ELECTRONIC NEWSLETTER ON
REASONING ABOUT ACTIONS AND CHANGE

Issue 98068 Editor: Erik Sandewall 14.9.1998

Today

Today's issue contains David Poole's answer to Uffe Kjaerulff's questions re his ETAI submitted article.

ETAI Junction - New Webpage Structure for ETAI

A new structure and layout for the ETAI web pages has been developed during the summer months. The new design is now sufficiently complete to be released, and it has just replaced its predecessor at

`http://www.ida.liu.se/ext/etai/`

The main features of the new design are:

- More systematic presentation of auxiliary information, such as information for authors, information for guests
- A uniform structure, common for all the ETAI areas
- More professional layout and graphics

We hope you will appreciate this innovation, and as usual we welcome your comments about it.

We wish to reserve the acronym ETAI for the **Transactions**, that is, for our journal publication. It is therefore not very appropriate to talk about "the ETAI webpage" since it suggests that ETAI would be first and foremost "just" a webpage. We will therefore use the term **the ETAI Junction** for the webpage structure that contains all the information pertaining to the ETAI as a journal as well as the review dicussions, colloquia, and other surrounding information.

ETAI Publications

Discussion about received articles

Decision Theory, the Situation Calculus and Conditional Plans

by **David Poole**

Comment from: **David Poole**

Uffe Kjaerulff writes:

> *The focus of the paper seems to be on representational issues, which is fine. But, being a Bayesian-network person, I would find it interesting to see how computations of expected utilities of plans are performed in the ICL framework. This issue is only touched upon very briefly in the paper, mentioning that it involves enumerating all possible cases. In my understanding that amounts to constructing the corresponding decision tree (which may be huge), and then computing the probability and utility for each leaf of the tree.*
>
> *So, my question boils down to: Is there an efficient computational scheme for ICL?*

Thanks for the question! First, ICL theories can be seen as representations of Bayesian networks; there is a local translation between groundings of the ICL theories and Bayesian networks. ICL theories (with only choices by nature) can be seen as first-order rule-based representations of Bayesian networks. Thus any Bayesian network algorithms can, in principle, be used for the ICL. There are two main computational advantages of the ICL approach. The first is that they can exploit logic-based reasoning techniques for probabilistic reasoning (although it turns out that once the link is realized, it isn't restricted to ICL theories). Secondly they give us an opportunity to exploit, not only the graphical independence of a Bayesian network, but also the contextual independence that i natural in the rule based representations. I will expand briefly on each of these here.

The ICL (with only choices by nature) can be seen as a way to write first-order Bayesian networks as rules. There is a local translation between ICL theories and Bayesian networks. Note that a Bayesian network has nothing to say about how a node depends on its parents. The ICL lets us write the conditional probabilities as a set of rules. (This is often a more compact representation than conditional probability tables when there are contextual independencies).

To translate an ICL theory into a Bayesian network: ground the theory (substituting ground terms for the free variables); the ground atoms are the nodes of the Bayesian network; the atoms in the bodes of the rules that imply a become the parents of a. There is a lot more structure that can be expressed by the rules than in the network (in particular the contextual independencies). Note that the acyclicity of the logic program ensures that the resulting network is acyclic and (even when the resulting network is infinite) there are only finitely many ancestors for any node (which means you only need a finite subset of the graph to answer any conditional probability statement).

To translate a Bayesian network to an ICL theory, you write as rules how the parents of a influence a. The atomic choices are introduced as extra conditions on the rules to let us interpret the rules logically. For more details see (Poole 1993).

The first way that the ICL can be used computationally is adapting the logic-based algorithms for probabilistic reasoning. In particular, the algorithms for consistency-based diagnosis (such as the focusing algorithms of de Kleer), based on the notions of conflicts, can be adapted to compute posterior probabilities in the ICL (and so for Bayesian networks). This works well when there are skewed probability distributions. This has been explored in (Poole 1996).

