UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Psychological Evidence for Assumptions of Path-Based Inheritance Reasoning

Permalink

https://escholarship.org/uc/item/3053x0br

Journal

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

Authors

Hewson, Claire

Vogel, Carl

Publication Date

1994

Peer reviewed

Psychological Evidence for Assumptions of Path-Based Inheritance Reasoning

Claire Hewson & Carl Vogel
Centre for Cognitive Science
University of Edinburgh
2 Buccleuch Place
Edinburgh EH8 9LW
Scotland
voice: +44 31 650 4667
fax: +44 31 650 4587
{hewson,vogel}@cogsci.ed.ac.uk

Abstract

The psychological validity of inheritance reasoners is clarified. Elio and Pelletier (1993) presented the first pilot experiment exploring some of these issues. We investigate other foundational assumptions of inheritance reasoning with defaults: transitivity, blocking of transitivity by negative defaults, preemption in terms of structurally defined specificity and structurally defined redundancy of information. Responses were in accord with the assumption of at least limited transitivity, however, reasoning with negative information and structurally defined specificity conditions did not support the predictions of the literature. 'Preemptive' links were found to provide additional information leading to indeterminacy, rather than providing completely overriding information as the literature predicts. On the other hand, results support the structural identification of certain links as redundant. Other findings suggest that inheritance proof-theory might be excessively guided by its syntax.

Introduction

Inheritance reasoning is a variety of propositional default logic with graphical syntax. Inheritance reasoners purport to provide a psychologically plausible model of reasoning with defaults[1] and have been invoked in the semantic analysis of natural language generics. However, there is considerable debate in the AI literature about the "correct" definition of inheritance reasoning (Touretzky, Horty, & Thomason, 1987). Most of this discussion is based on logicians' introspective analyses of what conclusions can be drawn from any particular network of propositional default statements. Conflicting intuitions, perhaps prejudiced by interest in proof-theoretic features like computational complexity (cf. Selman & Levesque, 1989; Horty, Thomason, & Touretzky, 1990), are in part responsible for the lack of an accepted unifying semantics for inheritance reasoning (cf. Boutilier, 1989). Given the absence of a parameterized model theory, it is surprising that until very recently there have been no psychological investigations designed to elucidate the semantics of generics with respect to the idealizations of inheritance theory. Elio and Pelletier (1993) present results about the way people classify exceptional objects in light of default theories in relation to the way general default logics (alternative nonmonotonic systems) classify the same exceptional objects. They also present the first pilot study applying similar scrutiny to inheritance reasoners, but they do not consider other foundational claims of inheritance reasoning. This paper presents an experiment designed to ascertain whether assumptions of inheritance reasoning about transitivity, negative reasoning, structural preemption and structural redundancy are predictive of human reasoning with generics. This same experiment was also designed to test the degree of fit of extant theories of inheritance reasoning, but those results are reported elsewhere (Hewson & Vogel, 1994).

Inheritance Reasoning

Default inheritance reasoning is a nonmonotonic system for reasoning about generics like, "Birds fly." In these systems, generics are encoded as links in directed acyclic graphs, the nodes of which represent individuals, properties, or classes, and the links of which represent statements of positive or negative defaults. So, for example, Fig. 1 depicts a default inheritance network. Let the nodes of the graph labeled A, B and C represent penguins, birds and fliers, respectively. Thus, the network represents that birds fly, that penguins are birds, and that penguins do not fly.
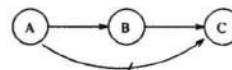

Figure 1: A Simple Inheritance Network

Inheritance reasoners define methods for reaching conclusions implicit in graph representations of sentences[2]. Implicit conclusions correspond to paths through the graph that are distinguished as *permitted*. As an example, an easily stated theory of inheritance is shortest path reasoning, in which the conclusion of a graph that is not simply linear is taken to be the conclusion that corresponds to the shortest linear path through the entire graph. Touretzky (1986) has shown this form of reasoning to be formally undesirable; however, most inheritance reasoners agree with shortest path reasoning in simple cases, and determine, for instance, that the potential path in Fig. 1 from A to C through the intervening node B is not permitted because it is *preempted* by the *more specific* information represented by the direct negative link from A to C. However, when it comes to more complex graph topologies, different inheritance reasoners diverge considerably on which paths should be permitted from a given graph (Touretzky, 1986; Touretzky et al., 1987; Boutilier, 1989).

