dimensionality Rule of Presentations

□ *Rule:* Dim(Info) ≤ Dim(Carrier) + Free Channels(Carrier) + Internal Semantic Dim(Substrate)

In addition, we have found that different rules apply to information of differing dimensions. With respect to dimensionality, we divide information into four classes as follows:

- *Simple:* Simple atomic items of information, such as an indication of the presence or absence of email.

  □ *Rule:* As carrier, use a modality with a dimension value of 0D.
  □ *Rule:* No special restrictions on substrate.

- *Single:* The value of some meter such as the amount of gasoline left. Associated rule is:

  □ *Rule:* No special restrictions on substrate.

- *Double:* Pairs of information components, such as coordinates (graphs, map locations), or domain-range pairs in relations (automobile × satisfaction rating, etc.).

  □ *Rule:* As substrate, use modalities with internal semantic dimension of 2D.
  □ *Rule:* As substrate, use modalities with discrete granularity (e.g., forms and tables) if information-class of both components is discrete.
  □ *Rule:* As substrate, use modalities with continuous granularity (e.g., graphs and maps) if information-class of either component is dense.
  □ *Rule:* As carrier, use a modality with a dimension value of 0D.

- *Multiple:* More complex information structures of higher dimension, such as home addresses. It is assumed that information of this type requires more time to consume (hence the last rule in this group).

  □ *Rule:* As substrate, use modalities with discrete granularity if information-class of all components is discrete.
  □ *Rule:* As substrate, use modalities with continuous granularity if the information-class of some component is dense.
  □ *Rule:* As carrier, use a modality with a dimension value of at least 1D.
  □ *Rule:* As substrate and carrier, do not use modalities with the temporal endurance value transient.

491

- *Complex:* Information with internal structure that is not decomposable, such as photographs.
  - □ *Rule:* Check for the existence of specialized modalities for this class of information.

Transience: Transience refers to whether the information to be presented expresses some current (and presumably changing) state or not. Presentations differ according to:

- *Live:* The information presented consists of a single conceptual item of information (that is, one carried item) that varies with time (or in general, along some linear, ordered, dimension), and for which the history of values is not important. Examples are the amount of money owed while pumping gasoline or the load average on a computer. Most appropriate for *live* information is a single exhibit.
  - □ *Rule:* As carrier, use a modality with the temporal endurance characteristic transient if the update rate is comparable to the lifetime of the carrier signal.
  - □ *Rule:* As carrier, use a modality with the temporal endurance characteristic permanent if update rate is much longer.
  - □ *Rule:* As substrate, unless the information is already part of an existing exhibit, use the neutral substrate.
- *Dead:* The other case, in which information does not reflect some current state, or in which it does but the history of values is important. An example is the history of some stock on the stock market; though only the current price may be important to a trader, the history of the stock is of import to the buyer.
  - □ *Rule:* As carrier, use ones that are marked with the value permanent temporal endurance.

Urgency: Some information may be designated *urgent*, requiring presentation in such a way that the consumer's attention is drawn. This characteristic takes the values *urgent* and *routine*:

- *Urgent:* This situation is exemplified in emergencies, whether they be imminent meltdowns or a warning to a person crossing the road in front of a car. Rules of modality allocation are:
  - □ *Rule:* If the information is not yet part of a presentation instance, use a modality whose default detectability has the value high (such as an aural modality) either for the substrate or the carrier.
  - □ *Rule:* If the information is already displayed as part of a presentation instance, use the present modality but switch one or more of its channels from fixed to the corresponding temporally varying state (such as flashing, pulsating, or hopping).
- *Routine:* The normal case.
  - □ *Rule:* Choose a modality with low default detectability and a channel with no temporal variance.

Density: The difference between information that is presented equally well on a graph and a histogram and information that is not well presented on a histogram is a matter of the density of the class to which the information belongs. The former case is *discrete* information; an example is the various types of car made in Japan. The latter is *dense* information; an example is the prices of cars made in Japan.

- *Dense:* A class in which arbitrary small variations along a dimension of interest carry meaning. Information in such a class is best presented by a modality that supports continuous change:

  □ *Rule:* As substrate, use a modality with granularity characteristic continuous (e.g., graphs, maps, animations).

- *Discrete:* A class in which there exists a lower limit to variations on the dimension of interest. Appropriate modalities are as follows:

  □ *Rule:* As substrate, use a modality with granularity characteristic discrete (e.g., tables, histograms, lists).

