On Epistemic Spaces for Iterated Belief Revision

Sébastien Konieczny[1,*]

[1]CRIL, CNRS - Université d'Artois, 62300 Lens - France

Abstract

Belief revision is a fundamental ability for an autonomous agent, that should be able to correct her beliefs when evidence contradicts them. In this talk we will recall the framework of iterated belief revision, following Darwiche and Pearl's approach and its developments. In particular this approach works on abstract objects called epistemic states. We will discuss what can be some interesting candidates for representing these abstract objects, through the definition of epistemic spaces. And we will see if TPO (total-pre-orders on interpretations) or OCF (Ordinal Conditional Functions) are acceptable candidates to be considered as potential canonical representations of these epistemic states.

Keywords

Belief Revision, Iterated Revision, Ordinal Conditional Functions

1. Summary of the talk

Belief revision is a fundamental ability for an autonomous agent, that should be able to correct her beliefs when evidence contradicts them.

The core of the belief revision theory is the seminal work by Alchourrón, Gärdenfors and Makinson [1, 2], where they define axiomatically the operators of revision and contraction (contraction is an operation where we want to remove some belief from the beliefs of the agent, whereas revision is the operation where we want to correct some belief), and where they provide a representation theorem in terms of partial-meet selection functions. A representation theorem is an equivalence between a set of axiomatic properties (called postulates) and a constructive way to define a family of operators.

One interesting representation theorem is the one given by Katsuno and Mendelzon [3], and that provides, in the standard propositional logic setting, a representation of revision operators in terms of faithful assignments (total pre-orders on interpretations), where the revision amounts to select, amongst the models of the new information, the ones that are the most plausible with respect to plausibility relation provided by the total pre-order on interpretations.

But all these classical works on belief revision only address one step of the process, and do not provide any constraint on the *iteration* of the revision. Then iterated belief revision was a central topic of belief revision theory in the 90s. And what is usually considered as the solution to this issue is the work of Darwiche and Pearl [4], where they propose four additional postulates in order to ensure a correct behaviour for the iteration of revision, and where they change the representation of the beliefs of the agent. In the standard AGM framework the beliefs of the agent are represented by a theory (a logically closed set of formulas). In Darwiche and Pearl approach the beliefs of the agent are represented by an *epistemic state*, that is an abstract object, to which a projection function Bel allows to associate a propositional formula, that represent the current beliefs of the agent. This formula correspond to the information that is used in the standard AGM framework, but epistemic states allow to encode also other information important for the iteration. In particular epistemic states allow to encode conditional information [5], and the preservation of this conditional information through the iteration is one of the main issue solved by Darwiche and Pearl approach.

This approach has been improved and extended in several ways. One can for instance note the class of *admissible operators*, defined in both [6] and [7], that allows to guarantee a better behaviour for iteration. One can also mention *improvement operators* [8, 9], that are weaker change operators, where the plausibility of the new information is improved in the epistemic state, but where it is not forced that this new information is believed after the change.

Darwiche and Pearl also provide a representation theorem, showing that we can associate to each epistemic state a total pre-order on interpretation, and that the beliefs after the revision will be the most plausible model of the new formula with respect to this corresponding pre-order.

This kind of representation theorem can make people think that epistemic states are just total pre-orders on interpretations (let's call these TPO), and that iterated revision are just functions that associate a new TPO to any TPO (representing the old beliefs of the agent) and any formula (representing the new piece of information).

And, even if this subclass of revision operators defined

ENIGMA-23, September 03–04, 2023, Rhodes, Greece

*Corresponding author.

konieczny@cril.fr (S. Konieczny)

http://www.cril.fr/~konieczny/ (S. Konieczny)

0000-0002-2590-1222 (S. Konieczny)

on TPO is interesting and can prove to be useful for many applications, it is not the whole class of Darwiche and Pearl's iterated revision operators. Some works [10, 11] show that interesting iterated revision operators can not be instantiated (represented) using the TPO representation.

