

Automated Decision Support in a Complex Information Space

M.E. Senglaub, PhD

ICCRTS Symposium June 17-19, 2008
Bellevue, Wa

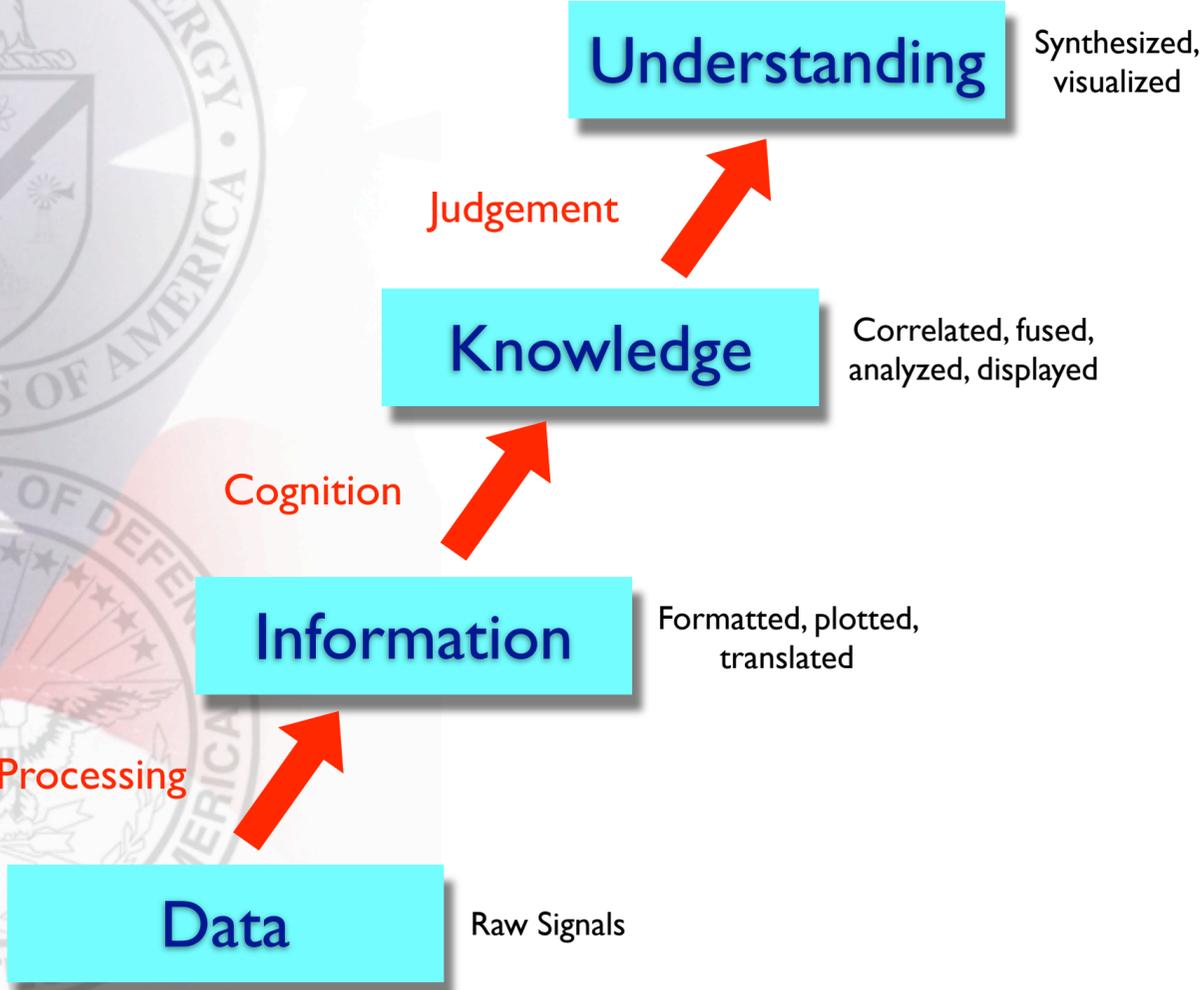
Briefing Structure

- Paper vs. Briefing Detail
- A Paradigm Shift
 - Information Models
 - Heuristics
 - Decision Support
- Foundations Needed to Understand Briefing Focus
 - Knowledge Representation
 - Peircean Philosophy
- Induction
- Abduction
- Applications to Date

