

Modelling an Air Traffic Control Environment Using Bayesian Belief Networks

Dr. Martin Neil, Professor Bob Malcolm & Roger Shaw
RADAR Group, Department of Computer Science, Queen Mary, University of London

Keywords: Bayesian Networks, Air Traffic Control, ATC

Abstract

This paper describes the process of constructing a Bayesian Belief Network for an Air Traffic Control environment. The network provides a high level model of ATC operations spanning a number of defensive barriers from airspace design, through tactical control, the operation of aircraft safety net features to a potential accident. Some socio-technical factors that impact the effectiveness of some of these barriers are also mentioned. The paper concludes with an example of how the net can be used to investigate changes in overall risk exposure based on a number of exploratory scenarios.

Background

The EPSRC funded SCORE project (Sensing Changes in Operational Risk Exposure) was set up in 2001 to investigate how the risk posed by poor organisational culture (ref. 1-2) might be measured and monitored. The project is a collaborative venture between the RADAR Group in the Department of Computer Science, Queen Mary, University of London and the Safety Research Group within the Department of Psychology, Liverpool University. Two studies were included in the project. One addressed Operational Risk in financial environments and the other focused on assessing the risk of collision in an Air Traffic Control environment. This paper describes the ATC study and, in particular, the Bayesian Network that was produced to predict accident risk and show how different factors influence the change in risk exposure.

Bayesian Networks

Bayesian Networks (also known as Bayesian Belief Networks, Causal Probabilistic Networks, Causal Nets, Graphical Probability Networks, Probabilistic Cause-Effect Models, and Probabilistic Influence Diagrams) provide decision-support for a wide range of problems involving uncertainty and probabilistic reasoning. The underlying theory of BNs is Bayesian probability theory and the idea of evidence propagation through a network structure.

A BN is a directed graph, together with an associated set of probability tables called Node Probability Tables or NPTs. The graph consists of nodes and arcs as shown in Figure 1. The nodes represent variables, which can be discrete or continuous. For example, the node *Faults in Test/Review* is discrete having values 0,1,2, whereas the node *System Safety* might be continuous (such as the probability of failure on demand). The arcs represent causal/influential relationships between variables. For example, the *Correctness of Solution* and *Accuracy of Testing* influence the number of *Faults in Test/Review*; hence this relationship is modelled by drawing appropriate arcs as shown. The key feature of BNs is that they enable uncertainty to be modelled and reasoned about. BNs also offer considerable analytical power to the modeller. They can show which variables contribute most to establishing a hypothesis and can show how sensitive variable valuation is in determining a hypothesis. In more complicated networks variables may contribute in different proportions to establishing a hypothesis. For example, knowing how sensitive a

variable is in determining a hypothesis will provide guidance on how much effort should be devoted to accurately determining the value of the variable. If small changes in the value of a variable dramatically alter the value of the hypothesis variable then it should be determined as accurately as possible; if the hypothesis is not that sensitive to the value of the variable then less effort can be devoted to its determination.

As part of the process of developing a BN, Node Probability Tables (NPTs) need to be constructed. Variables with no parents represent our prior assumptions regarding their likelihood while variables with parents will have probabilities conditioned upon those parent nodes. Determining the NPTs is part of the knowledge elicitation process, which traditionally has proved to be a time consuming and often difficult task. Although not discussed here the SCORE project has developed an approach to facilitate the collection of this information from experts using a small set of graphical tools to support the process.

The benefits of BNs may be summarised as follows:

- Provides a sound method for reasoning under uncertainty.
- Permit the combining of diverse data, including subjective beliefs and empirical data.
- Predictions can still be secured even when evidence is incomplete.
- Permits powerful “what-if” analysis to test the sensitivity of conclusions.
- Incorporate a visual reasoning tool, which aids documentation and explanation.

