NOVAMENTE

A Practical Architecture for Artificial General Intelligence

Ben Goertzel, PhD
Novamente LLC
Biomind LLC
Artificial General Intelligence Research Institute
Virginia Tech, Applied Research Lab for National and Homeland Security

The Novamente Project

- **Long-term goal**:
  - creating "artificial general intelligence" approaching and then exceeding the human level
- **Novamente AI Engine:** an integrative AI architecture
  - Overall design founded on a unique holistic theory of intelligence
  - Cognition carried out via computer science algorithms rather than imitation of human brain
  - efficient, scalable C++/Linux implementation
- Currently, isolated parts of the Novamente codebase are being **used for commercial projects**
  - natural language processing
  - biological data analysis

Overview Papers

- *The Novamente AI Engine*
  - – IJCAI Workshop on Intelligent Control of Agents, Acapulco, August 2003

- *Novamente: An Integrative Architecture for Artificial General Intelligence*
  - – AAAI Symposium on Achieving Human-Level Intelligence Through Integrated Systems and Research, Washington DC, October 2004

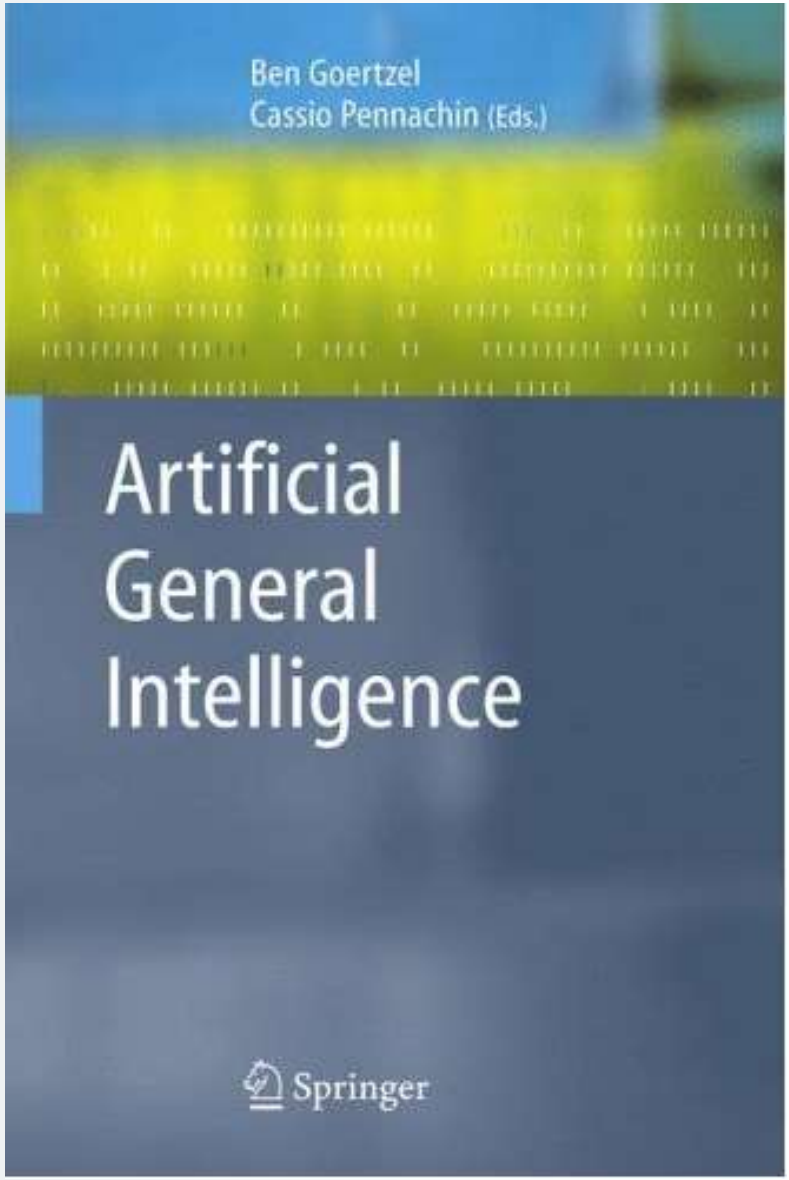
- *Patterns, Hypergraphs and General Intelligence*
  - – World Congress on Computational Intelligence, Vancover CA, July 2006

- Chapter on Novamente in
  - – *Artificial General Intelligence* volume, Springer Verlag, 2006

This edited volume -- the first ever to focus exclusively on Artificial General Intelligence -- is edited by Dr. Ben Goertzel and Cassio Pennachin and contains chapters by AGI researchers at universities, corporations and research institutes around the world.

A partial author list includes:

- Ben Goertzel (Novamente LLC)
- Cassio Pennachin (Novamente LLC)
- Marcus Hutter (IDSIA)
- Juergen Schmidhuber (ISDIA)
- Pei Wang (Temple University)
- Peter Voss (A2I2)
- Vladimir Redko (Keldysh Institute)
- Eliezer Yudkowsky (SIAI)
- Lukasz Kaiser (Aachen Univ. of Technology)



NOVAMENTE LLC
ARTIFICIAL GENERAL INTELLIGENCE

Novamente AI Engine

Components of the system have been commercially deployed

- Biomind OnDemand product for bioinformatic data analysis
- ImmPort: NIH Web portal with Biomind/Novamente based analytics on the back end
- INLINK language processing system developed for INSCOM (Army Intelligence)

NOVAMENTE LLC
ARTIFICIAL GENERAL INTELLIGENCE

The Grand Vision

- Conceptual Background
- Teaching Approach
- Knowledge Representation
- Software Architecture
- Cognitive Processes
- Emergent Mental Structures

The Current Reality

- Implemented Components
- Simulation-World Experiments

The Path Ahead

Novamente:
The Grand Vision

Conceptual Background:

Patternist Philosophy of Mind

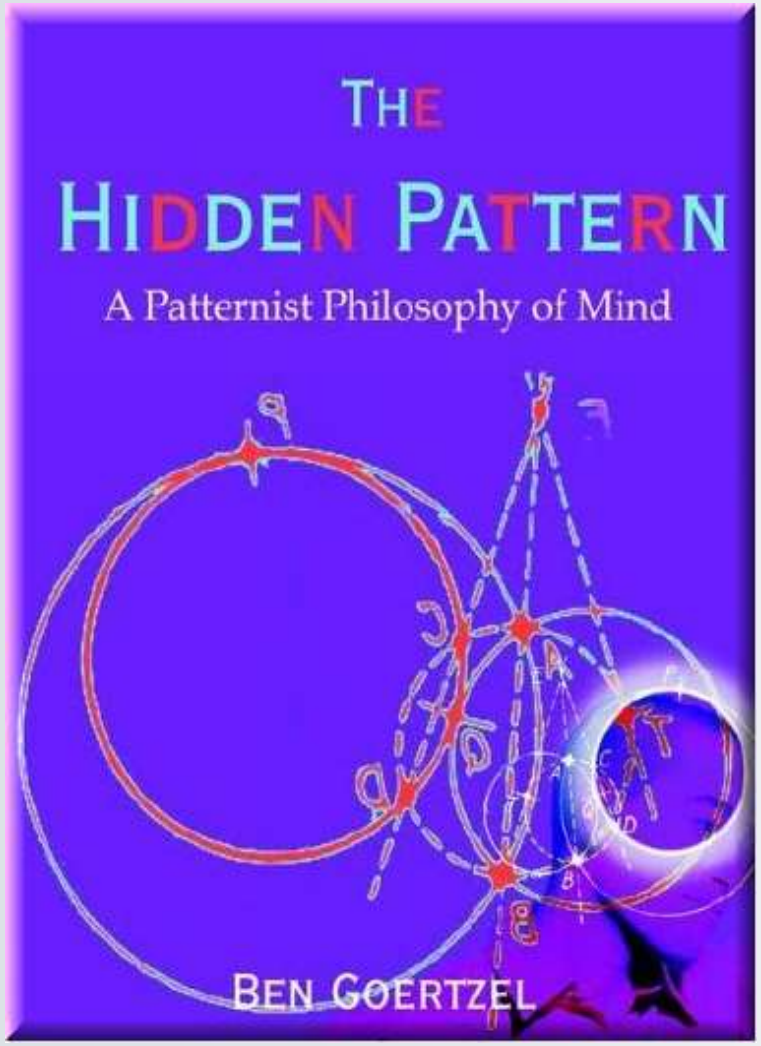
- An intelligent system is conceived as a system for recognizing patterns in the world and in itself
- Probability theory may be used as a language for quantifying and relating patterns
- Logic (term, predicate, combinatory) may be used as a base-level language for expressing patterns
- The reflexive process of *flexibly recognizing patterns in oneself and then improving oneself based on these patterns* is the "basic algorithm of intelligence"
- The phenomenal self, a key aspect of intelligent systems, is the result of an intelligent system *recognizing itself as a pattern* in its (internal and external) behaviors

Conceptual Background:

Definition of Intelligence

- Intelligence is considered as the ability to achieve complex goals in a complex environment
- Goals are achieved via recognizing probabilistic patterns of the form *“Carrying out procedure P in context C will achieve goal G.”*

Patternist Philosophy



- Minds are systems of patterns that achieve goals by recognizing patterns in themselves and the world
- AI is about creating software whose structures and dynamics will lead to the emergence of these pattern-sets

Prior, Conceptually Relevant Book Publications

The Structure of Intelligence, Springer-Verlag, *1993*

The Evolving Mind, Gordon and Breach, 2003

Chaotic Logic, Plenum Press, 1994

From Complexity to Creativity, Plenum Press, 1997

Creating Internet Intelligence, Kluwer Academic, 2001

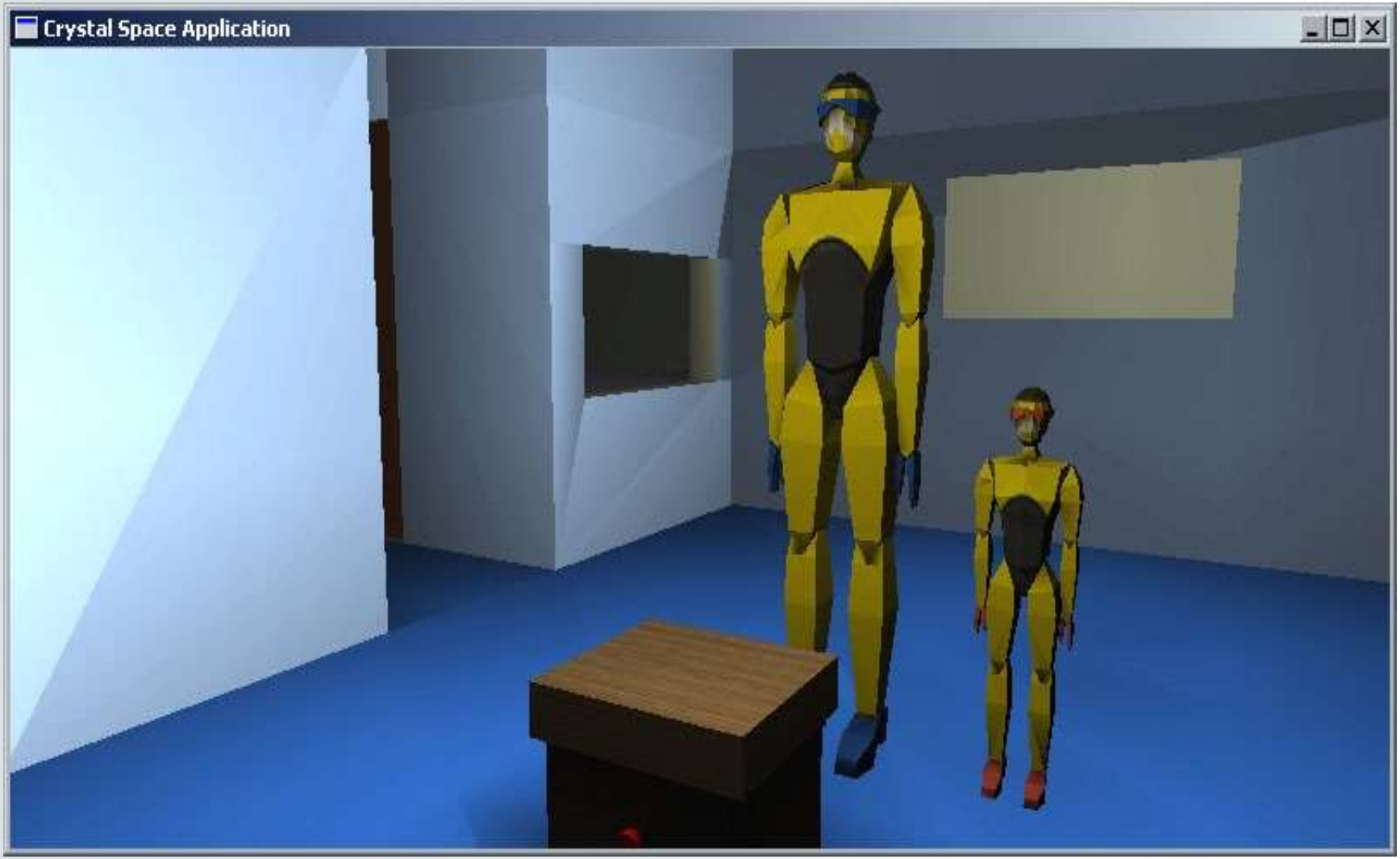
Novamente-Related Books-in-Progress

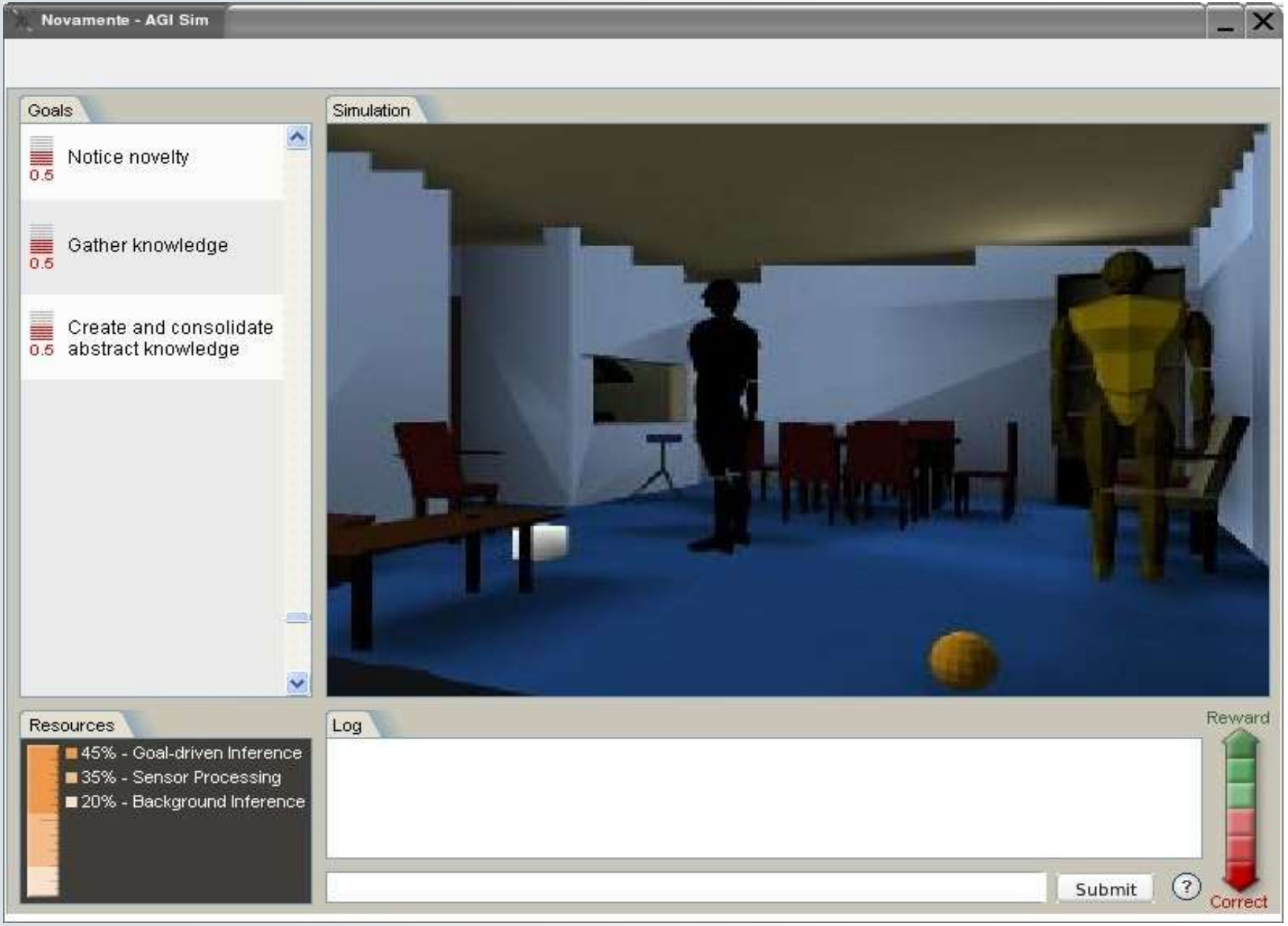
- *Probabilistic Term Logic*
  - *–In final editing stage; to be submitted 2006*
- *Engineering General Intelligence*
  - *–In final editing stage*
  - *–Reviews the overall NM design*
  - *–May or may not be submitted for publication (AI Safety and commmercial concerns)*
- *Artificial Cognitive Development*
  - *–Developmental psychology for Novamente and other AGIs*
  - *–In preparation*

