Taxonomy of Learning and Performance Integration

John Laird, Shiwali Mohan, Bryan Stearns

Soar Workshop 2019

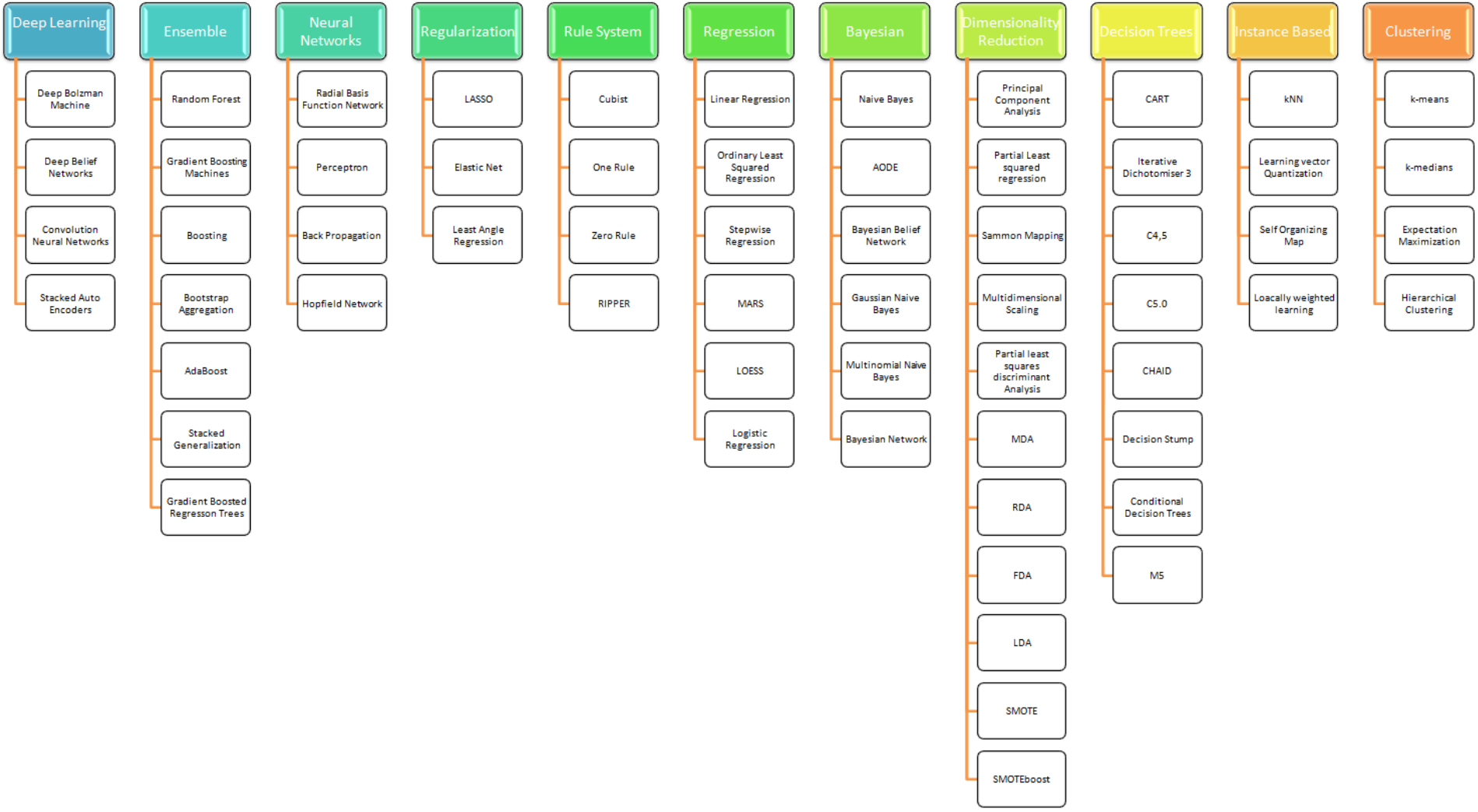
Funded by AFOSR

Possible Taxonomies of Learning

- **Information type** as input to learner
  - Continuous, discrete, symbolic, …
- **Feedback** available for learning
  - Unsupervised, supervised, semi-supervised, reinforcement
- **Algorithm and representation**
  - Regression, clustering, decision trees, version spaces, neural networks, K-nearest neighbors, Q learning, support vector machines, Bayesian networks, …
- **Type** of learned knowledge
  - Classification, decision making, planning, …

Categories of Machine Learning: Algorithm

Trymachinelearning.com



Alternative Taxonomy: How is Learning Integrated with Performance?

- Forthcoming CCC AI Roadmap will call for "Science of Integration in AI."

- Our proposal for human-level systems:
  - Level 1: Architectural Mechanisms
    - Automatically capture ongoing experience.
    - Innate, effortless, online, always active
    - Diverse learning mechanisms for diverse long-term memory structures
  - Level 2: Knowledge-based Strategies
    - Metareasoning that deliberately creates experiences for L1 mechanisms to learn.

- Project goal is better understand of integration of learning within autonomous systems.

Initial Research: Take a Step Back

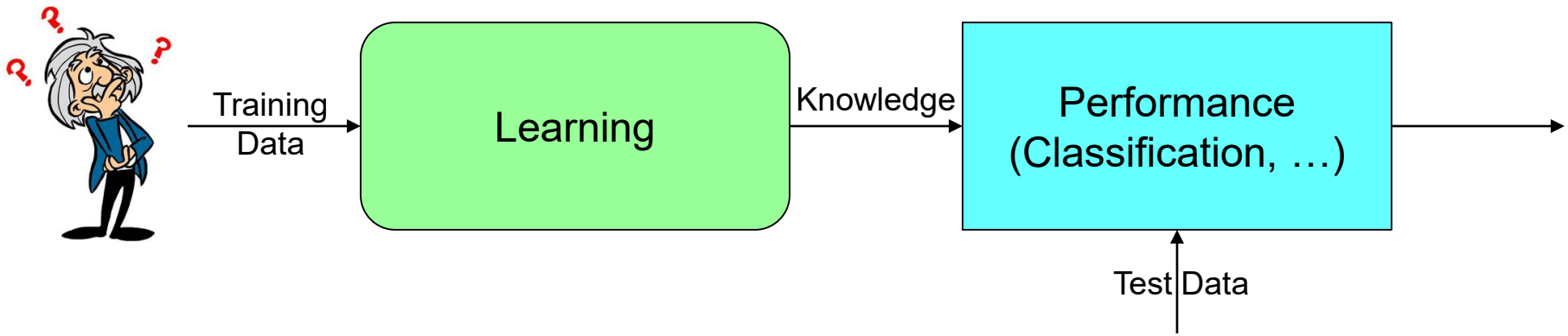
- Flesh out taxonomy of learning and performance integration.
  - Based on analysis of neuroscience, animal behavior, cognitive psychology, educational psychology, and AI/ML.
- Focused on AI/ML because of more variations.
  - Reviewed ~50 AI/ML learning systems and cognitive architectures.
  - Why is there so much variation in AI/ML?
- What are the critical dimensions of that taxonomy?

Integration Dimensions: *Data and Control*

- **Data Source:** Source of data used for learning.
  - External to agent / internal processing.
- **Learning Control:** Control of when learning occurs
  - External system / internal.
- **Experience Control:**
  - Internal (autonomous) / External (slave)
- **Internal Learning Control:**
  - Direct deliberate / automatic.
- **Internal Learning Control Goals:**
  - Innate / external direct / external indirect / internal.

