Adaptive Information Fusion in Asymmetric Sensemaking Environment

Paul Munya
Celestine Ntuen, Ph.D.

Center for Human Machine Studies
NC A&T State University

Outline

- Motivation
- Sensemaking in the Context of C2
- Illustration by Example
- Bayesian Abduction Model
  - Bayesian Probabilistic Reasoning
  - Peircean Abduction Reasoning
  - BAM
- Modeling Approach
- Simulation
- Conclusion

Motivation

- Asymmetric battlespace environments call for strategy rethink
  – Complex and “wicked” environment
  – Disparate information sources coupled with Uncertainty, ambiguity and dynamicity
- Deliberate MDMP is not sufficient
  – Linearity assumptions for non-linear asymmetric situations
- Generating COA must be progressive and opportunistic
  – Recalibration of the usual prescriptive-normative models of judgment and choice to fight unknown adversaries
- SENSEMAKING: Precursor to MDMP
  – “Connecting dots” to disparate information
  – Seeking explanations to unexpected evolving situations
  – Dynamic re-planning and re-tasking based on prospective and retrospective analysis

Sensemaking in the Context of C2

- How battle staff reduce uncertainty or ambiguity during decision making processes
- Aggregation of fragmentary battle space information (deriving meaning from fragmentary cues)
- Dynamic re-planning and re-tasking to account for the evolving asymmetric battlespace environments
- Aiding the commanders situation awareness by capturing the evolving states of battle dynamics, the information equivocality and the commander's intent.

Illustration by Example

A hypothetical scenario: Analyzing the Iraq insurgency

- The Battlestaff start out with various hypotheses regarding a perceived desired *end state* of an insurgent operation, H_o
- To achieve this *end state*, the insurgents have various *operational foci, h_i*. Examples of this could be funneling money and weapons to a particular cell, attacking soft targets to draw out coalition forces etcetera.
- For each *operational focus* there is a *motive X_i* or *motives* that avail themselves to the insurgents
- The *operational focus* and the *motivation* are uniquely effected by a pre-identified *influence pathway, S_i*. The influence pathway is a unique action or sets of actions that will be used to influence operations to achieve the desired *end state*

Illustration by Example

- In this case, an *influence pathway* could be the use of inflammatory religious sermons, political pressure-Al Sadr withdrawing from the unity government, arming militias, etcetera.
- For each of the unique *influence pathways* there is a specific set of *targets, m_i* to be attacked and *targeted actions* designed to collectively bring about the desired *end state* (Mosques, Bridges, Coalition Ops Bases, Kidnappings, etcetera)