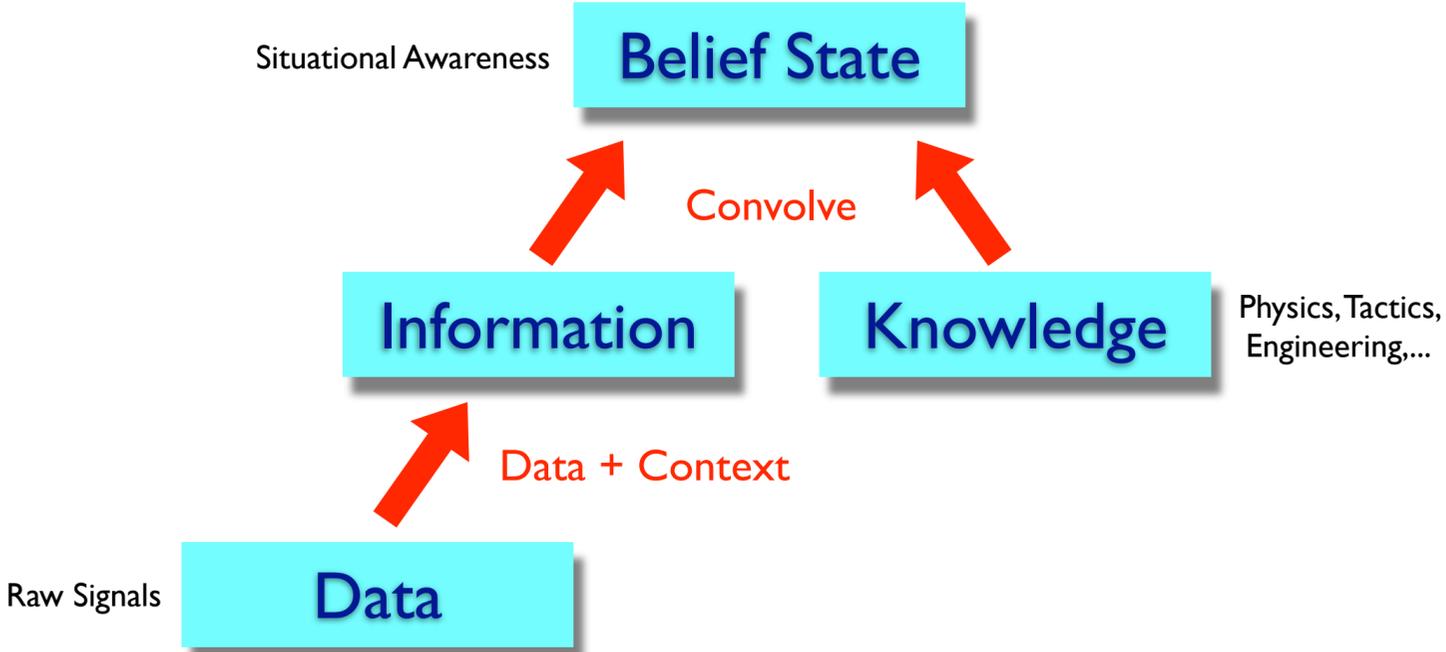
SAND2008-3168C

Information Models

Cognitive Hierarchy



Reasoning Hierarchy



Heuristics vs. Physics

Heuristic

Formulations based on observation

Heat Transfer

energy flow = $F(\text{temperature difference})$

Information Operations

reasoning = $F(\text{logic, philosophy, and mathematics})$

Physics

Algorithms based on physical theory

Statistical Physics

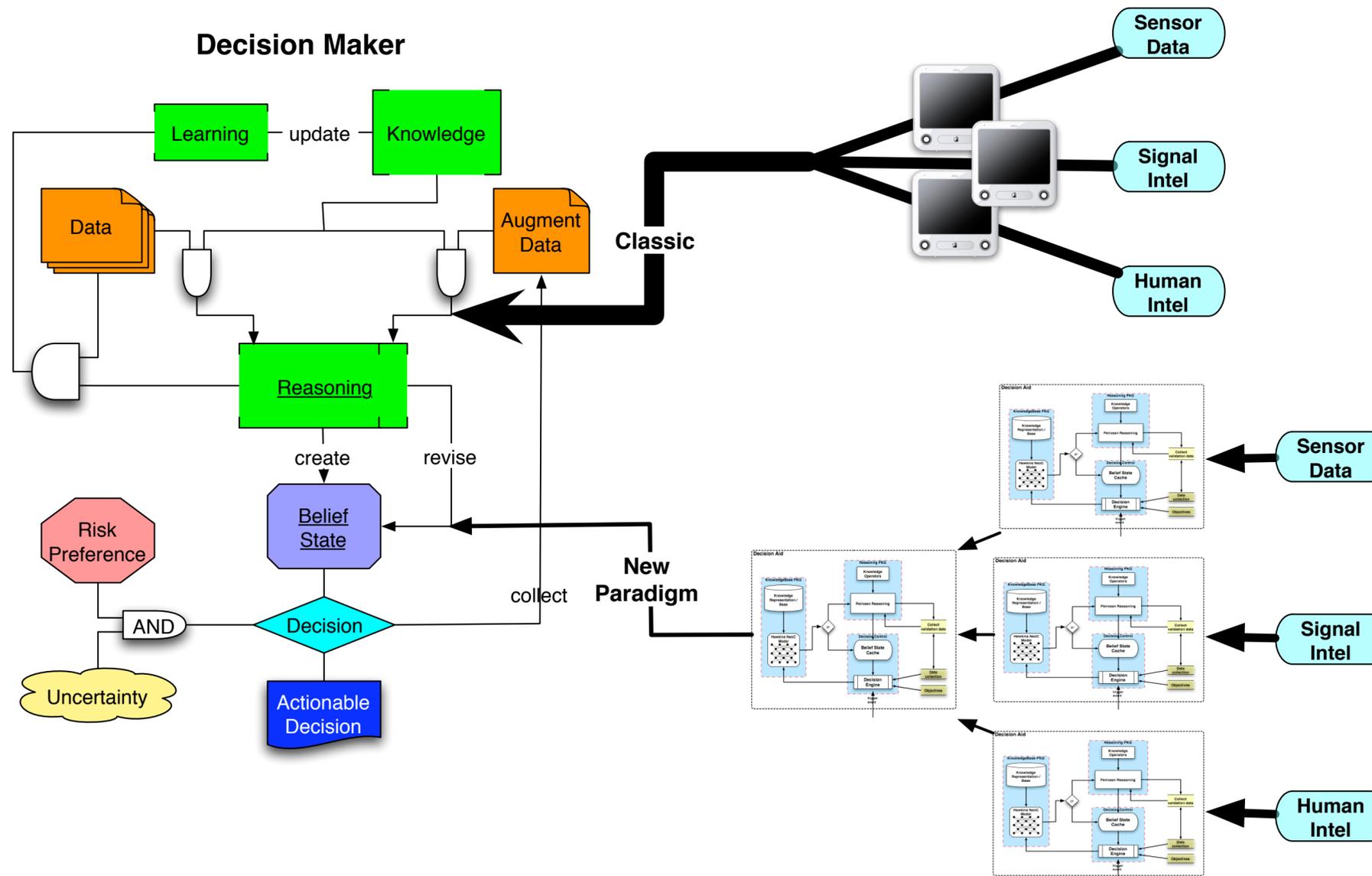
energy flow = $F(\text{particle microstates})$

Neuro-physiology

cognition = $F(\text{neural architecture/dynamics})$

SAND2008-3168C

Decision Making Paradigm



SAND2008-3168C

Knowledge Representation

SAND2008-3168C

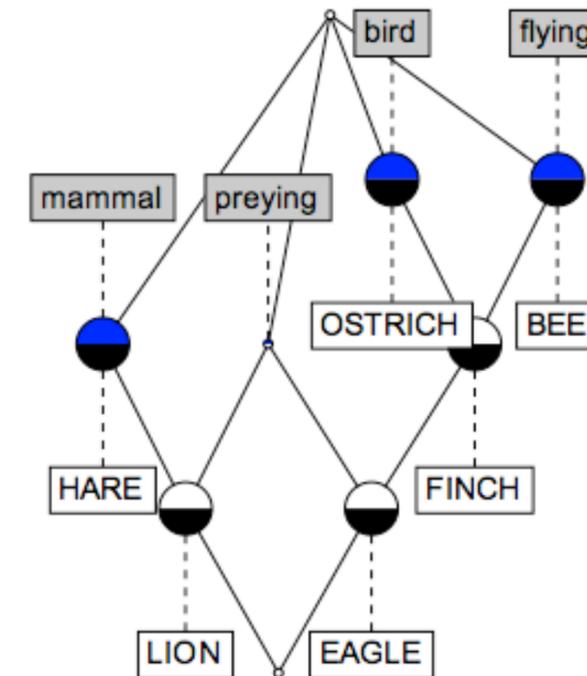
Formal Concept analysis (FCA)

$$K = \{ G, M, \bar{I} \}$$

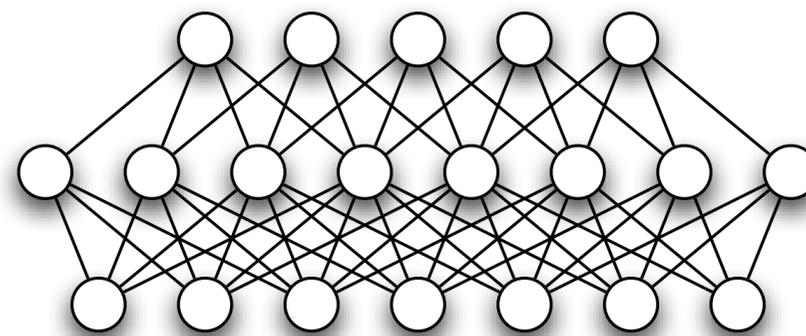
Cross Table

Animals	Preying	Flying	Bird	mammal
Lion	x			x
Finch		x	x	
Eagle	x	x	x	
Hare				x
Ostrich			x	
Bee		x		

FCA Lattice



Neural Processor



SAND2008-3168C

Attribute Output Layer

Feature Input Layer

Formal Concept Analysis* (Theory)

- **Ordered Set Theory**, a structure: $P = (P, \leq)$
 - A set P , and a binary relation \leq
 - An ordered set is reflexive, antisymmetric and transitive.
- **Formal Context**
 - A triple consisting of a set of objects, a set of attributes and a binary relation between the 2 sets.
- **The 'Prime' operator**
 - An operator that returns the set of attributes when applied to a set of objects or the set of objects when applied to a set of attributes.
- **Concepts**
 - A set of sets, the first term a set of objects, the second term the set of common attributes.
 - E.g. the concept (A, B) ; $A = \{\text{Earth, Mars}\}$, $B = \{\text{near sun, small, has-a-moon}\}$
- **'Begriff'**
 - The set of all concepts in a context.
 - Concept ordering based on subset-superset structure and a Galois connection.