Volume: A batch of information may contain various amounts of information to be presented. If it is a single fact, we call it *singular*; if more than one fact but still little relative to some some task- and user-specific threshold, we call it *little*; and if not, we call it *much*. This distinction is useful because not all modalities are suited to present *much* information.

- *Much:* The relatively permanent modalities such as written text or graphics leave a trace to which the consumer can refer if he or she gets lost doing the task or forgets, while transient modalities such as spoken sentences and beeps do not. Thus the former should be preferred in this case.

  □ *Rule:* As carrier, do not use a modality the temporal endurance value transient.
  □ *Rule:* As substrate, do not use a modality the temporal endurance value transient.

- *Little:* There is no need to avoid the more transient modalities when the amount of information to present is *little*.
- *Singular:* A single atomic item of information. A transient modality can be used. However, one should not overwhelm the consumer with irrelevant information. For example, to display information about a single ship, one need not draw a map.

  □ *Rule:* As substrate, if possible use a modality whose internal semantic system has *low* baggage.

4 An Example

We present three simple tasks in parallel.

Given: the task of presenting Paris (as the destination of a flight, say).

Available information (three separate examples): the coordinates of the city, the name *Paris*, and a photograph of the Eiffel Tower.

Available modalities: maps, spoken and written language, pictures, tables, graphs, ordered lists.

The modality characteristics are listed among those in Table 1. The information characteristics are listed in Table 3.

The allocation algorithm classifies information characteristics with respect to characteristics of modalities, according to the rules outlined in Section 3. The modality with the most desired characteristics is then chosen to form the exhibit.

Handling the coordinates: As given by the rules mentioned in Section3, information with a *dimensionality* value of *double* is best presented in a substrate with a *dimension* value of *2D*. This means that candidate substrates for the exhibit are maps, pictures, tables, and graphs. Since the *volume* is *little*, *transient* modalities are not ruled out. The value *dense* for the characteristic

	Coordinates	Name	Photograph
Information	48N 2E	Paris	Eiffel Tower
Dimensionality	*double*	*single*	*single*
Volume	*little*	*singular*	*singular*
Density	*dense*	*discrete*	*discrete*
Transience	*dead*	*dead*	*dead*
Urgency	*routine*	*routine*	*routine*

Table 3: Example information characteristics.

density rules out tables. The values for *transience* and *urgency* have no further effect. This leaves tables, maps, and graphs as possible modalities. Next, taking into account the rules dealing with the internal semantics of modalities, immediately everything but maps are ruled out (maps' internal semantics denote spatial locations, which matches up with the denotation of the coordinates). If no other information is present, a map modality is selected to display the location of Paris.

Handling the name: The name Paris, being an atomic entity, has the value *single* for the *dimensionality* characteristic. By the appropriate rule (see Section 3), the substrate should be the neutral substrate or natural language and the carrier one with *dimension* of *0D*. Since the *volume* is *singular*, a *transient* modality is not ruled out. None of the other characteristics have any effect, leaving the possibility of communicating the single word Paris or of speaking or writing a sentence such as *"The destination is Paris"*.

Handling the photograph: The photograph has a *dimensionality* value *complex*, for which appropriate rules specify modalities with *internal semantic dimension* of *3D*, and with *density* of *dense* (see Section 3) — animation or pictures. Since no other characteristic plays a role, the photograph can simply be presented.

5 Conclusion

We realize full well that this paper does not present an actual cognitive theory of how people represent and communicate information to highlight various characteristics. However, based on the pervasiveness and regularities in the use of multiple modalities in communication, we believe that any adequate cognitive theory will have to include the types of considerations and characteristics we discuss. That is to say, we believe that notions such as *dimensionality, urgency*, and *granularity* have an irrefutable cognitive reality, thanks to their essential role in the cognitive process of interhuman communication. Future refinements of these terms and identifications of others will help to uncover some of the ways in which people represent and manipulate various types of information.

References

[Bretin 83] Bretin, J. *Semiology of Graphics*, trans. by J. Berg. University of Wisconsin Press, 1983.

[Feiner 88] Feiner, S. An Architecture for Knowledge-Based Graphical Interfaces. ACM/SIGCHI workshop on *Architectures for Intelligent Interfaces: Elements and Prototypes*, Monterey, CA, 1988.

[Hovy & Arens 90] Hovy, E. & Y. Arens. When is a Picture Worth a Thousand Words?—Allocation of Modalities in Multimedia Communication. *AAAI Symp. on Human-Computer Interaction*, Stanford U., 1990.

[Mackinlay 86] Mackinlay, J. *Automatic Design of Graphical Presentations*. Ph.D. dissertation, Department of Computer Science, Stanford University, 1986.