  – Figure 1 illustrates this example

Illustration by Example

Figure 1

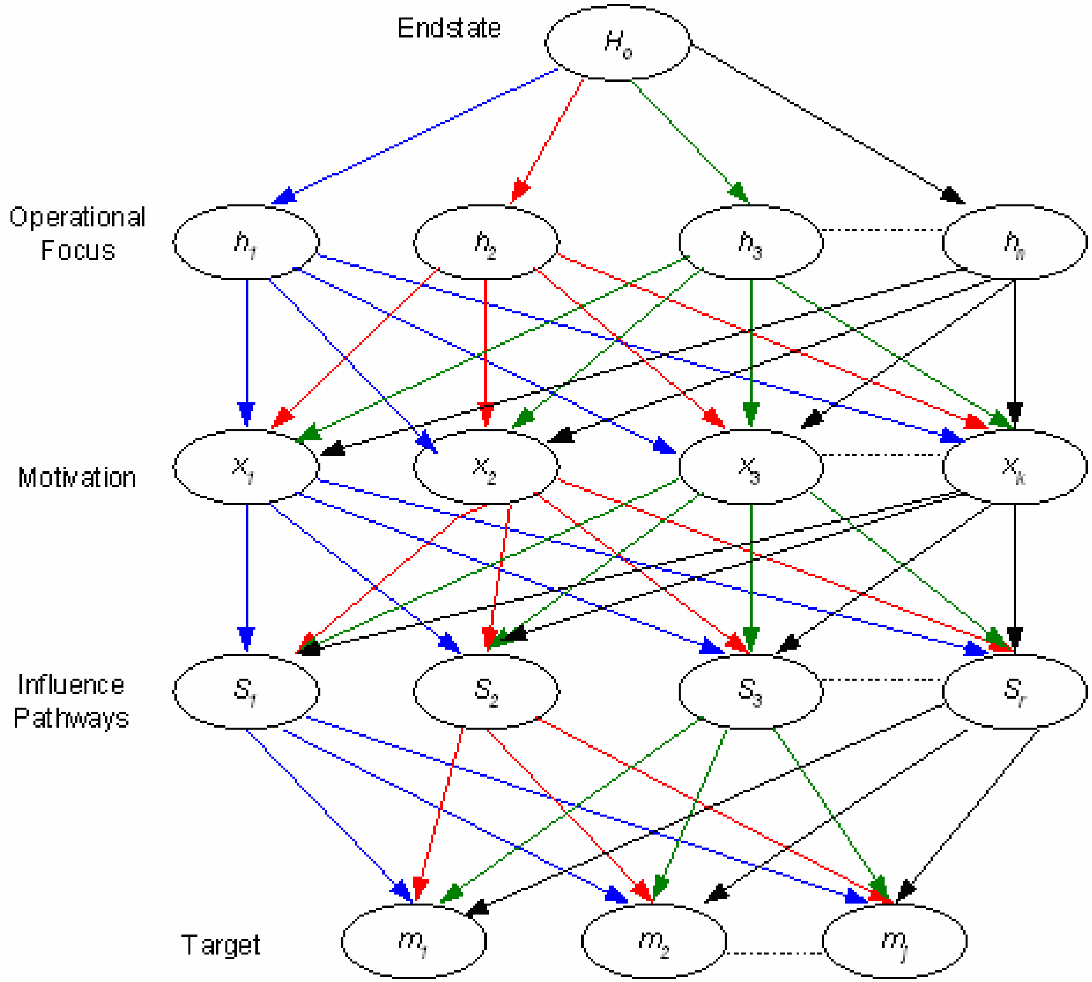


Illustration by Example

Construct a *network* to represent all the *variables* in the scenario.

Issues for analytical sensemaking:

- For a simple hypothetical scenario note the multiplicity of causal linkages!!
- Complexity increases with increasing variables; in real life battle space environments we expect a large number of variables and multiple linkages; We may not even be able to identify all of them; Some are interrelated, some are latent

Illustration by Example

Of interest for C2 sensemaking:

- What happens when new information arrives to the intelligent analyst?
- How does the network behave?
- What variables are affected?
- Are the effects serious enough to warrant immediate changes in the existing COA?

Illustration by Example

- Examples: The adversaries change their attack methods (armor penetrating IEDS);
- What is the most likely *target*?
- What is the *influencing* factor? (Sourced from Iran?);
- What is the likely change in *operational focus*?(From soft targets to armored coalition patrols).
- Does it represent an operational shift from low level attritional attacks to bold guerilla style hit and run tactics?
- If so, what *end state* does the adversary hope to achieve by focusing on these particular variables?

Bayesian Abduction Model

- The existing COA and planning models not flexible to handle the types of scenario described above
- We have proposed the Bayesian abduction model that combines sensemaking with Peircean abduction reasoning to model complex situations where information ambiguity, equivocality and dynamicity are dominant.
- Using this model, an intelligence analyst is able to fuse information from disparate sources in real time to identify variables and causal links of interest from the multiplicity of factors in the complex battlespace environment.
- The analyst can then use abductive reasoning to form plausible explanations for the situation of interest

