Explanations in Bayesian Networks using Provenance through Case-based Reasoning

Anders Kofod-Petersen, Helge Langseth, and Agnar Aamodt

Department of Computer and Information Science,
Norwegian University of Science and Technology,
7491 Trondheim, Norway
{anderpe|helgel|agnar}@idi.ntnu.no

Abstract. Bayesian Networks are useful for solving a wide range of problems in many domains. However, they are exposed to one important challenge when structural and parametrical changes occur. As Bayesian networks lack memory regarding their changes over time, there is currently no good way of maintaining a history of changes and their provenance. Thus, any variance in the network's problem solving behaviour will not be explainable to a user. Within the context of systems that integrate CBR and BN, we suggest to add a case-based reasoning functionality that will retain changes and their provenance, as well as approaches to explain any unexpected problem solving behaviour.

1 Introduction

Explanations have been identified as one of the most important properties of intelligent systems (see e.g. [1,2]). One important source of knowledge for generating useful explanations is the information related to changes in used knowledge sources and changes in reasoning processes.

Although providence-related information has been included in some CBR systems since the early days of the field, and often referred to as book-keeping information, justifications, or meta-information, a systematic treatment of providence and its role was introduced at the ICCBR conference in 2007 [3]. They identified two main roles of provenance in CBR; to compensate for delayed feedback in the REVISE phase of the CBR cycle [4], and to improve case-base maintenance. The role suggested here is an additional one, namely to produce explanations for unexpected results that stem from changes in parts of the domain model. In a philosophical treatment of provenance and explanation [5] the focus is on "how" and "why" type explanations related to provenance, and in a treatment of explanation and provenance within a Description Logics framework [5] an empahsis is on how to track "how-provenance" in a system, referring to how sources of provenance contribute to the observed data. In the work reported here the doman model could in principle be a case-specific domain model as well as one containing generalized knowledge, but in the present paper we focus solely on changes in the general domain knowledge, and more specifically in the Bayesian Network component of a system.

In our previous work [6], we have been looking into the benefits of combining Bayesian networks [7,8] and Case-based reasoning into a hybrid system, with the goal to utilise the Bayesian networks either as a reasoning engine in its own right, or to help with the case matching in the integrated system. This is an endeavour we have taken up again recently, and although we in this paper will exemplify the reasoning processes using only a Bayesian network, we do still have the total hybrid system in mind. Bayesian networks constitute one of the most popular modelling framework for uncertain knowledge, and with its formal grounding in probability theory, they have found applications in areas as diverse as genetics, failure detection, recommender systems, and speech recognition. One important argument for using Bayesian networks is the frameworks ability to adapt both its structure [9] as well as its parameters [10] as new training data is presented to the system; properties which are crucial in changing or partly unknown domains.

The nature of Bayesian networks allows for some explanations to be given regarding the reasoning process (see, e.g., [11] for a review of explanations in Bayesian networks). However, there is one important aspect that these models cannot explain, and that is when the conceived behaviour of the system has changed via learning (via adaption of the Bayesian network, or – in our hybrid system – by adding cases to the case-base).

As our system traditionally retain no history of the provenance of such changes made during learning, any variation in results coming from such changes are inexplicable.

The work presented here suggests the use of case-based reasoning to address this issue. In brief, the we suggest to construct cases based on changes in the Bayesian network and use this in combination with a model of the user to construct explanations. Sørmo et al. [12] identified five explanation goals: Explain How the System Reached the Answer (Transparency), Explain Why the Answer is a Good Answer (Justification), Explain Why a Question Asked is Relevant (Relevance), Clarify the Meaning of Concepts (Conceptualization), and Teach the User About the Domain (Learning). Our focus is here on the transparency and justification goals.

The rest of this paper is organised as follows: Section 2 introduces a motivational example for augmenting Bayesian networks with a case-based reasoning provenance part; Section 3 describes how case-based reasoning can be applied and how the cases can be constructed; The paper ends with an outlook on future work.

2 Motivational Example

To exemplify the need for provenance of changes in a Bayesian network we will use a simple example. A very simple Bayesian network, depicted in Figure 1 can be used to tell a user whether to go to the beach or not.

