NARS

an Artificial General Intelligence (AGI) Project

Peter Isaev & Patrick Hammer

Temple University

What is not Intelligent

- Innate behavior, or instinct
- Exhaustive search
- Information retrieval
- Repeated routines
- Algorithm following numerical calculation, sorting, fixed mapping, …

“INTELLIGENCE” INTERPRETED

- Mainstream AI treats “Intelligence” as a collection of problem-specific and domain-specific parts
- AGI takes “Intelligence” as a general-purpose capability that should be treated as a whole
- AGI research still includes different research objectives and strategies

BASIC ASSUMPTION

"Intelligence" is *the capability of a system to adapt to its environment and to work with insufficient knowledge and resources*

Assumption of Insufficient Knowledge and Resources (AIKR):

- To rely on *finite* processing capacity
- To work in *real time*
- To *open* to unexpected tasks

REASONING SYSTEM FRAMEWORK

- a *language* for representation
- a *semantics* of the language
- a set of inference *rules*
- a *memory* structure
- a *control* mechanism

Advantages:

- domain independence
- rich expressing power
- justifiability of the rules
- flexibility in combining the rules

FUNDAMENTAL ISSUES

Under AIKR, the system cannot guarantee absolute correctness or optimum anymore. Now what is the standard of *validity* or *rationality*?

Validity and rationality become *relative* to the available knowledge and resources.

Desired features: *general, adaptive, flexible, robust, scalable*

Non-Axiomatic Reasoning System

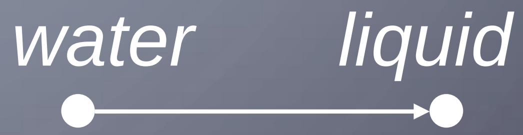
- NARS has a logic part and a control part, with a “logic” in the original sense
- NARS is fully based on AIKR
- NARS is a normative model built on a descriptive foundation
- NARS has a designed meta-level and an acquired object-level

TERM AND STATEMENT

Term: word, as name of a concept

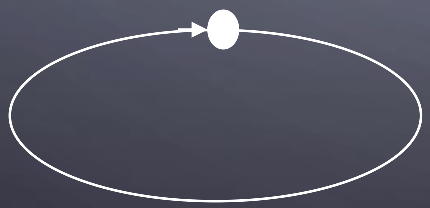
Statement: subject-copula-predicate

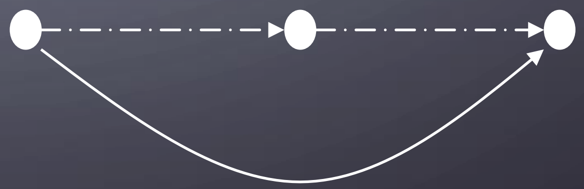
S → P



as specialization-generalization

Copula *inheritance* is reflexive and transitive





BINARY TRUTH-VALUE

- Experience ***K***: a finite set of statements
- Beliefs ***K****: the transitive closure of ***K***
- A statement is *true* if

  either it is in ***K****

  or it has the form of $X \rightarrow X$

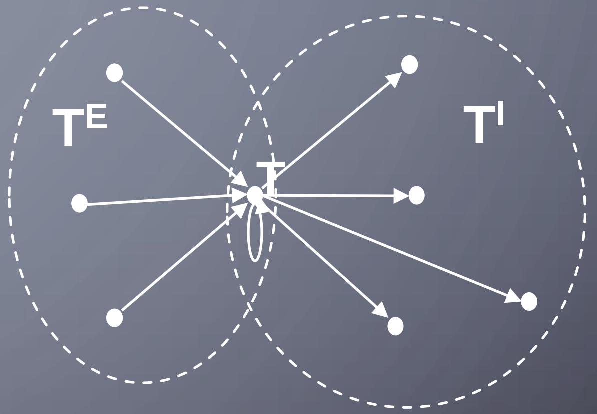
otherwise it is *false*

EXTENSION AND INTENSION

For a given term T,

its *extension* $T^E = \{x \mid x \rightarrow T\}$

its *intension* $T^I = \{x \mid T \rightarrow x\}$



Theorem:

$$(S \rightarrow P) \Leftrightarrow (S^E \subseteq P^E) \Leftrightarrow (P^I \subseteq S^I)$$