Bayesian Abduction Model

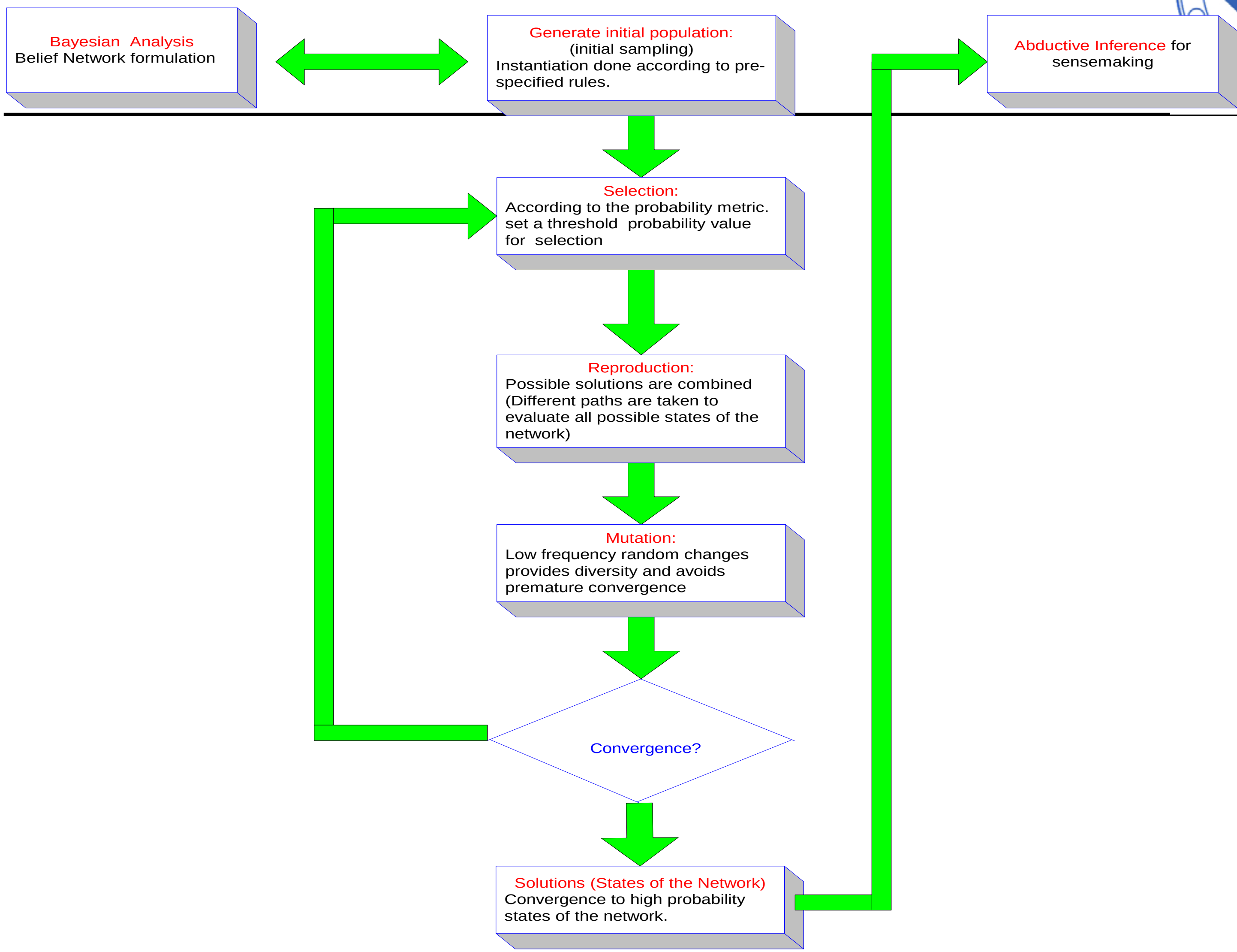
Features:

- Generates a list of exhaustive and mutually exclusive hypotheses regarding a scenario of interest.
- Represents all the variables of interest in the scenario as nodes to generate a belief network. Links from a parent node to a child node are causal links.
- Uses Bayesian analysis to evaluate all the possible states (solutions) for the network.
- Applies Peircean abduction reasoning to infer to the best explanation. (E is your collection of evidence; Hypothesis h_i explains E ;No other explanation explains E as well as h_i; therefore h_i is probably correct)

Bayesian Abduction Model

- Uses Genetic algorithm (GA) to perform fast and efficient search for plausible alternatives presented as possible states of the network
- The analyst makes a judgment call based on: How strong h_i as compared to other alternatives; independent of all h, how good is h_i ?How confident are you in the accuracy of E ?; How thorough is the search for other plausible alternatives ?.

