Risk assessment using Bayesian decision networks

Tali Boneh

Ann Nicholson, Kevin Korb

(Monash University)

John Bally, Gary Weymouth , Peter Newham

(Bureau of Meteorology)

July 2006

Forecasting fog for Melbourne airport

- important task
- done by the forecasters; information used by airlines
- prior knowledge enables airlines:
    - carry more fuel on aircrafts
    - prepare for diversions (fly crews and planes)
        - ➢ avoid crash, save time and money
- false alarms are expensive and should be minimised

Problems

- uncertainty in observations
- incomplete knowledge
- forecasts are subjective

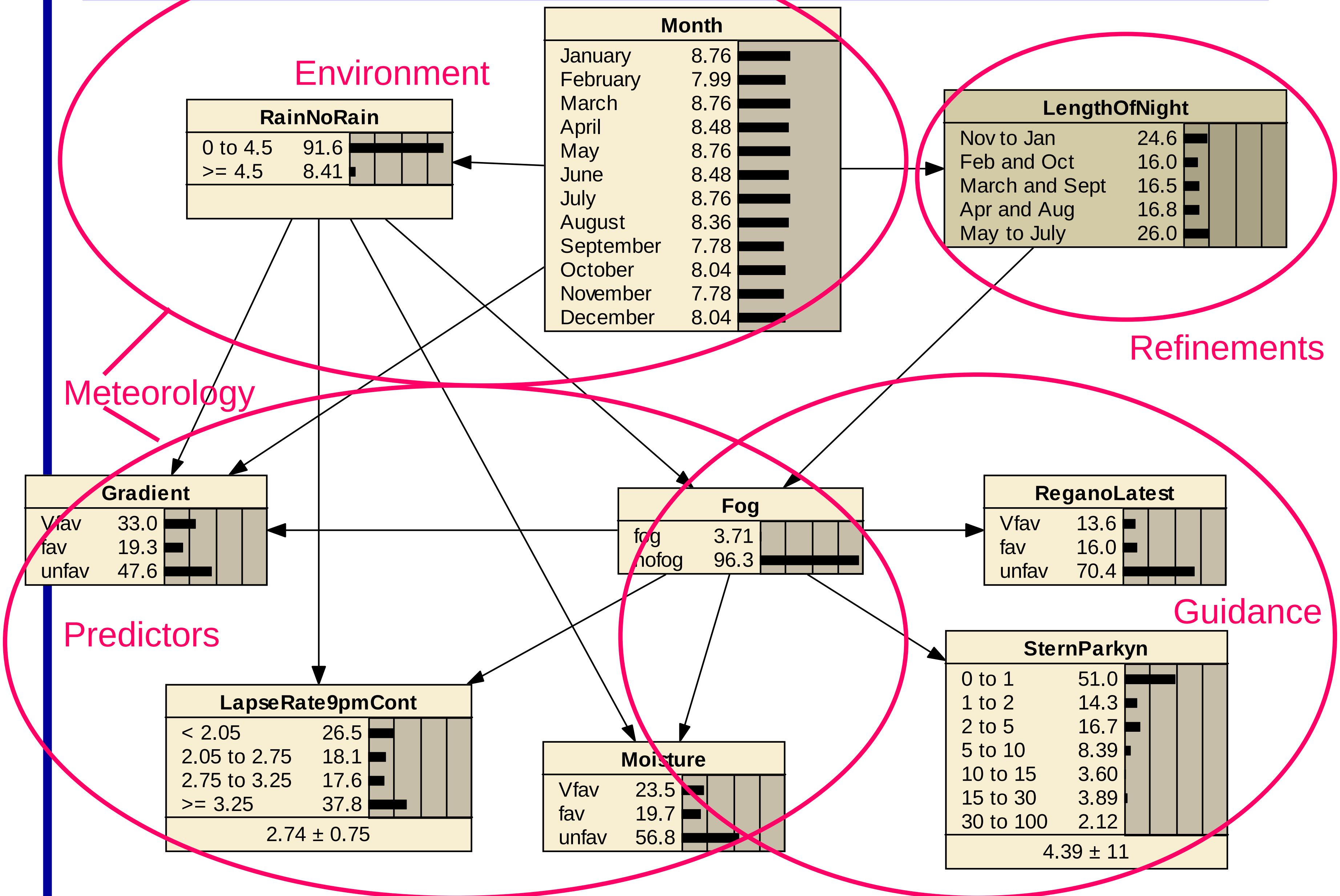
Aim of project

