Safety and Reliability – Bedford & van Gelder (eds)
© 2003 Swets & Zeitlinger, Lisse, ISBN 90 5809 551 7

Aviation causal model using Bayesian Belief Nets to quantify management influence

A.L.C. Roelen & R. Wever
National Aerospace Laboratory NLR

R.M. Cooke & R. Lopuhaä
Delft University of Technology, Faculty of Information Technology and Systems

A.R. Hale & L.H.J. Goossens
Delft University of Technology, Safety Science Group

ABSTRACT: The authors have recently completed a causal model for aviation safety under contract with CAA the Netherlands and the US Federal Aviation Administration. The goal was to develop a comprehensive model of aviation safety and to judge the potential impact of management decisions. The overall modeling tool chosen for this task was Bayesian Belief Nets (BBNs). The reason for choosing BBNs above more conventional tools (fault and event trees) were: 1) BBNs better enable the problem owner to recognize his problem; 2) Software support is such that the top level representation of the problem is at once the user interface for performing calculations; 3) The influence of management decisions is readily factored in using structured expert judgement. Perhaps the most satisfying aspect of the BBN approach is that it provides a comprehensive model which is maximally data-driven, yet which includes assessments of potential impacts of contemplated decisions.

1 OBJECTIVE

The main objective of the project is to develop a demonstration model, showing how causal modeling can be applied to aviation safety.

2 MODELING APPROACH

The overall modeling tool chosen for this project was Bayesian Belief Nets (BBNs). More conventional models have attempted the use of Fault Trees and Event Trees (DNV 2002). While Fault Trees and Event Trees have demonstrated to be useful techniques for modeling complex technical structures, they fail to adequately represent the complexities and uncertainties with respect to causal relations in the air transport system, especially as regards human factors and organizational behavior. Human behavior is difficult, if not impossible, to capture in AND and OR gates when modeling takes place at an aggregated system level. A BBN is able to represent the fact that elements may influence and react with each other. In addition, BBNs better enable the problem owner to recognize its problem.

Among the difficulties associated with Bayesian Belief Nets are:

– The vague notion of 'influence' may be understood in different ways by different team members.
– A combinatorial explosion threatens the assessment of probabilities in realistic problems.
– Assessments must be well documented to avoid them becoming slipshod and non-traceable.

The first issue was addressed by including all relevant problem-sharers in the modeling team. We also found satisfactory ways of dealing with the second and third issue, which are described in this paper.

3 A METHOD TO CALCULATE CONDITIONAL PROBABILITIES

Among the difficulties associated with the use of Bayesian Belief Nets is the fact that a combinatorial explosion threatens the assessment of probabilities in realistic problems. To solve this problem, we must have a simplifying assumption which reduces the assessment of probabilities of the form:

PROB{undesirable event | cause 1 and cause 2, … and cause n}

to a function of the probabilities

PROB{undesirable event | cause i}, i = 1, … n.

This enables the quantification of the Bayesian Belief Net with probabilities extracted from data. This simplifying assumption is derived below.

To quantify the model we need to know the probabilities of events X and Y (P(X) and P(Y)), and we must know the conditional probabilities P(A | X), P(A | Y) and P(A | X, Y). In many cases P(A | X) and P(A | Y) can be obtained from observations. P(A | X, Y) however is more difficult to determine from observations. Therefore we would like to model P(A | X, Y) as a function of P(A | X) and P(A | Y).

Let X and Y be events, not necessarily independent, which are capable of causing undesirable event A with high probability. We assume:

$$P(A^c|X,Y) = P(A^c|X)\ P(A^c|Y). \quad (1)$$

In which A^c is the complement of A, that is the set of all elements (outcomes) not in A.

Manipulations with Bayes' theorem yield:

$$P(X,Y|A^c)=P(X|A^c)\ P(Y|A^c)P(A^c)P(X,Y)/P(X)P(Y).$$

If X and Y are independent and $P(A^c) \sim 1$, then

$$P(X,Y|A^c) \sim P(X|A^c)\ P(Y|A^c)$$

i.e. X and Y are nearly independent conditional on A^c.