Bayesian Abduction Model



North Carolina Agricultural & Technical State University

Bayesian Probabilistic Reasoning

Rationale:

- Intelligence analysts assign subjective conditional probabilities to variables of interest in order to analyze their impact in a given scenario.
- The conditional probabilities are based on the "belief state" of the analyst, not classical probability.
- The analyst starts of by assigning a conditional probability to hypothesis *h* *apriori* based on his/her expertise and knowledge. Upon obtaining some new evidence *D*, the apriori epistemic state *P* (state of knowledge) is revised by Bayes theorem into a conditional probability given by

Bayesian Probabilistic Reasoning

$$P(h \mid D) = \frac{P(D \mid h)P(h)}{P(D)}$$

- *P(h)* denotes the initial probability that hypothesis *h* holds ,before we incorporate any new data.
- *P(D)* denotes the probability that evidence data *D* will be observed. *P(D)* represents the probability of evidence *D* given no knowledge about which hypothesis holds.
- *P(D|h)* denote the probability of observing data *D* given some world in which hypothesis *h* holds
- We are interested in the probability *P(h|D)* that *h* holds given the observed data *D*

Peircean Abduction Reasoning

A process of reasoning that tries to form a plausible explanation for new and anomalous data.

- Classification of a given data set into potentially relevant elementary explanatory hypotheses.
- Given an observation *d* and the knowledge that *h* causes *d*, it is an abduction to hypothesize that *h* occurred.
- Given a proposition q and the knowledge that $p \rightarrow q$, it is an abduction to conclude p.
- Is inherently *uncertain* since information or data supporting abduction process is *dynamic* in nature, leading to human construction of multiple and often competing hypotheses.

Modeling Approach

- We have a certain problem space or world *P(w)* comprising of certain events of interest *P(E).*
  - *Let P(w) = ∑P(E)* where *E* is an explanation of world *W*
- Assuming independent events

$$P(E) = \prod_{h \in E} P(h)$$

$$P(w \mid E) = \frac{P(w \,\&\, E)}{P(E)}$$

- The Abduction process in sensemaking is: Given *E,* explain *E,* then try to infer *w* from these explanations
- Extend the model to account for uncertain information. An uncertain consequence corresponds to an event *E*, along with the probability α that *E* did not happen,

$$P(w \mid E, \alpha) = \alpha P(w \mid E) + (1-\alpha) P(w \mid \overline{E})$$

Modeling Approach

- In the case of a set of alternatives *E_i, $i= 1,2\ldots n$*, one of which is true, we extend the above equation thus

$$P(w\,|\,\{(E_i,\alpha_i)\}_{i=1\ldots n}) = \sum_{i=1..n} \alpha_i P(w\,|\,E_i)$$

- Formulate the problem as a belief network showing all the causal linkages together with the associated conditional probabilities.
- Once the state of the network is determined with all the instantiated variables determined, it is straightforward to perform backward or forward inference.
- Use a fast search algorithm such as the genetic algorithm (GA) to perform the search and computation for the most probable hypothesis-Abductive inference in belief networks is *NP-hard;* The more complex the network, the harder the computation.

Modeling Approach

- A Genetic algorithm is an adaptation procedure based on the mechanics of natural genetics and natural selection. GA's search from a population, not a single point and use randomized operators as opposed to deterministic rules.
- GA's can handle very complex network problems.
  - Perform fast and efficient computation over large search spaces
  - Inference is performed as a search in a large discrete multi-dimensional space
  - Adaptive search facilitates the discovery of network states with high probability instantiations
  - Represent multiple states for each variable depending on the cardinality we select for the genetic coding.