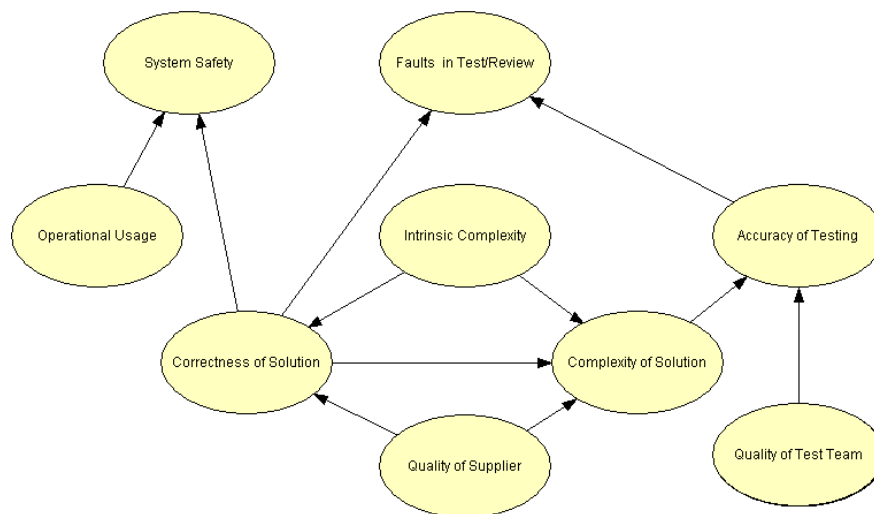


Figure 1: Example Bayesian Network

The use of BNs, as outlined above, is based on a decade old research programme, which started with the DTI/EPSRC project DATUM (ref. 3) in 1992. This project investigated how to integrate different sources of evidence together into a coherent and quantitative argument. The motivation was the production of what are called Safety Cases. Various approaches were investigated and BNs were chosen as the most suitable way of achieving this objective. With this decision made, work progressed on the EC funded SERENE project (1996-1999) (ref. 4) to develop a method for representing safety arguments using BNs. The results from this project included a fully developed method, a supporting tool for the method and a number of case studies. For those interested in reading further on the subject of Bayesian Networks reference should be made to (ref. 5-6).

Developing a Bayesian Network

Developing a Bayesian Network requires careful planning and execution. Firstly it is necessary to decide on the intent of the network. This involves identifying the hypothesis variable. In the case of the ATC example the hypothesis variable was the likelihood or probability of a collision between two aircraft under air traffic control. The second stage is to develop a causal or influence model that allows the hypothesis variable to be determined. Developing such a network requires a number of experts to be brought together to identify the factors (modelled as evidence and intermediate variables) that influence the hypothesis variable. For a complex problem where the answer is not well understood, yet alone agreed upon, debate will ensue and a consensus needs to be arrived at regarding the network structure. In these circumstances it is a good idea to formulate a structure and then discuss it with at least two groups of experts. This may then be followed up by a workshop involving all the experts. In this way a consensus can generally be secured. Once the structure has been agreed conditional probabilities need to be elicited from the experts. This is a long but essential part of the process as the probability space is sampled and conditional probabilities elicited from the experts. A range of experts should be used and results compared. Quite often, when the conditional probabilities are being discussed, changes are suggested to the network structure and these have to be carefully evaluated and, if agreed, introduced. When the network is fully populated it should be explored by introducing evidence, propagating it through the network, and observing the changes arising in the hypothesis variable. If these changes seem reasonable then historical data from the domain should be used to test the predictions of the model against what has already been observed in practice. By this means the network can be calibrated to improve its predictive power. Once this task has been successfully completed the model can be cautiously introduced into use.

The ATC Study

The ATC Bayesian Network was developed from a series of workshops with representatives from the Air Traffic Control authority. The network was developed to represent operations spanning the complete environment from airspace design, through the operation of tactical controllers to the possibility of a mid air collision. The model embodies dependencies between root causes (people, technology, processes) and safety performance measured in terms of actually recorded safety significant events. Further, the model quantifies the degree of importance or influence that a causal factor has on overall performance. For example, the model currently incorporates the influence that STCA and TCAS have on collision avoidance. Models such as these provide transparent mechanisms for “what if” analysis by allowing changes in the ATC environment to be incorporated into the network in order to determine their impact on the overall risk of collision. Finally, investigations of the network can be facilitated by the use of decision support tools, which permit risk monitoring to be conducted over time and changes in overall risk exposure to be discerned and analysed.

