Title
The Effect of Syntactic Form on Simple Belief Revisions and Updates

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 16(0)

Authors
Elio, Renee
Pelletier, Francis Jeffry

Publication Date
1994

Peer reviewed

The Effect of Syntactic Form on Simple Belief Revisions and Updates

Renée Elio
Department of Computing Science
University of Alberta
Edmonton, Alberta T6G 2H1
ree@cs.ualberta.ca

Francis Jeffry Pelletier
Departments of Philosophy and
Computing Science
University of Alberta
Edmonton, Alberta T6G 2H1
jeffp@cs.ualberta.ca

Abstract

In this paper we report preliminary results on how people revise or update a previously held set of beliefs. When intelligent agents learn new things which conflict with their current belief set, they must *revise* their belief set. When the new information does not conflict, they merely must *update* their belief set. Various AI theories have been proposed to achieve these processes. There are two general dimensions along which these theories differ: whether they are syntactic-based or model-based, and what constitutes a minimal change of beliefs. This study investigates how people update and revise semantically equivalent but syntactically distinct belief sets, both in symbolic-logic problems and in quasi-real-world problems. Results indicate that syntactic form affects belief revision choices. In addition, for the symbolic problems, subjects update and revise semantically-equivalent belief sets identically, whereas for the quasi-real-world problems they both update and revise differently. Further, contrary to earlier studies, subjects are sometimes reluctant to accept that a sentence changes from false to true, but they are willing to accept that it would change from true to false.

Introduction

As epistemic agents, both people (and intelligent machines) hold a set of beliefs about the world, they learn new things about the world, and sometimes come to recognize that new information conflicts with or extends their existing belief set. When new information creates an inconsistency with prior beliefs, rational agents would do some *revision* of their belief state. This revision involves identifying which of the old and new beliefs clash to create the inconsistency, deciding whether in fact to accept the new information, and, if that is the choice, to eliminate certain old beliefs in favor of the new information. Alternatively, new information may not create any inconsistency with old information at all. In this case, the agent would do a belief state *update*—adding new information to the current set of beliefs, along with whatever additional alterations this might entail.

Although this is an intuitively attractive picture, the "logic" behind the belief revision picture is neither well-understood nor agreed-upon. Belief revision as portrayed above is an important problem from several perspectives. For example, one area of AI research concerns the development of knowledge bases as a kind of intelligent database: one enters information into the knowledge base and the knowledge base itself constructs and stores the consequences of this information, a process which is non-monotonic in nature (i.e., consequences of previously believed information are abandoned). Not only is this seen as the future of traditional databases, but any computer system that calls for dynamic decision making can be viewed as requiring the ability to revise its beliefs concerning the state of the world. More generally, belief change—the process by which a rational agent makes the transition from one belief state to another—is an important component for most intelligent activity and as a psychological activity it bears on issues in science, law, business decision theory, artificial intelligence, and many other disciplines (Yates, 1990; Katsuno & Mendelson, 1991; Gärdenfors & Rott, 1992).

From the current AI perspective, a belief state is represented by a set of sentences, taken to be expressed in the propositional logic. In this language, sentence letters (p, q, r, etc.) represent simple, atomic sentences; more complex sentences are formed by using the logical connectives $\neg$, &, $\vee$, $\rightarrow$, $\leftrightarrow$ (respectively, 'not', 'and', (inclusive) 'or', 'if–then–', 'if and only if'). If a sentence is in the belief set, then it is believed and its negation is disbelieved. If the negation of a sentence is in the belief set, then the negation is believed and the sentence is disbelieved. If neither the sentence nor its negation is in the belief set, then they are uncertain—neither believed nor disbelieved.