AI Teaching Methodology

- Embodiment
- Post-embodiment
- Developmental Stages

Embodiment in AGISim Simulation World



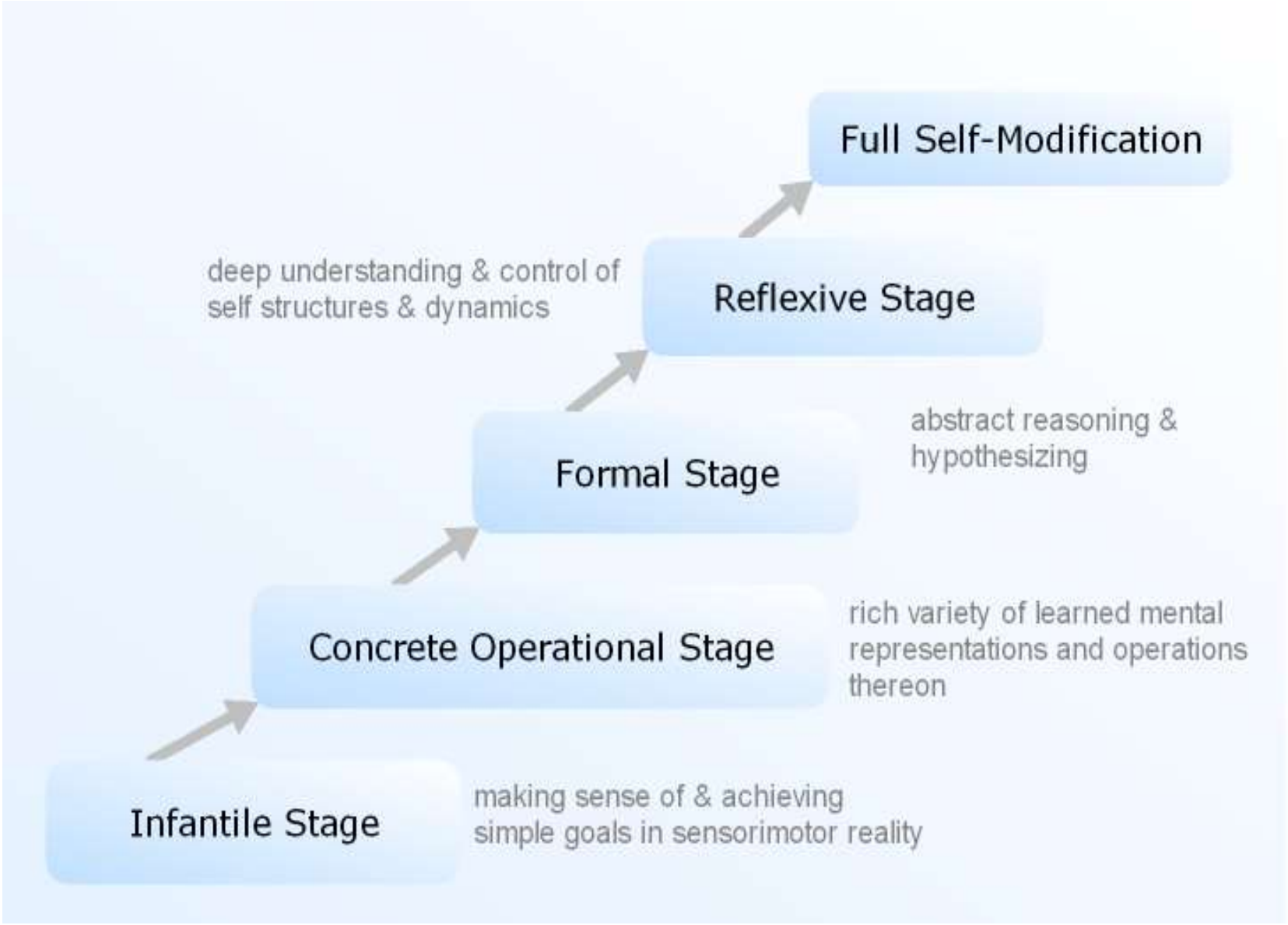

Novamente - AGI Sim
Goals
0.5
Notice novelty
0.5
Gather knowledge
0.5
Create and consolidate abstract knowledge
Simulation
Resources
45% - Goal-driven Inference
35% - Sensor Processing
20% - Background Inference
Log
Submit
Reward
Correct

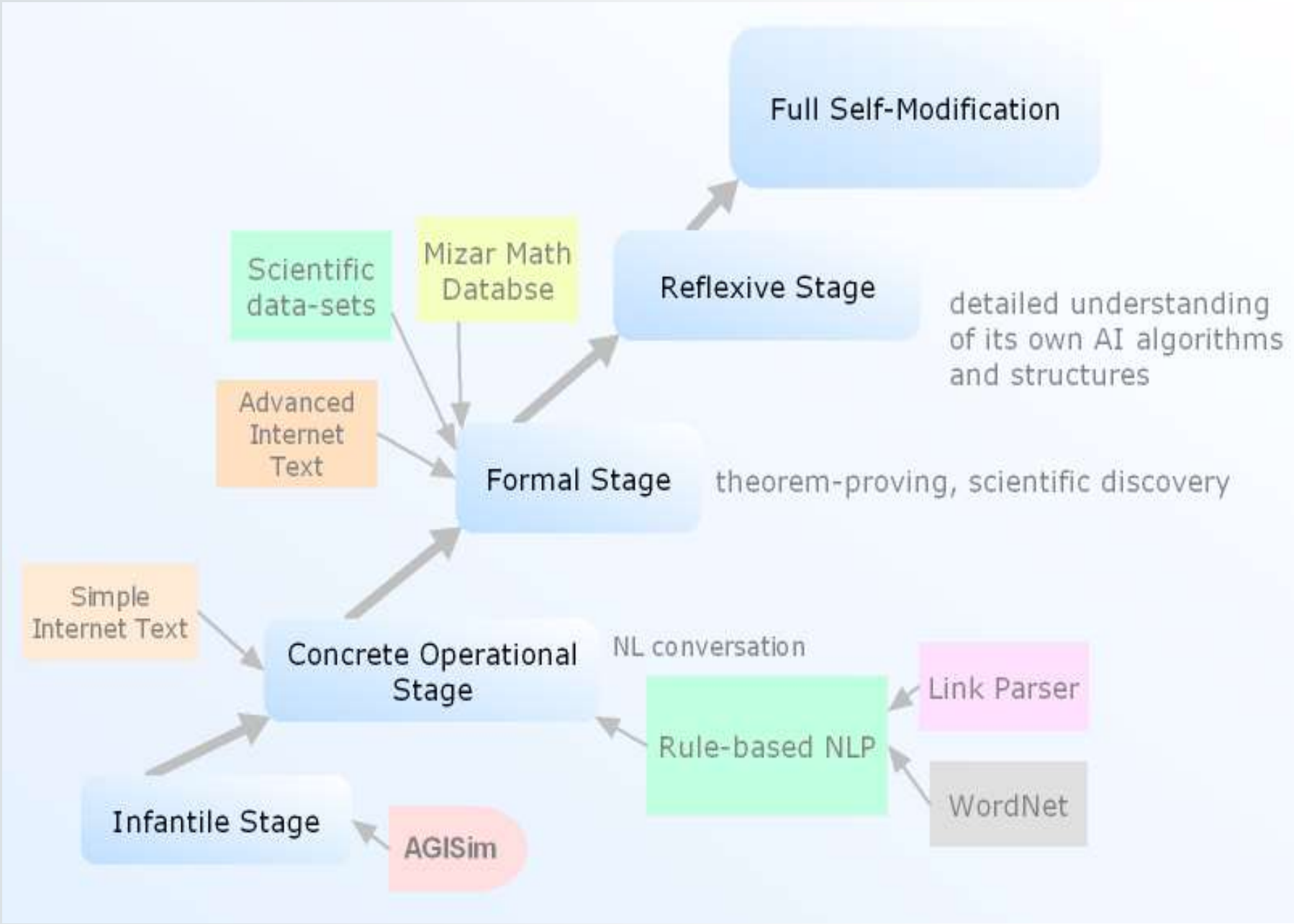
Post-Embodied AI

AI systems may viably synthesize knowledge gained via various means

- virtually embodied experience
  - AGISim
- physically embodied experience
  - Robotics
- explicit encoding of knowledge
  - in natural language
  - In artificial languages such as Lojban, Lojban++
- ingestion of databases
  - WordNet, FrameNet, Cyc, etc.
  - quantitative scientific data

Stages of Cognitive Development




Full Self-Modification
Scientific data-sets
Mizar Math Databse
Reflexive Stage
detailed understanding of its own AI algorithms and structures
Advanced Internet Text
Formal Stage
theorem-proving, scientific discovery
Simple Internet Text
Concrete Operational Stage
NL conversation
Link Parser
Rule-based NLP
WordNet
Infantile Stage
AGISim

Knowledge Representation

Novamente's "Atom Space"

- Atoms = Nodes or Links
- Atoms have
  - Truth values (probability + weight of evidence)
  - Attention values (short and long term importance)
- The Atomspace is a weighted, labeled hypergraph

Novamente's "Atom Space"

- Not a neural net
  - No activation values, no attempt at low-level brain modeling
  - But, Novamente Nodes do have "attention values", analogous to time-averages of neural net activations
- Not a semantic net
  - Atoms may represent percepts, procedures, or parts of concepts
  - Most Novamente Atoms have no corresponding English label
  - But, most Novamente Atoms do have probabilistic truth values, allowing logical semantics

Attention Values

	Low Long-term Importance	High Long-term Importance
Low Short-term Importance	Useless	Remembered but not currently used (e.g. mother's phone #)
High Short-term Importance	Used then forgotten (e.g. most precepts)	Used and remembered

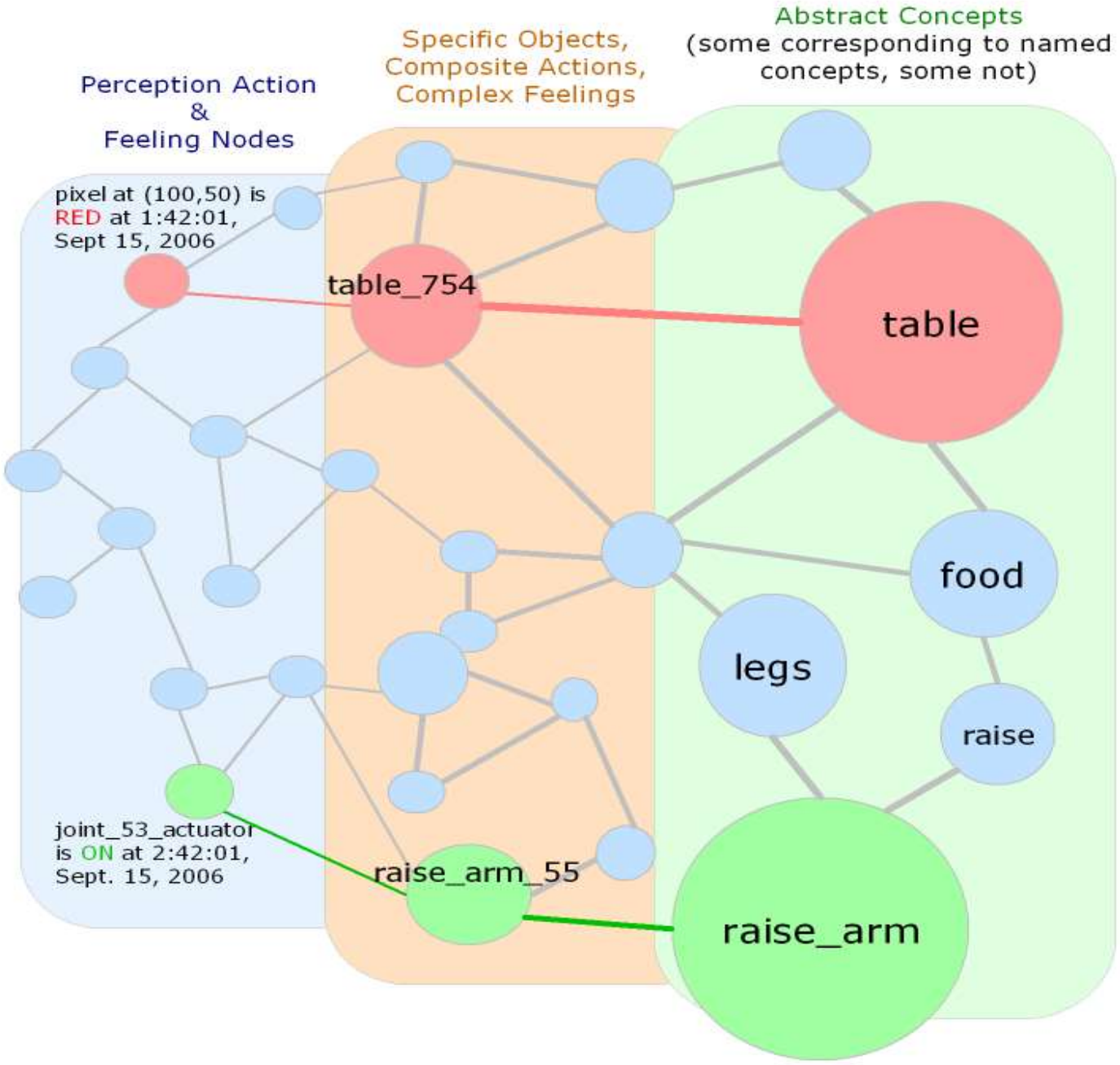
Truth Values

	Strength low	Strength high
Weight of evidence low	Weakly suspected to be false	Weakly suspected to be true
Weight of evidence high	Firmly known to be false	Firmly known to be true

