ON GRAPHICAL MODELS FOR DYNAMIC SYSTEMS

Lekuona, A.; Lacruz, B.; Lasala, P.
Departamento de Métodos Estadísticos
Universidad de Zaragoza
50009 Zaragoza (Spain)
☎ 34-76-357170
Fax 34-76-565852
e-mail {lekuona/lacruz/lasala}@cc.unizar.es

ABSTRACT

It is widely recognized that probabilistic graphical models provide a good framework for both knowledge representation and probabilistic inference (e.g., see [2],[14]).

The dynamic behaviour of a system which changes over the time needs an implicit or explicit time representation. In this paper, an implicit time representation using dynamic graphical models is proposed.

Our goal is to model the state of a system and its evolution over the time in a richer and more natural way than other approaches together with a more suitable treatment of the inference on variables of interest.

KEYWORDS: dynamic models, bayesian inference, graphical models, chain graphs

1 INTRODUCTION

It is widely recognized that probabilistic graphical models provide a good framework for both knowledge representation and probabilistic inference (e.g., see [2],[14]).

The dynamic behaviour of a certain system which changes over the time needs an implicit or explicit time representation. To model such systems is a very important task: the initial structure of the model and its propagation over the time, the probabilities attached to the structure, the qualitative and quantitative interrelations among variables in different time slices, etc., are several issues to take into account.

In this paper an implicit time representation using dynamic graphical models is focused.

The goal is to model the state of a system and its evolution over the time in a richer and more natural way than other approaches together with a more suitable treatment of the inference on variables of interest, according to bayesian methodology.

The qualitative and quantitative issues in the construction of probabilistic dynamic models are described in detail in Section 2. In Section 3, a formulation of the sequential procedure for making inference over variables is proposed. A briefly comment about related work is shown in Section 4. Finally, in Section 5, conclusions and future work are discussed.

2 GENERAL ISSUES IN THE CONSTRUCTION OF PROBABILISTIC DYNAMIC MODELS

In the following, a probabilistic dynamic model is a sequence of graphs indexed by the time, representing the temporal evolution of a system. Each graph symbolizes the system state and the dependences among its components at a given time. The dynamic behaviour of the system components is described by a set of temporal dependences among these components in different time slices.

Furthermore, these dependences are quantified by conditional probability tables associated with the system components.

In order to make the management of such models easier a set of restrictions must be considered, for both its qualitative and quantitative aspects.

2.1 QUALITATIVE ISSUES

The components that describe any dynamic system to be modeled can be split up into observables and non-observables and they will be depicted by discrete random variables (since the discretization of a continuous variable is always possible, this is not a very restrictive assumption). Observable components collect the evidence of the system measured in each time slice and the main goal will be to make inferences about the present or future state of non-observable components, given the accumulated information.

For the purposes of this paper, the above classification is adequate. However, another more specific types of observable variables could be considered, in order to exploit their features: variables which are independent of time (e.g., sex, chromosomic deficiencies, etc.); variables which tend to keep the same value (e.g., "smoker/non smoker" status, economic status, etc.).

Two types of relationships among variables must be considered: those restricted to the same time slice and those between different time slices.

2.1.1 Relationships in the same time slice

The relationships among variables in the same time slice will be explicited according to the following criteria:

- relationships between observable and non-observable variables will be described by directed edges from the former to the latter (see Figure 1), meaning that the evidence received in the current time slice modifies the beliefs about the states of the non-observable components.
- relationships between non-observable variables will be described by non-directed edges. There is no reason in assuming either causality or sequentiality because these features cannot be observed in such variables.
- relationships between observable variables won't be considered. A parsimonious construction of the model and assuming that every observable variable will always be observed, justify not to consider these relationships.

In Figure 1, a graphic scheme of these criteria is represented. Shaded nodes represent observable variables. The remainder are non-observable variables.

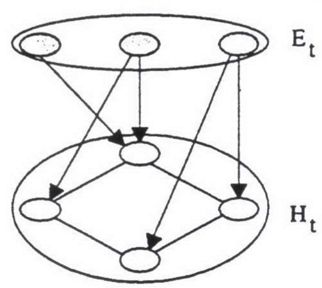

Figure 1

In the first time slice, the special time where the system begins to be observed, the graphical model represents a summary of the behaviour of the system until that moment and the structure in that time (together with the intertemporal structure) will determine the evolution of the system.