Simulation

Consider the hypothetical scenario previously described

- Code all the variables as a finite length string (in this case, cardinality 2 so that the set {0,1} is sufficient to represent all the states of the variables)
- At any instance, the state of the network is fully determined by a vector *a,* where

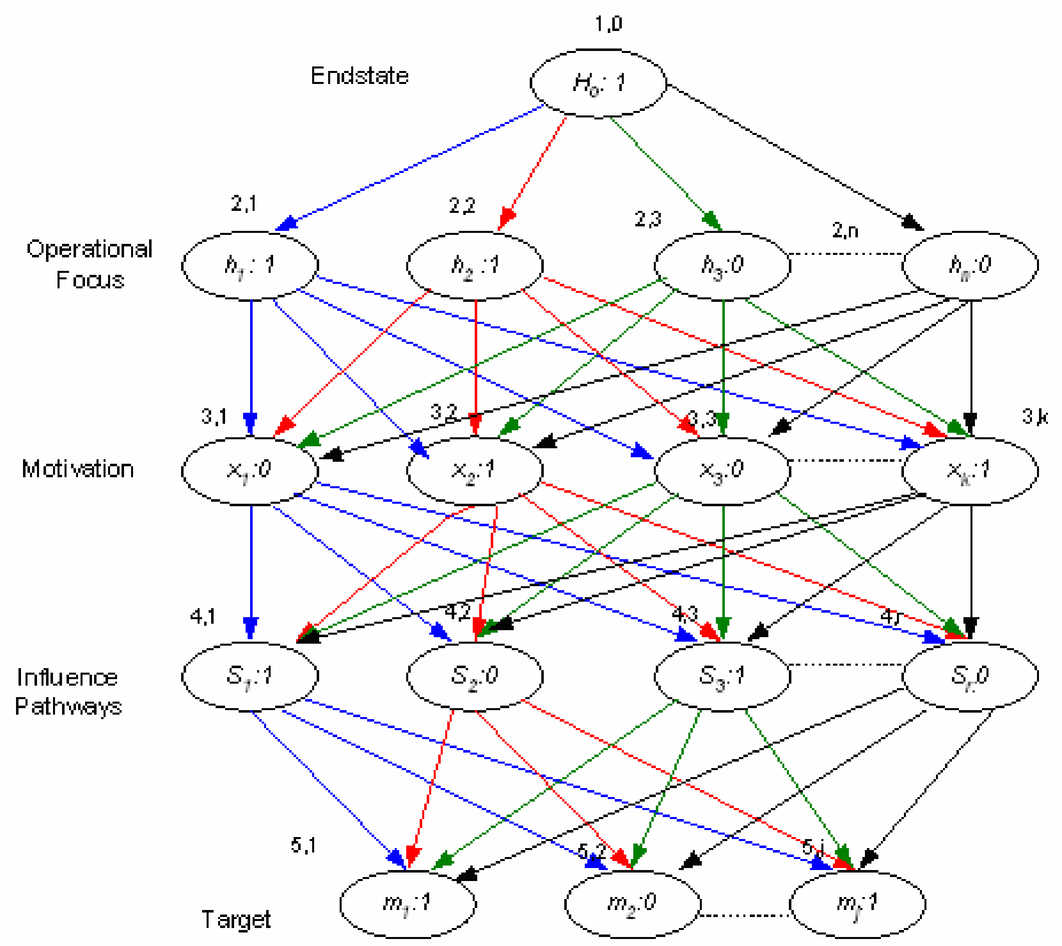
$$a = \begin{cases} 1 & \text{if a node } C_{kj} \text{ is instantiated} \\ 0 & \text{otherwise} \end{cases}$$

- The resulting network representation for all nodes is a binary pair $\{C_j, a\}$ for all nodes k.
- The initial population is generated by coding each of the variables with a {0,1} depending on the state of the instantiation

See Figure 2

Simulation

Figure 2



Simulation

- Subject the initial population to genetic operators {mutation, crossover, reproduction}
- The fitness function to determine propagation is calculated based on the defined Bayesian operators
- Start by assigning some apriori conditional probabilities such as

$$\boxed{P(H_o) = 0.4}$$

Implying we are only 40% confident that our chosen hypothesis regarding the end state is plausible.