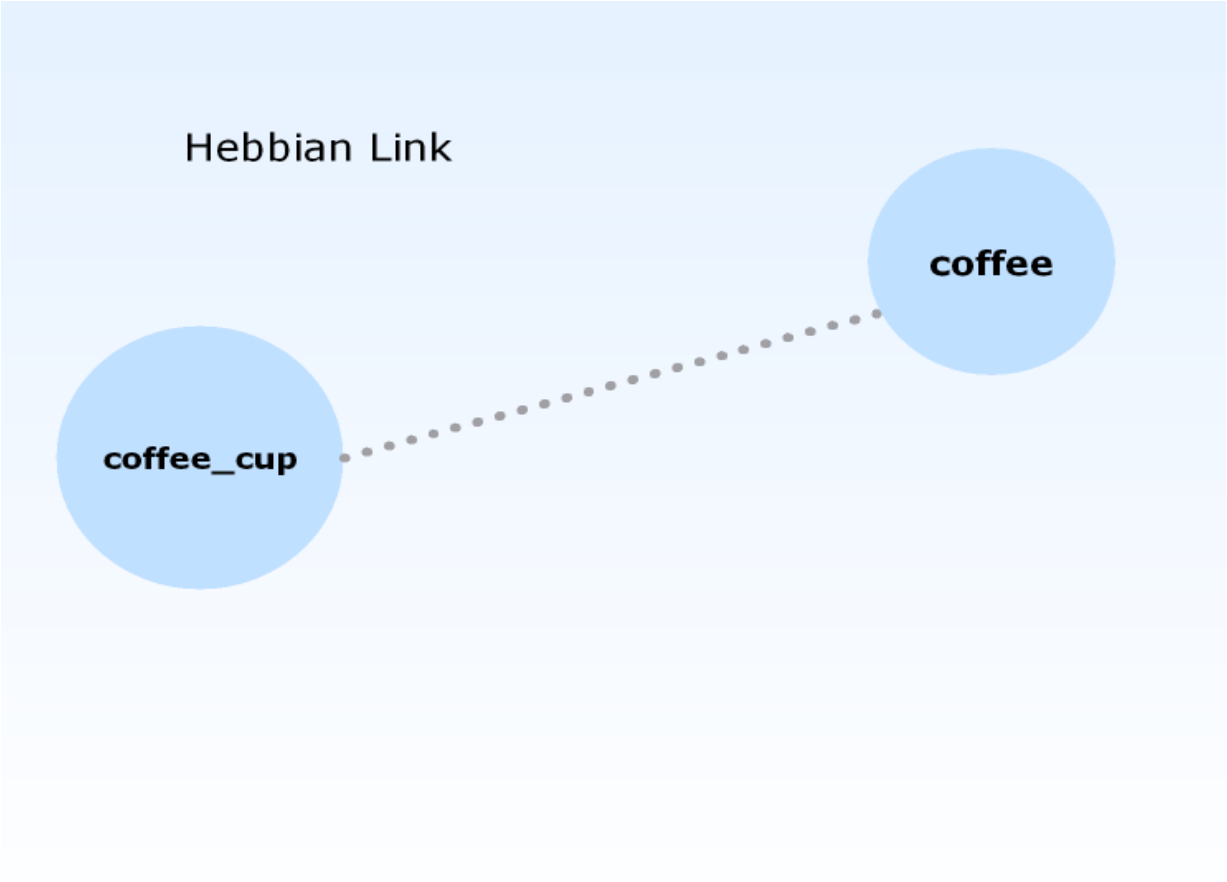
Atoms Come in Various Types

- ConceptNodes
  - “tokens” for links to attach to
- PredicateNodes
- ProcedureNodes
- PerceptNodes
  - Visual, acoustic percepts, etc.
- NumberNodes

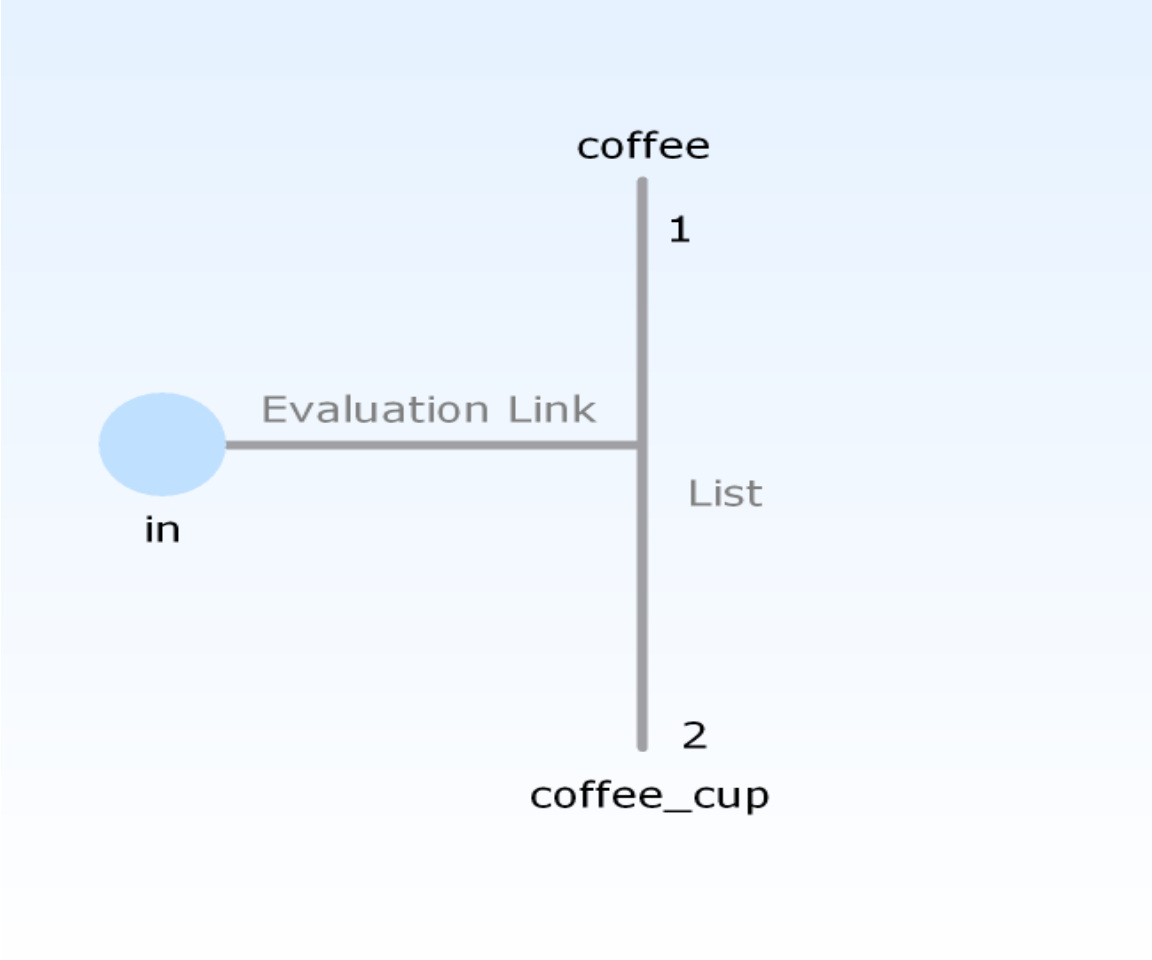
- Logical links
  - InheritanceLink
  - SimilarityLink
  - ImplicationLink
  - EquivalenceLink
  - Intensional logical relationships
- HebbianLinks
- Procedure evaluation links


Perception Action
&
Feeling Nodes
Specific Objects,
Composite Actions,
Complex Feelings
Abstract Concepts
(some corresponding to named
concepts, some not)
pixel at (100,50) is
RED at 1:42:01,
Sept 15, 2006
table_754
table
food
legs
raise
joint_53_actuator
is ON at 2:42:01,
Sept. 15, 2006
raise_arm_55
raise_arm

