

Washington University in St. Louis

Washington University Open Scholarship

All Computer Science and Engineering
Research

Computer Science and Engineering

Report Number: WUCS-90-22

1990-06-18

Workshop on Defeasible Reasoning with Specificity and Multiple Inheritance

R. P. Loui, M. Kahn, and G. Simari

A workshop on defeasible reasoning with specificity was held in under the Arch in St. Louis during April 1989, with support fro AAI and McDonnell Douglas, and the assistance of Rockwell Science Center Palo Alto and the Department of Computer Science of Washington University. The document includes a report on the proceedings and parts of the workshop notes that can be distributed. These include the schedule, lists of participants, synopses of systems, and benchmark problems.

Follow this and additional works at: https://openscholarship.wustl.edu/cse_research

Recommended Citation

Loui, R. P.; Kahn, M.; and Simari, G., "Workshop on Defeasible Reasoning with Specificity and Multiple Inheritance" Report Number: WUCS-90-22 (1990). *All Computer Science and Engineering Research*. https://openscholarship.wustl.edu/cse_research/697

Department of Computer Science & Engineering - Washington University in St. Louis
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

**WORKSHOP ON DEFEASIBLE REASONING
WITH SPECIFICITY AND MULTIPLE
INHERITANCE**

R. P. Loui, M. Kahn and G. Simari

WUCS-90-22

**Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899**

Abstract

A workshop on defeasible reasoning with specificity was held in under the Arch in St. Louis during April 1989, with support from AAAI and McDonnell Douglas, and the assistance of Rockwell Science Center Palo Alto and the Department of Computer Science of Washington University.

This document includes a report on the proceedings and parts of the workshop notes that can be distributed. These include the schedule, list of participants, synopses of systems, and benchmark problems.

Part of this report to appear in AI Magazine, 1990.

Workshop on Defeasible Reasoning with Specificity and Multiple Inheritance

R. P. Loui

Depts. of Computer Science and Philosophy
Washington University

M. Kahn

Depts. of Internal Medicine and Computer Science
Washington University

G. Simari

Dept. of Computer Science
Washington University

June 18, 1990

1

A workshop on defeasible reasoning with specificity was held in under the arch in St. Louis during April 1989, with support from AAAI and McDonnell Douglas, and the assistance of Rockwell Science Center Palo Alto and the Department of Computer Science of Washington University.

The workshop brought together all the proposers of systems of non-monotonic or defeasible reasoning that exhibited subclass or specificity defeat. There were twenty invited people. Also attending, in equal number, were persons responding to the call who somehow demonstrated their need to attend to the program committee. This committee consisted of Dave Etherington, Hector Geffner, and David Poole. One third of the attendees of the workshop came from abroad.

Twenty-six participants submitted three-page notes on their current thinking, and there were synopses of a dozen existing formal systems. These notes were edited and available at the workshop, but, by agreement with the authors, will not be distributed.

The workshop was recorded. An edited version of the recordings can be obtained by contacting the workshop organizer (first author). The proceedings

have been transcribed and will be made available soon.

The workshop was supposed to provide a venue for challenges to each system and to allow the airing of disputes already on record. It was however discovered that many of the disputes are no longer pointed. In controversy's stead there was a free exchange of half-baked, new ideas, and the development of more general perspectives on emerging work. Review of the tapes shows a dense, rich exchange, especially on methodology.

The workshop program consisted mostly of panels. There were also provocations by persons with extreme positions, a poster session, and a problem session for working various problems with various systems. Surprisingly, there was little concern over which systems solved which problems. The workshop was essentially a series of partly planned presentations: each person associated with a system or a contribution had twenty minutes to explain himself to a community that needed no background and would countenance no salesmanship. The pace was frantic, with tightly scheduled activities consuming all but eight hours a day. By weekend's end, all participants were exhausted.

Peculiar to this workshop was the participation of some senior philosophers of science and philosophical logicians, who were able to lend perspective. The workshop also contained previews of several of the remarks heard at the First International Conference on Principles of Knowledge Representation and Reasoning (KR89), held in Toronto during May 1989, including remarks by Kautz and Selman, Doyle, Pearl, Etherington, Neufeld, and Poole.

2

The workshop began with a session titled "What is this thing we're trying to formalize?" David Etherington noted there is temptation to say that Geffner's system (in *Knowledge Representation and Defeasible Reasoning*, H. Kyburg, *et al.*, eds., Kluwer 1990) just is the kind of defeasible reasoning we've all been trying to formalize. But, he reminded us, that's what we thought of Touretzky's system for defeasible inheritance in 1985. Ron Loui tried to establish a convention regarding notation in inheritance hierarchies and couldn't. (Thereafter no subsequent attempt to standardize syntax or vocabulary was made by anyone.)

Donald Nute reminded us that defeasible reasoning is not always motivated by probabilistic concerns; L. Thorne McCarty gave an example from British civil law. Judea Pearl thought that the example could be made probabilistic with consideration of utilities. Ben Grosz pointed out that one can always find a utility model under which adoption of a rule is justified. Nute explained objective criteria for adopting rules, other than truth conditions and justification conditions, including assertability conditions and compliance conditions. One can comprehend the conditions for complying with a rule, *e.g.*, a maxim for chess playing, without knowing the conditions for its justification. David Poole thought that this supported his proposed requirement that inventors of systems

write user's manuals.

Pearl felt that the disputes about system behavior had to do with behaviors on which we lack strong intuitions. He wondered if we could perceive of an era when decisions on the esoteric questions could make a drastic difference.

3

Six impromptu talks were given the first evening in parallel with a problem session that elicited discussion on some fifty benchmark problems that had appeared in the literature. A list of these problems is available upon request.

4

The next morning's session focused on research methodology. Poole claimed that reasoning is based on arguments; people reason amongst themselves in this way; his program, THEORIST, tries to be the simplest argument system. He claimed that he just uses logic; if you don't like the conclusion, criticise the premises. Kurt Konolige ferretted out the admission that much of this logic occurs at a meta-level. This deliberate confusion of logic and meta-logic caused Poole's abuse at KR89, where the remarks were reiterated.

Fahiem Bacchus preached a purely statistical view of defeasible reasoning. John Pollock, Konolige, and Etherington were concerned that defeasible reasoning underlies the statistical reasoning to which Bacchus appealed. (Grosz had given a poster talk on this point the night before.) Bacchus thought there was room to call rules for selecting reference classes policies, rather than reasoning; this is a long-standing view of Henry Kyburg.

Jim Delgrande explained that in his system, default instantiation relies on an assumption of least exceptionality of the world. Poole felt that if you're willing to make assumptions, you don't need conditional logic to define possible worlds in the first place.

Jeff Horty displayed an inheritance net and asked what conclusions each of us drew from the net. When asked what the links meant, Horty said they meant something like tendency. In one of the sharpest exchanges, Shastri pressed on the meaning of tendency: "You don't tell me what the links mean then you ask me what conclusions to draw!" Horty quipped, "Sometimes that happens." Horty claimed rhetorically that there were exactly 72 possible theories for inheritance and contemplated the translation of each to classical non-monotonic logics. Matt Ginsberg asked if the path-based work would go away if the translation were successful. Horty thought they would, except for implementation. Rich Thomason thought that translations reminded him of model theory for proof theories.

5

The most glib session asked the senior philosophers to comment on the role of convention in the design of logic, and on defeasible reasoning's place in philosophical logic's future. David Israel introduced the speakers, but first remarked that inheritance was supposed to be simple and had nothing to do with problems such as the Yale Shooting Problem. Perhaps we ought not to talk about all of defeasible reasoning at once.

Jim Fetzer claimed that defeasible reasoning is inductive reasoning when you don't know what you're doing. He reviewed Hempel's framework for scientific reasoning and asked us to clarify things in this framework that are unclear in defeasible reasoning. Examples include the purpose of inquiry and the seriousness of making mistakes.

Henry Kyburg thought it's not going to be easy to decide whether or in what sense the world dictates one logic or another. He listed for our consideration classical, temporal, modal, deontic, causal, intuitionistic non-monotonic, default, inductive, and probabilistic logics—all are in the same boat. Then he suggested that there are nevertheless principles for choosing among conventions: not semantic basis nor intuitive persuasiveness, but simplicity, power, familiarity, elegance. Kyburg also felt that specificity construed as subsets is inadequate—specificity taken as logical strength is just the total evidence requirement of induction.

Pollock told us not to worry about semantics until we have a better idea what the target is. He also noted that theoretical reasoning is skeptical whereas practical reasoning is credulous; this eliminates one of the problems in Touretzky et.al.'s "Clash of Intuitions" (IJCAI 1987). These are similar to the remarks he made at the AAAI Spring Symposium.

Nute added to the assault on semantics, stating that Frege and Russell probably weren't worried about completeness when they were working on what we now know as God's logic.

Thomason pretended to be a computer scientist and recommended keeping the philosophers at arm's distance. He felt that philosophical logic now belongs in computer science and that most of the former's new directions will come from the latter. Philosophical imperialism is bad for us; conventionalism can't be right—we can't just go around choosing our logic. Fortunately, the needs of users force us to keep our feet on the ground. Thomason joined Poole in requiring that manuals for systems be written. He also suggested that someone try to put defeasible and probabilistic reasoning together in one system.

6

The afternoon began with a session on defeasible reasoning and probability. Ben Grosz, as moderator, explained the view that inheritance arises in probabilis-

tic reasoning, so probabilities can't be the underlying semantics for defeasible reasoning. This idea had been mentioned earlier by several people and Grosz clarified it.

Bacchus tried to defend a system in which only a single defeasible link can be used in any chain of reasoning.

Eric Neufeld defended his system, motivated by qualitative probabilistic relationships, which reasons about shifts in beliefs. The topology encodes an underlying probability distribution that is factorable like influence diagrams, that licenses inferences that are not in other systems. For instance, it is reversible for diagnosis. Ginsberg found it counterintuitive, stating that it was "an attempt to model my reasoning in a certain mathematical way that has as a result something I find totally counterintuitive."

Neufeld also discussed Simpson's paradox and the lottery paradox in the context of defeasible reasoning. Pearl thought it was a great example of convention versus probability.

Hector Geffner talked about the conservatism of Ernest Adams' rules (in *Aspects of Inductive Logic*, J. Hintikka and P. Suppes, eds., Elsevier 1966), noting that they are shared by Lehmann, Delgrande, Makinson, and Maggidor. He explained that the last rule in his system assumes that the antecedent of a defeasible rule holds in the presence of other evidence. This raised a discussion on irrelevance, and justifying extensions to the core inference rules that are widely shared.

7

John Pollock was then invited to provoke the audience. Pollock felt that defeasible reasoning has a more complicated logical structure than has been appreciated in inheritance. Specificity defeaters are an incomplete generalization of subset defeaters. "Most penguins do not fly" does not entail "Most birds fly," and both are legitimate parts of the antecedents of reasons, in the way that Pollock writes reasons. Pollock also voiced the need for a projectibility constraint to prevent inheritance from disjunctive classes, for instance.

8

Saturday ended with a session on issues and principles. Horty continued to bemoan the numerous choices in inheritance about which he has no preference. He also raised contraposition as an issue of contention. Lin Padgham suggested that contrapositives be added as rules at a lower priority after drawing primary default inferences. Everyone seemed to agree, though there was a split among those who wanted it and those who did not.

Ginsberg debunked a naive argument against contraposition. He went on

with an attack on Gelfond and anyone else who would describe inheritance reasoning in a “god-awful way in their frameworks just to boost their confidence in their tools.”

Poole focused on the distinction between background and contingent evidence, noting that it arises in probabilistic approaches and path approaches, as well as in logic-based approaches like THEORIST.

Loui tried to carve a distinction between two paradigms for defeasible reasoning systems: irrelevance-based, which has axioms and a monotonic proof theory, and argument-based, which is always done in the meta-language. Pearl summed it as “Persons born inhibited need to be encouraged; if you are promiscuous, then your behavior needs to be limited.” But Loui felt the two paradigms behave differently under computational limitations. The session ended abruptly during musings on the relation between irrelevance and lack of defeat.

9

Sunday morning began with a discussion that was supposed to be on path-based versus model-based approaches to defeasible reasoning. Gerhard Brewka told us that defeasible reasoning can be construed as inconsistency-handling. He enjoined us to compute (preferred) maximal consistent subsets, since we have to solve the problem of handling inconsistencies anyway: in this approach, specificity is implicit.

Geffner felt that a focus on inheritance means commonalities are obscured. Many solutions to our problems have been understood only in terms of particular systems – inheritance is the biggest offender. For example, irrelevance is common to defeasible reasoning systems.

Michael Gelfond argued for auto-epistemic logic and reductions thereto, because he believes most formalisms reduce to auto-epistemic logic. He complained that not to translate systems makes it hard to see mathematical properties of new systems. Etherington wondered whether the exercise of translating inheritance into auto-epistemic logic tells us anything about inheritance or about auto-epistemic logic. Ginsberg reiterated his point that there are general lessons to be learned in a general framework, which are lost by translating to special frameworks, *e.g.*, preference of more premises over fewer. He asked if Gelfond believed in that, and Gelfond did not.

Brian Haugh depicted arguments as just richer relations between antecedents and consequents than links. Also, model-based theories with specificity have problems: apparently, successful theories must refer to links or arguments – they must refer to syntactic objects.

10

Sunday continued with remarks on implementation and complexity. Henry Kautz was aghast that there would be a separate session devoted to these topics. AI doesn't take computation as a metaphor, he inveighed, but takes thought as a specific kind of computation.

Randy Goebel related a story about THEORIST's development: deKleer said it couldn't be done but they had it up and running. He also felt that hands-on work would simplify defeasible reasoning systems more than continued mathematical work.

Thomason noted that the main reason to be interested in inheritance is its close connection with computationally effective deliverables.

Lynn Stein argued that tractability is irrelevant because formalism is just now being produced. We still have no agreement on what the right answer is, which is typical in commonsense reasoning. So we say of research like Touretzky's that its contribution was a specification that was more right than shortest path, if still not completely right. Its NP-completeness was not what made it a contribution. Grosz recommended to a weary audience that "perhaps we should inherit this discussion from AI in general."

Padgham then discussed the implementation of her inheritance system. Finally, Bart Selman discussed his various tractability results from collaboration with Levesque and Kautz.

11

Jon Doyle was asked to provoke the crowd. His work with Mike Wellman indicated that the search for a unified logic to resolve preferences in desirable ways is impeded by Arrow's Theorem, made famous in economics and social choice theory. One possible way out is to limit the domain of preferences. Etherington, Grosz, and Konolige all cried that the pessimistic result would not apply to skeptical reasoners? Doyle felt that even if it did not apply to skeptical reasoners, it applied to choice between skepticism and credulity. McCarty suggested that we can learn from what happened in social choice theory. The ideal is unattainable, but in lieu of the ideal, we can and should ask how to cope with the impediments to achieving the ideal.

12

In the summary session, Etherington was still concerned that we couldn't distinguish progress from motion, and that we had agreed on no explicit requirements for future papers. Konolige talked about having more "cribs in the deck," *i.e.*, places to file the various defeasible reasoning systems that have been developed, in an orderly fashion. He felt that there weren't any more general principles to

be found and that domain-dependent, flexible specification of priorities was the next area to investigate. Konolige also brought up the idea of resource bounds.

Thomason related his experience with the plethora of modal theories, where there are just as many alternatives as in defeasible reasoning; people learn to live with the alternatives. Theoretical progress has been really remarkable in defeasible reasoning, said Thomason, but he again reminded us to connect responsibly with “technology,” that is, with implemented systems. A lot of the problems might also disappear if we were to focus on decision-making. Generally, he pled for unity.

Doyle reminded us that basing our systems on intuition is fine if we are willing to live with intuitions that differ between people and change over time. He felt that the underlying intuitions invoked economic issues. Earlier, he had referred to social choice theory; here he mentioned the costs of reasoning under limited resources. Are there 72 theories, as Horty says, or 72 different utilities?

David Israel then led a discussion on limited rationality. He linked it to the pressures to commit to belief, to defaults, and to intentions. Finally Israel took a larger view to assuage some worries. He asked us to imagine what might have happened if complexity theory had arisen before 1928. He thought that even without “the crazy stuff about semantics,” paths were very good for inheritance, and arguments were very nice for defeasible reasoning. He recalled that in 1957 there was a symposium titled “Is There One Correct Modal Logic?” which seems ludicrous today. But he warned, there might yet be a difference –it was obvious that all the modal logics invoked different notions. Our clashes over defeasible reasoning somehow seem more substantive.

13

Overall, the workshop was a success despite the following considerations. It was exhausting. It did not achieve its original goals. More worrying is that the published output in this area has diminished drastically since the workshop.

But interest remains high. The workshop settled many of the disputes out of print. It established an understanding among the active researchers what would be considered old and new. It stands as a watermark of defeasible reasoning, one decade after the ruminations on non-monotonicity began.

14

The remainder of this document includes parts of the workshop notes that can be distributed. These include the schedule, list of participants, synopses of systems (though Brian Haugh says his system is inaccurately portrayed), and benchmark problems.

Defeasible Reasoning with Specificity and Multiple Inheritance

Workshop held in
Saint Louis April 7-9, 1989

*Sponsored by AAI and McDonnell Douglas
Co-Sponsored by The Washington University Department of Computer Science and
Rockwell Science Center Palo Alto*

DEFEASIBLE REASONING WITH SPECIFICITY
AND MULTIPLE INHERITANCE

Schedule

St. Louis April 7 - 9, 1989

Sponsored by AAAI and McDonnell Douglas
Co-Sponsored by The Washington University Department of Computer Science and
Rockwell Science Center Palo Alto

Friday April 7

5:00pm. Cheshire Inn Bus leaves Embassy Suites for Washington University.

5:45pm - 7:15pm. Women's Building Formal Lounge, Washington University.

Led Discussion: What Is This Thing We're Trying to Formalize?
David Etherington, Donald Nute, Judea Pearl, Erik Sandewall.
Ronald Loui, *Moderator*. Coordinated by Poole.

7:30pm. Hors D'oeuvres Reception at the Washington University Faculty Club.

8:45pm. Bus leaves Faculty Club for Embassy Suites.

9:00pm - 11:00pm. Embassy Suites Director's Suites.

Suite 1. *Poster Session*. Wine and Cheese.
Fifteen minute presentations welcome. Coordinated by Etherington.
Suite 2. *Problem Session*. Beer and Chips.
Problems and systems welcome. Coordinated by Loui.

Saturday April 8

8:00am. Breakfast in the Embassy Suites Atrium.

9:00am - 10:30am. Embassy Suites Terrace E.