Developing a Bayesian Network is a creative process informed by domain expertise. Therefore, the first meetings held with ATC staff concentrated on gaining a good understanding of the ATC environment. This background is briefly set out below.

The purpose of an air traffic control (ATC) system is to manage the trajectories of aircraft in such a way as to minimise the risk of collision. This objective is achieved through a range of measures that include:

- The design of the airspace (the virtual corridors within which the aircraft fly).
- Procedures for controlling access to, and use of, the airspace.

- The use of highly skilled and trained air traffic controllers.
- The use of collision warning systems such as the Short Term Conflict Avoidance system (STCA) and the Traffic Alert and Collision Avoidance System (TCAS).
- The skills of the pilots themselves.

Despite the care and attention devoted to the design of such a complex socio-technical system there is always a risk that a collision will occur. Recent events such as the mid air collision between a TU154 of Bashkirian Airlines and a Boeing 757 operated by DHL over Ueberlingen on the 1st July 2002 illustrates the reality of such risks.

In order to better understand the nature of the collision risk involved in an ATC environment it is necessary to have a model showing the means taken to avoid collisions and via this model identify possible factors that might arise and lead to an accident. The model adopted for this exercise is a barrier model, sometimes referred to as a “defence-in-depth” model (ref. 7). This model depicts a number of defences, which collectively aim to prevent accidents arising.

In order for a collision to occur various “events” must take place. The aircraft must have been on a collision course, the automated collision warning systems must have failed in some way, air traffic control staff must have failed to note the collision trajectory and rectify matters and the crew of the aircraft must have failed to detect and avoid the collision. Thus, when accidents do arise they can be seen as resulting from the breach of a number of defences.

Figure 2 shows a schematic of the barrier model taken as the starting point on this project. The first barrier is the design of the airspace the intent of which is to channel aircraft in such a way that they are kept apart, and to minimise the risks associated with crossing points where corridors meet (density and complexity). As well as serving safety objectives air space design has to serve economic and efficiency requirements as well.

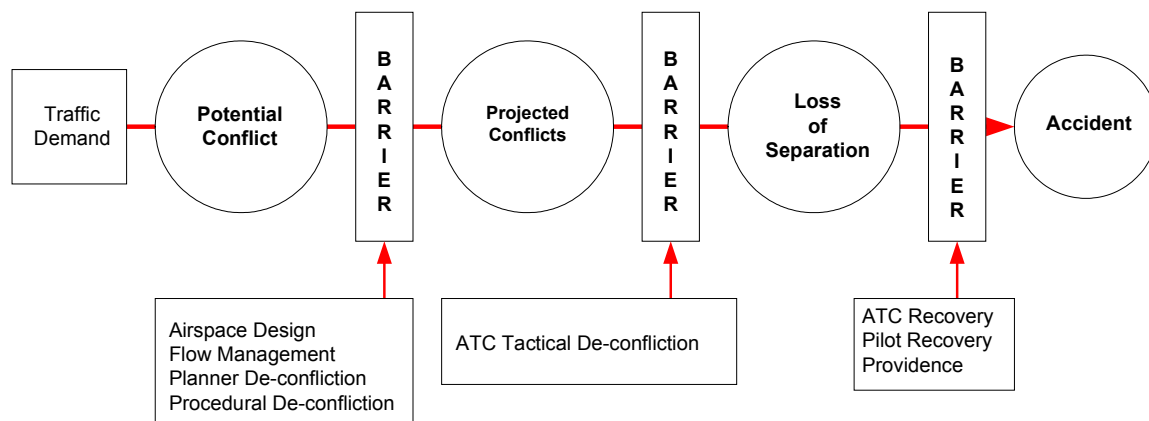


Figure 2: Schematic Barrier Model

Given a designed airspace, and given the demands of airlines and airports upon that airspace, flight planners have the task of assigning flight-paths to aircraft. They may take more or less cognisance of the risks arising from the combination of routes they assign.

