

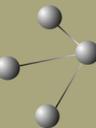
# Understanding the Situation Awareness of Pilots

Steven Mascaro\*, Kevin Korb\* and Ken McAnally#

(With much help from Alex Black and Thao Hoang Xuan)

\* Bayesian Intelligence

# Defence Science and Technology Organisation



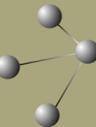
# Situation Awareness

What is situation awareness?

- Pilot's perception and model of their current environment

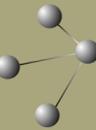
Why model it?

- Can improve pilot performance through:
  - Improved aircraft and cockpit design
  - Improved pilot training



# Our Situation

During flight, the pilot notices a nearby aircraft and uses context and evidence to assess whether it is **friendly**, **hostile** or neutral.



# Modelling Approach

Identify the Pilot's Model

or

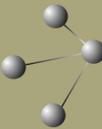
Model the Pilot's Behaviour



Better for expert elicitation

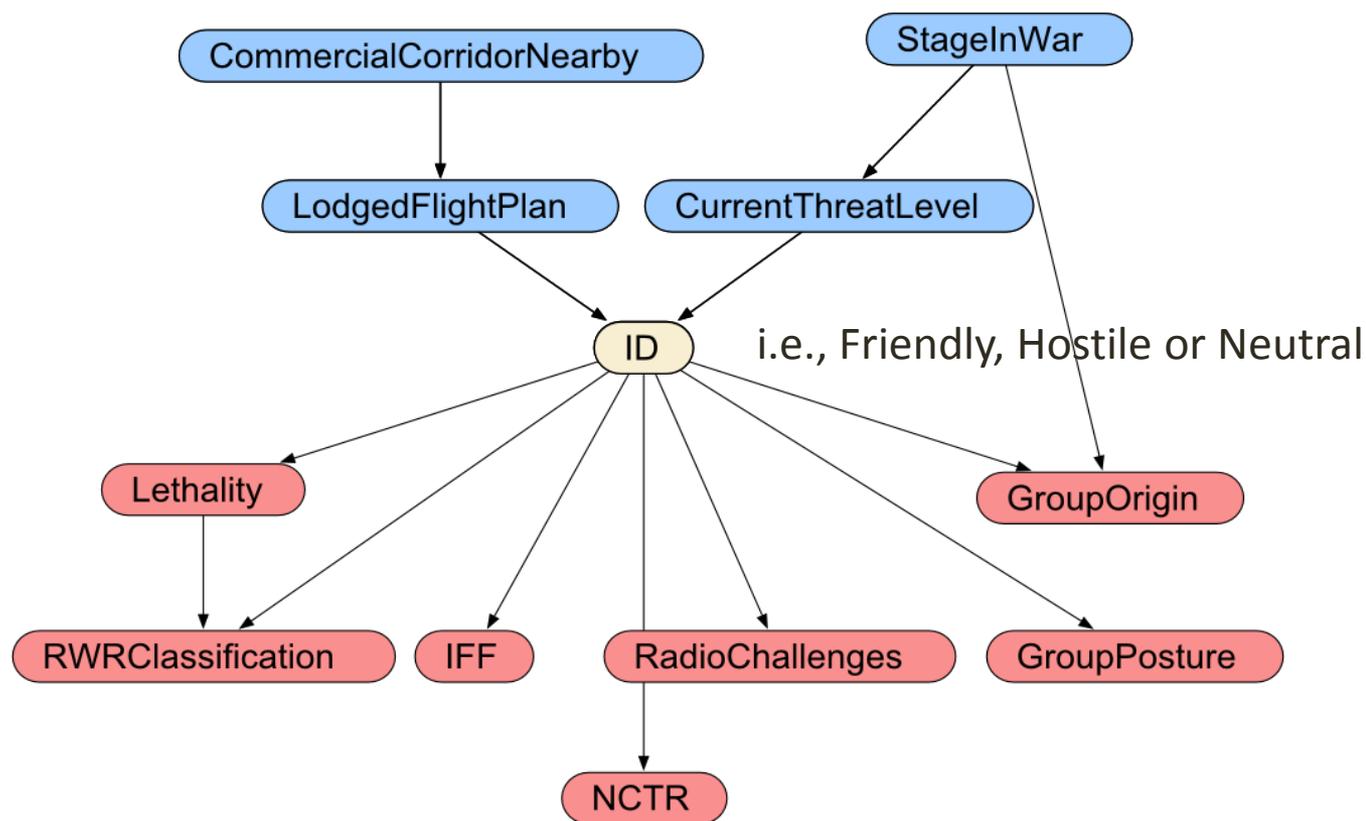


Better for machine learning & analysis



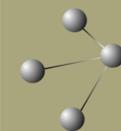
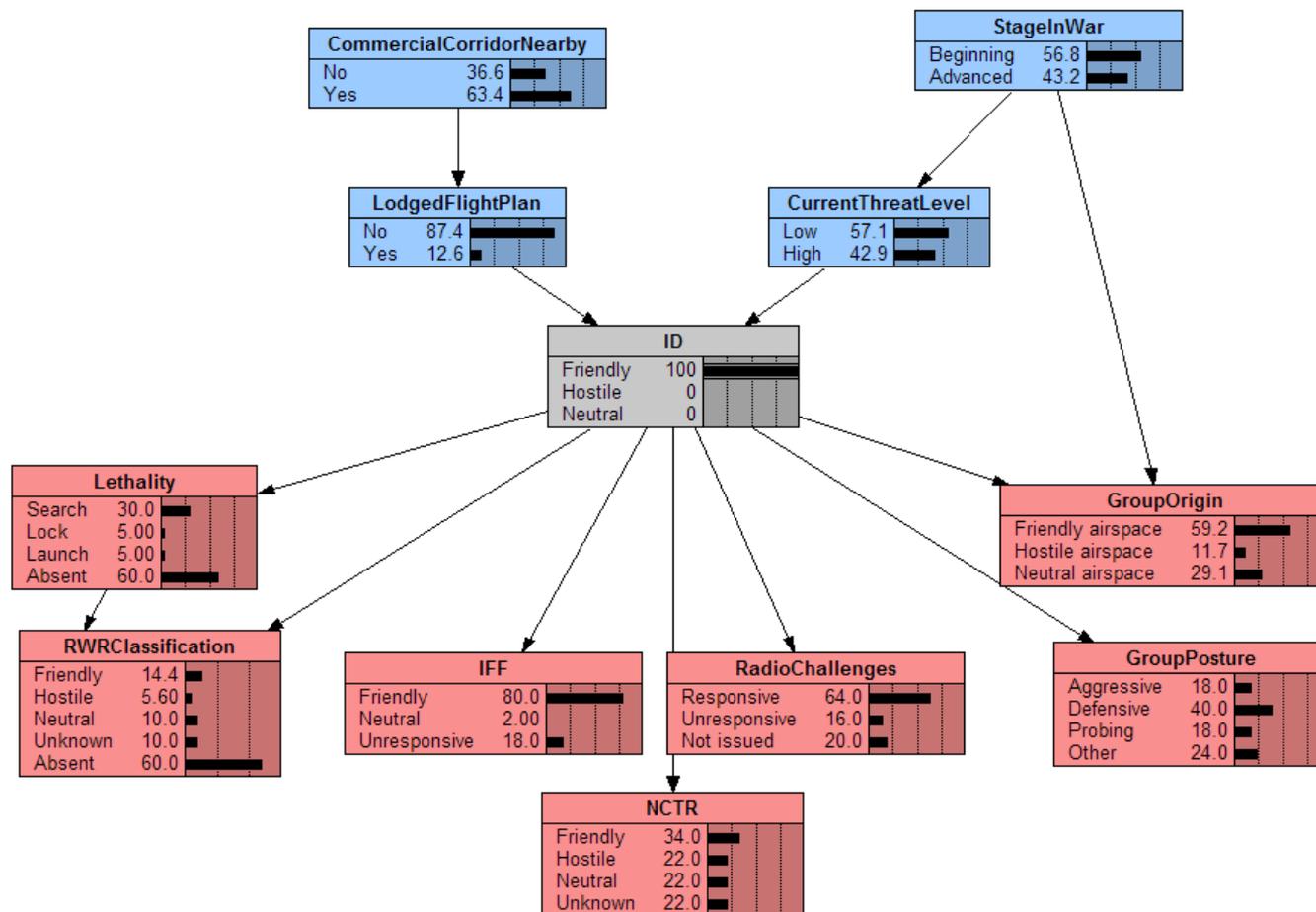
# Pilot's Model

- Expert-created structure



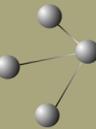
# Pilot's Model

- Elicited Parameters

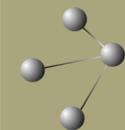
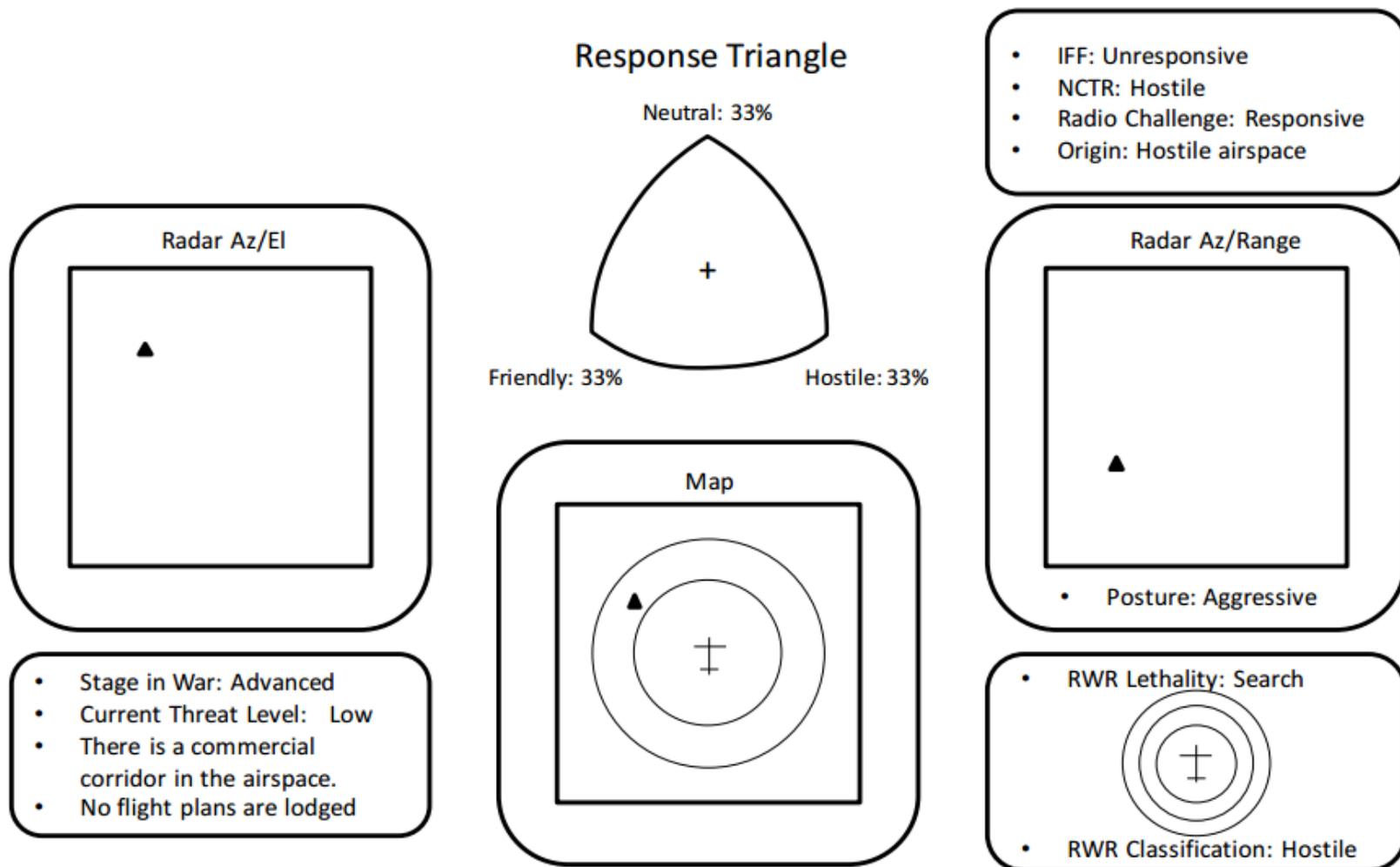


