Australian Government
Department of Defence
Defence Science and Technology Organisation

A Framework to Model and Measure System Effectiveness

Neill Smith
Consultant to DSAD
Mission Software P/L

Introduction

Motivation for current study

1. Definitions of Measures of Effectiveness
2. A simple thought experiment
3. Required properties required of MoE.
4. Approaches from decision theory
5. Applied in 2 examples
6. Applicability of Framework
7. Concluding remarks

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Motivation for study

Want to predict effectiveness and

- Measure comparative effectiveness
- Earlier work highlighted major issues
  - Difficulty in relating effectiveness to performance
  - Difficulty in aggregating effectiveness measures
  - Need to deal with a system in its context
  - Failure to predict impact of disruptive technology
  - Uncertainty in data, interactions and contributions
  - Immaturity of field, no widely agreed definitions or methods
  - Need to include effect of qualitative impacts

These are partially addressed in paper

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Definitions : Measures of Effectiveness

Sproles (2001) postulated that Measures of Effectiveness (MoE) are required to answer the question **“Does this meet my need?”** and hence defined MoE as

***“standards against which the capability of a solution to meet the needs of a problem may be judged. The standards are specific properties that any potential solution must exhibit to some extent. MOEs are independent of any solution and do not specify performance or criteria*”.**

He distinguishes between Measures of Performance (MoP) and MoE by declaring that MoP measures the internal characteristics of a solution while MoE measure external parameters that are independent of the solution – a measurement of how well the problem has been solved.

Definitions : Measures of Effectiveness

Dockery's (1986) definition:

"A measure of effectiveness is any mutually agreeable parameter of the problem which induces a rank ordering on the perceived set of goals".

Smith and Clark, (2004) definition:

*"A measure of the ability of a system to meet its specified needs (or requirements) from a particular viewpoint.
This measure may be quantitative or qualitative and it allows comparable systems to be ranked.
These effectiveness measures are defined in the* ***problem-space****. Implicit in the meeting of problem requirements is that threshold values must be exceeded."*

A Simple Thought Experiment

To explore these issues a small thought experiment was developed,
Needing these characteristics

- Easy to measure attributes
- Obvious interpretation of effectiveness
- Easy to perturb in such a way that equivalence could be maintained
- Complex in terms of resource usage and interactions
- Able to be measured at varying levels of resolution

An obvious candidate is a computer program, as not only can
the above be achieved.

A Simple Thought Experiment

MoP2

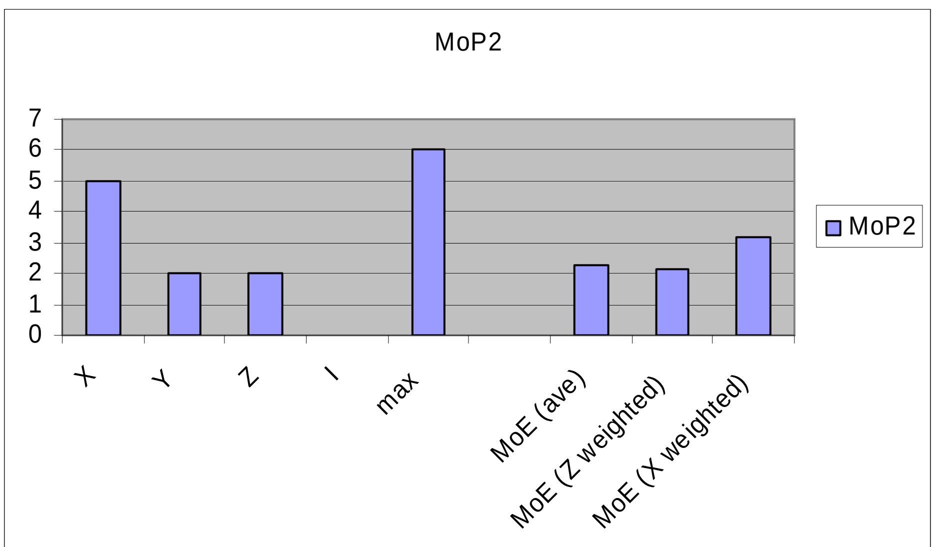
```
Begin
    X :=0; y:= 0;
    While (X<101) do
      Begin
          Y := Y + X;
          Z := Z– X;
          X := X + 1;
          output(Z);
      End; { of while}
End.
```

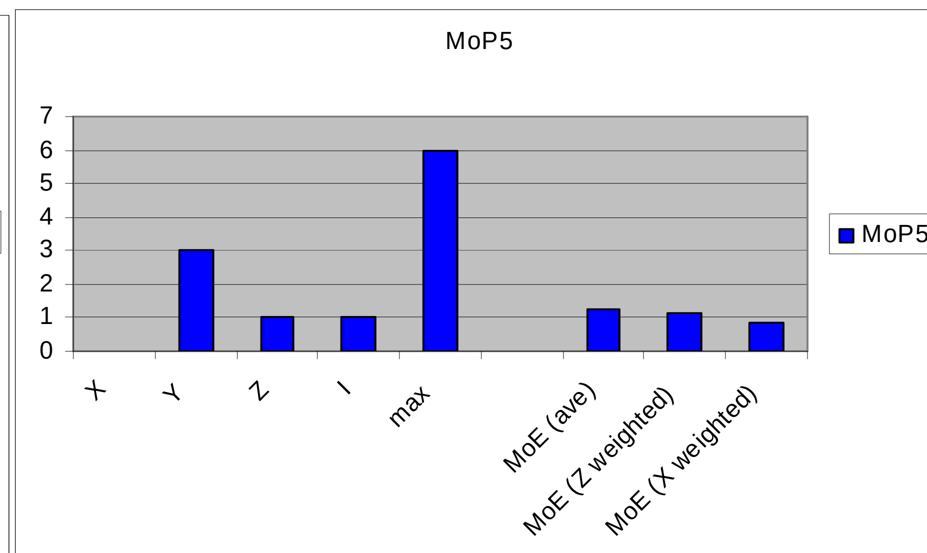
MoP3

```
Begin
 X :=0; y:= 0;
 While (X<101) do
    Begin
        Y := Y + X;
        X := X + 1;
        Z := Z– X;
        output(Z);
    End; {of while }
End.
```