Flow management is a function of air traffic control which, given warning of upcoming flight-plans, aims to maintain the volume of traffic through the various ATC sectors within manageable levels, so called “target sector flows”.

In some sectors there is a distinction between “planning controllers” and “tactical controllers”. While this is not true of all sectors, there is a useful distinction to be made between the two roles, even if performed by the same person. The planning controller is concerned with routing through the sector, while the tactical controller is concerned with the moment-by-moment management of aircraft, given that routing. Planners, given the anticipated flow of aircraft, attempt “planner de-confliction” by routing them through a sector in such a way as to minimise the difficulties faced by the tactical air traffic controllers - who have the job of maintaining separation of aircraft, normally through “procedural de-confliction” - i.e. following procedures designed to maintain separation.

After the tactical air traffic controllers, the next barrier is the Short Term Conflict Avoidance system (STCA). Where installed STCA automatically detects impending loss of separation and raises a warning to the air traffic controllers to attempt ‘tactical de-confliction’. STCA alerts may arise expectedly or unexpectedly. In the former case a controller is expecting an alert to arise and has already planned manoeuvres that will remove the conflict. In the latter case the controller becomes aware belatedly of the conflict and manoeuvres then need to be planned to de-conflict the situation.

If, despite the barriers thus far, a potential conflict situation arises between two aircraft, then if they both have Secondary Surveillance Radar (SSR) transponders and at least one has a system called TCAS (Traffic Alert and Collision Avoidance System) installed, then TCAS should automatically provide recommendations to the pilot(s) for avoiding action(s) so as to achieve “aircraft de-confliction”. TCAS is designed to advise one aircraft to gain altitude and the other to lose altitude, thus avoiding the collision.

If an incident is still impending despite these barriers, then the situation is left for “last minute see and avoid action” by the pilot or providence (the trajectories (situation geometry) are such that an accident is avoided). If this last barrier fails then an accident may ensue.

In order to determine the likelihood of an incident historical data was used to indicate how often breaches took place. Incident data is currently collected when Safety Significant Events (SSEs) arise and this data is linked to loss of separation. Loss of separation is divided into two bands:

- **Band 1:** Separation \leq 66% of prescribed value or, if no prescribed value, then \leq 2 nautical miles and 600 feet.
- **Band 2:** Separation $>$ 66% of prescribed value or, if no prescribed value, $>$ 2 nautical miles and 600 feet.

SSEs were then defined relative to these bands.

- **SSEP:** SSE Potentials occurred for any Band 2 incident.
- **SSE4:** An SSE4 is registered if a Band 1 event occurred and is effectively resolved by a controller who was providing the service when the event was initiated and no system or procedure failure affected the resolution. Two sub categories are defined, the first if the controller was aware in advance that the event would occur and the second if he or she became belatedly aware.
- **SSE3:** In this case the Band 1 event is detected and resolved by ATC but either not by the controller providing the service, or the event was detected by another controller, STCA or

the pilot, or it was not resolved in a timely or effective manner or a system or procedure failure affected the resolution.

- **SSE2:** In this case, once again Band 1, the pilot resolved the event, or it was not resolved by the aircraft safety net or it was resolved by the aircraft safety net.
- **SSE1:** Finally, the Band 1 event was not resolved by timely pilot action or there was a high risk that any action taken was ineffective. Matters are resolved by providence.

The first project meeting, following lengthy discussions about the design of airspace and the various roles within an ATC environment, proposed an initial Bayesian Network model based upon the following reasoning.