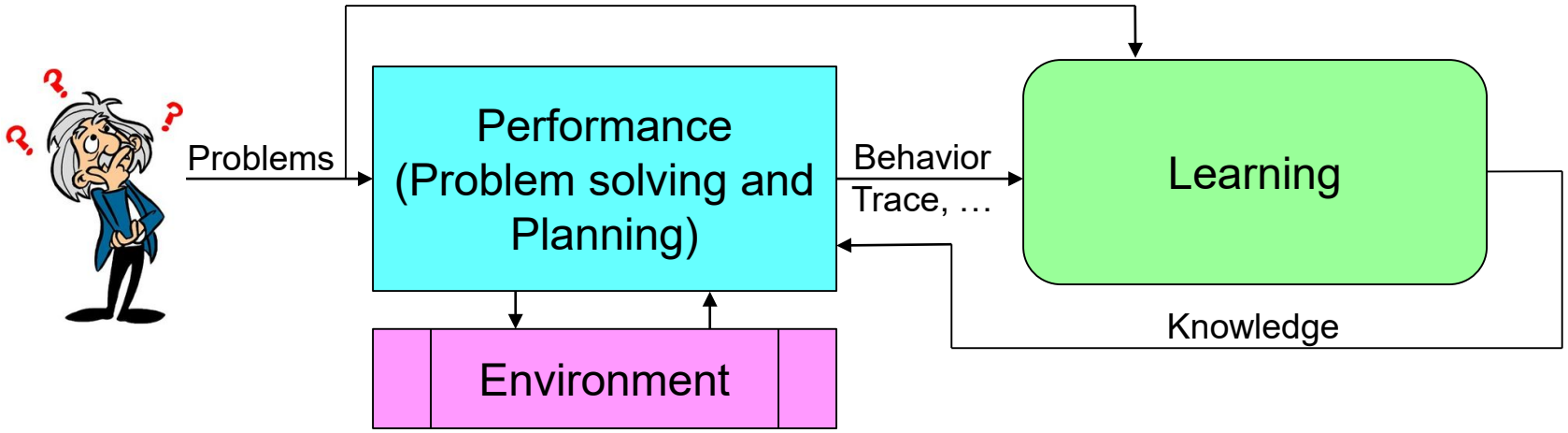
Data Source External: *Batch-*

- Learning and performance are separate.
- **External source of learning data: training data.**
- Majority of ML systems: classification, many learning by demonstration,…

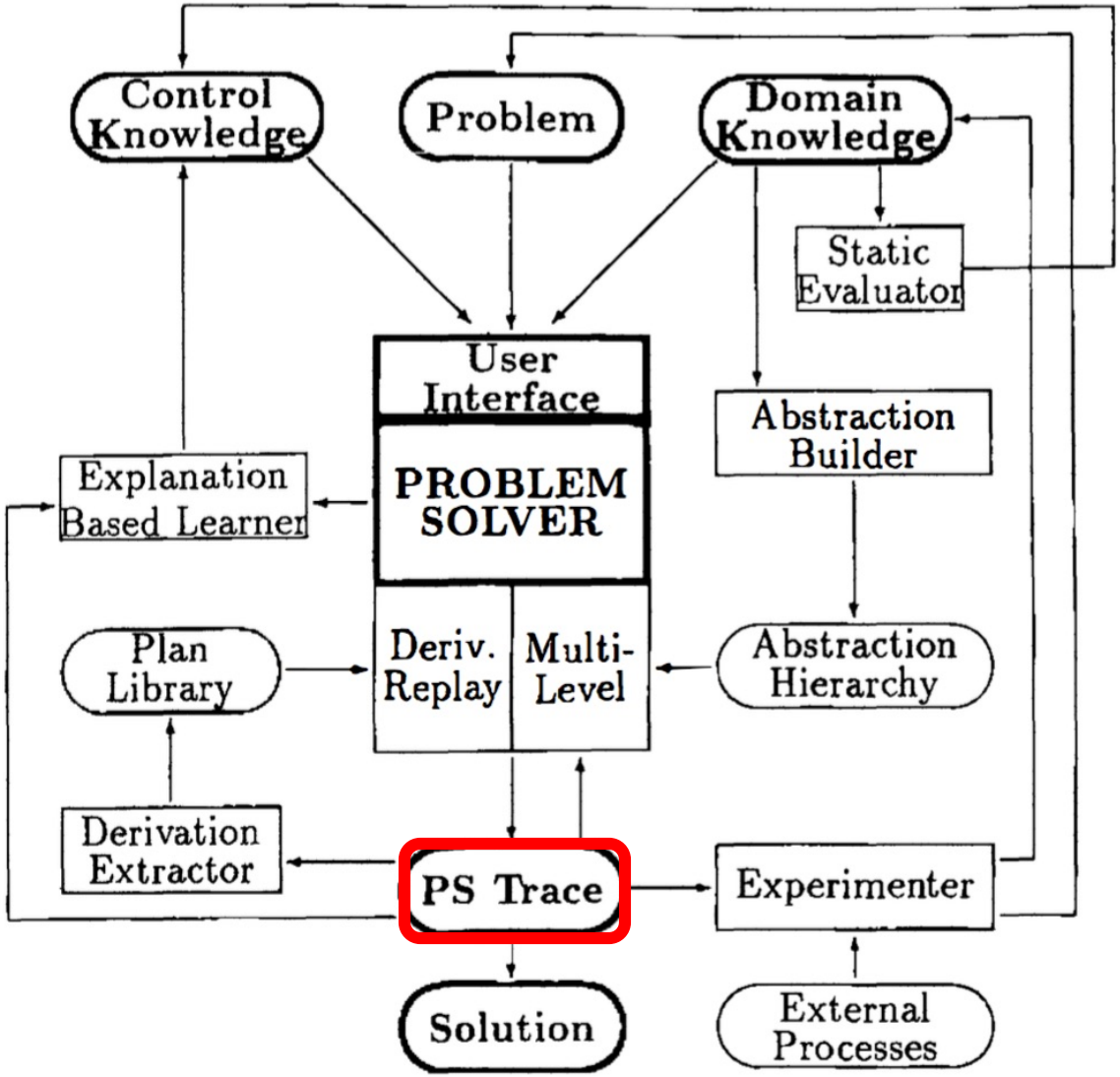


Offline Performance Learning: *Segregated-*

- Learning and performance still separate.
- **Internal source of learning data: behavior trace.**
- External control of learning experience.
- Examples: Prodigy, FORR, CBR, …