MoP5

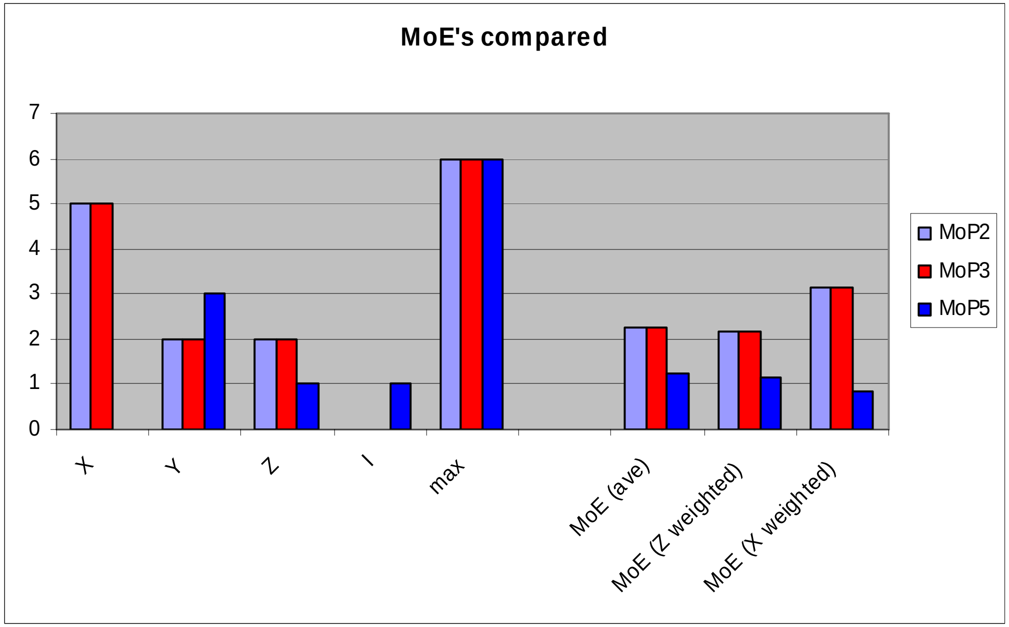
```
Begin
   X :=0; y:= 0;
   For I:=1 to 100 do
      Begin
           Z := – Y;
           Y := Y + I ;
           output(Z);
      End; {of for I }
End.
```





A Simple Thought Experiment

Australian Government
Department of Defence
Defence Science and
Technology Organisation



OUTCOMES

- •Performance Measures can produce misleading effectiveness measures
- •Internal measures can help quantify effectiveness but choice difficult
- •Aggregation is method unclear
- •More holistics may distinguish between MoP2 and MoP3
- •Regardless of measure MoP5 best
- •Comparative effectiveness is a useful concept

Properties required of comparative effectiveness measures

- Increase with improvement (of effectiveness)
- Total score cannot exceed score of ideal score
  should be zero for non-compliance (system not effective)
- Should support system decomposition and aggregation
- Should be normalised to [0,1] to facilitate comparison
  between systems
- Measures should be ratio scales

Two approaches were investigated which met these requirements.

1. **MUAT and VFT**
2. **BN and probability based measures**

Approaches from decision theory

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Utility functions are widely used in Decision theory (Multi-attribute Utility Theory, MAUT) meet all these requirements. Value Focussed Thinking (VFT) is often used to derive the utility function.

Need to also determine what objectives need to be met for a system to be effective and this is a primary goal of VFT.

VFT and MAUT provide:

- well-grounded, consistent mathematical framework
- hierarchical and network decompositions
    - values hierarchy (with specified properties)
    - means-end network
- strong emphasis on problem domain, NOT the solution domain
    - (by) focussing on values and fundamental objectives

BUT assumes that decision maker knows what they want and can specify values.

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Approaches from decision theory

The **Value Focussed Thinking** approach provides a mechanism to guide this process.

Where fundamental objectives should have these properties:

- **Essential**: indicate consequences in terms of the fundamental reasons for interest in the situation
- **Complete**: include all fundamental aspects of consequences
- **Measurable**: to define objectives precisely and to specify the degree to which objectives may be achieved
- **Operational**: to render the collection of information required for an analysis reasonable considering the time and effort available
- **Decomposable**: to allow separate treatment of different objectives in the analysis
- **Non-redundant**: to avoid double counting
- **Concise**: to reduce no. of objectives needed for analysis
- **Controllable**: address consequences influenced by choice of alternatives
- **Understandable**: to facilitate generation and communication of insights.

Approaches from decision theory

Bayesian Networks and Influence Diagrams provide and support:

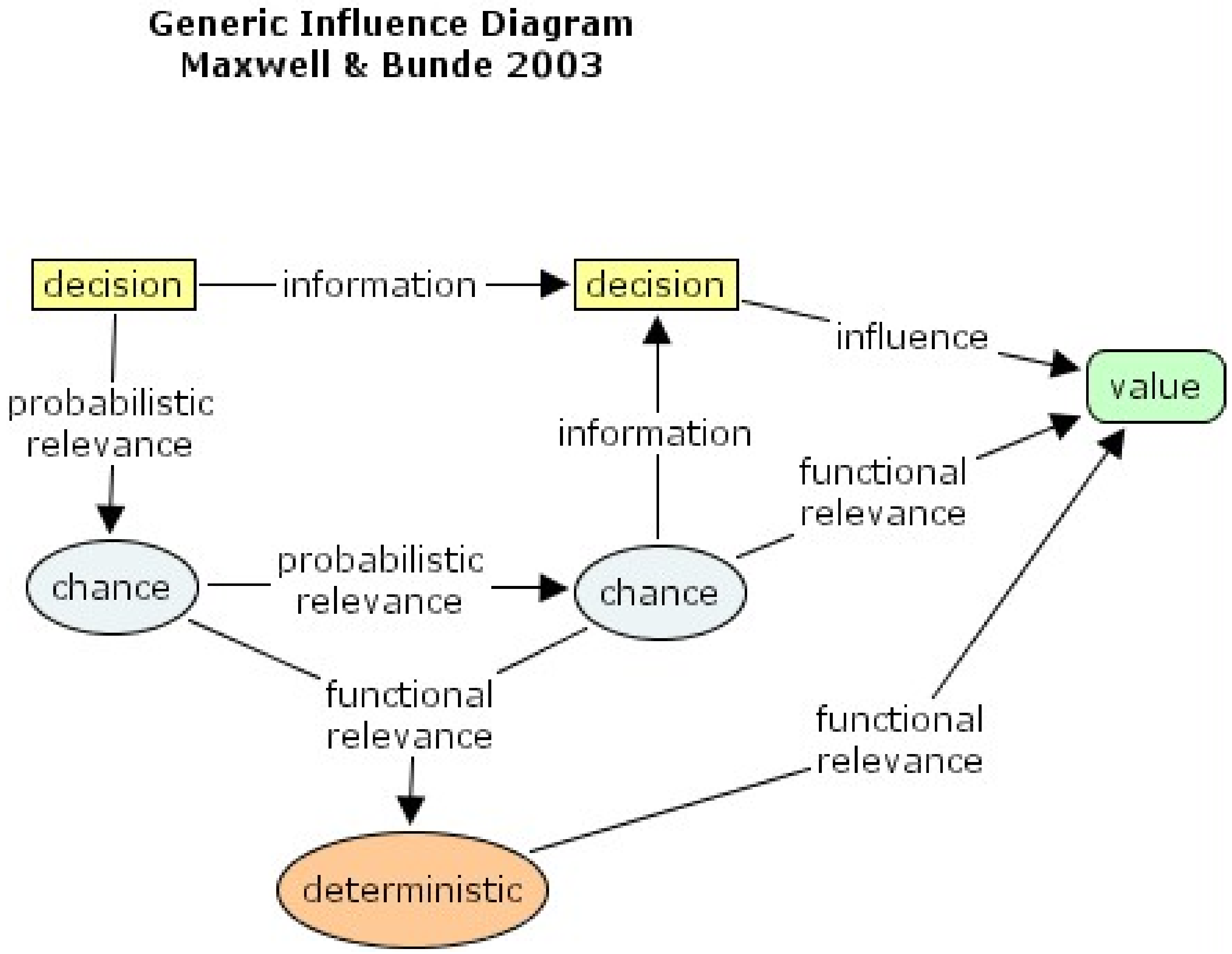
- a well-grounded, consistent mathematical framework
- hierarchical and network decompositions
    - values hierarchy
    - means-end network
- restricted to acyclic graphs