# Model of Pilot Behaviour

- Run experiments to collect data from pilots
- Use machine learning to analyse data and develop models
- Experiment:
  - Present pilots with context
  - **Pilots can then choose order to interrogate variables**
  - Pilots report assessment after each piece of evidence



# Collecting Data: Interface



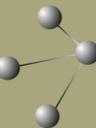
# Collecting Data

interrogation	CommercialCorridorNearby	StageInWar	LodgedFlightPlan	CurrentThreatLevel	ID
0	No	Beginning	No	Low	Neutral
1	No	Beginning	No	Low	Neutral
2	No	Beginning	No	Low	Neutral
3	No	Beginning	No	Low	Neutral
4	No	Beginning	No	Low	Neutral
5	No	Beginning	No	Low	Neutral
6	No	Beginning	No	Low	Neutral
7	No	Beginning	No	Low	Neutral

• • •

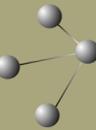
Lethality	RWRClassification	NCTR	IFF	RadioChallenges	GroupPosture	GroupOrigin	Neutral	Friendly	Hostile
*	*	*	*	*	*	*	0.25	0.38	0.36
*	*	*	*	*	Other	*	0.2	0.35	0.45
*	*	*	Neutral	*	Other	*	0.43	0.23	0.34
*	*	*	Neutral	*	Other	Neutral_airspace	0.62	0.19	0.19
*	Friendly	*	Neutral	*	Other	Neutral_airspace	0.45	0.33	0.22
*	Friendly	Neutral	Neutral	*	Other	Neutral_airspace	0.42	0.45	0.14
*	Friendly	Neutral	Neutral	Responsive	Other	Neutral_airspace	0.42	0.45	0.14
Search	Friendly	Neutral	Neutral	Responsive	Other	Neutral_airspace	0.55	0.32	0.13