1. Conflicts are made more or less likely as a result of airspace design, thus airspace complexity had to be represented in the model.
2. Flow management, the first ATC task, may be impacted by airspace complexity and, in turn, could make the task of the Planner ATC easier or harder. Thus flow management entered the model.
3. The Planner ATC's activities were impacted by the inherited flow characteristics and in turn this impacts the work of the Tactical ATC.
4. The Tactical ATC would manage the moment-by-moment sector movements and would de-conflict potential conflicts using planned procedures.
5. However, either as a result of planned activities or otherwise, the STCA system signals a separation breach in which case de-confliction (planned or otherwise) was required to avoid an incident.
6. If de-confliction failed then, if aircraft carried TCAS a proximity warning would be given to the pilot who could then take avoiding action.
7. The final barrier, if all else fails, is that provided by geometry and providence. If this final barrier fails then an accident might occur.

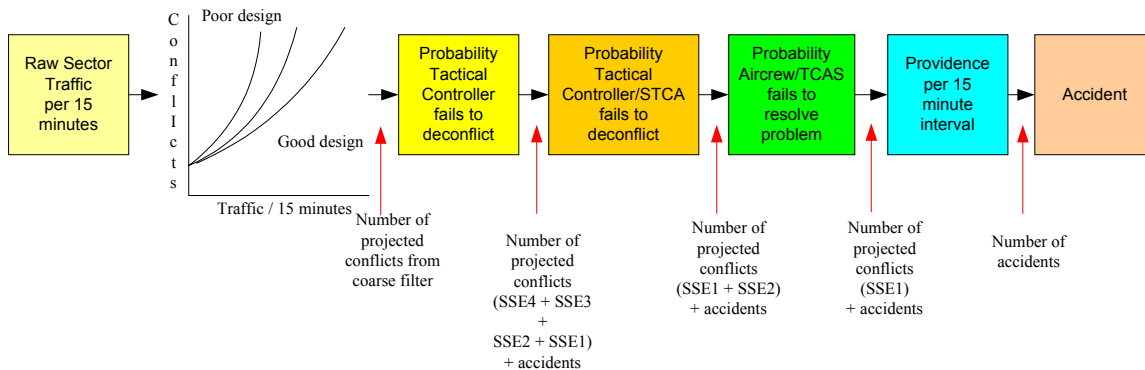


Figure 3: Initial Modelling Scenario

This model turned out to be too complicated. The flow management (bullet point 2) was removed because of a lack of data about how this process operates. After further consideration the role of the Tactical Planner (bullet point 5) was re-modelled to include their contribution in managing TCAS thus leading to two nodes involving the Tactical Planner, one modelling the performance of the Tactical ATC in normal operation and one when dealing with conflict alert situations. This revised perspective is shown in Figure 3, a diagram emerging from one of the workshop. The new model had the advantage that the SSE scheme could be used to provide useful data for entry into the network.

Based on this revised perspective a new Bayesian Network was produced into which was added structural details addressing socio-technical factors and how they impact the performance of the various roles in the network (Planner ATC staff, Tactical ATC staff) through their training, competencies and technical provisions. Additionally, pilot skills were also modelled in the network along with the manner in which they, and the Tactical ATC staff, interact and in the process impact one another's performance.

The complete model is too large to be shown here. It is made up of some 41 nodes and in excess of 2000 conditional probability values. To provide some insight into the network a small part will be examined.

Figure 4 shows a simplified view of the four barriers modelled. Starting from the left, Planner ATC Performance models the manner that traffic presentation (complexity/crossings/volume etc.) is altered by the Planner Controller, given their capabilities. Presented with a complex traffic pattern, a planner with excellent capabilities should make a big improvement to the situation (remove many potential conflicts). Given low capabilities the controller may worsen the situation. The capabilities node models such factors as individual competence, the prevailing safety culture and the quality of technical support provided. The upper node titled "events prior to tactical intervention" models the number of future potential conflicts that have been left in, or created, by the Planner Controller, based on the traffic scenario presented to the Tactical Controller.

