Artificial General Intelligence

3. Basic Inference Rules

Dr. Pei Wang

Dr. Patrick Hammer

Dr. Robert Johansson

Digital Futures (KTH, SU, RISE)

Stockholm, Fall 2022

Non-Axiomatic Logic (NAL)

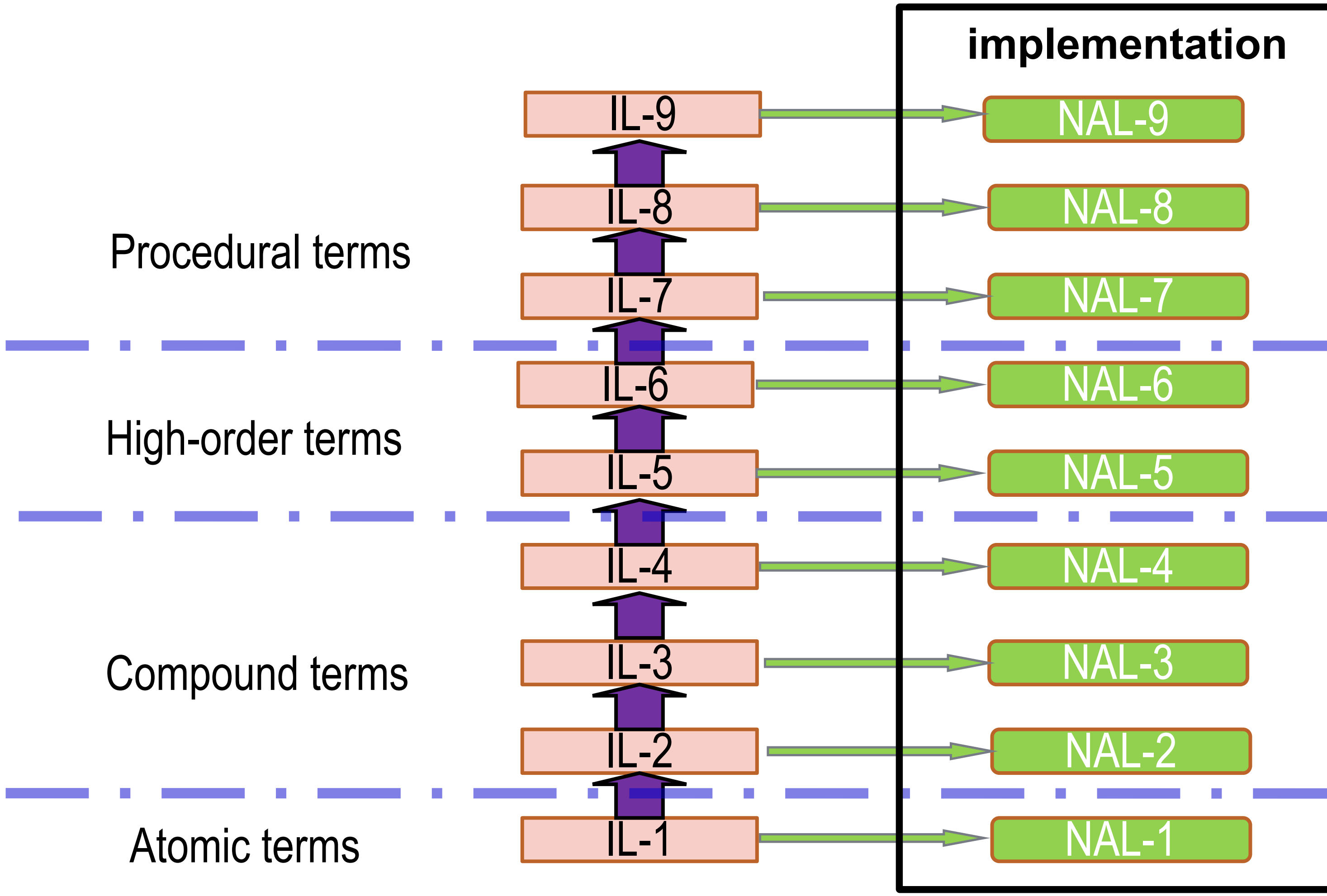
To get a **logic** for *adaptation under AIKR*, we need

- a formal language for knowledge representation
- a set of formal inference rules

To uniformly cover various cognitive functions

- starting from a core, then incrementally adding grammar and inference rules layer-by-layer
- at each layer, starting with an idealized form, then extending the model under AIKR

Constructing NALs



Two traditions of logic

Name	Traditional logic Term logic	Mathematical logic Predicate logic
Founder	Aristotle	Frege
Target	everyday reasoning	theorem proving
Statement	categorical	functional
Inference	syllogism	truth table

Inheritance Logic, Layer 1 (IL-1)

Statement: ‹*subject-term*, *copula*, *predicate-term*›

- Term: internal identifier, a string of character
- Copula: *inheritance*, $\rightarrow$, reflexive and transitive
- Intuitive meaning: specialization-generalization
- Examples: *water* $\rightarrow$ *liquid*, *raven* $\rightarrow$ *bird*

Network: terms as nodes, copulas as links

- Experience: a network (or set) of statement, K
- Knowledge: K^*, the transitive closure of K

Semantics of IL-1

- Truth-value of a statement:
  - it is *true* if it is in K^*, or is $T \longrightarrow T$
  - otherwise, it is *false*
- Meaning of term T : ‹extension T^E, intension T^I›
  - $T^E = \{x \mid x \longrightarrow T\}$, all known specializations of T
  - $T^I = \{x \mid T \longrightarrow x\}$, all known generalizations of T

Theorem 2.4. *If both S and P are in V_K, then*

$$(S \to P) \Longleftrightarrow (S^E \subseteq P^E) \Longleftrightarrow (P^I \subseteq S^I)$$

Semantics of NAL-1

Definition 3.1. *For an inheritance statement "S → P",* its *evidence* are terms in its *evidential scope*, S^E and P^I. Among them, terms in $(S^E \cap P^E)$ and $(P^I \cap S^I)$ are *positive evidence*, and terms in $(S^E - P^E)$ and $(P^I - S^I)$ are *negative evidence*.