---

[1] Partially motivated by the idea that tangled hierarchies are ubiquitous in the organization of information.
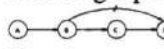
[2] In what follows we sometimes describe sets of sentences in terms of structural properties of those sets when represented in graphic notation; in these terms, theories are directed acyclic graphs.

Touretzky et al. (1987) outline a space of proof-theoretic variations in inheritance reasoners. According to the axes of variation that they identify, the reasoner of Horty et al. (1990) is *restrictedly skeptical, off-path preempting,* and *upwards chaining.* Hewson and Vogel (1994) investigate the degree to which human reasoning with sets of abstract generics corresponds to the predictions made by these properties. In this paper, we examine more foundational proof-theoretic assumptions assumed by a wider class of reasoners and not identified as a point of conflict among reasoners by Touretzky et al. (1987) nor investigated by (Elio & Pelletier, 1993); these properties are: *transitivity, negativity, preemption,* and *redundancy.*
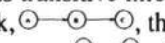
Transitivity. This property is realized in inheritance reasoners that admit chaining of explicit links into paths corresponding to implicit conclusions. Chains of statements can be either positive or negative, and can have arbitrary (but finite) length. Not all inheritance systems admit general chaining, notably the statistically based ones (Bacchus, 1989), because it is not a statistically valid inference for defaults, although it is a practically tenable inference. Thus, it is interesting to test both whether people draw conclusions consistent with transitivity at all and whether there is an interesting limit on the maximum length.
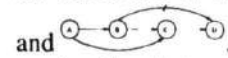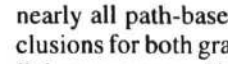
Negativity. The literature defines a negative path as one in which the final link is a negative link and the preceding links (if any) are all positive. This reflects the intuition that one cannot reason beyond a negative assertion of the form *As are normally not Bs*[3]: if As are normally not Bs, then As stand in no transitive relation to anything else that Bs might be. Nonetheless, general sequences that include non-final negative links can be labeled *negative chains.* It is possible that people reason with such statements as if negativity is a *feature,* so that if As are normally not Bs and Bs are normally Cs, then As would be considered normally not Cs. Also possible in this light is that 'double negations' can cancel or intensify each other. Both potential responses are classically invalid, but negativity propagation is more pragmatically misguided than chaining: no inheritance reasoner builds in these features. Hence, it is interesting to test the distinction between negative paths and negative chains.

Preemption. Preemptive links are treated uniformly in the literature as providing specific information that should override longer paths whose endpoints they connect. For example, ⊙—⊙—⊙—⊙, represents a network in which the inheritance literature nearly unanimously agrees to license the conclusion that As are not Ds. A test of the plausibility of preemption is simultaneously a test of instances of applicability of shortest path reasoning.

Redundancy. The priority of explicit links has been identified as a controversial issue with respect to stable reasoning (Boutilier, 1989). But, since Touretzky's thesis (Touretzky, 1986) it has been accepted in the literature that certain topologically identifiable links in inheritance networks are redundant since they convey no information that is not already present in longer paths, through transitivity. Essentially, certain explicit links are deemed redundant with respect to implicit

links, in particular, when the explicit link expresses the same conclusion that can be drawn from a longer path. In terms of graph topology, a 'redundant' link is a direct link that connects the endpoints of another path with the same polarity. For instance, a reasoner that implements transitive inference would likely conclude from this network, ⊙—⊙—⊙, that As are normally Cs. Thus, this related network, , is deemed to contain no additional information. In a reasoner that incorporates transitivity, the conclusions implicit in both graphs are the same regardless of the information-supplying status of the redundant link; however, when graphs like these are embedded in larger networks such as, ⊙—⊙—⊙—⊙, and , a reasoner that assumes redundancy (as nearly all path-based reasoners do) will reach the same conclusions for both graphs, but reasoners that assume each direct link conveys novel information may sanction different conclusions for each graph.