$$\begin{aligned} (i) & a \leq a \\ (ii) & a \leq b \text{ and } b \leq a \Rightarrow a = b \\ (iii) & a \leq b \text{ and } b \leq c \Rightarrow a \leq c \end{aligned}$$

$$K_{FC} = (G, M, I)$$

$$\begin{aligned} (A)' & \equiv \{m \in M \mid (g, m) \in I, \forall g \in A\} \\ (B)' & \equiv \{g \in G \mid (g, m) \in I, \forall m \in B\} \end{aligned}$$

$$\begin{aligned} (A, B) & \xrightarrow{fc} (G, M, I) \\ & \Leftrightarrow \\ & A \subseteq G, B \subseteq M \\ (A)' & = B \ \& \ (B)' = A \end{aligned}$$

$$\begin{aligned} (A_1, B_1) & \leq (A_2, B_2) \\ & \Leftrightarrow \\ A_1 & \subseteq A_2 \ \vee \ B_2 \subseteq B_1 \end{aligned}$$

SAND2008-3168C

Peircean Philosophy

SAND2008-3168C

Peircean Philosophy

- C.S. Peirce; Sept 10, 1839 - April 19, 1914.
- Writings cover 1857 to 1914
 - 12000 printed pages, ~80000 handwritten pages.
- His work touches/impacts:
 - Mathematics and Philosophy, Phenomenology, the Normative Sciences, and Metaphysics
- Peircean Reasoning based on his 'Method of Scientific Inquiry'.
- Consists of 3 forms of logic:
 - Deduction
 - Induction
 - Abduction
 - Analogical reasoning
- Impact on information theory is beginning to be understood.

SAND2008-3168C



Induction

SAND2008-3168C

Induction

Philosophical Perspective

- Induction: *Is a mode of reasoning which adopts a conclusion as approximate.*
 - It is reasoning from a sample to the whole.
- Peirce Identifies Three Forms of Induction.
 - Weak form.
 - When sample and population sizes can be determined.
 - Strong form.
 - Examples where a single counter example can be identified.
 - Find one liberal that is not intellectually bankrupt.
 - 'Gradual' form.
 - Like the strong form but conclusions are updated with new samples.
- Induction Serves Three Functions.
 - Inductive Reasoning
 - Throw a ball in the air N times, observe it falling to the ground each time. If I throw it up the $(N + 1)$ st time, I expect it to fall to the ground again.
 - Inductive Learning
 - Draw a rule from observation and use that to predict. (Law of convective heat transfer)
 - Inductive Hypothesis Verification

SAND2008-3168C

Induction

Psychological Perspective*

- Heit's article discusses research on Induction in the psychological communities.
- There appears to be a minimalist understanding / deference to the work of the philosophical communities.
 - Basic definitions of Induction.
 - No reference to which role/function of Induction is being researched.
 - Flawed interpretation of Hume's philosophical dilemma associated with arguing the existence of Induction using inductive arguments.
 - Pragmatism provides a basis for arguing the existence of Induction.
- There may be(?) an attempt to explore the forms of Induction as defined by Peirce but it is unclear in the write-up.
 - Weak form, Strong form
 - The property dimension does not clarify Induction, rather it confounds results with prior knowledge.
- The analysis of the 'weak form' of induction fails to consider prior knowledge.
- No consideration of negative examples.
- The last issue is the lack of developing a research hypothesis and designing the experiment that would validate the hypothesis if true.
 - Regression analysis on the presumptive variables of the experiment adds no new insights into the theory or functionality of Induction

SAND2008-3168C

* E. Heit 'Properties of inductive reasoning'

Operational Considerations

- Three roles for Induction: Reasoning, Learning, and Hypothesis verification.
- Flinn's method addresses inductive reasoning.
 - Entire 'training' set is used to evaluate a new unknown.
- Peircean Decision Aid (PDA) defines an Inductive learning engine.
 - Based on Finn's Inductive reasoning algorithm.
 - Uses training set to develop a heuristic representation of the underlying knowledge.
 - Categorical knowledge of the heuristic is used to assess new unknown blocks of data.

SAND2008-3168C

Finn's Method

- Finn's method is based on Mill's canons
 - 'Method of agreement'
 - 'Method of differences'
 - 'Indirect method'
 - 'Method of concomitant variation'
 - 'Method of residues'
- The Finn algorithm is based on the first canon.
 - If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon.
- Algorithm is based on training sets of positive and negative examples associated with a goal attribute.
 - The examples provide a basis for defining positive & negative Begriffs.
 - The algorithm for determining a Begriff is presented in the next slide.

SAND2008-3168C

Construction of the 'Begriff'

- Begriff
 - Set of all concepts associated with a context.
 - Identify the intent of the Context
 - The set of all attributes.
 - Next construct the power set of the Intent, set S on the left.
 - Use the 'Prime' operator to identify the extent for each member of the power set.
 - Each member of the power set is a potential intent in the Begriff.
 - Result is a set of concepts for the context.
 - Will have multiple intents for each extent.
 - Construct a union of all intents for each unique extent.
 - Result is the final Begriff of a context.

$$S = \{Intent\}_{formal\ context}$$

$$P_s = \{p\} = \{\mathcal{P}(S)\}$$

$$\mathcal{B}^* = \sum_k C_k^*(\mathcal{E}_k, I_k) = \{\sum_k \{(p_k)'\}, \{p_k\}\}$$

in which

$$C_k^* = C_k(\mathcal{E}_k, I_k^1 \cup I_k^2 \cup I_k^3 \dots \cup I_k^n)$$

Finn's Method of Inductive Reasoning

- Based on J.S. Mill's first Canon
 - Method of Agreement
 - Procedural description
 - Identify or select a goal attribute, g .
 - E.g. the terrorist group responsible for an incident or the tactics used.
 - Evaluate an unknown against the positive and negative Begriff's to determine if it is a positive, negative or unresolved example.
 - If the intent of an unknown exists in 2 or more concepts of the positive context and does not appear in any concept of the negative context, the unknown can be classed as a positive example. Similarly for a negative classification.
 - g_8 is positive, g_9 negative, g_{10} indeterminate
- Two issues with Method:
 - Must use full sample set for each Unknown.
 - Heuristic of 2 examples does not belong in theory.