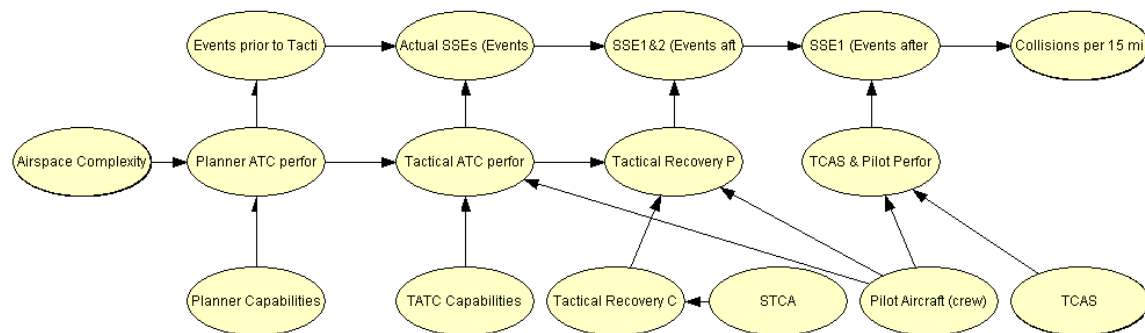


Figure 4: The Four Barrier Model

Tactical ATC Performance models the manner in which traffic presentation, given to the Tactical Controller by the Planner Controller, is altered. Once again the performance of the Tactical Controller is influenced by their capabilities, modelled as before. If the Planner ATC performs well a Tactical Controller with excellent capabilities should make a big improvement to the situation (and remove most of the potential conflicts). Again, in the presence of low capabilities the controller may worsen the situation. If the Planner ATC worsens the situation this means that more potential conflicts will be presented to the Tactical Controller. The node "actual SSEs" models the number of SSE1 – 4s and accidents that remain in the airspace following intervention by the tactical controller. These are events that have not yet reached the stage of tactical recovery.

Tactical Recovery Performance models how effectively the Tactical Controller is in recovery mode, given their performance and capabilities and those of the pilot. In the presence of an incompetent pilot, recovery can be very poor even if the Tactical Controller has excellent capabilities, because any recovery is ultimately performed by the pilot (albeit usually under a

controllers instruction). Similarly, a highly skilled pilot should be listening to the radio (situational awareness) so should respond quickly to any actions (especially “avoiding action” issued), without delay. The node “SSE1&2” represents the actual number of SSE1s, SSE2s and accidents remaining in the system. The Tactical Controller’s endeavours at recovery have failed or they were non-existent.

On occasions TCAS may give a nuisance alert and therefore mislead a pilot. Pilots are trained to follow TCAS advice hence when TCAS gives a nuisance/false alert even a highly skilled pilot may follow it although being highly skilled, they are more likely to have good awareness of surrounding aircraft from radio etc. A highly skilled pilot may also add a turn to whatever advice is given by TCAS (which only separates vertically). A poor pilot may be slow to react or ignore the advice given by TCAS. The node above yields the actual number of SSE1s and accidents remaining in the system. Recovery has been ineffective. Finally, to the far right of the network, the node “collisions per 15 minutes” yields that actual number of collisions/accidents in a fifteen minute interval for a single sector.

The description given above is qualitative. What are of interest are the predictions that the overall network yields, given particular input values. Table 1 contains some summary statistics based on evidence entered into the network. Much of the evidence noted below is entered into nodes that are not shown in Figure 4 and thus have not been discussed in this paper.

Scenario	Summary Statistics	All SSEs	SSE1 & 2	SSE1	Collisions
1. Status Quo	P(event)	1.99E-05	5.75E-06	1.65E-07	1.65E-08
	Mean events per year	17.84	5.16	0.15	0.01
	Mean years between events			6.73	67.28
2. ATC Challenged	P(event)	4.23E-05	1.93E-05	4.45E-07	4.46E-08
	Mean events per year	38.00	17.35	0.40	0.04
	Mean years between events			2.5	24.98
3. Pilots Degraded	P(event)	2.49E-05	7.62E-06	1.26E-06	1.26E-07
	Mean events per year	22.40	6.85	1.13	0.11
	Mean years between events			0.89	8.87
4. Pilots Degraded + Mitigations	P(event)	2.16E-05	9.32E-06	3.18E-07	3.18E-08
	Mean events per year	19.40	8.87	0.29	0.03
	Mean years between events			3.50	35.02