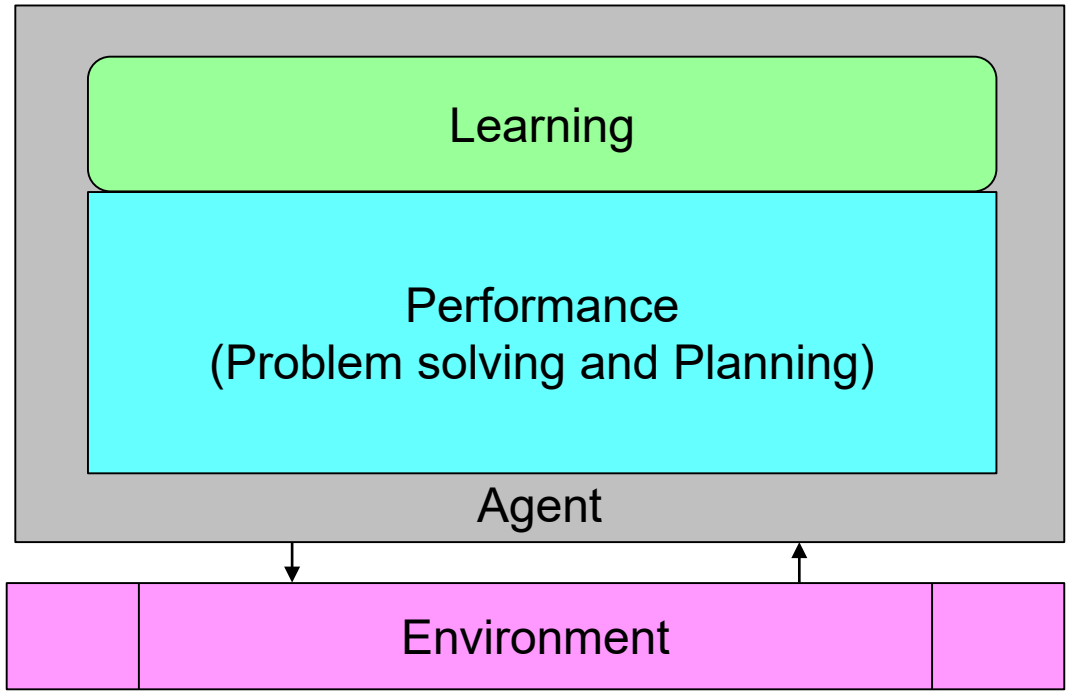


Prodigy


Control Knowledge
Problem
Domain Knowledge
Static Evaluator
User Interface
PROBLEM SOLVER
Explanation Based Learner
Abstraction Builder
Plan Library
Deriv. Replay
Multi-Level
Abstraction Hierarchy
Derivation Extractor
PS Trace
Experimenter
Solution
External Processes

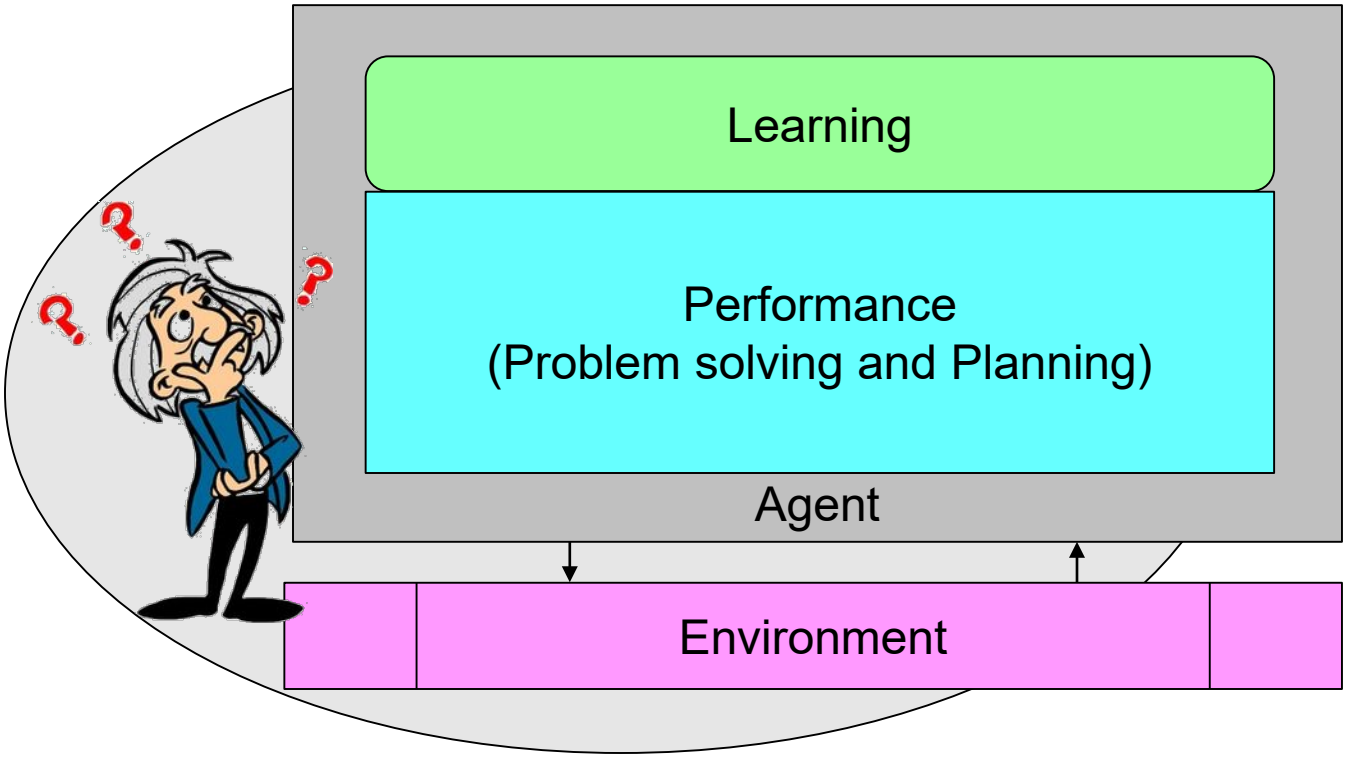
Online Learning over Performance:
Autonomous Learning Agents+

- Learning and performance integrated within an agent.
- **Internal source of learning data: agent's experience.**
- **Internal control of learning**
- Autonomous learning agents: Reinforcement learning (RL) agents, cognitive and robot architectures