The basic idea behind the AI theories of belief revision is that one tries to maintain as much as possible of the earlier belief state while nonetheless accommodating the new information. The initial work in this area was done by Alchourrón, Gärdenfors, & Makinson (1985) and since that time, various AI belief revision theories have emerged as alternative proposals about how belief revision should proceed. These theories diverge on two general issues. The first concerns what constitutes a "minimal change" to the initial belief set (reviewed in Katsuno & Mendelson, 1991). The second is more basic and has more profound implications: the distinction between the so-called syntax-based approaches and the semantic (or model-based) approaches. In a syntax-based approach, the belief set contains just those sentences that are explicitly believed, and

Original KB	Update	Revision Alternatives								
MP: p→q, p, q	~q	1. p→q	~p	~q	2. ~(p→q)	~q	?p	3. ~(p→q)	p	~q
MT: p→q, ~p, ~q	p	1. p→q	p	q	2. ~(p→q)	p	?q	3. ~(p→q)	p	~q
BCR: p↔q	p	1. p↔q	p	q	2. ~ (p↔q)	p	?q	3. ? (p↔q)	p	?q
BCD: (p & q) ∨ (~p & ~q)	p	1. p↔q	p	q	2. ~ (p↔q)	p	?q	3. ? (p↔q)	p	?q

Notes: MP = modus ponens; MT= modus tollens; BCR = biconditional rule form; BCD=biconditional disjunctive form; ? = undecided about

Figure 1: Belief Update Problem Set

contains them in exactly the form in which they are believed. It does not contain logical consequences of what is believed (unless they too are explicitly believed), nor does it necessarily contain any logically equivalent version of the explicitly-believed sentences. Thus, exactly what syntactic form the belief takes is important, because that in turn impacts what constitutes a "minimal change" during update. The semantic-based approach defines a belief state to be what is explicitly believed together with all its logical consequences. That is, the semantic approach talks about the *world* that is described by the beliefs and to what *the believer is logically committed*. It therefore does not matter what precise syntactic form any particular belief sentence takes, for if one form is in the set then all its logical equivalents are too. It seems clear that the semantic approach involves a certain amount of theoretical idealization that is not present in the syntactic approach. Yet its basic tenet—that the result of an update must be independent of how the original knowledge base was stated, as well as independent of the syntax of the new information itself—has been vigorously defended (Dalal, 1988; Yates, 1990; Katsuno & Mendelson, 1991). All that is relevant is how the world is, or would be, if the beliefs were true.

The alternative AI theories of belief revision present a starting set of proposals concerning the belief revision principles a rational agent should follow. Our general goal to investigate these and other alternative prescriptions of belief update and revision in people. However, the syntactic-vs.-model based distinction[1] is so fundamental to defining what the old and new belief states actually are, that any investigation of belief revision must include as its starting point whether and to what extent the syntactic form of the belief sets will affect revision and update. Clearly, people are finite creatures and their attention span and logical abilities are in fact quite limited; so the semantic approach with its inclusion of all logical consequences in the belief state might seem to be completely unreasonable as an account of how people *actually* behave. Still though, it could turn out to be a useful model in some hybrid form (e.g., "mental models," Johnson-Laird & Byrne, 1990); this is a matter yet to be investigated. In studying this issue, we wish to know whether people are insensitive to the syntax *of what they are attending* (as opposed to all syntax) in the context of integrating new and old information. Alternatively, if they *are* sensitive to syntax, how does is this manifest in belief-revision choices?

In this paper, we describe a portion of a larger study on belief revision that addresses the issue of syntactic form. Both our study and most AI belief revision theories presume that the new information *is* to be incorporated. There is a huge literature indicating that people are in general very reluctant to change their current belief sets; and that they are much more likely to reject, ignore, or reinterpret the new information which conflicts with their current beliefs rather than attempt to add it to their beliefs and make the necessary adjustments (Edwards, 1968; Einhorn & Hogarth, 1978; Ross & Lepper, 1980; Hoenkamp, 1988). Although it is true that people are reluctant to engage in the revision, our investigations take as a starting point that any inertia against changing a belief set has been overcome.

The Problem Set

Figure 1 gives the schematic versions of the four problems we investigated: modus ponens, modus tollens, and two logically-equivalent forms of a biconditional. There has been considerable research investigating logical reasoning involving these types of problems (for surveys, see Rips, 1990; Johnson-Laird & Bryne, 1991]. However, these investigations were not aimed at determining the effect of different syntactic forms given the task of belief revision—and it is that condition we wish to consider. So to be clear, our aim in this is *not* to compare reasoning with modus ponens versus modus tollens in the consistent case. That has already been studied thoroughly. Nor is it our aim to argue directly for a semantic-based (model-based) account of beliefs vs. a syntactic-based account of beliefs in the abstract. (It seems that the simple fact that the robustness with which

[1] A related distinction is the foundationalist vs. coherentist distinction. See Furhman (1991) and Nebel 1991).