• • •

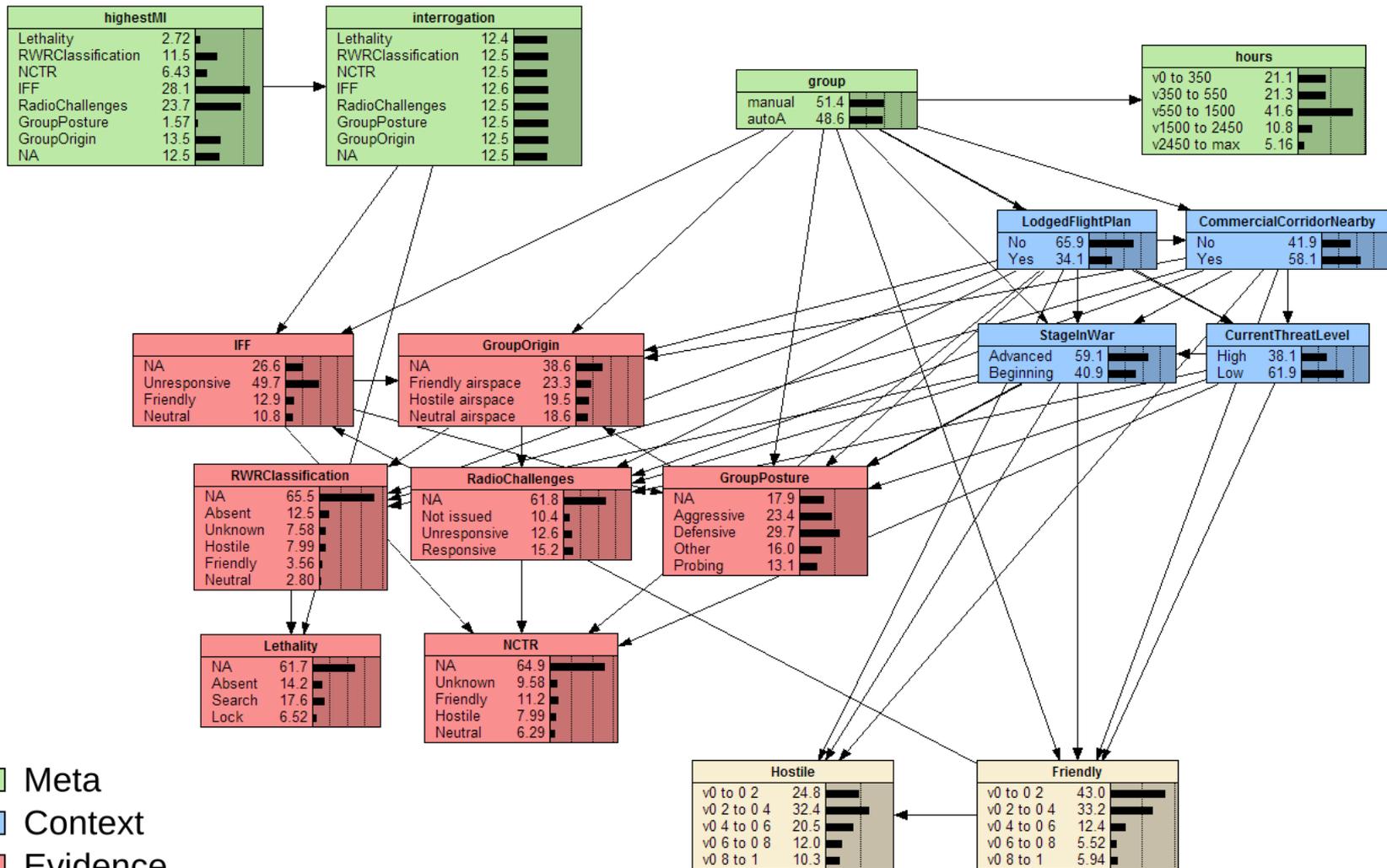


# Learning Models of Pilot Behaviour

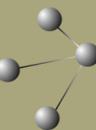
- Used CaMML to learn models from the data
- Several different models, including:
  - Unrestricted structure
  - Evidence variables last
  - ID variables last
  - DBN of choices
- For curiosity, also used data to parameterise pilot's model



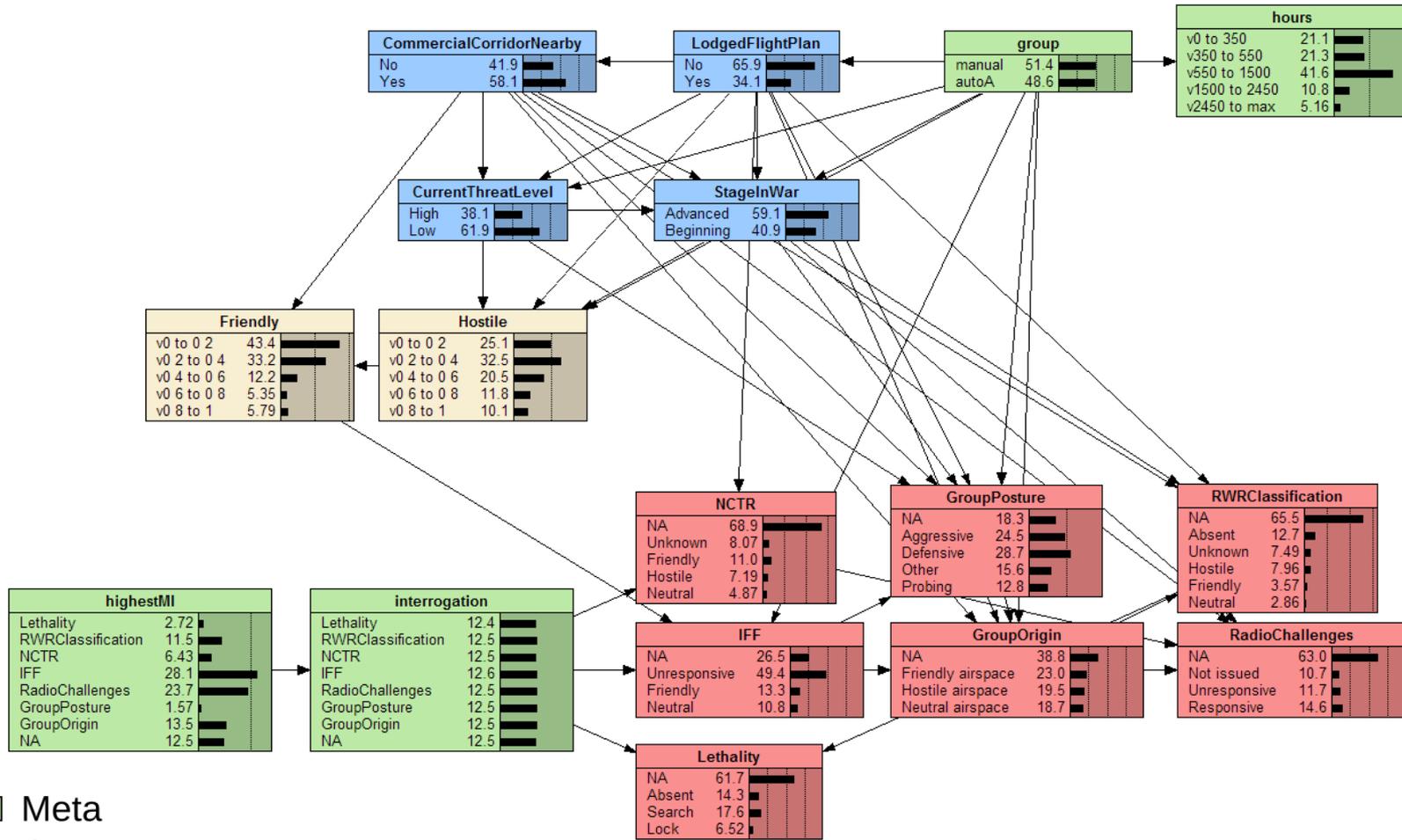
# Learning: Unrestricted Structure



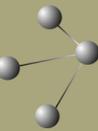
Includes priors to restrict arcs in and out of highestMI (as do all nets)