Based on Cox's (1946,1961) work: **subjective probability obeys rules of probability theory**

BN can be used to propagate effectiveness measures (Effectiveness constrained to [0,1])

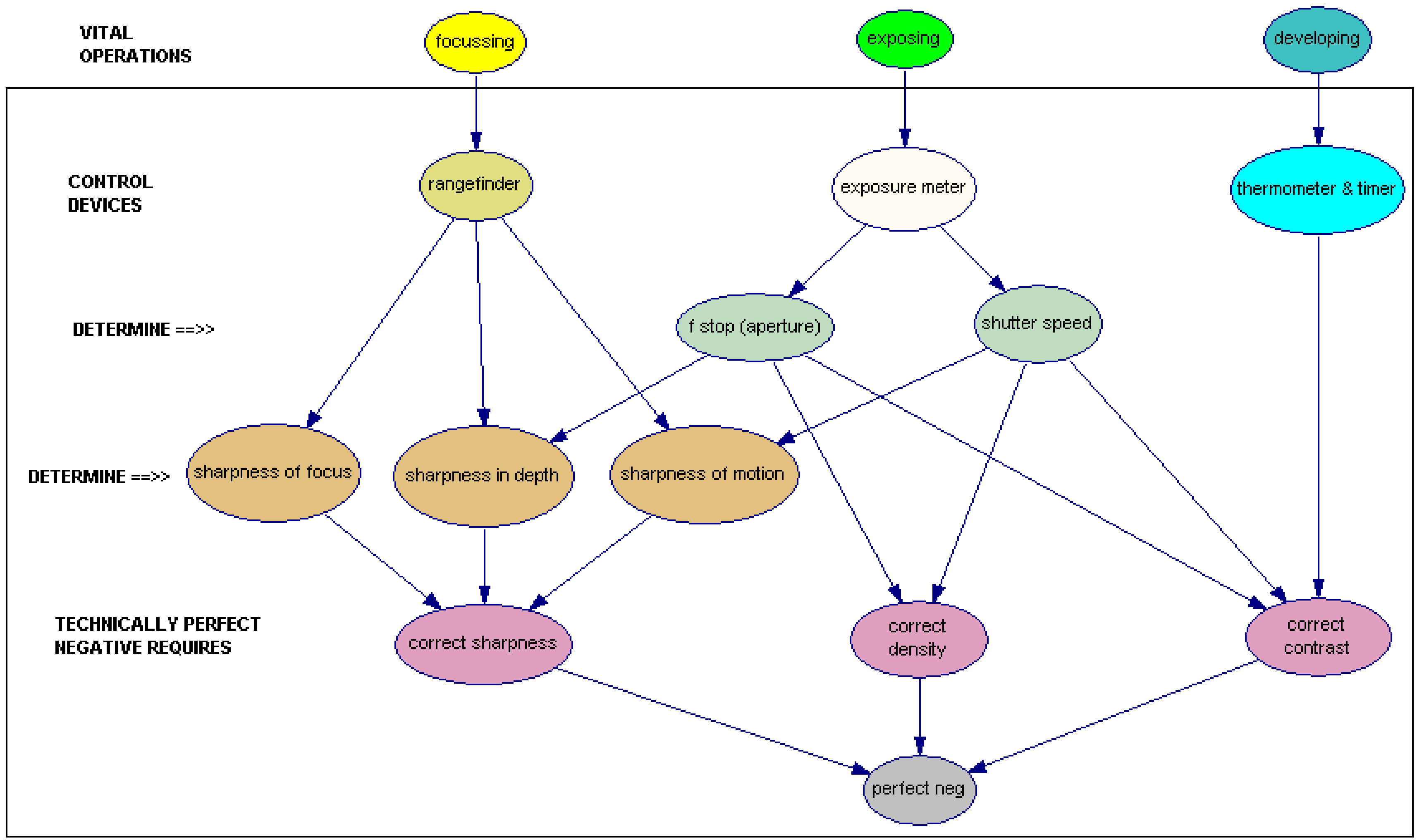
Australian Government
Department of Defence
Defence Science and Technology Organisation

Influence Diagrams



Feininger's Perfect Negative

Australian Government
Department of Defence
Defence Science and Technology Organisation



Initial Node Values (Genie)

Australian Government
Department of Defence
Defence Science and Technology Organisation

Node properties: exposure meter

General | Definition | Observation Cost | Format | Documentation | User properties

Add | Insert

correct	0.9
incorrect	0.1

Node properties: shutter speed

General | Definition | Observation Cost | Format | Documentation | User properties

