# Learning Background Knowledge through Instruction

Shiwali Mohan

shiwali@umich.edu

The Soar Group at University of Michigan

# Outline

- Motivation
- Properties of Instruction
- Semantic Memory and Storage
- Learning from Instruction in Infinite Mario
- Results
- Nuggets and Coal

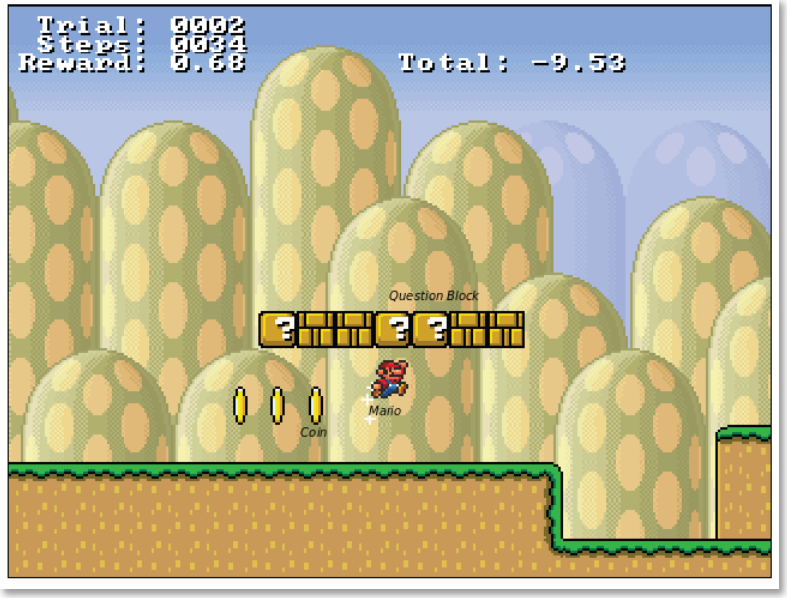
# Motivation

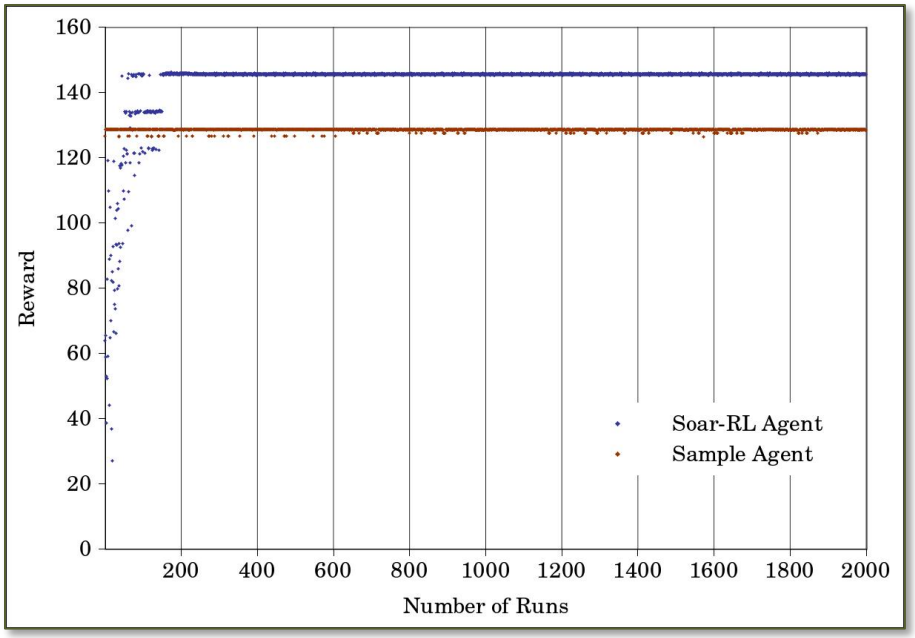
- What is the source of procedural knowledge?
  - Innate
  - Experience with the world
    - Learning to ride a bicycle

  - Specific Instructions
    - Cooking

# Motivation

- Previously,
  - Agent that learn behaviors from experience
    - State-action value function
- Further questions
  - Which behavior is associated with which object?
    - tackle-monster is to be applied when an object ‘monster’ is nearby
      - Problem of matching the FLO to corresponding object
  - Which parts of the input are important ?
    - What features should be learned over?
    - *Speed, type, relative distance* etc…
      - Learning the correct structure of FLO
- Till now knowledge is hand-coded
  - As rules in the Soar Implementation