Session: Research Methodologies: Individual Perspectives.
John Harty, *Individual*. David Poole, *Individual*.
Fahiem Bacchus, *Individual*. James Delgrande, *Individual*.
Benjamin Groszof, *Session Chair*. Coordinated by Poole.

10:30am. Morning Break.

11:00am - 12:15pm. (Continuing).

Panel: Philosophy of the Future or More of Conventionalism's Curse?
James Fetzer, Henry Kyburg, John Pollock, Donald Nute,
Rich Thomason.
David Israel, *Moderator*. Coordinated by Loui.

12:15pm. Unstructured Lunch on Laclede's Landing.

1:45pm - 3:00pm. Embassy Suites Terrace E.

Panel: Probabilistic Approaches.
Fahiem Bacchus, Hector Geffner, James Hawthorne, Eric Neufeld.
Judea Pearl, *Moderator*. Coordinated by Geffner.

3:00pm. Afternoon Break.

Saturday April 8 (*continued*)

3:30pm - 4:15pm. (Continuing).

Provocation: Specificity is The Wrong Generalization.
John Pollock, *Provocateur*. Coordinated by Loui.

4:15pm - 5:30pm. (Continuing).

Led Discussion: Issues and Principles.
Matt Ginsberg, John Horty, Ronald Loui, David Poole.
Paul Morris, *Moderator*. Coordinated by Etherington.

5:30pm. Free Time.

8:00pm. Drinks on the Riverboat Goldenrod.

8:30pm. Buffet Dinner on the Riverboat Goldenrod.

Sunday April 9

8:30am. Breakfast in the Embassy Suites Atrium.

9:15am – 10:30am. Embassy Suites Terrace E.

Led Discussion: Path, Argument, and Model-Based Approaches.

Hector Geffner, Michael Gelfond, Brian Haugh, Erik Sandewall.
Gerhard Brewka, *Moderator*. Coordinated by Geffner.

10:30am – 11:00am. Morning Break.

11:00am – 12:00pm. (Continuing).

Led Discussion: Complexity and Implementability.

Randy Goebel, Lin Padgham, Bart Selman, Lynn Stein.
Henry Kautz, *Moderator*. Coordinated by Poole.

12:00pm. Lunch on the Riverboat Robert E. Lee.

2:00pm – 2:45pm. Embassy Suites Terrace E.

Provocation: Some Impossibility Results.

Jon Doyle, *Provocateur*. Coordinated by Loui.

2:45pm – 4:00pm. (Continuing).

Led Discussion: Questions of Substance or Mere Clashes of Intuition?

Jon Doyle, David Israel, Kurt Konolige, Rich Thomason.
David Etherington, *Moderator*. Coordinated by Etherington.

DEFEASIBLE REASONING WITH SPECIFICITY AND MULTIPLE INHERITANCE

Participants

St. Louis April 7 - 9, 1989

Dr. Fahiem Bacchus
Department of Computer Science University of Waterloo
Waterloo, Ontario Canada N2L-3G1
fbacchus@watdragon.waterloo.edu 519-888-4670

Dr. William Ball
Department of Computer Science Washington University
St. Louis, Missouri 63105
wball@tardis.wustl.edu 314-889-6101

Dr. James Beard
Medical Informatics Program Washington University
St. Louis, Missouri 63105
beard@wuibc2.wustl.edu 314-889-4726

Dr. Gerhard Brewka
Gesellschaft fuer Mathematik und Datenverarbeitung
Schloss Birlinghoven Postfach 12 40 5205 Sankt Augustin F. R. G.
brewka@gmdzi.ucp 49-2241-142687

Mr. Steve Cousins
Department of Computer Science Washington University
St. Louis, Missouri 63105
sbc@wucs1.wustl.edu 314-362-6968

Dr. James Delgrande
School of Computing Science Simon Fraser University
Burnaby, B.C. Canada V5A 1S6
jim%lccr.sfu.cdn@relay.ubc.ca 604-291-4335

Mr. Patrick Doherty
Institut foer Datavetenskap Universitetet i Linkoeping
S-581 83 Linkoeping, Sweden
pdy@ida.ltu.se 013-2426 46-13-28-2426

Dr. Jon Doyle
Laboratory for Computer Science MIT
545 Technology Square Cambridge, MA 02139
doyle@zermatt.lcs.mit.edu 617-253-3512

Dr. David Etherington
AT&T Bell Laboratories
600 Mountain Ave. Murray Hill, NJ 07974-2070
ether@research.att.com 201-582-6914

Dr. James Fetzer
Department of Philosophy University of Minnesota
Duluth, MN 55812
phil@ub.d.umn.edu 218-726-8548 218-724-2706

Ms. Rosanne Fulcomer
Department of Computer Science Washington University
St. Louis, Missouri 63105
rmf@wucs1.wustl.edu 314-362-6124

Ms. Patricia Geerts
University of Antwerp Department of Math-Computer Science
Universiteitsplein 1 B2610 Wilrijk, Belgium
vermeir@banuia52.bitnet

Mr. Hector Geffner
Computer Science Department UCLA
Los Angeles, CA 90024-1596
hector@cs.ucla.edu 213-825-7367

Dr. Michael Gelfond
Department of Computer Science University of Texas
El Paso, TX 79968-0518
gelfond%utep.uucp@cs.utexas.edu cv00@utep.bitnet 915-747-5470

Dr. Matt Ginsberg
Department of Computer Science Stanford University
Stanford, CA 94035
ginsberg@polya.stanford.edu 415-723-1239

Dr. Randy Goebel
Department of Computing Science University of Alberta
Edmonton, Alberta Canada T6G 2H1
goebel@alberta.cs.ca 403-492-2683

Mr. Moisés Goldszmidt
Cognitive Systems Laboratory UCLA Computer Science Department
Los Angeles, CA 90024-1596
moises@cs.ucla.edu 213-206-3809

Mr. Scott Goodwin
Department of Computing Science University of Alberta
Edmonton, Alberta Canada T6G 2H1
scott@alberta.uucp 403-432-2821

Dr. Eric Gregoire
Faculte Des Sciences Appliques Unite d'Informatique
Place Sainte-Barbe 2 B-1348 Louvain-la-Neuve Belgium
egl@info.ucl.ac.be 32-10-47-2415

Mr. Benjamin Grosf
T.J. Watson Research Center IBM
P.O. Box 704 Yorktown Heights, NY 10598
grosf@ibm.com 914-789-7783

Dr. Brian Haugh
Martin Marietta Laboratories
1450 South Rolling Road Baltimore, MD 21227-3898
haugh@mmlai.uu.net haugh@mimsy.cs.umd.edu 301-247-0700

Dr. James Hawthorne
Honeywell Systems and Research Center
Minneapolis, MN 55418
hawthorn@src.honeywell.com 612-870-1312 612-782-7402

Dr. John Horty
Department of Computer Science CMU
Pittsburgh, PA 1521
horty@ri.cmu.edu john.horty@b.gp.cs.cmu.edu 412-268-3825

Dr. David Israel
SRI International
333 Ravenswood Menlo Park, CA 94025-3493
israel@ai.sri.com israel@csl.stanford.edu 415-859-4254

Dr. Michael Kahn
Department of Internal Medicine Washington University
St. Louis, Missouri 63105
kahn@wucsl.wustl.edu 314-362-6968

Dr. Henry Kautz
AT&T Bell Laboratories
600 Mountain Avenue Murray Hill, NJ 07974-2070
kautz@allegro.att.com 201-582-2815

Mr. Bruce Kirby
Center for Cybernetic Studies University of Texas
CBA 5.202 Austin, TX 78712-1177
kirby@cs.utexas.edu 512-471-1821

Dr. Kurt Konolige
SRI International
333 Ravenswood Menlo Park, CA 94025-3493
konolige@ai.sri.com 415-859-2788

Dr. Henry Kyburg
Departments of Computer Science and Philosophy University of Rochester
Rochester, NY 14627
kyburg@cs.rochester.edu 716-275-4105

Dr. Ronald Loui
Departments of Computer Science and Philosophy Washington University
St. Louis, MO 63130
loui@ai.wustl.edu 314-889-6102

Dr. J.-J.Ch. Meyer
Faculteit Wiskunde en Informatica Vrije Universiteit
De Boelelaan 1081 1081 HV Amsterdam The Netherlands
jules@cs.vu.nl.uucp 20-5485779 71-313258

Dr. L. Thorne McCarty
Department of Computer Science Rutgers University
New Brunswick, NJ 08903
mccarty@aramis.rutgers.edu 201-932-2657 201-821-8028

Dr. Paul Morris
IntelliCorp
1975 El Camino Real West Mountain View, CA 94040
morris@intellicorp.com 415-965-5723

Dr. Eric Neufeld
Department of Computer Science University of New Brunswick
Fredericton, New Brunswick Canada E3B 5A3
eric@unb.ca 506-453-4566

Dr. Donald Nute
Advanced Computational Methods Center University of Georgia
Athens, GA 30602
dnute%uga.bitnet@cunyvm.cuny.edu 404-542-0358

Ms. Lin Padgham
Institut foer Datavetenskap Universitetet i Linkoeeping
S-581 83 Linkoeeping Sweden
linpa@majestix.ida.liu.se +46 13 282066

Dr. Judea Pearl
Cognitive Systems Laboratory Computer Science Department UCLA
Los Angeles, CA 90024-1596
judaea@cs.ucla.edu 213-825-3243

Mr. Marcelino Pequeno
Department of Computing Imperial College University of London
180 Queen's Gate, London SW7 2BZ U.K.
jbc%doc.ic.ac.uk@cunyvm.cuny.edu 44-01-589-5111 x4985

Dr. John Pollock
Department of Philosophy University of Arizona
213 Social Sciences Building #27 Tucson, AZ 85721
pollock@jvaz.ccit.arizona.edu 602-323-0571

Dr. David Poole
Department of Computer Science University British Columbia
Vancouver, British Columbia Canada V6T 1W5
poole@vision.cs.ubc.ca 604-228-6254

Mr. T. K. Prasad
Department of Computer Science SUNY Stony Brook
Stony Brook, NY 11794
prasad@sbcs.sunysb.edu 516-632-7677

Dr. Halina Pryzmusinski
Department of Computer Science and Mathematics University of Texas
El Paso, TX 79968-0518
fp00@utep.bitnet 915-747-5470

Dr. Erik Sandewall
Institut foer Datavetenskap Universitetet i Linkoeeping
S-581 83 Linkoeeping Sweden
enea.lisbet.liu.se!e-sandewall@uunet.uu.net

Mr. Bart Selman
Department of Computer Science University of Toronto
Totonto, Ontario Canada M5S 1A4
bart@neat.ai.toronto.edu 416-978-5182

Dr. Lokendra Shastri
Department of Computer and Information Science University of Pennsylvania
Philadelphia, Pennsylvania 19104-6389
shastri@cis.upenn.edu 215-898-2661

Mr. Guillermo Simari
Department of Computer Science Washington University
St. Louis, MO 63130
grs@cics.wustl.edu 314-889-6124

Ms. Lynn Stein
Department of Computer Science Brown University
Box 1910 Providence, RI 02912
las@cs.brown.edu 401-863-7664

Ms. Helga Smith Taylor
Mcdonnell Douglas AI Support Center P.O. Box 516-0922232
St. Louis, Missouri 63166
314-925-7904

Dr. Rich Thomason
Intelligent Systems Program University of Pittsburgh
Pittsburgh, PA 15260
rich.thomason@ri.cs.cmu.edu 412-624-5791

Dr. Dirk Vermeir
University of Antwerp Department of Math-Computer Science
Universiteitsplein 1 B2610 Wilrijk, Belgium
vermeir@banuia52.bitnet

Part I: Abstracts and Papers

Brief working notes and abstracts are compiled in this section. The writings appear in alphabetic order by first author's last name. Writings that appears in this section are not for redistribution—they are present here for workshop participants to review positions adopted by other members.

Fahiem Bacchus
James Delgrande
Jon Doyle and Michael Wellman
James Fetzer
Hector Geffner
Moisés Goldszmidt and Judea Pearl
Eric Grégoire
Brian Haugh
James Hawthorne
Henry Kautz
Bruce Kirby
Henry Kyburg
Ron Loui
L. Thorne McCarty
Eric Neufeld
Donald Nute
Lin Padgham
Marcelino Pequeno
John Pollock
David Poole
T. K. Prasad
Bart Selman and Hector Levesque
Lokendra Shastri
Guillermo Simari and Ron Loui
Lynn Stein
Dick Vermeir, Donald Nute, and Paul Geerts

Part II: System Descriptions

Brief descriptions of various defeasible reasoning systems are compiled in this section. The systems appear in alphabetic order by author's last name.

Delgrande
Geffner
Haugh
Horty and Thomason
Loui
Neufeld
Nute
Padgham
Pollock
Poole
Sandewall
Touretzky

Synopsis of Delgrande (AAAI 87)

synopsis by Loui

From sentences such as:

$$\begin{aligned} \text{Raven}(x) &\Rightarrow \text{Black}(x) \\ \text{Raven}(x) \wedge \text{Albino}(x) &\Rightarrow \neg \text{Black}(x) \end{aligned}$$

derive:

$$\text{Raven}(x) \wedge \text{Albino}(x) \wedge \text{HasWings}(x) \Rightarrow \neg \text{Black}(x)$$

by axioms of N :

$$\begin{aligned} \alpha &\Rightarrow \alpha && \text{(ID)} \\ \alpha &\Rightarrow \beta, \alpha \Rightarrow \gamma \therefore \alpha \Rightarrow \beta \wedge \gamma && \text{(right conj)} \\ \alpha &\Rightarrow \beta, \alpha \wedge \beta \Rightarrow \gamma \therefore \alpha \Rightarrow \gamma && \text{(bayes)} \\ \alpha &\Rightarrow \neg \beta, \alpha \Rightarrow \gamma \therefore \alpha \wedge \neg \beta \Rightarrow \gamma && \text{(bayes)} \\ \alpha &\Rightarrow \gamma, \beta \Rightarrow \gamma \therefore \alpha \vee \beta \Rightarrow \gamma && \text{(disjunction)} \end{aligned}$$

(I simplify $\neg(\alpha \Rightarrow \beta)$ with $\alpha \Rightarrow \neg\beta$; presumably the latter entails the former)

or by convention I:

$$\begin{aligned} \alpha &\Rightarrow \gamma \text{ is supported if } \exists \beta \text{ s.t.} \\ &\alpha \vdash \beta \\ &\beta \Rightarrow \gamma \\ &\forall \delta . \text{ if } \alpha \vdash \delta \text{ and } \delta \Rightarrow \neg \gamma \text{ then } \beta \vdash \delta. \end{aligned}$$

For any rule $\alpha \Rightarrow \gamma$, and any wff w ,
 derive $\alpha \wedge w \Rightarrow \gamma$ if it is supported,
 otherwise, derive $\alpha \wedge \neg w \Rightarrow \gamma$.

or by convention II:

$$\begin{aligned} &\text{derive } \alpha \rightarrow \gamma \\ &\text{if } \alpha \Rightarrow \gamma \\ &\text{and } K \vdash \alpha \\ &\text{and } \forall \beta . \text{ if } K \vdash \beta \text{ and } \beta \Rightarrow \neg \gamma \\ &\text{then } \alpha \vdash \beta. \end{aligned}$$

(the actual procedure iterates).

If K is all you know, then conclude p if you can derive $K \Rightarrow p$.

Synopsis of Geffner (*SEP*, 88)

synopsis by Loui

Start with defaults, D ,

$$\begin{aligned} E(x) &\stackrel{d}{\dashv} Gray(x) \\ RoyalE(x) &\stackrel{d}{\dashv} \neg Gray(x) \end{aligned}$$

and a background context, L ,

$$\forall x. RoyalE(x) \rightarrow E(x)$$

and evidence, E ,

$$RoyalE(Clyde).$$

If E is all the evidence, and p is a putative conclusion, then the object is to derive $E \vdash p$.

Part of the proof theory is inherited from Adams and is ϵ -sound:For any A, B, h :

1. (Defaults) If $A \stackrel{d}{\dashv} h$ then $A \vdash h$.
2. (Logic) If $L \cup A \vdash h$ then $A \vdash h$.
3. (Triangularity) If $A \vdash h$ and $A \vdash B$ then $A, B \vdash h$.
4. (Bayes) If $A \vdash B$ and $A, B \vdash h$ then $A \vdash h$.
5. (Disjunction) If $A \vdash h$ and $B \vdash h$ then $A \vee B \vdash h$.

Irrelevance is the best way to introduce sentences on the left-hand side of \vdash .It depends on $I(h, e; A)$ which says that e is irrelevant to h in the context of A .

Roughly, this means that e does not provide an argument for $\neg h$ that was not already possible from A .

$I(\neg Gray(Clyde), E(Clyde); RoyalE(Clyde))$ because even though $E(Clyde)$ provides an argument for $Gray(Clyde)$, this argument was already possible from $RoyalE(Clyde)$.

- 6.1. (Explicit) If $a \stackrel{d}{\dashv} b$ and $I(\neg b, e; a)$ then $a, e \vdash b$.
- 6.2. (Implicit) If $a \stackrel{d}{\dashv} b$ and $\exists s$ s.t.
 - $a \vdash s$ and $a, e \vdash s$ and $I(\neg b, e; a \wedge s)$
 then $a, e \vdash b$.

Synopsis of Haugh (AAAI, 88)

synopsis by I. Flanigan

Introduce five kinds of links:

- isa_x for explicitly declared *isa*.
- isa_p for possible links.
- ab_x for explicitly declared *ab*.
- ab_d for directly *ab*.
- ab_i for inherited *ab*.