creating an <u>objective</u> guidance to support the prediction of Fog at Melbourne Airport

Reasoning under uncertainty: probability theory

- probability theory is one of the scientific ways of dealing with reasoning under uncertainty
  - applying formal statistics can yield better results, compared with subjective judgment
  - the final output of the process is objective and is based on a solid mathematical basis
- for applying probability theory we have used Bayesian Networks technology

Bayesian Network for forecasting fog



Evaluation: using a ROC curve

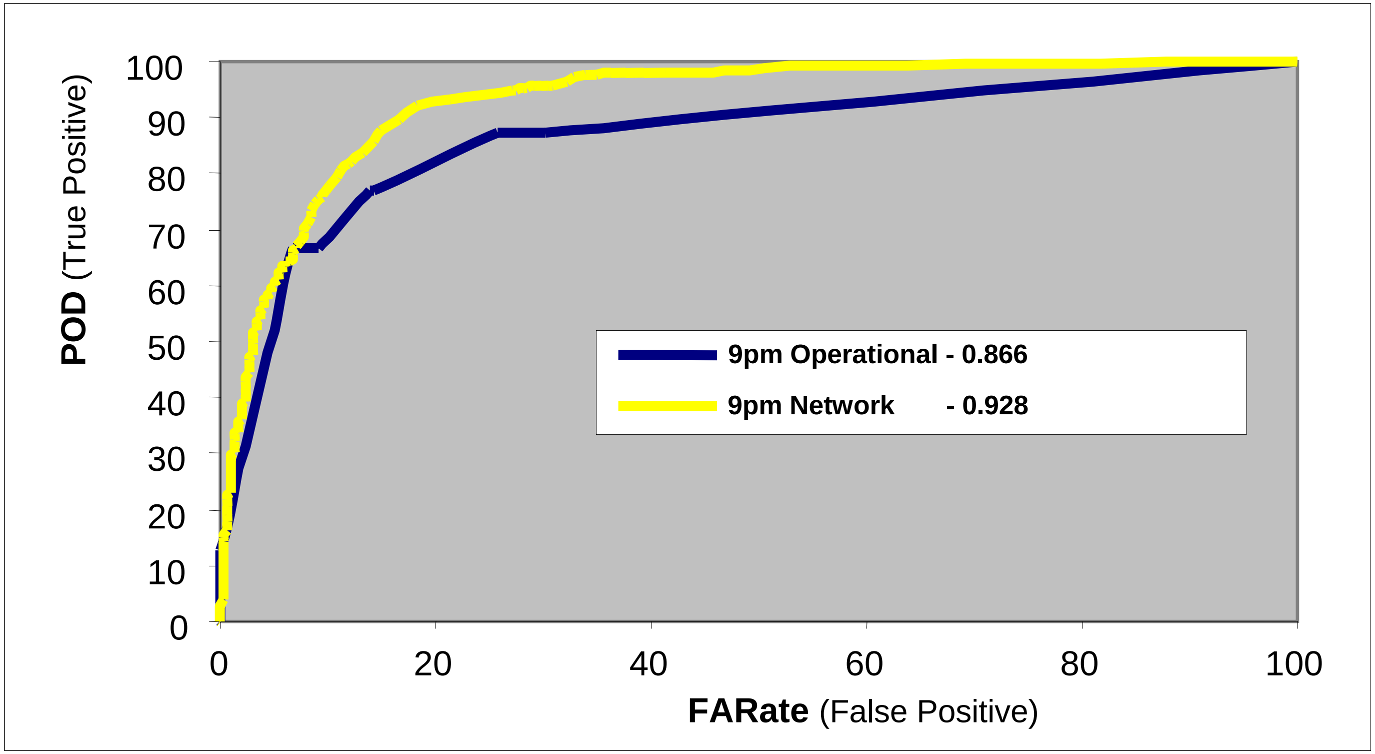
- Receiver Operating Characteristic (ROC) curves
- P(actual event predicted as event)

  vs.