Experiment

Seventy-two subjects were presented with 40 problems that were designed to elicit responses which would determine whether people reason in accord with particular inheritance reasoners.

Materials. Each problem presented a set of default statements about abstract classes, followed by a question in multiple choice format; the question asked what conclusions could be drawn, based on the stated information, about the relationship between two of the classes represented (see Fig.2).

The problems were presented either in graphical form, in sentential form, or with both graphical and sentential forms together; this created one between subjects factor—*mode of presentation*—which had three levels (graph, sentence, and graph+sentence). Within the inheritance literature the kind of information represented in such problems tends to be presented graphically; the purpose of constructing a factor 'mode of presentation' was to determine whether the responses elicited by subjects would be affected by this factor. This has important methodological implications for investigating the psychological plausibility of inheritance reasoners. Fig. 2 shows the graph+sentence version of problem No. 1.

Subjects. 52 Canadian students between 7th and 12th grade participated as a result of a call for subjects posted to SchoolNet, an electronic network of Canadian schools; 20 post-secondary school individuals from North America and Europe participated by responding individually to a call for subjects sent to an assortment of internet newsgroups.

Design and Procedure. Each subject was presented with a questionnaire containing all 40 problems, randomly ordered; two random presentations (one random order and its reverse) were used and subjects were randomly assigned to receive either of these. The questionnaire contained full instructions to subjects on how to answer the problems; it was stressed that there were no right or wrong answers, and subjects were to say what *they* thought could be concluded from the information given. The factor mode of presentation created three experimental conditions: in one condition subjects received each problem in graph format, in a second condition subjects received each problem in sentence format, and in a third condition subjects received each problem in both sentence and

---

[3]It is important to remember the difference between this and the weaker: As are not normally Bs.
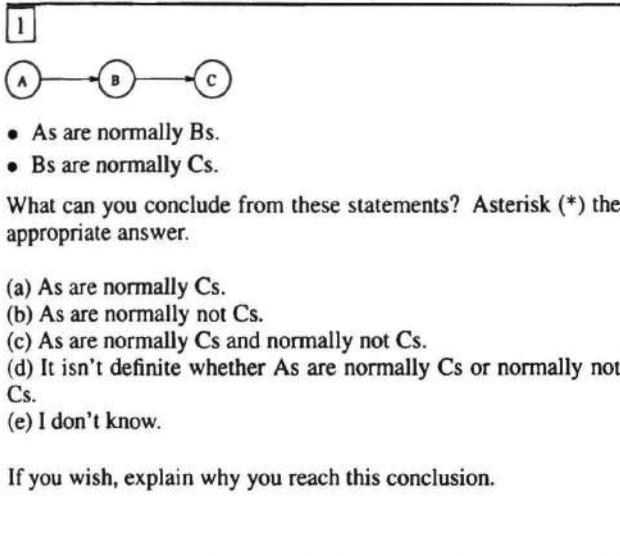
1

- As are normally Bs.
- Bs are normally Cs.

What can you conclude from these statements? Asterisk (*) the appropriate answer.

(a) As are normally Cs.
(b) As are normally not Cs.
(c) As are normally Cs and normally not Cs.
(d) It isn't definite whether As are normally Cs or normally not Cs.
(e) I don't know.

If you wish, explain why you reach this conclusion.

Figure 2: An Example Question

graph formats together.