Axioms:

- A1. (*X* is related in some way to *Q*)
 $isa(X, Q) \equiv isa_x(X, Q) \vee \exists P(isa(X, P) \wedge isa_x(P, Q) \vee \neg ab(X, P, Q))$
- A2. (*X* is abnormal to *Q* through *P*)
 $ab(X, P, Q) \equiv ab_d(X, P, Q) \vee ab_i(X, P, Q) \vee ab_c(X, P, Q) \vee ab_x(X, P, Q)$
- A3. (explicit link contradicts possible *isa*)
 $ab_d(X, P, Q) \equiv isa(X, P, Q) \wedge isa_x(P, Q) \wedge isa_x(X, \neg Q)$
- A4. (inherited abnormalities)
 $ab_i(X, Q, R) \equiv \exists P(isa(X, P) \wedge [ab_d(P, Q, R) \vee ab_c(P, Q, R) \vee ab_x(P, Q, R)])$
- A5. (Nixon diamonds)
 $ab_c(X, P, Q) \equiv \exists R(isa(X, P) \wedge isa_x(P, Q) \wedge isa(X, R) \wedge isa_x(R, \neg Q) \wedge \neg ab_{dix}(X, P, Q) \wedge \neg ab_{dix}(X, R, \neg Q))$
- A6. (derivation of *ab_dix*)
 $ab_{dix}(X, P, Q) \equiv ab_d(X, P, Q) \vee ab_x(X, P, Q) \vee \exists S(isa(X, S) \wedge [ab_d(S, P, Q) \vee ab_x(S, P, Q)])$
- A5'. (ambiguity-propagating)
 $ab_c(X, P, Q) \equiv \exists R(isa_p(X, P) \wedge isa_x(P, Q) \wedge isa_p(X, R) \wedge isa_x(R, \neg Q) \wedge \neg ab_{dix}(X, P, Q) \wedge \neg ab_{dix}(X, R, \neg Q))$
- A7. (possible *isas* for A5')
 $isa_p(X, R) \equiv isa_x(X, R) \vee \exists T(isa_p(X, T) \wedge isa_x(T, R) \wedge \neg ab_{dix}(X, T, R))$

For *Clyde*, start with

- $isa_x(Clyde, RoyalE)$
- $isa_x(RoyalE, E)$
- $isa_x(E, Gray)$
- $isa_x(RoyalE, \neg Gray)$

and derive

- $isa(Clyde, E)$ (Axiom A1)
- $ab_d(RoyalE, E, Gray)$ (Axiom A3)
- $ab_i(Clyde, E, Gray)$ (Axiom A4)

Synopsis of Horty and Thomason(AAAI, 88)

synopsis by Loui and H. Smith Taylor

Let a path σ in a network have two parts, μ , the maximal strict end segment, and δ , the remainder.

We say $\Gamma \triangleright \sigma$, or just $\triangleright \sigma$, when path σ is inheritable in a network Γ .

The simple rules are:

Case A: $\sigma \neq \delta$ and $\sigma \neq \mu$. Then $\triangleright \sigma$ iff $\triangleright \delta$ and $\triangleright \mu$.

Case B: $\sigma = \mu$. Then $\triangleright \sigma$ iff each link in σ is in Γ .

Case C-I: $\sigma = \delta$ and σ is a direct link. Then $\triangleright \sigma$ iff σ is in Γ .

Case C-II: is too complex to reproduce here, but here is an example of its application:

$$\begin{array}{l} \sigma \text{ is } Clyde \longrightarrow RoyalE \not\rightarrow Gray \\ \mu \text{ is null,} \\ \delta \text{ is } Clyde \longrightarrow RoyalE \not\rightarrow Gray \\ \sigma \text{ is of the form } \pi(Clyde, RoyalE) \not\rightarrow Gray. \end{array}$$

a. $\triangleright \pi(Clyde, RoyalE)$? yes, by Case B.

b. $RoyalE \not\rightarrow Gray \in \Gamma$? yes, direct $\not\rightarrow$ link.

c. $Clyde \longrightarrow Gray \notin \Gamma$? yes.

$Gray \notin \kappa_{\Gamma}(Clyde) = \{ Clyde, RoyalE, E \}$? yes.

d. only need to worry about

$\triangleright \pi(Clyde, RoyalE, E)$ because of $E \longrightarrow Gray$ and $Gray \in \kappa_{\Gamma}(Gray)$.

But this is ok because we have

$\triangleright \pi(Clyde, RoyalE, E)$ and $RoyalE \not\rightarrow Gray \in \Gamma$ and $Gray \in \kappa_{\Gamma}(Gray)$.

So $\triangleright \pi(Clyde, RoyalE, Gray)$. $Clyde$ is not $Gray$.

There is evidence, in EK , and there are relations between sentences, of the supporting kind, \succ , and of the interfering kind, \vdash . Defeasible conclusions find their way into DK .

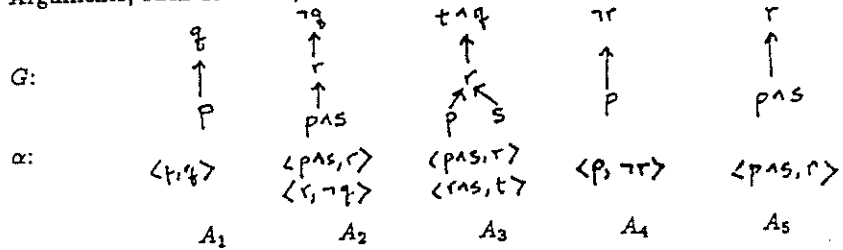
Loui

Loui

For example,

$EK = \{p, s, t \rightarrow q, \dots \text{entailments} \dots\}$
 $p \succ q$
 $p \wedge s \succ r$
 $r \succ \neg q$
 $r \wedge s \vdash t$
 $p \vdash r$

Arguments, such as below, are constructed:



A_3 is an argument because
 it's an acyclic digraph with unique sink.
 no two node labels are logically equivalent under EK .
 sources are in EK .
 nodes are jointly consistent with EK .
 all rules in α are used in G .
 if a source, s , is in the support of a node, r , then $\langle support(r), r \rangle \in \alpha$.
 $\langle support(r), r \rangle \in \alpha$ or $support(r), EK \vdash r$.
 if $\langle r \wedge s, t \rangle \in \alpha$ and $not(r \wedge s \succ t)$, then t is the sink.

A_3 and A_4 are interfering arguments because they use interfering reasons. Others are supporting arguments. A_4 is a counter-argument to A_2, A_3 , and A_5 , each, because $\neg r$ is not EK -consistent with some node in each. A_4 disagrees with A_5 because EK -inconsistency occurs at the sink.

A_2 uses more evidence than A_1 because of the logical strength of the sources. A_1 is more direct than A_2 because some radius in G_1 can be obtained from some radius in G_2 by deletion, dilution, and negation of the terminus, with at least one deletion; and not vice versa. A_3 is more specific than A_2 because they have respective rules $\langle r \wedge s, t \rangle$ and $\langle r, q \rangle$ where $EK \cup \{t, q\}$ is inconsistent.

A_3 reflects A_2 because at least one of more evidence, more specificity, and more directness holds.

A_5 is a sub-argument of A_2 , but not of A_3 . A_5 defeats A_4 because it is a counter-argument of A_4 and it reflects every sub-argument of A_4 with which it disagrees.

Following Pollock:

- any interfering argument is a level-0 I-argument.
- any supporting argument is a level-0 I-argument and level-0 S-argument.
- a supporting argument is a level-($n+1$) S-argument iff
 $not(\exists$ a level- n I-argument that is a counter-argument to it).
- an argument is a level-($n+1$) I-argument iff
 $not(\exists$ a level- n I-argument that defeats it).

q is justified because there is some argument, A_1 , s.t. \exists level- m beyond which it is always an S-argument.

$q \in DK$ because it is justified and does not belong to a minimal inconsistent set, all of whose members are justified. DK is also stipulated to be closed under \vdash .

Synopsis of Neufeld (*Uncertainty Workshop, 88*)
synopsis by J. Mehta

$a \rightarrow b$ means $p(b|a) > p(b)$
 $a \Rightarrow b$ means $p(b|a) = 1$
 $a \dashv\rightarrow b$ means $p(\neg b|a) > p(\neg b)$
 $a \dashv\Rightarrow b$ means $p(\neg b|a) = 1$

UCi = unconditionally independent
 Ci = conditionally independent
 Cf = confirms

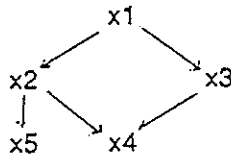
a Cf b if $p(b|a) > p(b)$
 a UCi b if $p(a|b) = p(a)$
 a Ci $b | c$ if $p(a|bc) = p(a|c)$

Ci in influence diagram (Pearl)

Defn: x Ci $y | S$, i.e. separated by S , where $S = \{a_1, a_2, \dots, a_m\}$
 if all paths between x and y are blocked

Defn: a path is blocked if it has
 (i) atleast one head-tail $\rightarrow a \rightarrow$ or tail-tail $\leftarrow a \leftarrow$ and $a \in S$
 (ii) atleast one head-head $\rightarrow a \leftarrow$ and $a \in S$

example:



$x2$ Ci $x3 | \{x1\}$
 since path $x2 \leftarrow x1 \rightarrow x3$ is blocked as $x1 \in \{x1\}$
 and path $x2 \rightarrow x1 \leftarrow x3$ is blocked as $x4 \in \{x1\}$

$x2$ \neg Ci $x3 | \{x4\}$ or $\{x1, x4\}$
 since path $x2 \leftarrow x1 \rightarrow x3$ is not blocked as $x1 \notin \{x4\}$
 and path $x2 \rightarrow x4 \leftarrow x3$ is not blocked as $x4 \in \{x1, x4\}$

symmetry

a	Cf	b
b	Cf	a

negation

a	Cf	b
$\neg a$	Cf	$\neg b$

logical resolution

a	Cf	c
b	Cf	d
a	\vdash	b
c	Cf	ab

logical strengthening

a	\vdash	b
b	$\dashv\vdash$	a
a	Cf	b

probabilistic resolution

$\exists c$	a	Cf	c
	b	Cf	c
	a	Ci	$b c$
	a	Cf	b

conjunction of events

irrelevance

a	Cf	c
a	Ci	$b c$
a	Cf	bc

relevance

a	Cf	c
b	Cf	c
a	Ci	$b c$
ab	Cf	c

lemmas

specificity

$\neg a$	Cf	b
a	Cf	c
b	$\dashv\vdash$	c
a	Cf	$\neg bc$

Synopsis of Nute (SEP, 88)

synopsis by J. Mehta

Nute

absolute rule $A \rightarrow p$

defeasible rule $A \Rightarrow p$

defeater rule $A \rightarrow p$

rule set R

literal set K

partial order on R \leq

theory $T = \langle R, K, \leq \rangle$

" $r_1 \leq r_2$ " for " r_2 superior to r_1 "

"e p" for "evidently, p"

proof line $L = \langle q, v, A \rangle$

where, q = sentence

v = + proved | - disproved

A = premise set for this triple.

proof of $\langle e p, +, K \rangle$ from $\langle R, K \rangle$

proof = $\{L\}_M \langle e p, +, K \rangle$

proved absolutely from A

$$C1 \quad \frac{q \in A}{\therefore \langle q, +, A \rangle}$$

$$C2 \quad \frac{\exists B \rightarrow q \text{ such that} \\ \forall b \in B \langle b, +, A \rangle}{\therefore \langle q, +, A \rangle}$$

disproved absolutely from A

$$C3 \quad \frac{q \in A \\ \forall B \rightarrow q \text{ such that} \\ \exists b \in B \langle b, -, A \rangle}{\therefore \langle q, -, A \rangle}$$

proved evidently from K

$$C4 \quad \frac{\langle q, +, K \rangle}{\therefore \langle e q, +, K \rangle}$$

$$C5 \quad \begin{array}{l} a) \neg q \in K \\ b) \exists B \rightarrow q \text{ such that} \\ \quad \forall b \in B \langle e b, +, K \rangle \\ c) \forall C \rightarrow \neg q \text{ such that} \\ \quad \exists c \in C \langle e c, -, K \rangle \end{array}$$

$$\therefore \langle e q, +, K \rangle$$

$$C6 \quad \begin{array}{l} a) \neg q \in K \\ b) \exists B \Rightarrow q \text{ such that} \\ \quad \forall b \in B \langle e b, +, K \rangle \\ c) \forall C \rightarrow \neg q \text{ such that} \\ \quad \exists c \in C \langle e c, -, K \rangle \\ d) \forall C \Rightarrow \neg q \text{ such that} \\ \quad (i) \exists c \in C \langle c, -, K \rangle \\ \quad \text{or, (ii) } \forall c \in C \langle c, +, B \rangle \\ \quad \exists b \in B \langle b, -, C \rangle \end{array}$$

$$\therefore \langle e q, +, K \rangle$$

disproved evidently from K

$$C7 \quad \frac{\langle q, -, K \rangle \\ \langle \neg q, +, K \rangle}{\therefore \langle e q, -, K \rangle}$$

$$C8 \quad \begin{array}{l} a) \langle q, -, K \rangle \\ b) \forall B \rightarrow q \text{ such that} \\ \quad (i) \exists b \in B \langle e b, -, K \rangle \\ \quad \text{or, (ii) } \exists C \rightarrow \neg q \text{ such that} \\ \quad \quad \forall c \in C \langle c, +, B \rangle \\ c) \forall B \Rightarrow q \text{ such that} \\ \quad (i) \exists b \in B \langle e b, -, K \rangle \\ \quad \text{or, (ii) } \exists C \rightarrow \neg q \text{ such that} \\ \quad \quad \forall c \in C \langle e c, +, K \rangle \\ \quad \text{or, (iii) } \exists C \Rightarrow \neg q \text{ such that} \\ \quad \quad \forall c \in C \langle c, +, K \rangle \\ \quad \text{and, (i) } \exists c \in C \langle c, +, K \rangle \\ \quad \quad \text{or, (ii) } \forall b \in B \langle b, +, C \rangle \end{array}$$

$$\therefore \langle e q, -, K \rangle$$

Nute

Synopsis of Padgham(AAAIL, 88)

synopsis by Loui

Clyde is represented by:

$$\begin{aligned}
 \text{Core}(\text{Royal}E) &\sqsupseteq \text{Core}(E) \\
 \text{Core}(\text{African}E) &\sqsupseteq \text{Core}(E) \\
 \text{Default}(E) &\sqsupseteq \text{Core}(\text{Gray}) \\
 \text{Core}(\text{Royal}E) \sqcup \text{Core}(\text{Gray}) &\Rightarrow \kappa \\
 \text{Core}(\text{Royal}E) &\in \text{NOT}(\text{Core}(\text{Gray})) \\
 \text{Core}(\text{Gray}) &\in \text{NOT}(\text{Core}(\text{Royal}E)) \\
 \text{Desc}(\text{Clyde}) &\sqsupseteq \text{Core}(\text{African}E) \\
 \text{Desc}(\text{Clyde}) &\sqsupseteq \text{Core}(\text{Royal}E)
 \end{aligned}$$

where κ stands for inconsistency in at least one feature value.

\sqsupseteq stands for chaining relationships in the diagram. ϵ stands for the default assumption that the object at the head of the path is not only a member of some class, but a typical member of that class. $\in \text{NOT}$ can terminate a path and indicate that members of one class are NOT members of another class.

The inferences are

$$\begin{array}{l}
 \text{Desc}(\text{Clyde}) \sqsupseteq \\
 \quad \text{Core}(\text{African}E) \sqsupseteq \\
 \quad \quad \text{Core}(E) \epsilon \\
 \quad \quad \quad \text{Default}(E)
 \end{array}$$

and

$$\begin{array}{l}
 \text{Desc}(\text{Clyde}) \sqsupseteq \\
 \quad \text{Core}(\text{Royal}E) \sqsupseteq \\
 \quad \quad \text{Core}(E)
 \end{array}$$

and

$$\begin{array}{l}
 \text{Desc}(\text{Clyde}) \sqsupseteq \\
 \quad \text{Core}(\text{Royal}E) \in \\
 \quad \quad \text{NOT}(\text{Core}(\text{Gray}))
 \end{array}$$

Synopsis of Pollock (*Cognitive Science*, 87)

synopsis by Loui

- (R1): a 's tend to be b 's and x is an a IS-REASON-FOR x is a b .
 (R2): c 's tend to be $\neg b$'s and x is an c IS-REASON-FOR x is a $\neg b$.
 (R3): c is a subset of a IS-REASON-FOR undercutting (R1).

- (R1) defeats (by rebuttal) (R2).
 (R2) defeats (by rebuttal) (R1).
 (R3) defeats (by undercutting) (R1).

Note that rebuttal is always symmetric and undercutting is always asymmetric.

Arguments can be linear collections of reasons, but there are also non-linear arguments:

e.g.

< context, assertion, justification, dependency >

1. $\langle \{RoyalE(Clyde)\}, E(Clyde), PREMISE, 0 \rangle$
2. $\langle \{RoyalE(Clyde)\}, RoyalE(Clyde) \rightarrow E(Clyde), FOUNDATION, 0 \rangle$
3. $\langle \{RoyalE(Clyde)\}, E(Clyde), (deductive) REASON, \{1, 2\} \rangle$
4. $\langle \{RoyalE(Clyde)\}, Gray(Clyde), (defeasible) REASON, \{3\} \rangle$
5. $\langle \{\}, RoyalE(Clyde) \rightarrow Gray(Clyde), CONDITIONALIZATION, \{4\} \rangle$

All arguments are level-0 arguments.

An argument is level- $(i + 1)$ iff

no level i argument defeats any of its reasons.

An argument is ultimately undefeated iff

$\exists m. \forall n > m.$

the argument is level- n .

Collective defeat occurs for each member of Γ when

there are undefeated arguments for each member of Γ and

$\forall P \in \Gamma.$

$\exists \Gamma_P,$ a finite subset of Γ s.t.

$\Gamma_P \vdash \neg P.$

Synopsis of Poole (IJCAI 85)

synopsis by Loui

Facts, F , are divided into necessary facts, F_N , and contingent facts, F_C .
 F_C is a subset of possible contingent facts, F_P .

For example,

$$F_N = \{\forall x. RoyalE(x) \rightarrow E(x)\}$$

$$F_C = \{RoyalE(Clyde)\}$$

$$F_P = \{RoyalE(Clyde), \neg RoyalE(Clyde), E(Clyde), \dots\}$$

There are defeasible rules in D .

A Theory, $T_1 = \langle D_1, g_1 \rangle$, such as

$$T_1 = \{\text{ASSUME } E(x) \rightarrow Gray(x)\}$$

which explains $Gray(Clyde)$

is more general than a theory, $T_2 = \langle D_2, g_2 \rangle$, such as

$$T_2 = \{\text{ASSUME } RoyalE(x) \rightarrow \neg Gray(x)\}$$

which explains $\neg Gray(Clyde)$

just in case

$\exists f \in F_P$ s.t.

$$f \cup D_1 \cup F_N \vdash g_1$$

(f is adequate for g_1)

$$\text{not}(f \cup D_2 \cup F_N \vdash g_2)$$

(f is not adequate for g_2)

$$\text{not}(f \cup D_2 \cup F_N \vdash g_1)$$

(f is non-trivial)

$T_1 = \langle D_1, g_1 \rangle$ is more specific than $T_2 = \langle D_2, g_2 \rangle$

just in case

S_2 is more general than S_1 and it is not the case that S_1 is more general than S_2 .

T_2 actually is more specific than T_1 because of $f = \{E(Clyde)\}$.

Sandewall

Syntax of Propositions

$$isax(x, y, s) \quad isa(x, y, s) \quad precl(x, y, z, x) \quad cntr(x, y, z, s)$$

where $x, y,$ and z are nodes, and s is either of the symbols $+$ or $-$. $Isax(x, y, s)$ is used to indicate given facts (from "is-axiom"), while isa represents both given and derived facts (see rule 1).