$$M = \{a, b, c, d, e\}$$

$$G_+ = \{g_1, g_2, g_3, g_4\}$$

$$G_- = \{g_5, g_6, g_7\}$$

Positive & negative examples

$$g_1^+ = \{a, b, c\}$$

$$g_2^+ = \{a, b, d\}$$

$$g_3^+ = \{a, b, e\}$$

$$g_4^+ = \{a, c, e\}$$

$$g_5^- = \{a, c, d\}$$

$$g_6^- = \{b, c, d\}$$

$$g_7^- = \{a, d, e\}$$

Unknowns

$$g_8^? = \{a, b, c, e\}$$

$$g_9^? = \{c, d, e\}$$

$$g_{10}^? = \{a, b, c, d\}$$

PDA Inductive Learning

- Based on Finn's Method
- Process description:
 - Accept entire training context.
 - Construct positive Begriffs based on the selected goal attribute.
 - To define the classification context for a goal 'value', no heuristics are built into the theory.
 - Biases are part of the engineering implementation and user 'set-able'.
 - Remove negative concepts from the positive Begriff of goal value k.
 - Using thresholds and 'likelihood' estimates we can bias the system in the direction of false positives or false negatives.
- Result is best estimate of attributes describing an object associated with a goal

$$\begin{array}{|l} B_k \equiv (A_k, B_k) | A_k \subseteq A; B_k = (A_k) \\ \text{with} \\ k = \text{goal attribute} \end{array}$$

$$B_k^* \equiv B_k - \sum_{m \neq k} B_m$$

$$lh_{k,m} = \text{SizeOf}(C_{m,E}) / \text{SizeOf}(B_k^*)$$

Abduction

SAND2008-3168C

Abduction

- Means by which new knowledge is created.
- Abduction begins with the observation of some 'surprising' fact.
 - Surprise implies your belief state is in doubt, rationality dictates that we want to eliminate doubt.
- Hypotheses are generated to explain the fact.
 - ...if theory A is correct then observation C would follow...
- Peirce identified a process/criteria for finding an optimal hypothesis.
 - Criteria involve explanation, verifiability, and economy.
 - Peirce's 'economy of research' involves the intrinsic value of the proposed hypothesis, and the effect on future research.

SAND2008-3168C

Abduction

(Hypotheses based on observations/data)

Knowledge + Data = Situational Understanding

Knowledge

ClassName	Type	2004	2005	2006	2007	2008	2009	2010	AssignedNorth	AssignedEast	AssignedSouth	SubmergedSpeed	SurfaceSpeed	SonarCM_Level	Sonar_Level	Noise	BatteryPoints	WeapRounds	FwdTubes	AftTubes	TorpedoRating	Mines	Hull_Numbers	SOF_Teams	Power
Type094	SSBN	0	0	0	1	1	1	2	0	0	0	27	18	CMLev_4	SonarLev_6	VeryQuiet	12	20	6_tubes	No_AftTubes	20	36	Hull_409_412	2_teams	NS
Type093	SSGN	1	1	2	2	3	3	4	1	0	0	30	18	CMLev_4	SonarLev_6	VeryQuiet	12	20	6_tubes	No_AftTubes	20	36	Hull_407_408_410_411_413_414	2_teams	NS
Xia092	SSBN	1	1	1	1	1	1	1	0	0	0	22	12	CMLev_3	SonarLev_5	Average	9	20	6_tubes	No_AftTubes	20	36	Hull_406	2_teams	NS
Han091	SSN	5	5	5	5	5	5	5	1	0	0	26	12	CMLev_2	SonarLev_5	Average	9	20	6_tubes	No_AftTubes	20	36	Hull_401_402_403_404_405	2_teams	NS
KiloII(AIP)	SS	0	0	0	0	0	0	0	0	0	0	17	12	CMLev_4	SonarLev_6	VeryQuiet	120	18	6_tubes	No_AftTubes	20	36	Hull_381_382_383_384_385_386	1_team	DES
KiloII(636)	SS	2	2	6	10	10	10	10	0	1	0	17	12	CMLev_4	SonarLev_6	VeryQuiet	24	18	6_tubes	No_AftTubes	20	24	Hull_366_367_368_369_370_371_372_373_374_375	1_team	DES
KiloII(877)	SS	2	2	2	2	2	2	2	0	1	0	17	12	CMLev_3	SonarLev_5	Quiet	24	18	6_tubes	No_AftTubes	20	24	Hull_364_365	1_team	DES
SongII(039C)	SSG	6	7	7	9	9	11	11	0	1	1	22	15	CMLev_4	SonarLev_6	VeryQuiet	180	12	6_tubes	No_AftTubes	20	24	Hull_321_322_323_324_314_315_316_317_318_319_325	1_team	DES
SongII(039)	SSG	1	1	1	1	1	1	1	0	1	0	22	15	CMLev_4	SonarLev_5	Quiet	36	12	6_tubes	No_AftTubes	20	24	Hull_320	1_team	DES
MingIII(035AIP)	SS	6	6	6	6	6	6	6	0	0	1	18	15	CMLev_4	SonarLev_6	VeryQuiet	120	18	6_tubes	2_AftTubes	20	326	Hull_361_305_306_307_308_309	1_team	DES
MingIII(035C)	SS	12	12	12	12	12	12	12	1	0	0	18	15	CMLev_3	SonarLev_5	Quiet	24	16	6_tubes	2_AftTubes	11	326	Hull_342_352_353_354_355_356_357_358_359_360_362_363	1_team	DES
MingIII(035)	SS	2	0	0	0	0	0	0	0	0	0	18	15	CMLev_2	SonarLev_4	Quiet	24	12	6_tubes	2_AftTubes	11	24	Hull_233	1_team	DES
RomeoII(033G)	SSG	1	0	0	0	0	0	0	0	0	0	12	15	CMLev_2	SonarLev_4	Quiet	39	14	6_tubes	2_AftTubes	11	28	Hull_351	1_team	DES
RomeoII(033)	SS	21	18	12	6	6	6	6	1	1	1	12	15	CMLev_3	SonarLev_5	Quiet	39	14	6_tubes	2_AftTubes	11	28	Hull_293_294_295_296_297_298_299_300_301_302_303_304_343 to 349_3	1_team	DES
ReserveRomeo	SS	15	15	15	15	12	9	6	1	1	1	12	15	CMLev_2	SonarLev_4	Quiet	39	14	6_tubes	2_AftTubes	11	28	Hull_268_269_270_271_272_275_276_277_278_279_280_286_287_291_292	1_team	DES
Moth.Romeo	SS	12	12	12	12	12	12	12	0	0	1	12	15	CMLev_2	SonarLev_3	Quiet	39	14	6_tubes	2_AftTubes	11	28	Hull_249_250_251_252_253_254_255_256_257_258_259_260	1_team	DES
Gulf(Auxiliary)	SSBA	1	1	1	1	1	1	1	0	0	0	15	15	CMLev_2	SonarLev_3	Average	14	265	6_tubes	4_AftTubes	11	44	Hull_200	1_team	DES
Sang-00xpt	SSM	1	1	1	1	1	1	1	0	0	1	8	8	CMLev_2	SonarLev_4	VeryQuiet	4	0	No_tubes	No_AftTubes	0	8	Hull_351	1_team	DES

Data

```

<UNK_root>
<Control input="xtable" > </Control>
  <Intel_Report>
    <Object> Initial_finding <Class type="generic">
      <Attr> noiseLevel <value type="label"> VeryQuiet</value></Attr>
      <Attr> speed <value type="label"> SubmergedSpeed.Q4 </value></Attr>
      <Attr> areaOfOps <value type="label"> AssignedSouth </value></Attr>
    </Class> </Object>
  </Intel_Report>
</UNK_root>

```

Hypotheses

2/15/2008 @ 0:0:00 <=> hypthesis <=> Initial_finding ^ MingIII(035AIP)

Support of Hypothesis:

[AssignedSouth, SubmergedSpeed.Q4, VeryQuiet]

Implicit data based on hypothesis:

[1_team, 2_AftTubes, 6_tubes, BatteryPoints.Q2, BatteryPoints.Q3, BatteryPoints.Q4, BatteryPoints.Q5, CMLev_4, Hull_361_305_306_307_308_309, Mines.Q3, Mines.Q4, Mines.Q5, SS, SonarLev_6,