Another approach is to exploit the structure that can be expressed in a Bayesian network as well as the rule structure of the ICL. This is ongoing work with Nevin Zhang. I'll first sketch how the efficient structured Bayesian network algorithms (such as clique-tree propagation and more query-oriented algorithms such as D'Ambrosio's SPI, Zhang and Poole's VE or Dechter's bucket elimination) can be seen as ways to exploit the factorization of the joint probability distribution. The basic operation is to sum out a (non-observed, non-query) variable. Algorithms such as clique-tree propagation are efficient because they can compile this operation into an secondary structure (the clique tree or join tree) that works by message passing in the join tree so that the posterior distribution for all variables can be obtained in twice the time it takes to compute the posterior on one. The basic operation of the message passing from one clique to another is to sum out the variable that appears in one clique that doesn't appear in the other.

It turns out that you can carry out a similar operation on the rule base directly. We have been working on a rule-based variation of the query-oriented algorithms I call probabilistic partial evaluation (Poole 1997). Eliminating a variable turns out to be like resolution. When the rules give no structure beyond that of the Bayesian network (i.e., there is no contextual independence, and there is a rule for each assignment of variables to a parent), the algorithm reduces to the standard query-oriented Bayesian network algorithms. But when there is contextual independence (as there is in many ICL theories) the algorithm lets us exploit that. This also seems to be very promising for approximation, where different approximations can be used in different contexts (Poole 1998).

Aside: One main difference between clique tree propagation and the query-oriented algorithms is that clique tree propagation acts on marginal distributions (each clique represents the marginal distribution of the variables in the clique), whereas the query-oriented algorithms act on conditional probabilities. Probabilistic partial evaluation relies on manipulating conditional probabilities. The contextual independencies get lost quickly when you reason with marginal distributions.

Finally, there is also a promise of structured dynamic programming using the rules. The most comment methods for solving sequential decision problems is to do dynamic programming. Traditional dynamic programming explicitly reasons over state space.s and thus suffers from what Bellman calls the "curse of dimensionality". We should be able to do much better by reasoning in terms of the variables (or propositions) that define the state space and exploiting the rule-structure in dynamic programming. In the ICL, dynamic programming looks much like regression planning. We can either do this in POMDPs where we regress values through conditional plans (Boutilier and Poole 1996) or where we are building plans conditional on history (as in influence diagrams), and we can use the rule structure to determine states that all have the same utility (Poole 1995).

Thus in answer to your question, this is an active area of research by myself and others who are looking at exploiting contextual independence for probabilistic inference and sequential decision making. One of the reasons I am optimistic that this will work is that it promises nice methods of approximation, where we can approximate differently in different contexts. We will have to see if we can deliver on this promise.

References

The following are references to some of my papers that give more details (the HTML version gives hyperlinks to the actual papers):

C. Boutilier and D. Poole, "Computing Optimal Policies for Partially Observable Decision Processes using Compact Representations", Proc. Thirteenth National Conference on Artificial Intelligence (AAAI-96), Portland, Oregon, August 1996.

D. Poole, "Probabilistic Horn abduction and Bayesian networks", Artificial Intelligence, 64(1), 81-129, 1993.

D. Poole, "Exploiting the Rule Structure for Decision Making within the Independent Choice Logic", Proceedings of the Eleventh Conference on Uncertainty in AI, Montreal, August 1995, 454-463.

D. Poole, "Probabilistic conflicts in a search algorithm for estimating posterior probabilities in Bayesian networks", Artificial Intelligence, 88, 69-100, 1996.

D. Poole, "Probabilistic Partial Evaluation: Exploiting rule structure in probabilistic inference", Proc. Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97), Nagoya, Japan, August 1997, pp. 1284-1291.

D. Poole, "Context-specific approximation in probabilistic inference", Proc. Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI-98), Madison, Wisconsin, July 1998.

N.L. Zhang and D. Poole, "Exploiting Causal Independence in Bayesian Network Inference", Journal of Artificial Intelligence Research, 5, 301-328, 1996.

This newsletter is issued whenever there is new news, and is sent by automatic E-mail and without charge to a list of subscribers, in plaintext format. The present postscript issue is a bona fide rendering of the same information in alternative typography. Instruction for contributors and other additional information is found at the following URL:

`http://www.ida.liu.se/ext/etai/actions/njl/`