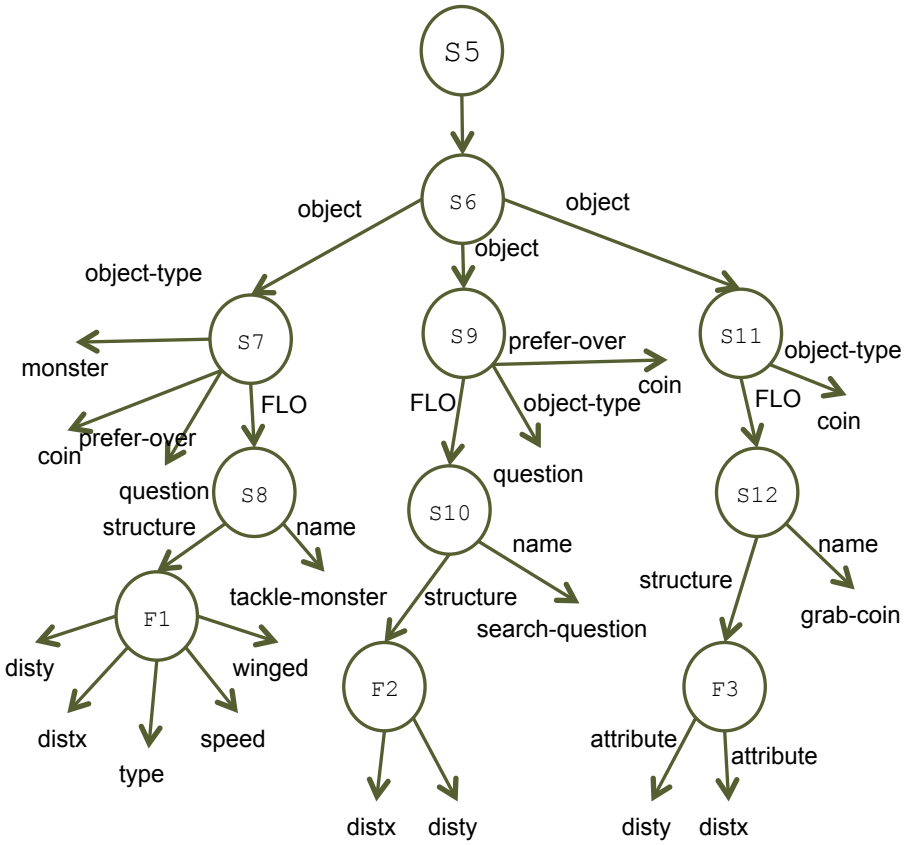
# Properties of Instruction

- Situation specific
  - instructions  are provided for a specific task in a particular situation
- Incremental
  - knowledge is elicited incrementally
  - it directly addresses points where agent's knowledge is lacking
- Knowledge-level interaction
  - objects, features and actions
  - not to lower-level symbols.
- Agent initiated communication

# Semantic Memory

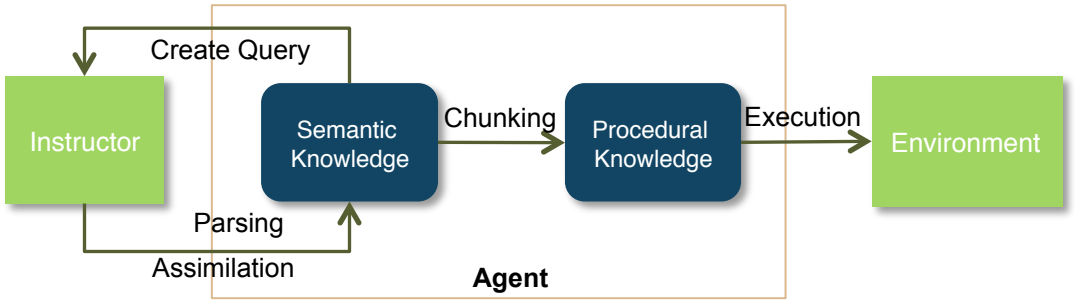
- Semantic Memory
  - Declarative memory
    - General, context-free knowledge of the world
  - Organized information
    - Salient properties of an object
      - What it looks like?
      - How it moves?
      - What are its components
  - Can be acquired in different ways
    - Can be built up from experiences
    - Through explicit instruction

- Proposed Object-Oriented Representation in Semantic Memory
  - Structures associating FLOs with objects
  - Learn preferences between objects
  - Structures associating features to FLOs