Re-writing (1), we find:

$$P(A|X,Y) = P(A|X) + P(A|Y) - P(A|X)P(A|Y). \quad (2)$$

Note that for positive numbers, x and y, less than one:

$$x + y - xy = x(1 - y) + y < 1.$$

More generally, for $X_1 \ldots X_k$ causing A with high probability, we find

$$P(A|X_1 \ldots X_k) = \Sigma_i\ P(A|X_i) - \Sigma_{i<j}\ P(A|X_i)P(A|X_j) + \Sigma_{i<j<h}\ (A|X_i)P(A|X_j)P(A|X_h) - \ldots\text{etc.} \quad (3)$$

The advantage of this approach is that it does not require the estimation of joint probabilities and that it cannot yield values outside the [0, 1] interval, even if conditional independence does not hold.

4 EXAMPLE: MISSED APPROACH

4.1 *Description of missed approach*

When during the approach to the landing runway a certain situation exists or arises, which would make the continuation of the approach and consecutive landing 'unsafe', the flight crew should initiate a missed approach. The purpose of the missed approach procedure is to reject flying into unsafe conditions or under unsafe circumstances and to enable the crew to carry out a new approach and landing under safe(r) circumstances.

A missed approach (Figure 1) is ultimately initiated by the flightcrew, based on their mental representation of the current situation. The main factors that are important in the missed approach decision making are considered to be the following:

- Weather
- Air traffic situation
- Aircraft status
- Aircraft trajectory
- Flight crew.0

This leads to a 'model' of missed approach, expressed in the form of a Bayesian Belief Net, as shown in Figure 2.

4.2 *Quantification for missed approach*

Conditions for missed approach obtain if any of the following hold:

Weather conditions (W) are LIMITING (l) or
Air traffic Situation (AT) is NOT OK (nok) or
Trajectory (T) is OUTSIDE CORRIDOR (oc) or
Aircraft status (AS) is NOT OK (nok)

Execute Missed Approach (XMA) = yes is not an undesirable event, but it is unlikely, unconditionally. Each of the above conditions for missed approach is also unlikely. Following equation (3) we can restrict ourselves to univariate conditionalization. In other words, we compute

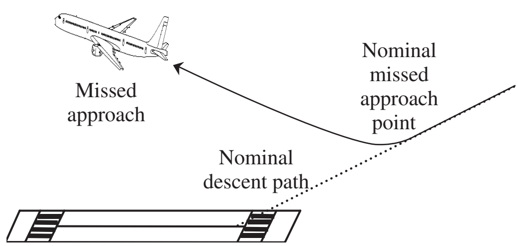

Figure 1. Missed approach.

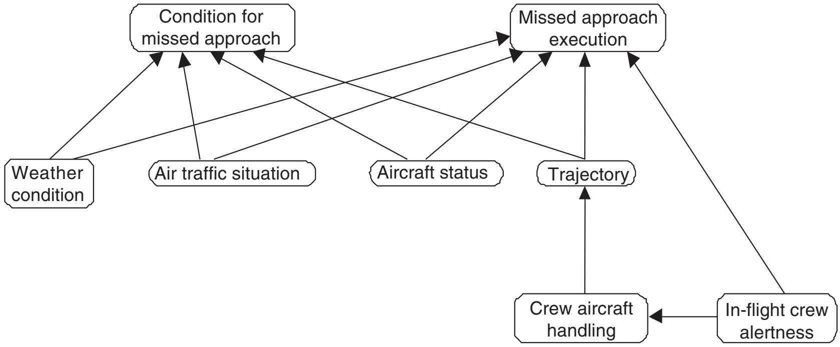

Figure 2. Missed approach model.