Links may denote generic association …



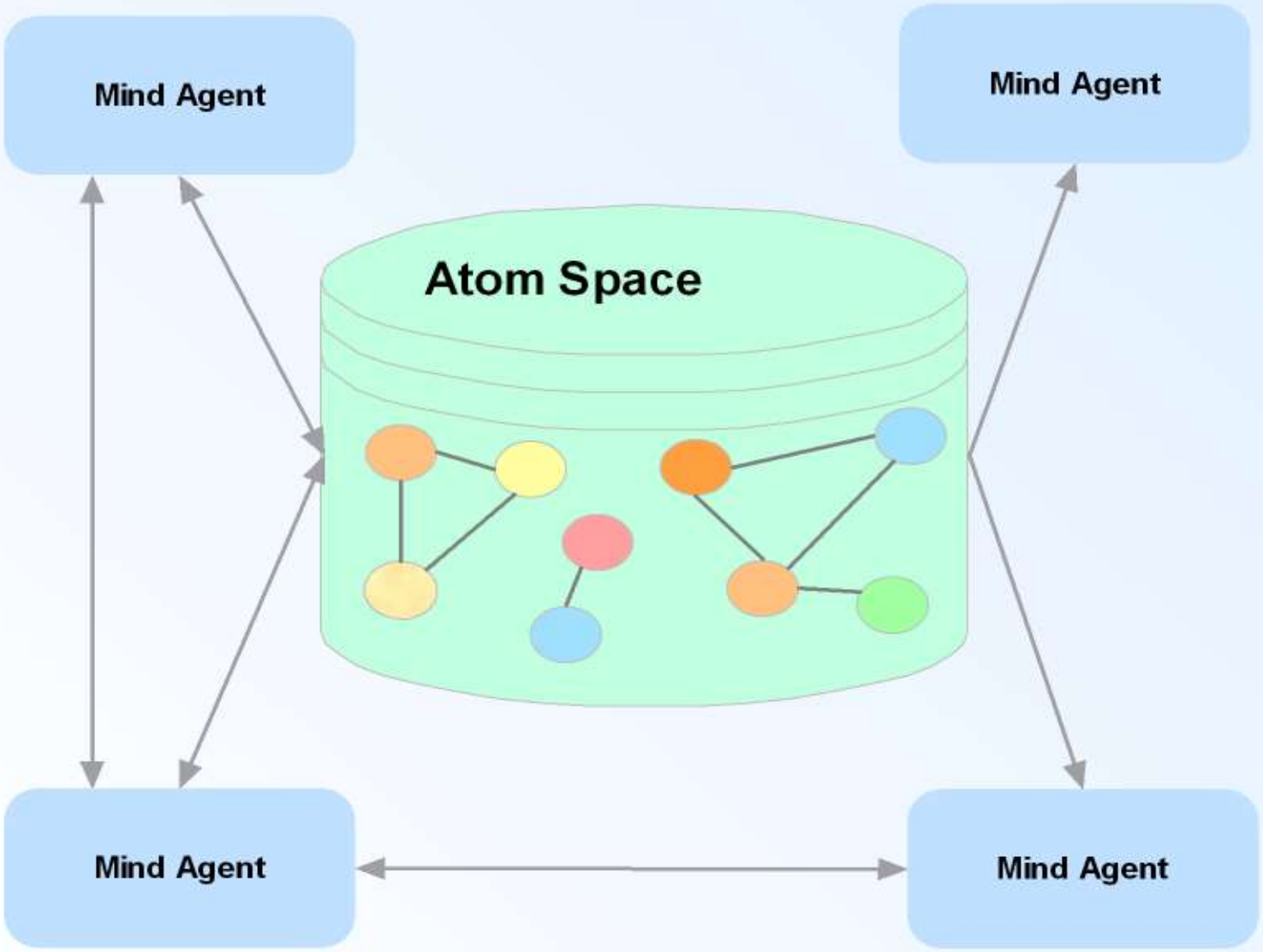
…or precisely specified relationships

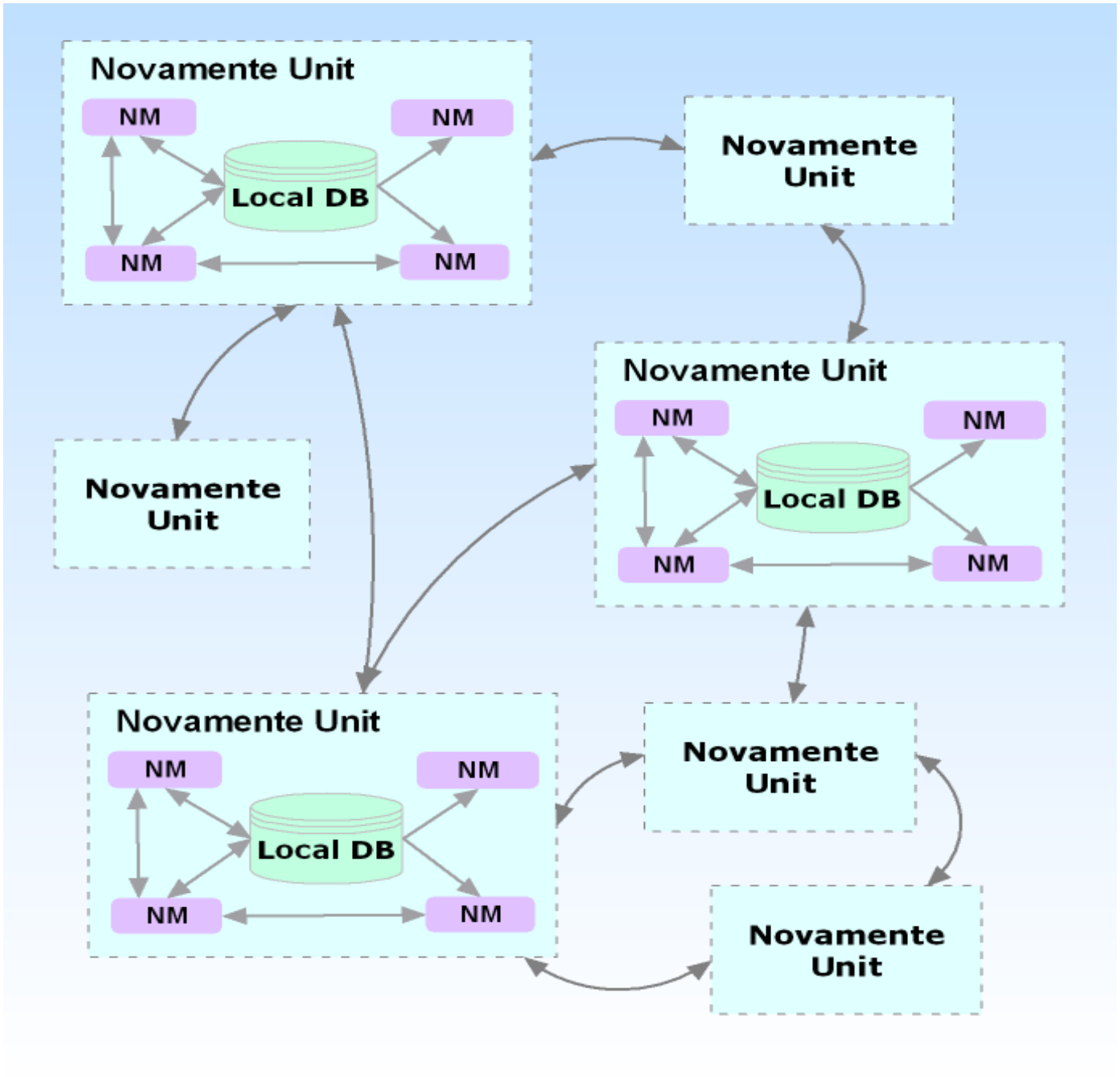



people
1
drink
Evaluation Link
1
List
Evaluation Link
2
from
coffee
2
coffee_cup

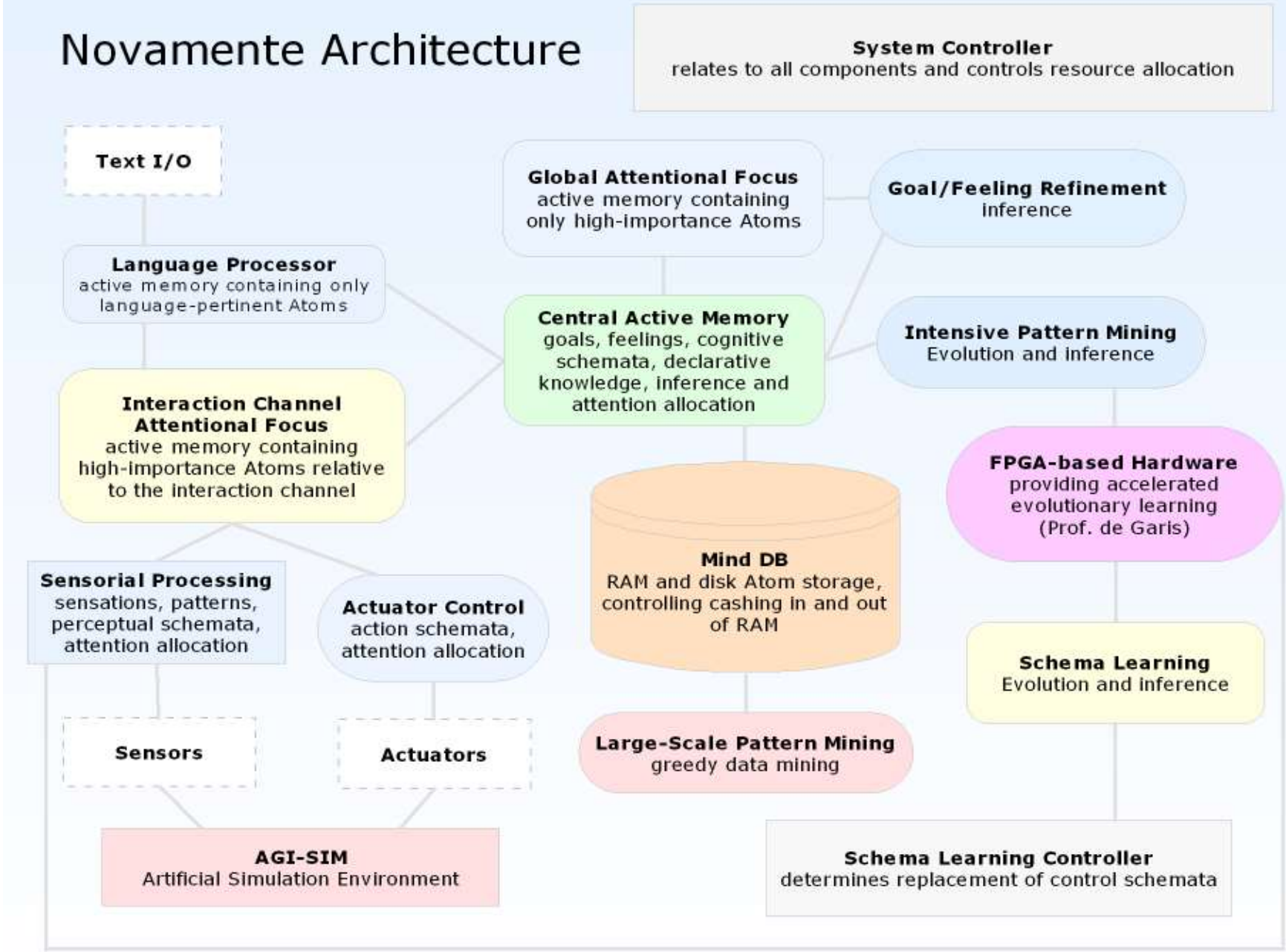
Software Architecture & Cognitive Architecture

Novamente Machine

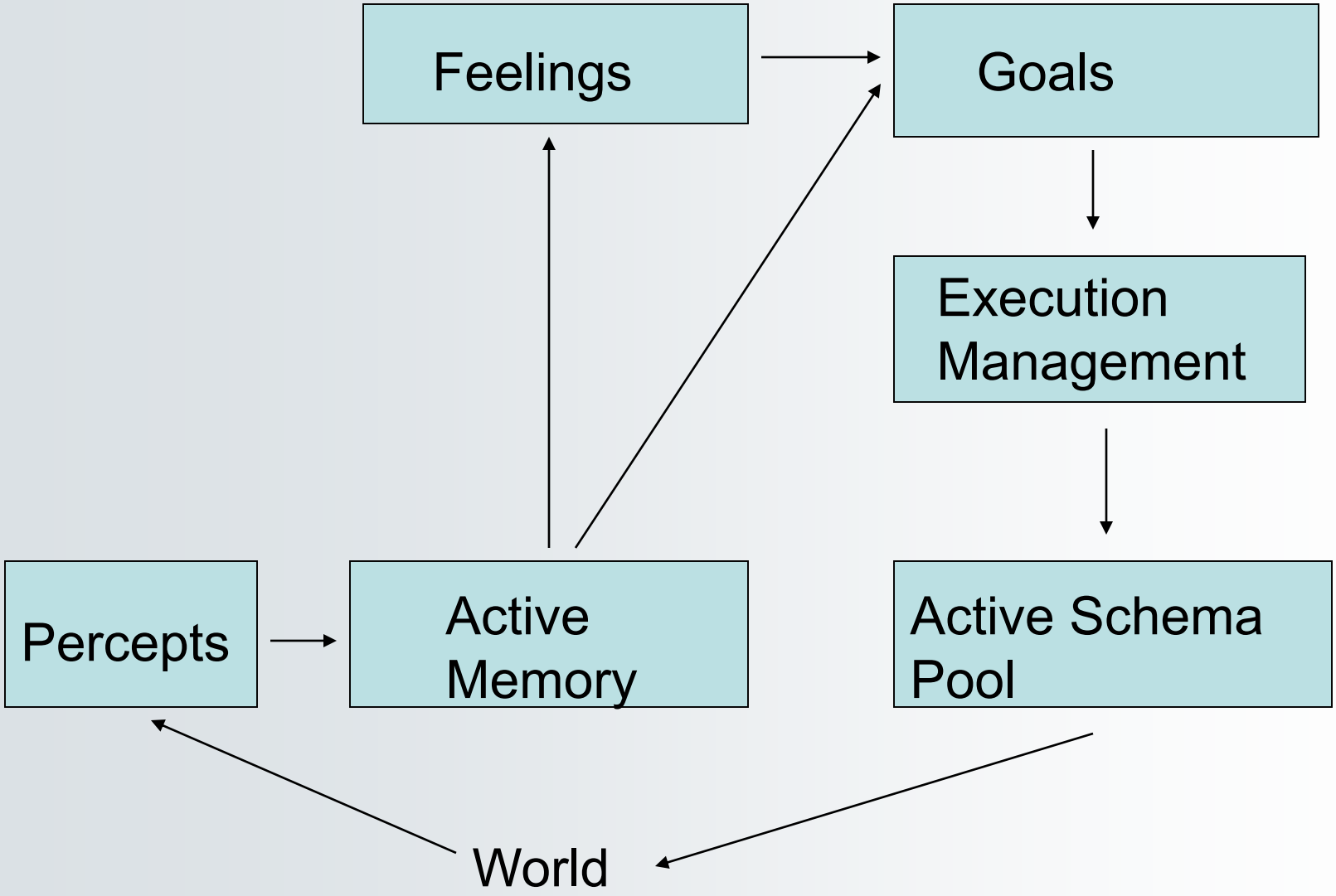



Novamente Unit
NM
NM
Local DB
NM
NM
Novamente Unit
Novamente Unit
Novamente Unit
NM
NM
Local DB
NM
NM
Novamente Unit
Novamente Unit
NM
NM
Local DB
NM
NM
Novamente Unit
Novamente Unit

Novamente Architecture



Simplified Workflow



Cognitive Processes

Typology of Cognitive Processes

Global processes

- MindAgents that periodically iterate through all Atoms and act on them
- “Things that all Atoms do”

Control Processes

- Execution of actions
- Maintenance of goal hierarchy
- Updating of system control schemata

Focused processes

- MindAgents that begin by selecting a small set of important or relevant Atoms, and then act on these to generate a few more small sets of Atoms, and iterate
- Two species:
  - Forward synthesis
  - Backward synthesis

Global Cognitive Processes

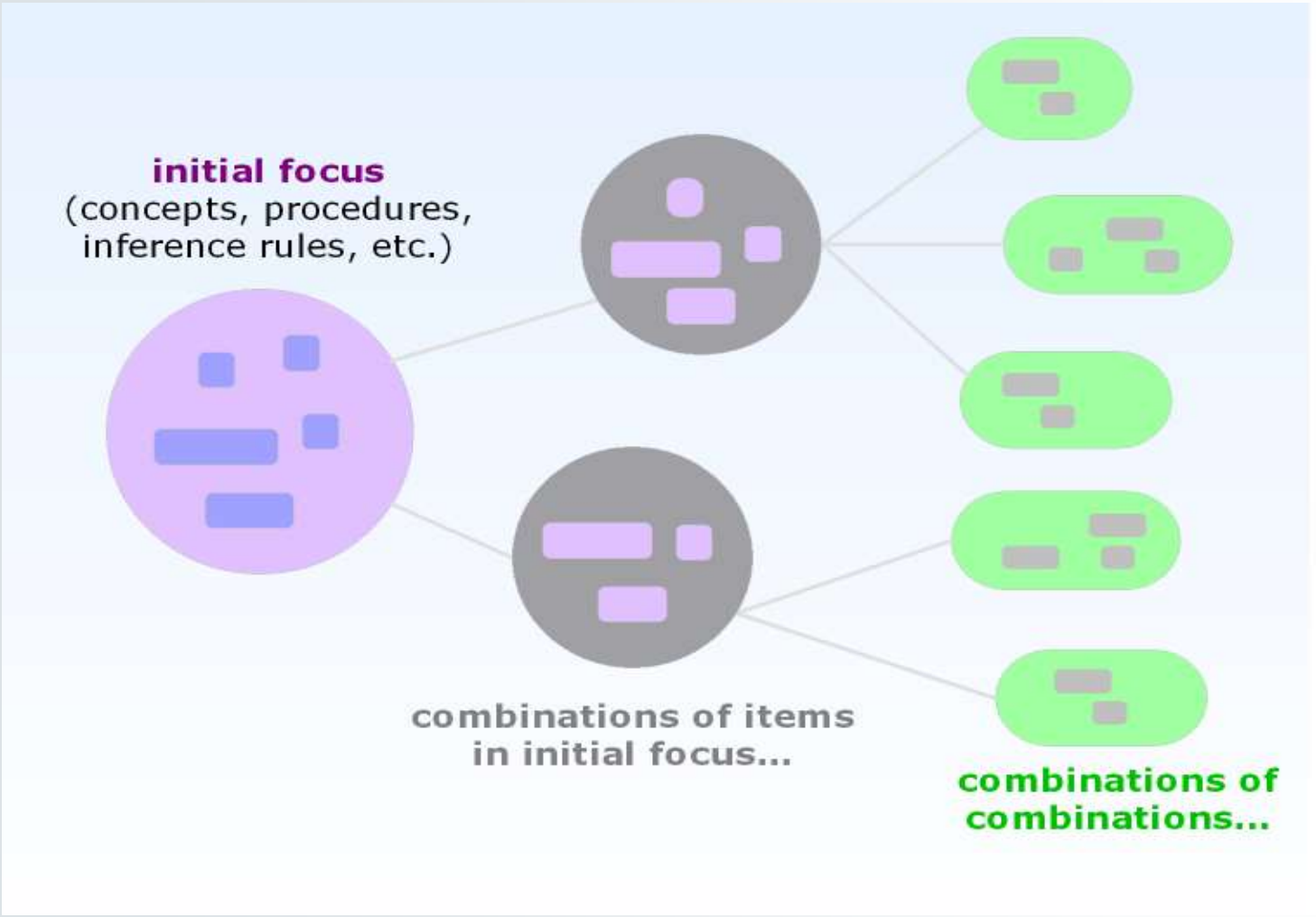
- **Attention Allocation**
  - Updates short and long term importance values associated with Atoms
  - Uses a “simulated economy” approach, with separate currencies for short and long term importance
- **Stochastic pattern mining of the AtomTable**
  - A powerful heuristic for predicate formation
  - Critical for perceptual pattern recognition as well as cognition
  - Pattern mining of inference histories critical to advanced inference control
- **Building of the SystemActivityTable**
  - Records which MindAgents acted on which Atoms at which times
  - Table is used for building HebbianLinks, which are used in attention allocation

Control Processes

- **Execution of procedures**
  - "Programming language interpreter" for executable procedures created from NM Atoms
- **Maintenance of "active procedure pool"**
  - Set of procedures that are currently ready to be activated if their input conditions are met
- **Maintenance of "active goal pool"**
  - Set of predicates that are currently actively considered as system goals

Global Cognitive Processes, Part I

Forward Synthesis



Forward Synthesis Processes

- **Forward-Chaining Probabilistic Inference**
  - Given a set of knowledge items, figure out what (definitely or speculatively) follows from it
- **Concept/Goal Formation**
  - “Blend” existing concepts or goals to form new ones
- **Map formation**
  - Create new Atoms out of sets of Atoms that tend to be simultaneously important (or whose importance tends to be coordinated according to some other temporal pattern)

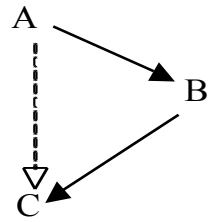
Forward Synthesis Processes

- **Language Generation**
  - Atoms representing semantic relationships are combined with Atoms representing linguistic mapping rules to produce Atoms representing syntactic relationships, which are then transformed into sentences
- **Importance Propagation**
  - Atoms pass some of their “attentional currency” to Atoms that they estimate may help them become important again in the future

“Probabilistic Logic Networks” (PLN) for uncertain inference

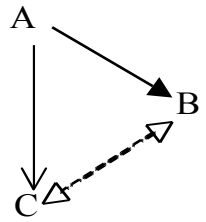
Example First-Order PLN Rules Acting on ExtensionalInheritanceLinks