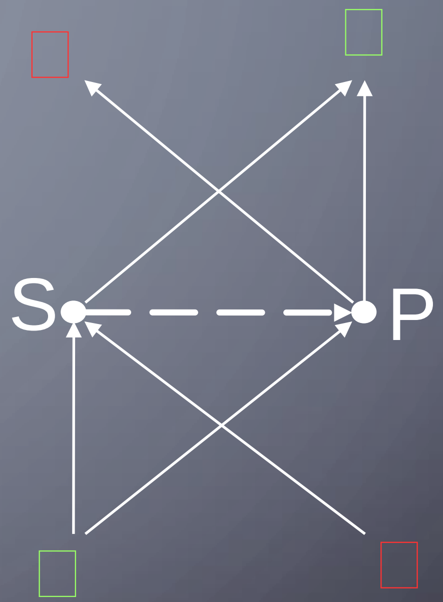
EVIDENCE

Positive evidence of $S \rightarrow P$:

$\{x \mid x \in (S^E \cap P^E) \cup (P^I \cap S^I)\}$

Negative evidence of $S \rightarrow P$:

$\{x \mid x \in (S^E - P^E) \cup (P^I - S^I)\}$



Amount of evidence:

positive: $w^+ = |S^E \cap P^E| + |P^I \cap S^I|$

negative: $w^- = |S^E - P^E| + |P^I - S^I|$

total: $w = w^+ + w^- = |S^E| + |P^I|$

Meaning of Truth

- Correspondence: the “truth-value” of a statement measures its agreement with the reality, how close it is to an objective fact

- Experience-grounded: the “truth-value” of a statement measures its evidential support, indicates how close it is to the evidence

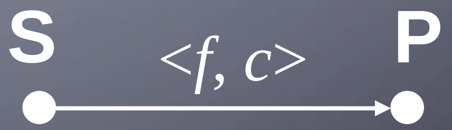
TRUTH-VALUE DEFINED

In NARS, the truth-value of a statement is a pair of real numbers in [0, 1], and measures the evidential support to the statement.

$S \rightarrow P <f, c>$

frequency: $f = w^+/w$

confidence: $c = w / (w + 1)$



TRUTH-VALUE PRODUCED

- Actual experience: a stream of statements with truth-value, where the confidence is in (0, 1)
- Each inference rule has a truth-value function, and the truth-value of the conclusion is determined only by the evidence provided by the premises

TRUTH-VALUE FUNCTION DESIGN

1. Treat all involved variables as Boolean

2. For each value combination in premises, decide the values in conclusion

3. Build Boolean functions among the variables

4. Extend the operators to real-number:

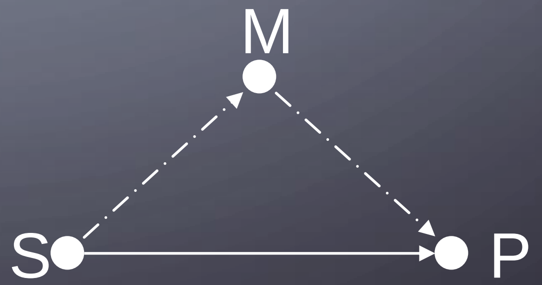
$$\text{not}(x) = 1 - x$$
$$\text{and}(x, y) = x * y$$
$$\text{or}(x, y) = 1 - (1 - x) * (1 - y)$$

DEDUCTION

$$\frac{\begin{array}{l} M \rightarrow P\ [f_1, c_1] \\ S \rightarrow M\ [f_2, c_2] \end{array}}{S \rightarrow P\ [f, c]}$$

$$\frac{\begin{array}{l} \text{bird} \rightarrow \text{animal}\ [1.00, 0.90] \\ \text{robin} \rightarrow \text{bird}\ \ [1.00, 0.90] \end{array}}{\text{robin} \rightarrow \text{animal}\ [1.00, 0.81]}$$

$$f = f_1 * f_2$$
$$c = c_1 * c_2 * f_1 * f_2$$



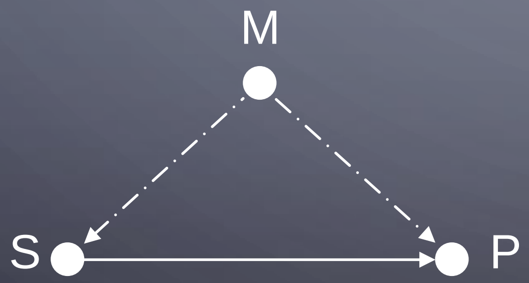
INDUCTION

$$\frac{\begin{array}{l} M \rightarrow P\ [f_1, c_1] \\ M \rightarrow S\ [f_2, c_2] \end{array}}{S \rightarrow P\ [f, c]}$$