Inference Rules (given an assertion set Γ that initially contains only axioms)

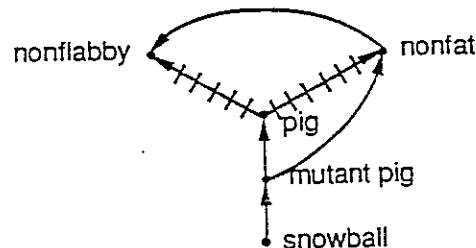
- | | |
|---|---|
| <p>1) if $isax(x, y, s) \in \Gamma$ then add to Γ:
 $isa(x, y, s)$</p> <p>2) if $isa(x, y, s) \in \Gamma$ then add to Γ:
 $\neg isa(x, y, -s)$</p> <p>3) if the following are members of Γ:
 $isa(x, y, +)$ $isa(y, z, s)$
 and the following are not in Γ:
 $isa(x, z, -s)$ $cntr(x, y, z, s)$
 then add to Γ:
 $isa(x, z, s)$ $precl(x, y, z, s)$
 $\neg cntr(x, y, z, s)$</p> <p>4) if the following are members of Γ:
 $precl(x, y, z, s)$ $isa(x, v, +)$
 $isa(v, y, +)$</p> | <p>then add to Γ:
 $precl(x, v, z, s)$</p> <p>5) if the following are members of Γ:
 $precl(x, y, z, s)$ $isa(y, z, -s)$
 then add to Γ:
 $\neg isa(y, z, -s)$</p> <p>6) if the following are members of Γ:
 $isa(x, y, +)$ $isa(y, z, +)$
 $isa(x, z, -)$ $isa(z, v, s)$
 and the following is not in Γ:
 $isax(y, v, s)$
 then add to Γ:
 $cntr(x, y, v, s)$</p> |
|---|---|

Extensions

Extensions are expanded versions of Γ that are closed under the above rules. Self-contradictory extensions are disallowed; plural distinct (but mutually exclusive) extensions are allowed, so long as each is self-consistent. Rules 5 and 6 purposely introduce contradictions into a theorem set in order to disqualify undesired types of extensions from consideration.

Example: Mutant pigs

- $isax(pig, nonfat, -)$
- $isax(pig, nonflabby, -)$
- $isax(mutant\ pig, nonfat, +)$
- $isax(nonfat, nonflabby, +)$
- $isax(snowball, mutant\ pig, +)$
- $isax(mutant\ pig, pig, +)$



$Isa(mutant\ pig, nonfat, -)$ can never be derived because it is contradicted by c. $Isa(snowball, nonfat, -)$ cannot be part of a consistent extension for more complex reasons (effects of Rule 1 are assumed):

- | | |
|---|-----------------|
| 1. $isa(snowball, pig, +)$ | e, f, Rule 3 |
| 2. $isa(snowball, nonfat, -)$ | 1, a, Rule 3 |
| 3. $precl(snowball, pig, nonfat, -)$ | " |
| 4. $precl(snowball, mutant, nonfat, -)$ | 3, e, f, Rule 4 |
| 5. $\neg isa(mutant, nonfat, +)$ | 4, c, Rule 5 |

Since 5 contradicts c, no extension containing statement 2 can be self-consistent.

What about $isa(snowball, nonflabby, -)$? There are at least two extensions: in one, $isa(snowball, nonflabby, -)$; in another, $isa(snowball, nonflabby, +)$. The "earlier" specificity of mutant $\rightarrow nonfat \rightarrow nonflabby$ (compared to mutant $\rightarrow pig \rightarrow \neg nonflabby$) does not appear to have an effect.

Comment

Unless a graph exactly matches a familiar pattern, learning what extensions may result from it requires traversing a large number of paths through "theorem space." Many such paths are eliminated only retroactively, by intentional generation of an inconsistency. This makes identifying extensions a tedious and unsure process — at least for humans.

Synopsis of Touretzky (*TMOIS*, 86)*synopsis by Loui*

Graphs have links such as $\langle +Penguin, +Bird \rangle$. Φ is the set of inheritable sequences. We define conditions of inheritability on Φ and look for fixed-points.

$conc(\Phi) = \{ \langle x, y \rangle : \langle x, \dots, y \rangle \in \Phi \}$.

Φ contradicts $\langle x_1, \dots, x_n \rangle$ iff $\exists i. \langle x_1, -x_i \rangle \in conc(\Phi)$.

e.g. $\langle +Clyde, -Gray \rangle$ contradicts $\langle +Clyde, +RoyalE, +E, +Gray \rangle$.

y is an intermediary to $\langle x_1, \dots, x_n \rangle$ in Φ iff

$\exists i. y = x_i$ or

\exists some sequence $\sigma \in \Phi$ from which $\langle x_1, \dots, x_i, y, x_{i+1} \rangle$ can be obtained by deletion.

e.g. $+RoyalE$ is an intermediary to $\langle +Clyde, +RoyalE, +E, +Gray \rangle$ in the obvious Φ .

Φ precludes $\sigma = \langle x_1, \dots, x_n \rangle$ iff

$\exists y. \langle y, -x_n \rangle \in \Phi$ and

y is an intermediary to σ in Φ .

$\sigma = \langle x_1, \dots, x_n \rangle$ is inheritable in Φ iff

$not(\Phi$ contradicts $\sigma)$ and

$not(\Phi$ precludes $\sigma)$ and

(choose one)

(double chaining):

$\langle x_1, \dots, x_{n-1} \rangle \in \Phi$ and $\langle x_2, \dots, x_n \rangle \in \Phi$.

(backward chaining):

$\langle x_1, x_2 \rangle \in \Phi$ and $\langle x_2, \dots, x_n \rangle \in \Phi$.

(forward chaining):

$\langle x_1, \dots, x_{n-1} \rangle \in \Phi$ and $\langle x_{n-1}, x_n \rangle \in \Phi$.

Φ is closed iff it contains every inheritable sequence in Φ .

Φ expands a graph iff Φ contains all the graph's links and Φ is closed.

Φ is grounded in a graph if every sequence in Φ not among the graph's links is inheritable in Φ .

Grounded expansions of a graph's links are like extensions.

Part III: Workshop Problems

- 1 Royal African Elephants
- 2 Jesse Jackson if not Michael Dukakis
- 3 Garfield and People
- 4 University Students
- 5 Adults under 22
- 6 Dancers and Ballerinas
- 7 Noticing Danger Mouse
- 8 Gullible Citizens
- 9 Fred after unaltered Yale shooting
- 10 Fred after altered Yale shooting
- 11 Fred doesn't get wet
- 12 Unrefined Big Blocks
- 13 Say Randy flies
- 14 Nixon's Political Motivation
- 15 Nixon's Anti-militarism
- 16 Lottery Paradox
- 17 Billboard's hot 100
- 18 Integers 80-100
- 19 Doctors are Medical?
- 20 Backwater CS PhD's
- 21 Tom's waking
- 22 Genetically Altered Pigs
- 23 Hernandez Tries to Homer
- 24 Hermann's Pennsylvania Dutch
- 25 Fiats are Fast?
- 26 Nice Guy Lawyers
- 27 Party for Friends
- 28 Museum Director's Testimony
- 29 Carl, Alice and Betty
- 30 Nute's Nixon Diamond

Benchmark Problems

1 Royal African Elephants

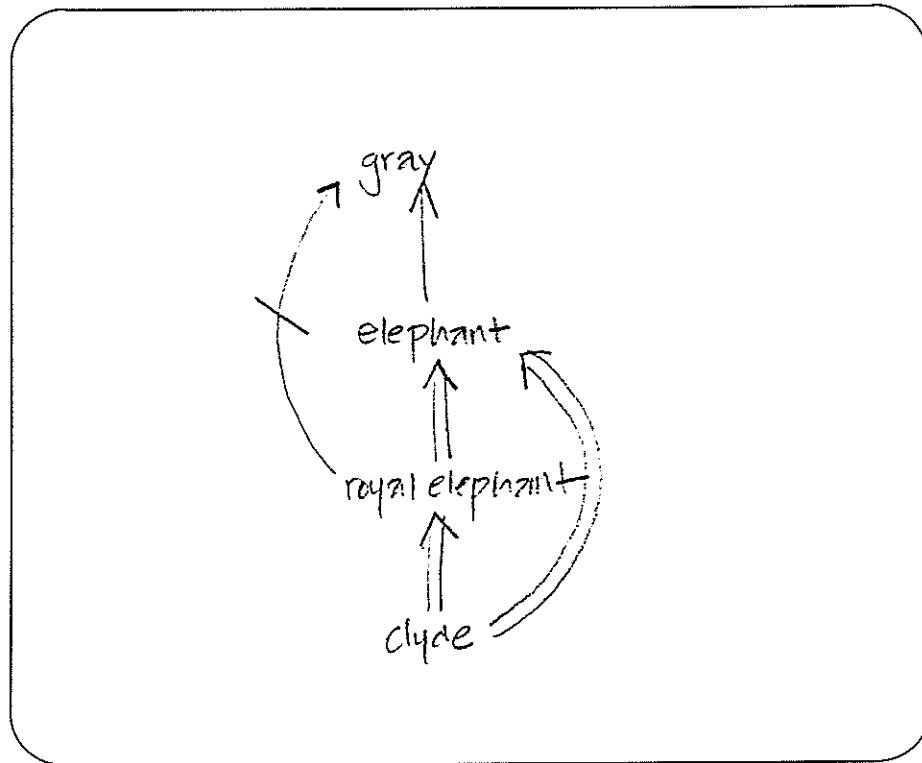
Locus of Preemption (Sandewall)

Elephants tend to be gray

Royal elephants tend to be non-gray

Clyde is a royal elephant and an african elephant

Is Clyde non-gray?



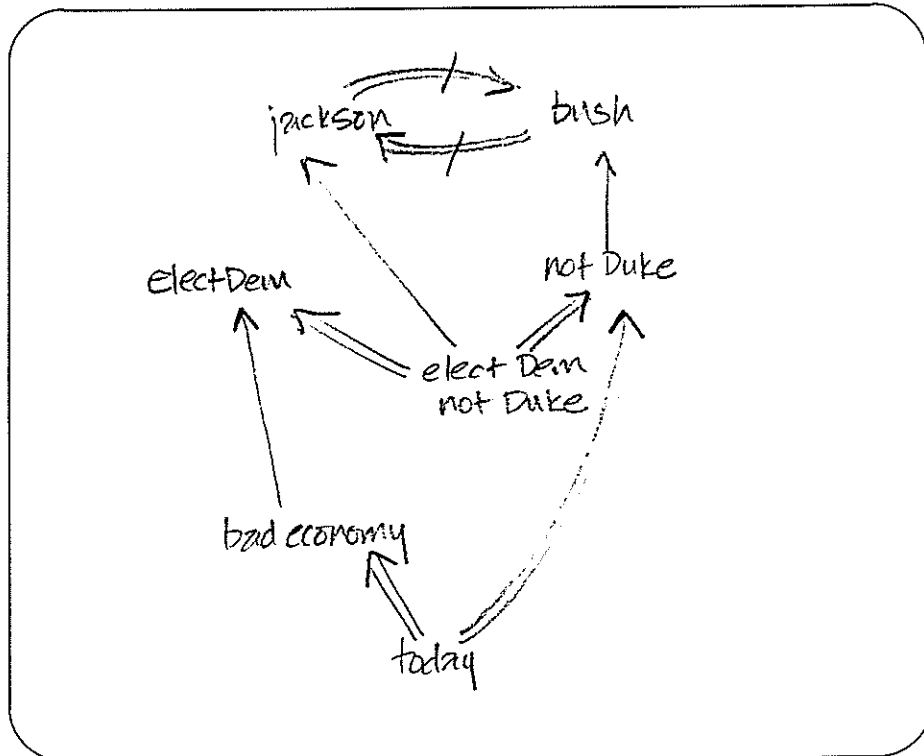
Benchmark Problems

2 Jesse Jackson if not Michael Dukakis

Irrelevance (Baker)

Likely, if the economy is bad, we'll elect a democrat
Likely, if we elect a democrat and it's not Dukakis, it'll be Jackson
Likely, if it is not Dukakis, it'll be Bush
The economy is bad
It's not Mike Dukakis
It can't be both Bush and Jackson

It is Jackson?



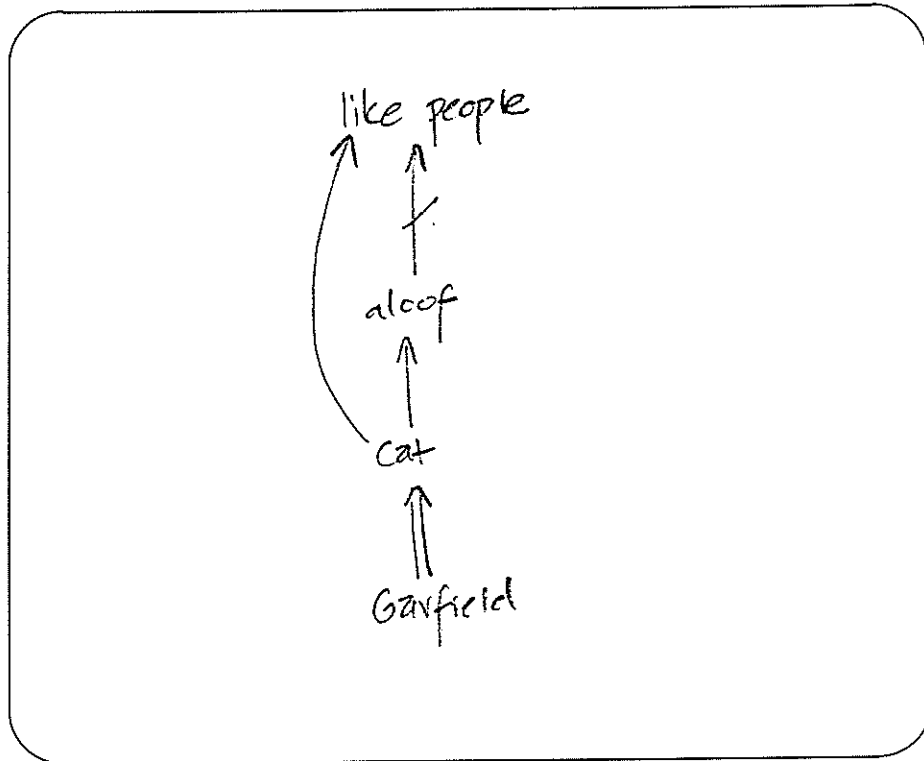
Benchmark Problems

3 Garfield and People

Directness (Loui)

Cats tend to be aloof
Aloofness tends to indicate dislike of people
Cats tend to like people
Garfield is a cat

Does Garfield like people?



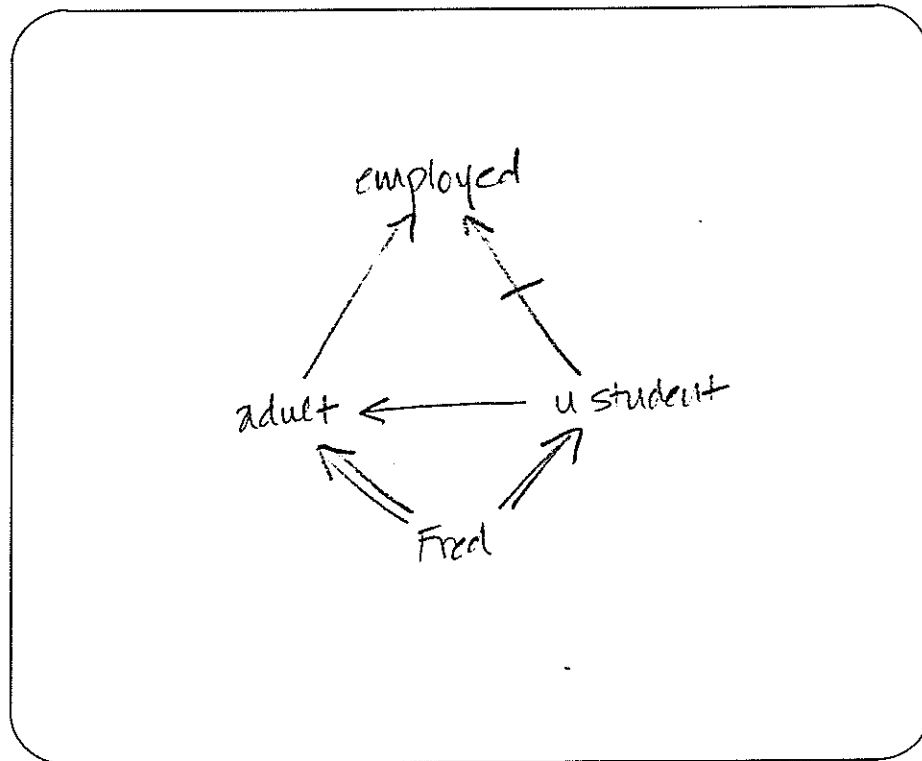
Benchmark Problems

4 University Students

Strong Specificity (Delgrande)

Adults tend to be employed
University students tend to be unemployed
University students tend to be adults
Fred is a university student and an adult

Is Fred unemployed?



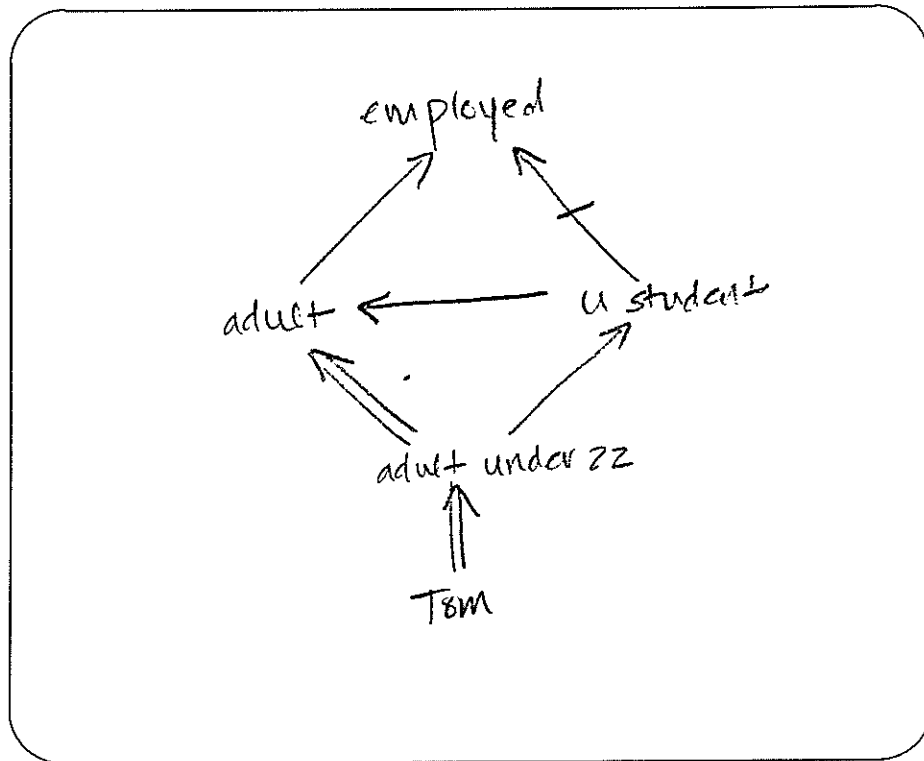
Benchmark Problems

5 Adults under 22

(Geffner-Pearl)

Adults tend to be employed
University students tend to be unemployed
University students tend to be adults
Adults under 22 tend to be university students
Tom is an adult under 22

Is Tom unemployed?