The network can observe the current weather and if it is raining or cloudy will tell the user to stay at home, and if it is sunny it will tell the user to go to the beach. The column marked 'Historical' in Table 1 shows the mapping between

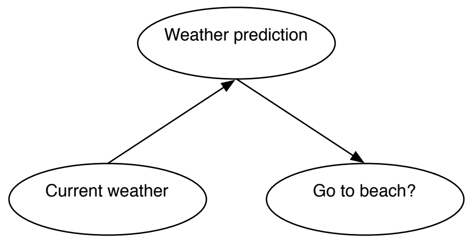

Fig. 1. Bayesian network

weather conditions and expected suggestion. This explicit mapping is naturally not visible for the user. However, we can assume that a user who uses any system on a regular basis will develop some expectation as to how the system behaves.

Table 1. Condition and response

Condition	**Historical recommendation**	**New recommendation**
Rain	Stay at home	Stay at home
Cloudy	Stay at home	Go to beach / stay at home
Sunny	Go to beach	Go to beach

Let us assume that the system now gets equipped with a barometer, which to some degree can be used to predict the weather. The new Bayesian network, including the barometer is depicted in Figure 2. Using this new device the network is able to predict that cloudy weather in some situation will improve and become sunny. Thus, the system will sometimes suggest the user to go to the beach even when it is cloudy, for example when the barometer shows a trend towards higher pressure. This new observed behaviour is shown in the column marked 'New' in Table 1.

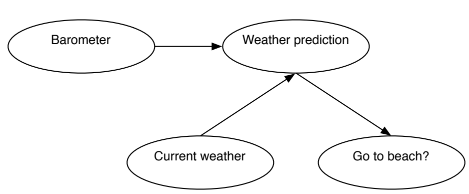

Fig. 2. Modified Bayesian network

The first time the system makes an *unexpected* prediction, i.e. suggest going to the beach even when it is cloudy, the users would likely request an explanation. The current Bayesian network can supply an explanation regarding the relation between the observed variables (the current weather and the barometer reading), the weather forecast, and the query variable. This type of explanation is clearly not sufficient as it does not include the real provenance of the change in the reasoning process. The user is quite likely to request a justification of the change in the expected outcome.

3 Case-based reasoning as explanation mechanism

The work presented here suggests to augment Bayesian networks with a case-based reasoning component to explain changes in behaviour. This is achieved by retaining the provenance of either parametric or structural changes in the problem-solving network.

When a request for an explanation is received, the case-based reasoning system will retrieve a case containing a relevant part of the network that solved the problem (in the example predicting whether or not to go to the beach), and the current network. The case base contains snapshots of partial BNs where the solution the BN gave triggered an explanation question from the user. This indicates an unexpected solution or insufficient confidence in the suggested solution, for which a transparency and/or justification explanation is called for. Figure 3 depicts an example case.

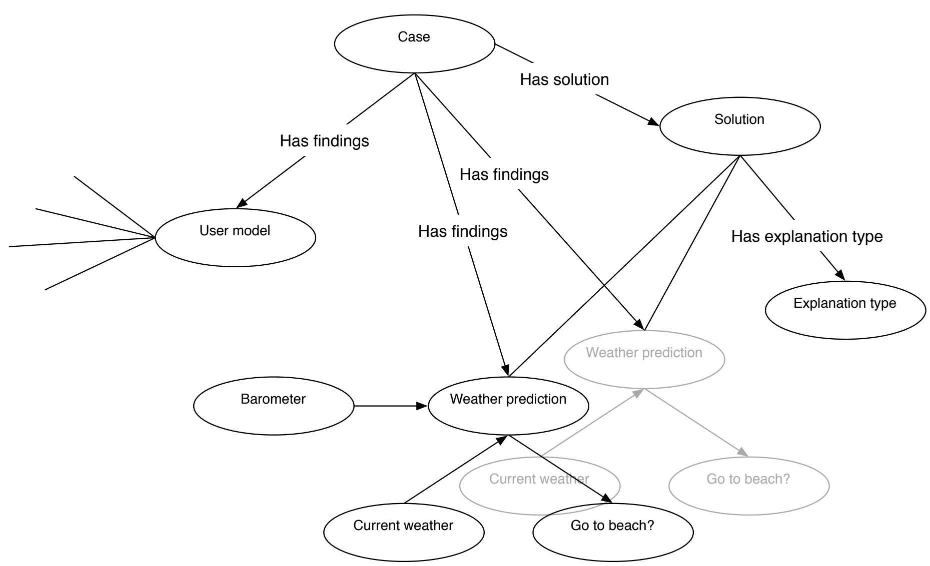

Fig. 3. Case structure

In the problem description part of the a case, the findings are: a user model, which can be used to influence the explanation goal to be solved and the way to present it; the current network; and the last known accepted network (marked in light grey). The solution to the case contains the explanation goal and the two Bayesian networks. The latter can be used as a knowledge container [13] where relevant knowledge to construct the explanation can be acquired.