The materials were distributed electronically. As the Canadian secondary school students did not have facilities for previewing graphics, they were assigned to the sentence only condition. Text files containing the questionnaire in both orderings were emailed to the relevant teachers who randomly assigned them to students and organized their return. The graph only and mixed modes of presentation were offered to people with access to internet news, upon individual request. For each of these conditions, files containing the same orderings of the problems as in the sentence only condition and corresponding answersheets were made available via FTP in a unique location[4], the answersheets returned to us electronically upon completion. 10 subjects received the graph only condition and 10 received the mixed modes of presentation. No limit was imposed on the time subjects spent on each problem.

Results and Discussion

The results were analyzed by picking out sets of problems that enabled conclusions to be drawn regarding the conformance of subjects' responses to specific predictions of inheritance reasoners. By this method it was possible to examine the extent to which people reasoned in accord with particular isolable features of these models.

Responses were coded in terms of the multiple choice answer categories (a–e, as shown in Fig.2), thus making it possible to directly compare subjects' responses with the predictions of inheritance reasoners. Option 'c' classifies an assertion of definite inconsistency represented by a set of generic statements, but option 'd' expresses indeterminacy. This distinction is important to the inheritance literature since most inheritance logics classify the statements in Fig. 3.a as inconsistent, but the ones represented by Fig. 3.b and Fig. 3.c are deemed inconclusive. However, both answers 'c' and 'd'

---

[4]So that no participant had access to other possible conditions in which to participate accidentally.

can be grouped together to form a general response category meaning "not 'a' or 'b'," since the alternative option 'e', "I don't know," was also provided and exercised. In analysis of the data, categories 'c' and 'd' were collapsed into a single response category, and the category 'e' (I don't know) was excluded; thus the three response categories used were 'a', 'b', and 'c/d'. We refer to subjects' responses as the *predicted* answer (the answer to a problem as predicted by H90 when it is 'a' or 'b'), the *complement* answer (when it is 'a' or 'b'), or an *indeterminate* answer (when it is 'c or d').
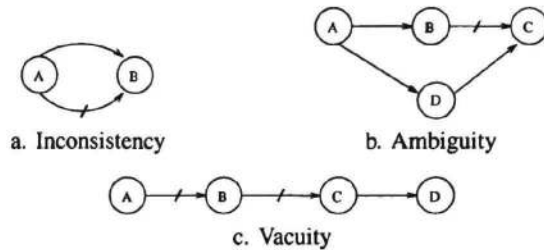

a. Inconsistency   b. Ambiguity

c. Vacuity

Figure 3: Category 'c/d': Indeterminacy

Transitivity

To test whether people reason in accord with transitivity subjects' responses to the problems with just one linear path, or a linear path with a redundant link, were compared with the responses predicted by the inheritance literature. It was found that 66% of subjects' responses conformed to the literature prediction; of the remaining 35% of responses, 11% fell into the complement category and 23% fell into category 'c/d'. A chi square analysis showed this difference in the proportion of responses in each category to be significant $(\chi^2(2) = 199.6, p < .01)$.

These results indicate that people do tend to reason in accord with transitivity. This means that people tend to conclude from the facts that As are normally Bs and Bs are normally Cs that As are normally Cs. Having said this, there were still a remaining 35% of responses which did not accord with the transitive conclusion; this result may reflect the influence of considered statistical validity—indeed, a number of subjects indicated this as the motivation for their answers. Inheritance reasoners generally admit transitivity, although as noted above in the *Inheritance Reasoning* section, some systems limit the length of chaining. We found for linear graphs that the tendency to reason transitively fit up to the maximum length we tested—three links (i.e. subjects' responses did not vary as a function of number of links). This finding supports the basic assumption of inheritance proof theory that some chaining should be admitted.

