Available online at www.sciencedirect.com
ScienceDirect
Journal of Applied Logic 5 (2007) 586–601
JOURNAL OF APPLIED LOGIC
www.elsevier.com/locate/jal

Preferential theory revision

Pierangelo Dell'Acqua [a,b], Luís Moniz Pereira [b,*]

[a] *Department of Science and Technology – ITN, Linköping University, 60174 Norrköping, Sweden*
[b] *Centro de Inteligência Artificial – CENTRIA, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*

Available online 18 April 2006

Abstract

Employing a logic program approach, this paper focuses on applying preferential reasoning to theory revision, both by means of preferences among existing theory rules, and by means of preferences on the possible abductive extensions to the theory. And, in particular, how to prefer among plausible abductive explanations justifying observations.

Keywords: Theory revision; Preferences; Abduction; Diagnosis; Logic programs

1. Introduction

Logic program semantics and procedures have been used to characterize preferences among the rules of a theory [5]. Whereas the combination of such rule preferences with program updates and the updating of the preference rules themselves [4] have been tackled, a crucial ingredient has been missing, namely the consideration of abductive extensions to a theory, and the integration of revisable preferences among such extensions. The latter further issue is the main subject of this paper.

We take a theory expressed as a logic program under stable model semantics, already infused with preferences between rules, and we add a set of abducibles constituting possible extensions to the theory, governed by conditional priority rules amongst preferred extensions. Moreover, we cater for minimally revising the preferential priority theory itself, so that a strict partial order is always enforced, even as actual preferences are modified by new incoming information. This is accomplished by means of a diagnosis theory on revisable preferences over abducibles, and its attending procedure.

First we supply some epistemological background to the problem at hand. Then we introduce our preferential abduction framework, in the context of a specific logic program approach, and proceed to apply it to exploratory data analysis. Next we consider, within the same approach, the diagnosis and revision of preferences, theory and method, and illustrate it on the data exploration example. Finally, we exact general epistemic remarks about the general approach.

* Corresponding author.
E-mail addresses: pier@itn.liu.se (P. Dell'Acqua), lmp@di.fct.unl.pt (L.M. Pereira).

1570-8683/$ – see front matter
doi:10.1016/j.jal.2006.03.010

1.1. Preferences, rationality, theory revision, and AI

(I) The theoretical notions of *preference* and *rationality* with which we are most familiar are those of the economists'. *Economic preference* is a comparative choice between alternative outcomes whereby a *rational* (economic) *agent* is one whose expressed preferences over a set of outcomes exhibits the structure of a complete pre-order.

However, actual preferences may change. Viewing this phenomena as a comparative choice, however, entails that there are meta-level preferences whose outcomes are various preference rankings of beliefs and that an agent chooses a change in preference based upon a comparative choice between the class of first-order preferences [6]. But this is an unlikely model of actual change in preference, since we often evaluate changes—including whether to abandon a change in preference–based upon items we learn *after a change in preference is made*. Hence, a realistic model of preference change will not be one that is couched exclusively in decision theoretic terms. Rather, when a conflict occurs in updating beliefs by new information, the possible items for revision should include both the set of conflicting beliefs *and* a reified preference relation underlying the belief set. The reason for adopting this strategy is that we do not know, *a priori*, what is more important—our data or our theory. Rather, as Isaac Levi has long advocated [14], rational inquiry is guided by pragmatic considerations not a priori constraints on rational belief. On Levi's view, all justification for change in belief is pragmatic in the sense that justification for belief fixation and change are rooted in strategies for promoting the goals of a given inquiry. Setting these parameters for a particular inquiry fixes the theoretical constraints for the inquiring agent. The important point to stress here is that there is no conflict between theoretical and practical reasoning on Levi's approach, since the prescriptions of Levi's theory are *not* derived from minimal principles of rational consistency or coherence [14].

In [19], arguments are given as to how epistemic entrenchment can be explicitly expressed as preferential reasoning. And, moreover, how preferences can be employed to determine belief revisions, or, conversely, how belief contractions can lead to the explicit expression of preferences.

[6] provides a stimulating survey of opportunities and problems in the use of preferences, reliant on AI techniques.

(II) Suppose your scientific theory predicts an observation, o, but you in fact observe $\neg o$. The problem of carrying out a principled revision of your theory in light of the observation $\neg o$ is surprisingly difficult. One issue that must be confronted is what the principle objects of change are. If theories are simply represented as sets of sentences and prediction is represented by material implication, then we are confronted with *Duhem's Problem* [7]: If a theory entails an observation for which we have disconfirming evidence, logic alone won't tell you which among the conjunction of accepted hypotheses to change in order to restore consistency. The serious issue raised by Duhem's problem is whether disconfirming evidence targets the items of a theory in need of revision in a principled manner.

The AGM [1] conception of belief change differs to Duhem's conception of the problem in two important respects. First, whereas the item of change on Duhem's account is a set of sentences, the item of change on the AGM conception is a belief state, represented as a pair consisting of a logically closed set of sentences (a belief set) and a selection function. What remains in common is what Sven Hansson [11] has called the *input-assimilating* model of revision, whereby the object of change is a set of sentences, the input item is a particular sentence, and the output is a new set of sentences. But one insight to emerge is that the input objects for change may not be single sentences, but a sentence-measure pair [15], where the value of the measure represents the entrenchment of the sentence and thereby encodes the ranking of this sentence in the replacement belief set [15,19,20]. But once we acknowledge that items of change are not *beliefs simpliciter* but belief and order coordinates, then there are two potential items for change: the acceptance or rejection of a belief and the change of that belief in the ordering. Hence, implicitly, the problem of preference change appears here as well.

(III) Computer science has adopted logic as its general foundational tool, while Artificial Intelligence *AI* has made viable the proposition of turning logic into a *bona fide* computer programming language. At the same time, *AI* has developed logic beyond the confines of monotonic cumulativity, typical of the precise, complete, endurable, condensed, and closed mathematical domains, in order to open it up to the non-monotonic real-world domain of imprecise, incomplete, contradictory, arguable, revisable, distributed, and evolving knowledge. In short, *AI* has added dynamics to erstwhile statics. Indeed, classical logic has been developed to study well-defined, consistent, and unchanging mathematical objects. It thereby acquired a static character. *AI* needs to deal with knowledge in flux, and less than perfect conditions, by means of more dynamic forms of logic. Too many things can go wrong in an open non-mathematical world, some of which we don't even suspect. In the real world, any setting is too complex already for us to define exhaustively each time. We have to allow for unforeseen exceptions to occur, based on new incoming information.