  P(actual no-event predicted as event)
- Area Under Curve (AUC) can be used as a measurement of the accuracy of the model (perfect model: AUC = 1)
- ROC curve can be used to find cutoff values

ROC evaluation of the Melbourne network

- Bayesian Networks yield probabilities
- evaluation by ROC curves



Forecasters' evaluation measures

- POD (True Positive Rate)

$$\frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})} = \#\text{fog events}$$

- False Positive Rate

$$\frac{\text{False Positive}}{(\text{False Positive} + \text{True Negative})} = \#\text{no-fog events}$$

- False Positive Ratio (FAR)

$$\frac{\text{False Positive}}{(\text{False Positive} + \text{True Positive})} = \#\text{fog was forecasted}$$

POD & FAR – operational (current) *vs.* network

we take Bayesian Network output and apply cut-offs to make fog forecasting decisions

- 1% cutoff was used for Code Grey (low probabilities of fog)
- 20% cutoff was used for TAF (high probabilities of fog)

Forecast	**Operational**		**Network**	
	POD (%)	**FAR (%)**	**POD (%)**	**FAR (%)**
3pm TAF	**56**	**73**	**67**	**76**
3pm TAF and Code Grey	**87**	**90**	**94**	**89**
9pm TAF	**67**	**73**	**69**	**78**
9pm TAF and Code Grey	**87**	**90**	**95**	**89**

TAF - **T**erminal **A**erodrome **F**orecasts (high prob of fog >= 30%), **Code Grey** – (low prob of Fog < 30%)

POD – Power Of Prediction = True Positive Rate, FAR – False Alarm Ratio

Refining decision using forecasters cut-offs

Forecast Decision	CUT-OFF
No fog	0-1
<5% Code Grey	1-5
5% Code Grey	5-10
10% Code Grey	10-15
20% Code Grey	15-20
Fog on TAF	20-100