# Two step learning

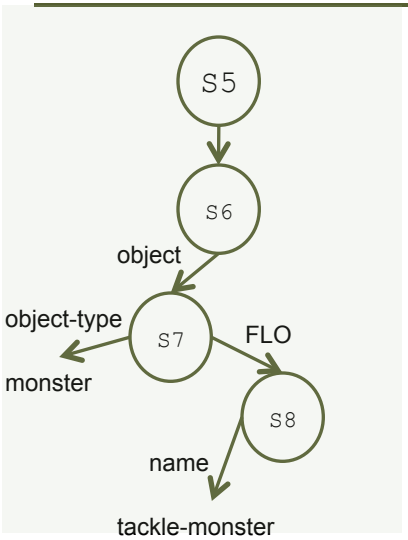
- Semantic Learning
  - Detect incomplete structures in semantic memory
  - Convert internal representations in to a query
  - On receiving a reply, parse the answer
  - Assimilate
  - Incremental
- Procedural Learning
  - On forming new semantic structures, use generic ways to convert them into procedural knowledge
  - Learn general rules from specific examples

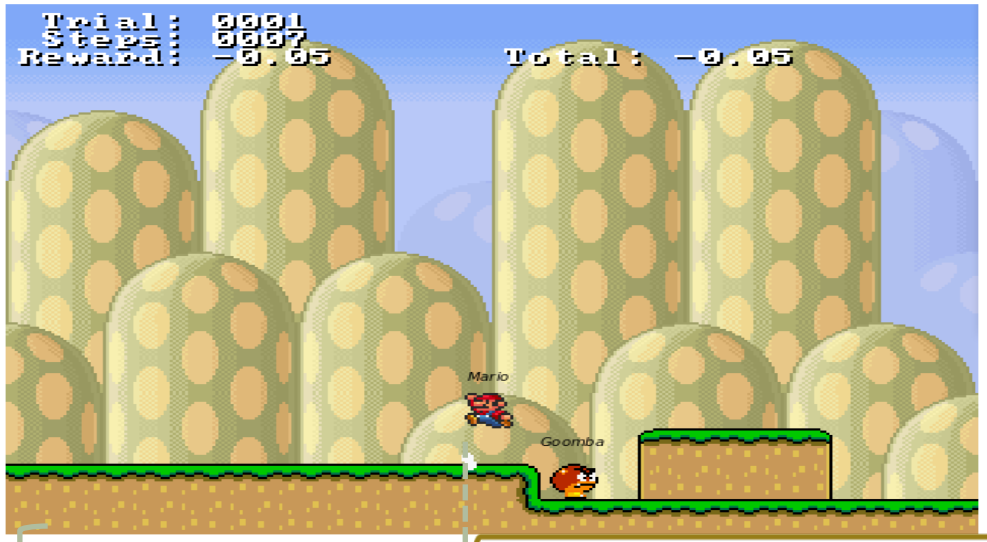


# Learning from Instruction in Infinite Mario

- Semantic learning using Soar-SMem
  - No natural language support
  - Agent communicates using its internal representations
- Procedural Learning
  - Chunking – explanation based learning
    - Learn new rules at a state by solving a problem in the substate
- Reduced learning problem – example
  - Match the FLO with correct object
  - Assume the agent knows symbolic preferences for FLOs
  - Assume the agent posses procedural knowledge to apply FLO
  - Once the correct FLO is selected, the agent knows which features are important
    - for tackle-monster, features – speed, type, winged, distx, distxy

# Progression




*If object of type monster isreachable, propose tackle-monster*

*If object m1 of type monster isreachable, propose tackle-monster*

Goal: apply correct FLO

state: mario-soar
action: attend-object(m1)

m1 isreachable
yes

action: apply(FLO(m1))

Goal: retrieve correct structure from smem

state: attend-object(m1)
action: query-smem(m1)

failure

action: ask-instructor(m1)

action: query-smem(m1)

success
Retrieved: FLO(m1)

