Detecting Inconsistencies in Revision Problems

Fabian Schmidt, Jörg Gebhardt and Rudolf Kruse

Abstract When dealing with complex knowledge, inconsistencies become a big problem. One important aspect of handling inconsistencies is their detection. In this paper we consider approaches to detect different types of inconsistencies that may occur in the formulation of revision problems. The general discussion focuses on the revision of probability distributions. In our practical analysis, we refer to probability distributions represented as Markov networks.

1   Introduction

One important aspect of maintaining knowledge for knowledge based systems is the ability to react to changes in beliefs quickly and frequently. Therefore, methods have been developed to properly adapt knowledge to new beliefs. One important aspect of proper adaptation is formulated in the **principle of minimal change** [9], which states that in order to incorporate given new beliefs, only absolutely necessary changes have to be made in a knowledge base. This means, after the incorporation of the new beliefs, the knowledge base should be as close to the original one as possible, in an information theoretic sense. The **revision operation** has been introduced as a belief change operation that applies new beliefs respecting this principle [7]. From the perspective of knowledge based systems, further properties a revision operation should satisfy have been formulated as postulates in [1, 5, 13]. How to approach revision algorithmically has been outlined in [6], and computational considerations have been made in [18]. Our work focuses on the revision of probability distributions as it has been introduced in [10]. In this context the revision operation has been

F. Schmidt (✉) · J. Gebhardt
ISC Gebhardt, Celle, Germany
e-mail: schmidt@isc-gebhardt.de

J. Gebhardt
e-mail: gebhardt@isc-gebhardt.de

R. Kruse
Otto-von-Guericke University, Magdeburg, Germany
e-mail: kruse@iws.cs.uni-magdeburg.de

successfully implemented for Markov networks [2, 11] using iterative proportional fitting [21, 23]. This method is well known in the area of statistics and shows beneficial properties for our context. Markov networks are a suitable tool to decompose high-dimensional probability spaces into a number of smaller low-dimensional probability distributions. They belong to a group of techniques called graphical models [15, 16, 19, 24].

The growing complexity and interconnectedness of knowledge bases and an increasing number of new beliefs lead almost inevitably to inconsistencies in the formulation of revision problems. In almost any type of knowledge based systems, inconsistencies render the underlying upon useless and should consequently be addressed. In this contribution we focus on inconsistencies during the revision of probability distributions. This is a multi-facet problem and different aspects of it have been introduced in [22]. Furthermore, two types of inconsistencies and a revision control algorithm have been described in [12].

In this work we focus on the important aspect of detecting the presence of inconsistencies in a given revision problem. In Sect. 2 of this paper, we will formally introduce the revision operation, specify what a revision problem is, and define revision inconsistencies. Section 3 then discusses how the problem of detecting inconsistencies can be approached, deals with different classes of possible solutions as well as a short analysis on the usability of the given classes in our scenario. In Sect. 4 we look at the detection of inconsistencies from the point of view of an application using Markov networks. Section 5 then concludes the paper and provides some ideas for future research.

2   Fundamentals

In this section we will describe the revision operation, define the revision problem, and specify what inconsistencies are in that context.

2.1   The Revision Operation

This work focuses on the revision of probability distributions and we therefore define it in this context.

As mentioned before, the goal of (probabilistic) revision is to compute a posterior probability distribution which satisfies given new distribution conditions, only accepting a minimal change of the quantitative interaction structures of the underlying prior distribution.

More formally, in our setting, a revision operation (see [2, 12]) operates on a joint probability distribution $P(V)$ on a set $V = \{X_1, \ldots, X_n\}$ of variables with finite domains $\Omega(X_i), i = 1, \ldots, n$. The purpose of the operation is to adapt $P(V)$ to new sets of beliefs. The beliefs are formulated in a so-called **revision structure** $\Sigma =$

$(\sigma_s)_{s=1}^S$. This structure consists of **revision assignments** σ_s, each of which represents a low dimensional (conditional) probability assignment. The pair $(P(V), \Sigma)$ is called **revision problem**.

The result of the revision, and solution to the revision problem, is a probability distribution $P_\Sigma(V)$ which

- satisfies the revision assignments (the postulated new probabilities)
- preserves the probabilistic interaction structure as far as possible.