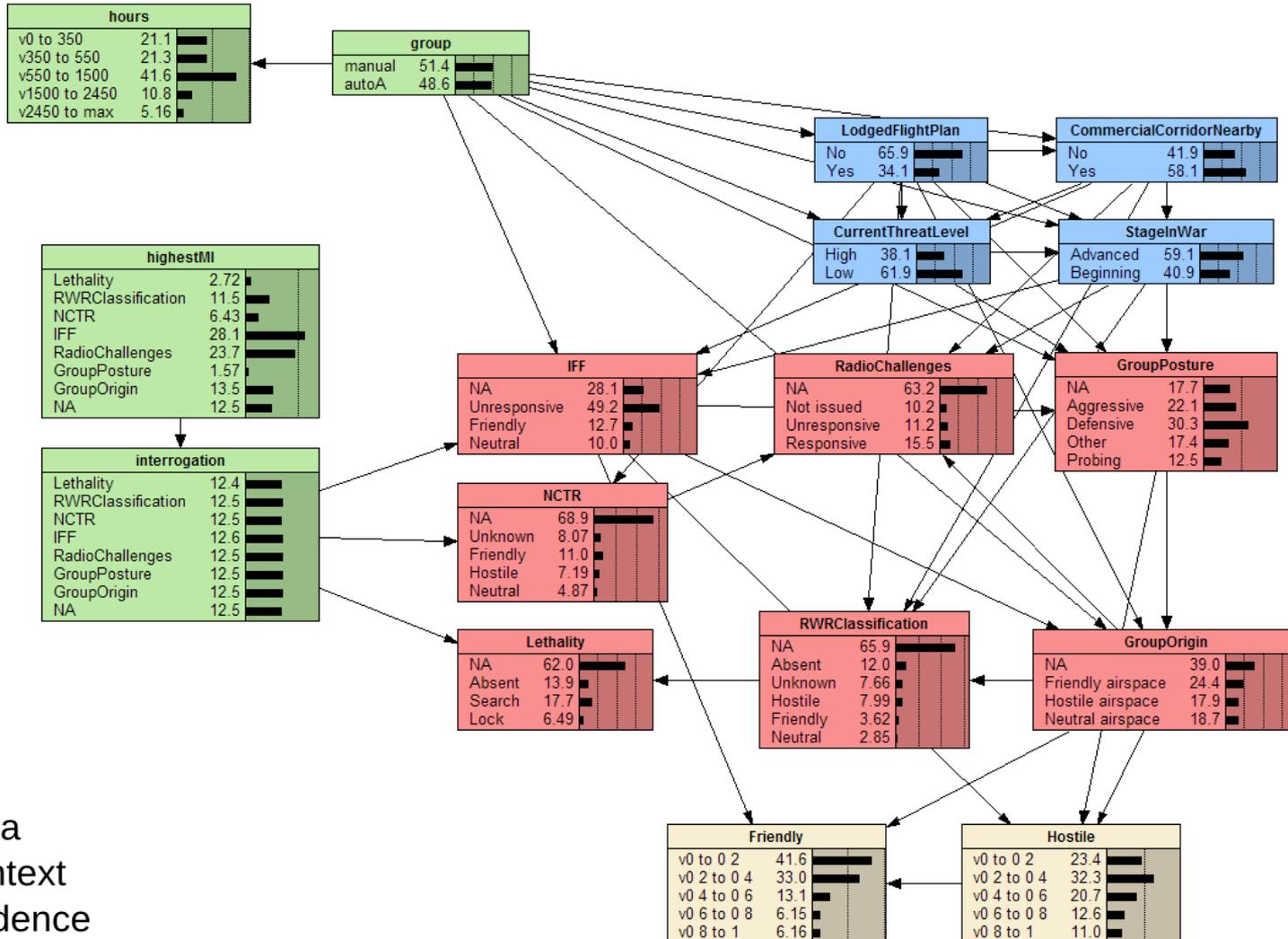
# Learning: Evidence Last



- Meta
- Context
- Evidence
- ID

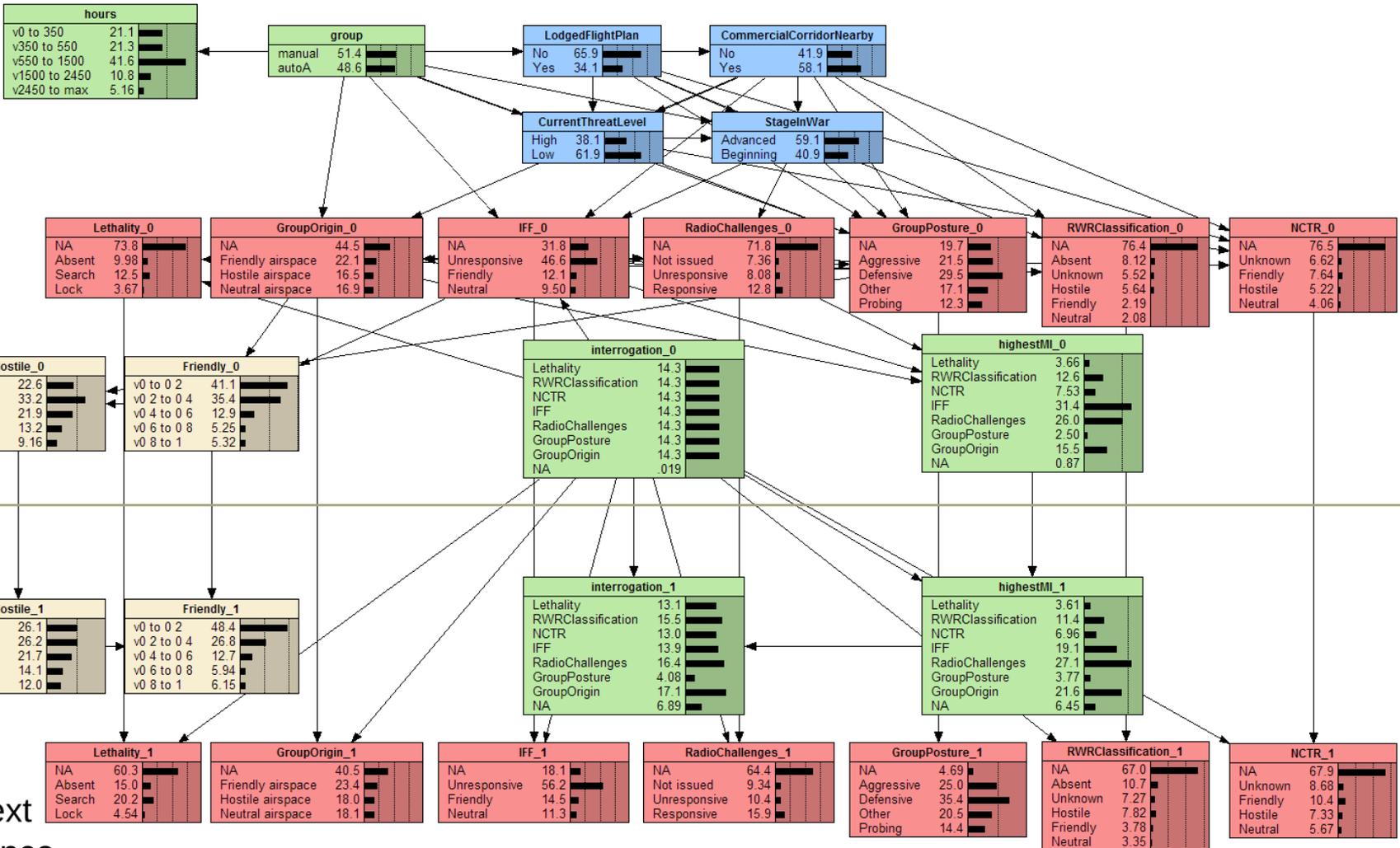


# Learning: ID Last



- Meta
- Context
- Evidence
- ID

# Learning: ID Last DBN

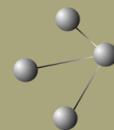


# Model Performance

- Unrestricted matches data best (as expected)
- ID Last comes close second

<b>Net</b>	<b>KL</b>	<b>RMSE</b>