- The amounts of positive and negative evidence, w^+ and w^-, are the sizes of the sets, respectively
- The definition of a measurement is not necessarily the method by which its values are determined

Grammar of NAL-1

Table 3.2. The grammar rules of NAL-1.

⟨*sentence*⟩	::=	⟨*judgment*⟩ | ⟨*question*⟩
⟨*judgment*⟩	::=	⟨*statement*⟩ ⟨*truth-value*⟩
⟨*question*⟩	::=	⟨*statement*⟩ | ? ⟨*copula*⟩ ⟨*term*⟩ | ⟨*term*⟩ ⟨*copula*⟩ ?
⟨*statement*⟩	::=	⟨*term*⟩ ⟨*copula*⟩ ⟨*term*⟩
⟨*copula*⟩	::=	$\rightarrow$
⟨*term*⟩	::=	⟨*word*⟩

The representation language of NAL, *Narsese*, is extended incrementally by layer

Inference in NAL-1

- Actual experience is a stream of sentence
- Truth-value and meaning are explained by the **experience-grounded semantics**, and are subjective, uncertain, and change over time
- Inference rules are justified by the semantics, and are both syntactic and semantic
- Truth-value functions are designed by extending the Boolean operators from {0, 1} to [0, 1]

Local rules in NAL-1

Local rules do not produce new statements

- **Revision**: when pooling distinct evidence, the *amounts of evidence* are summations, *frequency* is a weighted average, and *confidence* increases
- **Choice**: to pick a "better" answer for a question
  - by *confidence* for an evaluative question ("*yes/no*")
  - by *expectation* for a selective question ("*what/which*")

Expectation of future frequency: $e = c\,(f - 0.5) + 0.5$

Forward rules in NAL-1

Table 4.4. The basic syllogistic rules.

$J_2 \backslash J_1$	$M \rightarrow P \langle f_1, c_1 \rangle$	$P \rightarrow M \langle f_1, c_1 \rangle$
$S \rightarrow M \langle f_2, c_2 \rangle$	$S \rightarrow P \langle F_{\text{ded}} \rangle$ $P \rightarrow S \langle F'_{\text{exe}} \rangle$	$S \rightarrow P \langle F_{\text{abd}} \rangle$ $P \rightarrow S \langle F'_{\text{abd}} \rangle$
$M \rightarrow S \langle f_2, c_2 \rangle$	$S \rightarrow P \langle F_{\text{ind}} \rangle$ $P \rightarrow S \langle F'_{\text{ind}} \rangle$	$S \rightarrow P \langle F_{\text{exe}} \rangle$ $P \rightarrow S \langle F'_{\text{ded}} \rangle$

Truth-value functions for syllogism

Table 4.7. The truth-value functions of the syllogistic rules.

Function name	Boolean version	Truth-value version
F_{ded}	$f = and(f_1, f_2)$	$f = f_1 \times f_2$
(deduction)	$c = and(f_1, c_1, f_2, c_2)$	$c = f_1 \times c_1 \times f_2 \times c_2$
F_{abd}	$w^+ = and(f_1, c_1, f_2, c_2)$	$f = f_2$
(abduction)	$w = and(f_1, c_1, c_2)$	$c = \frac{f_1 \times c_1 \times c_2}{f_1 \times c_1 \times c_2 + k}$
F_{ind}	$w^+ = and(f_2, c_2, f_1, c_1)$	$f = f_1$
(induction)	$w = and(f_2, c_1, c_2)$	$c = \frac{f_2 \times c_1 \times c_2}{f_2 \times c_1 \times c_2 + k}$
F_{exe}	$w^+ = and(f_1, c_1, f_2, c_2)$	$f = 1$
(exemplification)	$w = and(f_1, c_1, f_2, c_2)$	$c = \frac{f_1 \times c_1 \times f_2 \times c_2}{f_1 \times c_1 \times f_2 \times c_2 + k}$

Backward rules in NAL-1

Table 4.8. The backward syllogistic rules of NAL-1.

$J \backslash Q$	$M \rightarrow P$	$P \rightarrow M$
$S \rightarrow M\ \langle f, c \rangle$	$S \rightarrow P$ $P \rightarrow S$	$S \rightarrow P$ $P \rightarrow S$
$M \rightarrow S\ \langle f, c \rangle$	$S \rightarrow P$ $P \rightarrow S$	$S \rightarrow P$ $P \rightarrow S$

Related traditional problems

- Hume's problem of induction
- Hempel's confirmation paradox
- Wason's selection task
- Paradoxes of material implication
- The limitation of term logic
- Defeasible reasoning
- Conjunction Fallacy
- Randomness and fuzziness

Suggested Readings

- Robin Smith, Aristotle's logic
- George Englebretsen, Syllogistic: Old Wine in New Bottles
- John Vickers, The Problem of Induction
- Igor Douven, Peirce on Abduction
- Pei Wang, Non-Axiomatic Logic, Chapter 4
- Pei Wang, Rigid Flexibility, Chapter 3, 9