$P(XMA = n \mid W = l, T = oc) =$
$P(XMA = n \mid W = l) + P(XMA = y \mid T = nok) -$
$P(XMA = y \mid W = l)\, P(XMA = y \mid T = nok).$

With respect to the event Crew Alert (CA) the situation is different. Crew not alert (CA = n) is not in itself a high probability cause of XMA = n. Indeed, XMA = n is not an unlikely event and CA = n by itself will not raise its probability significantly. Hence, $P(XMA = n \mid CA = n)$ will be of the same order as $P(XMA = n) \sim 1$. However, if one of the conditions for missed approach is present, then XMA = n should be unlikely, but CA = n might make it significantly more likely. Thus, it may be anticipated that

$P(XMA = n \mid W = l, CA = n) \neq$
$P(XMA = n \mid W = l) + P(XMA = n \mid CA = n) -$
$P(XMA = n \mid W = l)\, P(XMA = n \mid CA = n) \sim 1.$

In this case, there is reason to expect that the contributions to XMA = n from causes W = 1 and CA = n do not simply add as foreseen in (1).

Suppose we know all the binary causal conditions:

$P(XMA = n \mid W = l, CA = n),$
$P(XMA = n \mid AT = nok, CA = n)$
$P(XMA = n \mid T = oc, CA = n)$
$P(XMA = n \mid AS = nok, CA = n).$

We may apply an analogue of equation (3), whereby we consider the above binary conditions as one event, hence:

$P(XMA = n \mid W = l, AT = nok, CA = n) =$
$P(XMA = n \mid W = l, CA = n) +$
$P(XMA = n \mid AT = nok, CA = n) -$
$P(XMA = n \mid W = l, CA = n)\, P(XMA = n \mid AT = nok, CA=n).$

And similarly for higher order causal conditions. For example, the event

$\{W = l, AT = nok, T = oc, AS = nok, CA = n\}$

is written as $V \cap X \cap Y \cap Z$ with

$V = \{W = l, CA = n\}$
$X = \{AT = nok, CA = n\}$
$Y = \{T = oc, CA = n\}$
$Z = \{AS = nok, CA = n\}$

and we apply (3) to the events V, X, Y, Z.

5 HIDDEN NODES

5.1 *Problem description*

There may be nodes influencing events, which are not represented in the graph. If these nodes influence multiple events, they may induce dependencies between nodes in the graph which appear to be independent, and this can cause problems.

5.2 *Example*

The wind condition is limiting if cross and tailwind components exceed limits prescribed in the Aircraft Operating Manual. These limits depend on aircraft type. Similarly, the aircraft status depends on the fuel state. Whether the fuel state is 'not ok' depends on the flight route and therefore on aircraft type.

Suppose for the sake of example that there are two aircraft types, heavy and light, and that 2/3 of all flights concern heavy aircraft. Suppose that limiting wind conditions apply only to light aircraft and that a light aircraft approach has 3/100 chance of experiencing these limiting conditions. Suppose that fuel 'not ok' occurs only with heavy aircraft, and has a 3/100 probability of occurrence per approach. The probability of wind limiting is $(2/3) \times (3/100) = 0.02$; the probability of fuel state 'not ok' is $(1/3) \times (3/100) = 0.01$. These events are not connected in the graph, but the probability of wind limiting AND fuel not ok is not

$0.02 \times 0.01 = 0.0002$, but is zero, since no aircraft is both heavy and light.

If we suspect that hidden nodes like 'aircraft type' could induce significant correlations then we must either add them to the model, or specialize the model to aircraft types which are sufficiently similar that the correlations do not arise.