Unrestricted	0.1361	0.1866
ID Last	0.1517	0.1989
Evidence Last	0.1849	0.2227
Expert Structu	0.1920	0.2128
Expert All	0.5893	0.3336

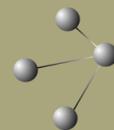
Smaller = Better



# Model Performance

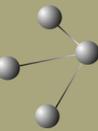
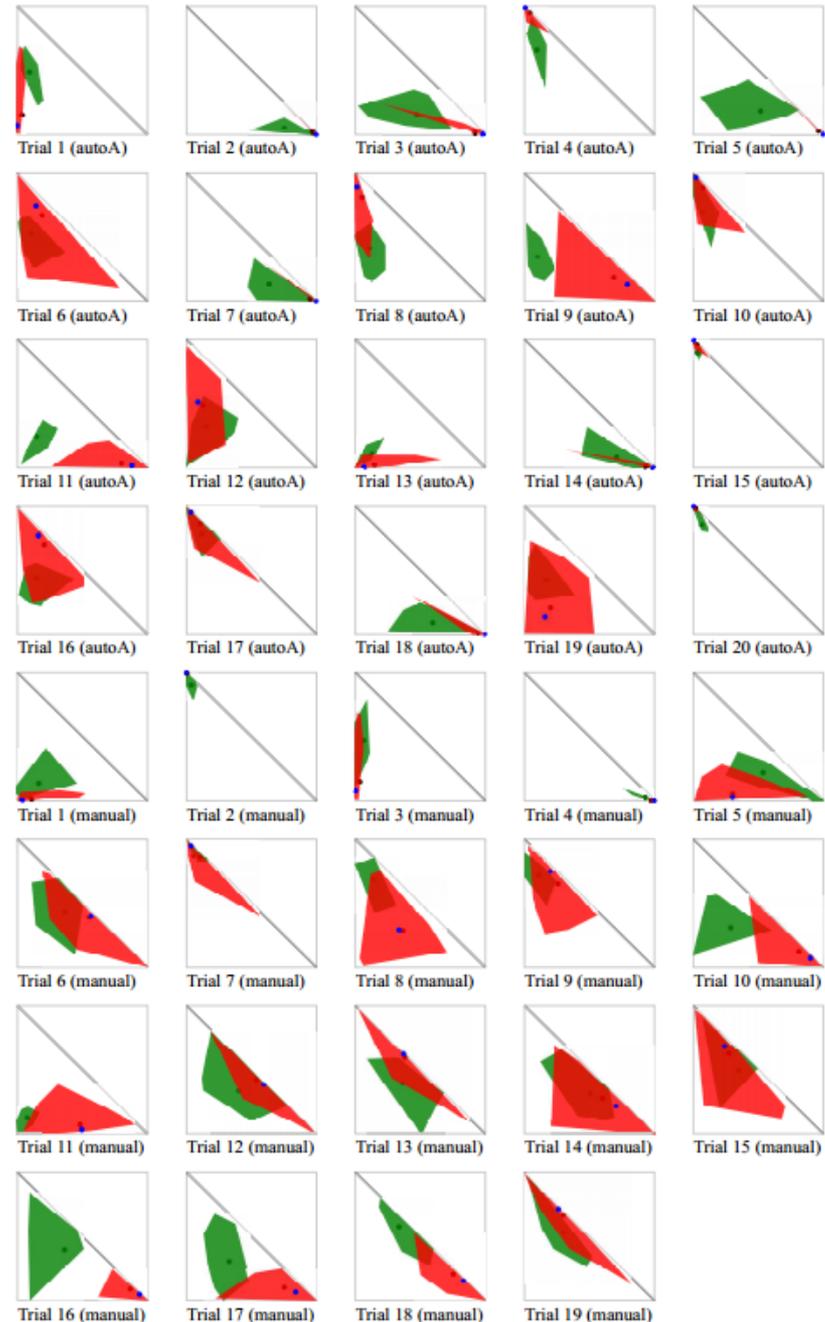
- KL divergence between pilots and network predictions increases as pilots progress through the trial (for all networks)