2.1.2 Relationships between different time slices

The relationships among variables in different time slices will be explicited according to the following criteria:

- the relationships will be drawn using directed edges, from the older time to the newer time, meaning the natural time sequence.
- only relationships among non-observable variables will be considered (as depicted in Figure 2). On the one hand, the influence of the past in the observable variables is irrelevant, because they will be certainly known sooner or later. On the other hand, we consider that the influence of the past in the the non-observable variables is well described only through the non-observable variables (for instance, let us think in a graph representing the evolution of a patient, who shows certain symptoms (observable variables). These symptoms lead to a set of possible diseases (non-observable variables). It's clear that the information of the symptoms shown in a certain day, must modify the beliefs about the possible diseases in posterior days but this information will be irrelevant if the disease is known). Therefore, the relationships won't be direct but indirect.

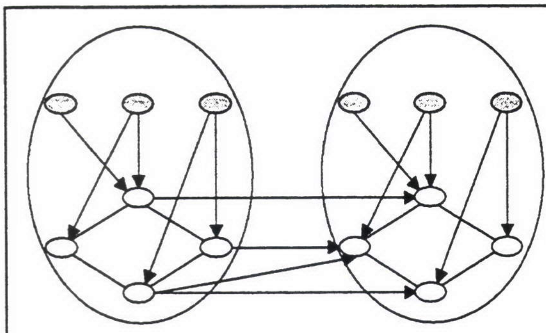

Figure 2

From our point of view, the arrows of the directed edges don't mean cause-effect relationships. They reflect either sequential relationships, inherent to every system evolving over the time, or a natural way of reasoning (first, to observe the evidence and after to guess about the non-observable).

Combining directed and undirected edges forces us to consider chain graphs and to use their Markov properties, described in [7].

2.2 QUANTITATIVE ISSUES

Let us denote

E_t = { observable variables at time t}

H_t = { non-observable variables at time t}

S_t = { E_t, H_t } (it describes the whole system at time t} (see Figure 1)

In order to simplify the model and to make it manageable, a markovian behaviour is attached to $\{S_t\}_{t\geq 0}$, i.e.,

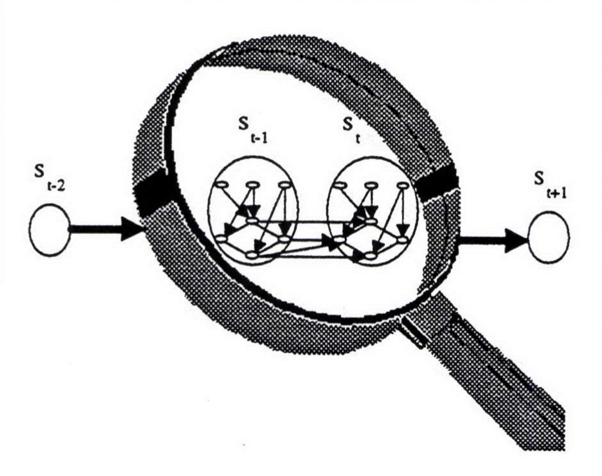

Figure 3

$$S_{t_n} \perp_P S_{t_0},\ldots,S_{t_{n-2}} \mid S_{t_{n-1}},\ \forall t_0 < t_1 < \ldots < t_n$$

where P is a probability distribution properly defined in some product space. This property allows to observe all the past history completely summarized in the more recent past.

Taking into account the markovian condition, arrows from two previous time slices to a posterior time slice won't simultaneously exist in the graphical representation. Furthermore, all the arrows from the past arise from the immediately preceding time slice (see Figure 3).

Finally, all the probabilities and structures discussed in Section 2 must be properly elicitated (using, if possible, a database and/or an expert). In this paper, we don't focus this elicitation.

3 INFERENCE OVER THE NON-OBSERVABLE VARIABLES

We describe the sequential procedure which allows inference and forecasting about the present or future states of the non-observable variables. Let us represent the accumulated evidence up time t by $D_t=\{E_t, D_{t-1}\}$(with $D_0=\{E_0\}$). Four steps are considered:

- *Step 1*: Measure the observable variables in the time t and, therefore, consider D_t.
- *Step 2*: Infer the beliefs of the states of the non-observable variables in the time slice t (see Figure 4). By Bayes:

$$P[H_t|D_t] \propto P[H_t|D_{t-1}]\, P[E_t|H_t, D_{t-1}]$$

and that means a modification of the forecasted beliefs in the previous iteration. We defend to modify the beliefs after receiving new evidence instead of building it again.

