Process modelling using Bayesian networks

Rainer Deventer Bernhard Kahles
Heinrich Niemann
Universität Erlangen-Nürnberg
Lehrstuhl für Mustererkennung
Martensstr. 3
D-91058 Erlangen
e-mail:deventer@informatik.uni-erlangen.de

1 Introduction

We used Bayesian networks for process modelling. The main advantage of such a process model is a deeper comprehension of the process, the usage of the model to search optimal input for the process to guarantee an optimal result and the prediction of output-parameters, e. g. the quality of a workpiece. We tested our approach by modelling injection moulding which can be divided in the subprocesses preparation (heat the workpiece), moving the workpiece into the machine, and spraying the synthetics around the heated metal. After a discussion with one of the project partners (Lehrstuhl für Kunststofftechnik) the structure of a Bayesian network was developed. An application in an mechanical engineering has the advantage that there is a clear distinction between input and output parameters. Thus we can assume that the input parameters are independent. A second advantage of this domain is that the given data can be considered as complete. Nevertheless we decided to allow missing data because this enables the introduction of hidden variables to simplify the model which results in less parameters to be trained. This is important because each data set had to be gained by an experiment which causes additional costs. Before starting with the training the continuous variables are replaced by discrete ones using vector quantization. During the training the missing data of the hidden variables is completed using the most probable configuration. Thus the training algorithm is similar to the 'complete' case, but resulting in an iterative procedure. After the conditional probabilities are learned the Bayes net is transformed into a junction tree as described in [Jen96]. The algorithms described in that book allows the easy calculation of the maximal probable configuration, arbitrary marginal distributions, and easy input of evidences.

2 Model evaluation

After training the models were evaluated using three different inference strategies listed in the following table:

Most probable configuration	$(x_1, \cdots, x_k) = \max\limits_{(x_1,\cdots,x_k)} P(x_1, \cdots, x_k \mid E)$
Maximum a posteriori	$x_i = \max\limits_{x_i} P(x_i \mid E)$
Maximum mutual information	$(x_1, \cdots, x_k) = \max\limits_{(x_1,\cdots,x_k)} \frac{P(x_1,\cdots,x_k \mid E)}{P(x_1,\cdots,x_k)}$
E	Given evidence, i. e. desired value of output parameters
x_i	Input variables, to gain the wanted output

For model evaluation we used a set of 385 training data, divided into 231 training data set and 154 for the model evaluation. To get an impression of the model accuracy we used the net for prediction of one missing variable using the different inference strategies mentioned above. The obtained accuracy is normally between 4% and 16%. For one variable there is an error of 328%, which is caused by the fact that this variable takes on values near by zero. These results are obtained using maximal mutual information as inference mechanism, which gives best results in most of the cases.

Beside the experiments done with discrete Variables we modeled our process also with a complete continuous network and a hybrid network with discrete input variables and continuous output variables. The tests including continuous variables are done with Bugs, using Gibbs-Sampling as learning algorithm. The results for the prediction of one missing variable are comparable to that of the discrete model.

Additionally the continuous model is used to test the ability of the model to select optimal values for the input variables to guarantee the best output. This test is done by learning a normal distribution, (i. e. mean and dispersion) using Gibbs-sampling. The suggested input of the expert of the project-partner is in all cases within the interval defined by means and variance.

3 Open problem

It was shown, that Bayesian networks can be used for selection of the best process input for a static process, i. e. Bayesian networks are a suitable mean for process control. To come to an continuous control of dynamic process, it is important to expand this approach with the possibility of feedback. That means that not only the desired output is used as input, but also the difference between desired and actual value. Thus the modelling of time is necessary and the calculation of the correcting variable has to be done in realtime.

Besides the modelling of the time dependency of a dynamic system it is from advantage to be able to use continuous variables directly, so it is planned to expand the system with an inference mechanism as described in [Lau92].

References

[Jen96] Finn V. Jensen, 'An introduction to Bayesian Networks', UCL Press Limited, University College London,

[Lau92] Steffen L. Lauritzen, 'Propagation of Probability, Means, and Variances in Mixed Graphical Association Models', Journal of the American Statistical Society December 1992

Contributed for poster presentation (Presentation by R. Deventer)