$f = f_1$

$c = f_2 * c_1 * c_2 \ / \ (f_2 * c_1 * c_2 + 1)$

$$\frac{\begin{array}{ll} \text{swan} \rightarrow \text{bird} & [1.00, 0.90] \\ \text{swan} \rightarrow \text{swimmer} & [1.00, 0.90] \end{array}}{\text{bird} \rightarrow \text{swimmer}\ [1.00, 0.45]}$$

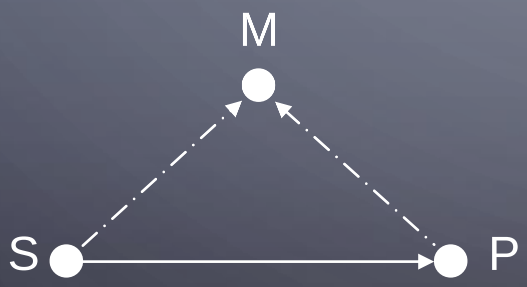


ABDUCTION

$$\frac{\begin{array}{c} P \rightarrow M\ [f_1, c_1] \\ S \rightarrow M\ [f_2, c_2] \end{array}}{S \rightarrow P\ [f, c]}$$

$$f = f_2$$

$$c = f_1 * c_1 * c_2 \,/\, (f_1 * c_1 * c_2 + 1)$$



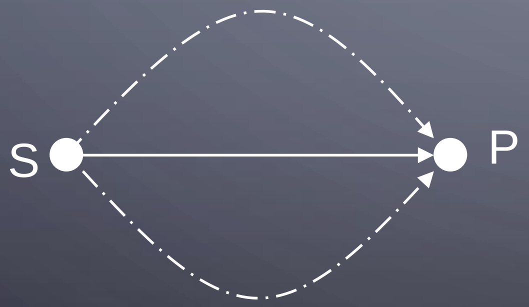
$$\frac{\begin{array}{r} \text{seabird} \rightarrow \text{swimmer}\ [1.00, 0.90] \\ \text{gull} \rightarrow \text{swimmer}\ [1.00, 0.90] \end{array}}{\text{gull} \rightarrow \text{seabird}\ \ [1.00, 0.45]}$$

REVISION

$$\frac{S \rightarrow P\ [f_1, c_1] \quad S \rightarrow P\ [f_2, c_2]}{S \rightarrow P\ [f, c]}$$

$$f = \frac{f_1 * c_1 * (1 - c_2) + f_2 * c_2 * (1 - c_1)}{c_1 * (1 - c_2) + c_2 * (1 - c_1)}$$

$$c = \frac{c_1 * (1 - c_2) + c_2 * (1 - c_1)}{c_1 * (1 - c_2) + c_2 * (1 - c_1) + (1 - c_2) * (1 - c_1)}$$



bird → swimmer [1.00, 0.62]
bird → swimmer [0.00, 0.45]

---

bird → swimmer [0.67, 0.71]

