Extended Abstract which appeared at:
IJCAI-93 Workshop on Computational Models of
Conflict Management in Cooperative Problem-Solving
Version date: March 8, 1993

New Prioritization Methods for Conflict Management

Benjamin N. Grosof
IBM T. J. Watson Research Center
P.O. Box 704, Yorktown Heights, NY 10598
(914) 784-7100 ; Internet: grosof@watson.ibm.com

Worshop Themes: theoretical underpinnings, generic aspects.

Short Abstract

We address theoretical underpinnings and generic aspects of conflict management in cooperative problem-solving.

Management of conflict often involves *aggregating* multiple transitive preference criteria into a single overall preference criterion: e.g., when reconciling several agents' negotiating positions. We observe that many different kinds of conflict and conflict management can be viewed in these terms: including about goals, desires, and beliefs.

We offer a new method for preference aggregation: our generalized concept of *prioritization* [Grosof, 1991] [Grosof, 1992]. This generalizes the usual concept of lexicographic combination, and is applicable very widely as a qualitative "weak method".

We apply our generalized prioritization to the realm of managing conflict about beliefs. An important task in this realm is to maintain a common blackboard-spirit theory that is shared by cooperating participants, e.g., to maintain a collaborative design rationale. We offer a new method for organizing conflicting beliefs into such a common theory: the Prioritized Argument Management System (PAMS). PAMS revolves around a new logical formalism for non-monotonic reasoning: a variant of prioritized circumscription. PAMS provides more functionality, and a higher level of abstraction, than Truth Maintenance Systems cf. Doyle [1979] and de Kleer [1986]. PAMS provides a computational model with strong, model-theoretic semantics.

1 Representation of Conflicting Preferences

What is conflict? One simple and basic notion of conflict is inconsistency among preferences. E.g., one preference may say: "prefer a to b", and another may say "prefer b to a". Many kinds of goals, desires, and beliefs can be represented as preference criteria, i.e., binary preference relations. A basic coherence property of a preference relation is transitivity: in particular, transitivity makes the notion of "maximally preferred" be well-defined.

Considering a collection of agents, each agent may have one or more preference criteria. E.g., Joe may have one preference criterion corresponding to desire for higher income, and another preference criterion corresponding to desire for more leisure time. Conflict may occur within a single agent, or between multiple agents that are attempting to co-operate. E.g., Joe's two preference criteria conflict when he compares working his current job versus moonlighting at a second job as well.

As another example: beliefs in default reasoning can often be viewed in terms of preferences, as well. A premise-status default belief, e.g., $Quaker(x) \supset Pacifist(x)$ can be viewed as corresponding to a preference criterion: believing the default as a conclusion is preferred to believing its negation as a conclusion. (Formally, this preference criterion might be defined over models, e.g., as in the model-preference approach described by [Shoham, 1988].) E.g., when Nixon is known for-sure to be a Quaker, concluding that Nixon is a Pacifist is preferred to concluding that Nixon is not a Pacifist. Conflict between defaults can then be viewed as conflict between preference criteria. For example, suppose an agent has not only the above premise default but also a second premise default that $Republican(x) \supset \neg Pacifist(x)$. When Nixon is also known for-sure to be a Republican, then the two defaults conflict. The default conclusions are inconsistent; the preference relation corresponding to the Quaker default prefers the conclusion that Nixon is a Pacifist, while the preference relation corresponding to the Republican default prefers the conclusion that Nixon is not a Pacifist.

More generally, when several agents are cooperating, one agent may have desire, goal, or belief preferences that conflict with those of other agents. E.g., in the domain of business planning, one agent may have the goal of increasing market share, while another may have the goal of increasing short-term operating profits. E.g., in the domain of international politics, one agent may believe (the premise default that) that military strength deters aggression and thus war, while another may believe that military build-up provokes distrust and thus precipitates war.

Preferences may have a cardinal aspect as well as an ordinal aspect, of course. E.g., in Joe's desires above, money and time might be real-numbered quantities. In this paper, however, we concentrate mainly on ordinal aspects of preference and conflict.

2 Principles for Managing and Resolving Conflicts

What is managing conflict? It is desirable for a collection of cooperating agents to have a coherent overall set of desires, goals, and beliefs. One basic kind of conflict management is thus: to take a collection of preference criteria, and to **aggregate** (combine) them into a single overall preference criterion. E.g., a "starting" collection of goal preference criteria is aggregated into an overall goal preference criterion. E.g., a collection of defaults is aggregated into an overall default theory. This aggregation view was pioneered social choice theory within mathematical economics to represent desires of consumers and voters (see, e.g., [Arrow, 1963]). In the non-monotonic reasoning literature, this view has been applied to beliefs, notably by Doyle & Wellman [Doyle and Wellman, 1991].