A → B
B → C
|-
A → C



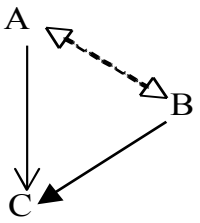
Deduction

A → B
A → C
|-
B → C



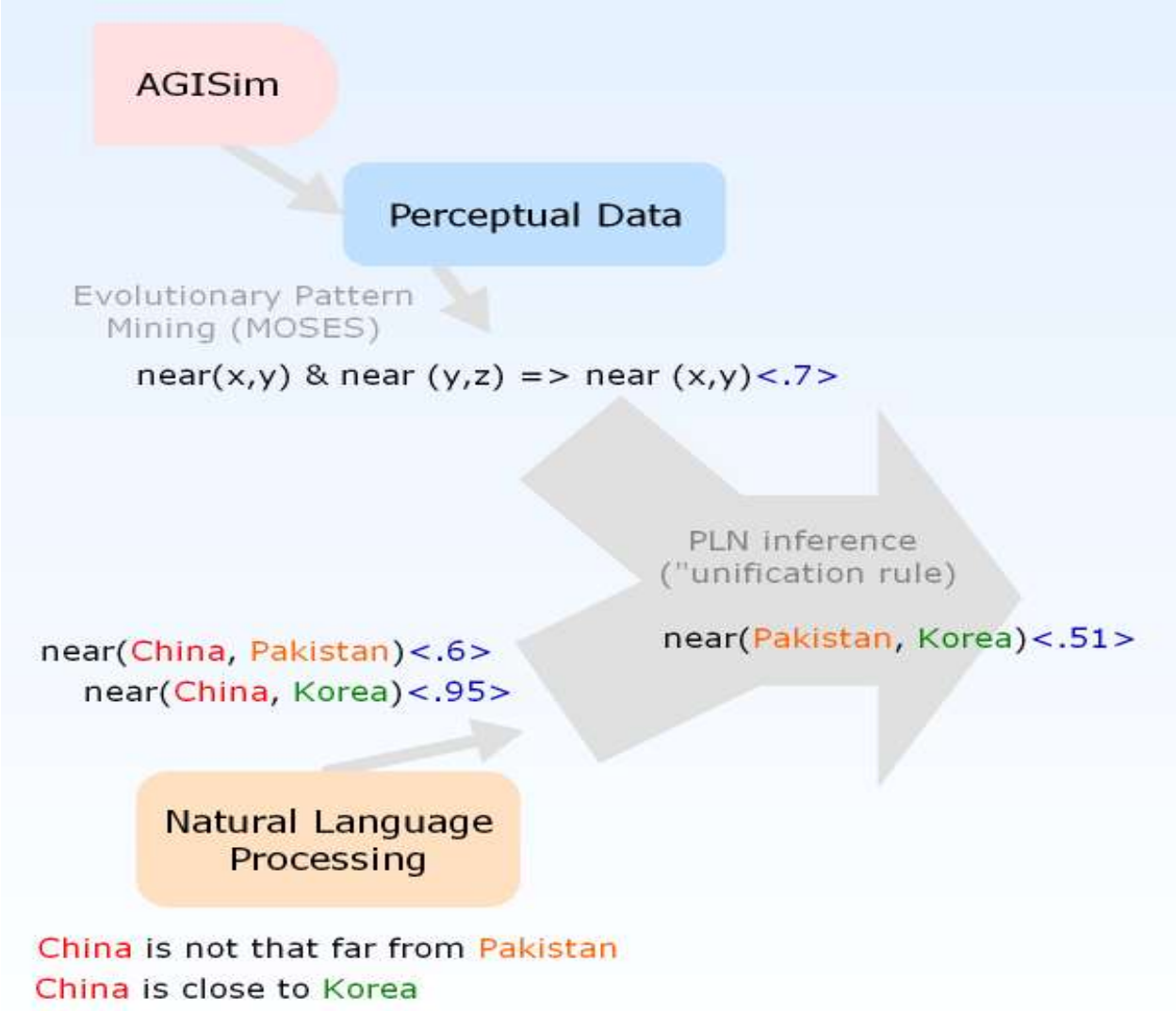
Induction

A → C
B → C
|-
A → B

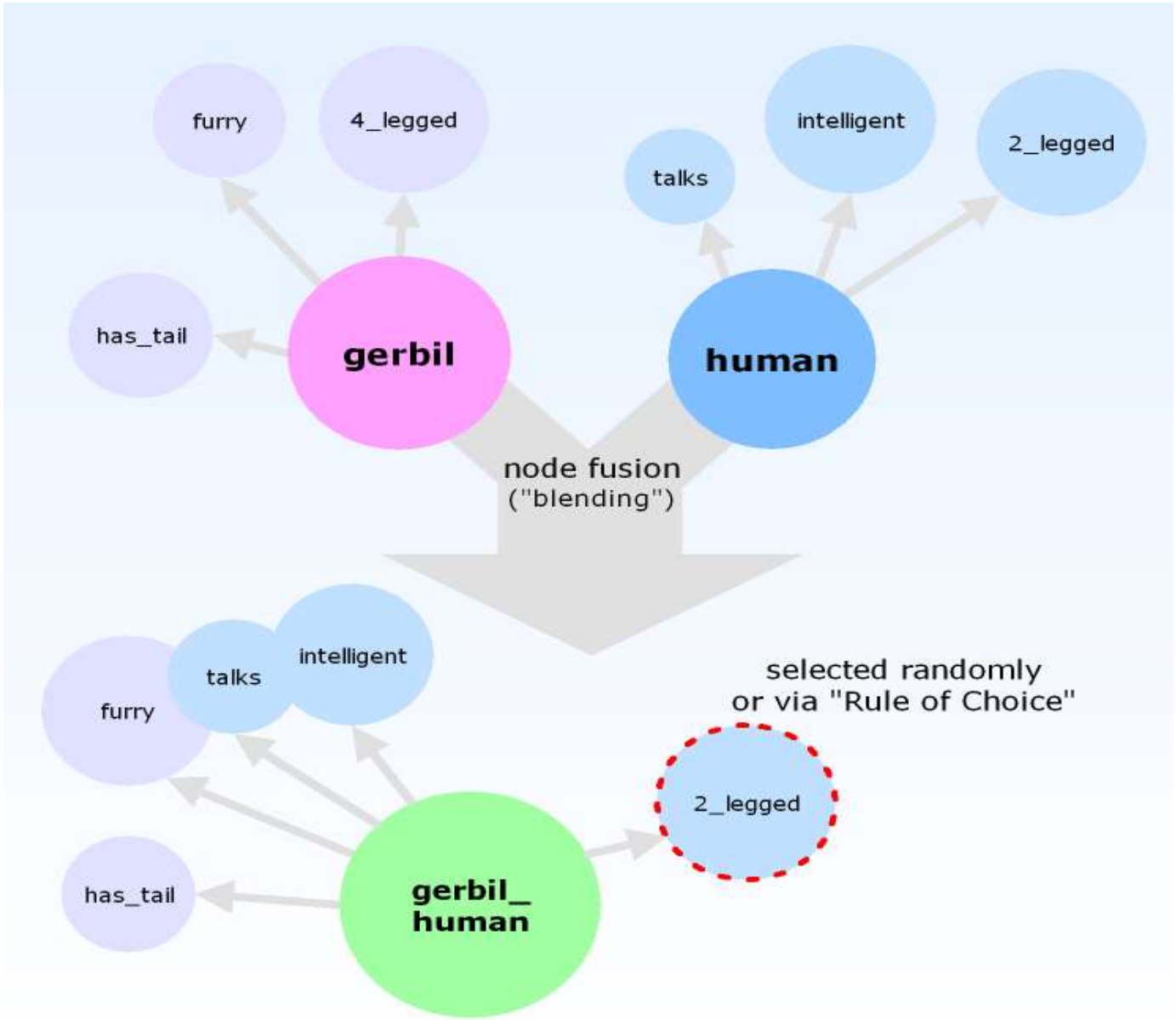


Abduction

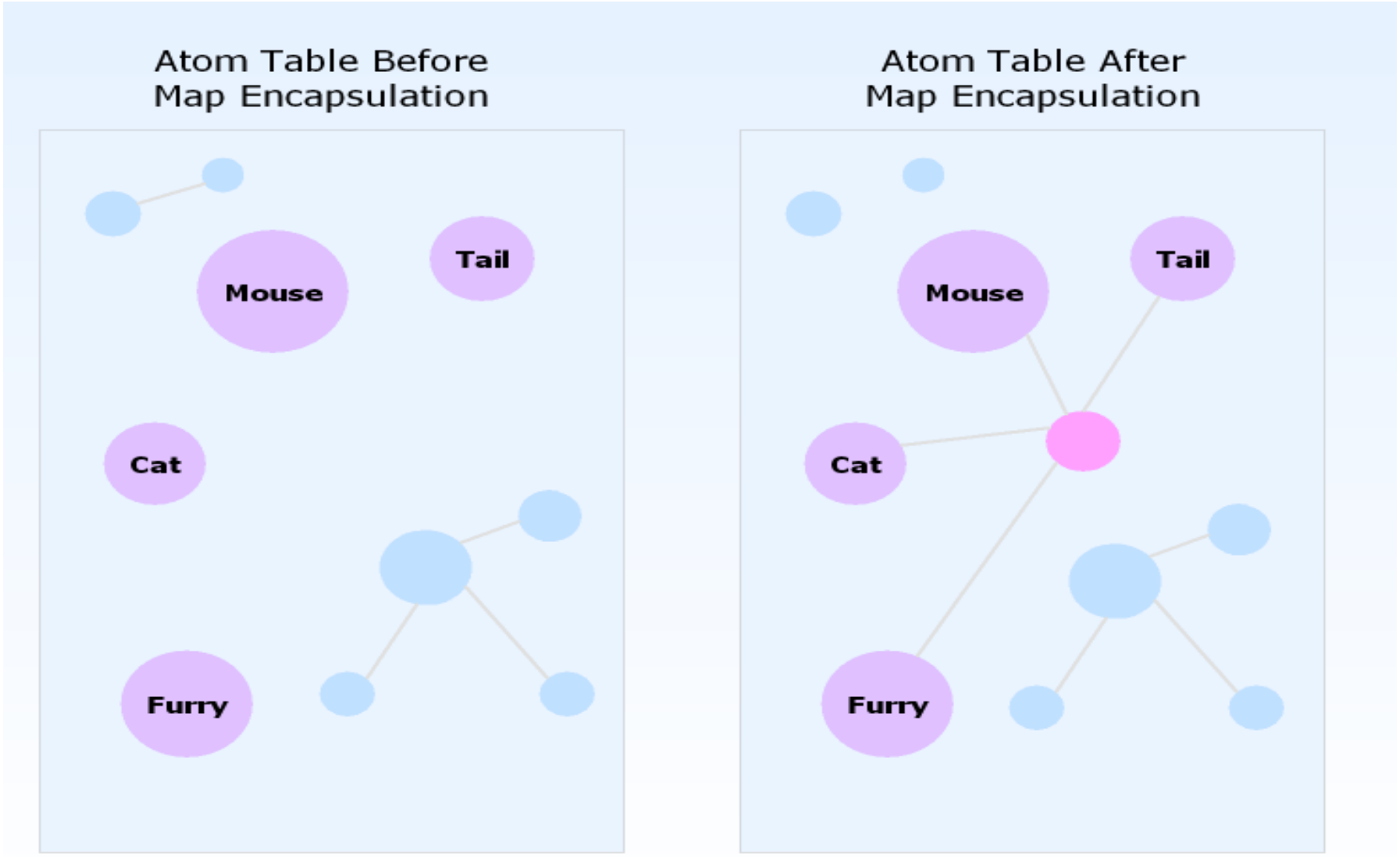
Grounding of natural language constructs is provided via inferential integration of data gathered from linguistic and perceptual inputs



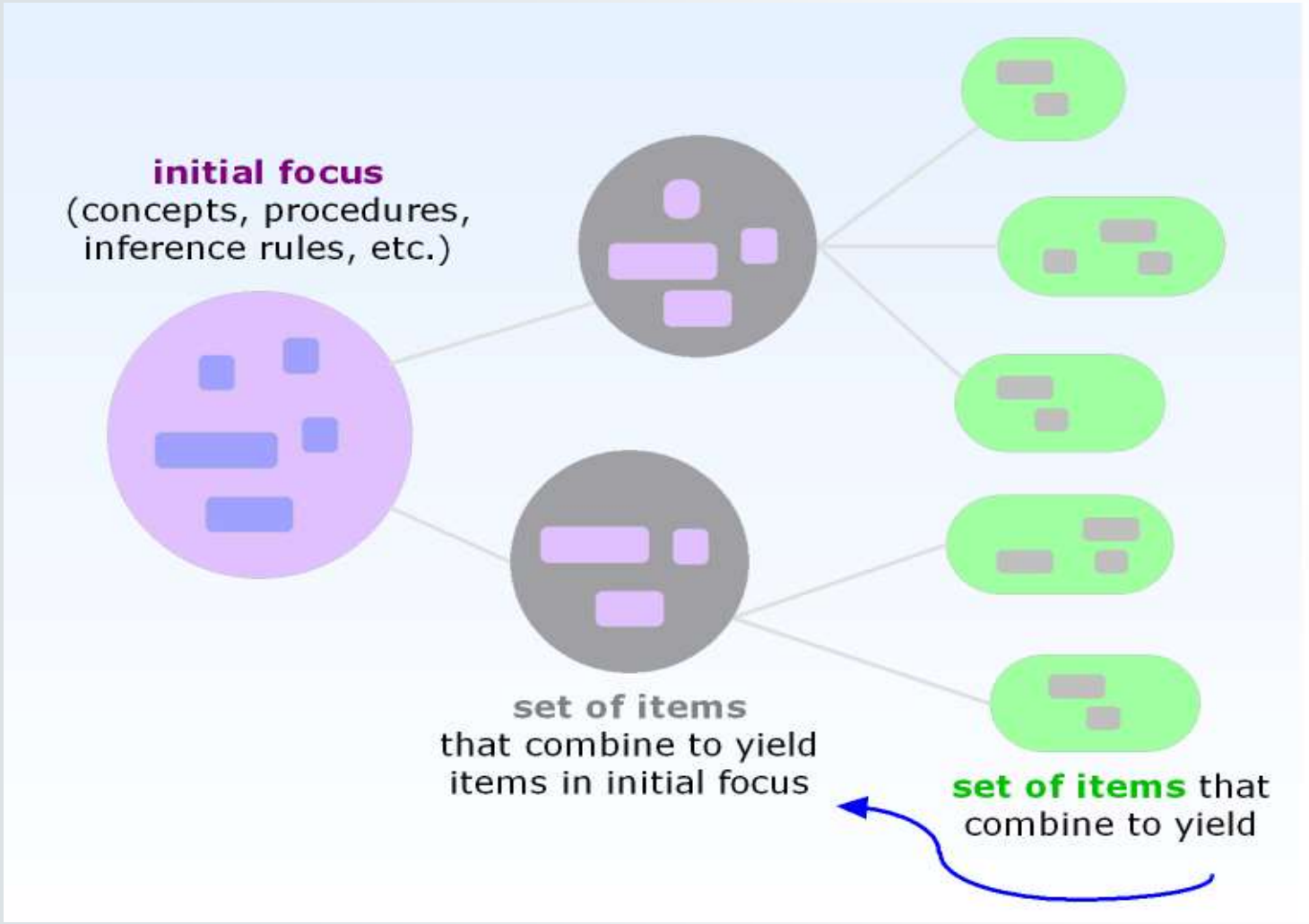
Novamente contains multiple heuristics for Atom creation, including "blending" of existing Atoms



Atoms associated in a dynamic “map” may be grouped to form new Atoms: the Atomspace hence *explicitly representing patterns in itself*



Backward Synthesis



Backward Synthesis Processes

- **Backward-chaining probabilistic inference**
  - Given a target Atom, find ways to produce and evaluate it logically from other knowledge
- **Inference process adaptation**
  - Given a set of inferential conclusions, find ways to produce those conclusions more effectively than was done before
- **Predicate Schematization**
  - Given a goal, and knowledge about how to achieve the goal, synthesize a procedure for achieving the goal
- **Credit Assignment**
  - Given a goal, figure out which procedures' execution, and which Atoms' importance, can be expected to lead to the goal's achievement
- **Goal Refinement**
  - Given a goal, find other (sub)goals that imply that goal


Novamente - AGI Sim
File
View
Help
Goals
0.5
Notice novelty
0.5
Gather knowledge
0.5
Create and consolidate abstract knowledge
Simulation
Resources
45% - Goal-driven Inference
35% - Sensor Processing
20% - Background Inference
Log
Teacher: Where is the toy?
Teacher: Open it
Teacher: Good! The toy is there!
Submit
?
Reward
Correct

(Partial) PLN Backward-Chaining Inference Trajectory for Piagetan A-not-B Problem

Target:

Eval found_under(toy_6,$1)

Step 1

ANDRule:

Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)
|-
AND <1.00, 0.98>
Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)

Step 2

Unification:

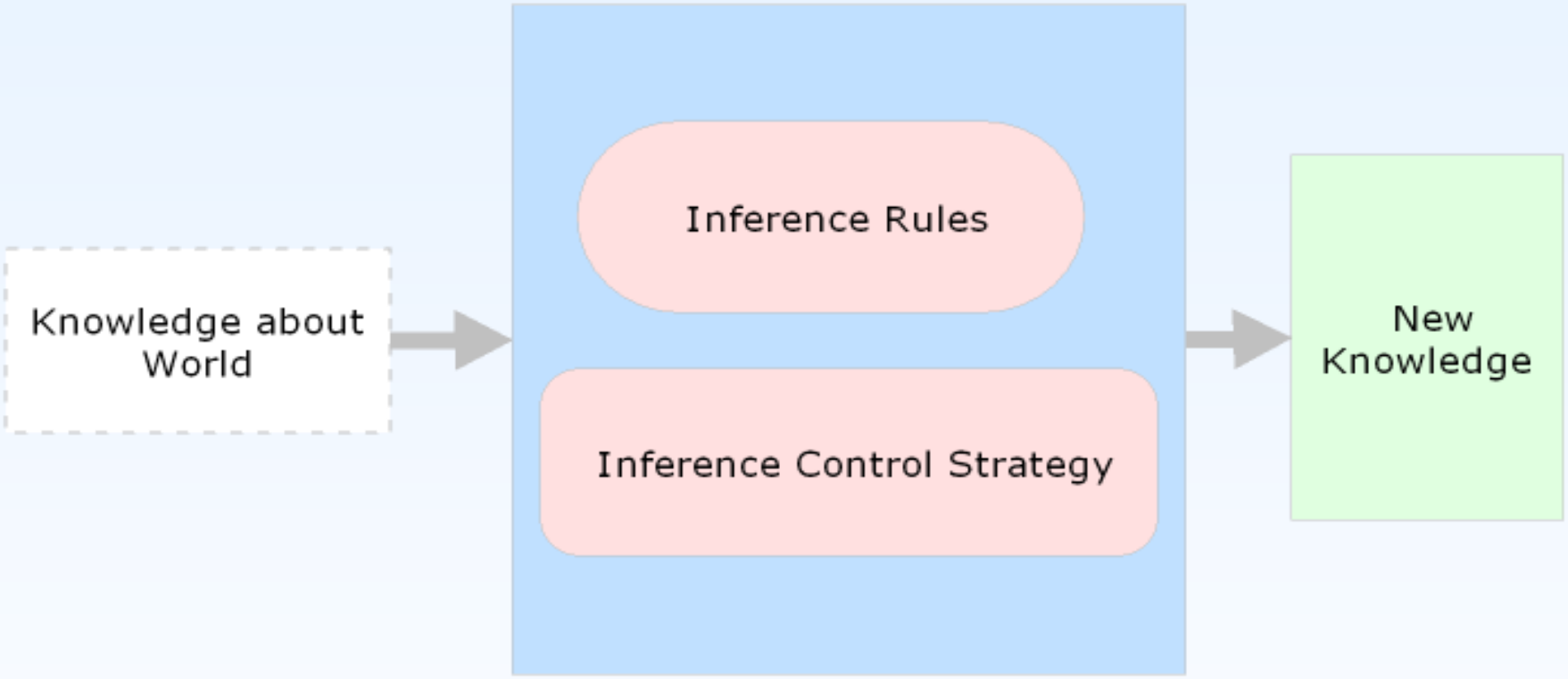
Imp <1.00, 0.95>
AND
Inh($t,toy)
Inh($b,bucket)
Eval placed_under($t,$b)
Eval found_under($t,$b)
AND
Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)
|-
Imp <1.00, 0.94>
AND
Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)
Eval found_under(toy_6,red_bucket_6)

Step 3

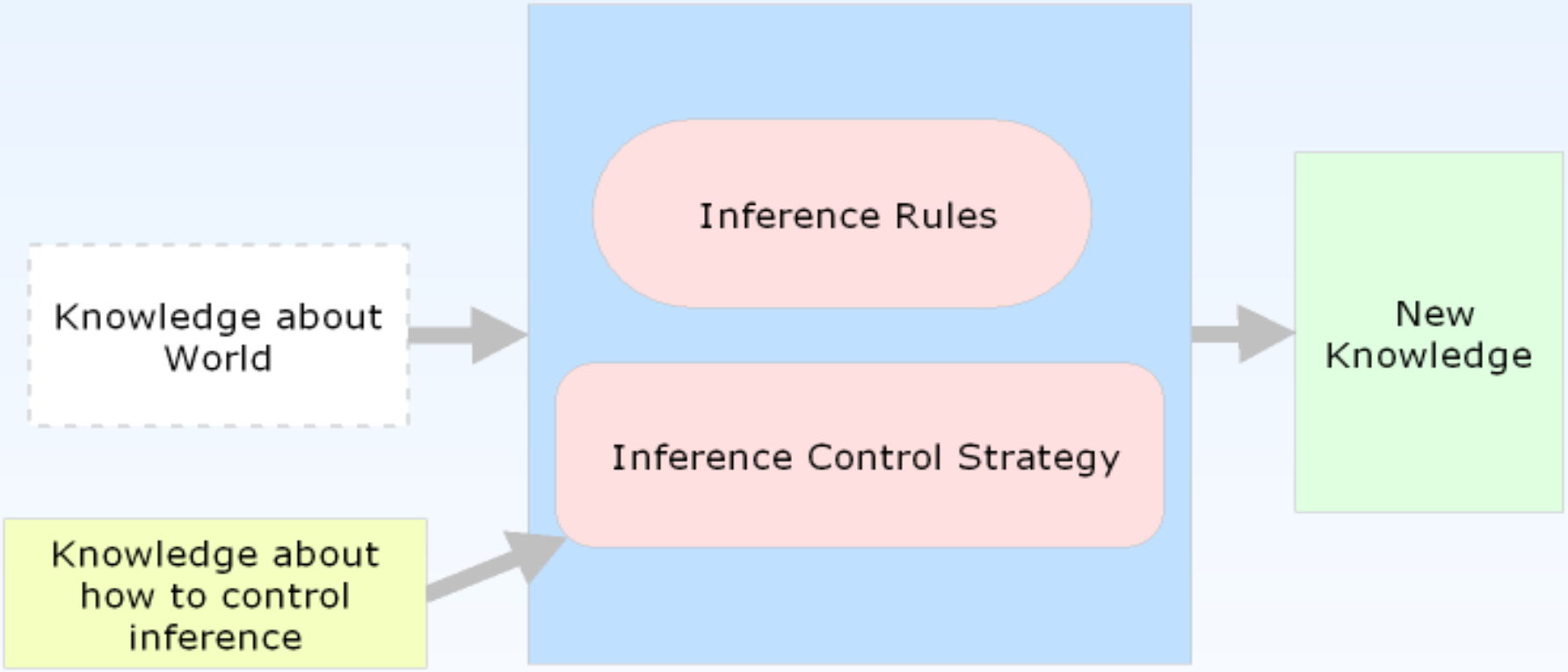
Modus Ponens

Imp <1.00, 0.94>
AND
Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)
Eval found_under(toy_6, red_bucket_6)
AND <1.00, 0.98>
Inh (toy_6,toy)
Inh (red_bucket_6,bucket)
Eval placed_under(toy_6,red_bucket_6)
|-
Eval found_under(toy_6, red_bucket_6) <1.00, 0.93>

Stage 1: Infantile

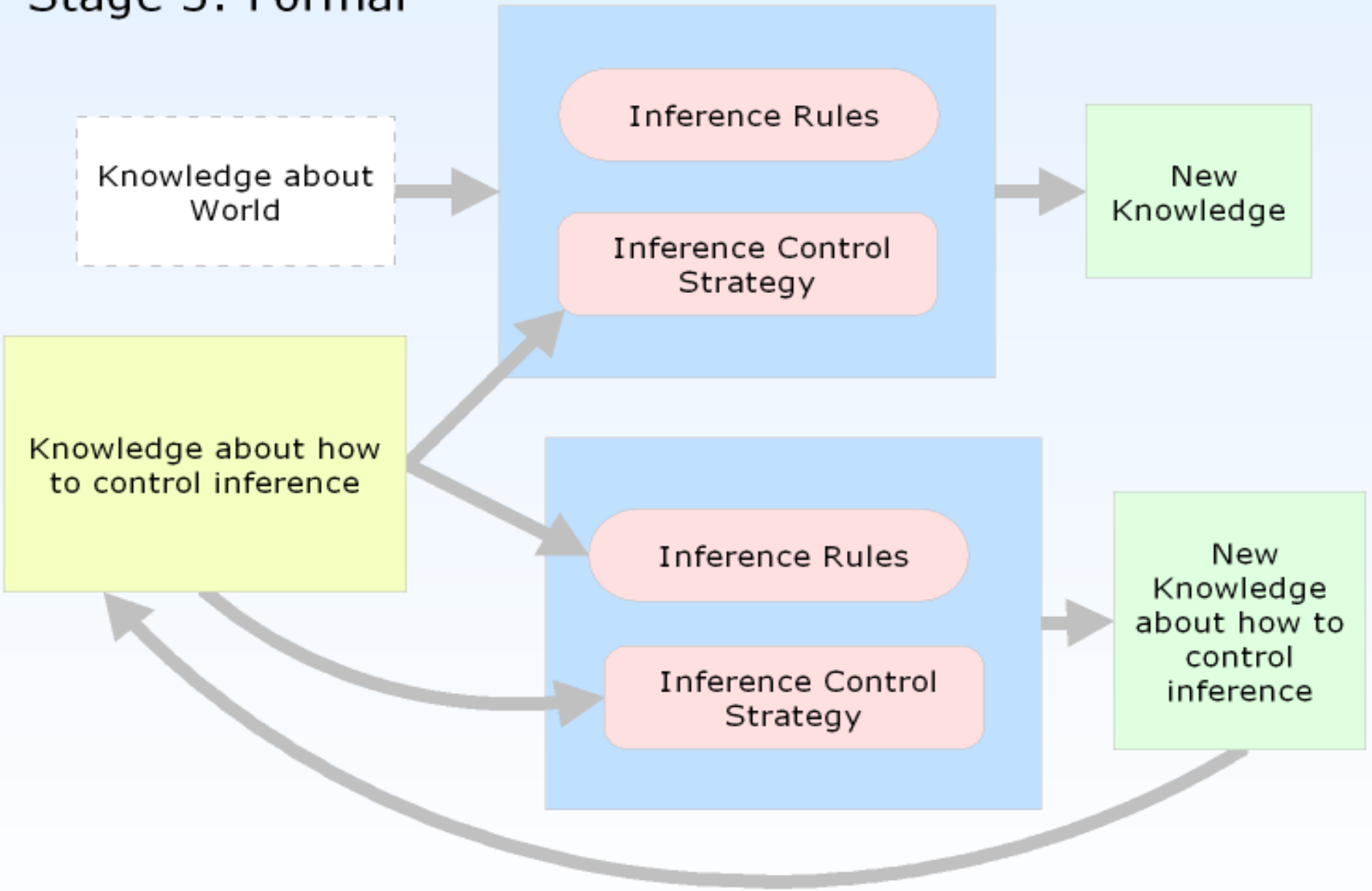


Stage 2: Concrete Operational

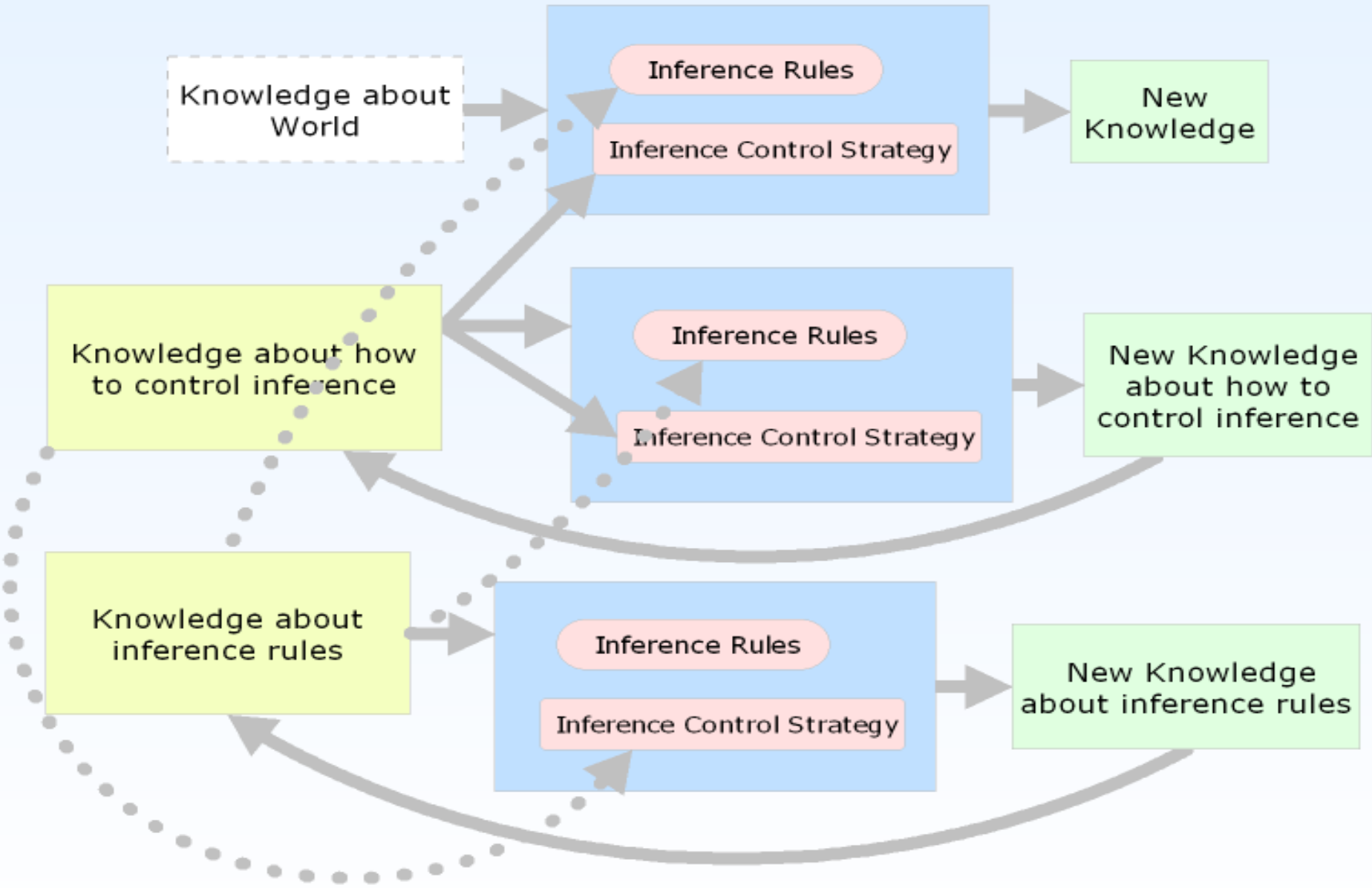


The system may study its own inference history to figure out inference control patterns that would have let it arrive at its existing knowledge more effectively. This is a type of backward synthesis that may lead to powerful iterative self-improvement.