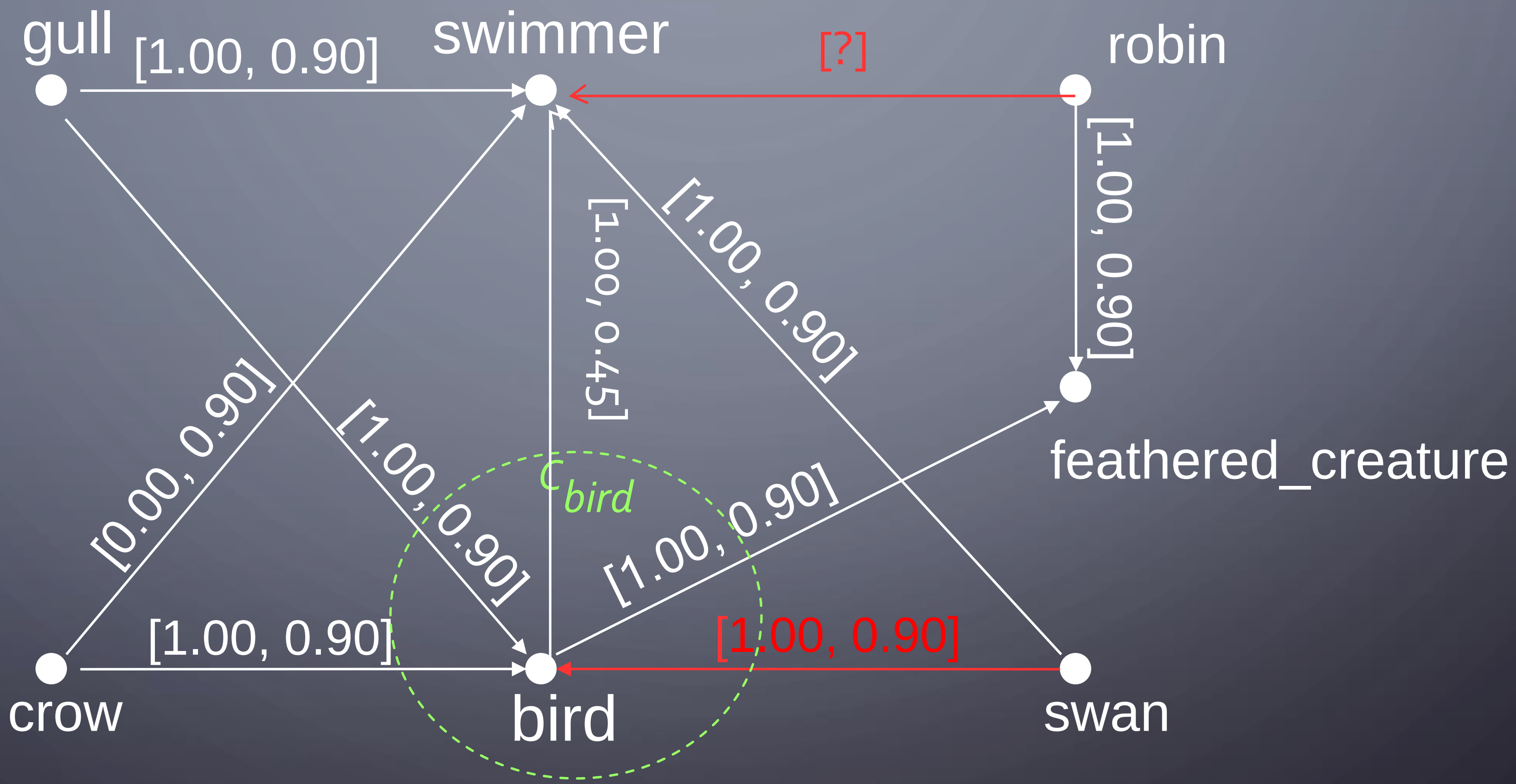
TYPES OF INFERENCE

- **Local Inference**: revising beliefs or choosing an answer for a question
- **Forward inference**: from existing beliefs to new beliefs (deduction, induction, abduction, …)
- **Backward inference**: from existing questions and beliefs and to derived questions

MEMORY STRUCTURE

- A *task* is either a question or a piece of new knowledge
- *A belief* is accepted knowledge
- The tasks and beliefs are clustered into *concepts*, each named by a *term*
- Concepts are prioritized in the memory; tasks and beliefs are prioritized within each concept