Table 1: Summary Statistics

Scenario 1 is derived from the basic net but with a node modelling the sector unit culture set to good. When this evidence is propagated it can be seen that the mean number of collisions per year is 0.01 with a mean of 67.28 years between collisions. The probability of a collision is 1.65E-08 per year and the probability of the identified SSE events is as shown. Scenario 2 models the following node settings – “unit culture” = good, “ATC planner” = below average, “traffic volume” set at 30% above target sector flow, “traffic mix” set as challenging and “traffic characteristics” to complex. This scenario yields, for example, a mean number of collisions per year of 0.04 with a mean number of years between collisions of 24.98. Scenario 3 models the situation where the “pilot/crew mix” is below par. By “pilot/crew mix” is meant that the capabilities of the pilots, the quality of their English, their attentiveness to the radio, their effectiveness in responding to ATC clearances/requests (how fast do they respond and perform

necessary manoeuvres, do they question the need for avoiding action etc.). Finally, scenario 4 sets “pilot/crew mix” to below par, “TCAS” to highly supportive and “unit culture” to strong.

Each node in a Bayesian Network has an associated probability table called a node probability table (NPT). Figure 5 shows four nodes from the ATC network and their associated NPTs. The node “unit culture” has five states ranging from strong to destructive. The initial node probability table indicates that this variable is as likely to be in any of those five states thus we end up with a near uniform distribution. The “pilot aircraft” node has five states ranging from highly skilled to poor. However, in this table we see the distribution of capability suggests that around 24% are highly skilled, 37% professional and so on. This represents the distribution of capability likely to be encountered in practice (the prior probability). With these settings the network predicts the likelihood of a single accident as 5.56E-08. If evidence is now entered into these tables, as shown below the oval nodes in Figure 5, thus the unit culture is strong, the pilots and crew are below par and TCAS is highly supportive, then the risk of a single collision changes to 3.18E-08. Compare this with the value entered into Table 1 above. By identifying specific variables in the network and entering evidence, the changes in accident risk, associated with that evidence, can be investigated. In this way “what if” analysis can be conducted. Finally, the intent of the model is not to predict accurately the risk of an incident, although reasonable accuracy is required, but to investigate the overall risk exposure based on the different scenarios investigated.