Log linear analysis showed that mode of presentation did affect responses to problems for which the predicted answer was 'a' $(\chi^2(4) = 33.87, p < .01)$, but not those for which the predicted answer was 'b'. For the former, the graphical conditions (graph only and graph+sentence) elicited mainly predicted responses, no complement responses, and very few indeterminate responses; however, in the sentence only condition subjects were less likely to give the predicted answer and more likely to give both the complement and indeterminate answers. Note that for each mode of presentation answers were more likely to conform to the predictions of

transitivity, but the 'cleaner' conformance in the graph-only condition is stark. It suggests that the graphs lend interpretive strategies in addition to simply denoting the equivalent generics. This suggests that the natural semantics of graphs interacts with reasoning about problems when they are expressed as graphs, but problems with certain structures might lead to significantly diverging responses. This is an important point because inheritance proof theory has been developed largely with topological features of graphs in mind, and for that reason may have been misled into devices that conflict with human reasoning with generics. Although there was no significant effect of mode of presentation on responses to the problems for which the predicted answer was 'b', a similar trend was observed.

Negative Paths

We distinguish between negative paths as defined above and negative chains as sequences of links containing one or more negative links but which are not also negative paths. To test whether people reason in accordance with this distinction, responses to those problems with one negative path only were compared with responses to those with one negative chain only.
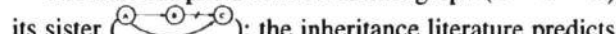
For both chains and paths subjects gave very few 'a' responses; however, whereas for negative paths there were a lot more 'b' (65%) than 'c/d' (27%) responses, for negative chains there were roughly equal numbers of 'b' (41%) and 'c/d' (47%) responses. Subjects were more likely to say 'b' and less likely to say 'c/d' for negative paths than for negative chains ($\chi^2(6) = 32.91, p < .01$), thus indicating that they do distinguish between the two. This suggests that people differentiate the validity of transitivity and general negative chaining. However, the results present some surprises: although people largely reason in accord with predictions for negative paths, the answers elicited with respect to negative chains show greater deviation from literature predictions. The equal proportion of 'c/d' (predicted) and 'b' answers to negative chains indicates a substantial tendency for people to opt for the incorrect answer (where set theoretic interpretations arbitrate). One speculation, which provides a basis for follow-up studies, is that less abstract problems would alter this pattern of responses.

Log linear analysis revealed that mode of presentation had an effect on subjects' responses for negative chains ($\chi^2(4) = 19.36, p < .01$) but not for negative paths. For chains, in the graph only condition subjects responded 'c/d' (64%) more often than 'b' (33%), but in the sentence only condition there were equal numbers of responses in these categories (43%). (The graph/sentence condition had slightly more 'c/d' (54%) responses than 'b' (42%) responses). This result reinforces the trend observed in the preceding discussion of transitivity for responses to be more polarized in the graph conditions than in the sentence condition.
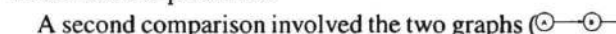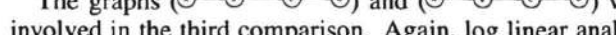
Preemptive Links

As described in the *Inheritance Reasoning* section, preemptive links are treated almost uniformly in the literature as providing more specific information that should override longer paths whose endpoints they connect. To test whether subjects dealt with preemptive links as predicted by the literature, 5 pairs of problems were compared; each comparison involved a graph

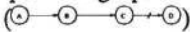and its sister graph which was identical except for the addition of one preemptive link.

The first comparison involved the graph (⊙—⊙—⊙) and its sister (⊙—⊙—⊙); the inheritance literature predicts that subjects should answer 'b' to the first of these and 'a' to its sister. Subjects gave mainly the predicted answer (68%) to the first graph, but there were also a fair number of indeterminate responses (27%); however, for the sister graph subjects were as likely to give an indeterminate response (42%) as the predicted response (43%). Log linear analysis showed that the difference in distribution of responses to each of these problems was significant ($\chi^2(6) = 63.31, p < .01$); thus, with the addition of the preemptive link there was a reduction of 'b' (predicted) responses which was reflected in an increase in both 'a' (predicted) and 'c/d' responses (with the increase in predicted responses being greater). This result suggests that the effect of a preemptive link is not to override the existing path, as the inheritance literature argues, but rather to add extra information which is considered along with the existing path.
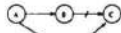
Log linear analysis showed an interaction between the difference in subjects' responses to each of the problems and mode of presentation ($\chi^2(4) = 13.07, p < .01$): for the graph only and graph+sentence conditions there was a very large drop in the number of 'b' responses (from 90% and 78%, both to 0%) between the graph and its sister, accounted for by an increase in both 'a' and 'c/d' responses. However, for the sentence only condition the observed drop in 'b' responses was a lot less (61% to 20%), and was accounted for primarily by an increase in 'a' responses. Again this result confirms the observation that the distribution of subjects' responses tends to be more polarized when graphical rather than sentential information is presented.