SubmergedSpeed.Q2, SubmergedSpeed.Q3, SurfaceSpeed.Q2, SurfaceSpeed.Q3, SurfaceSpeed.Q4, SurfaceSpeed.Q5, WeapRounds.Q1, WeapRounds.Q2, WeapRounds.Q3]

Un-resolved collected data:

[]

2/15/2008 @ 0:0:00 <=> hypthesis <=> Initial_finding ^ SongII(039G)

Support of Hypothesis:

[AssignedSouth, SubmergedSpeed.Q4, VeryQuiet]

Implicit data based on hypothesis:

[1_team, 6_tubes, AssignedEast, BatteryPoints.Q3, BatteryPoints.Q4, BatteryPoints.Q5, CMLev_4, Hull_321_322_323_324_314_315_316_317_318_319_325, Mines.Q1, Mines.Q2, Mines.Q3, No_AftTubes,

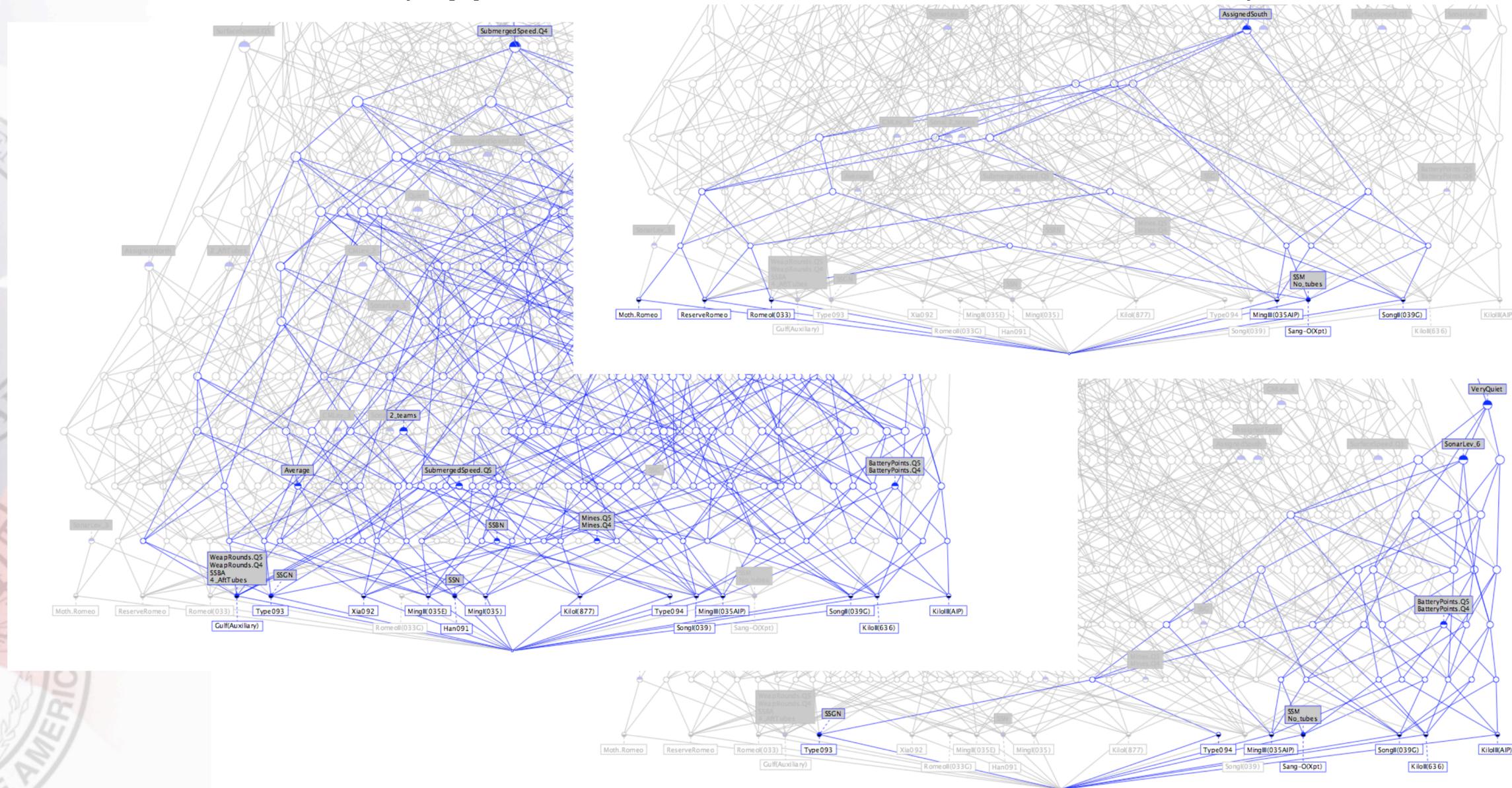
SSG, SonarLev_6, SubmergedSpeed.Q2, SubmergedSpeed.Q3, SurfaceSpeed.Q2, SurfaceSpeed.Q3, SurfaceSpeed.Q4, SurfaceSpeed.Q5, WeapRounds.Q1, WeapRounds.Q2, WeapRounds.Q3]

Un-resolved collected data:

SAND2008-3168C

Abduction

(Hypotheses based on observations/data)



Set Intersection of concepts associated with { VeryQuiet }, { AssignedSouth }, { SubmergedSpeed.Q4 }

$$\begin{aligned} & \{ \text{Gulf(Auxiliary)}, \text{Type093}, \text{Xiao92}, \text{MingII(035E)}, \text{Han091}, \text{MingI(035)}, \text{KiloI(877)}, \text{Type094}, \text{SongI(039)}, \text{MingIII(035AIP)}, \text{SongII(039G)}, \text{KiloII(636)}, \text{KiloIII(AIP)} \} \\ & \cap \{ \text{Moth.Romeo}, \text{ReserveRomeo}, \text{RomeoI(033)}, \text{MingIII(035AIP)}, \text{Sang-O(Xpt)}, \text{SongII(039G)} \} \\ & \cap \{ \text{Type093}, \text{Type094}, \text{MingIII(035AIP)}, \text{Sang-O(Xpt)}, \text{SongII(039G)}, \text{KiloII(636)}, \text{KiloIII(AIP)} \} \\ & = \{ \text{MingIII(035AIP)}, \text{SongII(039G)} \} \end{aligned}$$

SAND2008-3168C

Applications Intel Analysis

SAND2008-3168C

Convolve Information with Knowledge

- Intelligence collected:
 - Loss/Theft of explosives
 - North American AO
 - Link of explosives & bomb making

```
UNKsoc2.txt
<UNK_root>
<Control input="xtable" > </Control>
  <Intel_Report>
    <Object> material <Class type="generic">
      <Attr> explosive <value type="real"> 200 </value></Attr>
    </Class> </Object>
    <Object> speculation_capabilities <Class type="generic">
      <Attr> Tactic <value type="label"> Bombing </value></Attr>
      <Attr> Location <value type="label">North_America</value></Attr>
      <Attr> skill <value type="label"> bomb_making </value></Attr>
      <Attr> Tactic <value type="label"> explosives </value></Attr>
    </Class> </Object>
  </Intel_Report>
</UNK_root>
```

- Single hypothesis:
 - Group is likely to be HAMAS

```
InfoBase.BeliefCache
publishBeliefCache PUBLISH:
8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ HAMAS
Support of Hypothesis:
  [Bombing, North_America]
Implicit data based on hypothesis:
  [ArmeAssault, Killed, Middle_East, Military, Wounded, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  [bomb_making, explosives]
Publish Belief Cache Complete
```