Original Knowledge Base:		S's have A's if and only if S's have B's. (This means both these things are true at the same time: If an S has A, then an S has B. If an S has B, then an S has A.)
Update:		Lex is an S. Lex has A.
Revision #2	Believe Disbelieve Undecided about	Lex is an S. Lex has A. S's have A's if and only if S's have B's. Whether Lex has B.

Figure 2: Biconditional Rule (BCR) Problem, Symbolic Condition

modus tollens is so much more difficult for people to perform than is modus ponens shows that a full-fledged theory of model-based belief states cannot hold for people.)

Our first two problems presented new information that created an inconsistency with (a) a modus ponens knowledge base, and (b) a modus tollens knowledge base. Modus ponens and modus tollens are just two different sides of the same coin when considered strictly from the semantic or model-theoretic point of view: they differ only in their syntactic expression. The second two problems involve (a) a biconditional $(p \leftrightarrow q)$ expressed as "if and only if" and (b) a biconditional expressed as its equivalent disjunction. In these two problems, the new information does not introduce an inconsistency and hence does not require a revision to the original belief set, but rather constitutes an update to the belief set. Since the two forms of the biconditional are logically equivalent, there should be no difference in the updates if the semantic point of view is correct.

Each problem consisted of an original knowledge base or belief set, an update, and then three alternative theories. The task for subjects was to chose one of the alternative theories as their preferred revised/updated belief set. (For this study, all alternative revision or update choices require acceptance of the update information; this was necessary given the broader intent of the study. Current studies relax this constraint in the choices provided to subjects.)

Method

Stimulus Set

Figure 2 illustrates more closely the actual content of the problems that subjects solved, showing the biconditional-rule form problem presented under what we called our symbolic condition (described in greater detail below), with one of the possible revision choices.

In the problems that subjects solved, the original knowledge base was labeled as "the well-established knowledge at time 1." The update information was introduced with the phrase, "By time 2, knowledge had increased to include the following." Each alternative belief-set revision was called a "theory" and consisted of statements labeled "Believe", "Disbelieve" or "Undecided About." A theory could have statements of all these types, or of just some of these types.

Since that people find modus tollens more difficult to perform than modus ponens, these problems included the correct conclusions already drawn in the initial belief set (see Fig. 1). Hence, subjects were not called upon to perform these inferences when constructing the original knowledge base which must then be revised. For the biconditional rule-form problem, subjects were also told (as part of the original knowledge base) what "if and only if" means, namely that it means the conjunction of the two conditionals (see Fig. 2).

For the modus ponens and modus tollens problems, we *explicitly indicated* that the update knowledge conflicted with the original knowledge. For the two biconditional problems, the problem stated, "The increased knowledge at time 2 does not conflict with the initial knowledge held at time 1. However, there are alternative ways to update the initial knowledge with the increased knowledge."

Design

There was one between-subjects factor, namely the presentation form of the problem: symbolic form vs. science-fiction form. Figure 2 illustrates the symbolic condition, which used letters and nonsense syllables to instantiate the basic problems. We also wanted a problem-presentation form that might mitigate the extent to which subjects shifted into "logical problem solving mode" and

Original Knowledge Base:		If an ancient ruin has a protective force field, then it is inhabited by the aliens called Pylons. The tallest ancient ruin does not have a protective force field. The tallest ancient ruin is not inhabited by Pylons.
Update:		The tallest ancient ruin has a protective force field.
Revision #3	Believe Disbelieve	The tallest ancient ruin has a protective force field. If an ancient ruin has a protective force field, then it's inhabited by the aliens called Pylons. The tallest ancient ruin is inhabited by Pylons.