Step	Unrestricted	ID Last	Evidence Last
0	0.0502	0.0521	0.0494
1	0.0901	0.0947	0.1114
2	0.1131	0.1260	0.1583
3	0.1249	0.1472	0.1909
4	0.1492	0.1702	0.2178
5	0.1647	0.1858	0.2289
6	0.1823	0.2028	0.2474
7	0.2142	0.2345	0.2751
<b>Avg</b>	<b>0.1361</b>	<b>0.1517</b>	<b>0.1849</b>



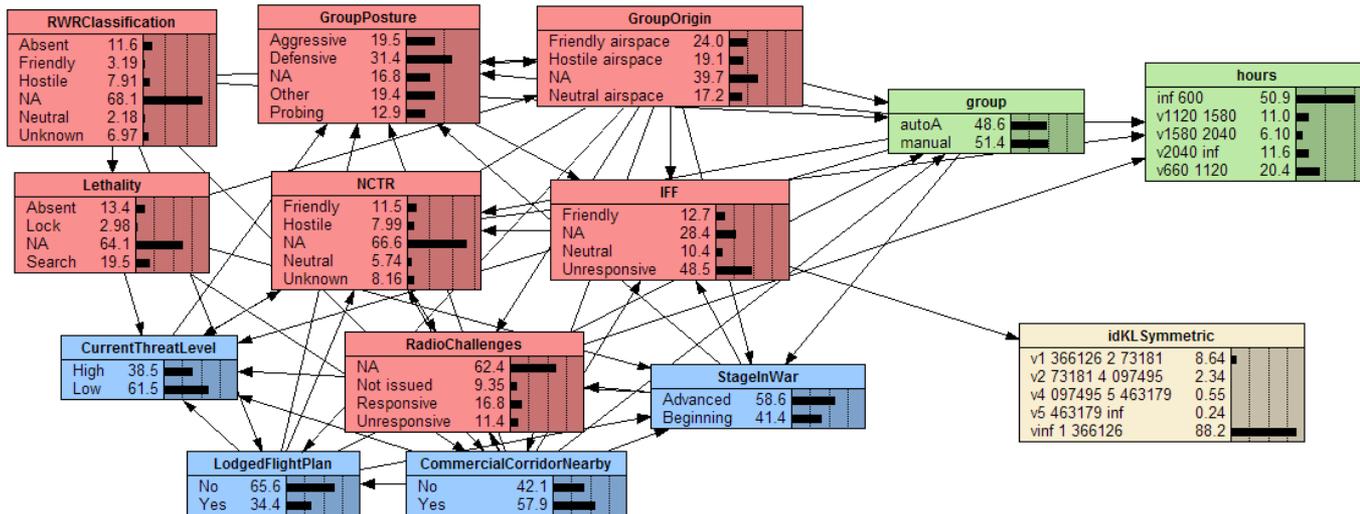
# Elicited Pilot's Model vs Pilot Responses

- Elicited model gets several trials very wrong



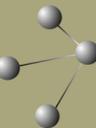
# What affects the ID Elicited score?

Learn a network to find out:



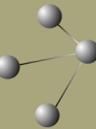
MI of variables to KL(Pilot,Elicited)

Node	MI
IFF	0.12042
GroupPosture	0.01551
GroupOrigin	0.01026
NCTR	0.00779
group	0.00488
RadioChallenges	0.00404
RWRClassification	0.00309
Lethality	0.00195
CommercialCorridorNe	0.00107
hours	0.00091
LodgedFlightPlan	0.00082
StagelnWar	0.00008
CurrentThreatLevel	0.00007



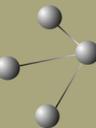
# Gaining Insight into Pilot Behaviour

- How do pilots approach the evidence?
- Which variables are most decisive?
- Do pilots choose in conformance with MI?
- Do pilots update in response to new information?



