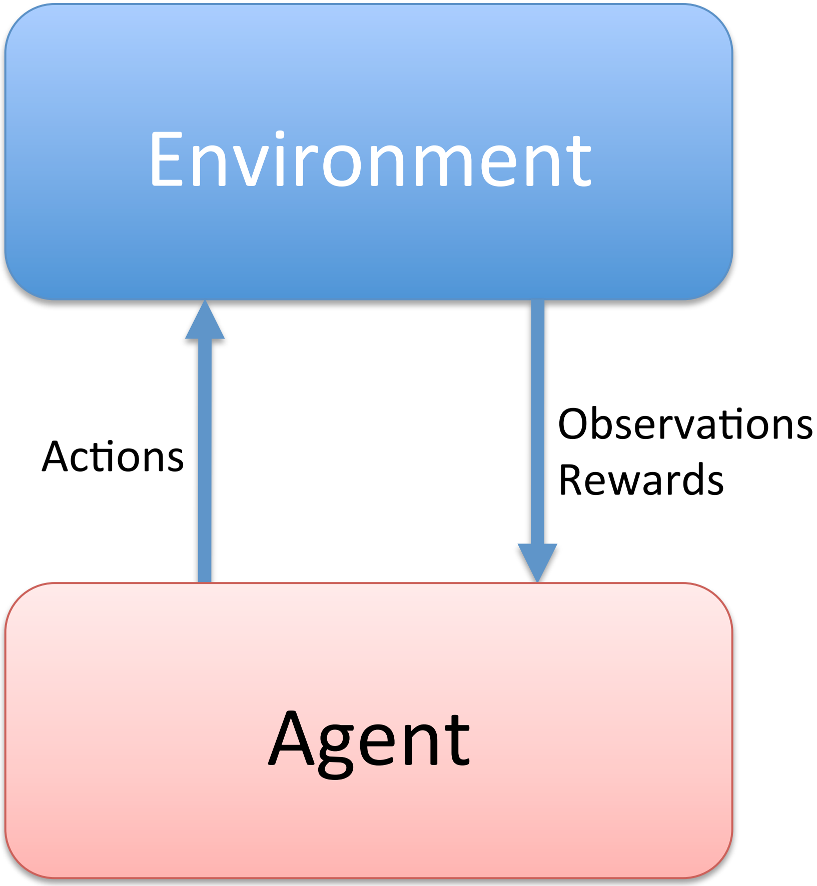
Soar-RL and Reinforcement Learning

Introducing talks by
Shiwali Mohan, Mitchell Keith Bloch
& Nick Gorski

Reinforcement Learning

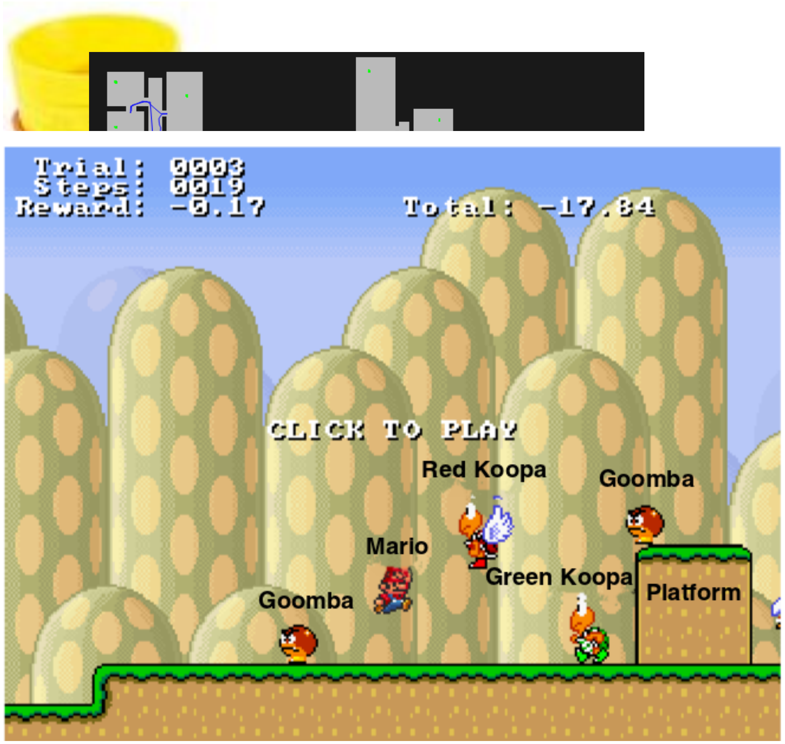


Environment

Actions

Observations
Rewards

Agent