Alternate / New Information Hypotheses

Collected

```
<UNK_root>
<Control input="xtable" > </Control>
<Intel_Report>
  <Object> material <Class type="generic">
    <Attr> explosive <value type="real"> 200 </value></Attr>
  </Class> </Object>
  <Object> speculation_capabilities <Class type="generic">
    <Attr> Tactic <value type="label"> Bombing </value></Attr>
    <Attr> Location <value type="label"> North_America </value></Attr>
    <Attr> skill <value type="label"> bomb_making </value></Attr>
    <Attr> Tactic <value type="label"> explosives </value></Attr>
  </Class> </Object>
  <Object> speculation_capabilities <Class type="generic">
    <Attr> Tactic <value type="label"> ArmeAssault </value></Attr>
  </Class> </Object>
</Intel_Report>
</UNK_root>
```



```
InfoBase.BeliefCache

8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ ArmelIslamic_Group
Support of Hypothesis:
  [ArmeAssault]
Implicit data based on hypothesis:
  [Africa, Civilian, Kidnap, Killed, Shooting, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  []

8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ HAMAS
Support of Hypothesis:
  [ArmeAssault]
Implicit data based on hypothesis:
  [Bombing, Killed, Middle_East, Military, North_America, Wounded, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  []

8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ Khmer_Rouge
Support of Hypothesis:
  [ArmeAssault]
Implicit data based on hypothesis:
  [Civilian, Far_East, Government, Kidnap, Killed, Military, ProjecteExpl, Wounded, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  []

8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ LTTE
Support of Hypothesis:
  [ArmeAssault]
Implicit data based on hypothesis:
  [Bombing, Central_Aisa, Civilian, Kidnap, Killed, Wounded, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  []

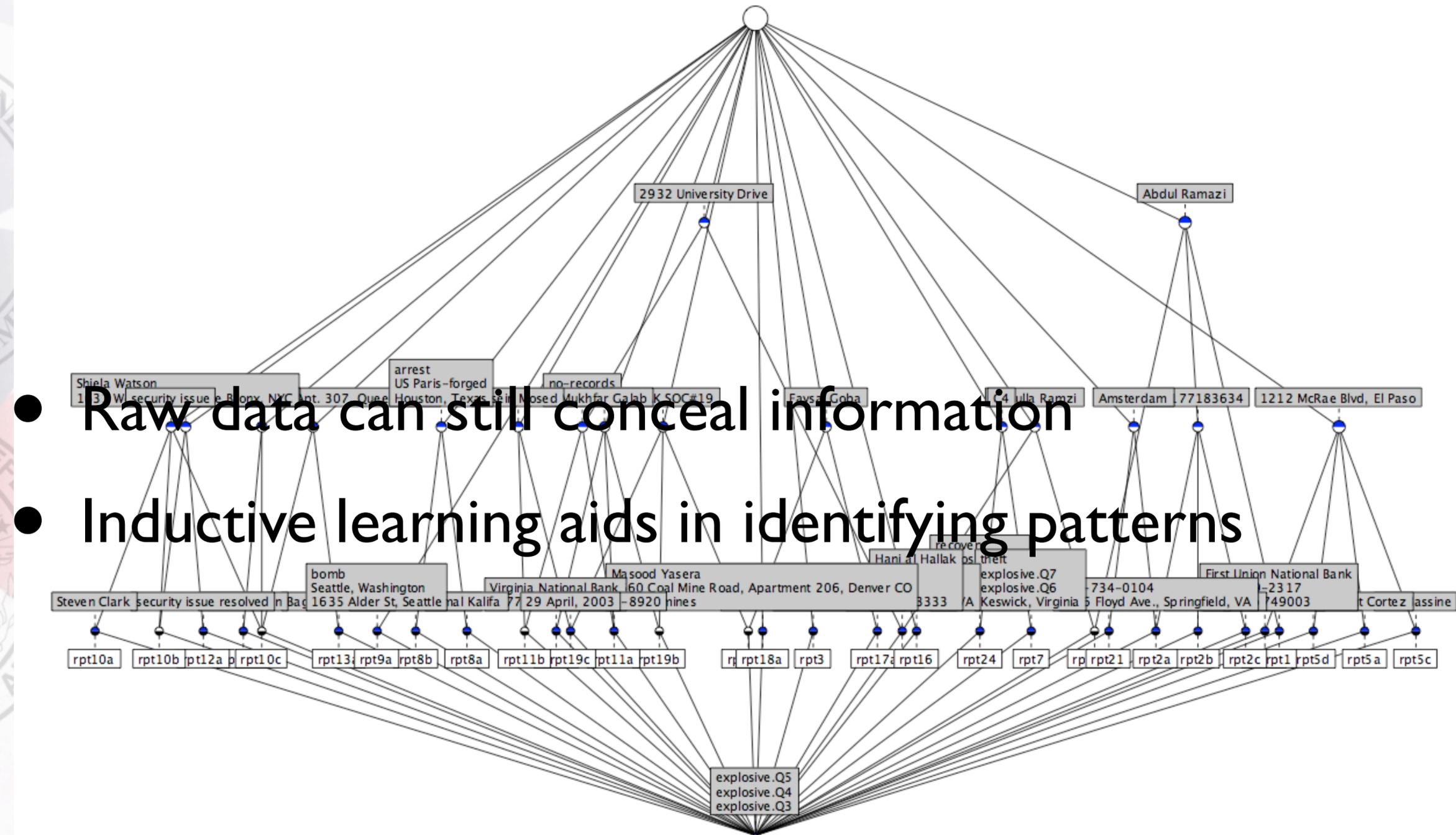
8/3/2007 @ 0:0:00 <=> hypthosis <=> speculation_capabilities ^ HAMAS
Support of Hypothesis:
  [Bombing, North_America]
Implicit data based on hypothesis:
  [ArmeAssault, Killed, Middle_East, Military, Wounded, result_UNK, tactic_UNK, target_UNK]
Un-resolved collected data:
  [bomb_making, explosives]
```



- Hypothesis testing:
 - Deductive - Inductive process
 - Severe Test Constrained (D. Mayo)