6 MANAGERIAL INFLUENCES

One of the requirements of the model was the ability to express the effect of managerial decisions on air transport safety. We have addressed this issue by adding a management model to the technical model described above. The technical model consists of 'chance' nodes which can have multiple states, nominally OK and NOT OK. The management model consists of 'decision' nodes which have two possible values, DO or DON'T DO. Management decisions are defined as concrete measures which have not yet been taken but which could be taken. The basic technical model is quantified conditional to an airport resembling Schiphol. In this situation the managerial decisions are not yet taken, but are under consideration. Hence the technical model is quantified conditional on DON'T DO. The impact of a decision is quantified by expert elicitation, saying how the DO value would change the probabilities of the nodes on which the decisions impact. In this first exercise our decision maker is the airport.

Generic delivery and learning systems that are used to model safety management were specified for each of the main actors (airline, airport, ATM organization and regulator) for the missed approach phase. This was combined with information from interviews that were conducted with KLM and ATC the Netherlands into a single list of 26 managerial influences on missed approach and flight crew alertness for use in the quantification process.

The impact of the decision is quantified by expert elicitation. It was therefore necessary to formulate the decisions so that they are sufficiently clear and concrete. This includes removing ambiguity between the purpose of the decision and actual concrete measures to be taken. Put otherwise, the decisions must be expressed as a statement of what actually will be done instead of a statement of purpose.

The protocol that we used for expert judgement elicitation has been developed by Delft University of Technology (Cooke & Goossens 2000). The goal of applying structured expert judgement techniques is to enhance rational consensus. A fundamental assumption of the approach we used is that future performance of experts can be judged on the basis of past performance, reflected in calibration variables. The performance of the experts on the calibration variables is taken as indicative of the performance on the variables of interest.

The goal of the preliminary study was to demonstrate the feasibility of using structured expert judgement to quantify the Bayesian Belief Net for causal models of aviation safety. The time constraints were very severe. The number of experts interviewed (4) is insufficient for a full expert judgement study, and the time spent familiarizing the experts with the study objectives and with probabilistic assessments was limited. This study involved 53 continuous items for which the experts gave their 5%, 50% and 95% quantiles. 18 of these items were calibration variables for which the true values were retrieved from the NLR Air Safety Database (Van Es & Van der Nat 1998). Experts were able to quantify their uncertainty with regard to the relevant variables in the causal model. Moreover, it was possible to define suitable calibration variables and to measure performance of the experts and of the various decision makers. In this sense the expert judgement study achieved its objectives.

At the same time, it must be acknowledged that the performance of the experts varied significantly. This together with the small number of experts (4), results in poor performance of the performance based combinations of the experts' assessments. The equal weight combination resulted in acceptable performance, although the information score is rather low. Since the Bayesian Belief Net quantification will use only the median values, the low information will not directly affect the Belief Net quantification. Thus, in spite of the drawbacks attaching to this preliminary study, it has resulted in a useable and defensible quantification of the Bayesian Belief Net.

To give an example of management influence type analysis, the model for missed approach is again reproduced in Figure 3. It consists of two main parts: 'Condition for Missed Approach' which represents the fact that the conditions dictate that a missed approach should be executed, and 'Missed Approach Execution', which represents the fact that a missed approach is indeed executed by the flight crew.

Failure to execute while the conditions require a missed approach can result in a collision with terrain, represented by the upper box in Figure 3. The model in Figure 3 also includes two decision nodes that represent managerial influences: 'CRM training' and 'Missed Approach training'. CRM training represents the frequency of Crew Resource Management Training for flight crew members. The default situation in this model is 'CRM training at the start of the flying career'. One decision is to increase the frequency to once every two year ('increase'), the other decision is to abolish CRM training completely ('never'). The decision node Missed Approach training has two options. In the default situation specific missed approach training is given twice a year, the alternative is to stop this training ('never').

