POULIN • HUGIN
PATTERNS & PREDICTIONS

Identifying Credit Risk Exposure

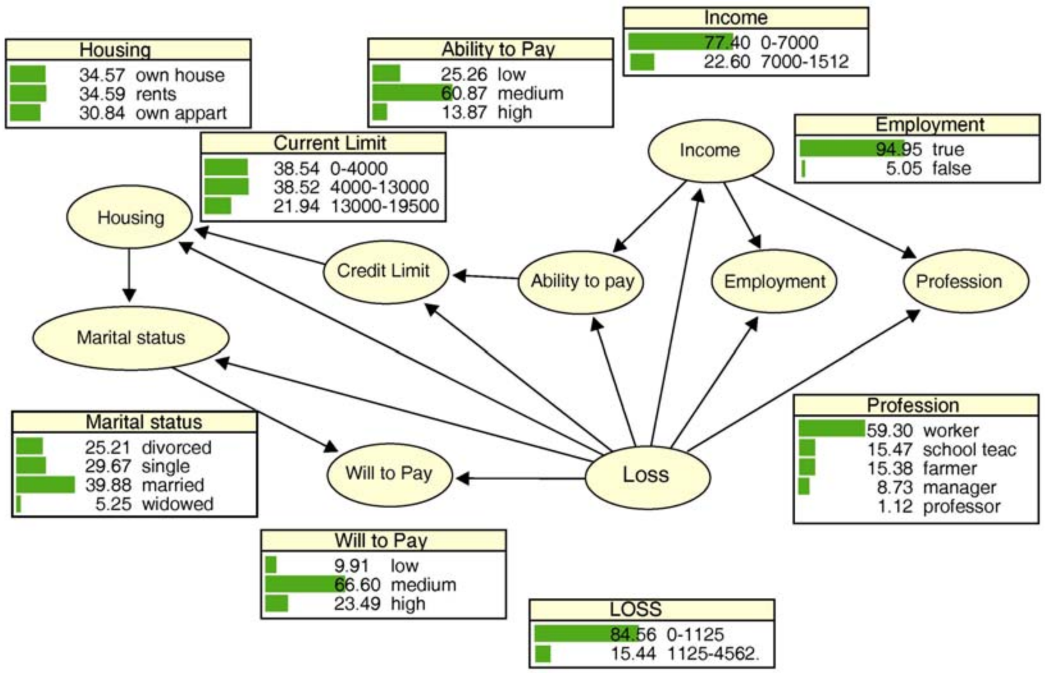
Bayesian networks can and have been used as components for reasoning under uncertainty in large and complex systems. In this case example (which is based on work we did for a large European property financing company) we illustrate how POULIN-HUGIN technology may be used for handling uncertainty in a bank's credit management system, enabling the bank to identify, assess, monitor, and better control risk.

The primary function of this model's knowledge base is to identify the clients that should be given particular attention, and to ensure that time is not spent monitoring more stable clients (automatically monitored by the system). The system may also be used to support the establishment of new client relationships, as well as to monitor the development of the client relationship on both an individual and group level.

The main component of this Bayesian network is a knowledge base designed to describe the most important attributes of a bank client with respect to credit management: a *credit-profile.* A knowledge base is a graphical model that usually describes dependence relations between entities (or factors) of the problem domain (which, in this case, is credit worthiness). The knowledge base consists of variables representing specific characteristics that determine the loan repayment behavior of the client, such as housing conditions, marital status, income, profession and conditions of employment.

As this knowledge base is based upon the same fundamental understanding of risk assessment, it will now be able to solve different tasks, such as establishing and monitoring credit. Additionally, the system will be partly self-learning and self-modifying. Finally, but importantly, the knowledge base is easily modified in the case of discovery of irregularities or external changes in the conditions fundamental to the risk evaluation.

Our technology is illustrated. The diagram below shows the structural correlation between the factors considered in the model. Each oval represents a specific characteristic of the bank's client and each link represents a direct correlation between two factors. The diagram also shows the possible values of some of the factors. For instance, the will of a client to repay her loan may be projected to be low, medium, or high whereas the loss to the bank is measured in Dollars (USD).