A second comparison involved the two graphs (⊙—⊙—⊙) and (⊙—⊙—⊙). This pair is symmetric to the preceding one, and elicited an almost identical pattern of results (in terms of predicted, complement, and indeterminate responses). As above, log linear analysis showed the observed difference in the distribution of responses to each of these two graphs to be significant ($\chi^2(6) = 64.26, p < .01$). This sustains the inference that people do not treat a preemptive link as overriding an existing path in the way the inheritance literature predicts. Surprisingly, in contrast to the previous case, no significant effect of mode of presentation was found for this comparison.

The graphs (⊙—⊙—⊙—⊙) and (⊙—⊙—⊙—⊙) were involved in the third comparison. Again, log linear analysis showed that the distribution of subjects' responses between each of these graphs was significantly different ($\chi^2(6) = 35.05, p < .01$), and this effect interacted with mode of presentation ($\chi^2(4) = 10.26, p < .05$). Subjects gave mainly the predicted answer (60%) to the first graph, though there were also about half as many indeterminate responses (31%); for the sister most responses fell into the indeterminate category (54%), and roughly half as many in each of the predicted (24%) and complement (22%) categories. In this case the addition of the preempting link revealed a decrease in 'a' (predicted) responses, accounted for by an increase in both 'b' and 'c/d' responses; however, the predicted increase in 'b'

responses was smaller than the unexpected increase in 'c/d' responses. Again this result suggests that the effect of a preempting link is not to override the existing information, but to add to it. As for the effect of mode of presentation, for both the graph-involving conditions there was a large drop in the number of 'a' responses—in the graph only condition this drop was accounted for by an increase in 'c/d' responses, and in the graph/sentence condition it was accounted for by an increase in both 'c/d' and 'b' responses. In the sentence only condition the decrease in 'a' responses was less marked, and was accounted for by an increase in 'c/d' and 'b' responses. Again the graph conditions show more polarized response.
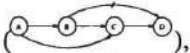
A fourth comparison examined a pair of graphs which are symmetric to the above pair, (⊙—⊙—⊙—⊙) and (⊙—⊙—⊙—⊙). The pattern of responses elicited by these graphs was very similar to the pattern described above (in terms of predicted, complement, and indeterminate responses), and again the observed difference in the distribution of responses to each of these graphs was significant under a log linear analysis ($\chi^2(6) = 36.21, p < .05$). This reinforces the same conclusion. However, unexpectedly, the effect of mode of presentation in this case was not significant though it did approach significance $\chi^2(4) = 8.8, p = .0642$.

A final analysis compared graphs ([graph]) and ([graph]); the distributions of responses between these two graph were significantly different ($\chi^2(5) = 22.00, p < .01$). To the first graph subjects responded mainly 'c/d' (65%) as predicted, and to a lesser extent, and each about equally, 'a'(15%) and 'b' (19%)); to its sister they responded mostly 'b' (53%) as predicted, but also 'c/d' (37%) and 'a' (10%). Thus, addition of a preemptive link in this case is reflected in a decrease in 'c/d' responses, accounted for by an increase in 'b' responses. Since this shift from 'c/d' to 'b' responses was only partial, where the literature predicts a complete shift from 'c/d' to 'b', this result also confirms the idea that addition of a preemptive link does not override the existing link. There was no effect on mode of presentation on this result. This test provides a specific point of comparison with the results of the pilot experiment run by Elio and Pelletier (1993); we found for the first of the two graphs that people mainly classified the graph as indeterminate (65%) and only 15% and 19% in each of the definite categories while in Elio and Pelletier's (1993) study of the same problem (presented with interpretations and with a different sort of question) roughly half of the people found the problem determinate, though people still split about equally between the two determinate categories.