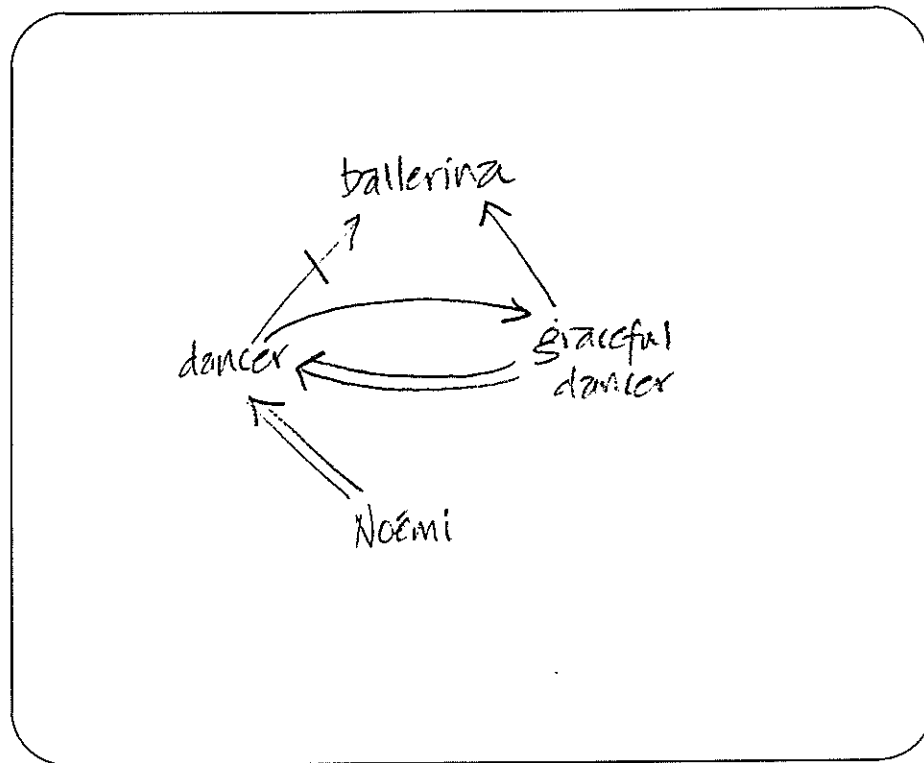
Benchmark Problems

6 Dancers and Ballerinas

Specificity versus Directness (Loui)

Dancers tend not to be ballerinas
Dancers tend to be graceful
Graceful dancers tend to be ballerinas
Noémi is a dancer

Ambiguous?



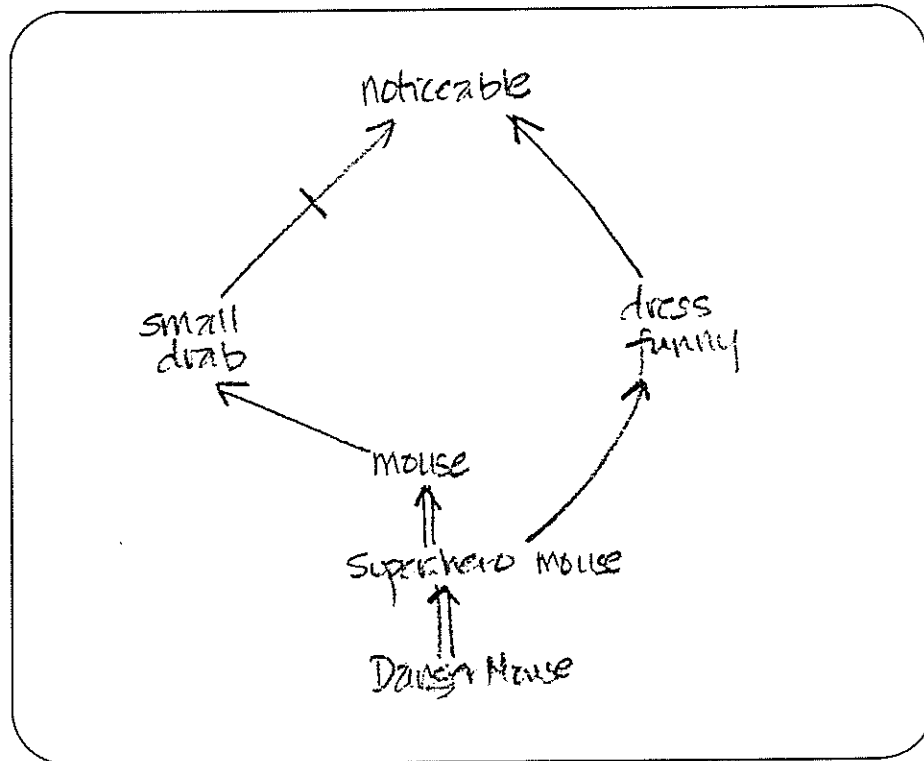
Benchmark Problems

7 Noticing Danger Mouse

Evidence without specificity (Loui)

Mice tend to be small and drab
Small and drab things tend to be not-noticeable
Super-hero mice tend to dress funny
Dressing funny tends to make one noticeable
Danger Mouse is a super-hero mouse

Is Danger Mouse noticeable?



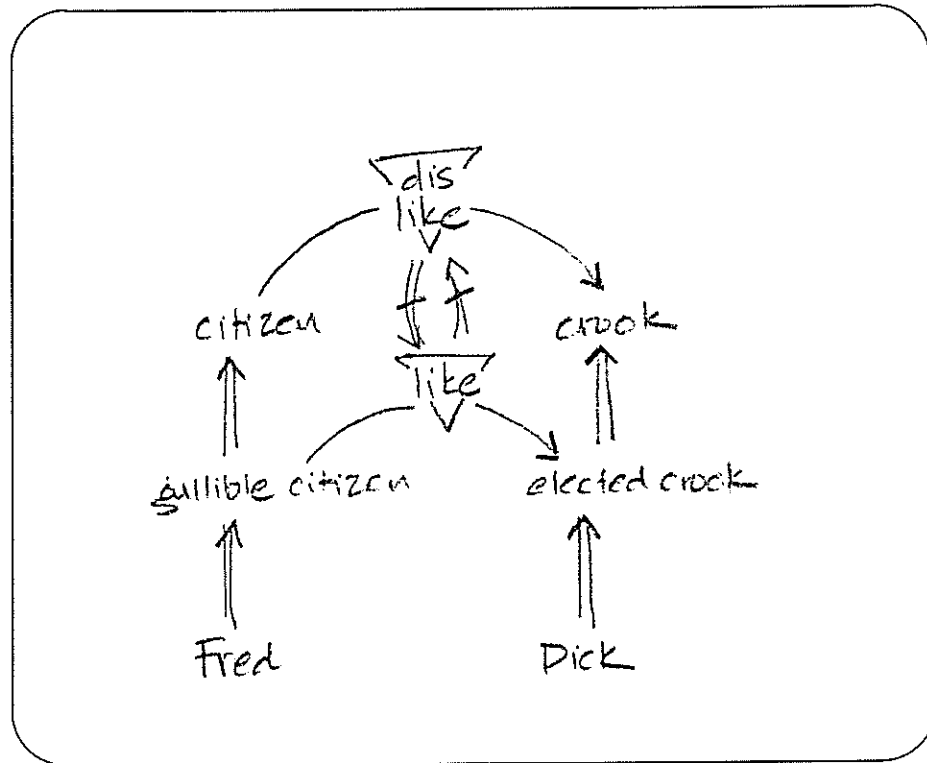
Benchmark Problems

8 Gullible Citizens

Relations(Touretzky)

Citizens tend to dislike crooks
Gullible citizens tend to like elected crooks
Fred is a gullible citizen
Dick is an elected crook

Is it the case that Fred likes Dick?



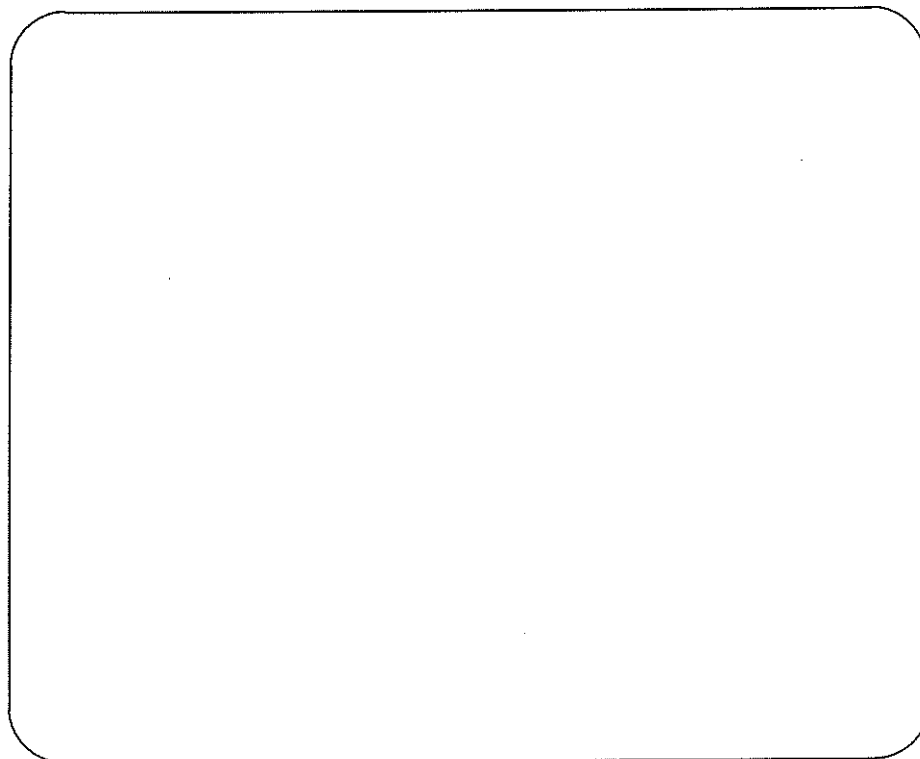
Benchmark Problems

9 Fred after unaltered Yale shooting

Temporal Preference (Hanks-McDermott)

Fred is alive in S_0
Guns are always loaded after loading events
Guns tend to persist in being loaded
Persons tend to persist in being alive
Firing a loaded gun always results in Fred's Death

Is Fred dead in $\langle fire \mid wait \mid load \mid S_0 \rangle$?



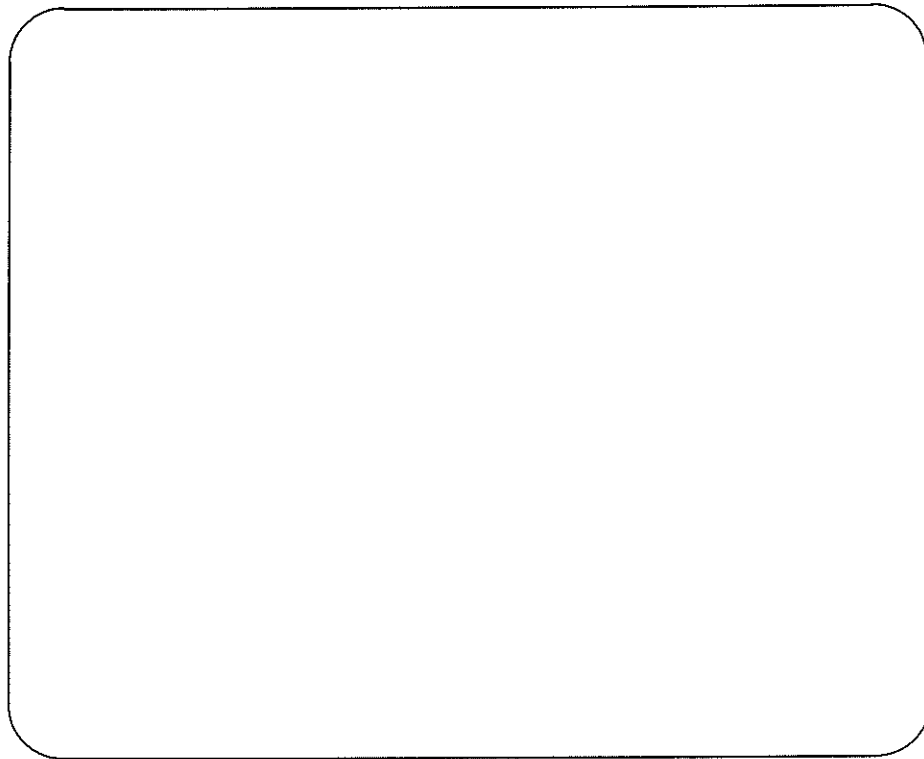
Benchmark Problems

10 Fred after altered Yale shooting

Solution to puzzle (Loui, Nute, Delgrande, Pearl)

Fred is alive in S_0
Guns are always loaded after loading events
Guns tend to persist in being loaded
Persons tend to persist in being alive
Firing a loaded gun tends to coerce Fred's Death

Is Fred dead in $\langle \text{fire} \mid \text{wait} \mid \text{load} \mid S_0 \rangle$?



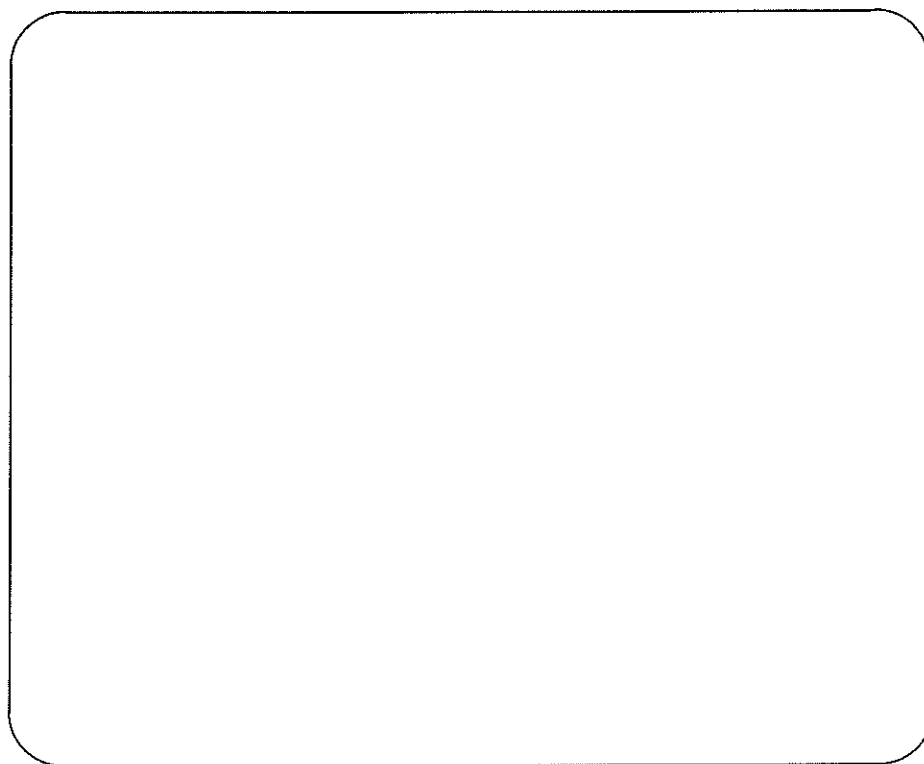
Benchmark Problems

11 Fred doesn't get wet

Non-monotonic chaining/solution to puzzle (Hanks-McDermott)

It's raining and Fred is alive in S_0
Guns are always loaded after loading events
Guns tend to persist in being loaded
Persons tend to persist in being alive
Raining tends to persist
Firing a loaded gun tends to coerce Fred's Death
Going out in the rain tends to result in getting wet

Is Fred wet in $\langle \text{leave} \mid \text{fire} \mid \text{wait} \mid \text{load} \mid S_0 \rangle$?



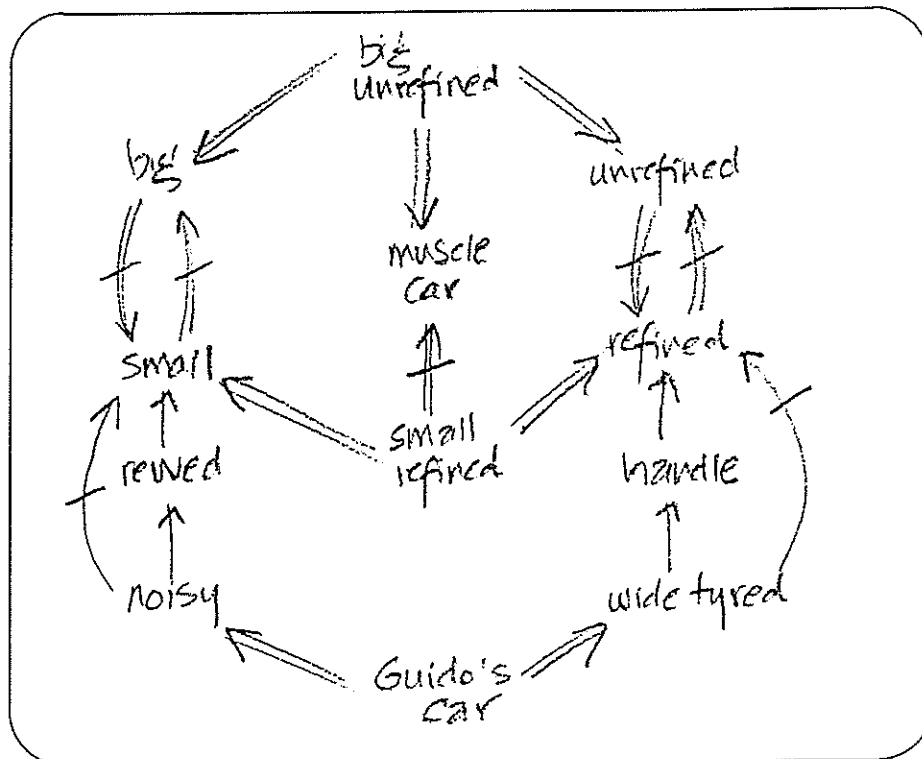
Benchmark Problems

12 Unrefined Big Blocks

Monotonic Chaining, Closure (Poole, Loui)

Noisy cars tend to be highly revved
Highly revved cars tend to be small
Noisy cars tend to have big block engines
Wide-tyred cars tend to handle well
Good handling tends to indicate refinement
Wide tyred cars tend to be unrefined cars
All big block unrefined cars are muscle cars
All small block refined cars are non-muscle cars
Guido's car is noisy with wide tyres

Is Guido's car a muscle car?