In order to exploit the new supplied information, a graph in time t must be considered. While there aren't evidence against the structure used in previous time slices, we propose to keep it (this is defensible in systems not abruptly changing). Therefore, D_t implicitly carries to inherit the structure of the past.

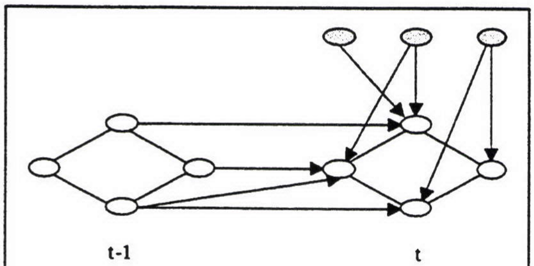

Figure 4

The meaning of the likelihood $P[E_t|H_t,D_{t-1}]$ is clear: given all the past information and assuming a certain state in the non-observable variables, which behaviour would the observable variables have? This likelihood allows the next decomposition:

$$P[E_t|H_t, D_{t-1}] \propto P[S_t|D_{t-1}] =$$
$$= \sum_{H_{t-1}} P[S_t|H_{t-1}]\, P[H_{t-1}|D_{t-1}]$$

where $P[S_t|H_{t-1}]$ can be factorized according to the chain graphs way:

$$P[S_t|H_{t-1}] = \prod_j P[E_t^j] \prod_k P[C_t^k | bd(C_t^k)]$$

where $E_t^j \in E_t\ \forall j$; C_t^k stands for a chain component in H_t and bd(C_t^k) stands for the boundary of that component, as defined in [7].

- *Step 3*: Forecast the states of the non-observable variables in the time slice t+1, assigning beliefs to them. Using the markovian property in $\{S_t\}_{t\geq 0}$, it turns out that

$$P[H_{t+1}|D_t] = \sum_{H_t} P[H_{t+1}|H_t]\, P[H_t|D_t]$$

that is, a weighted mean of transition probabilities associated to the non-observable variables from a period of time to other. The weightings are the actual beliefs in the state of such variables, given all the collected information.

In this forecasting step, the independence of non-observable variables in time t+1, given D_t, must be supposed (the relationships mustn't appear: we are forecasting the state of the variables in the future but we don't actually know (in time t) if the relationships among non-observable variables remain in the future) (see Figure 5). Accordingly, we can factorize the previous formula, using the local markovian property in a chain graph:

$$P[H_{t+1}|D_t] = \sum_{H_t}\left(\prod_j P[H_{t+1}^j|\pi_t(H_{t+1}^j)]\right) P[H_t|D_t]$$

with $H_{t+1}^j \in H_{t+1}$, $\forall j$ and $\pi_t(H_{t+1}^j) = \{H_t^k | \exists \text{ arrow } (H_t^k, H_{t+1}^j)\}$

- *Step 4*: t=t+1 and go to Step 1.

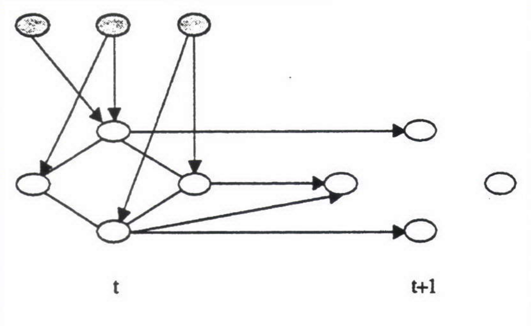

Figure 5

4 RELATED LITERATURE

Cooper, Horvitz and Heckerman [3] propose a severe set of restrictions to be considered in temporal probabilistic reasoning but they do not consider stochastic processes for modeling the temporal evolution.

Kjaerulff [8] proposes a schema for reasoning in dynamic probabilistic networks and he consideres Markov chains for modeling the evolution of the system.

Dagum and Galper [4], Dagum, Galper and Horvitz [5] and Dagum, Galper, Horvitz and Seiver [6] simplify the assessment of conditional probabilities in dynamic networks by employing simple parametric decompositions. Whenever certain dependence conditions don't hold such assumptions are not easily justifiable.

