Notes for Meeting 26
Metareasoning and Metacognition

Thinking about Thinking

Cognitive systems have the ability to think about the environment in which they operate.

- However, in some sense, they are part of that environment, which makes their own thinking a candidate topic for cognition.

Researchers often refer to humans’ ability to think about their own thinking as metacognition.

- This has been considered a legitimate topic of scientific study within psychology for decades.

- Metacognition has received less attention in AI, but there has been some work in the area.

The abstact, self-referential nature of metacognition qualifies it as a distinctively human ability.

Examples of Metacognition

Metacognition in humans arises in many different contexts, including:

- Reasoning about one’s own memory processes
  - When to take notes, how much to study for an exam
- Thinking about one’s ability to make decisions
  - How long it takes to buy soap or a car, explaining reasons
- Reasoning about one’s own problem-solving abilities
  - How good you are at solving puzzles, explaining solutions
- Thinking about one’s learning abilities and strategies
  - Which ways you learn most effectively, when use which technique

Each of these capabilities require cognitive structures that encode metacontent over which to operate.

Paradigms for Metacognition

Cox (2005) notes that four distinct AI communities have addressed the topic of metacognition:

- Procedural approaches and societies of mind (e.g., Minsky)

- Declarative approaches and formal logic (e.g., McCarthy)

- Knowledge-based and expert systems (e.g., Davis and Clancy)

- Model-based and case-based reasoning (e.g., Leake and Cox)

These paradigms make quite different assumptions about representations and processes that support metacognition.

Cox also claims that some approaches labeled as metacognition do not really satisfy his definition of the concept.

Representations for Metacognition

Systems that engage in metacognition must store information about their own processes; this can take the form of:

- Episodic traces of particular events
  - When one was (not) able to retrieve something from memory
  - How one went about making a specific decision / solving a problem
  - Situations in which a learning methods worked well or poorly
- Concepts that describe generalizations about such situations
- Skills or strategies that apply in such generalized situations

A cognitive system can encode such information using traditional list structures, but the content is more abstract.

These structures support the raw material that metacognition needs in order to operate.

Processes for Metacognition

Given content that describes traces of cognitive activities and general structures that match again them, a metacognitive system can:

- Match its general structures against the concrete traces

- Select among the structures that match these traces

- Apply the matched structures to alter its knowledge or behavior

The basic machinery of metacognition does not require anything beyond regular cognition.

The difference lies entirely in the nature of experience (traces of cognition) and the general structures used.

This suggests there should be no need for any levels of thinking above basic metacognition.

The MetaAQUA System

One example of a metacognitive system is MetaAQUA (Cox & Ram, 19950, which operates on top of AQUA, a story-understanding system.

The system extends the notion of explanation patterns to support meta-explanation patterns about reasoning and learning.

MetaAQUA uses metacognition to drive learning about the task of story understanding by:

1. Determining the cause of a reasoning failure through a form of introspective blame assignment;

2. Deciding what content to acquire by formulating explicit learning goals that drive this process; and

3. Selecting and ordering known learning methods in order to pursue its learning goals.

The system reasons about both errors related to memory retrieval and errors related to inference.

Computational Costs of Metacognition

The additional powers of metacognition come with costs, since it can require substantial processing.

Some systems incorporate metacognitive processing to decide which cognitive actions to take:

- Genesereth & Ginsberg’s (1985) MRS system reasons about which inferences to make.

- Minton’s (1988) Prodigy uses control rules to decide which planning operate to select, reject, or prefer.

Such proactive approaches may be costly enough that they offset the benefits of metacognition.

Thus, a more cautious approach resorts to metacognition only when the agent encounters a problem, as in Cox and Ram’s MetaAQUA.

Assignments for Meeting 27
Integrated and Unified Approaches to Intelligence

Read the articles:

* Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. Cognitive Systems Research, 10, 141-160. [required]

* Lewis, R. (1999). Cognitive theory, Soar. [optional]

* Anderson, J. R. et al. About ACT-R. [optional]