# In What Order Do Pilots Interrogate Variables?

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
<b>GroupPosture</b>	0.80239	0.11884	0.01174	0.04472	0.00712	0.00689	0.00723
<b>IFF</b>	0.11904	0.61265	0.18601	0.03724	0.03147	0.00769	0.00643
<b>GroupOrigin</b>	0.06342	0.18883	0.46023	0.09874	0.1506	0.01863	0.00865
<b>RadioChallenges</b>	0.00515	0.04603	0.05214	0.27144	0.15782	0.27034	0.18973
<b>Lethality</b>	0.0025	0.02031	0.07034	0.28052	0.17833	0.2853	0.15015
<b>RWRClassification</b>	0.0025	0.00445	0.04976	0.13061	0.37433	0.20728	0.22755
<b>NCTR</b>	0.0025	0.00445	0.16423	0.13017	0.09432	0.19698	0.40384



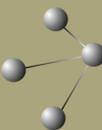
# Which Variables Influence Pilot Assessments the Most?

- CaMML regularly links Friendly/Hostile with IFF, and often with GroupPosture and GroupOrigin

Mutual Information values from the ID-Last network:

Hostile		Friendly	
Node	MI	Node	MI
GroupPosture	0.24456	IFF	0.23956
IFF	0.18034	GroupOrigin	0.20065
GroupOrigin	0.13638	GroupPosture	0.08756
StageInWar	0.04263	RWRClassification	0.03111
RWRClassification	0.02249	RadioChallenges	0.02146
RadioChallenges	0.01858	Lethality	0.01748
Lethality	0.01406	NCTR	0.00994
CurrentThreatLevel	0.01	CommercialCorridorNearby	0.00994
NCTR	0.009	CurrentThreatLevel	0.00682
CommercialCorridorNearby	0.00794	StageInWar	0.00131
LodgedFlightPlan	0.0079	LodgedFlightPlan	0.00036

Suggests pilots more likely to look at GroupPosture if they suspect hostile (and to some extent StageInWar)



# Do Pilots Use Mutual Information?

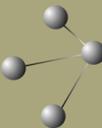
Interrogation Chosen

	Lethality	RWRClassificati	NCTR	IFF	RadioChallenge	GroupPosture	GroupOrigin
Lethality	0.89	0.01	0.08	0.01	0.01	0.01	0.01
RWRClassificatio	0.33	0.53	0.10	0.00	0.03	0.01	0.00
NCTR	0.04	0.04	0.90	0.00	0.00	0.00	0.00
IFF	0.02	0.02	0.00	0.44	0.01	0.39	0.11
RadioChallenges	0.19	0.20	0.14	0.00	0.41	0.00	0.06
GroupPosture	0.03	0.03	0.05	0.01	0.15	0.53	0.19
GroupOrigin	0.05	0.06	0.13	0.00	0.14	0.05	0.57

Highest MI

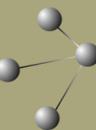
MI is calculated from the **ID\_Elicited** network. Interrogation distributions taken from ID last network.

Difficult to test because we have no real IDs.



# Do Pilots Choose the Highest MI at Each Step?

interrogationStep	highestMI	Lethality	RWRClassification	NCTR	IFF	RadioChallenge	GroupPosture	GroupOrigin	NA
Step 1	Lethality (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	RWRClassification (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	NCTR (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	IFF (98%)	0.00	0.00	0.00	0.12	0.00	0.81	0.06	0.00
Step 1	RadioChallenges (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	GroupPosture (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	GroupOrigin (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 1	NA (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 2	Lethality (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 2	RWRClassification (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 2	NCTR (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 2	IFF (86%)	0.02	0.00	0.00	0.70	0.00	0.07	0.21	0.00
Step 2	RadioChallenges (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 2	GroupPosture (4%)	0.03	0.03	0.03	0.03	0.42	0.36	0.08	0.03
Step 2	GroupOrigin (8%)	0.01	0.01	0.01	0.01	0.34	0.54	0.04	0.01
Step 2	NA (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 3	Lethality (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 3	RWRClassification (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 3	NCTR (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 3	IFF (26%)	0.06	0.10	0.01	0.70	0.08	0.00	0.03	0.00
Step 3	RadioChallenges (13%)	0.10	0.01	0.18	0.01	0.07	0.01	0.61	0.01
Step 3	GroupPosture (3%)	0.03	0.03	0.10	0.03	0.03	0.23	0.50	0.03
Step 3	GroupOrigin (56%)	0.07	0.03	0.23	0.00	0.04	0.00	0.62	0.00
Step 3	NA (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 4	Lethality (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 4	RWRClassification (5%)	0.37	0.02	0.28	0.02	0.15	0.11	0.02	0.02
Step 4	NCTR (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Step 4	IFF (7%)	0.22	0.18	0.02	0.45	0.05	0.02	0.05	0.02
Step 4	RadioChallenges (59%)	0.39	0.13	0.18	0.00	0.28	0.00	0.02	0.00
Step 4	GroupPosture (4%)	0.02	0.02	0.07	0.02	0.02	0.74	0.07	0.02
Step 4	GroupOrigin (23%)	0.06	0.16	0.02	0.01	0.39	0.01	0.35	0.01
Step 4	NA (0%)	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12





# Summary

- Expert elicited model did not match pilot behaviour too badly
  - Biggest difference: IFF
- Focussing on behaviour rather than guessing the pilot's model performed better for analysis
- Pilots focus on a few key pieces of information
- Pilots seem to be mostly rational
- Future work: more extensive experiments