Redundant Links

The effect of redundant links on subjects' responses was examined by comparing the responses to pairs of graphs that were identical apart from one redundant link; if subjects reason in accord with the predictions of the literature then their answers should not be affected by the addition of a redundant link. These paired comparisons could be broken into two groups based on the balance of polarity among paths through the networks—those in which the original graph had an equal number of positive or negative paths and for which the additio-

nal link would have offset the balance and those in which the original graph had only paths of one polarity or the other and the additional link created just another path of the same polarity. As expected in the latter case, it was found, for each of the paired comparisons, that subjects' responses were not significantly affected by the addition of a redundant link, nor was there an interaction of mode of presentation. The former case is more interesting to the inheritance literature since it involves comparisons between graphs like (⊙—⊙—⊙—⊙) and (⊙—⊙—⊙—⊙), and if the inheritance literature is correct there will be no difference in response because the additional link contains no information that is not already in the original graph. An alternative proof theory in which explicit links are assumed to convey novel information might propose a method of 'path counting' in which the number of arguments in favor of one conclusion or the other determines the decision and would predict a different response between the two problems.

As it happens, there was not a significant difference in response, which in this case fails to support path counting. However, it cannot really be taken as evidence for the predictions of the inheritance literature because the inheritance literature predicted no change in answers between the two graphs, but further predicted a definite rather than indeterminate response to the graphs. The response patterns for the first graph were stated in the preceding section where it was pointed out that people behaved contrary to the predictions of the inheritance literature: 24% gave the response predicted by the literature ('b') and 22% gave the complementary response, while 54% classified it as indeterminate although the literature presents strong intuitions that people will conclude that As are normally Ds. People gave similar responses to the sister graph with the redundant link (62% indeterminate, 17% predicted and 21% complementary). Log linear analysis showed there to be no significant difference between the distributions of responses to each of these graphs. This supports the intuition presented in the inheritance literature that the additional link is in fact redundant, even though the predicted response to the graph is not borne out. To understand this, note that a 'path counting' system which incorporated a topological definition of redundancy would have predicted that both graphs be classified as indeterminate since they would have an equal number of non-redundant positive and negative paths.

Another comparison examined the pair, (⊙—⊙—⊙—⊙) and its sister (⊙—⊙—⊙—⊙) which are symmetric in polarity to the first comparison. As in that case, people answered contrary to the inheritance literature for the first graph: 66% indeterminate, 15% predicted ('a') and 19% complement. For the sister graph responses were: 53% indeterminate, 11% predicted, and 36% complement. Again there was no significant difference between the distributions of responses to each of these graphs. It is interesting that, though not significant, there was an observed trend in the responses to these two graphs such that with the addition of the redundant link responses shifted away from determinacy, not to the literature-predicted answer ('a'), but to the answer that would have been predicted by a path counting method ('b'). However, the nonsignificance of the difference gives support to the idea that the additional link does not convey novel information.

General Discussion

This study finds human reasoning consistent with the inheritance literature in terms of its basic assumption of transitive reasoning with defaults, but found that people do not satisfy the literature's predictions about negative chains. More fundamentally, we found a tendency for people to take 'preemptive' links as additional information leading to indeterminacy rather than preemption as predicted by the literature. On the other hand, we did find support for the idea that links which express the same information that is contained in longer paths do not change responses to problems when added to graphs and are effectively 'redundant,' as the literature predicts. Graphic presentation of the problems polarized responses to problems in a way that suggests that the syntax of inheritance reasoners has influenced its proof-theory beyond its initial motivations for capturing human reasoning with generics.