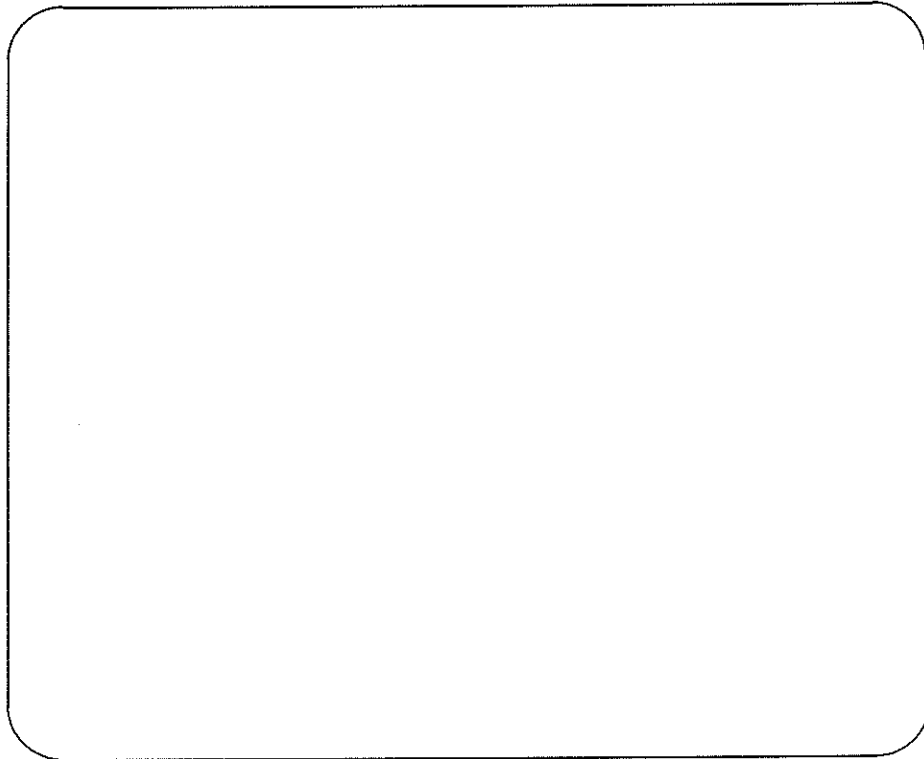
Benchmark Problems

13 Say Randy flies

Specificity without Contradiction (Poole)

The usual response for a bird is "flies"
The usual response for an emu is "runs"
Things at the emu farm tend to be emus
If soothsaying and the response for an x is y then say(x,y)
Presume soothsaying
Randy is at the emu farm

Say "Randy runs"?



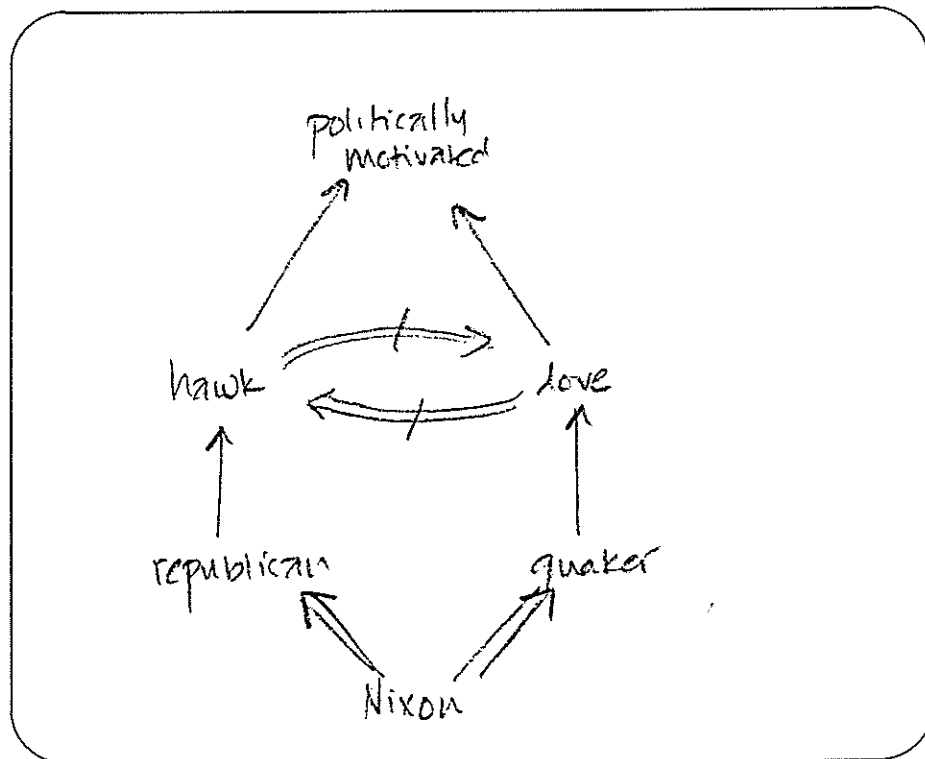
Benchmark Problems

14 Nixon's Political Motivation

Cases (Ginsberg)

Republicans tend to be hawks
Quakers tend to be doves
Nobody is both a hawk and a dove
Hawks tend to be politically motivated
Doves tend to be politically motivated
Nixon is a republican quaker

Is Nixon politically motivated?



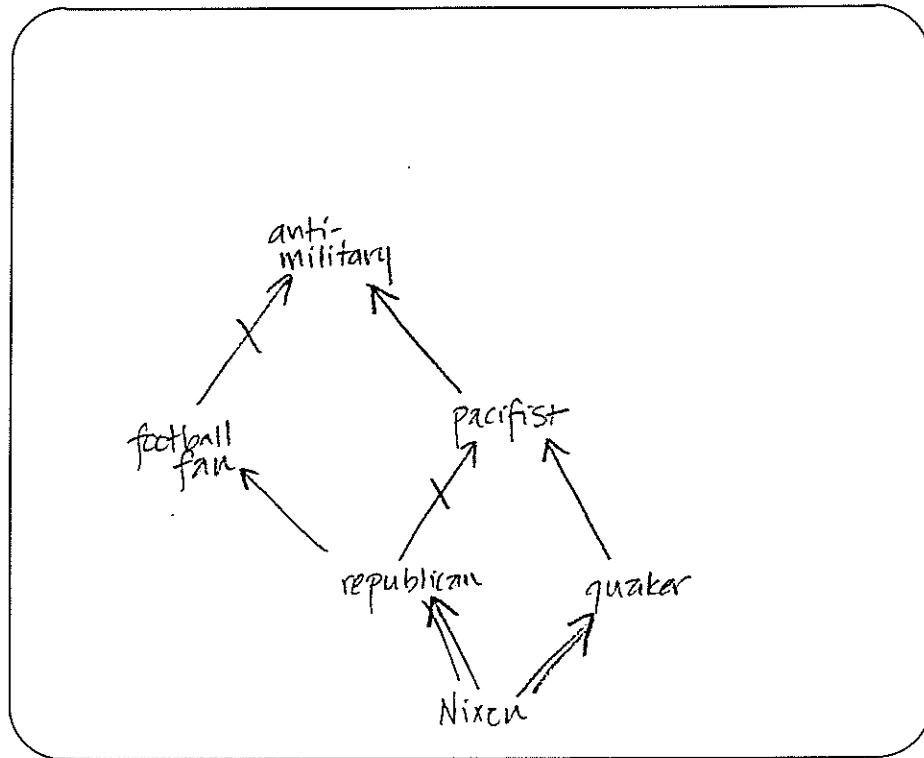
Benchmark Problems

15 Nixon's Anti-militarism

Cascaded Ambiguity (Touretzky)

Republicans tend to be non-pacifist and football fans
Quakers tend to be pacifist
Pacifist tend to be anti-military
Football fans tend to be non-anti-military
Nixon is a republican quaker

Ambiguous?



Benchmark Problems

16 Lottery Paradox

Probabilistic inconsistency, Collective defeat (Kyburg, Pollock)

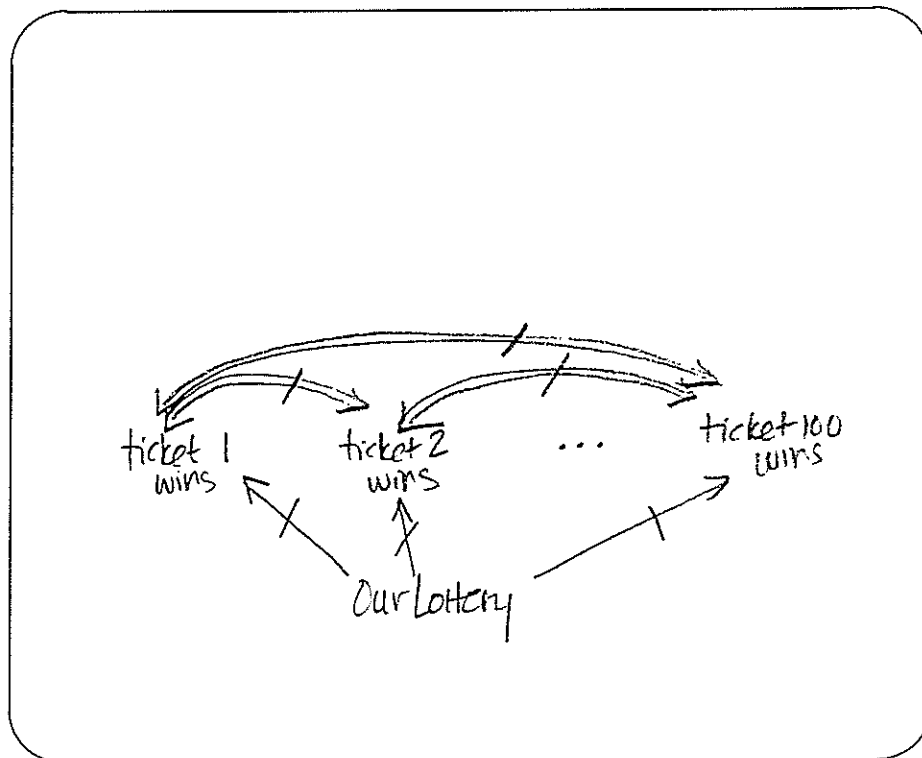
Probably, ticket 1 will not win

⋮

Probably, ticket 100 will not win

Ticket 1 or ... or ticket 100 will win

Any conclusion?



Benchmark Problems

17 Billboard's hot 100

Specificity versus Cases, Simpson's Paradox (Loui, Neufeld, Poole)

Records on the dance chart tend to be on the Hot100 chart

Records on the soul chart tend to be on the Hot100 chart

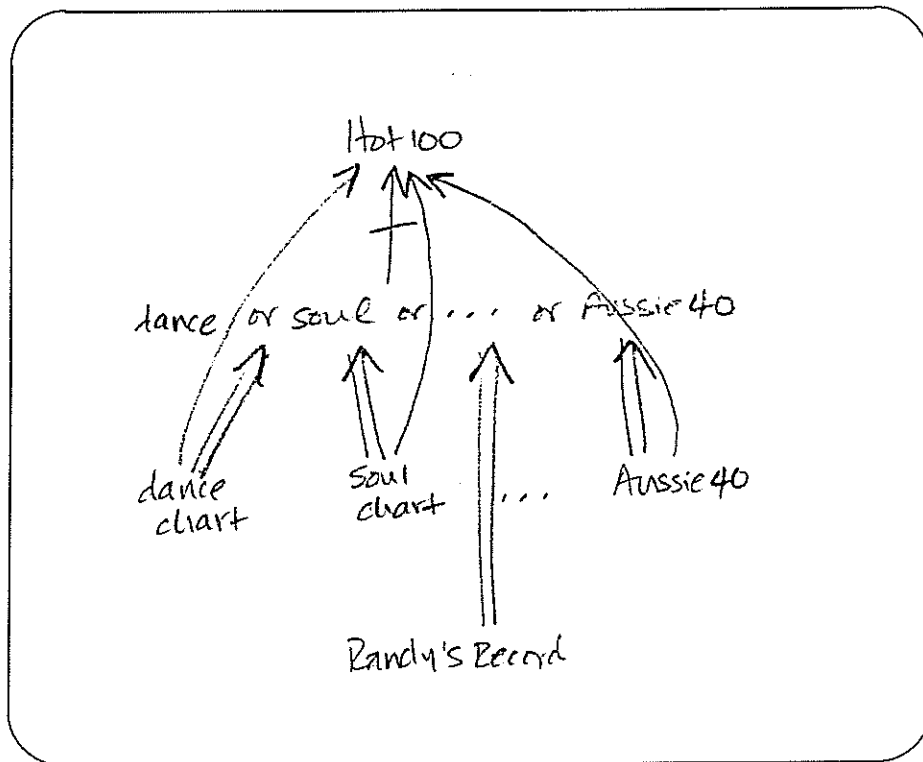
⋮

Records on the Aussie40 chart tend to be on the Hot100 chart

Records on one of the dance or soul or ... or Aussie40 chart
tend to be not on the Hot100 chart

Randy bought a record on the dance or soul or ... or Aussie40 chart

Is Randy's record not on the Hot100 chart?



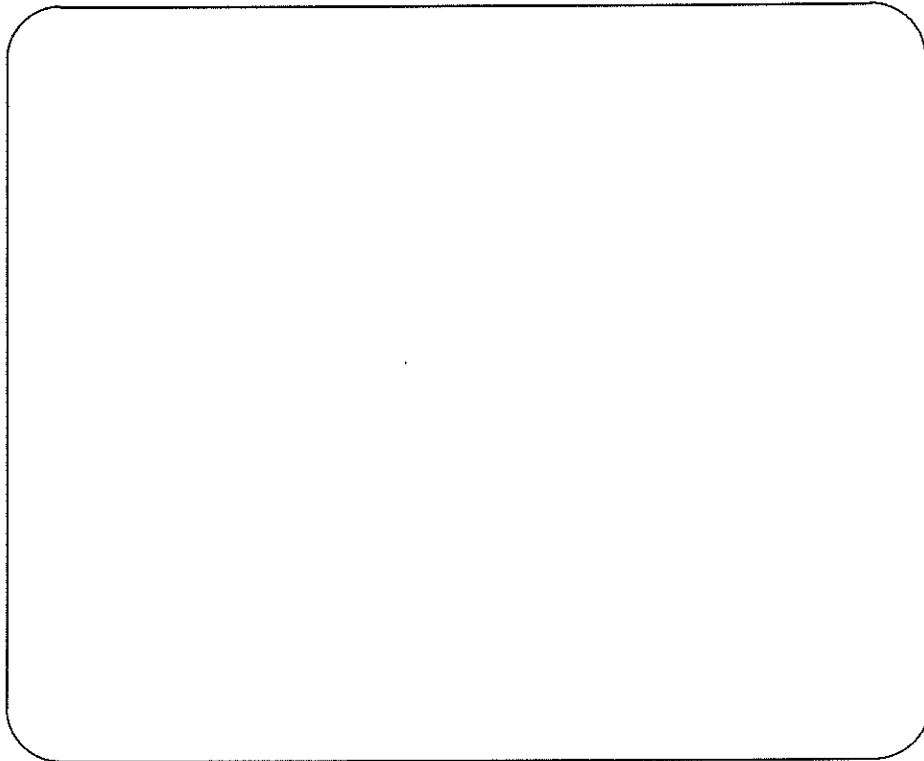
Benchmark Problems

18 Integers 80-100

Strong specificity versus Specificity: Statistical

Most $[0,100]$ are $[20,100]$
Most $[20,100]$ are $[36,100]$
Most $[36,100]$ are $[51,100]$
Most $[51,100]$ are $[61,100]$
Most $[61,100]$ are $[69,100]$
Most $[69,100]$ are $[75,100]$
Most $[75,100]$ are $[80,100]$
Most $[0,100]$ are not $[80,100]$
 x is in $[0,100]$ and in $[75,100]$

Is x in $[80,100]$?



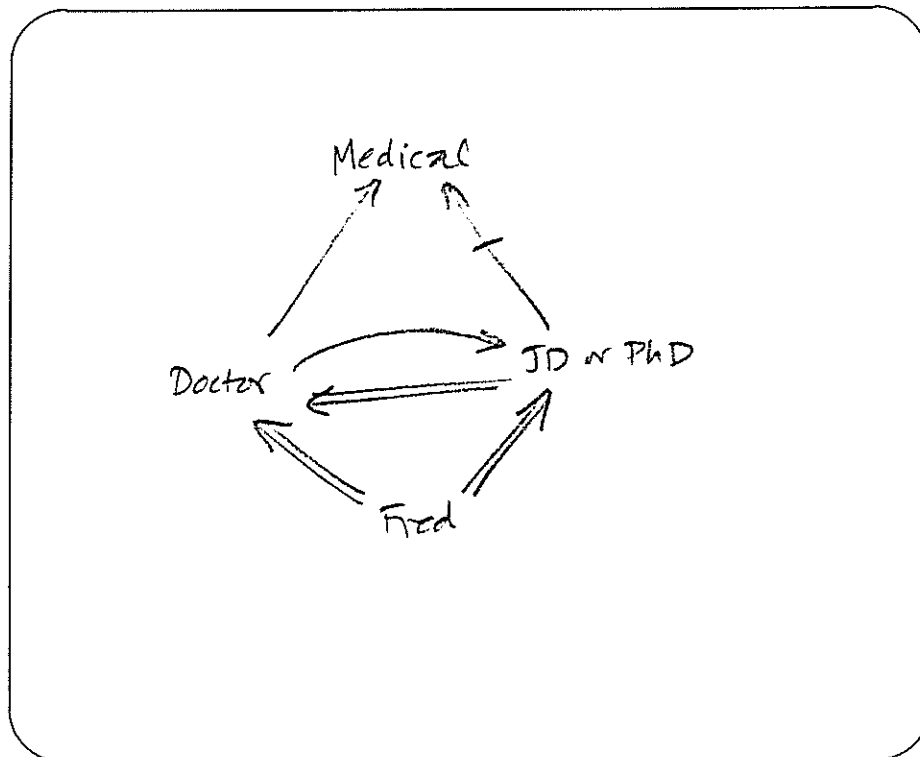
Benchmark Problems

19 Doctors are Medical?

Strong Specificity versus Specificity: Statistics and Implicature (Loui)

Doctors tend to be Medical
Most Doctors have the PhD or JD
Persons with the PhD or JD tend to be non-Medical
Fred has a PhD or JD

Is Fred non-Medical?