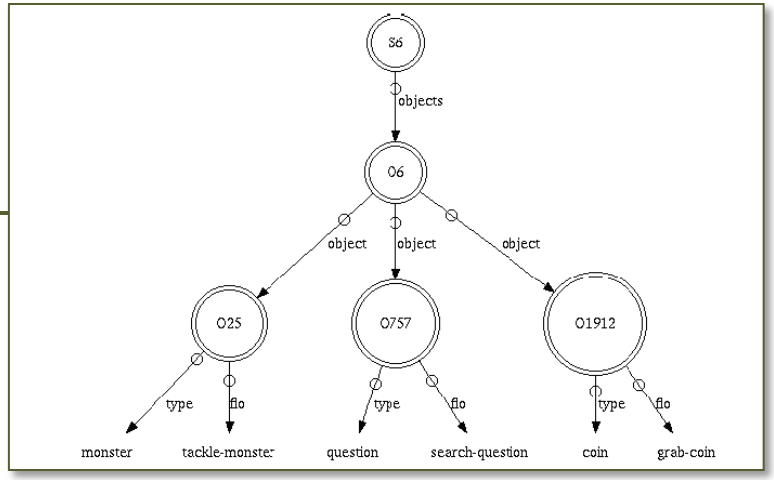
Goal: learn correct FLO from Instructor

state: ask-instructor(m1)
action: put-string(m1)

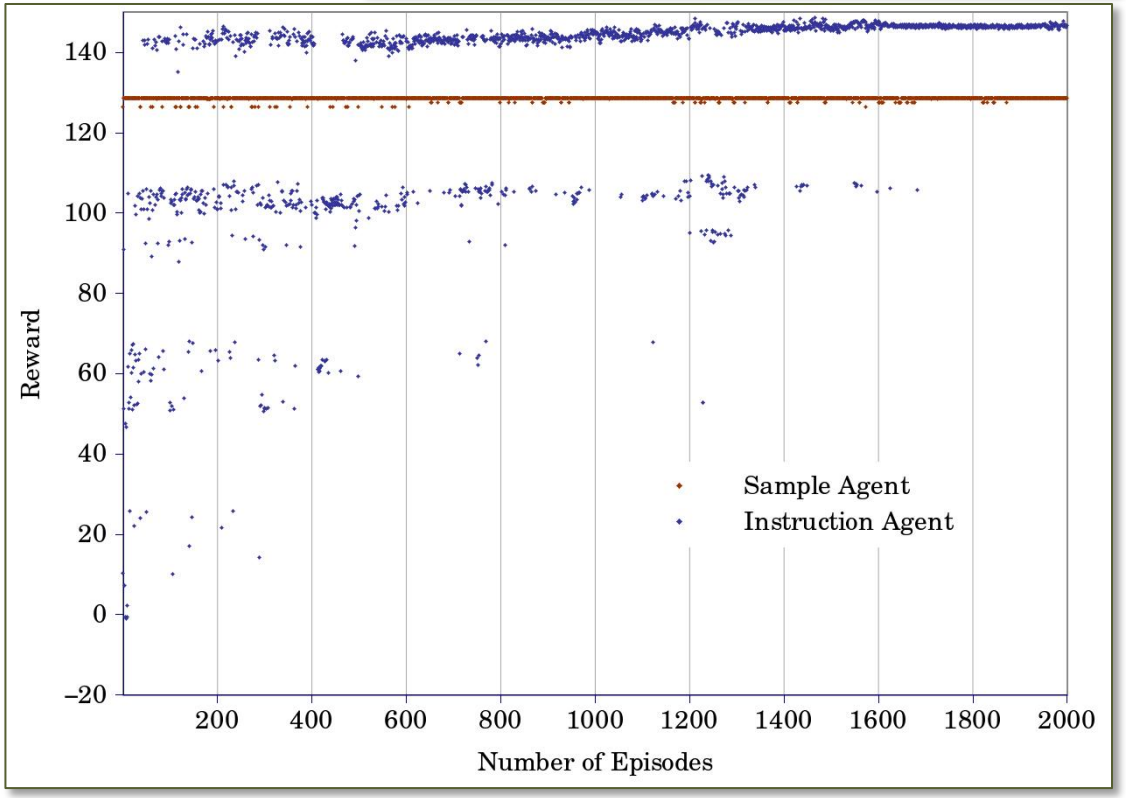
reply from instructor:
FLO(m1)

action: store-flo(m1)

# Results




- Learning algorithm - SARSA
- Learning rate – 0.3
- Discount rate – 0.9
- Exploration policy – Epsilon-greedy
- Epsilon – 0.01
- Reduction-rate – 0.99
- Performance similar to Agent 2
- Converges to a policy in 1600 runs
- Average reward earned by converged policy (last 100 runs) =144.34



# Nuggets and Coal

- Proposed a structure for learning from instruction.
- Proposed object-oriented storage in semantic memory that may aid learning correct procedural structures
- Still in initial phases