Berzuini [1], Provan [11] and Provan and Clarke [12] propose to use semi-markovian processes. Furthermore, in [11] y [12], they assume that each variable satisfies the Markov property and they infer the markovianity of the whole system (this is only supported under restrictive independence assumptions, from our point of view).

5 CONCLUSIONS AND FUTURE WORK

In this paper, a procedure for modeling dynamic systems has been proposed. We have justified the use of chain graphs instead of directed graphs. The elicitation of the structure of dynamic models using arrows can be very dificult from an expert and almost impossible from a database. Our model only consider arrows for temporal or sequential relationships.

The sequential beliefs updating in our dynamic model is closer to the natural way of thinking of the experts than other approaches that generalize classical time-series analysis or collapse the evidence of the past and present..

Obviously, there are many topics to be focused in the future. The most important are to use databases for acquiring probabilities and learning structures as much as possible; and relaxing certain restrictions (fixed structures; only relationships between two consecutive times; etc.).

ACKNOWLEDGEMENTS

This research was supported by CICYT grant TIC91-1041. The authors thank Peter Cheeseman for enjoyable and helpful discussions of this material.

REFERENCES

[1] BERZUINI, C. Modeling temporal processes via belief networks and Petri nets, with application to expert systems. Annals of Mathematics and Artificial Intelligence, 2 (1990), 39-64.

[2] CHEESEMAN, P.; OLDFORD, R.W. (Eds.) Selecting Models from Data. Artificial Intelligence and Statistics IV. Lecture Notes in Statistics nº 89. Springer-Verlag, 1994.

[3] COOPER, G. F.; HORVITZ, E.J.; HECKERMAN, D.E. A method for temporal probabilistic reasoning. Working paper KSL 88-30. Knowledge Systems Laboratory. Medical Computer Science. Standford University. Standford, California, 1989

[4] DAGUM, P.; GALPER, A. Forecasting sleep apnea with dynamic network models. In HECKERMAN, D.; MAMDANI, A. (Eds.) Uncertainty in Artificial Intelligence. Proceedings of the Ninth Conference (1993). Morgan Kaufmann, 1993, 64-71.

[5] DAGUM, P.; GALPER, A.; HORVITZ, E. Dynamic network models for forecasting. In DUBOIS, D.; WELLMAN, M.P.; D'AMBROSIO, B.; SMETS, P. (Eds.) Uncertainty in Artificial Intelligence. Proceedings of the Eighth Conference (1992). Morgan Kaufmann, 1992, 41-48.

[6] DAGUM, P.; GALPER, A.; HORVITZ, E.; SEIVER, A. Uncertain reasoning and forecasting. Report KSL-93-47. Knowledge System Laboratory. Standford University. Standford, California, 1993.

[7] FRYDENBERG, M. The chain graph Markov property. Scandinavian Journal of Statistics, vol 7, nº 4, 1990, 333-353.

[8] KJAERULFF, U. A computational scheme for reasoning in dynamic probabilistic networks. In DUBOIS, D.; WELLMAN, M.P.; D'AMBROSIO, B.; SMETS, P. (Eds.) Uncertainty in Artificial Intelligence. Proceedings of the Eighth Conference (1992). Morgan Kaufmann, 1992, 121-129.

[9] NICHOLSON, A. E.; BRADY, J.M. Sensor validation using dynamic belief networks. In DUBOIS, D.; WELLMAN, M.P.; D'AMBROSIO, B.; SMETS, P. (Eds.) Uncertainty in Artificial Intelligence. Proceedings of the Eighth Conference (1992). Morgan Kaufmann, 1992, 207-214.

[10] PROVAN, G.M.A. Tradeoffs in constructing and evaluating temporal influence diagrams. In HECKERMAN, D.; MAMDANI, A. (Eds.) Uncertainty in Artificial Intelligence. Proceedings of the Ninth Conference (1993). Morgan Kaufmann, 1993, 40-47.

[11] PROVAN, G.M.A. Model selection for diagnosis and treatment using temporal influence diagrams. In [2], 133-142.

[12] PROVAN, G.M.A.; CLARKE, J. R. Dynamic network construction and updating techniques for the diagnosis of acute abdominal pain. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 15, nº 3, 1993, 299-307.

[13] WEST, M.; HARRISON, J. Bayesian forecasting and dynamic models. Springer Verlag, 1989.

[14] WHITTAKER, J. Graphical models in applied multivariate statistics. John Wiley and Sons, 1990.