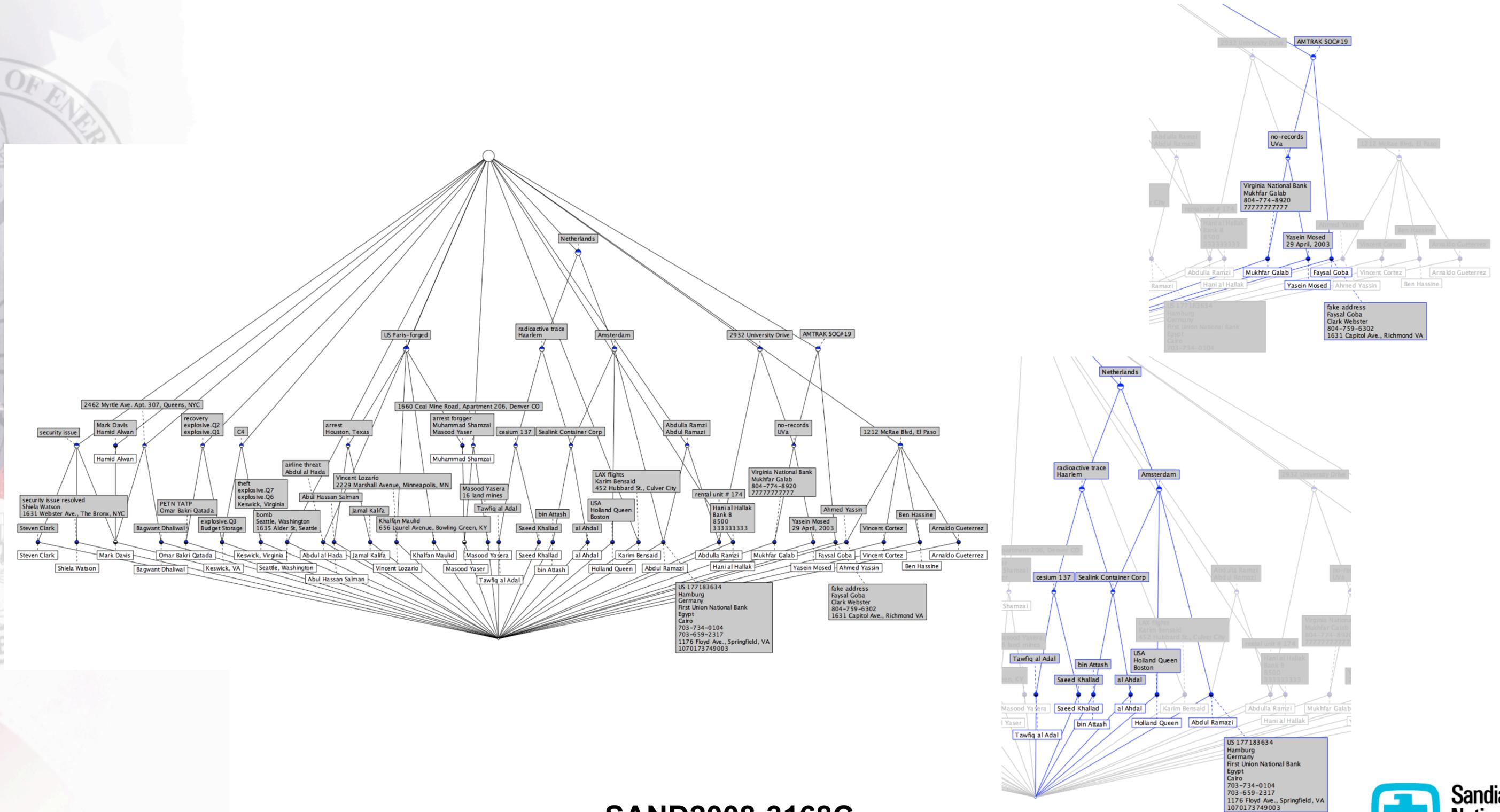
SAND2008-3168C

Intelligence Reports



SAND2008-3168C

Knowledge from Intel Reports



SAND2008-3168C

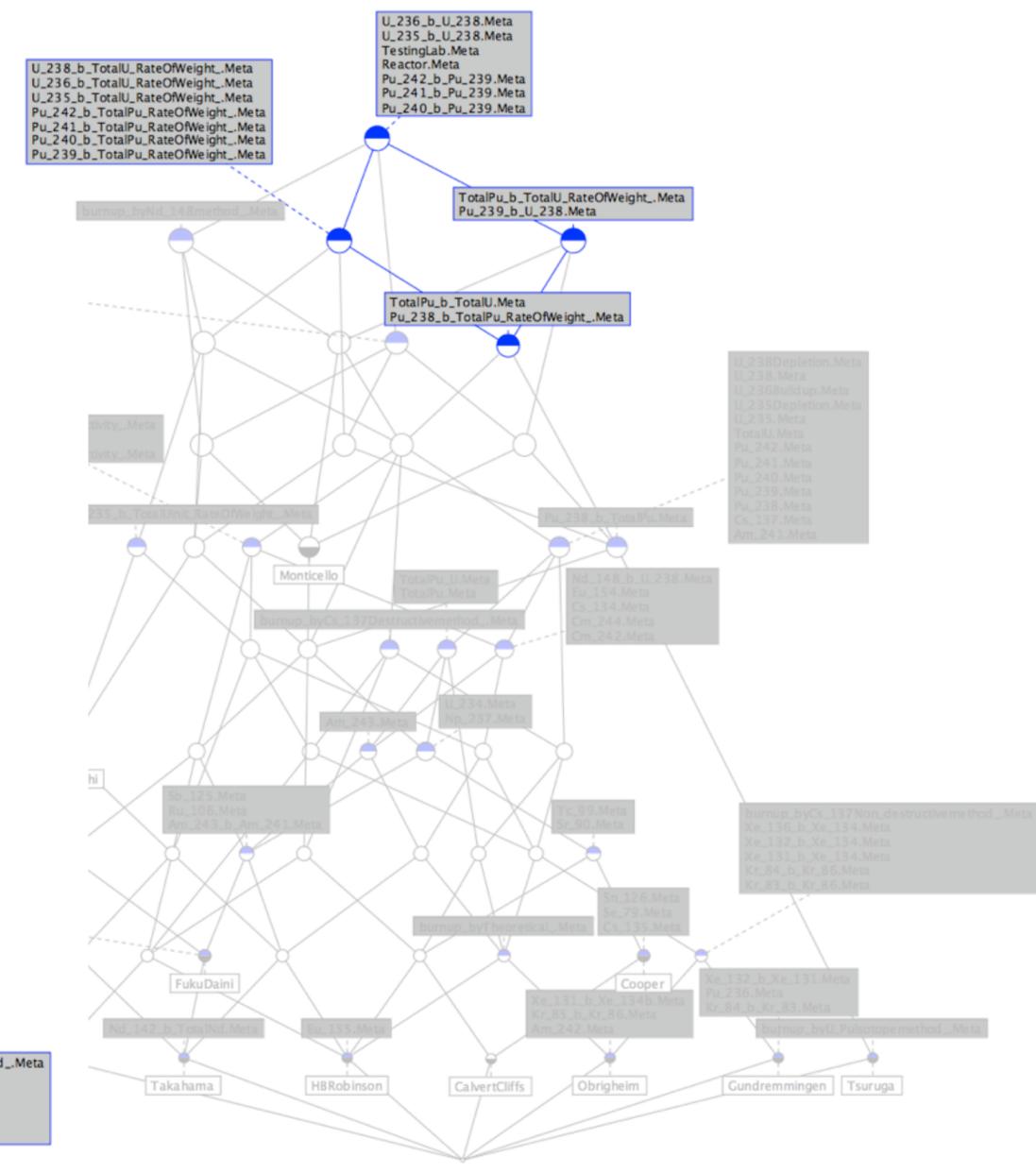
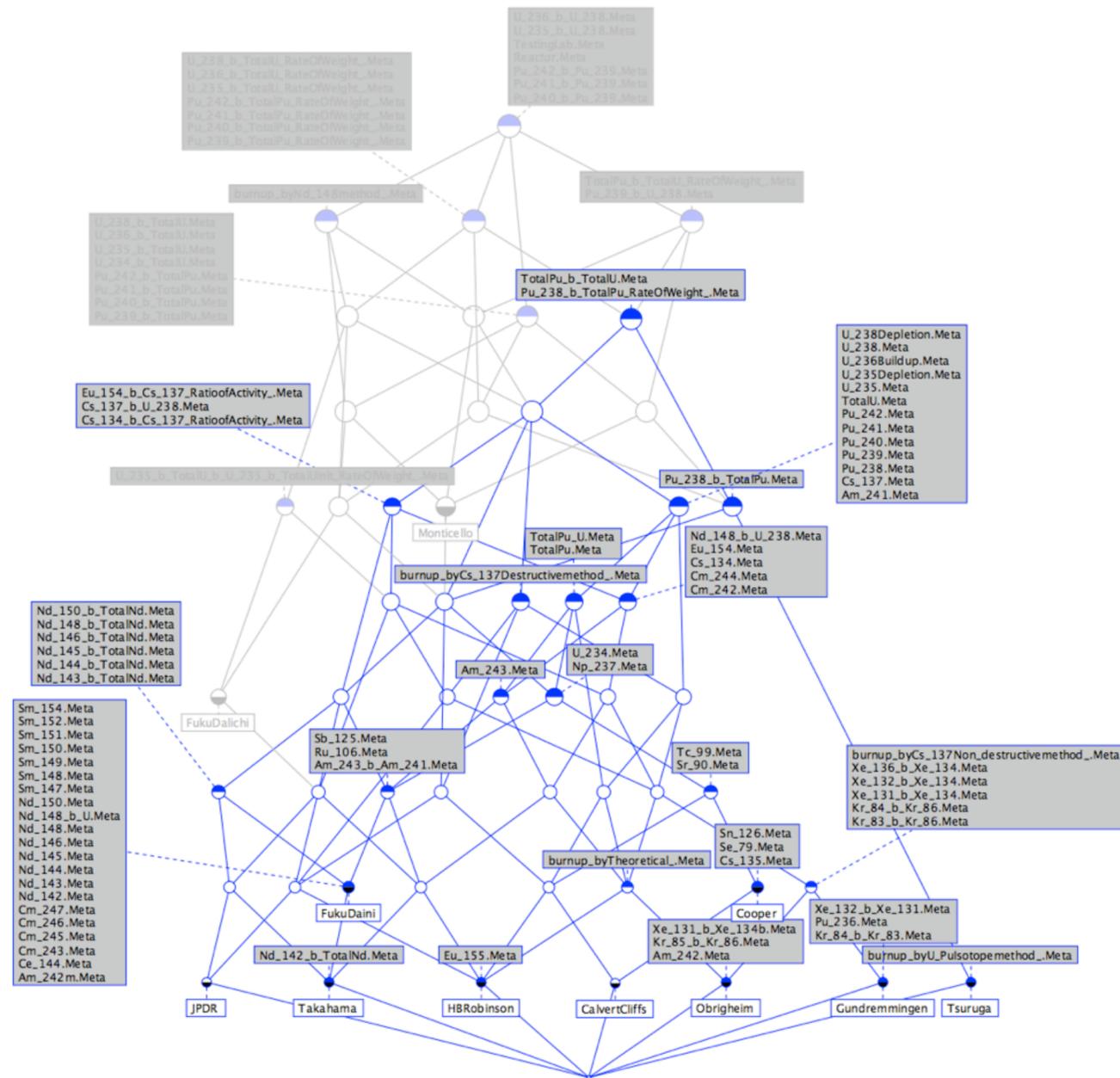