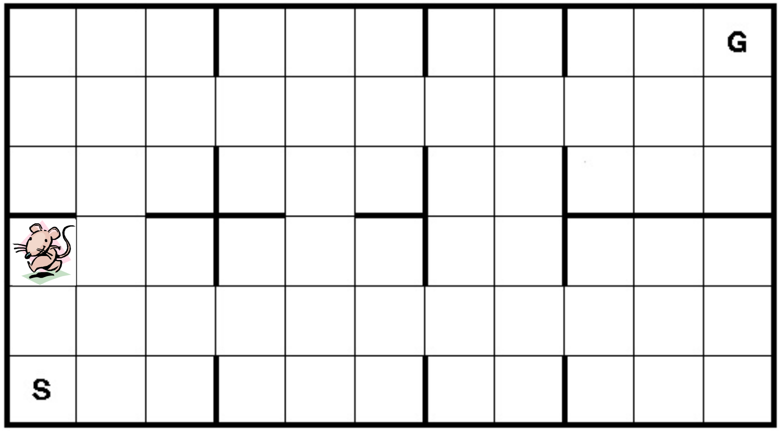
Value in Reinforcement Learning

- *Value*: future expected reward
- RL goal: maximize value
- RL agent: select actions with highest value

State and Observability

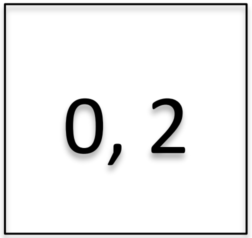
Agent observes a representation of world state