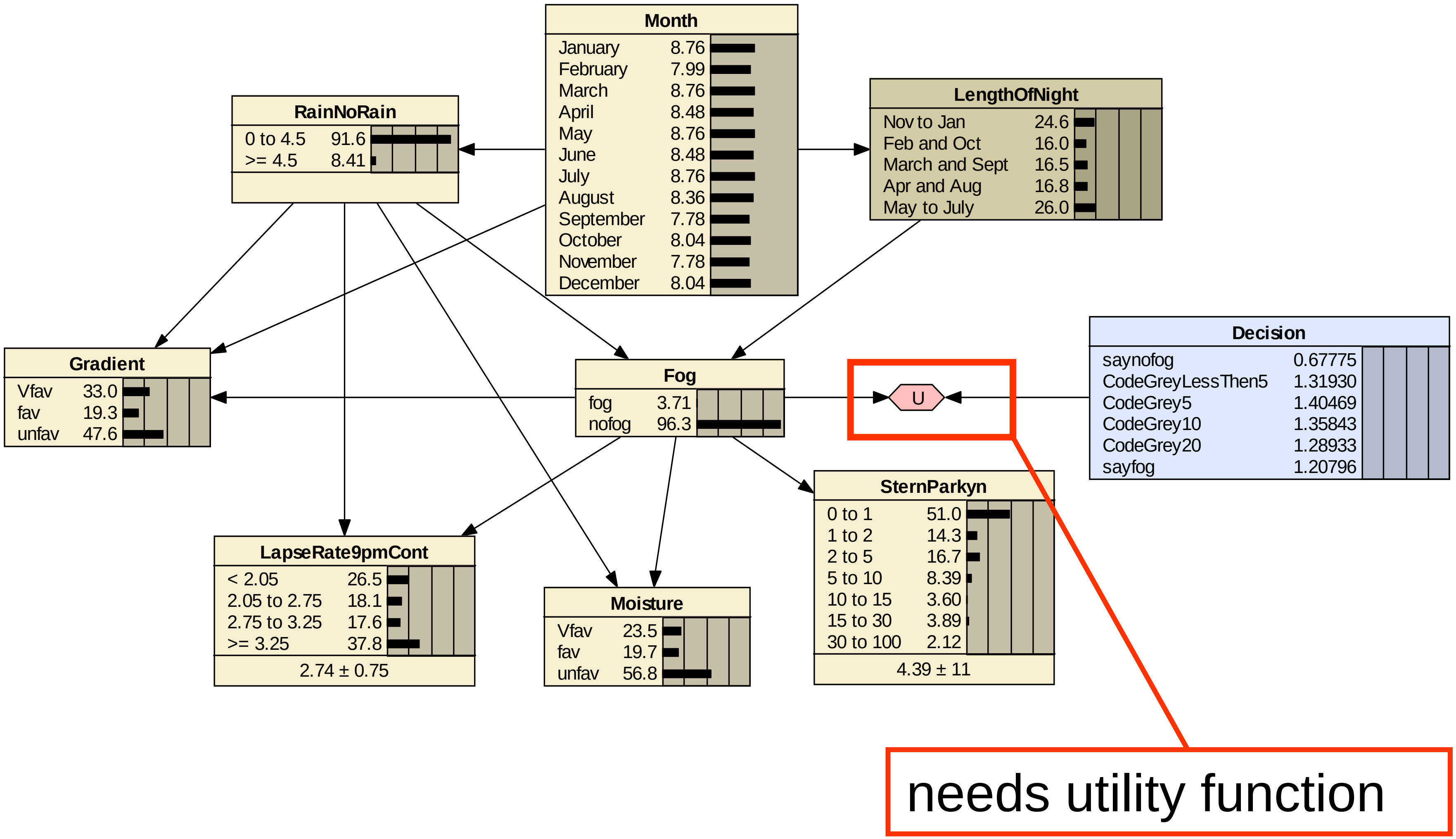
TAF - **T**erminal **A**erodrome **F**orecasts (high prob of fog >= 30%), **Code Grey** – (low prob of Fog < 30%)

POD – Power Of Prediction = True Positive Rate, FAR – False Alarm Ratio

Turning probability to decision: decision theory

- the calculation of the probability distribution of Fog, given observations, is based on Probability Theory
- Instead of using cut-offs on probabilities we apply Decision Theory
- we use Decision Theory to calculate the expected utility of a situation
- a utility function assigns a single number (preference) to a situation to express its desirability
- given new evidence, the decision network calculates the expected utility of the possible decisions
- a rational 'agent' chooses an action that maximises its expected utility

decision network for Fog



Different situations and forecasters' preferences

decision	actual	preferences
Say No Fog	Fog	1
Say Code Grey - less than 5% chance of fog	Fog	3
Say Code Grey - 5% chance of fog	Fog	36
Say Code Grey - 10% chance of fog	Fog	38
Say Code Grey - 20% chance of fog	Fog	39
Say TAF – Prob Fog	Fog	40
Say No Fog	No Fog	20
Say Code Grey - less than 5% chance of fog	No Fog	19
Say Code Grey - 5% chance of fog	No Fog	18
Say Code Grey - 10% chance of fog	No Fog	17
Say Code Grey - 20% chance of fog	No Fog	16
Say TAF – Prob Fog	No Fog	15

BUT, what are the aviation decision makers preferences?