By preserving the interaction structure we mean that, except from the modifications induced by the revision assignments in Σ, all probabilistic dependencies of $P(V)$ are to be invariant. This requirement ensures that modifications are made according to the principle of minimal change.

It can be proven (see, e.g. [2]) that in case of existence, the solution of the revision problem $(P(V), \Sigma)$ is uniquely defined. This solution can be determined using iterative proportional fitting [23, 24]. Starting with the initial probability distribution, this process adapts the initial probability distribution iteratively, one revision assignment at the time, and converges to a limit distribution that solves the revision problem, given there are no inconsistencies.

2.2 Inconsistencies in the Context of the Revision Operation

Inconsistencies in the context of revising probability distributions have been analysed in [12], and two types of inconsistencies of revision problems have been distinguished, which are *inner inconsistencies* and *outer inconsistencies*, respectively.

Inner consistency of a revision structure Σ is given, if and only if a probability distribution exists that satisfies the revision assignments of Σ; otherwise we refer to **inner inconsistencies** of Σ.

In Fig. 1, a simple example is shown where the given revision assignments contradict each other and hence do not form a single probability distribution. The filled entries in the left table represent the revision assignments. In the right table consequences for the rest of the table are shown and one conflict is highlighted.

Given that there is a probability distribution that satisfies Σ, it is still possible that due to the zero probabilities of $P(V)$ the revision problem $(P(V), \Sigma)$ is not
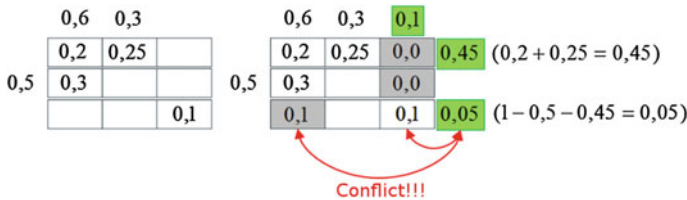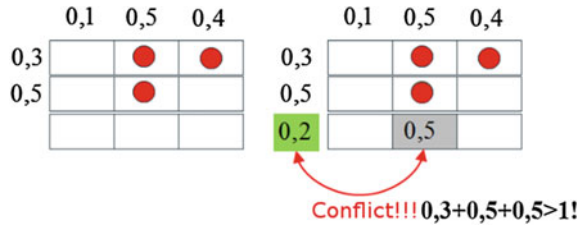

Fig. 1 Inner inconsistency

Fig. 2 Outer inconsistency

solvable. This is the case when one of those zero values would need to be modified in order to satisfy the revision assignments. Such a modification of the interaction structure of $P(V)$ is not permitted during a revision operation. Therefore, a second type of inconsistency is defined as follows:

Given that Σ has the property of inner consistency, the revision problem $(P(V), \Sigma)$ shows the property of **outer inconsistency**, if and only if there is no solution to the revision problem.

Figure 2 illustrates an outer inconsistency. In the left table again the numbers represent revision assignments. This time there are additional circles representing zero values that cannot be changed during the revision operation. As before, the right table shows consequences for the remaining table entries as well as an inconsistency.

3 Detection

Detecting the presence of inconsistencies amounts to calculating the posterior probability given some evidence and is therefore NP-hard [3, 25]. Hence, to determine consistency we have to attempt the construction of a posterior probability distribution. If the construction is successful, the revision problem shows the property of consistency. This is true for both types of inconsistencies we defined earlier. In fact both problems can be transformed into one another. If one can solve the first problem, one can solve the second problem by adding revision assignments representing the zero values. The second problem is actually a generalisation of the first one - there are simply no zero values present. Hence, by solving the second problem one can solve the first one as well.

From this observation, we can infer that both problems have roughly the same degree of complexity, where the first problem most likely needs less effort to calculate. In the literature we found two general approaches to construct a high dimensional probability distribution from lower dimensional probability statements, namely algorithms that find either an *approximating solution* or *exact solutions* if there is one.