Add | Insert

exposure meter	correct	incorrect
correct	0.95	0.3
incorrect	0.05	0.7

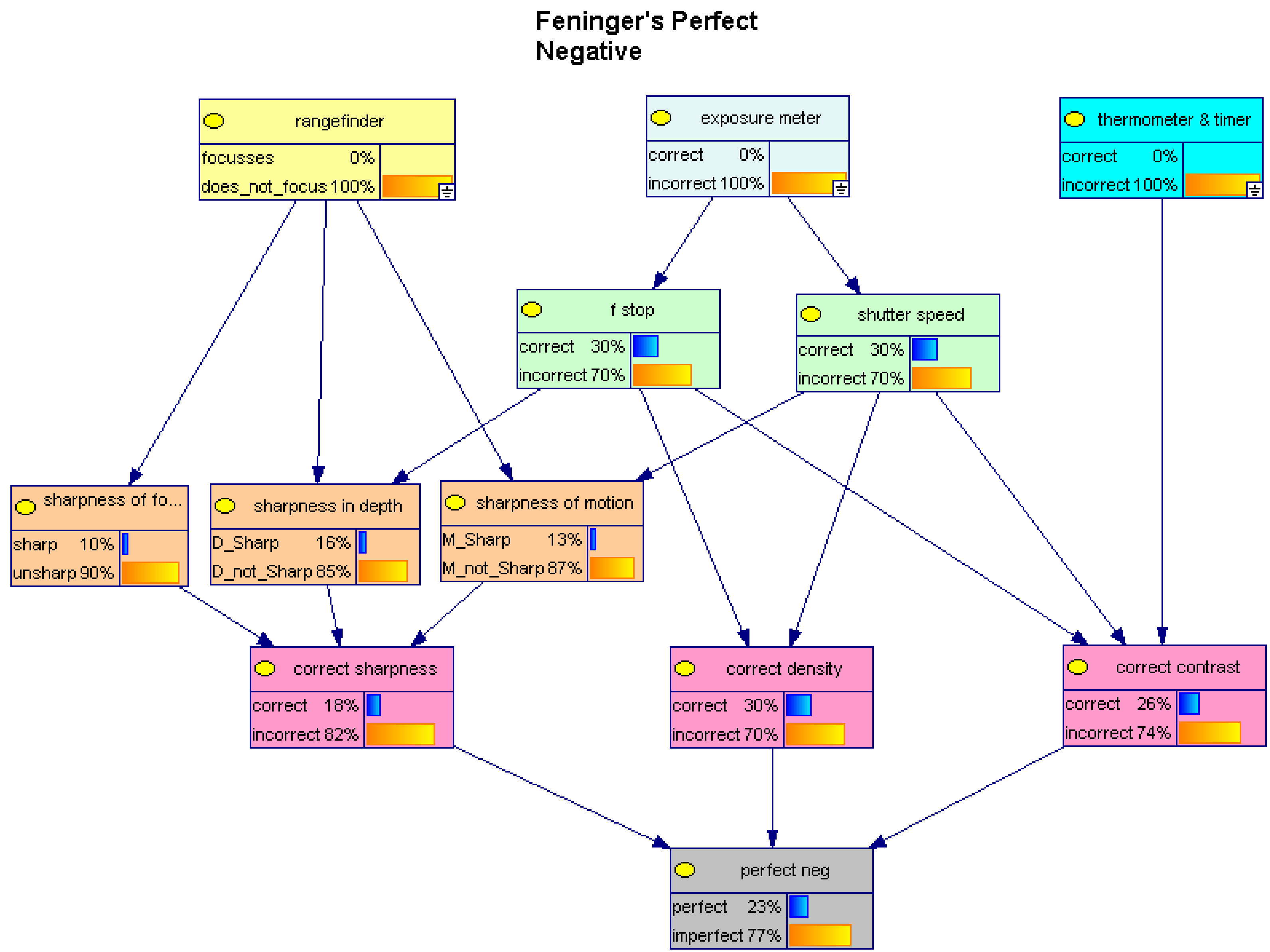
Node properties: sharpness of motion

General | Definition | Observation Cost | Format | Documentation | User properties