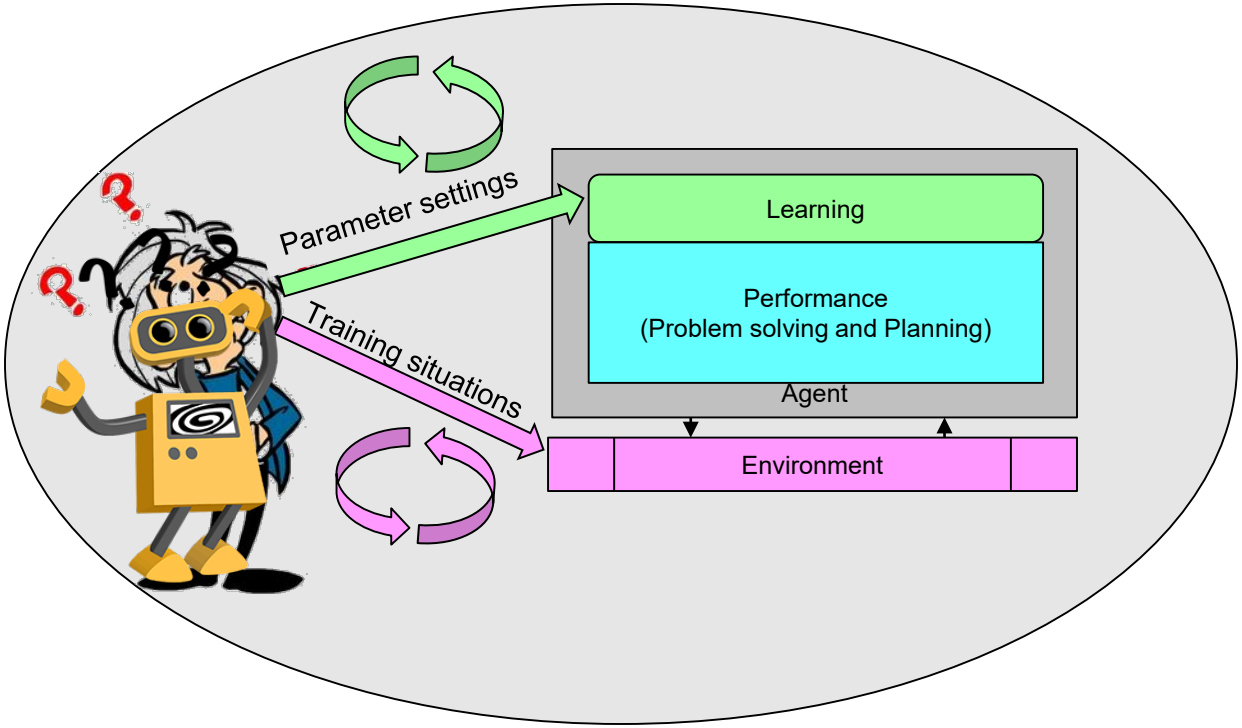
External Experience Control: *Slave-*

- Core agent is still an autonomous online learning system, but..

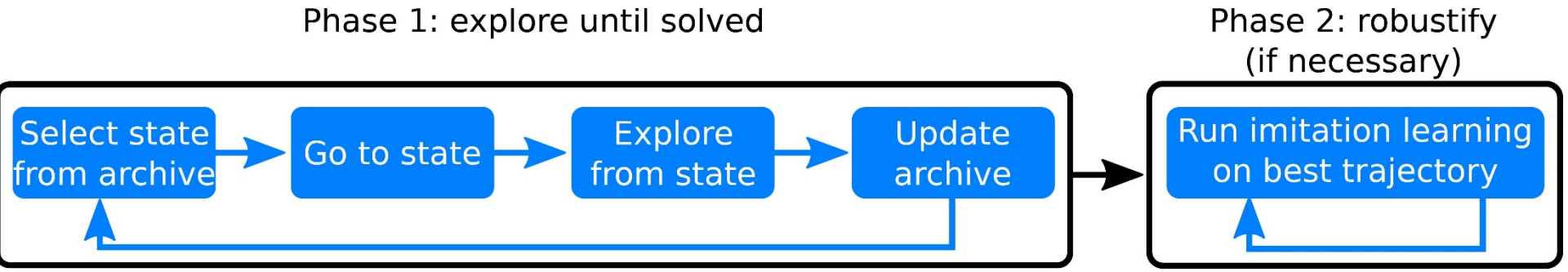


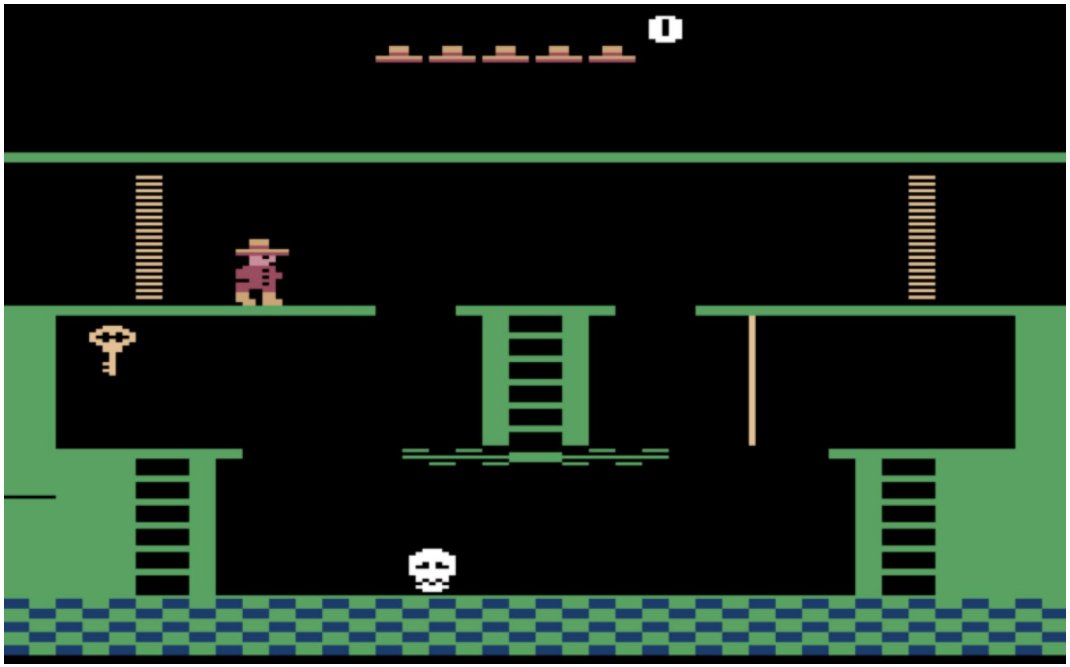
External Experience Control: *Slave-*

- Agent's experience is controlled external human (or program)
  - Curriculum training, parameter sweeps, reward functions, goals, …
- Many learning by demonstration and imitation
- Many RL systems: AlphaZero, AlphaStar, Tamer2, …
- Worth cataloging all of these manipulations…