Figure 3: Modus Tollens Problem, Science Fiction Condition

Table 1: Proportions of Selected Revisions for Modus Ponens and Modus Tollens

	Symbol	SciFi	Symbol	SciFi
	Modus Ponens		Modus Tollens	
Original KB:	p -> q, p, q		p -> q, ~p, ~q	
Update:	~q		p	
	Presentation Condition			
Revision Alternatives	Symbol	SciFi	Symbol	SciFi
1 accept update, apply rule to determine other	.25	.14	.34	.17
2 accept update, deny rule, uncertain about other	.38	.29	.38	.54
3 accept update, deny rule, leave other unchanged	.37	.57	.28	.29

might constitute a more real-world or natural-language context for the belief revision task. However, we did not want to venture very far into real-world knowledge for this initial study. Science-fiction cover stories was our compromise "real-world" condition. Figure 3 shows the content of the modus tollens problem with one of the science-fiction cover stories.

Subjects

One-hundred twenty subjects from the University of Alberta Psychology Department subject pool participated in the study. Equal numbers of subjects were randomly assigned to the symbol or science fiction problem-presentation condition.

Procedure

Problems were presented in random order in booklet form; all subjects received all four problems. The order of alternative theories within the answer section of each problem was counterbalanced across subjects. Different science fiction cover stories were created and appeared in equal numbers across the four problems for subjects in the science-fiction condition. Below are excerpts from the instruction, to help clarify how we presented this task to our subjects:

> "....The first part of the problem gives an initial set of knowledge that was true and well-established at **time 1** (that is, some point in time). There were no mistakes at that time. The second part of the problem presents additional knowledge about the world that has come to light at **time 2** (some later time). This knowledge is also true and well-established.... The world is still the same but what has happened is that knowledge about the world has increased........After the additional knowledge is presented, the problem gives two or more possible "theories" that reconcile the initial knowledge and the additional knowledge.
>
> ....Your task is to consider the time 1 and time 2 knowledge, and then select the theory that you think is the correct way to reconcile all the knowledge....The time 2 information should not necessary just "replace" the time 1 knowledge. It is not more reliable or trustworthy than the time 1 facts. On the other hand, time 1 knowledge is not necessarily the information to keep. You want to find the best balance between the initial information and the new information...."

Results

We converted the subject data into the proportions for which each revision alternative was selected; under this scheme, revision-alternative becomes another factor. Table 1 gives these proportions for the modus ponens and modus tollens problems. In Table 1, we have described the revision alternatives in a more general way: from Figure 1, it is clear that the three belief revision alternatives for modus tollens and modus ponens have a certain symmetry, even though the actual details of each revision are necessarily different. Under revision 1, the rule is applied to revise the truth status of the component not mentioned in the update. In Table 1, we call this unmentioned component "other." For the modus ponens problem, the update mentions p; and q is "the other." For the modus tollens problem, the update mentions q; and p is "the other." Under revision 2 for both problems, the rule is denied, and the status of the other component is uncertain. Under revision 3, the rule is denied and the other component retains whatever truth value it had in the original knowledge base.

For the data presented in Table 1, there was a main effect for revision alternative ($F(2, 234) = 6.35$, $p = .002$), an interaction between presentation mode and revision alternative that approached significance ($F(2, 234) = 2.69$, $p = .06$), a significant interaction between problem type and revision choice ($F(2, 234) = 5.78$, $p = .003$), and a marginally significant interaction between presentation

Table 2: Proportions of Update Choices for Equivalent Forms of Biconditions

			Rule Form			Disjunctive Form		
		Original KB:	(p if and only if q)			(p & q) ∨ (~p & ~q)		
		Update:	p			p		
			Presentation Condition					
Update Alternatives			Symbol	SciFi	Mean	Symbol	SciFi	Mean
1. Believe: biconditional, p, q			.68	.70	.69	.60	.44	.52
2. Believe: p	Disbelieve: biconditional	Uncertain: q	.10	.14	.12	.12	.15	.14
3. Believe: p	Uncertain: biconditional, q		.22	.17	.19	.28	.41	.34

mode, problem type, and revision alternative ($F(2, 234) = 2.91$, $p = .056$).[2] Symbol condition subjects gave the same response pattern across the two logically-equivalent problems and paired comparisons revealed no significant differences in the proportions given either within or across problems.[3] However, the proportion of science-fiction revision-3 choices for modus ponens (0.57) was significantly different from the proportion of revision-1 choices (0.14). For modus tollens, this pattern shifted: the proportion of revision-2 choices (0.54) was significantly higher than the proportion of revision 1 choices (0.17) and also higher than the proportion of revision-1 choices for the modus ponens problem (0.29). Basically, science-fiction subjects elected to change the status of the unmentioned component (*q*) from 'false' to 'uncertain' for modus tollens (a maximal change on the original knowledge base). For modus ponens, the unmentioned component (*p*) was true in the original knowledge base and science-fiction subjects preferred to leave its truth status alone.