MEMORY AS A NETWORK

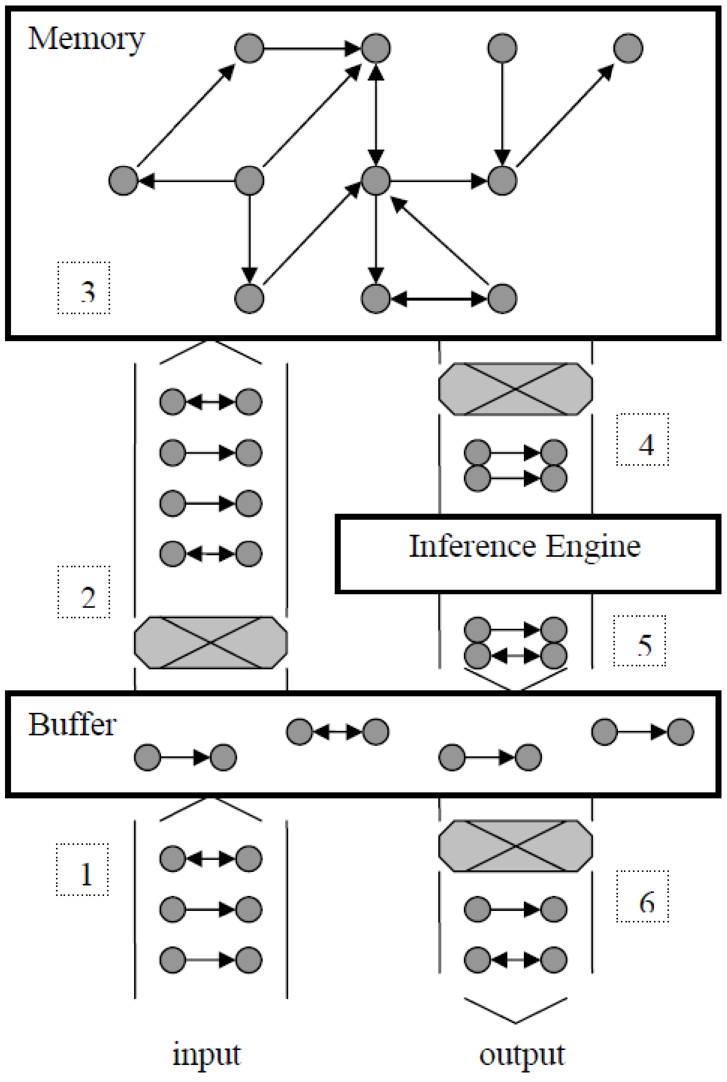


MEANING OF CONCEPT

Every concept in NARS is *fluid*:

- Its meaning is determined neither by reference nor definition, but by experienced relations
- Each relation is a matter of degree
- Meaning changes by history and context