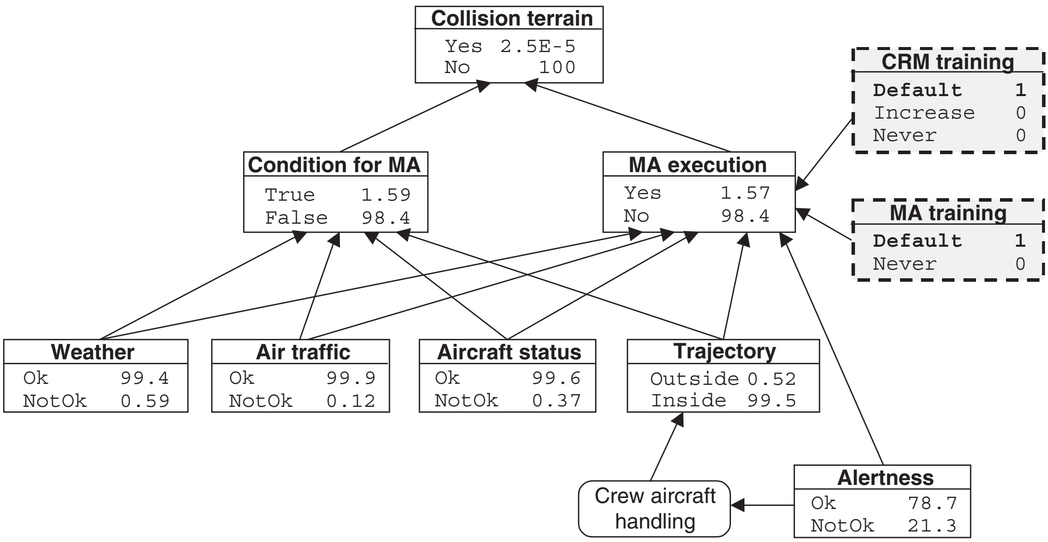

Figure 3. Collision with terrain model, default situation.

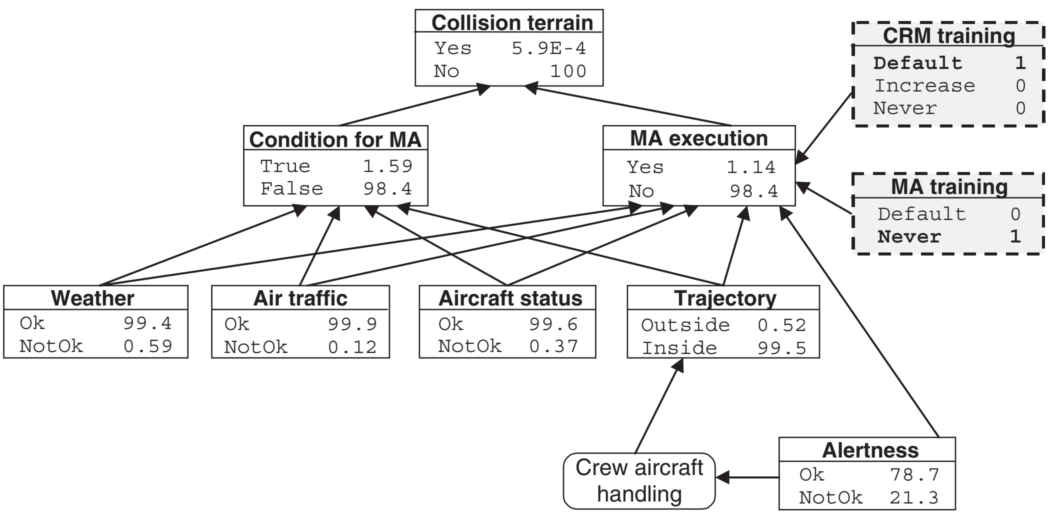

Figure 4. Collision with terrain model, no MA training.

Figure 3 includes all possible states of the chance nodes and their probabilities per flight (expressed as percentage values). These values were obtained from airport and airline data. The decision nodes were quantified by applying the above mentioned expert judgement elicitation process. In the default situation the missed approach execution probability is 1.57%. Conditions for missed approach have a probability of 1.59%. The small discrepancy between conditions and execution result in a collision with terrain probability of 2.5×10^{-5}% per flight.