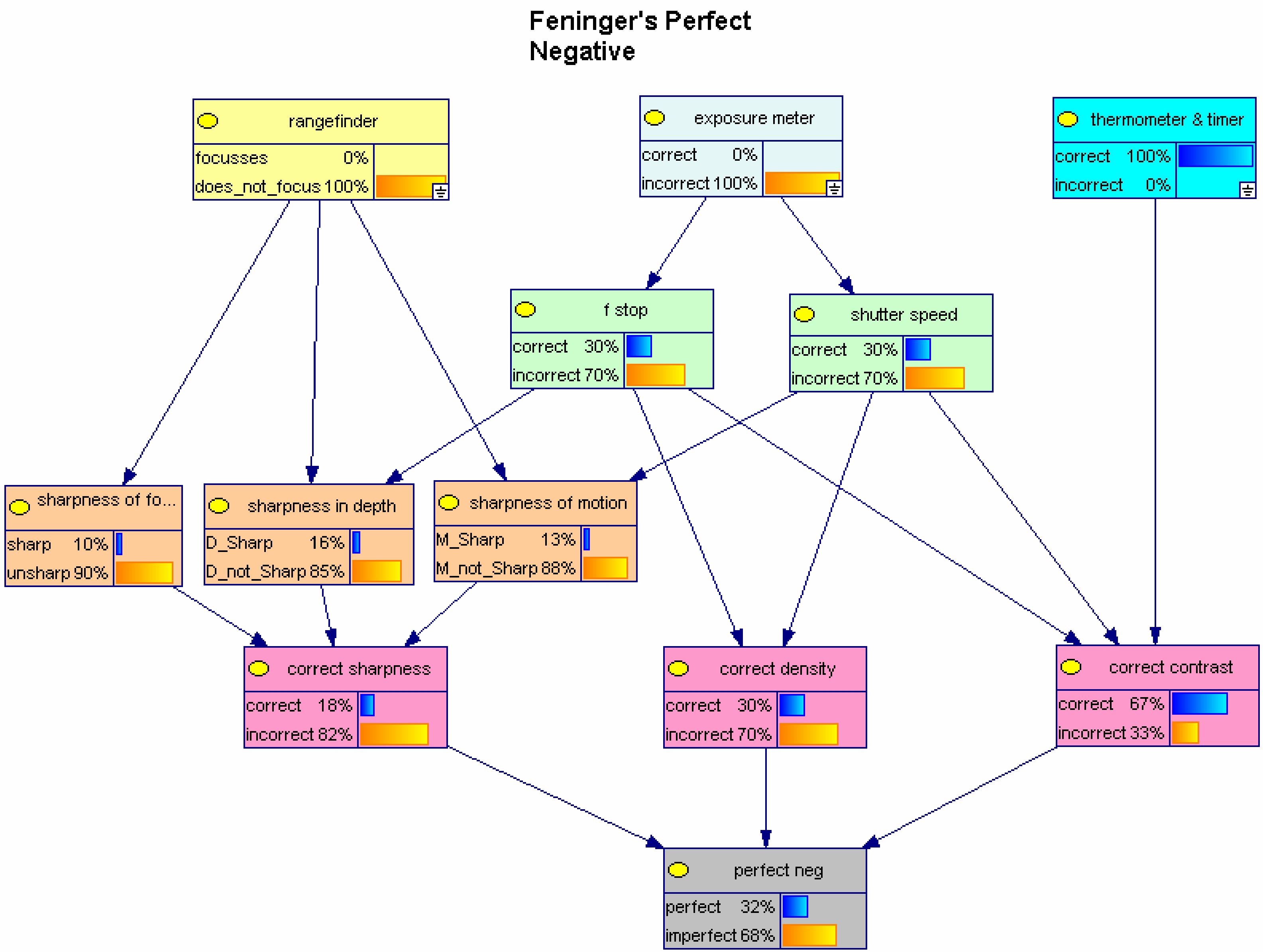
Add | Insert

rangefinder	focusses		does_not_focus	
shutter speed	correct	incorrect	correct	incorrect
sharp_in_M	1	0.2	0.3	0
nor_sharp_i...	0	0.8	0.7	1