Stage 3: Formal



Stage 4: Reflexive (Post-formal)



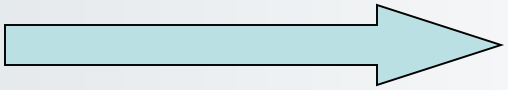
Predicate Schematization

Logical knowledge

EvPredImp <0.95, 0.3>
    Execution try(goto box)
    Eval near box

SimultaneousImplication
    Eval near box
    Eval can_do(push box)

EvPredImp <0.6,0.4>
  And
    Eval can_do(push box)
    Execution try(push box)
  Evaluation Reward



Predicate schematization

Executable procedure

repeat
    goto box
    near box
repeat
    push box
    Reward

(More)

Backward Synthesis Processes

- **Model-Based Predicate Generation**
  - Given probabilistic knowledge about what patterns characterize predicates or procedures satisfying a certain criterion, generate new predicate/procedures satisfying the criterion
- **Criterion-Based Predicate Modeling**
  - Building of probabilistic knowledge regarding the patterns characterizing predicates satisfying a certain criterion

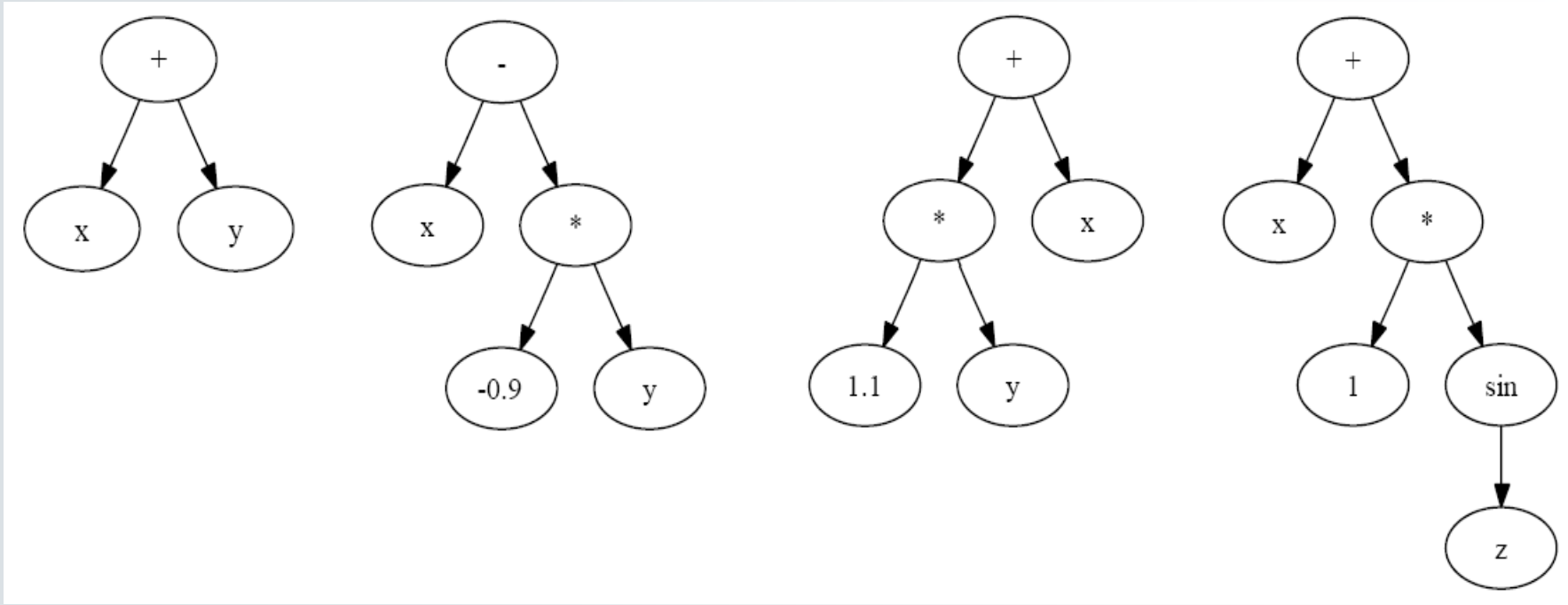
As shown in Moshe Looks' PhD thesis work, the combination of the above two processes may play the role of evolutionary programming, but with dramatically better performance on many problem cases, and an enhanced capability to carry out learning across multiple fitness functions (criteria).

MOSES: Meta-Optimizing Semantic Evolutionary Search

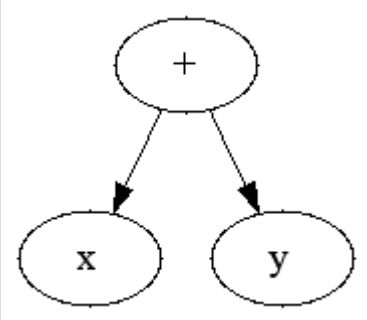
Bringing evolutionary programming and probabilistic inference together

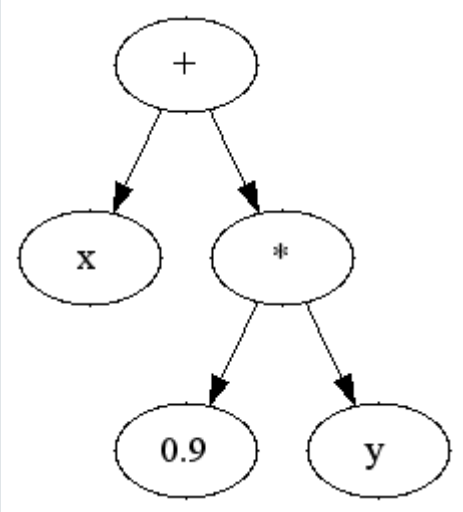
- MOSES evolved out of BOA Programming, which was an extension to program tree learning of the Bayesian Optimization Algorithm approach to probabilistic evolutionary learning
- May be fully integrated with PLN backward chaining inference as a special kind of "backward synthesis process"
  - Integration currently incomplete, to be completed in 2007
- Algorithm:
  - a population of procedure/predicate trees are evaluated
  - the best ones are simplified and normalized …
  - … and modeled probabilistically (Criterion-Based Predicate Modeling)
  - Then new trees are generated via instance generation based on these probabilistic models (Model-Based Predicate Generation)
- Moshe Looks PhD Thesis 2006, Washington University, St. Louis
  - www.metacog.org

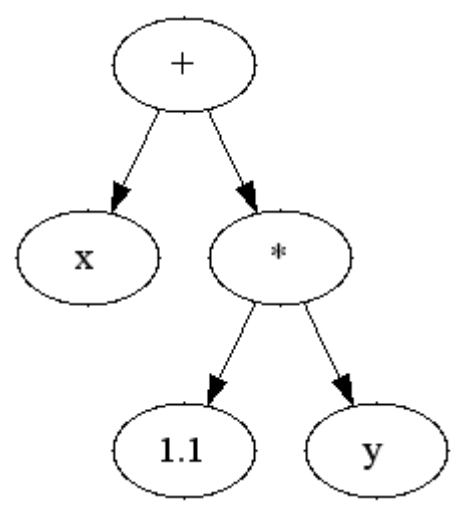
Simple Example: A MOSES Population of Arithmetic Procedures

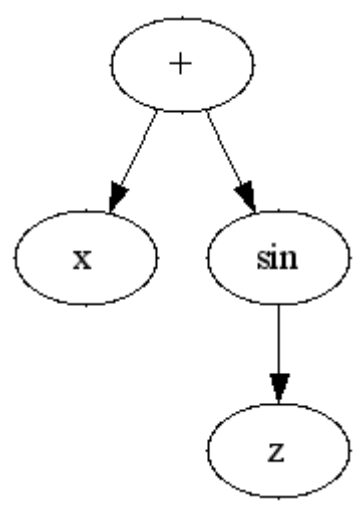


Simplification & Normalization

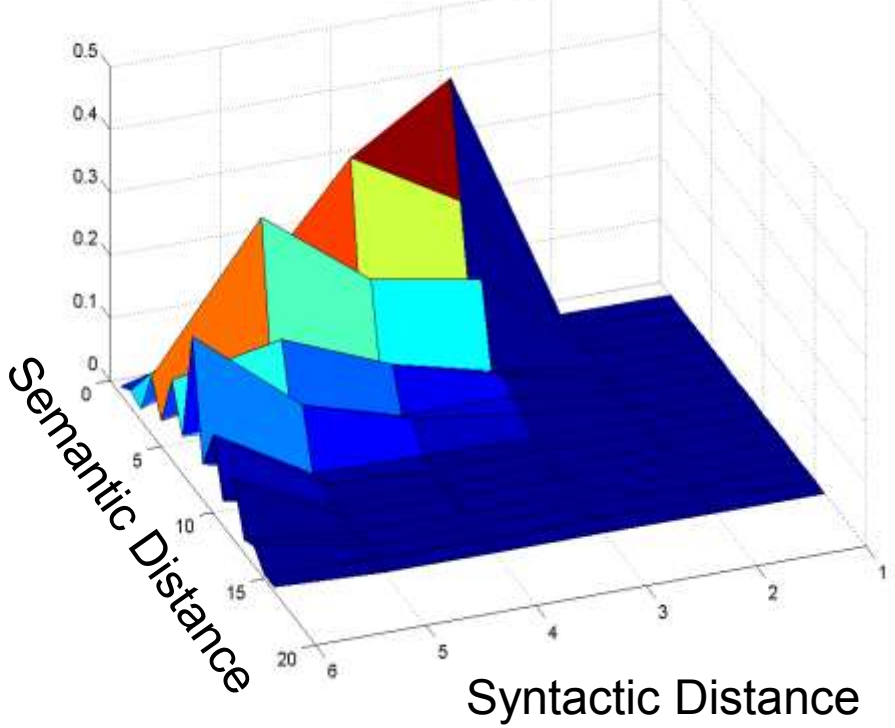








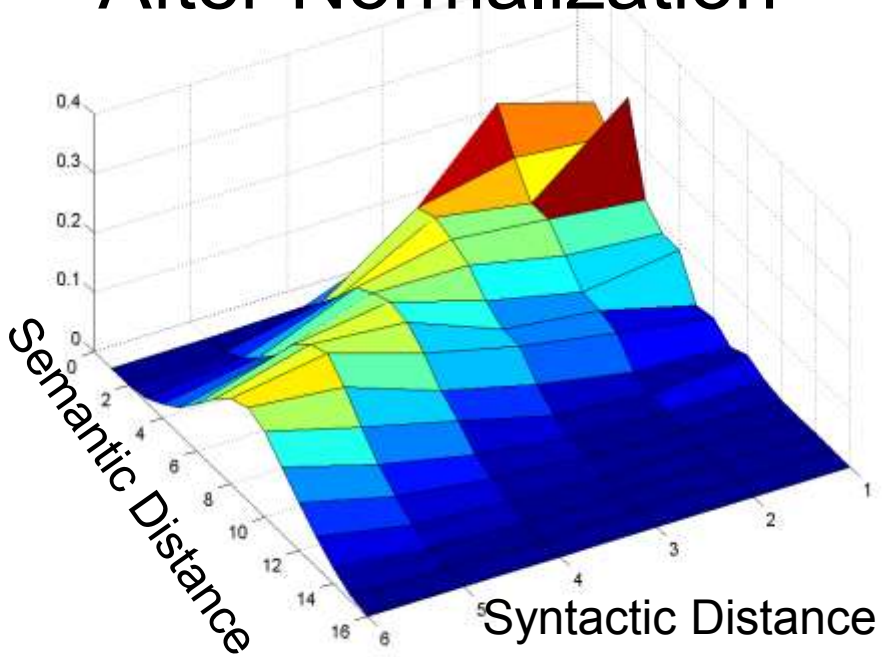
Before Normalization



Normalization of procedure/predicate trees harmonizes syntactic form with semantic meaning (I/O behavior)

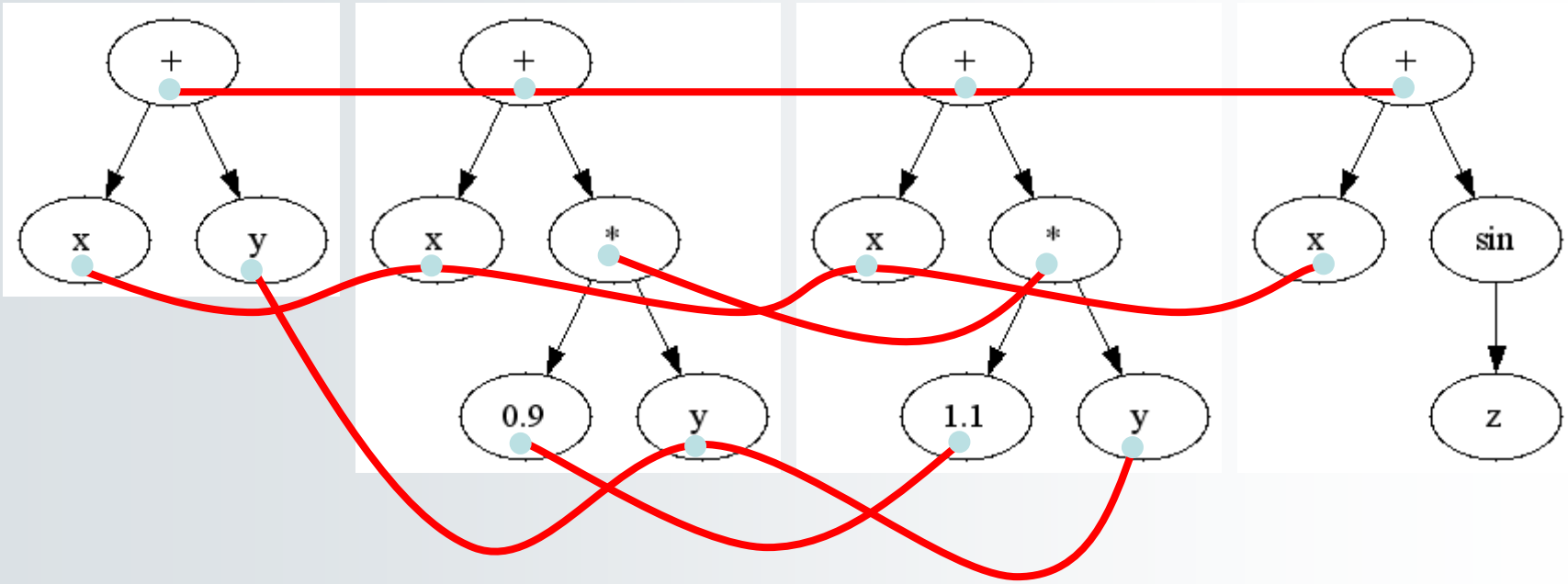
Graphs based on Boolean predicates; same phenomenon holds more generally