Then an interesting question is to investigate the possible candidates for the representation of these iterated revision operators, and to look for some canonical representation. It is possible to define a function of *instantiation* between two representations (that will be called epistemic spaces). These functions can be interesting to show for instance that an epistemic space is more general than another one. And then the question of the canonical representation is to find a top element for this relation.

In fact, Darwiche and Pearl's epistemic states being abstract objects, the whole class is very general, and one can show that no countable representation can be adequate. But if we add another sensible property we can reach such a result of a canonical representation. First, note that, whatever the chosen representation, every epistemic state can be viewed as a black box associating each *finite* sequence of formulae with a formula representing the beliefs of the agent after the successive revision of the epistemic state by each formula from the sequence. Based on that observation, two epistemic states are *strongly equivalent* according to a revision operator if they cannot be distinguished from each other by any such successive revision steps, which means that these epistemic states have the same behavior for that revision operator. Now the (very natural) assumption is that every epistemic state is reachable from an initial, "empty", epistemic state, through a finite succession of revisions. If we make this assumption, then OCFs are a possible candidate of canonical representation of epistemic states.

OCFs were proposed long ago by Spohn [12] for (iterated) change, but this representation need some numerical information, since an OCF is a function that associates to each interpretation a natural number (Spohn's original work uses ordinals instead of numbers but for most works numbers are sufficient) that represents its implausibility, with the constraint that some interpretations are associated to zero, and are considered as the current beliefs of the agent.

But if one supposes that we start from an initial "empty" epistemic state, then several operators can be define that produce OCFs without the need of extra numerical information [8, 11, 13]. So OCFs can be considered as the canonical representation of epistemic states (that are generated from an initial empty epistemic state).

Acknowledgments

The works corresponding to this talk are joint works with Nicolas Schwind (AIST - Tokyo - Japan - nicolas-schwind@aist.go.jp) and Ramon Pino Pérez (CRIL - CNRS, Université d'Artois - Lens - France - pinoperez@cril.fr).

This work has benefited from the support of the AI Chair BE4musIA of the French National Research Agency (ANR-20-CHIA-0028).

References

[1] C. E. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, Journal of Symbolic Logic 50 (1985) 510–530.

[2] P. Gärdenfors, Knowledge in flux, MIT Press, 1988.

[3] H. Katsuno, A. O. Mendelzon, Propositional knowledge base revision and minimal change, Artificial Intelligence 52 (1991) 263–294.

[4] A. Darwiche, J. Pearl, On the logic of iterated belief revision, Artificial Intelligence 89 (1997) 1–29.

[5] C. Boutilier, Iterated revision and minimal change of conditional beliefs, Journal of Philosophical Logic 25 (1996) 263–305.

[6] R. Booth, T. A. Meyer, Admissible and restrained revision, Journal of Artificial Intelligence Research 26 (2006) 127–151.

[7] Y. Jin, M. Thielscher, Iterated belief revision, revised, Artificial Intelligence 171 (2007) 1–18.

[8] S. Konieczny, R. Pino Pérez, Improvement operators, in: Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08), 2008, pp. 177–187.

[9] S. Konieczny, M. Medina Grespan, R. Pino Pérez, Taxonomy of improvement operators and the problem of minimal change, in: Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR'10), 2010, pp. 161–170.

[10] T. I. Aravanis, P. Peppas, M. Williams, Observations on Darwiche and Pearl's approach for iterated belief revision, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19), 2019, pp. 1509–1515.

[11] N. Schwind, S. Konieczny, Non-prioritized iterated revision: Improvement via incremental belief merging, in: Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR'20), 2020, pp. 738–747.

[12] W. Spohn, Ordinal Conditional Functions: A Dynamic Theory of Epistemic States, Springer Netherlands, 1988, pp. 105–134.

[13] N. Schwind, S. Konieczny, R. Pino Pérez, On the

Representation of Darwiche and Pearl's Epistemic States for Iterated Belief Revision, in: Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR'22), 2022, pp. 320–330.