3.1 Approximative Algorithms

There is a whole class of algorithms for finding entropy maximising solutions based on the uniform distribution. More specifically, for Markov networks there are, for example, parameter estimation methods based on maximum likelihood and maximum entropy [8, 15, 20]. These methods are potentially faster than the exact methods and always give a result (either the exact one in case of consistency or an approximation in case of inconsistencies). In order to use this kind of methods to detect inconsistencies, one can follow a two-step process:

1. Create a candidate probability distribution
2. Check whether all revision assignments (and zero values) are satisfied

The first step is potentially faster than using an exact method. The second step, which becomes necessary because we don't know whether we have an exact solution or an approximation, may require a significant number of checks.

3.2 Exact Algorithms

Methods based on iterative proportional fitting that do not use approximations to speed up the process find entropy maximising solutions, can be based on any probability distribution, not just the uniform distribution. However, in case of inconsistencies there are multiple limit distributions satisfying different subsets of revision assignments. A single unique solution can only be obtained in the case of consistency. In addition to this disadvantage, they are potentially slower since they are not sacrificing accuracy for performance.

From a mathematical point of view, detecting inconsistencies with these methods is straightforward. In case of consistency the iterative proportional fitting converges towards a single unique probability distribution, which then also solves the revision problem. Otherwise, it will find multiple limit distributions, each of which is satisfying a different subset of revision assignments. In practice, the problem is to decide which of the two cases is present.

3.3 Further Remarks

In practical applications, detection is often embedded in the process of revising probability distributions. For that reason, it is interesting to analyse whether the constructed distributions already sufficiently solve the actual revision problem.

The approximative methods always deliver a distribution, even if inconsistencies are present. This is a useful property for working with real world problems. However, those methods maximise entropy towards the uniform distribution which is not what

we need in our application. We found approaches in the literature that ,theoretically, would make those methods maximise towards a specific non-uniform distribution [4]. However, that would entail adding a large number of constraints to indicate all the deviations of the wanted prior distribution from the uniform distribution. We believe that the necessary effort then neglects the performance advantage due to the additional constraints.

The exact methods work with any kind of prior probability distribution and maximise entropy against those. If they find a unique solution, it is also a suitable solution for our revision problem. If inconsistencies are present, no unique solution can be obtained. Nevertheless, for the revision of Markov networks, an approach has been proposed in [14], that can resolve inconsistencies in a way that the resulting distribution solves the revision problem that is information theoretically closest to the original problem.

4 Practical Application Using Markov Networks

In our practical application we use Markov networks to efficiently represent probability distributions. In this application the detection of inconsistencies is not a separate processing step, but it is embedded in an overall revision control mechanism that detects inconsistencies, removes them and finally calculates the solution for the (then possibly) modified revision problem. Consequently, we use an exact approach based on iterative proportional fitting and the automatic elimination of inconsistencies proposed in [14].

Since we use the revision of Markov networks we can leverage the benefits of a decomposed probability distribution. This is done implicitly through the revision algorithm, which uses propagation. The propagation algorithm as described in [17] efficiently exploits the decomposition.

As mentioned previously, the problem of detecting inconsistencies in this setting is to decide whether the algorithm converges towards a single distribution or is oscillating between multiple competing distributions.

We identified several interconnected challenges when trying to decide whether convergence is reached. In industrial applications any algorithm has to deliver a result within a reasonable amount of time. Consequently, the number of iterations is usually limited. Therefore, after that limit, the algorithm has to decide whether convergence will be reached or not. We use a measure based on the sum of the differences between revision assignments and their actual value in the distribution. This method works well in many cases. However, we still have problems when the process converges slowly, or runs into a local minimum.

5 Conclusion

Detecting inconsistencies in revision problems is an important topic when using revision to adapt knowledge to new beliefs. In this work we discussed different approaches to detect inconsistencies in revision problems when using probabilistic revision. Both presented types of inconsistencies can be detected using very similar approaches. In this work we analysed two different classes of methods to detect inconsistencies using constructive approaches. Both classes have their advantages and disadvantages. In our setting we prefer the exact methods since, with slight modifications, they allow us to use the detection and elimination of the occurring inconsistencies in one step, and at the same time, they provide a usable solution to our revision problem. However, under different requirements approximative methods can potentially be better suited.