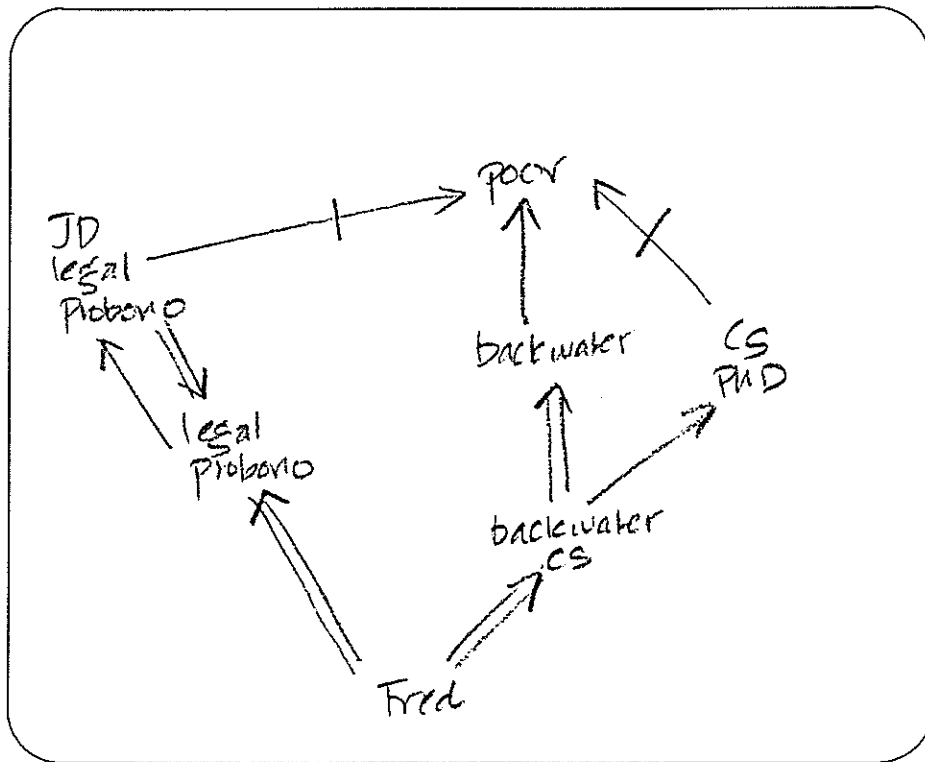
Benchmark Problems

20 Backwater CS PhD's

(Cross, Nute)

- Backwater teachers tend to be poor
- Backwater CS teachers tend to have CS PhD's
- CS PhD's tend to be rich
- Legal pro-bono people tend to be poor
- Legal pro-bono people tend to have JD's
- Legal pro-bono JD's tend to be rich
- Fred teaches CS at Backwater and does legal pro-bono work

Ambiguous?



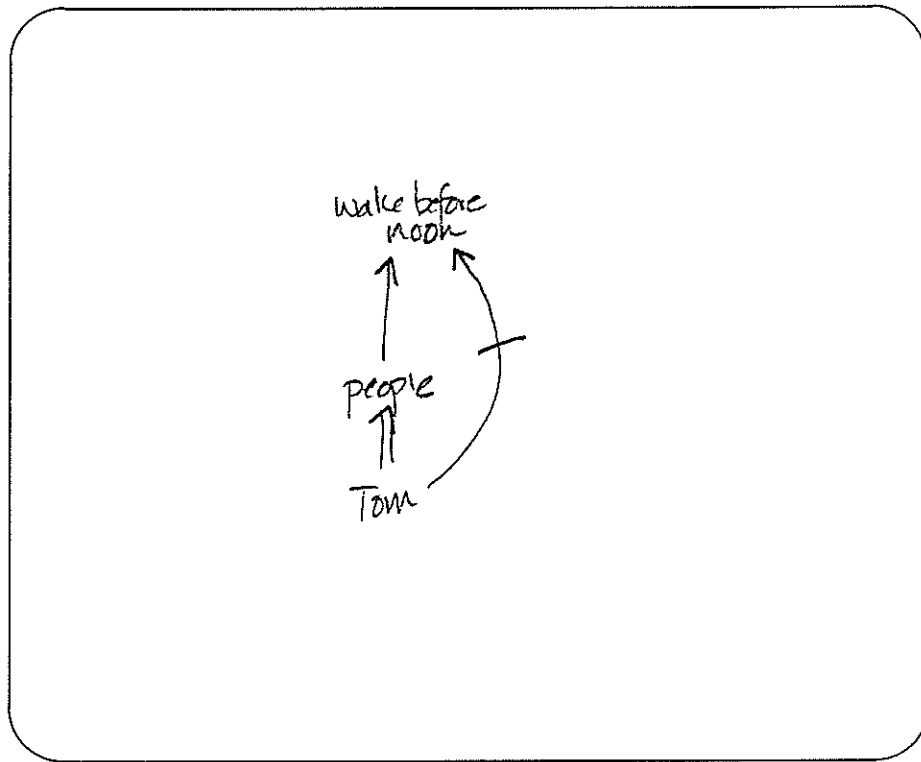
Benchmark Problems

21 Tom's waking

Hidden Reasons for presumptions (Geffner)

People usually wake before noon
Tom usually wakes after noon

Does Tom wake in the afternoon?



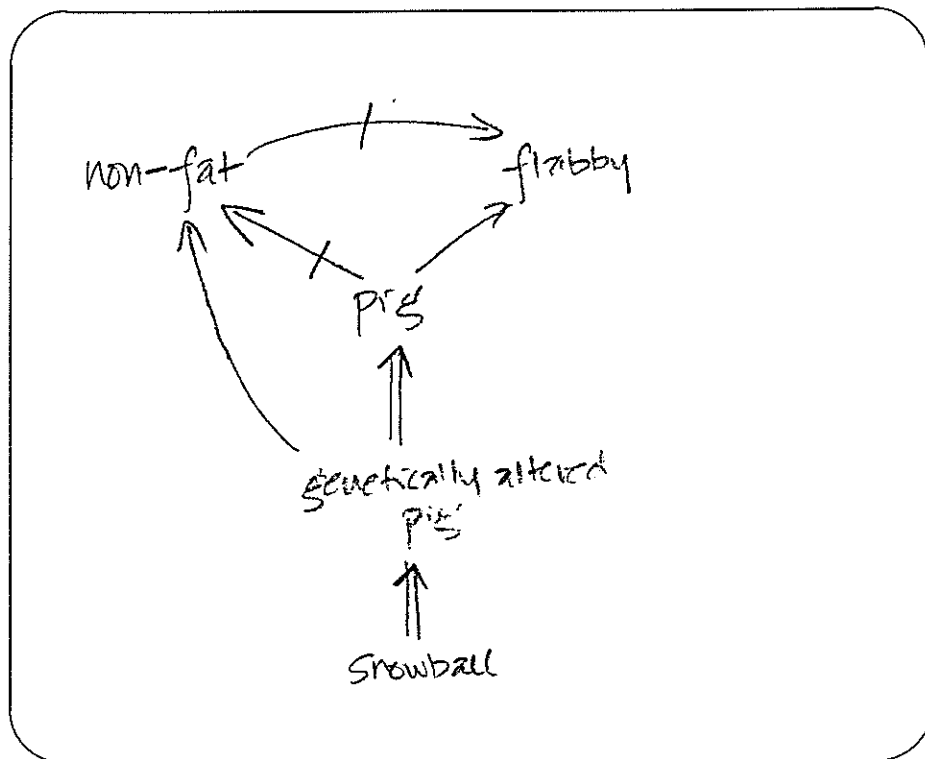
Benchmark Problems

22 Genetically Altered Pigs

Evidence versus Directness (Loui)

Pigs tend to be fat and flabby
Genetically altered pigs tend to be non-fat
Non-fat animals tend to be non-flabby
Snowball is a genetically altered pig

Is Snowball non-flabby?



Benchmark Problems

23 Hernandez Tries to Homer

Argument versus sub-argument (Loui)

Good hitters tend not to try to homer

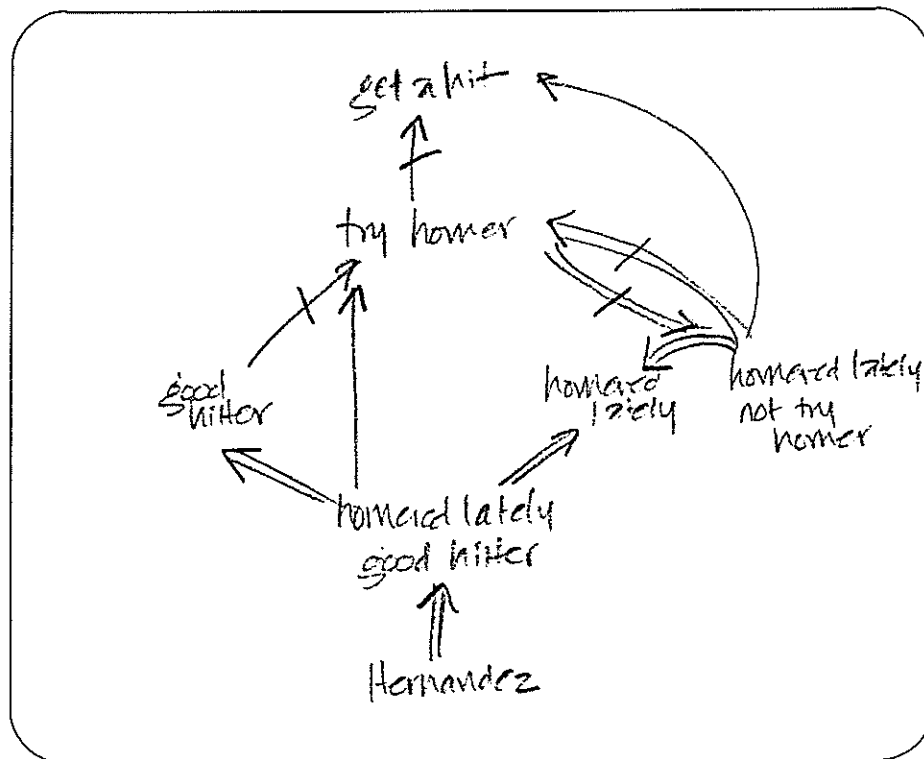
Good hitters who have homered lately tend to try to homer

Trying to homer tends to result in not getting a hit

Batters who have homered lately and aren't trying to homer tend to get hits

Hernandez is a good hitter who has homered lately

Does Hernandez fail to get a hit?



Benchmark Problems

24 Hermann's Pennsylvania Dutch

Mixing strict and defeasible links (Horty-Thomason)

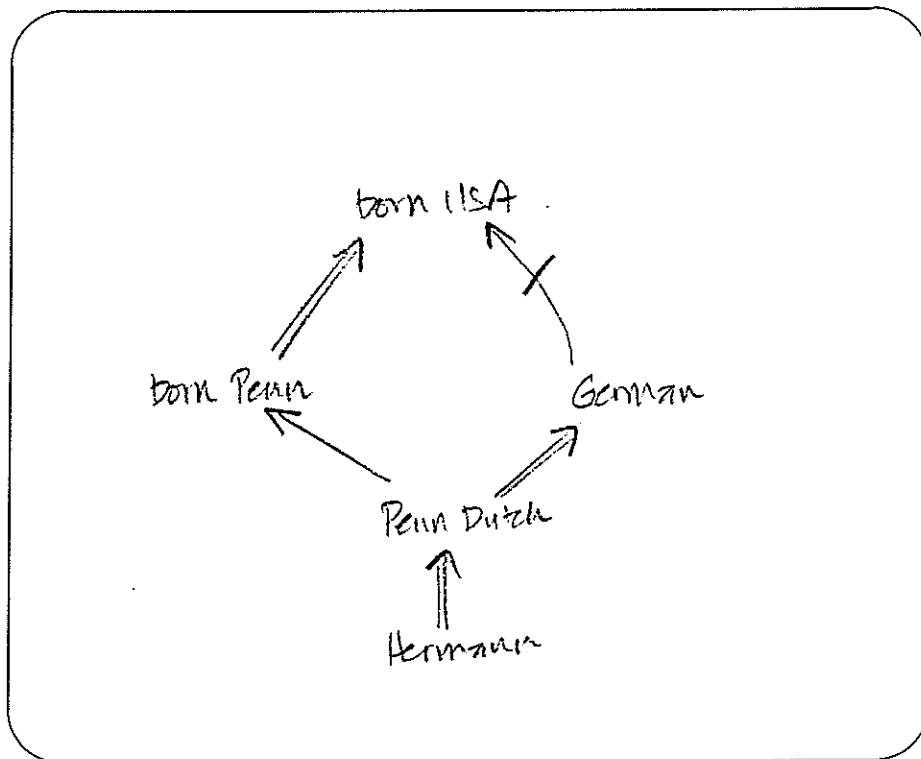
Native speakers of German tend not to be born in America

All native speakers of Pennsylvania Dutch are native speakers of German

Native speakers of Pennsylvania Dutch tend to be born in Pennsylvania

Hermann is a native speaker of Pennsylvania Dutch

Is Hermann born in America?



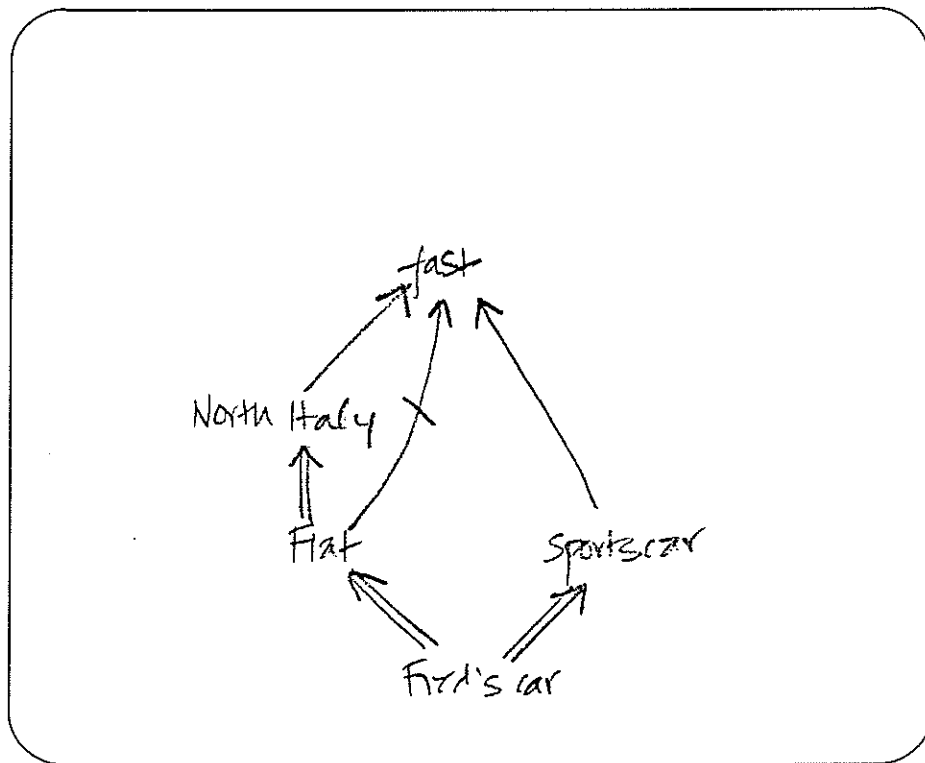
Benchmark Problems

25 Fiats are Fast?

Indirect strong specificity (Loui)

Fiats are North Italian cars
North Italian cars tend to be fast
Sports cars tend to be fast
Fiats tend to be slow
Fred's Fiat is a sports car

Ambiguous?



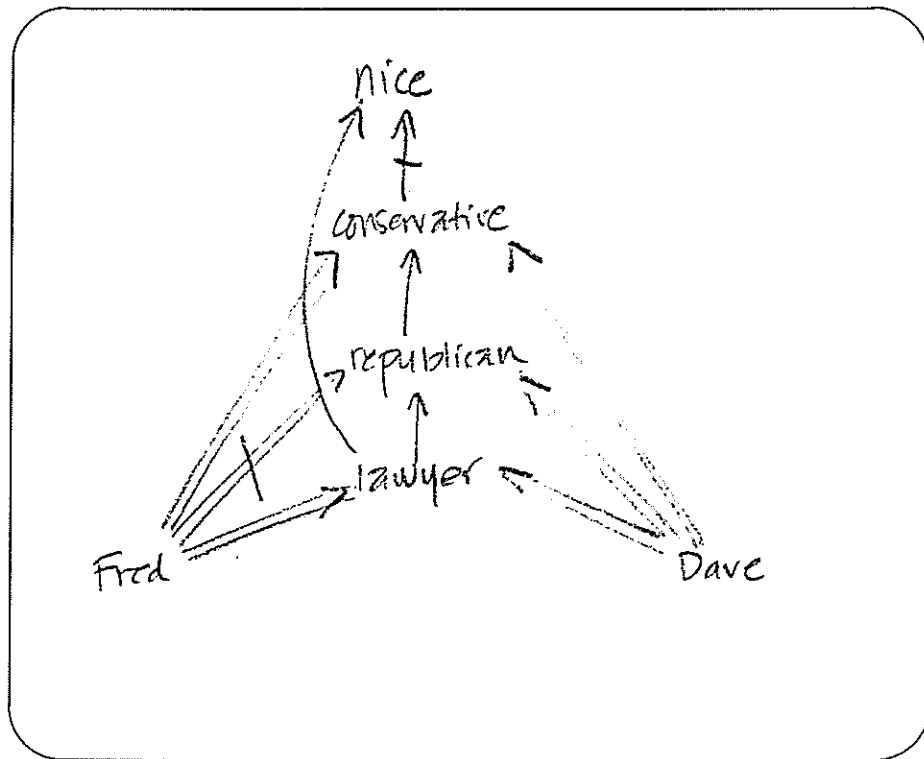
Benchmark Problems

26 Nice Guy Lawyers

Exception to class-based strong specificity (Ginsberg)

Lawyers tend to be republican and nice
Republican tend to be conservative
Conservatives tend to be not-nice
Dave is a conservative republican lawyer
Fred is a conservative non-republican lawyer

Both ambiguous?