Applications Nuclear Forensics

SAND2008-3168C

Meta Model Identification of Common Attributes

Node that links all objects

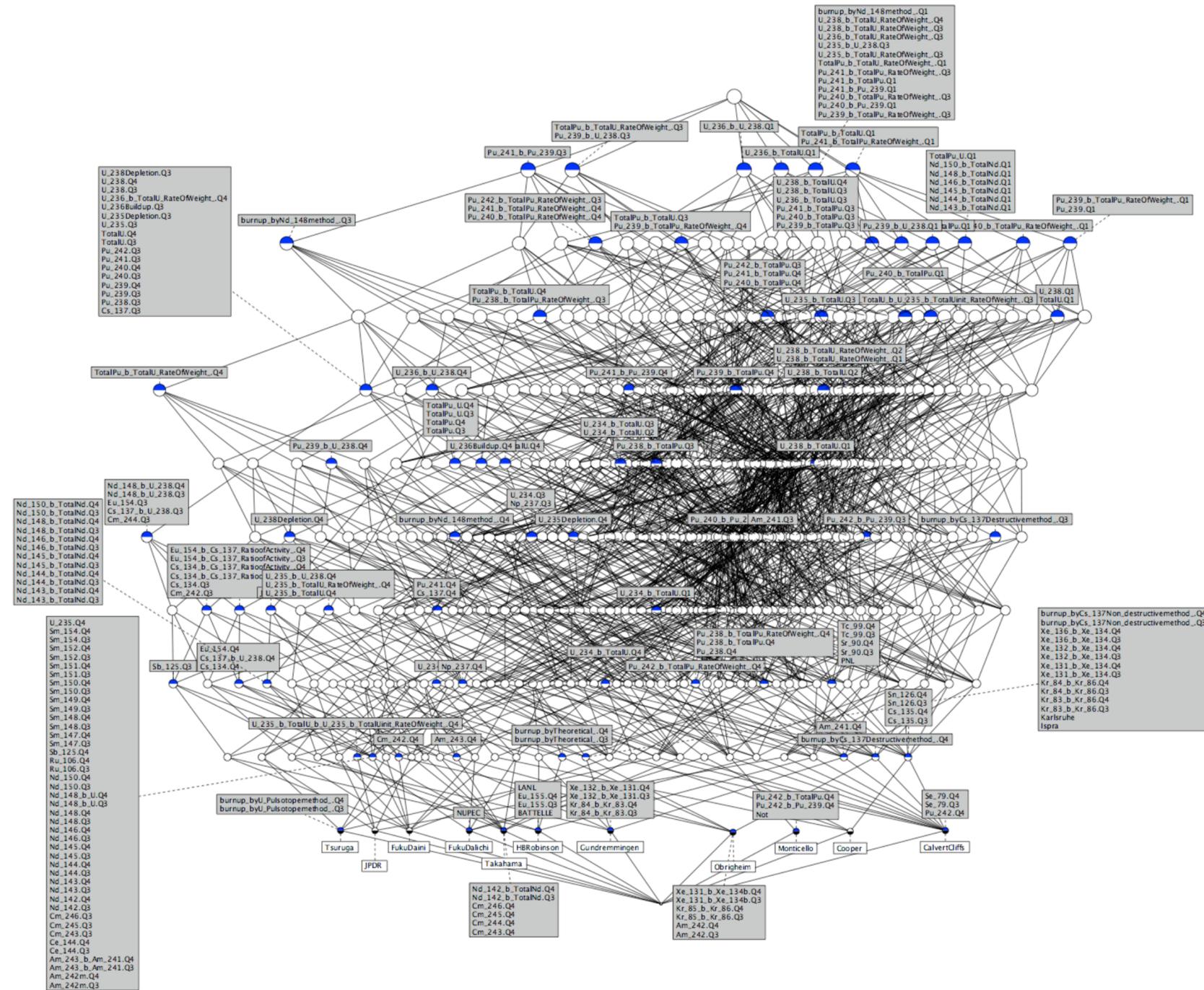


Attributes common to all objects

SAND2008-3168C



FCA Lattice of SFCOMPO Data



SAND2008-3168C

Status

- **Basic functional system complete.**
 - Production architecture implemented.
 - Abduction engine operational;
 - Inductive learning engine operational.
 - Accesses information in MySQL databases.
- **First order knowledge operators in place.**
 - Complete set identified.
- **Zero order modal logics implemented.**
 - Belief cache revision logic needed.
 - Enabling a more robust “what-if” capability.
- **Markov based Temporal attributes to be implemented.**
 - Enables a unique ‘state based’ knowledge representation.
 - Supported by Nuclear Forensics application.
- **Current lattice display is ‘third party’.**
 - 3-D display is in development.

SAND2008-3168C



Thats
All(...enough) Folks...

SAND2008-3168C