TAF - **T**erminal **A**erodrome **F**orecasts (high prob of fog >= 30%), **Code Grey** – (low prob of Fog < 30%)

Preferences elicitation

- direct preferences elicitation is not possible because we do not know actual risk-analysis models used in aviation industry
- instead, we need to find preferences using the information we have
- we know that
  - aviation industry would like POD = 100%, FAR = 0%
  - current operational forecast POD = 56, FAR = 73
  - situations can be ranked according to significance
    - missing fog may be catastrophic
    - false alarm is un-desirable

find utility function that improves POD and FAR

POD – Power Of Prediction = True Positive Rate, FAR – False Alarm Ratio

Expected utility function

- given the network structure, the Expected Utility (EU) of a particular decision is defined to be the sum of the possible outcomes of the decision weighted by their probabilities:

EU(decisionA)=

P(Fog)*u(Fog,decisionA)+P(noFog)*u(NoFog,decisionA)

(a simplified definition)

Finding utility function by solving constraints

Instead of direct preferences elicitation from experts the next steps are:

- make a set of constraints
- feed it to constraint solver
- use the new constraint solver-produced utilities

Set of constraints for preferences

order

- X1<X2<X12<X11<X10<X9<X8<X7<X3<X4<X5<X6

use cut-off probabilities + utility function definition to form constraints

(1<i<7 and 7<j<13)

- for 0<=P(fog)<1 =>
  P(fog)*X1 + P(nofog)*X7 > P(fog)*Xi + P(fog)*xj
  (for i<>1 and j<>7)
- for 1<=P(fog)<5 =>
  P(fog)*X2 + P(nofog)*X8 > P(fog)*Xi + P(fog)*xj
  (for i<>2 and j<>8)
- etc.

more constraints

- X2 < 9*X1 (X2 is not that great)
- X12 > 16*X1 (Worst false alarm is 16 times better than a complete miss)
- X3 < 80*X1 . . .

decision	actual	pref
Say No Fog	Fog	X1
Say Code Grey - less than 5% chance of fog	Fog	X2
Say Code Grey - 5% chance of fog	Fog	X3
Say Code Grey - 10% chance of fog	Fog	X4
Say Code Grey - 20% chance of fog	Fog	X5
Say TAF – Prob Fog	Fog	X6
Say No Fog	No Fog	X7
Say Code Grey - less than 5% chance of fog	No Fog	X8
Say Code Grey - 5% chance of fog	No Fog	X9
Say Code Grey - 10% chance of fog	No Fog	X10
Say Code Grey - 20% chance of fog	No Fog	X11
Say TAF – Prob Fog	No Fog	X12

Different situations and constraints preferences

decision	actual	forecasters	output of constraints solver
Say No Fog	Fog	1	1
Say Code Grey - less than 5% chance of fog	Fog	3	11
Say Code Grey - 5% chance of fog	Fog	36	62
Say Code Grey - 10% chance of fog	Fog	38	84.16
Say Code Grey - 20% chance of fog	Fog	39	96.75
Say TAF – Prob Fog	Fog	40	106.42
Say No Fog	No Fog	20	29.65
Say Code Grey - less than 5% chance of fog	No Fog	19	29.55
Say Code Grey - 5% chance of fog	No Fog	18	26.87
Say Code Grey - 10% chance of fog	No Fog	17	24.42
Say Code Grey - 20% chance of fog	No Fog	16	22.21
Say TAF – Prob Fog	No Fog	15	19.8

yields utility function for a decision network

relative importance of missing fog larger then forecaster estimate

Conclusions

- it is important to distinguish between the user of the Bayesian Network (forecasters) and the user of the information the network provides (aviation industry)

- inferring the utility function using available information enables the use of decision theory even if direct elicitation of preferences from experts is impossible

Risk assessment using Bayesian decision networks

Tali Boneh

Ann Nicholson, Kevin Korb

(Monash University)

John Bally, Gary Weymouth , Peter Newham

(Bureau of Meteorology)

July 2006