- Similarly prior probabilities of all instantiated variables can be determined by straightforward application of Bayes theorem, for example

$$\boxed{P(m_1) = \sum_{S_1,..S_r} P(m_1 \mid S_1, S_2, S_3, ..S_r)}$$

North Carolina Agricultural & Technical State University

Sample Results

Array 1: $P(h_i | H_o)$

$h_i \mid H_o$	$H_o = 1$
$H_1 = h_1$	0.8
$H_2 = h_2$	0.5
$H_3 = h_3$	0.3
$H_4 = h_4$	0.9

Array 2: $P(X_i|h_i)$

$x_i \mid h_i$	$H_1 = h_1$	$H_2 = h_2$	$H_3 = h_3$	$H_4 = h_4$
$X_1 = x_1$	0.7	0.2	0.6	0.1
$X_2 = x_2$	0.3	0.4	0.5	0.8
$X_3 = x_3$	0.9	0.3	0.6	0.1
$X_4 = x_4$	0.1	0.9	0.7	0.5

Array 3: $P(S_i|X_i)$

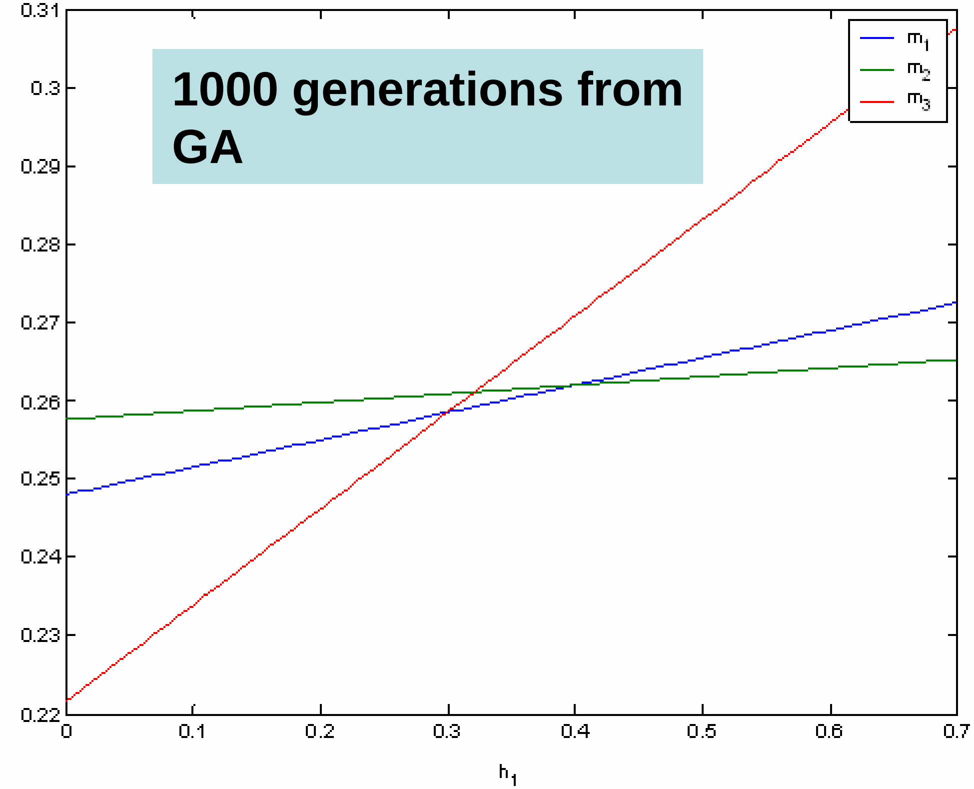
$S_i \mid x_i$	$X_1 = x_1$	$X_2 = x_2$	$X_3 = x_3$	$X_4 = x_4$
$S_1 = s_1$	0.5	0.6	0.9	0.3
$S_2 = s_2$	0.1	0.0	0.5	0.4
$S_3 = s_3$	0.9	0.1	0.3	0.5
$S_4 = s_4$	0.5	0.6	0.7	0.4

Array 4: $P(m_i|S_i)$

$m_i \mid S_i$	$S_1 = s_1$	$S_2 = s_2$	$S_3 = s_3$	$S_4 = s_4$
$M_1 = m_1$	0.6	0.3	0.8	0.1
$M_2 = m_2$	0.3	0.5	0.4	0.9
$M_3 = m_3$	0.1	0.9	0.2	0.6

Sample Results

Figure 3



Sample GA run. Variable h_1 is instantiated for different values and the resultant steady state probabilities of variable m_i are displayed.