New cases are constructed whenever a user asks a question to the system. In this example, when a request for a recommendation for visiting the beach is made. If the answer to the question and the input parameters are sufficiently similar to an existing case, no new case is retained. In case of the user requesting an explanation, as an example in the form of: "The last time it was cloudy, you suggested me to stay at home, but now to tell me to go to the beach, why?", the current case is matched against the case base and the lasted case that gave the 'correct' answer is retrieved. The current case contains the current Bayesian network, which is then matched against the network in the retrieved case. The difference between these two cases form the basis for the explanation.

To continue the example from Section 2 and augment it with the case-based reasoning provenance system, we can now supply an explanation as to why the answer differs from the expected one. Comparing the two networks in questions (see Figure 3) shows that the current network contains a new node, the barometer. This barometer affects the systems ability to predict the weather. Given that it is cloudy and the barometer says that there is a high pressure, it will presumably be sunny shortly. So the explanation offered by the system might be: "Since the last time that cloudy weather was observed, when I suggested to stay at home, I have received a barometer. The barometer tells me that even though it is cloudy, there is a high pressure. Thus, it is very likely that it will be sunny at the beach."

Augmenting the Bayesian network with the case-based reasoning system allows the system to supply a *transparency* explanation, to e.g. experts, and a *justification* explanation, to e.g. novices (for a description of the relationship between user competencies and transparency versus justification, see e.g. [14])

4 Further work

We are currently investigating suitable approaches to implement the suggestions made in this paper. A suitable implementation in a domain closer to the real-world must be carried out in order to test the validity of this Bayesian networks and case-based reasoning hybrid approach.

Further, it is likely that with a suitable structuring of the knowledge model, other types of explanation are possible, either directly from the Bayesian network or from the case-based reasoning provenance system.

Finally, other types of relevant provenance information, such as the context of the problem-solving system and the rationale for changes in the problem-solver's changes, could be combined with the suggested approach to further improve the users' trust and acceptability of reasoning systems.

References

1. Leake, D.B.: Evaluating Explanations: A Content Theory. Lawrence Erlbaum Associates (1992)
2. Roth-Berghofer, T.R., Cassens, J.: Mapping goals and kinds of explanations to the knowledge containers of case-based reasoning systems. In Muñoz-Avila, H., Ricci, F., eds.: Case Based Reasoning Research and Development – ICCBR 2005. Volume 3630 of LNAI., Chicago, Springer (2005) 451–464
3. Leake, D., Whitehead, M.: Case provenance: The value of remembering case sources. Case-Based Reasoning Research and Development (2007) 194–208
4. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications **7**(1) (1994) 39–59
5. Lockie, R.: Knowledge, Provenance and Psychological Explanation. Philosophy **79**(03) (2004) 421–433
6. Aamodt, A., Langseth, H.: Integrating Bayesian networks into knowledge intensive CBR. In: American Association for Artificial Intelligence, Case-based reasoning integrations; Papers from the AAAI workshop – Technical Report WS-98-15, Madison, WI., AAAI Press (1998) 1–6
7. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann (1988)
8. Jensen, F.V., Nielsen, T.D.: Bayesian Networks and Decision Graphs. 2nd edn. Springer (2007)
9. Friedman, N., Goldszmidt, M.: Sequential update of Bayesian network structure. In: Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann (1997) 165–174
10. Spiegelhalter, D.J., Lauritzen, S.L.: Sequential updating of conditional probabilities on directed graphical structures. Networks **20** (1990) 579–605
11. Lacave, C., Díez, F.: A review of explanation methods for Bayesian networks. The Knowledge Engineering Review **17**(2) (2002) 107–127
12. Sørmo, F., Cassens, J., Aamodt, A.: Explanation in case-based reasoning – perspectives and goals. Artificial Intelligence Review **24**(2) (October 2005) 109–143
13. Richter, M.M.: The Knowledge Contained in Similarity Measures. Invited Talk at the First International Conference on Case-Based Reasoning, ICCBR'95, Sesimbra, Portugal (1995)
14. Kofod-Petersen, A., Cassens, J.: Explanations and context in ambient intelligent systems. In kokinov, B., Richardson, D.C., Roth-Berhofer, T.R., Vieu, L., eds.: Modeling and Using Context, proceedings of the 6th International and Interdisciplinary Conference (CONTEXT 2007). Volume 4635 of Lecture Notes in Artificial Intelligence., Roskilde, Denmark, Springer (August 2007) 303–316