The object of this knowledge base (Bayesian network) is to predict monetary loss to the bank and identify patterns in customer behavior impacting the loss. So, the variables are joined in a network that describes relations between the loss and different characteristics of the customer. The loss relates directly to all predictors in the model. Some predictors are interdependent. For example, the ability to pay is directly related to income and credit limit. Income is, in addition, directly related by profession and conditions of employment, and so on. An extrapolated example, constructed only for the purpose of illustrating the underlying idea, is shown in the diagram on the previous page.

Each variable in the bank's knowledge base may take several values that are specific to the client in question (some are shown in the diagram). It is important to note that *the system does not need information about all the characteristics of the client to be useful.* However, the more known about a client, the more accurate the risk assessment will be.

Not only does the knowledge base describe possible interdependence relations between variables, it also quantifies the strengths of the relationships. The relations between variables (factors) are quantified by conditional probabilities. The data calculation box labeled "Loss" in the above diagram shows that for the average client there is a high probability (approximately 85%) of a loss less than $1,000.

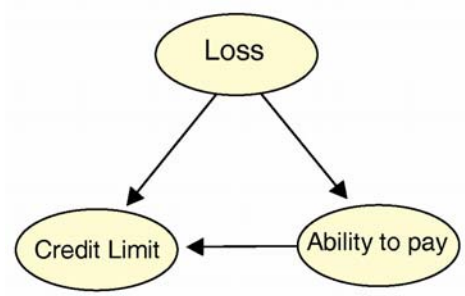
A knowledge base like this one can be constructed automatically from a database of past customer cases, by hand modeling, or by a combination of the two. The knowledge base may be used to make inferences about customers and perform different types of analysis such as value of information analysis and scenario based sensitivity analysis.

The first step of the model construction is to identify the values of the different factors under consideration. The values of three factors are shown below.

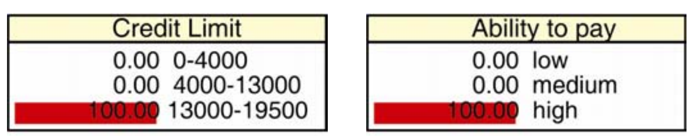
Employment	**Housing**	**Loss**	**Marital status**
False	Rents	$0	Single
True	Own house	$100	Married
	Own apartment	$4,500	Widowed/Divorced

Once the factors are identified our software identifies the structural relations between factors as shown in the Bayesian network diagram on the previous page.

Our software examines the structural correlations between ability to pay, loss, and credit limit, as identified through the factors used the model.



We may want to adjust the values based on expert knowledge. Therefore, assume that we adjust the bars to obtain the conditional probability distribution shown below where the likelihood of a high credit limit AND a high ability to pay THERFORE a high loss has been increased.



The distribution over loss for this client (in the original model) is shown in the below figure.

However, the distribution over loss for this client (in the model with a revised correlation between credit limit, ability to pay, and loss) is shown in the below figure.

Notice how the revision has increased the probability of a loss in the range of 1,125 to 4,562.

The real system should be imagined as somewhat more complex than this example. Additionally the system could be designed to be dynamic – with the possibility for the credit-profile to change over time. Based on the available information at a given time, the system will calculate the probabilities of loss of different amounts (as shown in the above figure). Based on this, a control-system could be built that would automatically give warnings if such probabilities exceed a certain limit. The system could also identify risky clients by other criteria.

Model Summary

We have illustrated how Poulin-Hugin components can be used for credit management. More than merely theoretic, a large Danish property financing company, Nykredit, is using our technology in this way as an efficient tool for risk management. The tool is **BayesCredit**, and in a benchmark evaluation, the overall performance of BayesCredit is in some cases better than that of the benchmark models developed by S&P (in comparison to the same database). For more information visit *www.Hugin.com*.

Beyond the credit space, Poulin-Hugin can be applied to almost all problem domains as it offers efficient, coherent, and sound handling of inherent uncertainty. Application of our decision-theoretic technology is used by leading companies to solve many problems of reasoning and decision making.

# # #