Equivalent Influence Diagram

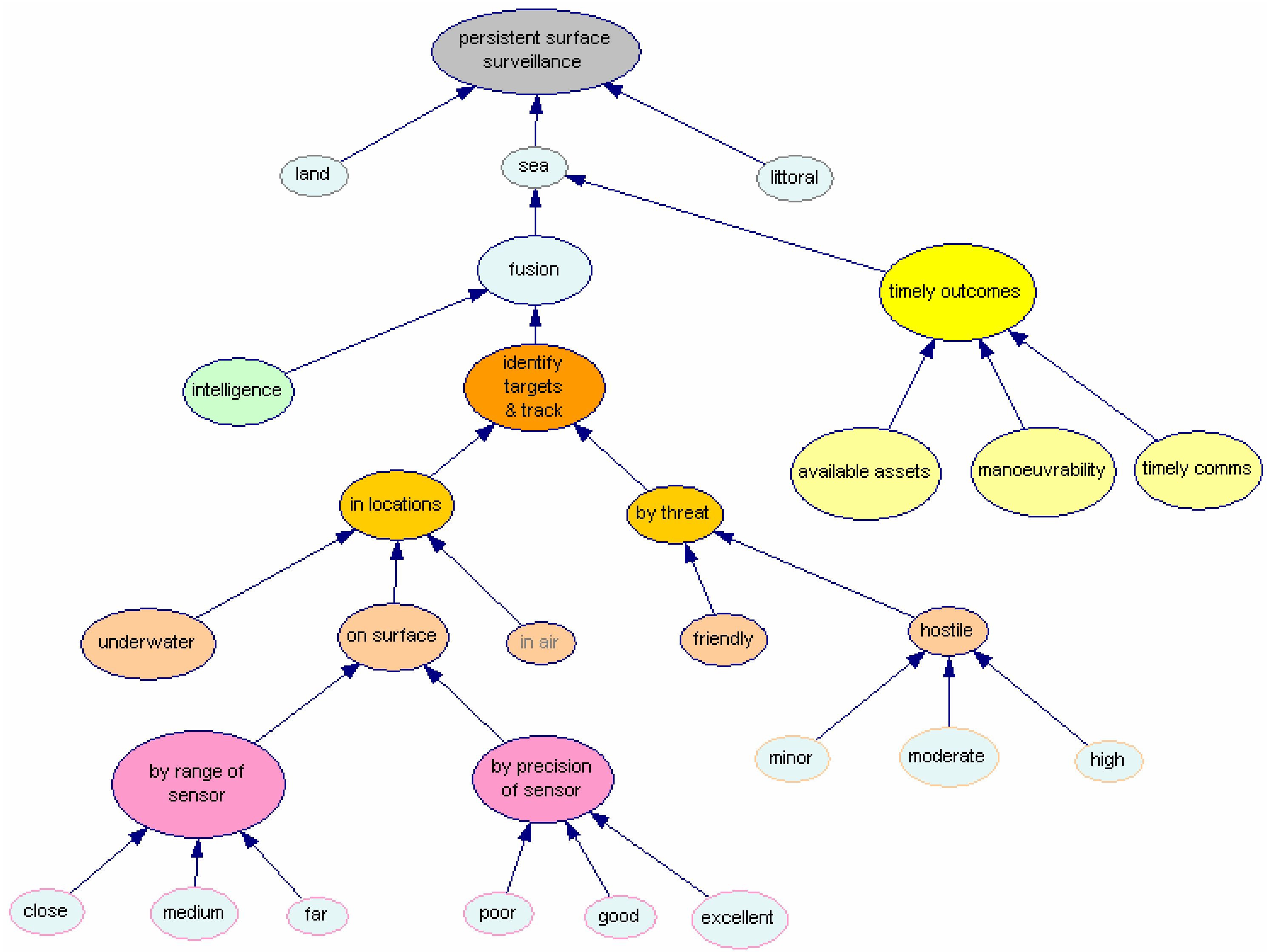


Equivalent Influence Diagram

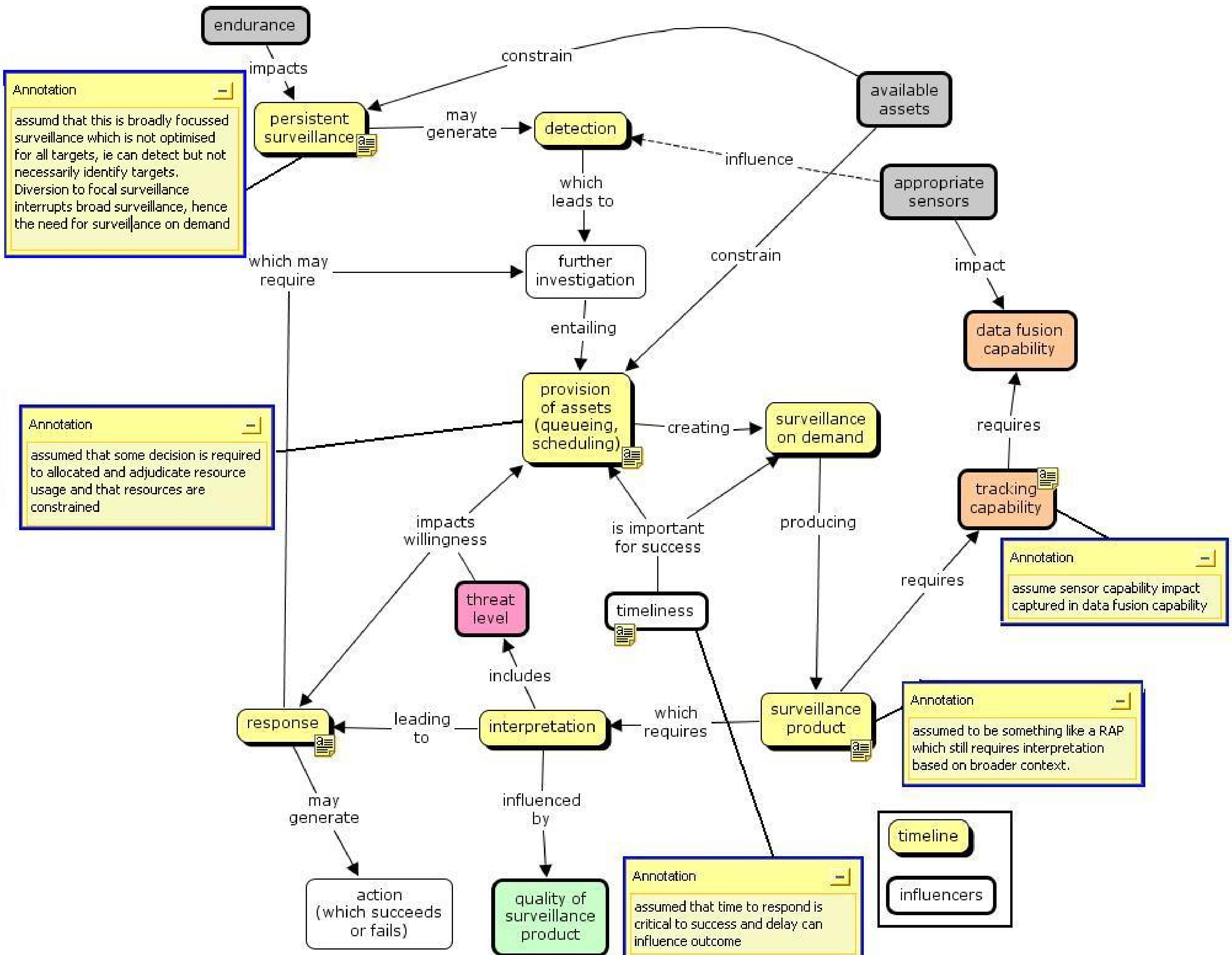


Value Model

Australian Government
Department of Defence
Defence Science and
Technology Organisation