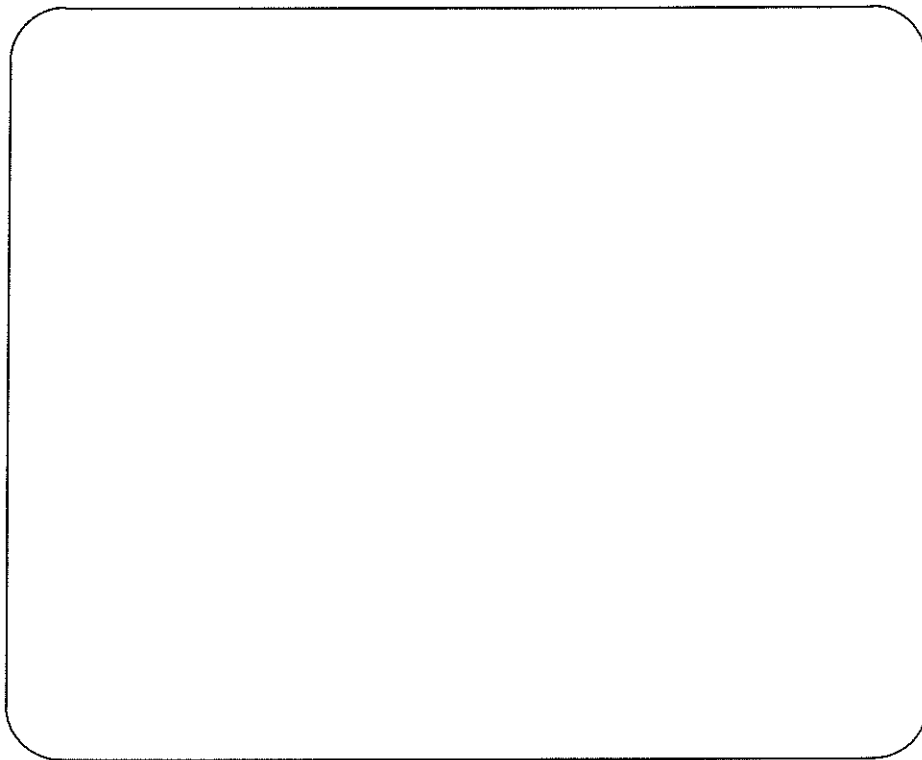
Benchmark Problems

27 Party for Friends

Implicature (Geffner, Myers, Baker)

By default, all my friends will show at my party
Tom and Mary are my friends
Tom says that if Mary shows, he won't show

Does Tom fail to show?



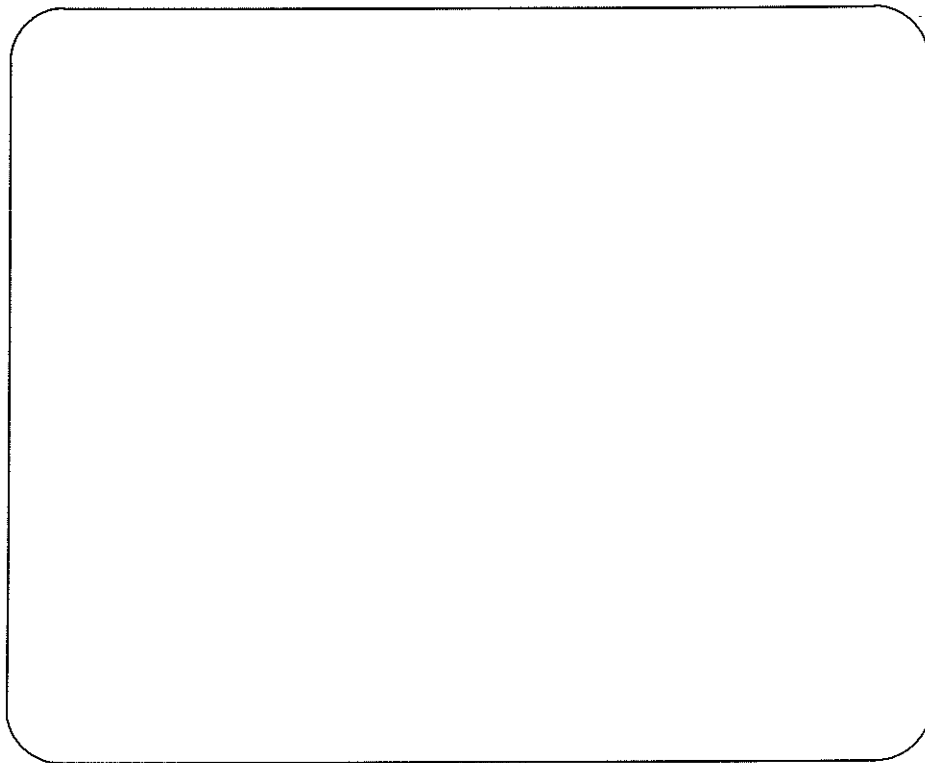
Benchmark Problems

28 Museum Director's Testimony

Hidden reason (De Coster)

An animal collector sends some bones to a museum curator. The bones are from the forelimb of an animal, and they clearly show that the animal has claws. The collector, who is normally reliable, reports that the bones were taken from the wing of a penguin. Now penguins are birds, but animals with foreclaws normally are not birds. (There is an exception, a primitive South American bird born with foreclaws which it loses when it learns to fly).

Is the animal in question a bird?



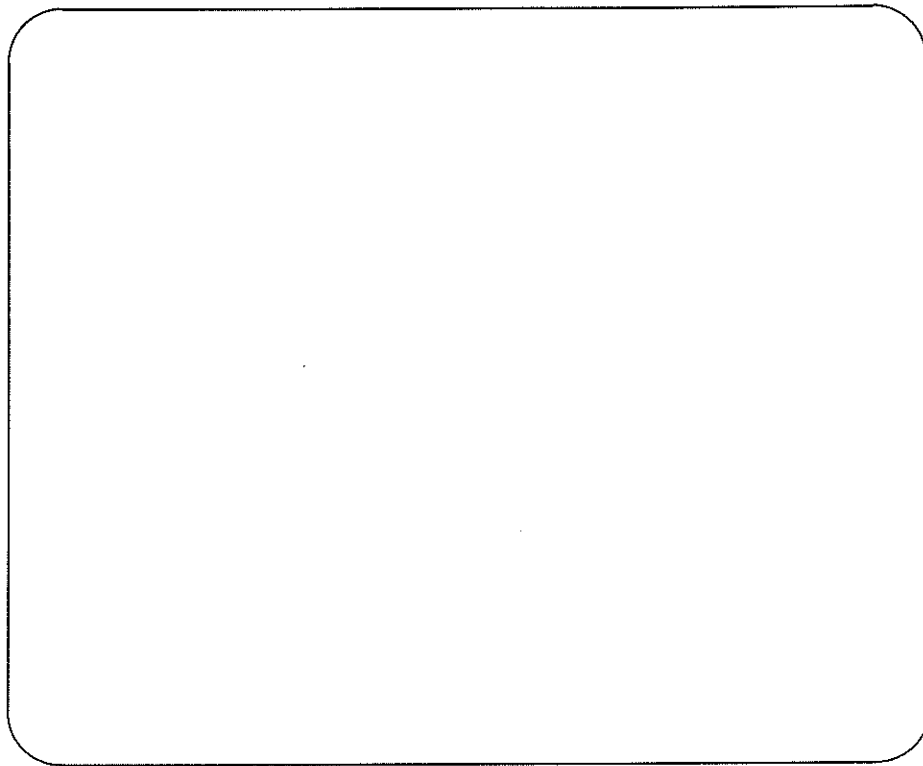
Benchmark Problems

29 Carl, Alice and Betty

Establishing Specificity (Nute)

Carl is dating both Alice and Betty. The girls don't know each other very well, and each believes that she is Carl's only girlfriend. Dave knows and likes Alice and Betty. He also knows Carl, knows about Carl's deceit, and dislikes him for it. With this in mind, we have four plausible rules about when Carl and Dave are likely to attend a party. If Alice is at the party, Carl is normally there too; but if Alice and Betty are both at a party, Carl prudently stays away. If Alice and Betty are both at a party, Dave normally doesn't attend. In a certain occasion, Alice, Betty and Carl do in fact attend the same party.

Does Dave attend this party?



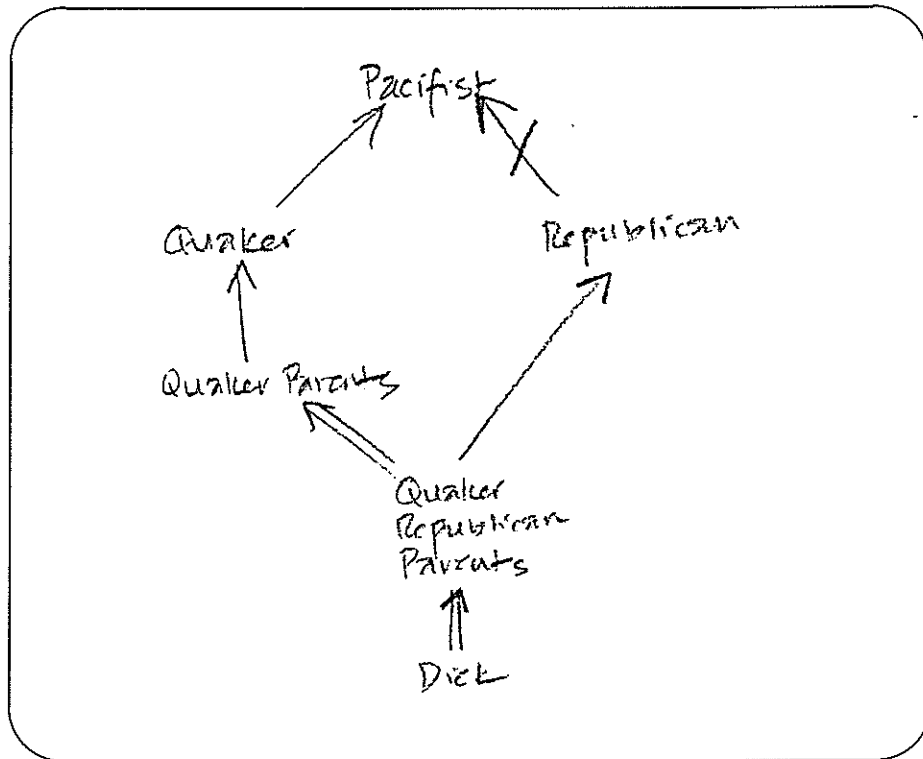
Benchmark Problems

30 Nute's Nixon Diamond

Evidence versus Specificity (Nute)

People whose parents are Quakers normally are Quakers themselves. People whose parents are both Quakers and Republicans normally are Republicans themselves. (This might not be true for people whose parents are both Catholics and Republicans, or whatever.) Furthermore, Quakers are normally pacifist and Republicans normally are not. Assuming all this, and assuming that Dick's parents are both Quakers and Republicans.

Is Dick a pacifist?



How Many Birds Fly?

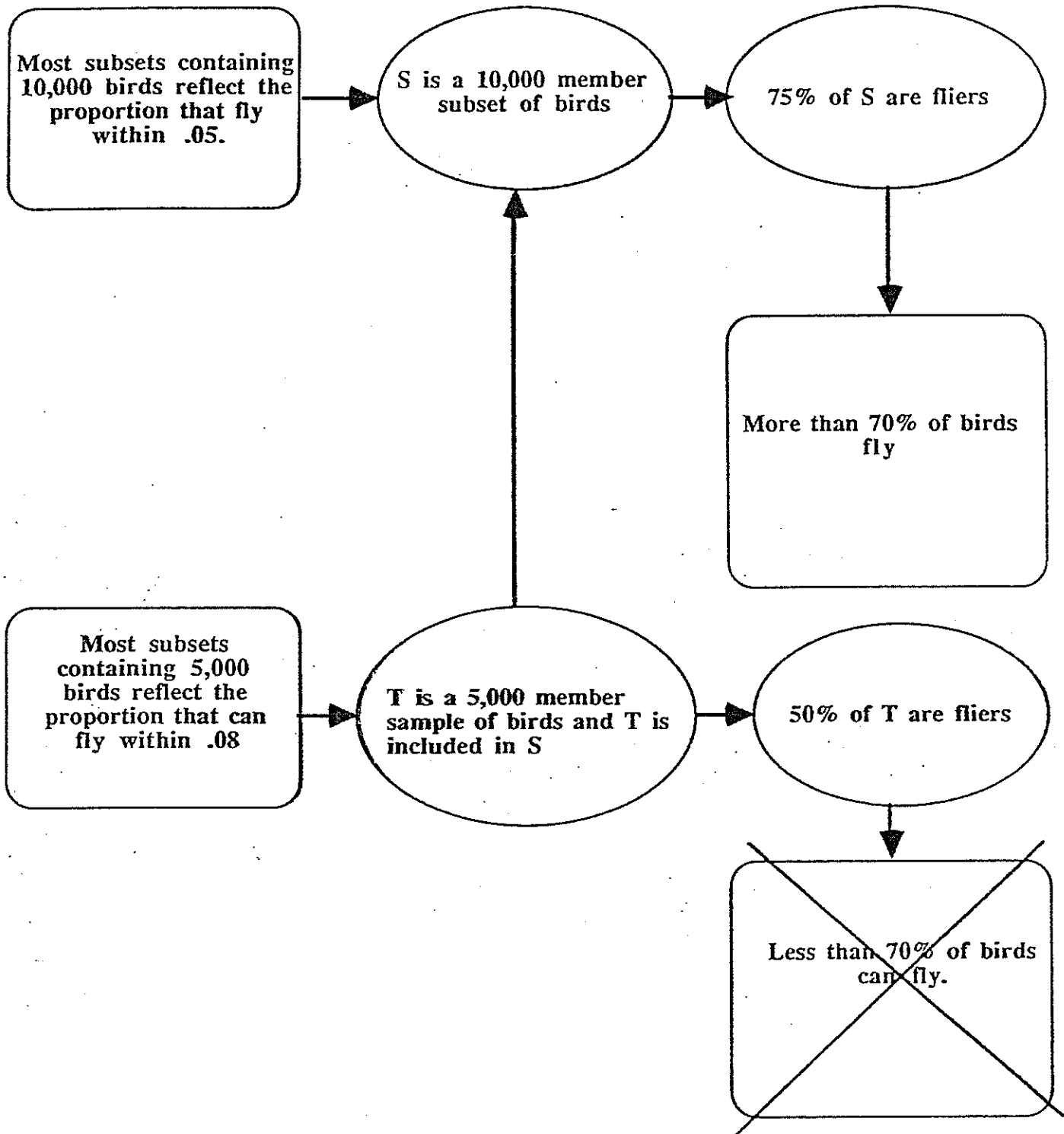


figure 4

Cages of Birds

The Aviary Contains 180 healthy crows, and 20 fat penguins.

Nine cages contain two penguins and a crow. The remaining cage contains 171 crows and two penguins

The Aviary Contains ten cages of birds.

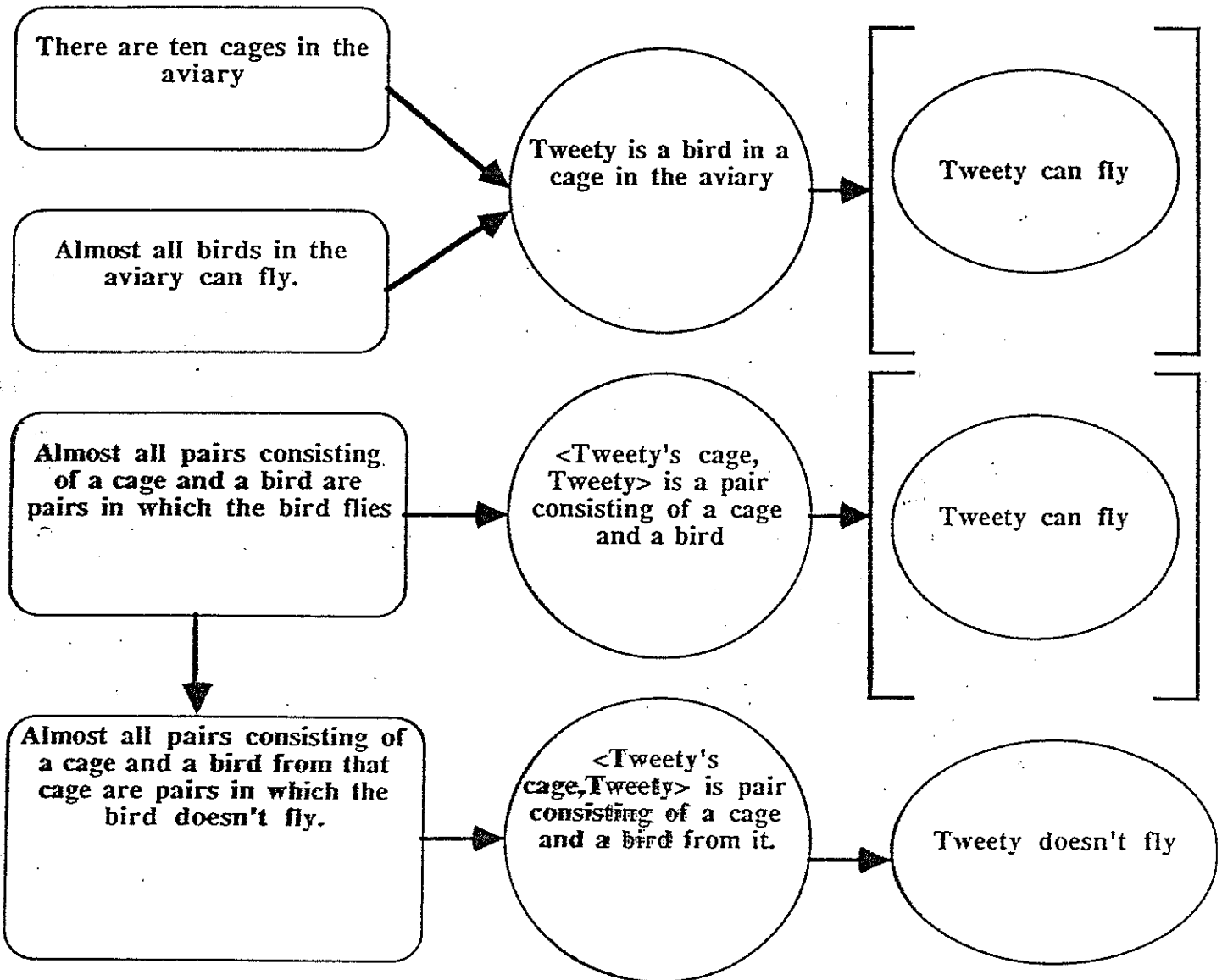


figure 3