Table 2 presents the proportions of update alternatives for the two forms of biconditional. There was a main effect for update alternative: when the new knowledge indicated that *p* was true, most subjects applied the biconditional rule to assert that *q* was also true (alternative 1). This effecy was qualified by an interaction with problem form: the proportion of update 1 choices (applying the biconditional to conclude *q* is also true) was lower in the disjunctive case than in the rule form, while the proportion of update 3 choices (accepting the update but changing the biconditional from 'true' to 'uncertain') was nearly twice as large in the disjunctive form than in the rule form ($F(2.0, 234)=7.51$, $p < .001$). Paired comparisons indicated that the proportion of update-1 choices was significantly lower in the disjunctive case that in the rule-form case (0.52 v. 0.69); the proportion of update-3 choices was higher in the disjunctive case than in the rule case (0.35 vs. 0.19). In addition, the update-3 proportion (0.35) was significantly higher than the update-2 proportion (0.14) under the disjunctive form of problem, but the corresponding means for the rule-form were not significantly different. Essentially, on disjunctive-form problems, subjects shifted from applying the biconditional (i.e., update-1 asserts *q* is true upon learning that *p* is true) to labeling both the biconditional and *q* as uncertain. This interaction is mostly due to a shift in choices in the science-fiction condition, although the impact of this factor was not statistically significant.

Discussion

We have a number of interesting findings. First, for the modus ponens and modus tollens cases, it was interesting that the syntactic form of these logically equivalent problems *did not* impact the choices when presented in their symbolic-form. This is contrary to what one might expect, given prior results that people have more difficulty on modus tollens than with modus ponens (Johnson-Laird & Byrne, 1991). (This may be a feature of the belief-revision process we investigated here versus the deductive problem-solving focus of other studies on these problems.) Second, syntactic form *did* affect revisions choices for our quasi-real world problems: science-fiction subjects chose a maximal change to the initial knowledge base to accommodate the new information on the modus tollens problems, but not on the modus ponens problems. For modus ponens, the truth status of the item unmentioned at update time was retained from the original belief set; for modus tollens, the status of the unmentioned item was altered. In the modus ponens case, the original status of this unmentioned component (*p*) was true and in the modus tollens case, the original status of the unmentioned component (*q*) was false. This may reflect a preference to retain positive (true) *statements* or, in lieu of that, to create something more positive (shifting a negative to an uncertain). But why would this occur under science-fiction stories and not under symbol form? Another interpretation is that, under quasi-real world scenarios, the

[2] These data are interval data, i.e., answers falling into one of three revision choices. ANOVA assumes that data are normally distributed, a questionable assumption in this case. Subsequent log-linear analyses will test specific components of the model indicated by the ANOVA.

[3] Paired-comparisons are based on Neuman-Keuls tests, $p < .05$.

$(p \longrightarrow q)$ relation might be interpreted not as a logical relation, but as something closer to a causal relation between p and q. For the modus tollens case, the original KB makes sense as an explanation of why q would be false (its only apparent cause, p, is false). The update that q's only apparent cause has become true makes suspect the initial "explanation" of q's original falsity. Hence, its status in the world is now uncertain. For modus ponens, one can make a more minimal change by simply denying the "causal" relation between p and q, and retaining p as true. Finding this pattern only for our science-fiction subjects underscores how thinking about real-world relations may bias the kinds of updates people are inclined to do under real-world conditions.