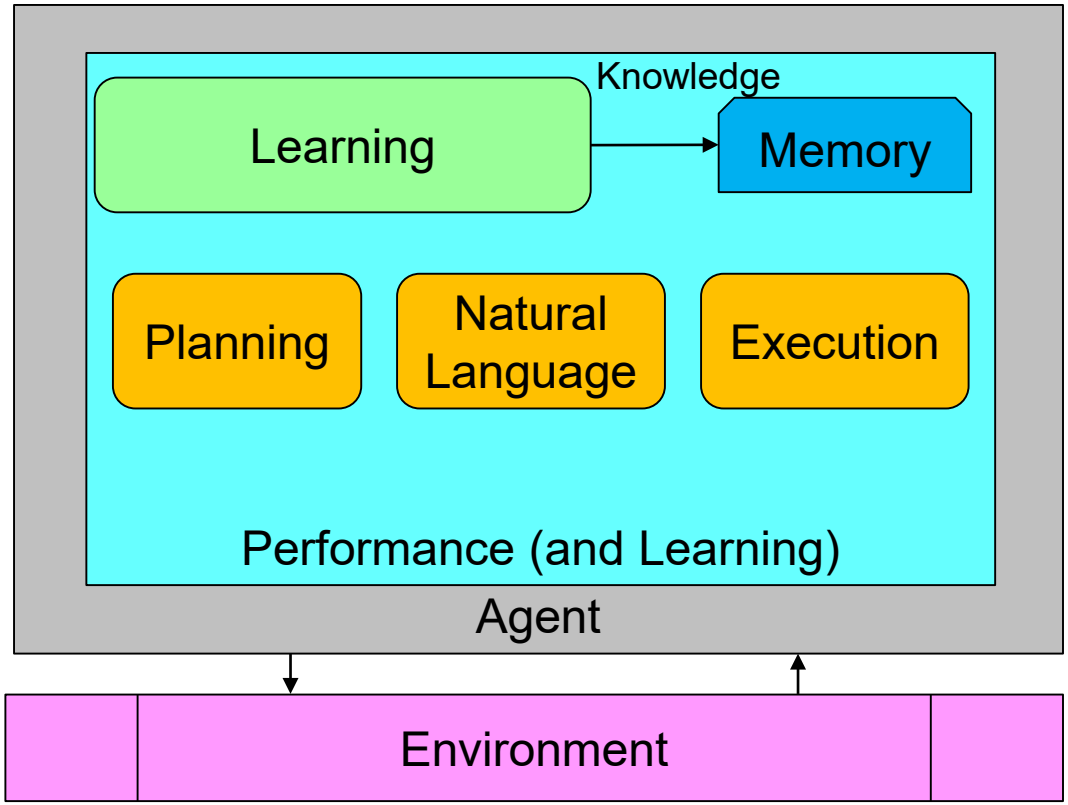
Go-Explore: Uber



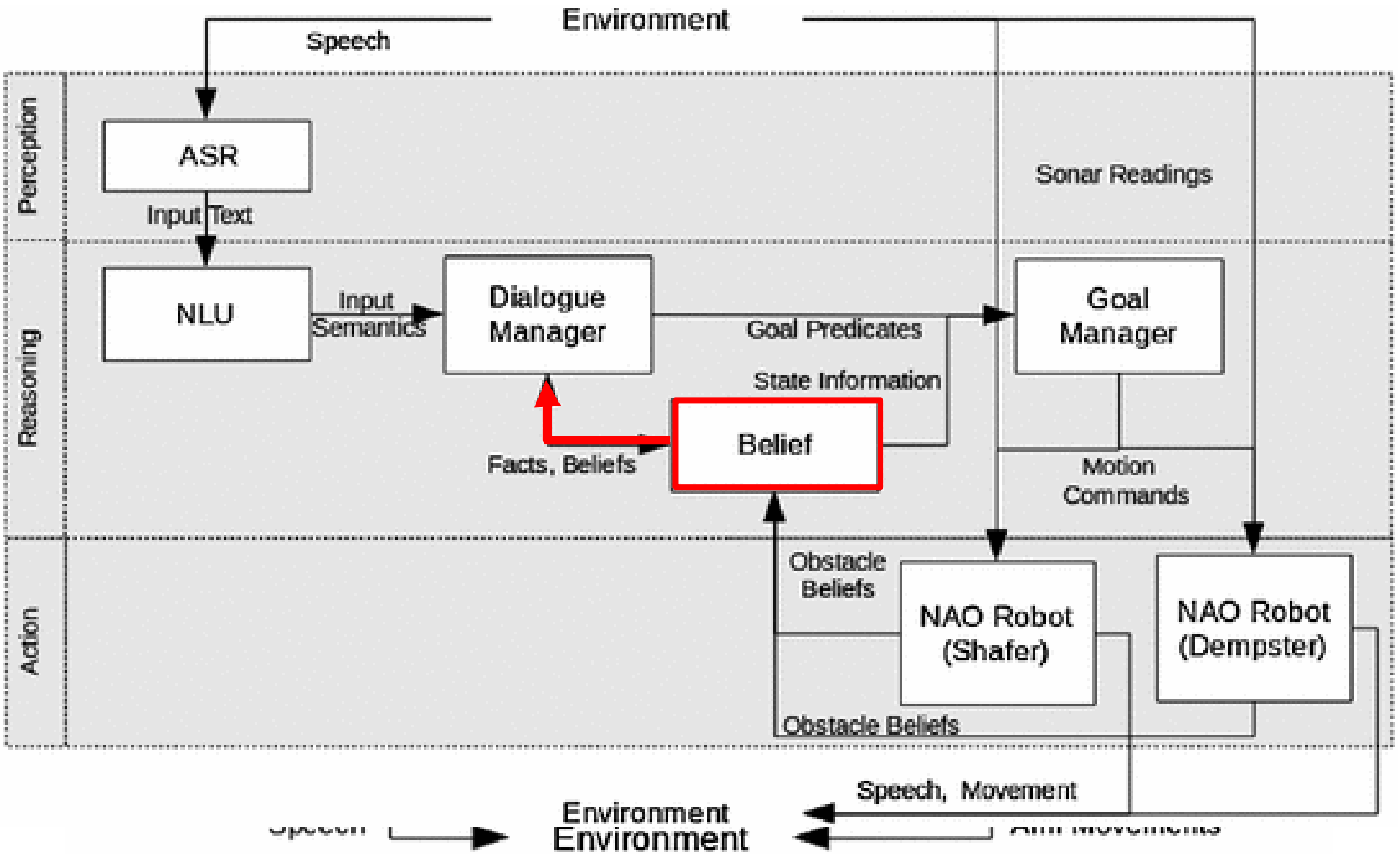


Autonomous Experience: *Deliberate Learner*

- Learning "task" modules deliberately store knowledge
- Cognitive frameworks: DIARC, CoBots, Blackboards, and learning robotic architectures, personal assistants (?)