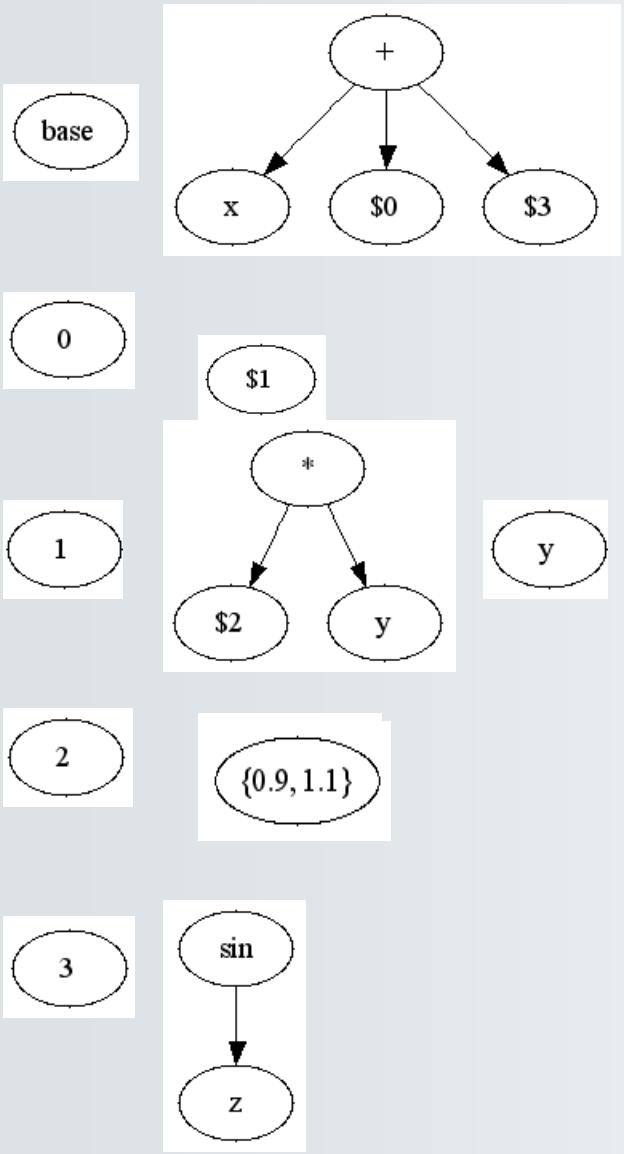
After Normalization

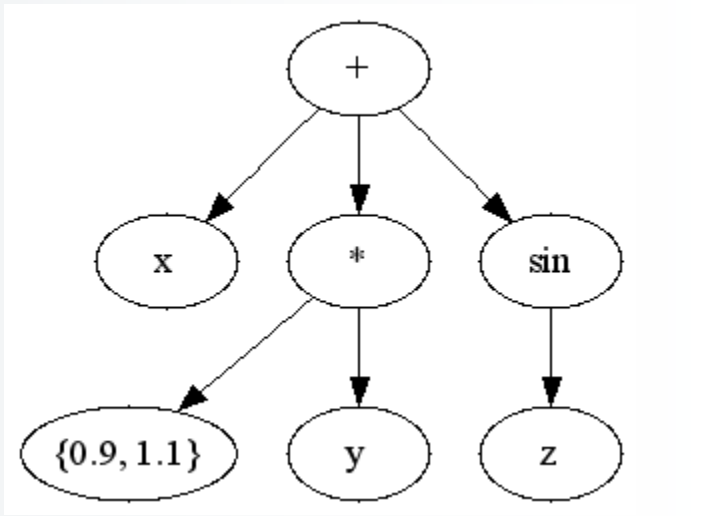


Alignment

(Recognizing common patterns)








```
    +(x,$0,$3)
0:  nil | $1
1:  *($2,y) | y
2:  {0.9,1.1}
3:  nil | sin(z)
```

Abstract trees (predicates) are created from the population of concrete ones

Example: MOSES learns program to play “fetch” in AGISim

```
ifelse
  holding
  ifelse

  facingteacher
    step
    rotate
  ifelse
    nearball
    pickup
    ifelse

  facingball
      step
      rotate
```

(More)

Backward Synthesis Processes

Language Comprehension

- **Syntax parsing**: given a sentence, or other utterance, search for assignments of syntactic relationships to words that will make the sentence grammatical
- **Semantic mapping**: Search for assignment of semantic meanings to words and syntactic relationships that will make the sentence contextually meaningful

Lojban / Lojban++

- Lojban is a constructed language with syntax and semantics founded on predicate logic
- Lojban++ is a variant of Lojban that incorporates English content words in certain roles
- In these languages, ambiguity is minimized relative to natural languages
- Parsing Lojban/++ is automatic and mechanical
- Semantic mapping into predicate logic is also fully mechanical -- but some contextual disambiguation of predicates may still be required

Lojban / Lojban++

English	I eat the salad with croutons
Lojban	mi citka le salta poi mixre lo sudnabybli
Lojban++	mi eat le salad poi mixre lo crouton mi eat le salad poi contain lo crouton

English	I eat the salad with a fork
Lojban	mi citka le salta sepi'o lo forca
Lojban++	mi eat le salad sepi'o lo fork

Lojban++

le dog pe mi uncle cu stupid

Needs contextual disambiguation

EvaluationLink stupid $D

InheritanceLink $D dog

AssociationLink $D $U

EvaluationLink uncle($U, Ben_Goertzel)

Holistic Cognitive Dynamics and Emergent Mental Structures

The Fundamental Cognitive Dynamic

Let X = any set of Atoms

Let F(X) = a set of Atoms which is the result of forward synthesis on X

Let B(X) = a set of Atoms which is the result of backward synthesis of X -- assuming a heuristic biasing the synthesis process toward simple constructs

Let S(t) denote a set of Atoms at time t, representing part of a system's knowledge base

Let I(t) denote Atoms resulting from the external environment at time t

$$S(t+1) = B(F(S(t) + I(t)))$$

The Fundamental Cognitive Dynamic

$$S(t+1) = B(F(S(t) + I(t)))$$

Forward: create new mental forms by combining existing ones

Backward: seek simple explanations for the forms in the mind, including the newly created ones. The explanation itself then comprises additional new forms in the mind

Forward: …

Backward: …

Etc.

… Combine … Explain … Combine … Explain … Combine …

The Construction and Development of the Emergent Pattern that is the "Phenomenal Self"

The self originates (and ongoingly re-originates) via backward synthesis

Backward chaining inference attempts to find models that will explain the observed properties of the system itself

The self develops via forward synthesis

Aspects of self blend with each other and combine inferentially to form new Atoms

These new Atoms help guide behavior, and thus become incorporated into the backward-synthesis-derived self-models

Self = A strange attractor of the Fundamental Cognitive Dynamic

The Construction and Development of the Emergent Pattern that is "Focused Consciousness"

- Atoms in the "moving bubble of importance" consisting of the Atoms with highest Short-Term Importance are continually combining with each other, forming new Atoms that in many cases remain highly important

- Sets of Atoms in the moving bubble of importance are continually subjected to backward synthesis, leading to the creation of compact sets of Atoms that explain/produce them -- and these new Atom-sets often remain highly important

Focused Consciousness = A strange attractor of the Fundamental Cognitive Dynamic

Why Will Novamente Succeed Where Other AGI Approaches Fail?

- Only Novamente is based on a *well-reasoned, truly comprehensive theory of mind*, covering both the concretely-implemented and emergent aspects
- The specific algorithms and data structures chosen to implement this theory of mind are efficient, robust and scalable
- So is the software implementation!

More specifically: Only in the Novamente design is the fundamental cognitive dynamic implemented in a powerful and general enough way adequate to give rise to self and focused consciousness as strange attractors.

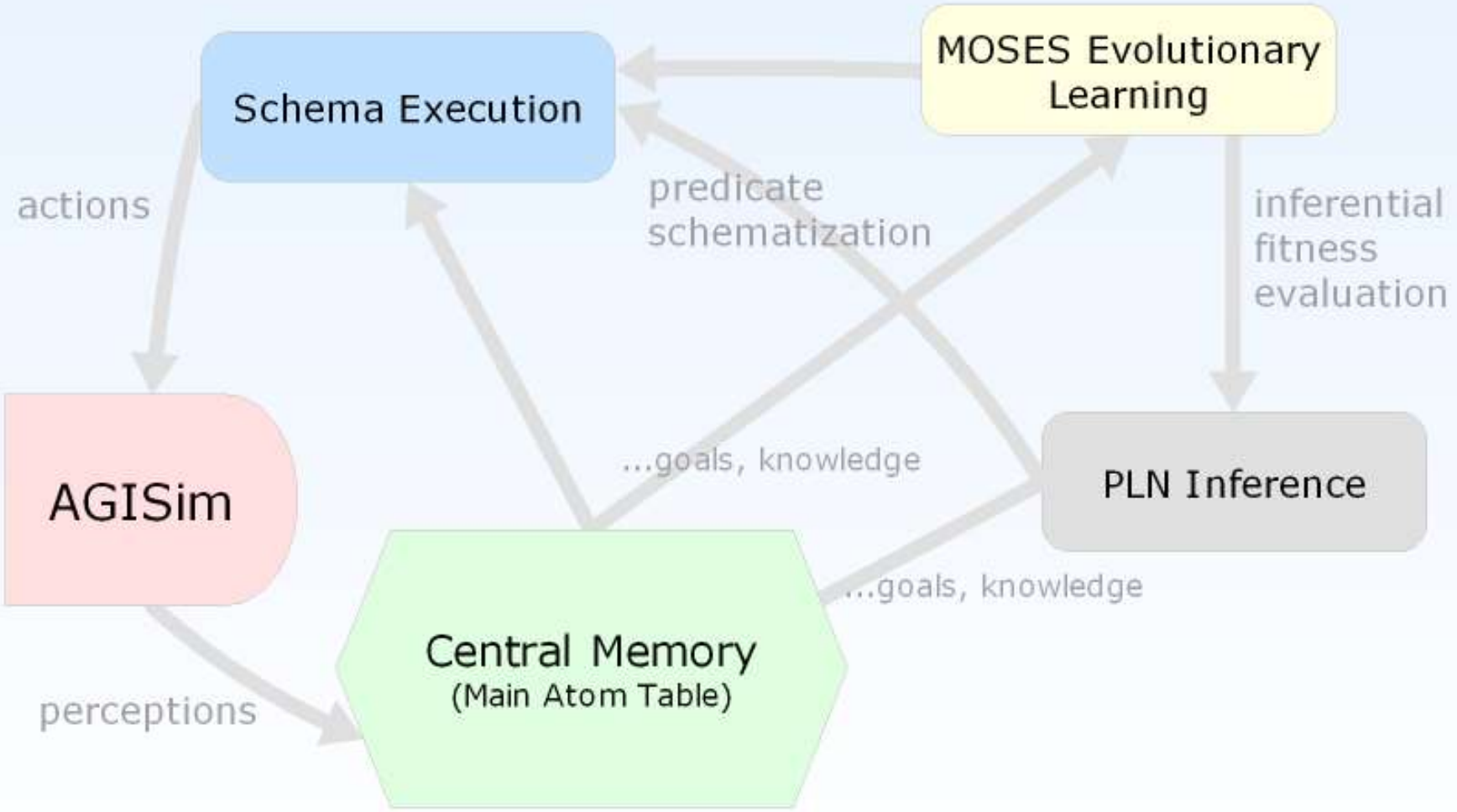
Novamente:

The Current Reality

Implemented Components

- **Novamente core system**
  - AtomTable, MindAgents, Scheduler, etc.
  - Now runs on one machine; designed for distributed processing
- **PLN**
  - Relatively crude inference control heuristics
  - Simplistic predicate schematization
- **MOSES**
  - Little experimentation has been done evolving procedures with complex control structures
  - Not yet fully integrated with PLN
- **Schema execution framework**
  - Enacts learned procedures
- **AGISim**
  - And proxy for communication with NM core
- **NLP front end**
  - External NLP system for “cheating” style knowledge ingestion

Current AGISim Learning Architecture

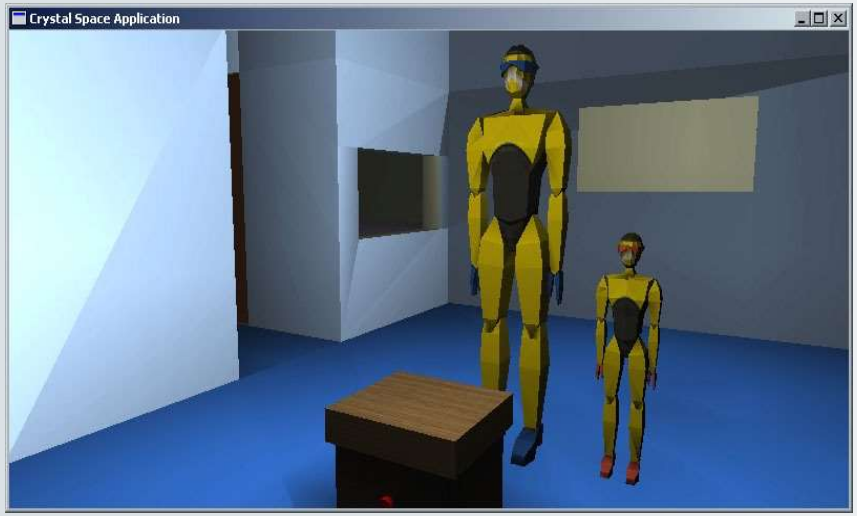


Simple, Initial AGISim Experiments

- Fetch
- Tag
- Piagetan A-not-B experiment
- Word-object association

Goal For Year One After Project Funding

Fully Functional Artificial Infant



Able to learn infant-level behaviors "without cheating" -- i.e. with the only instruction being interactions with a human-controlled agent in the simulation world

Example behaviors: naming objects, asking for objects, fetching objects, finding hidden objects, playing tag

System will be tested using a set of tasks derived from human developmental psychology

Within first 9 months after funding we plan to have the most capable autonomous artificial intelligent agent created thus far, interacting with humans spontaneously in its 3D simulation world in the manner of a human infant

NOVAMENTE LLC
ARTIFICIAL GENERAL INTELLIGENCE

Teaching the Baby Language

Artificial Infant + Narrow-AI NLP System =

AGI system capable of learning complex natural language

(Narrow-AI NLP system as “scaffolding”)

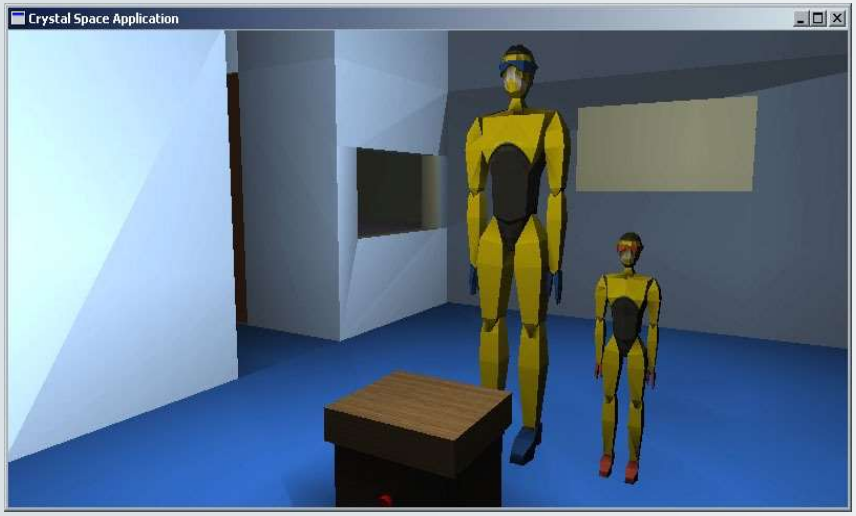
Narrow-AI NLP System =

Novamente’s RelEx English semantic analysis engine + a Lojban++ parser

(Parallel instruction in English and Lojban++ will accelerate learning dramatically)

Goal For Year Two After Project Funding

Artificial Child with Significant Linguistic Ability



Ability to learn from human teachers via linguistic communication utilizing complex recursive phrase structure grammar and grounded semantics

Linguistic instruction will be done simultaneously in English and in the constructed language Lojban++, which maps directly into formal logic

At this stage, the symbol groundings learned by the system will be commercially very valuable, and will be able to dramatically enhance the performance of natural language question answering products

NOVAMENTE LLC
ARTIFICIAL GENERAL INTELLIGENCE

Acknowledgements

The Novamente Team

- Bruce Klein – President, Novamente LLC
- Cassio Pennachin – Chief Architect, Novamente AI Engine
- Andre Senna – CTO
- Ari Heljakka – Lead AI Engineer
- Moshe Looks – AI Engineer
- Izabela Goertzel– AI Engineer
- Murilo Queiroz – AI Engineer
- Welter Silva – System Architect
- Dr. Matthew Ikle' – Mathematician

[2006 AGIRI.org Workshop](#)
Sponsored by Novamente LLC)

NOVAMENTE LLC
ARTIFICIAL GENERAL INTELLIGENCE

Thank You