The second important result was that syntax affected belief-set updates for logically equivalent forms of biconditionals. In the rule form $(p \leftrightarrow q)$, all subjects applied the rule upon learning that p is true, adding q to their belief set. But when the relation was stated as a disjunct, the update choices shifted to being uncertain about the biconditional relation and also about the unmentioned component q. The proportions in Table 2 show that this result was due mostly to a shift amongst the science-fiction subjects. Disjuncts are certainly more difficult for subjects to understand (Johnson-Laird & Byrne, 1991) and subjects may incorrectly think that the entire relation (rather than one disjunct or the other) is somehow supposed to hold, particularly in quasi-real world scenarios. Their preference to deny the biconditional could be due to the fact they received no *explicit* information about q (only about p) at update time. If they believe that the entire disjunct must hold, then not receiving explicit evidence about q along with p may seem fishy to them, particularly under a real-world scenario...despite the fact that they have *implicit* information about q. The not-strictly-logical approach to knowledge update and revision found here may be consistent with and extend the notion of pragmatic reasoning schemas, proposed by Cheng et al (1986) for deductive problem solving.

These preliminary results on patterns of belief-revision choices, given these types of problems as the "original knowledge base," constitute an important extension to our understanding of how people reason. They set the stage for understanding the principles by which people modify a *set* of beliefs. More directly, the effects of syntactic form and real-world scenarios set the stage for understanding and perhaps questioning the premise of "minimal change" on a prior set of beliefs as new information becomes available.

Acknowledgments

This research is supported by Canadian NSERC grant OPG A0089 to R. Elio and NSERC grant OPG 5525 to F.J. Pelletier. Pelletier is also a member of the Institute for Robotics and Intelligent Systems and he acknowledges the support of the Networks of Centres of Excellence Program of the Government of Canada and the participation of PRECARN Associates Inc. in this program. We thank our research assistant Sarah Hoffman, and the Department of Psychology, for the use of their subject pool.

References

Alchourrón, C., Gärdenfors, P. & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Jour. Symbolic Logic*, 50, 510-530.

Ashton, A. & Ashton, R. (1988). Sequential belief revision in auditing. *The Accounting Review*, 63, 623-641.

Chend, P. W., Holyoak, K. J., Nisbett, R.E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cog. Psych.*, 17, 391-416.

Dalal, M. (1988). Investigations into a theory of knowledge base revision: Preliminary Report. *Proc. 8th American Assn. for Artificial Intelligence*, pp.475-479.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal Representation of Human Judgment*. NY: Holt Rinehart & Winston.

Einhorn, H. & Hogarth, R. (1978). Confidence in judgment: Persistence in the illusion of validity. *Psych. Rev.* ,85, 395-416.

Fuhrmann, A. (1991). Theory revision through base revision. *Jour. Philosophical Logic*, 20, 175-203.

Gärdenfors, P. (1990) The dynamics of belief revision systems: Foundations vs. coherence theories. *Revue Internationale de Philosophie*,44, 24-46.

Gärdenfors, P. & Rott, H. (1992). Belief Revision. In D. Gabbay (Ed.), *Handbook of Logic in AI and Logic Programming, Vol. IV: Epistemic and Temporal Reasoning*.

Hoenkamp, E. (1988). An analysis of psychological experiments on non-monotonic reasoning. *Proc. 7th Biennial Meeting of the Canadian Soc. for the Computational Study of Intell.* (pp. 115-117).

Johnson-Laird, P.N. & Byrne, R.M.J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum.

Katsuno, H. & Mendelzon, A. (1991). Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52, 263-294.

Macnamara, J. (1986). *A border dispute: The place of logic in psychology*. Cambridge: MIT.

Nebel, B. (1991). Belief revision and default reasoning: Syntax-based approaches. *Proc. 2nd Conf. on Knowledge Representation*, (pp. 417-428).

Rips, L. J. (1990). Reasoning. *Annual Review of Psychology*, 41, 321-353.

Ross, L. & Lepper, M. (1980). The perseverance of beliefs: empirical and normative considerations. In R. Shweder (Ed.), *Fallible Judgment in Behavioral Research*. San Francisco: Jossey-Bass.

Winslett, M. (1988). Reasoning about action using a possible models approach. *Proc. 8th American Assn. for Artificial Intelligence*, (pp.89-93).

Yates, J.F. (1990). *Judgment and Decision Making*. Englewood Cliffs: Prentice Hall.