Preferences can be viewed as the underlying basis for **negotiating positions**. Aggregating preferences corresponds to **reconciling** a collection of various (agents') negotiating positions into some kind of decision outcome (which may perhaps be a stalemate).

Example principles/methods for preference aggregation include:

1. majority vote.
2. weighted voting. (Majority vote is the special case where the weights are equal.)
3. lexicographic combination. E.g., alphabetic ordering (or Arabic numerical ordering); here, each character or digit corresponds to one starting preference criterion. Earlier characters (digits) have higher **precedence**; there is a total order of precedence.

3 Generalized Concept of Prioritization

"Prioritization" is an often-used word in discussions of preference aggregation and conflict management. Sometimes, it is used to mean weighting in a voting or linear combination scheme, often in connection with cardinal preferences. Sometimes, it is used to mean precedence in a lexicographic combination scheme.

In [Grosof, 1991] (elaborated and studied in [Grosof, 1992]), we developed a generalized concept of prioritization. Here, we extend it a bit further, and show it is suitable for aggregating any starting collection transitive preference relations into a new overall transitive preference relation. Formally, the prioritization operation takes two inputs: the tuple of starting preference criteria, and a precedence ordering. The precedence ordering is itself any transitive relation, e.g., a partial order. Intuitively, when one starting criterion has higher precedence than another, then in case of conflict between the two criteria, the first one "wins".

This generalizes the usual concept of lexicographic combination, in two ways. Firstly, it permits the **precedence** ordering to be **partial** rather than total; *any* finite (or well-founded infinite) precedence *partial* order is permitted. Secondly, it requires only a very weak structure on preferences, namely transitivity, and is thus applicable very widely as a qualitative "weak method". In particular, a preference criterion may be partial in two ways: it may 1) abstain, or 2) tie, when comparing two given candidates. In addition, prioritization can be composed, i.e., applied recursively and hierarchically.

Our concept of prioritization thus appears an interesting tool for specification of conflict management for a wide variety of preference aggregation scenarios, e.g., to aggregate goals, desires, and/or beliefs.

Our generalized concept of prioritization was originally developed in order to represent precedence in non-monotonic reasoning (NMR), e.g., in the circumscription formalism [McCarthy, 1986] [Lifschitz, 1985]. Recently, we have shown [Grosof, 1993] that this concept subsumes and unifies several different concepts of precedence and specificity dominance, in various non-monotonic logical formalisms, from the NMR literature: each of those concepts is equivalent to a special case of our concept of prioritization.

4 Prioritized Argument Management System

Next, we show how prioritization can be applied to the realm of managing conflict about *beliefs* in cooperative problem-solving. We develop an approach, the Prioritized Argument Management System (PAMS), based on a logical formalism for default reasoning: the Circumscriptive Language of Defaults [Grosof and Russell, 1990] [Grosof, 1992].

Maintaining a Dynamic Common Theory: An important task in this realm is to maintain a common blackboard-spirit theory that is shared by cooperating problem-solvers (agents). An oft-used architectural style in Distributed AI is to have different agents "post" each's partial results, e.g., hypotheses or constraints, to the common theory, and to "read" others' from there as well, dynamically. In collaborative design, for example, agents may post their current design choices and constraints. Conflict may arise in at least two ways. Firstly, two or more agents' beliefs may be inconsistent with each other. E.g., in designing a car, the engine expert may post the design choice that the engine is a V-6 (which weighs 300 kilos); while the weight/fuel-mileage expert may post the design choice that the engine weigh less than 220 kilos). Secondly, beliefs may be updated dynamically in a manner that overrides previous beliefs. E.g., as various problem-solvers post hypotheses associated with their specialized activities, they may "change their mind" and retract their previously-posted hypotheses. E.g., the engine expert may, on the basis of a technology breakthrough, change his posted design choice to be that the engine is a new kind of 4-cylinder. In both of these cases, it is appropriate to represent beliefs as retractable/revisable in status, from point of view of forming a common design. From a declarative point of view, the reasoning is non-monotonic (NM); to us, this suggests using a NM logical formalism to represent the posted and common beliefs.

The PAMS: In a NM logical formalism, a theory has two parts: the set of premises and the set of conclusions. The set of premises can be updated only by accumulation (adding more). Conclusions, however, may be retracted as the set of premises is updated.

In the intended application of PAMS, agents contribute premises to the common theory by posting. These may lead to conflict, which is resolved and managed to form a common set of conclusions. Agents then may read the common conclusions.

As part of the common theory, there is also information specifying how conflicts are to be resolved. In PAMS, this information is simply a prioritization-type precedence ordering over the premises. Precedence may represent, for example, relative expertise or freshness of information.