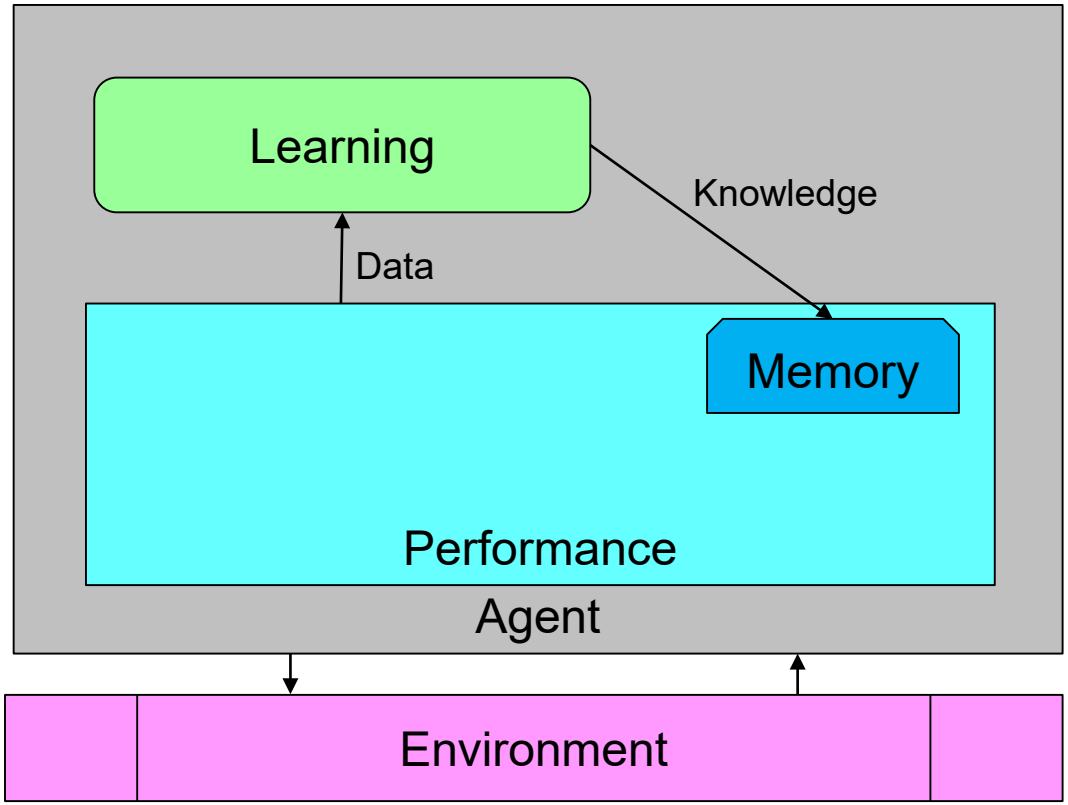


DIARC

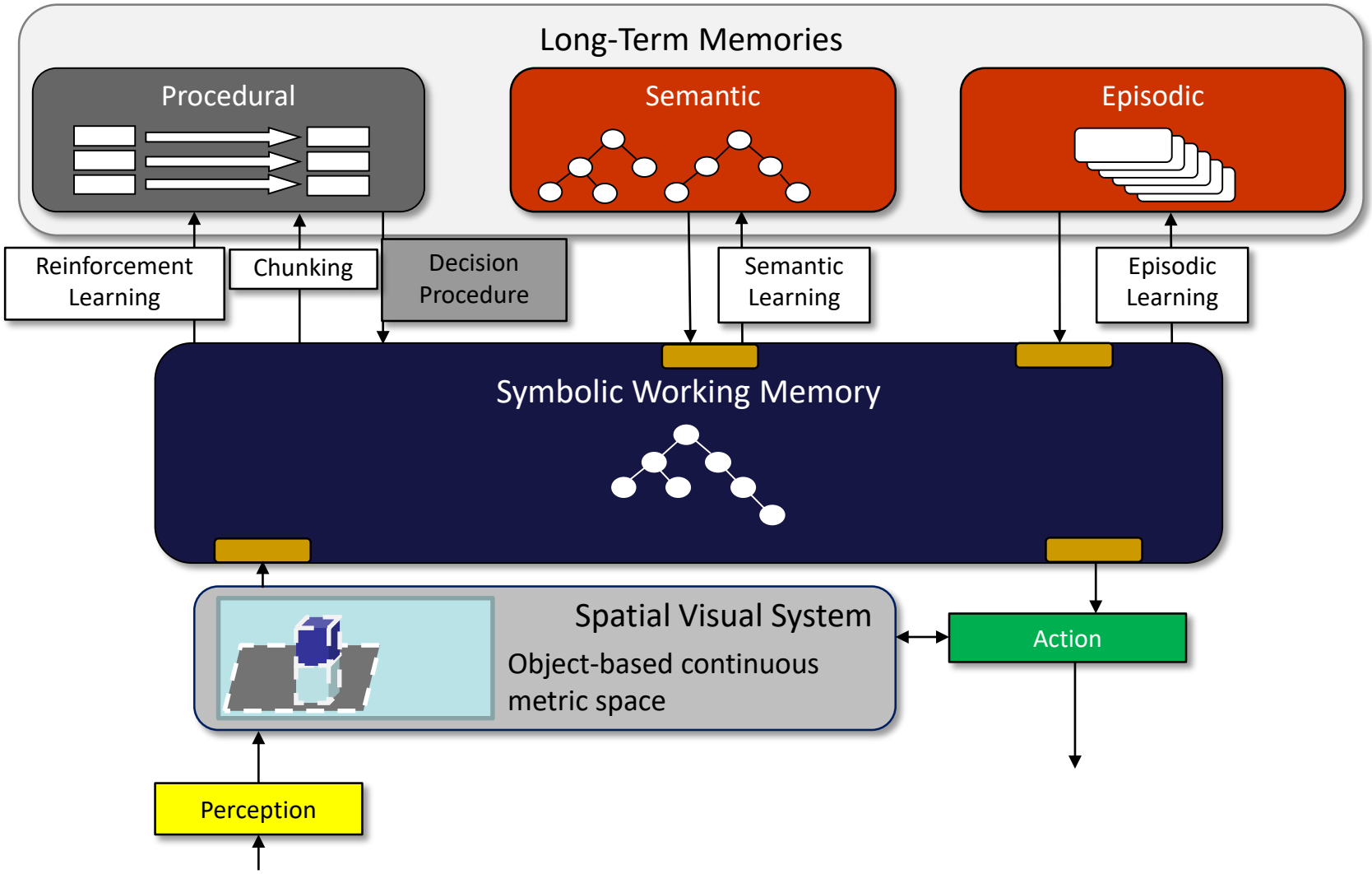

Environment
Speech
Perception
ASR
Input Text
Sonar Readings
Reasoning
NLU
Input Semantics
Dialogue Manager
Goal Predicates
Goal Manager
State Information
Facts, Beliefs
Belief
Motion Commands
Action
Obstacle Beliefs
NAO Robot (Shafer)
NAO Robot (Dempster)
Obstacle Beliefs
Speech, Movement
Environment
Environment

Autonomous Experience: *Automatic Learning*+

- Source data is agent's experiences
- Architectural learning mechanisms: L1
- Many cognitive architectures: Soar, ACT-R, ...
- Reinforcement learning (RL), SLAM, …