ARCHITECTURE AND ROUTINE



1. Input tasks are added into the task buffer.

2. Selected tasks are inserted into the memory.

3. Inserted tasks in memory may also produce beliefs and concepts, as well as change existing ones.

4. In each working cycle, a task and a belief are selected from a concept, and feed to the inference engine as premises.

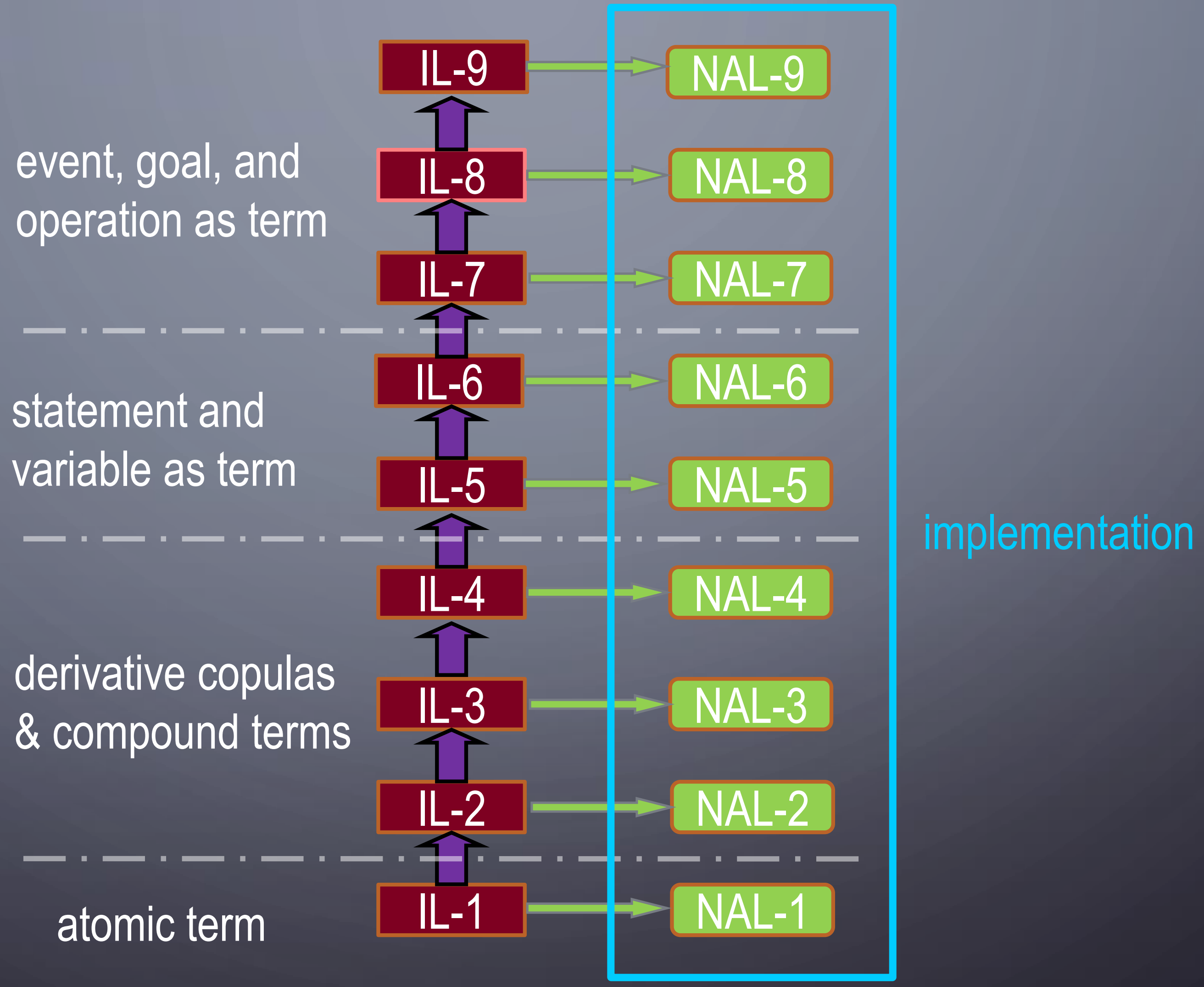
5. The conclusions derived from the premises by applicable rules are added into the buffer as derived tasks.

6. Selected derived tasks are reported as output tasks.

CONTROL STRATEGY

- In each step, a task interacts with a belief according to applicable rules
- The task and belief are selected probabilistically, biased by priority
- Factors influence the priority of an item: its quality, its usefulness in history, and its relevance to the current context

THE LAYERS OF THE LOGIC



COPULAS & COMPOUND TERMS

Ideas from set theory:

- Variants of the *inheritance* copula: *similarity*, *instance*, and *property*
- Compound terms: *sets*, *intersections*, *differences*, *products*, and *images*
- New inference rules for *comparison*, *analogy*, plus compound-term *composition* and *decomposition*

HIGHER-ORDER REASONING

Ideas from propositional/predicate logic:

- Copulas: *implication* and *equivalence*
- Compound statements: *negation*, *conjunction*, and *disjunction*
- *Conditional inferences as implication*
- Variable terms as symbols

NAL as a universal meta-logic

PROCEDURAL REASONING

Ideas from logic programming:

- *Events* as statements with temporal relations (*sequential* and *parallel*)
- *Operations* as executable events, with a sensorimotor interface
- *Goals* as events to be realized
- *Mental operations* are integrated into the inference process

UNIFICATIONS IN NARS

- Fully based on AIKR
- Unified representational language
- Complete inferential power
- *Reasoning as learning, planning, problem solving, decision making, …*
- Integrating with other software & hardware via plug-and-play

IMPLEMENTATION

- NARS has been mostly implemented in the open-source project *OpenNARS for research*
- Working examples exist as proof of concept, though only cover simple cases
- The system shows many human-like properties, though it is not a psychological model

POTENTIAL APPLICATIONS

NARS is not designed for any specific application, it can be considered as a general purpose tool

Suitable domains:

- AIKR is applicable
- Tasks expressible as reasoning
- Tools have compatible interface

- Publications & reports: http://www.cis.temple.edu/~pwang/

- Source code, examples, and documents: http://opennars.org/

- Participations and COLLABORATIONS are welcome!