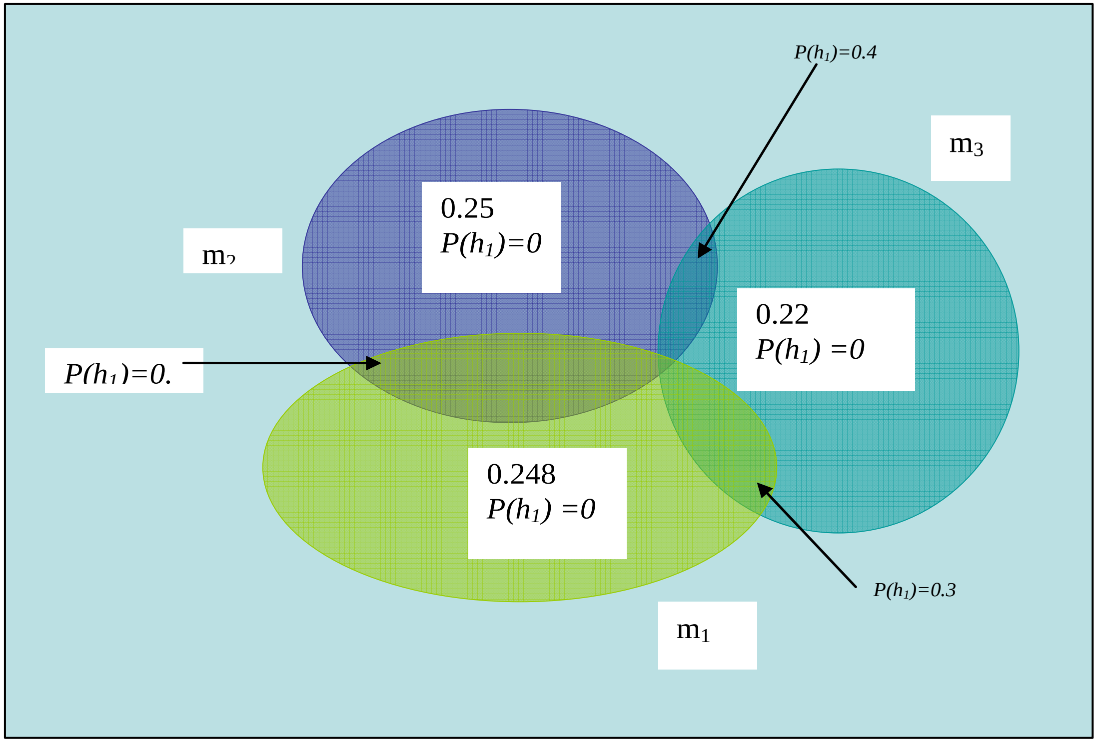
Results Discussion

- The graph shows how the most probable outcome varies as we manipulate the value of one variable h_1 .For example if the analyst believes there is a 70% chance that the *Operational Focus* of the adversary is node h_1 then there is a 30% chance that the targeted node is m_3.
- If 0% chance for node h_1,then the node with the highest probability of being targeted would be m_2 (26% chance).
- Notice also that with a 30% chance of occurrence for node h_1 both m_1 and m_3 are equally likely targets
- If the probability of h_1 occurring is increased to 0.4 then both m_1 and m_2 are equally likely targets. In this case, it is left to the analyst to look at other contributing factors before making inference

See Venn Diagram in Figure 4

Sample results

Figure 4



Solution space showing the feasible solutions for the sample problem

Conclusion

- This paper proposes an analytical sensemaking model to aid the C2 decision making process that combines Bayesian formalism with Peircean abduction reasoning .
- The Bayesian abduction model (BAM) has been implemented using GA .The developed model and algorithms will improve the design of sensemaking support systems for the Future Combat Force
- The aim of the modeling process is twofold: Foremost, retrospectively discovering or identifying variables or combinations therefore that can adequately explain observed adversary COA and secondly; Identifying variables and causal linkages that can aid in predicting an adversary's set of COA.
- The model provides an advantage to information fusion in a system characterized by dynamicity and complexity—evolving system states.

Questions??