Figure 4 shows how the situation will change when a decision is made to stop missed approach training for the flight crew. The missed approach execution probability drops to 1.14% and as a result the probability of collision with terrain rises significantly to 5.9×10^{-4}% per flight.

Figure 5 shows the results of the decision to increase the CRM training frequency to once every two years. The missed approach execution probability increases to 1.58% and the probability of a collision with terrain drops to 2.3×10^{-5}% per flight.

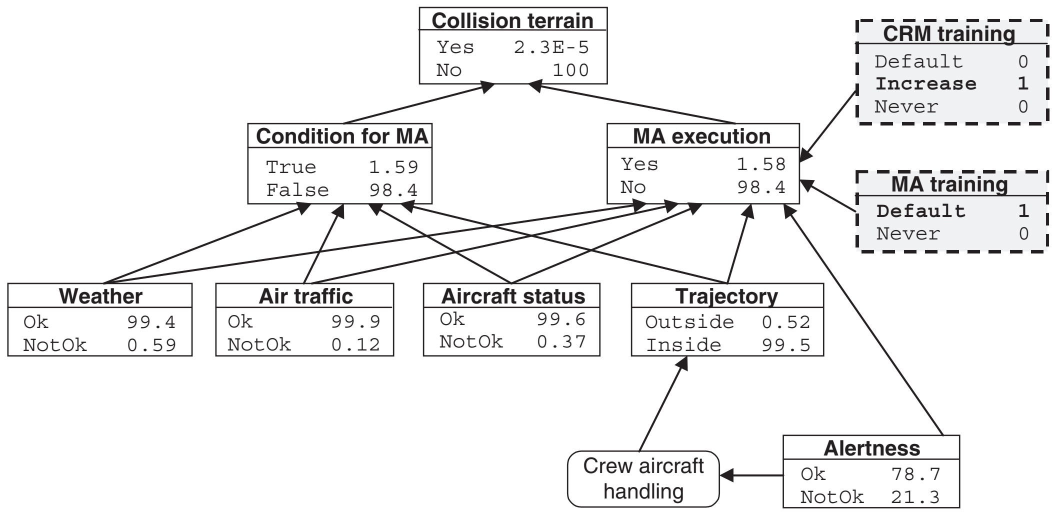

Figure 5. Collision with terrain model, increase CRM training.

In a similar way, all 26 managerial influences have been described as decision nodes, have been quantified and linked to the 56 chance nodes that constitute the core of the causal model (Roelen et al. 2002)

7 CONCLUSIONS

A Bayesian Belief Net is a proper way to express a causal model. It provides a comprehensive model which is maximally data driven, yet which includes expert assessments of potential impact of contemplated decisions.

Simplifying assumptions have successfully been used for quantification of the model with probabilities extracted from data. In special cases where this assumption was deemed unacceptable, probabilities under multiple causes were assessed with expert judgement. Finally, thanks in part to the role of the simplifying assumptions, a structured and traceable expert judgement process was possible. This enabled the project team to quantify probabilistically the effects of management influences related to airport equipment and procedures, training of ground and flight crew, changes in air traffic control and changes in airlines human resources programs.

REFERENCES

Cooke, R.M. & Goossens, L.H.J. 2000. *Procedures guide for structured expert judgement*, EURATOM document EUR 18820EN, European Commission, Directorate General for Research.

Det Norske Veritas. 2002. *Causal modelling of air safety, final report*, CM-DNV-016, DNV London.

Van Es, G.W.H. & Van der Nat, G.W.F.M. 1998. *The use and needs for air safety data in the Netherlands*, NLR-CR-98167, NLR Amsterdam.

Roelen, A.L.C., Wever, R., Hale, A.R., Goossens, L.H.J., Cooke, R.M., Lopuhaä, R., Simons, M. & Valk, P.J.L. 2002. *Causal modelling of air safety, demonstration model*, NLR-CR-2002-662, NLR Amsterdam.