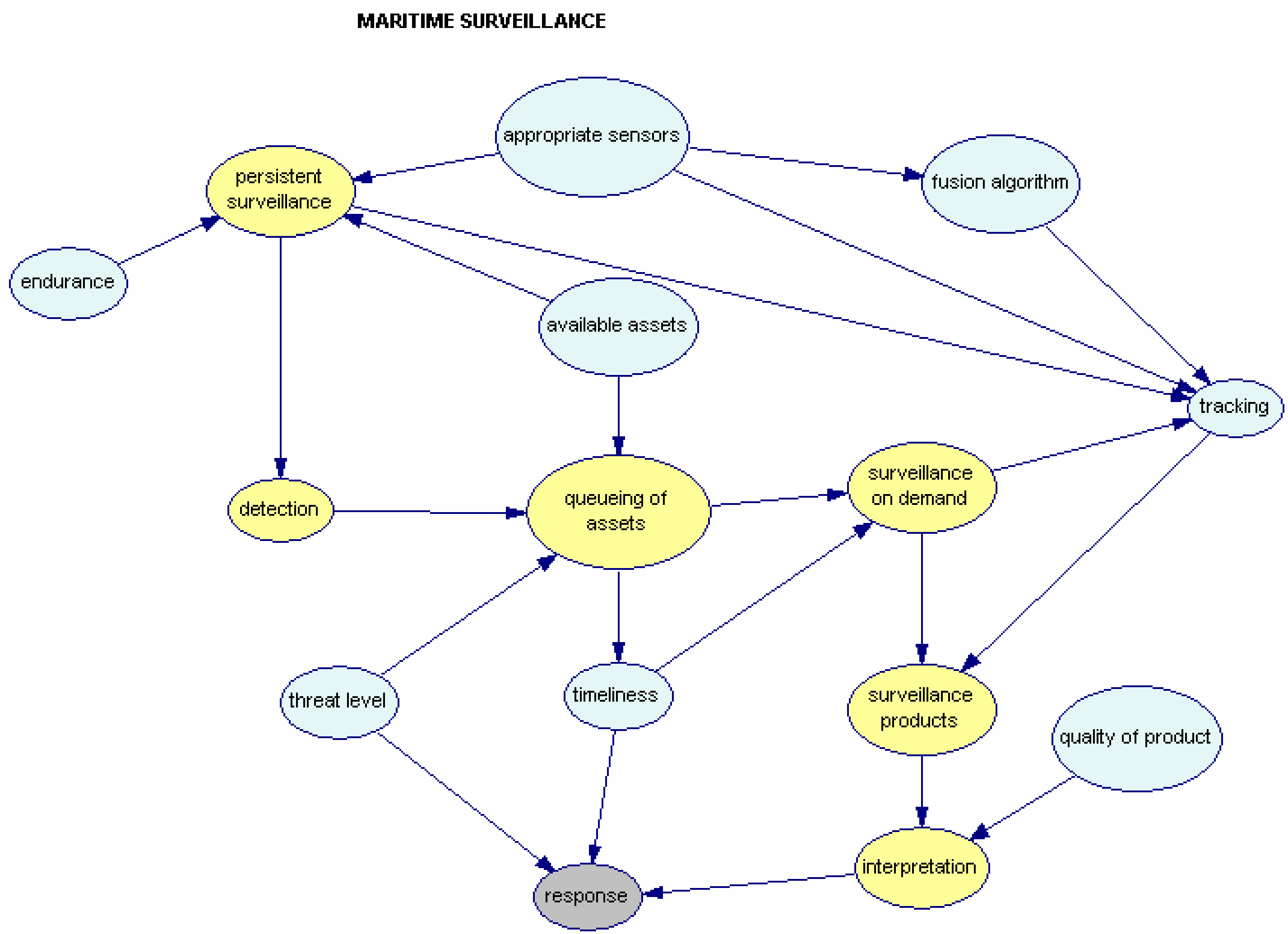


Causal model and assumptions



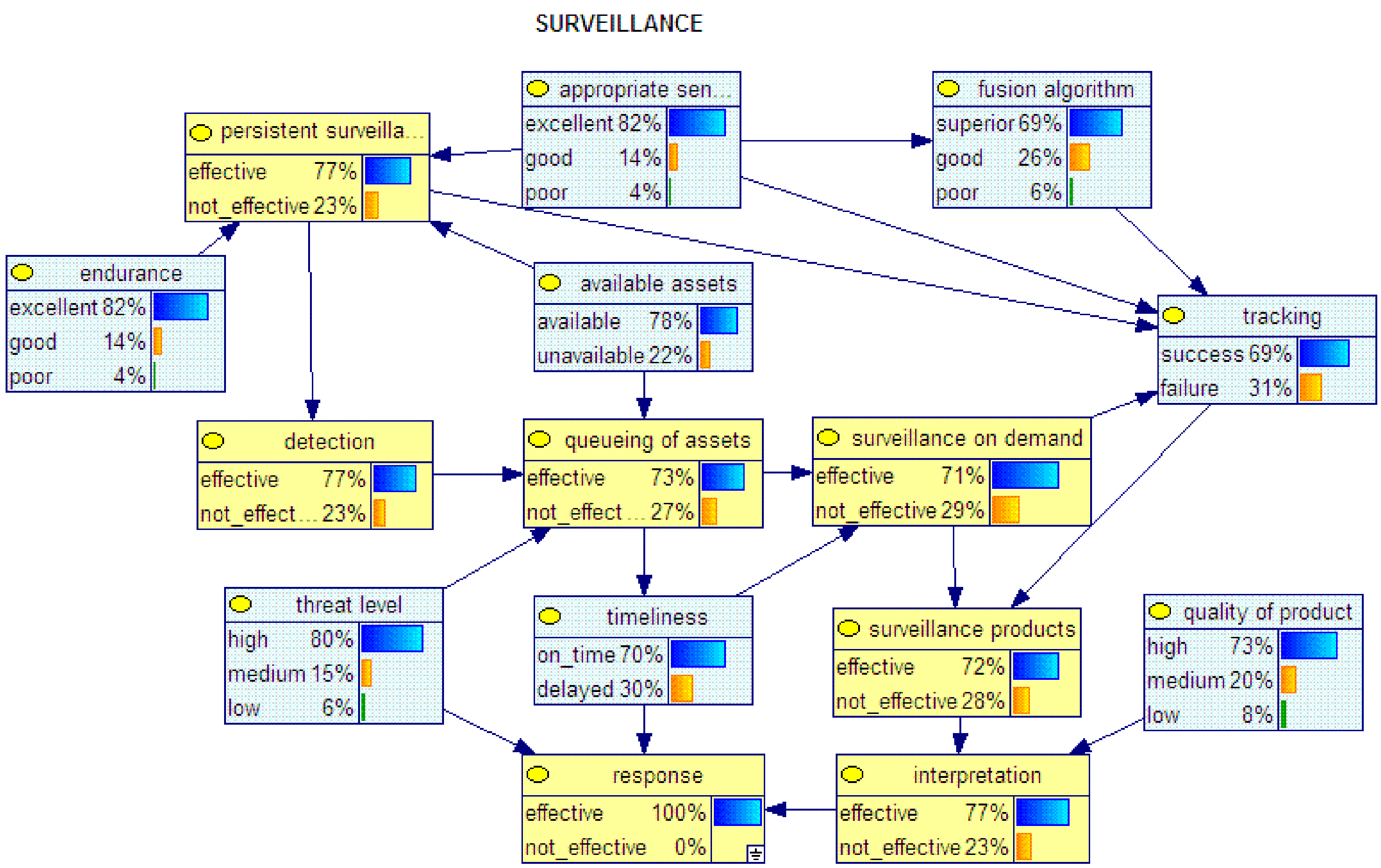
Australian Government
Department of Defence
Defence Science and Technology Organisation

Causal Model



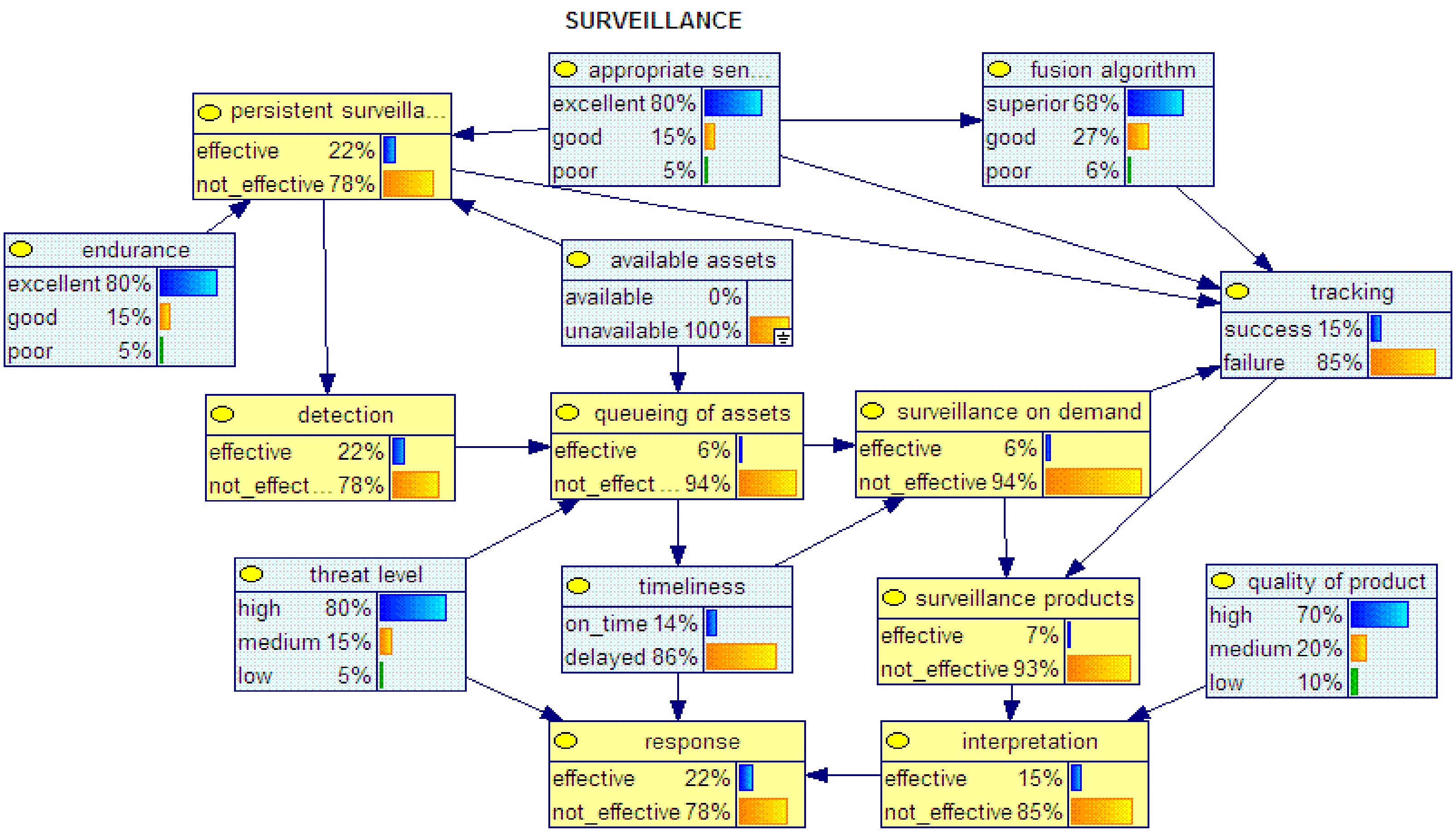
Australian Government
Department of Defence
Defence Science and Technology Organisation