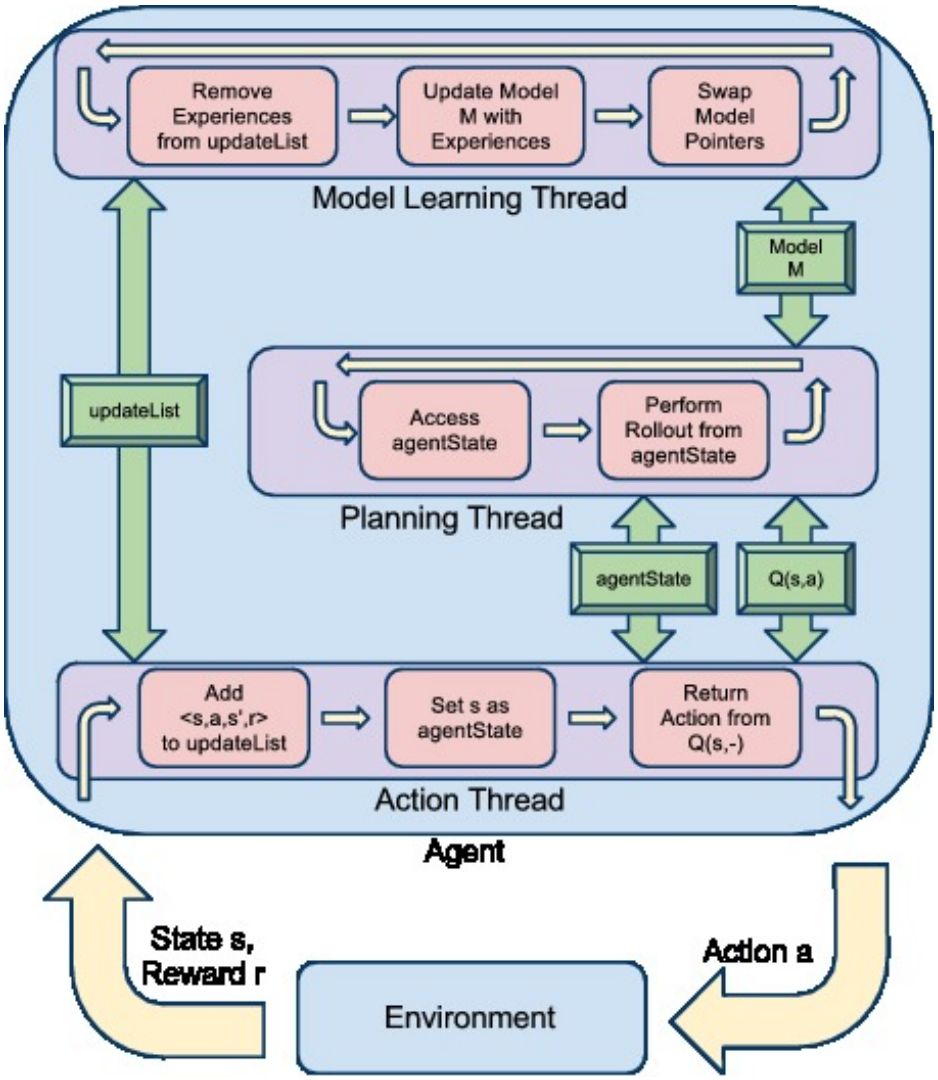


Soar 9 Structure



TEXPLORE Reinforcement Learning

Hester and Stone 2015



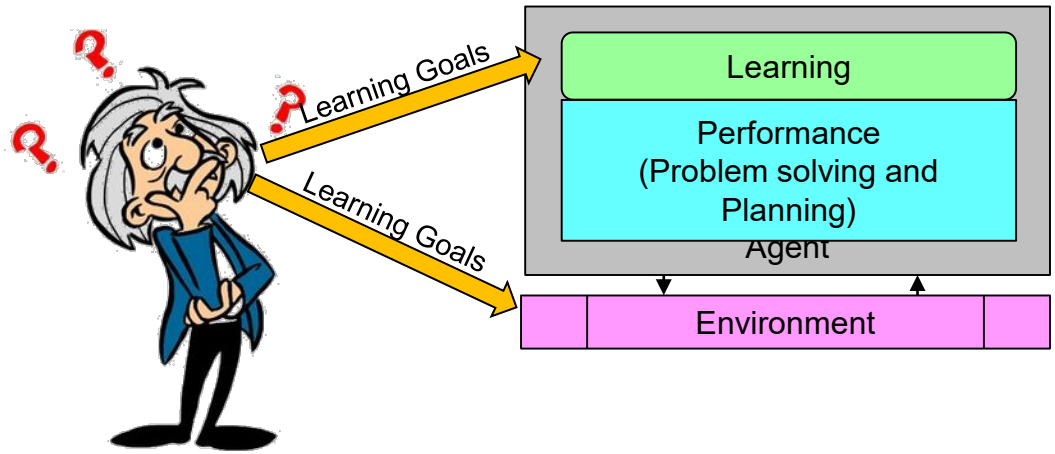
Final Dimension: Source of Goals for Learning

Decided what should the system learn?

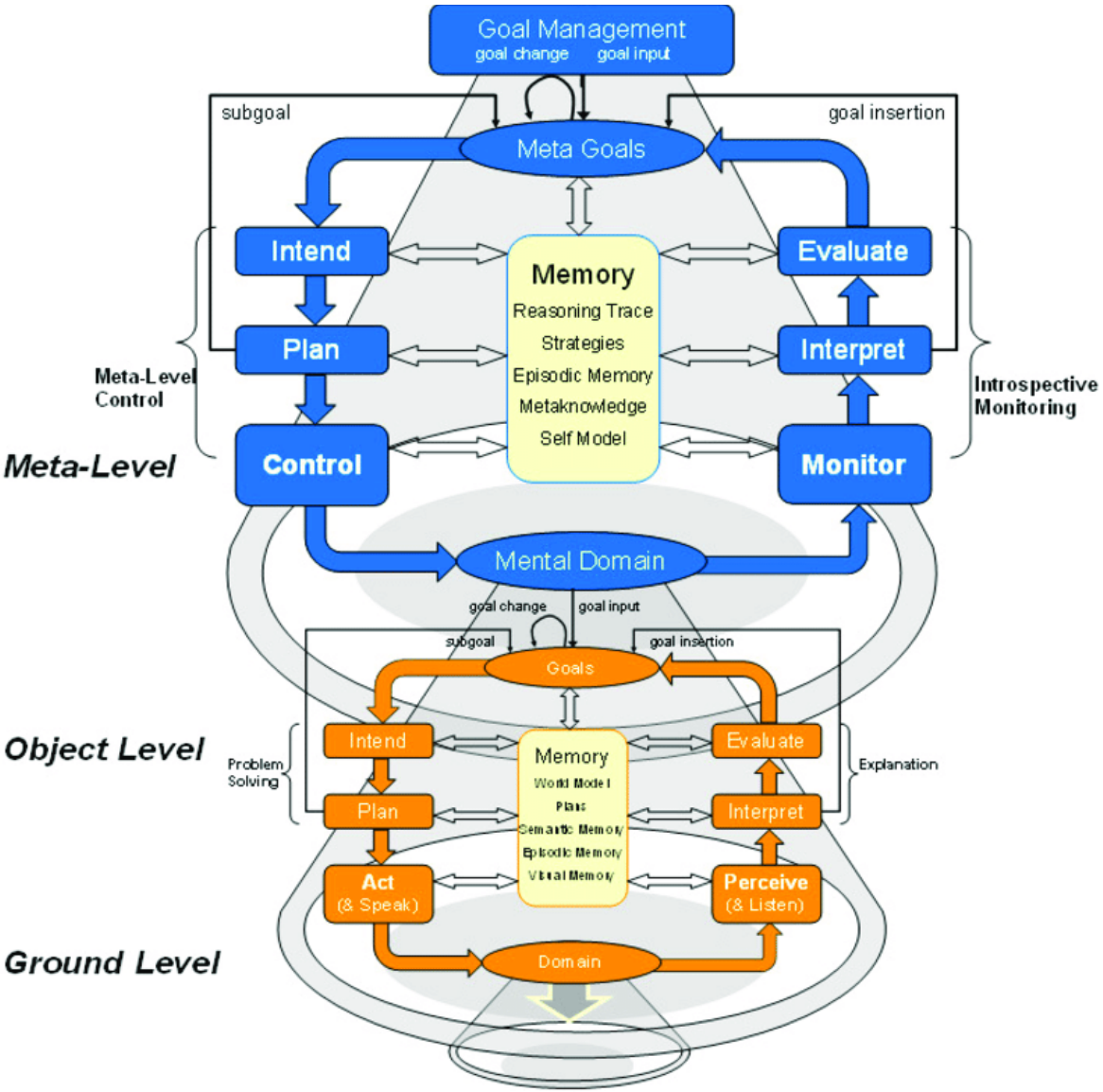
1. Innate: fixed set of what it will learn.
   - Batch, off-line learning, many on-line learning,...
2. External direct
   - Potentially *slave* systems. No known examples…
3. External indirect
   - Few systems can take a learning goal as input.
   - Some interactive task learning agents?
4. Internal meta-reasoning
   - Direct control of learning
   - Indirect control of learning through performance

Source of Learning Goals

- Direct^{-}: Inject into agent's learning system.
- Indirect^{+}: Communication through some interaction (language): Telling Siri to remember something…