An important challenge in maintaining such a common blackboard-spirit theory is to track dependencies. Suppose a conclusion was justified using several premises; if one of those premises is overridden by a premise update, then that conclusion should be retracted. **Dependency tracking** is important for **maintaining a design rationale**, for example: when some design choices D are revised, one would like to know which other design choices E were dependent on D, since E should now be re-considered.

Truth Maintenance Systems (TMS) cf. Doyle [1979] and de Kleer [1986] are a well-known and widely-used method for handling dependency tracking. In developing the PAMS, we take inspiration from these TMS ideas. PAMS has several advantages over TMSes; we discuss these below.

Formally, PAMS is defined as set of computational reasoning services over a NM theory.

The NM theory is expressed in terms of a NM logical formalism. The basic case of this NM logical formalism is the Circumscriptive Logic of Defaults. In the core version of this formalism, there are three kinds of premises: base axioms, default axioms, and prioritization axioms. Base axioms and default axioms have a formula part; each of these formulas is in some first-order (monotonic-logical) base language. A default axiom specifies that a formula is asserted as a default rule, i.e., with revisable status. A base axiom specifies that a formula is asserted with for-sure status. Each default axiom has a label (i.e., name). A prioritization axiom takes a pair of default labels as arguments, and specifies that the default with the first label has higher precedence than the default with the second label. The overall precedence ordering for the NM theory is then the transitive closure of the pairs specified by the prioritization axioms.

Example: Collaborative Design of a Car. Consider the following scenario of a common theory. Various experts assert design choices or implications among design choices. These are represented in PAMS as either default axioms, if revisable; or else as base axioms. The marketing expert asserts that the acceleration is .65 g's. The weight expert asserts that the weight is 1600 kilos. The engine expert asserts that if the acceleration is over .64 g's, and the weight is over 1500 kilos, then the engine type is V-6. The engine expert also asserts that if the acceleration is under .66 g's and the weight is under 1450 kilos, then the engine type is standard-4-cylinder. These four premises are represented as default axioms. In addition, the usual numerical quantity inequalities (e.g., 1600 kilos is over 1500 kilos) and obvious incompatibilities (e.g., the engine type cannot be both V-6 and standard-4-cylinder) are represented with base axioms. Let $\mathcal{A}_1$ stand for this initial set of axioms. So far there is no conflict; all of the defaults "go through". The common *conclusions* entailed (in the NM formalism CLD) by $\mathcal{A}_1$ include that the acceleration is .65 g's, the weight is 1600 kilos, and the engine type is V-6.

Next, there is a premise update. The weight expert asserts that the weight is 1400 kilos. This is represented as a new default axiom. The weight expert asserts furthermore that this new default has higher precedence than (the default corresponding to) his previous assertion that weight is 1600 kilos. I.e., the weight expert has changed his mind. This new default conflicts with that previous default. Higher prioritization-type precedence means that this new default overrides the previous one. Let $\mathcal{A}_2$ stand for the new axiom set (i.e., $\mathcal{A}_1$ augmented by the new default axiom and the new prioritization axiom). The new common *conclusions* entailed by $\mathcal{A}_2$ involve revision: now the weight is concluded to be 1400 kilos. Moreover, the previous conclusion about engine type that was justified by the previous conclusion about weight is also revised: now the engine type is concluded to be standard-4-cylinder.

This illustrates how the PAMS can gracefully and simply represent overriding updates involving conflict, and track dependencies appropriately.

PAMS can also represent stalemated conflict: e.g., when two defaults conflict and neither has higher precedence than the other. For example, let us continue updating the above theory: the mileage expert asserts that the engine type is wimpy-4-cylinder. This is represented as a new default axiom. Let $\mathcal{A}_3$ stand for the new axiom set. The new default conflicts with the previous defaults: about whether the engine type is standard-4-cylinder or wimpy-4-cylinder. However, in $\mathcal{A}_3$, there is no strict precedence established in either direction between the new default and any of the previous ones. Thus the conflict results in stalemate: neither "wins". In the new conclusions, the previous conclusion about engine type is retracted; the strongest new conclusion entailed about engine type is that it is *either* standard-4-cylinder *or* wimpy-4-cylinder. The conflict might be resolved definitively in favor of one or the other: by the assertion of prioritization axioms as an additional update. E.g., if the mileage expert's default were asserted to have higher precedence than the engine expert's default (say, because the mileage expert has greater authority than the engine expert), then it would "win" and wimpy-4-cylinder would be concluded as the engine type.

Bases for Prioritization-type Precedence: As we saw above, prioritization-type precedence may be used to represent, for example, relative expertise or freshness of information. Other bases for precedence among defaults include: specificity, reliability, authority, and decision-theoretic (differential probabilistically-expected utility).