In the future our findings need to be verified by running tests on data from different real world applications. Furthermore, although we did not find an approach to test for inconsistencies other than to attempt the construction of a probability distribution, there might be techniques in areas like statistics that obtain a solution faster and with less calculation. Additionally, the problems with slow convergence and local minima are of interest.

References

1. Alchourrón CE, Gärdenfors P, Makinson D (1985) On the logic of theory change: partial meet contraction and revision functions. J Symbolic Logic 50(02):510–530
2. Borgelt C, Kruse R (2004) Knowledge revision in Markov networks. Vm075031.Usc.Es, 11:93–107
3. Cooper GF (1990) The computational complexity of probabilistic inference using bayesian belief networks. Artif Intell 42(2–3):393–405 March
4. Csiszar I (1975) I-Divergence geometry of probability distributions and minimization problems. Ann Probab 3(1):146–158
5. Darwiche A (1997) On the logic of iterated belief revision. Artif Intell 89(1–2):1–29
6. Gabbay D (2003) Controlled revision—an algorithmic approach for belief revision. J Logic Comput 13(1):3–22
7. Gabbay D, Smets P (eds) (1998) Handbook of defeasable reasoning and uncertainty management systems, vol 3, Belief change. Kluwer Academic Press, Netherlands
8. Ganapathi V, Vickrey D, Duchi J, Koller D (2008) Constrained approximate maximum entropy learning of markov random fields. In: Proceedings of uncertainty on artificial intelligence, pp 196–203
9. Gärdenfors P (1988) Knowledge in flux: modeling the dynamics of epistemic states. MIT Press
10. Gebhardt J, Detmer H, Madsen AL (2003) Predicting parts demand in the automotive industry—an application of probabilistic graphical models. In: Proceedings of international joint conference on uncertainty in artificial intelligence
11. Gebhardt J, Klose A, Detmer H, Ruegheimer F, Kruse R (2006) Graphical models for industrial planning on complex domains. Decis Theory Multi-Agent Plann 482:131–143
12. Gebhardt J, Klose A, Wendler J (2012) Markov network revision: on the handling of inconsistencies. Computational intelligence in intelligent data analysis, vol 445 of Studies in computational intelligence. Springer, Berlin, pp 153–165

13. Katsuno H, Mendelzon AO (1991) Propositional knowledge base revision and minimal change. Artif Intell 52(3):263–294
14. Klose A, Wendler J, Gebhardt J, Detmer H (2012) Resolution of inconsistent revision problems in Markov networks. Synergies of soft computing and statistics for intelligent data analysis, vol 190 of Advances in intelligent systems and computing. Springer, Berlin, pp 517–524
15. Koller D, Friedman N (2009) Probabilistic graphical models: principles and techniques. MIT Press
16. Kruse R (2013) Computational intelligence., A methodological introductionSpringer, London
17. Lauritzen SL (1992) Propagation of probabilities, means and variances in mixed graphical association models. J Am Stat Assoc 87(420):1098–1108
18. Nebel B (1994) Base revision operations and schemes: representation, semantics and complexity. In: Proceedings of the eleventh European conference on artificial intelligence (ECAI94), Amsterdam, The Netherlands. Wiley, pp 341–345
19. Pearl J (1991) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann
20. Pietra SD, Pietra VD, Lafferty J (1997) Inducing features of random fields. IEEE Trans Pattern Anal Mach Intell 19(4):380–393
21. Pukelsheim F, Simeone B (2009) On the iterative proportional fitting procedure: structure of accumulation points and L1-error analysis. Structure (05):28
22. Schmidt F, Wendler J, Gebhardt J, Kruse R (2013) Handling inconsistencies in the revision of probability distributions. In: Pan J et al. (ed) HAIS13, vol 8073 of Lecture notes in computer science. Springer, Berlin, pp 598–607
23. Teh YW, Welling M (2003) On improving the efficiency of the iterative proportional fitting procedure, pp 0–7
24. Whittaker J (1990) Graphical models in applied multivariate statistics. Wiley
25. Yu H, Engelen R (2011) Symbolic and quantitative approaches to reasoning with uncertainty: 11th European conference, ECSQARU 2011, Belfast, UK, June 29–July 1, 2011. Proceedings, chapter importance. Springer, Berlin, pp 146–157