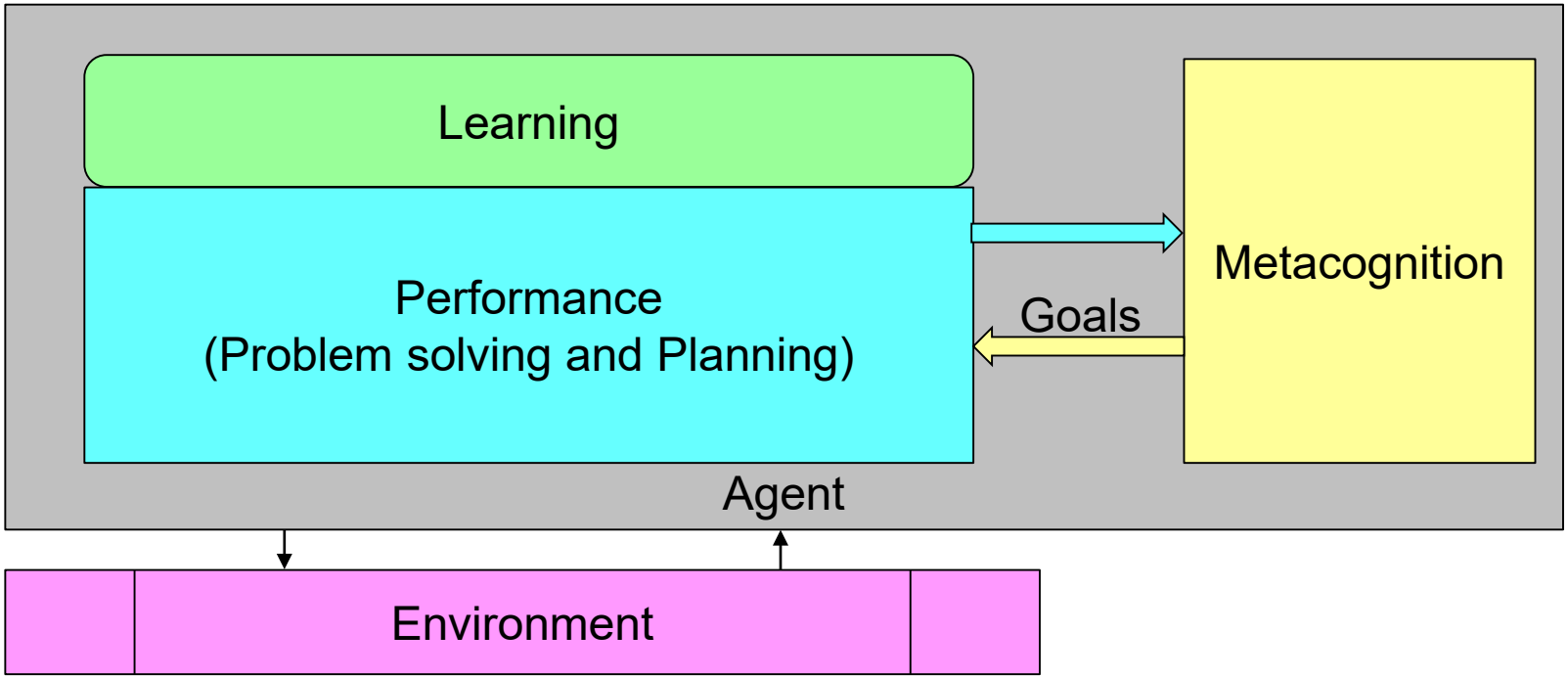


Internal Metareasoning Direct Control of Learning: MIDCA: Cox et al. (2016)



Internal Metareasoning: L2+

- Agent determines *performance* goals that indirectly influence learning.

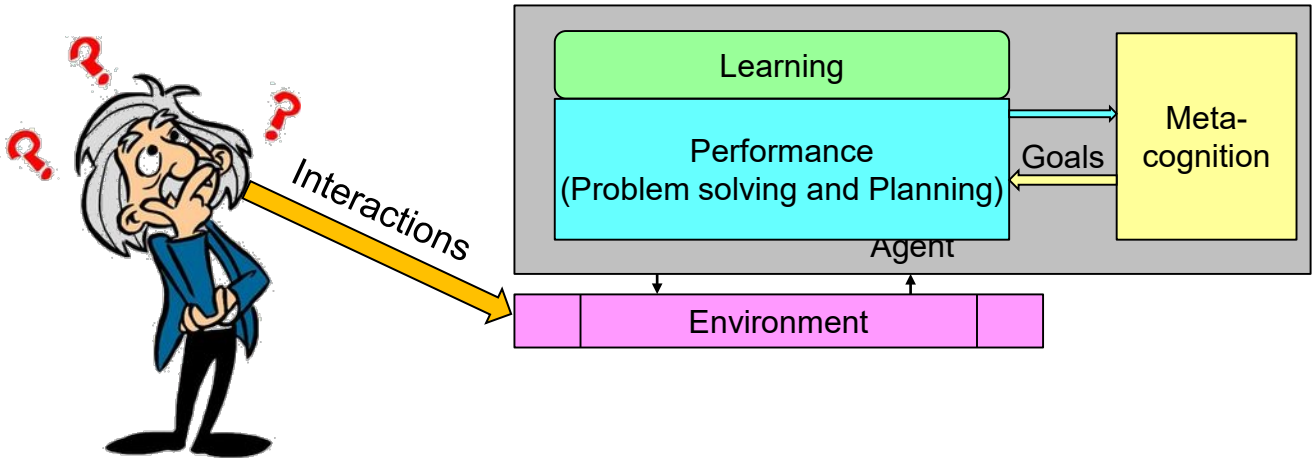


Potential L2 Strategies

- Repeat important experiences:
  - *practice, studying, training, …*
- Initiate novel experiences:
  - *explore, go to classes, lectures, …*
- Recall, replay, analyze prior experiences:
  - *self-explanation, retrospective analysis, …*
- Interact with other agents:
  - *learning by instruction, demonstration, …*
- Retrieve related knowledge:
  - *complex analogy*
- Imagine hypothetical situations:
  - *planning, mental preparation, and rehearsal*

Metareasoning and External Interaction (L2+)

- Interact with other agents:
  - *learning by instruction, demonstration, …*
  - "Maybe you should use flash cards?"



Level 1

Self-Explanation

Recognition

Discovery

Episodic Learning

Learning by Analogy

Category and Concept Learning

Learning by Instruction

Sequence Learning

Level 2

Learning by Demonstration

Rehearsal

Procedure Learning

Meta-Learning

Temporal-Difference Learning

Experimentation

Imitation Learning

Perceptual Learning

Practice & Rehearsal

Environmental and Agent Characteristics that may lead to L1 & L2 split

Properties that may engender:

- Continual, embodied autonomous existence.
- Computational limits on architectural learning mechanisms.

Capabilities that may enable:

- Metacognitive reasoning about how actions can enable learning.
  - Self-modeling of its own L1 learning capabilities.
  - Unclear what metacognitive capabilities are required for L2.
- Episodic memory that allows regularity detection and construction of self-model.
- Social interactions that share knowledge.

Expanded Levels

- L0: Evolution: creates L1 Mechanisms
- L1: Architectural Learning Mechanisms
- L2-: Innate Learning Strategies
  - Curiosity, imitation, play in young animals, …
- L2: Knowledge-based Learning Strategies
- L2+: Social Learning Strategies
  - Organized education, funded research, conferences!
  - Focus of much of educational psychology
  - Eliminates need for agent's own motivation and meta-knowledge
- L3: Modification of L1 Mechanisms
  - Rest, exercise, ingesting cognitive enhancing drugs (nootropics and stimulants)

Wild speculation: L2, L2+, L3 are unique to humans

Will future of ML be in Level 2 strategies?