Comparison with TMSes; Advantages of PAMS: PAMS provides more functionality, and a higher level of abstraction, than TMSes. Most notably, PAMS has a precedence mechanism, while TMSes do not. In effect, TMSes can be viewed as having an empty precedence ordering. However, PAMS has two further advantages over TMSes even apart from precedence. Firstly, in NMR terminology, PAMS is "skeptical": it maintains a single set of strongly justified common conclusions. By contrast, Doyle-style TMSes when used non-monotonically are "brave": there are many different possible IN-OUT labellings (IN means concluded) sanctioned for a given set of premise beliefs ("justifications").[1] Yet a Doyle-style TMS computes non-deterministically only one of these labellings. Skepticism corresponds to the intersection of all these labellings: a skeptical conclusion is one that is IN in all the brave labellings. de Kleer-style Assumption-based TMSes are fundamentally logically monotonic (however, they can be used as a sub-procedure to support NMR — see below). Secondly, PAMS is more "semantic" than TMSes in the following sense. PAMS avoids TMSes' need to generate and assert all relevant tautologies of the underlying monotonic base language (e.g., first-order logic) in which believed formulas are expressed. TMSes treat belief (sub-)formulas as "nodes", i.e., primitive propositions with opaque semantics. Interestingly, PAMS can mirror this "less semantic mode" of TMSes as a special case: i.e., by representing belief (sub-)formulas as primitive propositions.

Similarly to an ATMS, the reasoning services in a PAMS include not only the current conclusions, but also the "arguments" *pro* and *con* any given proposition, and conflict sets (i.e., sets of defaults that currently conflict).

Discussion: PAMS provides a computational model with strong, model-theoretic semantics. In effect, it provides ordinal degrees of belief (and skepticism) in order to specify handling of conflict. We have shown (partly reported in [Grosof, 1992]) that the effectively-propositional case (e.g., first-order with domain closure and uniqueness of names) of PAMS is **implementable** by building on top of previous NMR algorithms: e.g., employing Assumption-based TMSes [de Kleer, 1986] in two different roles at once. Like TMSes, the PAMS is **worst-case intractable** computationally; experimentation is needed to explore its practicality.

Future Work: Issues for future work include developing algorithms to implement PAMS more *efficiently* and in more expressive generality; extending its expressive capabilities to

[1]This proliferation is typically combinatoric in the number of conflicts, and in the worst case is exponential in the number of defaults instances.

include *meta-levels* of argumentation about precedence; and extending its expressive capabilities to include *cardinal* degrees of belief, e.g., probabilities.

References

[Arrow, 1963] Kenneth J. Arrow. *Social Choice and Individual Values.* Yale University Press, 2nd edition, 1963.

[de Kleer, 1986] J. de Kleer. An assumption-based truth maintenance system. *Artificial Intelligence*, 28:127–162, 1986.

[Doyle and Wellman, 1991] Jon Doyle and Michael P. Wellman. Impediments to universal preference-based default theories. *Artificial Intelligence*, 49:97–128, 1991. Revised and expanded from version in *Proceedings of the First International Conference on Principles of Knowledge Representation*, pp. 94-102, Toronto, Ontario, Morgan Kaufmann, 1989.

[Doyle, 1979] J. Doyle. A truth maintenance system. *Artificial Intelligence*, 12:231–272, 1979.

[Grosof and Russell, 1990] Benjamin N. Grosof and Stuart J. Russell. Shift of bias as non-monotonic reasoning. In Pavel Brazdil and Kurt Konolige, editors, *Machine Learning, Meta-Reasoning, and Logics.* Kluwer Academic, 1990. Based on the workshop held in Sesimbra, Portugal, in Feb. 1988. Also available as IBM Research Report RC14620.

[Grosof, 1991] Benjamin N. Grosof. Generalizing prioritization. *Proceedings of the Second International Conference on Principle of Knowledge Representation and Reasoning*, pages 289–300, April 1991. Also available as IBM Research Report RC15605, IBM T.J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598.

[Grosof, 1992] Benjamin N. Grosof. *Updating and Structure in Non-Monotonic Theories.* PhD thesis, Computer Science Dept., Stanford University, Stanford, California 94305, October 1992. Published by University Microfilms, Inc.. Also available as Research Report from Stanford University Computer Science Dept. and/or IBM T.J. Watson Research Center.

[Grosof, 1993] Benjamin N. Grosof. A unification of several concepts of precedence and specificity dominance between defaults: All are special cases of prioritization. Technical report, IBM T.J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598, Mar. 1993. Submitted.

[Lifschitz, 1985] V. Lifschitz. Computing circumscription. *Proceedings IJCAI-85*, pages 121–127, Los Angeles, CA, 1985.

[McCarthy, 1986] J. McCarthy. Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, 28:89–116, 1986.

[Shoham, 1988] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, Mass., 1988.