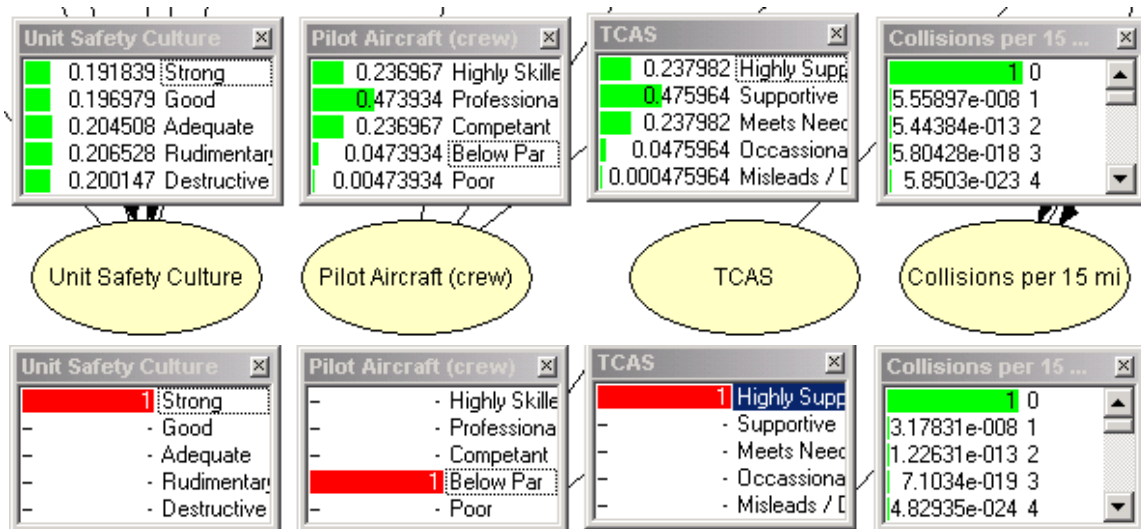


Figure 5: Node Probability Table

Conclusions

This paper has tried to show how the ATC Bayesian Network was developed and some of the reasoning that went into structuring the model. In addition a brief example has been given of inputting evidence into the model and determining the changes in risk exposure consequent upon those measures. The purpose of this network is to allow cultural and competence issues to be modelled and to investigate how changes in these variables impact the overall risk of collision. The Safety Research Group at Liverpool University has developed a range of instruments to allow safety culture to be measured and modelled. These measures feed into the parent variables of the culture node, briefly referred to above. Once an organisation’s safety culture has been measured the data can be entered into the network, along with other evidence, and propagated to determine whether it increases or decreases overall risk exposure. Models such as these are useful

for performing “what if” analysis to investigate, for example, the impact of planned changes on an organisation

References

1. Neil M, Shaw. R. et al. Measuring and Managing Culturally Inspired Risk. Current Issues in Safety Critical Systems, Proceedings of the Eleventh Safety-critical Systems Symposium, Bristol, UK, February 2003. Redmill F and Anderson T. ED., Springer. 2002.
2. Donald I and Shaw R. Safety Culture. Safety Systems The Safety-Critical Systems Club Newsletter. Volume 11, No. 3. May 2002.
3. Fenton N, Littlewood B, Neil M, Strigini L, Sutcliffe A, Wright D. Assessing Dependability of Safety Critical Systems using Diverse Evidence, IEE Proceedings Software Engineering, 145(1), 35-39, 1998.
4. SERENE Consortium, “SERENE (SafEty and Risk Evaluation using Bayesian Nets): Method Manual”, ESPRIT Project 22187, 1999.
5. Jensen F. Bayesian Networks and Decision Graphs. Springer-Verlag. 2001.
6. Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Francisco, 1988.
7. Fullwood R. R and Hall R. E. Probabilistic Risk Assessment in the Nuclear Power Industry. Pergamon Press. 1987.

Biography

Dr. Martin Neil. B.Sc., PhD., MIEE, C.Eng. RADAR Group, Department of Computer Science, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel. +44 20 7882 5221, fax +44 20 8980 6533, e-mail – martin@dcs.qmul.ac.uk.

Martin Neil is a Reader in the Computer Science Department of Queen Mary, University of London. His interests cover applications and theory in Bayesian probability to intelligent personalization of media content, software quality, system dependability and operational risk in finance. Martin is a Chartered Engineer and is also the chief technology officer at Agena.

Bob Malcolm., B.Sc., MIEE, C.Eng. RADAR Group, Department of Computer Science, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel. +44 20 7882 5223, fax +44 20 8980 6533, e-mail – bobm@ideo.co.uk.

Bob Malcolm is a Senior Research Fellow in the Computer Science Department of Queen Mary, University of London and Visiting Professor at City University, London. Bob is also the Managing Director of ideo ltd, a consultancy specializing in R&D strategy and management. He managed the UK Safety Critical Systems Research Programme and chairs the Steering Group of the Safety Critical Systems Club.

Roger Shaw. MBCS, MIEE, MSaRS. RADAR Group, Department of Computer Science, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel. +44 20 7882 5223, fax +44 20 8980 6533, e-mail – roger@dcs.qmul.ac.uk.

Roger Shaw is a Senior Research Fellow in the Computer Science Department of Queen Mary, University of London and is Head of Software Assurance with ERA Technology in Leatherhead, Surrey, UK.