Can be Markovian or partial

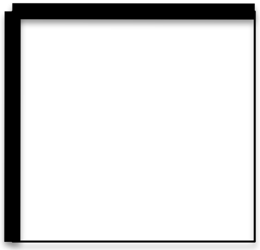


Agent doesn’t observe semantics of task

Must does learn the meanings of symbols and actions

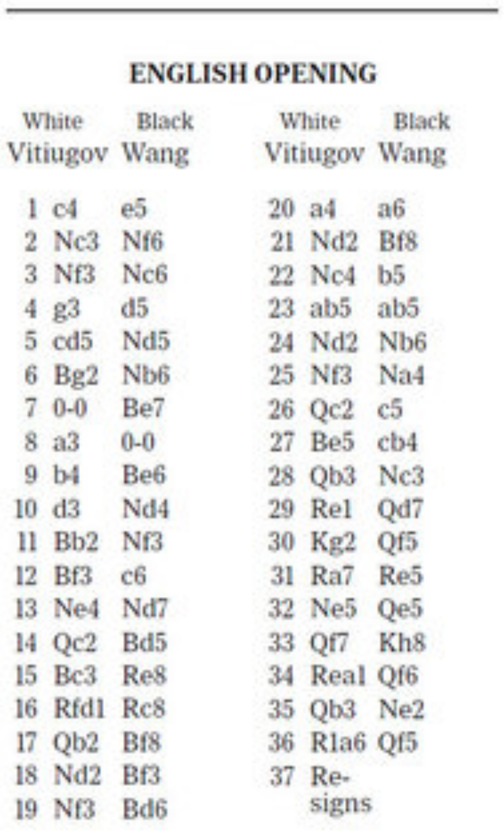


Markovian representation



Partial representation

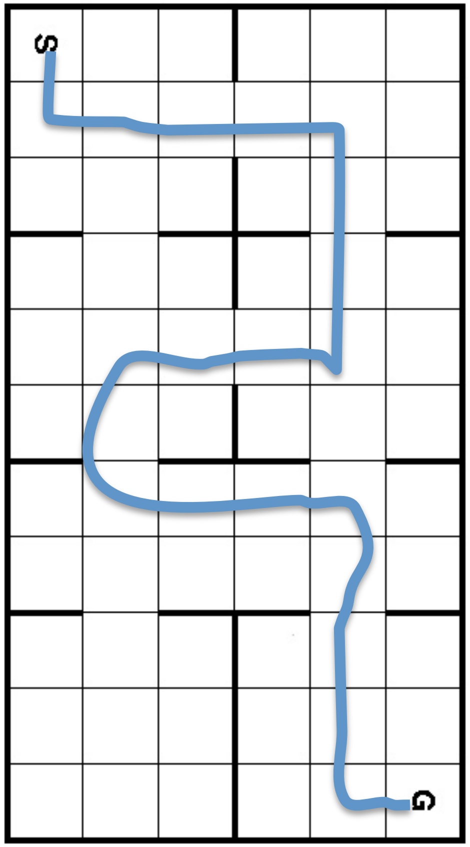
Why Reinforcement Learning Is Hard



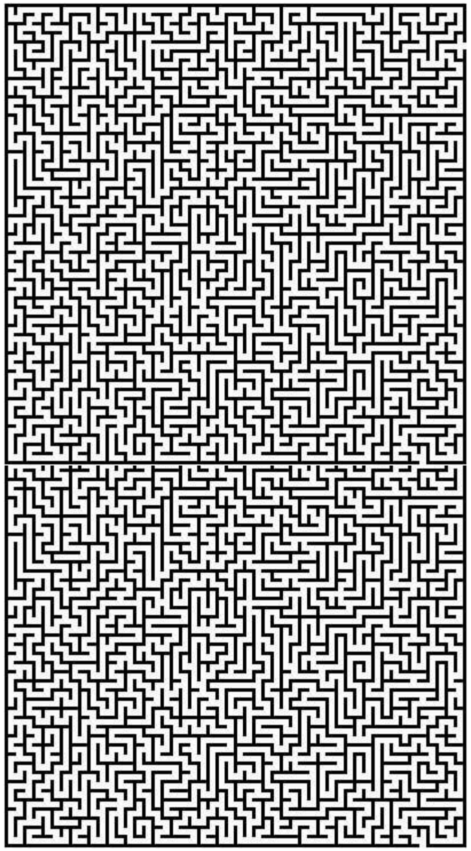
ENGLISH OPENING

	White	Black		White	Black
	Vitiugov	Wang		Vitiugov	Wang
1	c4	e5	20	a4	a6
2	Nc3	Nf6	21	Nd2	Bf8
3	Nf3	Nc6	22	Nc4	b5
4	g3	d5	23	ab5	ab5
5	cd5	Nd5	24	Nd2	Nb6
6	Bg2	Nb6	25	Nf3	Na4
7	0-0	Be7	26	Qc2	c5
8	a3	0-0	27	Be5	cb4
9	b4	Be6	28	Qb3	Nc3
10	d3	Nd4	29	Re1	Qd7
11	Bb2	Nf3	30	Kg2	Qf5
12	Bf3	c6	31	Ra7	Re5
13	Ne4	Nd7	32	Ne5	Qe5
14	Qc2	Bd5	33	Qf7	Kh8
15	Bc3	Re8	34	Rea1	Qf6
16	Rfd1	Rc8	35	Qb3	Ne2
17	Qb2	Bf8	36	R1a6	Qf5
18	Nd2	Bf3	37	Resigns	
19	Nf3	Bd6			

