MÖBUS, CLAUS & SEEBOLD, HEIKO, *A Greedy Knowledge Acquisition Method for the Rapid Prototyping of Bayesian Belief Networks*, in: Ch-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (eds), Artificial Intelligence in Education, Amsterdam: IOS Press, 2005, p. 875 - 877, ISBN 1-58603-530-4

A Greedy Knowledge Acquisition Method for the Rapid Prototyping of Bayesian Belief Networks

Claus MÖBUS, Heiko SEEBOLD
Learning Environments and Knowledge Based Systems
Department of Computing Science, University of Oldenburg
D-26111 Oldenburg, Germany

Abstract. Bayesian belief networks (BBNs) are a standard tool for building intelligent systems in domains with uncertainty for diagnostics, therapy planning and user-modelling. Modelling their qualitative and quantitative parts requires sometimes subjective data acquired from domain experts. This can be very time consuming and stressful - causing a knowledge acquisition bottleneck.

The main goal of this paper is the presentation of a new knowledge acquisition procedure for rapid prototyping the qualitative part of BBNs. Experts have to provide only simple judgements about the causal precedence in pairs of variables. From these data a new greedy algorithm for the construction of transitive closures generates a Hasse diagram as a first approximation for the qualitative model. Then experts provide only simple judgements about the surplus informational value of variables for a target variable shielded by a Markov blanket (wall) of variables. This two-step procedure allows for very rapid prototyping. In a case-study we and two expert cardiologists developed a first 39 variables prototype BBN within two days.

Keywords. Knowledge Acquisition, Acquisition of Uncertain Causal Knowledge, Greedy Construction of DAGs in Bayesian Network Models, Greedy Construction of Hasse Diagrams and Transitive Closures, Acquisition of Causal Precedence Relations

Introduction

BBNs are relevant for the success of intelligent systems in assessing or modelling uncertain knowledge. The classical procedure for the construction of BBNs under the knowledge based approach was published by Pearl as the boundary strata method (BSM) [1]. The BSM is presented in many textbooks [2] and online tutorials. Because of its cognitive demanding aspects it is unsuitable for domain experts without modelling experience. The most problematic aspect of the procedure is the determination of a minimal set of direct influencers for a selected variable under the constraint of independence properties. Our experts had problems distinguishing between influencers and direct influencers, especially when a forgotten variable had to be included in the model again. In that case direct influencers could become indirect influencers.

This led to the development of a computerized procedure with a new greedy algorithm for the determination of transitive closures at anytime. This algorithm controls the selection of pairs, guarantees that the data comprise a partial order relation (POR) and generates the Hasse diagram of the POR (Hasse model). In the best case the monitor acquires the Hasse model of the causal precedence relation in just one pass. The savings in pair-comparisons are then (1-2/n)*100%, the judgement complexity is O(n) and the computational complexity is O(n3). If the Hasse model also passes a Markov blanket independence test, the Hasse model is without further modifications the DAG of the BBN. In the worst case the monitor needs n(n-1)/2

MÖBUS, CLAUS & SEEBOLD, HEIKO, *A Greedy Knowledge Acquisition Method for the Rapid Prototyping of Bayesian Belief Networks*, in: Ch-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (eds), Artificial Intelligence in Education, Amsterdam: IOS Press, 2005, p. 875 - 877, ISBN 1-58603-530-4

comparisons. The judgement complexity is O(n2) and the computational complexity stays O(n3). If the Hasse model does not pass the Markov blanket test, there is a lack of influences (or links). These must then be added back into the Hasse model. The modified DAG is then considered as the qualitative model of the BBN. Despite its flexibility, the computational complexity of the greedy algorithm is only O(n3).

The new method was successfully used to design and implement a BBN-based eLearning system for problem oriented diagnostics in aortic stenosis [3]. The knowledge acquisition for the complete model of the first prototype with 39 nodes (pair-comparisons, Markov blanket tests and estimation of conditional probability tables) could be accomplished in a two-day crash-course workshop.

A New Greedy Method for the Acquisition of DAGs in BBNs

The greediness of the new method stems from the fact that after each data input it determines which not-yet-acquired pairs are informative for the construction of Hasse diagrams. The *best case data acquisition complexity* is O(n) and the *worst case computational complexity* $O(n^3)$.

When a pair (j, i) is presented subjects have to select a judgment from a set of alternatives {"i causes/precedes j", "i follows j", "i neither causes/precedes nor follows j" } internally abbreviated as {+(j,i), -(j,i), 0(j,i)}}. Though the greedy algorithm does not presuppose a special order in the data acquisition events, we selected a special order of pair comparisons along the main diagonal of the adjacency matrix. If it possible to order the variables according to some vague causal hypothesis we support the algorithm working along the main diagonal thus maximizing the number of inferences and reducing at the same time the pair comparison workload of the domain experts.