Instantiated Influence Diagram : success



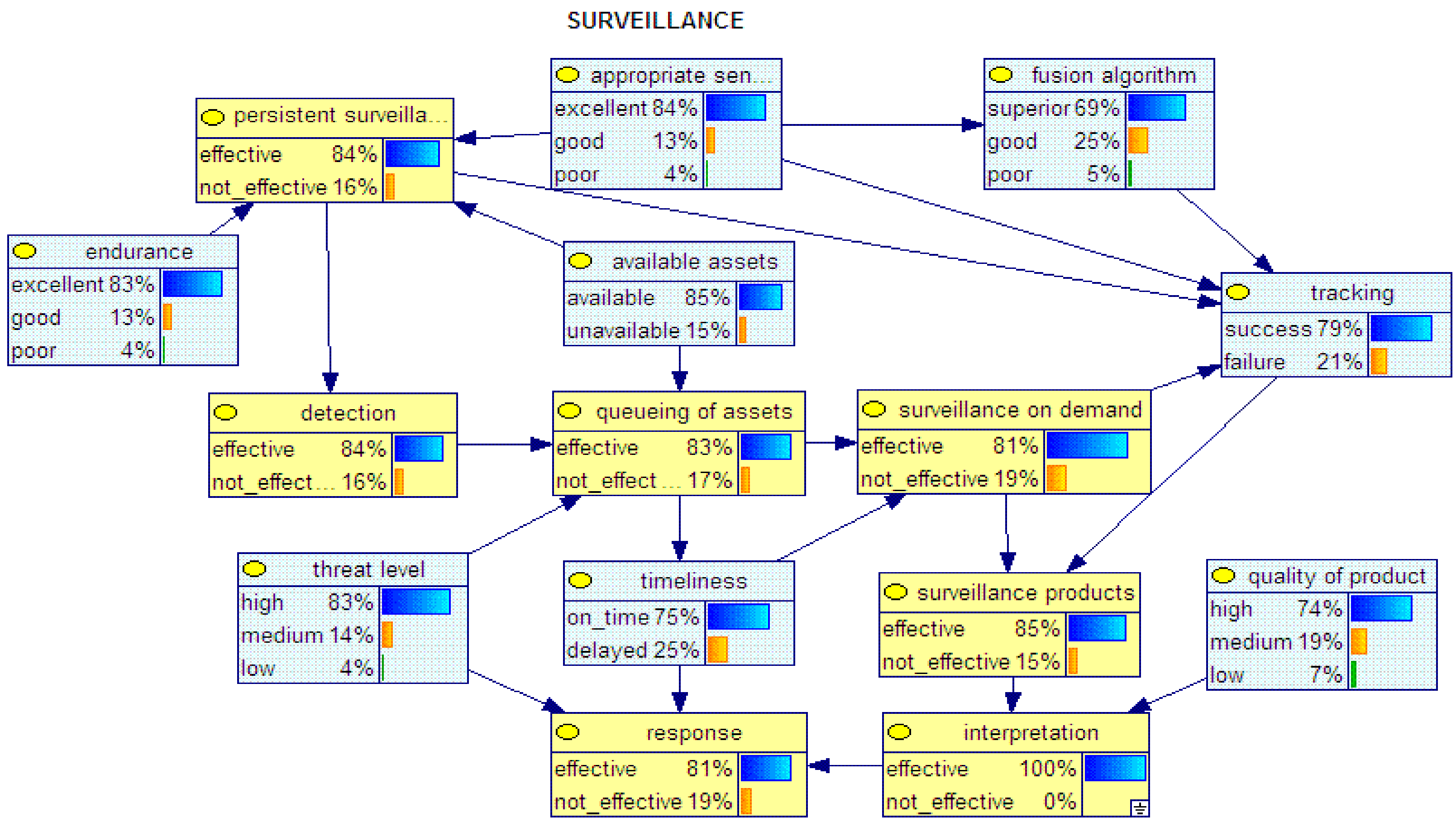
Instantiated Influence Diagram: no assets

Australian Government
Department of Defence
Defence Science and Technology Organisation



Instantiated Influence Diagram: interpretation success

Australian Government
Department of Defence
Defence Science and Technology Organisation



Concluding Remarks

The development of a value model is a non-trivial exercise; **the hierarchical relationship between values is a critical outcome of this process**.

This value model and any means-end network developed needs to be used to generate an BN which models the causal relationships between the nodes. The causal relationships are then quantified by specifying the impact of the effectiveness of predecessor nodes on consequent nodes.

Within the context of Cox's work the resulting BN captures the notion of how the effectiveness of a system influences the effectiveness of other systems.

MoE characterization

system type	Values Determined	Model Determined	MoE possible	Example
well defined interactions (causal relationships known)	yes	yes	yes	well known physics; known physical laws, for example: ballistic missile
undefined interactions	possibly (but not necessarily)	possibly	only at high level	NCW (now)
disruptive technology	no (wrong value structure from sustaining viewpoint))	partial	yes, but measures wrong attributes	Digital /film cameras. Steel mini-mills
sustaining technology	yes	yes	yes	improving radar technology
new approaches	no (no experience to determine values at low level)	no (lack of knowledge)	no (but maybe at higher level, borrowing value from comparable systems)	early stage of radar, or totally new surveillance technique. Information fusion.
evolving needs (assured technology)	partial	partial	no (partial even at higher level)	future car, where initial values are so achievable that they can effectively be ignored, but new needs are in a state of flux. (for example: NCW (now))