We have presented an experiment which provides data about human reasoning with generics and the degree to which human reasoning makes inheritance reasoning a plausible formal model. We focused on foundational assumptions of inheritance reasoning and tested them using abstract concepts and generic relations among them. Follow-up studies should investigate the interaction of less abstract interpretations. The present results factor out the difficult to control influence of personal knowledge and beliefs about real classifications like 'pacifist' or 'birds'. The inheritance literature would model the effect that specific background knowledge has on the conclusions derived from a set of generics involving concrete interpretations by encoding those beliefs as direct links and invoking preemption. Transitivity may accurately describe people's behavior in certain abstract cases as well as for interpreted instances in which inference is performed rather than direct recall: without further contextual information, people should reason transitively with the information represented in Fig. 4.a to conclude that penguins are fliers. The influence of world knowledge which might prevent this inference is modeled by inheritance reasoners with preemption by an explicit link, as represented in Fig. 4.b. However, we found little support for preemption in ungrounded reasoning.
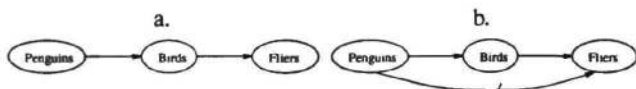

Figure 4: Representing Context

The nature of the conflicting intuitions on inheritance proof theory is such that it is informed more by studies that discriminate the influence of the abstract structure of a set of sentences than by the determination of which facts are explicitly represented and which are implicit. However, a more interesting set of contextual effects given the model of reasoning supplied by the inheritance literature would be those factors that have an impact on the abstract patterns of reasoning rather than the presence or absence of explicit information—for instance whether reasoning with known quantities (as opposed to using sets that have fuzzy cardinality, like 'chairs') eliminates the applicability of transitive reasoning. Studies of both abstract and grounded reasoning are important parts of the general problem of determining whether there is abstract systematicity in human reasoning with generic information.

Further studies should explore more of the proof-theoretic claims in greater detail. Hewson and Vogel (1994) consider on-path and off-path preemption as well as cascaded ambiguity (cf. Touretzky et al., 1987) in addition to the more foundational consideration of symmetry between positive and negative conclusions. The initial result presented here on redundancy should also be examined further, especially given the conflicting intuitions in the literature about what the appropriate topological definition of redundancy should be (cf. Boutilier, 1989). For example, Horty et al. (1990) use a more complex definition than the one we suggested using informal terms in this paper; in theirs certain links that are redundant in the simpler terms are in fact deemed by them to convey novel information. We have not tested these problems. Finally, the present analysis suggests that path-counting should also be studied in closer detail to determine its efficacy as a predictor of human reasoning with generics as a model of weighted evidence.

Acknowledgements

References

Bacchus, F. (1989). A Modest, but Semantically Well Founded, Inheritance Reasoner. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. Detroit, Michigan.

Boutilier, C. (1989). A Semantical Approach to Stable Inheritance Reasoning. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. Detroit, Michigan.

Elio, R. & Pelletier, F. J. (1993). Human Benchmarks on AI's Benchmark Problems. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, pp. 406–411. June 18-21, 1993. Boulder, Colorado.

Hewson, C. & Vogel, C. (1994). Psychological Evidence for Styles of Path-Based Inheritance Reasoning. Centre for Cognitive Science, University of Edinburgh. In preparation.

Horty, J., Thomason, R., & Touretzky, D. (1990). A Skeptical Theory of Inheritance in Nonmonotonic Semantic Networks. *Artificial Intelligence*, 42(2-3), 311–48.

Selman, B. & Levesque, H. (1989). The Tractability of Path-Based Inheritance. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. Detroit, Michigan.

Touretzky, D. (1986). *The Mathematics of Inheritance Systems*. Los Altos, CA: Morgan Kaufman.

Touretzky, D., Horty, J., & Thomason, R. (1987). A Clash of Intuitions: The Current State of Non-Monotonic Inheritance Systems. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*. Milan, Italy.