We demonstrate the algorithm with an example. First we take the DAG from Fig.1.1 as the "mental model" of the experts. Nodes are already numbered according an *ancestral ordering.*

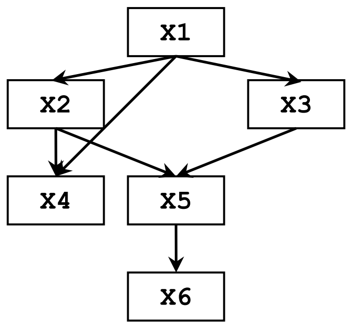

Fig. 1 - DAG of *true* model (TrM)

	1	2	3	4	5	6
1	/	$+_1$	$+_6$			
2		/	0_2	$+_7$	$+_{10}$	
3			/	0_3	$+_8$	
4				/	0_4	0_9
5					/	$+_5$
6						/

Reduction in number of pair comparisons: 33%

Tab. 1 – data acquired under *greedy* algorithm

The algorithm asks for data from the expert working above the main diagonal from top-left to bottom-right when the cell d(j,i) is empty. Diagonals move from the main diagonal in the middle of the matrix to the right upper corner. Cells are marked with "+(j,i)" (i causes/precedes j), "-(j,i)" (i follows j), "0(i,j)" (no order relation between i and j) and "/(i,j)" (transitive or reflexive cell: not necessary for Hasse diagram). Each cell entry in Tab. 1 has an index which marks the step number of the algorithm $\text{<step-nr>}_{\text{<entry>}}$. The behaviour of the algorithm is controlled by 13 *inference rules* (Tab. 2) which are triggered after any new data entrance in cell d(i,j), and which can trigger each other by *recursive calls. The rule*

MÖBUS, CLAUS & SEEBOLD, HEIKO, *A Greedy Knowledge Acquisition Method for the Rapid Prototyping of Bayesian Belief Networks*, in: Ch-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (eds), Artificial Intelligence in Education, Amsterdam: IOS Press, 2005, p. 875 - 877, ISBN 1-58603-530-4

set is complete and can be made commutative, if we enrich the conditions of the rules appropriately.

Nr. of rule	rule
	mirroring data and inferences
1	$+(i,j) \wedge \neg -(j,i) \Rightarrow -(j,i)$
2	$++(i,j) \wedge \neg --(j,i) \Rightarrow --(j,i)$
3	$-(i,j) \wedge \neg +(j,i) \Rightarrow +(j,i)$
4	$--(i,j) \wedge \neg ++(j,i) \Rightarrow ++(j,i)$
5	$0(i,j) \wedge \neg 0(j,i) \Rightarrow 0(j,i)$
	rowwise inferences k=1,...,n
6.1	$+(i,j) \wedge +(j,k) \wedge \neg ++(i,k) \Rightarrow ++(i,k)$
7.1	$+(i,j) \wedge ++(j,k) \wedge \neg ++(i,k) \Rightarrow ++(i,k)$
8.1	$++(i,j) \wedge +(j,k) \wedge \neg ++(i,k) \Rightarrow ++(i,k)$
9.1	$++(i,j) \wedge ++(j,k) \wedge \neg ++(i,k) \Rightarrow ++(i,k)$
	columnwise inferences k=1,...,n
6.2	$+(k,i) \wedge +(i,j) \wedge \neg ++(k,j) \Rightarrow ++(k,j)$
7.2	$+(k,i) \wedge ++(i,j) \wedge \neg ++(k,j) \Rightarrow ++(k,j)$
8.2	$++(k,i) \wedge +(i,j) \wedge \neg ++(k,j) \Rightarrow ++(k,j)$
9.2	$++(k,i) \wedge ++(i,j) \wedge \neg ++(k,j) \Rightarrow ++(k,j)$

Tab. 2 - inference rules for controlling the greedy algorithm

Tab. 1 shows that we need only 10 judgements, whereas a naïve acquisition of every possible pair would take n(n-1)/2 = 15 comparisons. This 33% more efficient. Taking only the +(j,i) order information from the transitive closure, (Tab. 3) we can reconstruct the Hasse diagram (Fig. 2).

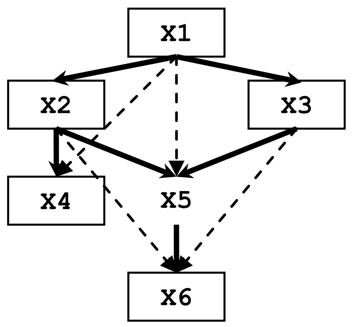

Fig. 2 - *Hasse* model reconstructed from transitive closure of input data

	1	2	3	4	5	6
1	/	+	+	++	++	++
2	-	/	0	+	+	++
3	-	0	/	0	+	++
4	--	-	0	/	0	0
5	--	-	-	0	/	+
6	--	--	--	0	-	/

Tab. 3 - transitive closure of *input* data generated by the *greedy* algorithm

References

[1] Pearl, J., 19982, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (Revised Second Printing), Morgan Kaufman Publishers, San Mateo, CA., ISBN 0-934613-73-7

[2] Russell, St., Norvig, P., 2003, Artificial Intelligence. A Modern Approach, Upper Saddle River, N.J.: Pearson Education, Inc., ISBN 0-13-080302-2

[3] Seebold, H., Lüdtke, A., Möbus, C., 2005, Bayesian Belief Network based Diagnostics in a Problem-oriented Learning Environment for Cardiology. Proceedings of Training, Education & Simulation International 2005 (TESI 2005), Maastricht, The Netherlands, http://www.tesi2005.com/bookofproceedings/