Temporal credit
Assignment problem

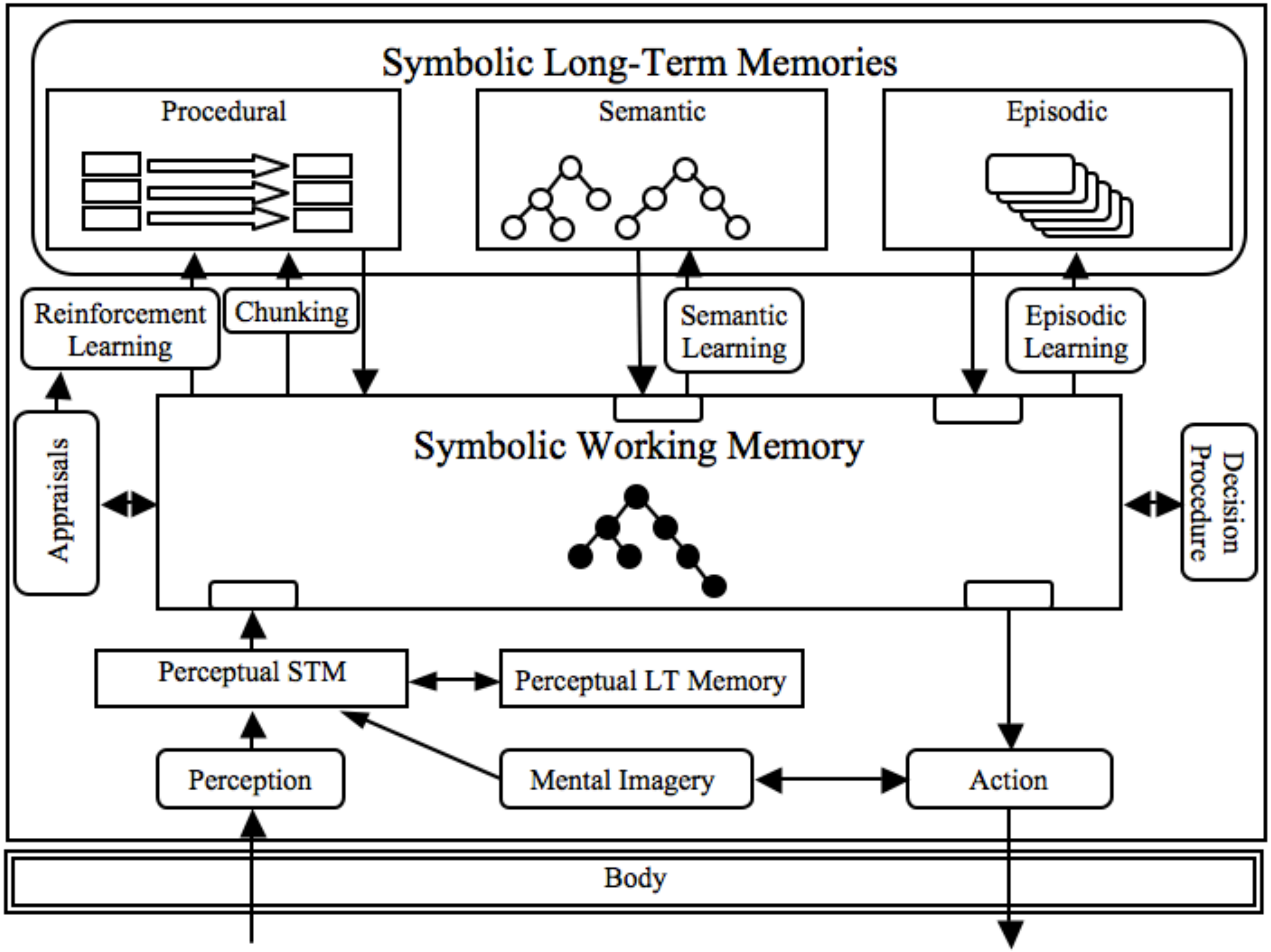


Exploration / exploitation tradeoff



Curse of dimensionality

Soar 9.3.1


Symbolic Long-Term Memories
Procedural
Semantic
Episodic
Reinforcement Learning
Chunking
Semantic Learning
Episodic Learning
Appraisals
Symbolic Working Memory
Decision Procedure
Perceptual STM
Perceptual LT Memory
Perception
Mental Imagery
Action
Body

Soar-RL

- Reward

```
<state>
  ^reward-link
     ^reward
        ^value float
```

- Value representation

```
sp {rl*move*left
   (state <s> ^name left-right
              ^operator <op> +)
   (<op> ^name move ^dir left)
-->
   (<s> ^operator <op> = -1.0) }
```

- Decisions

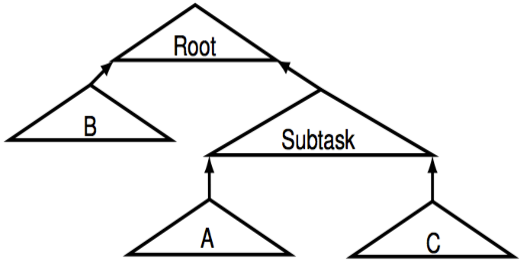
Move*left	-0.8
Move*right	-0.2
Move*sit	-1.2

- Adaptive behavior

Soar-RL Talks

- Modular RL in Soar
  Shiwali Mohan

- Improving Off-Policy HRL
  Mitchell Keith Bloch



- Learning to Use Memory
  Nick Gorski