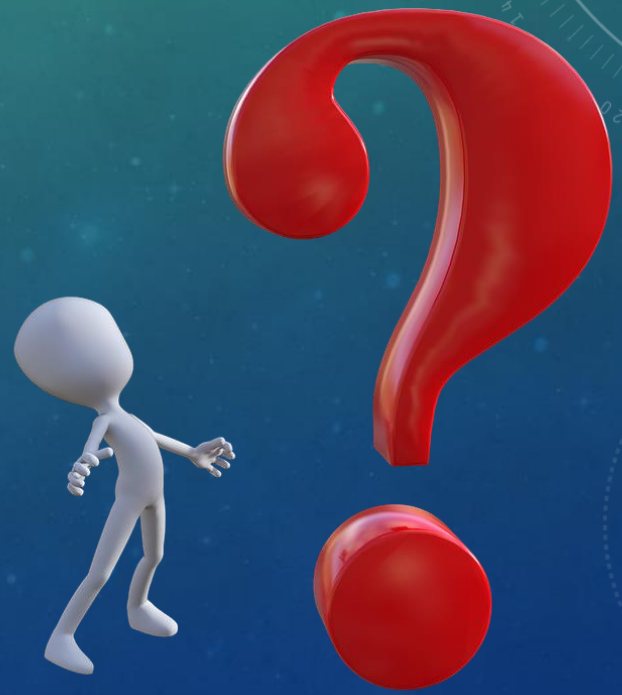
AGENTS IN REINFORCEMENT LEARNING

ANDER SALABERRIA

ENERITZ DOMINGUEZ

WHAT IS STARCRAFT II?

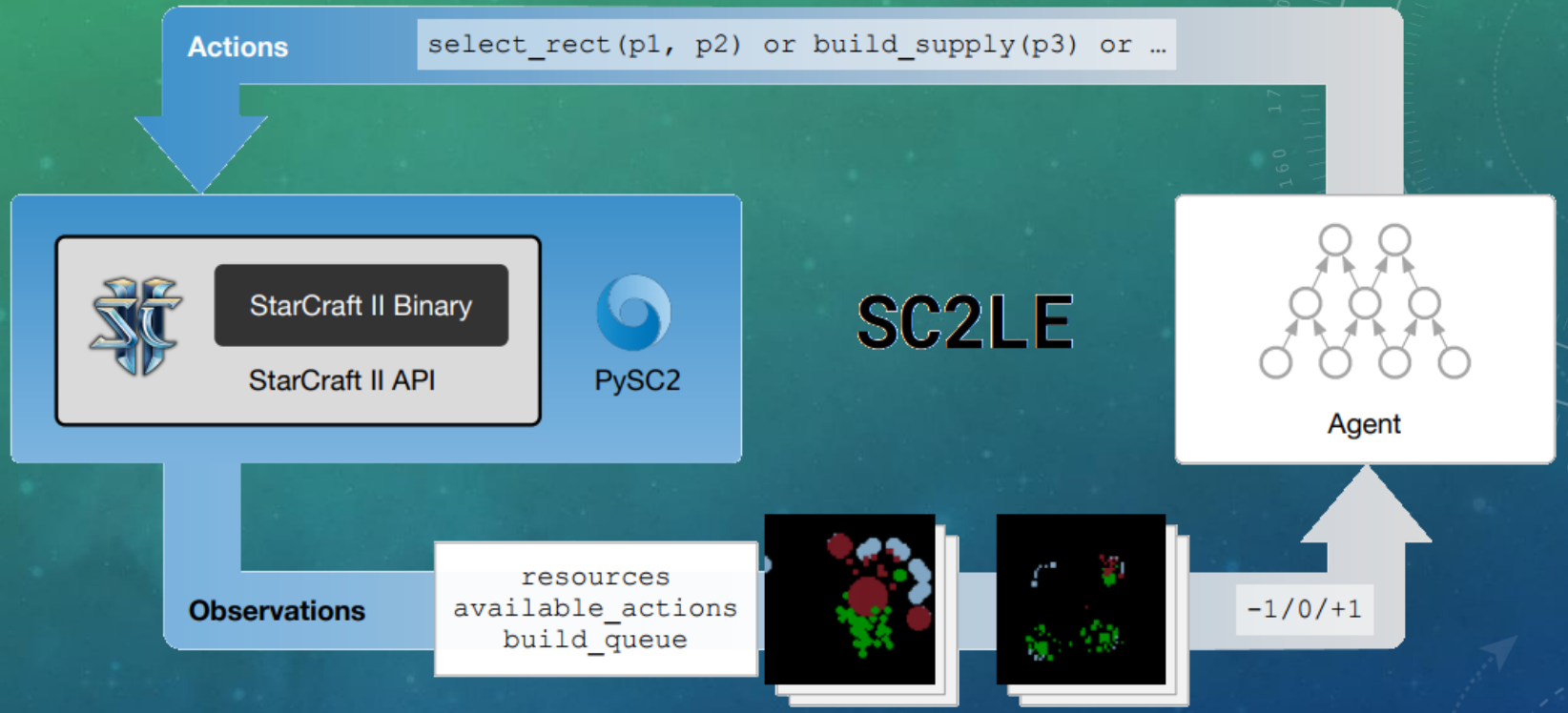
- Real Time Strategy game
- Gain resources and build armies to defeat the enemy
- Information is key to success



DEEP MIND

- Pushing the boundaries of AI since 2010
- Bought by Google (2014)
- Alpha Go beats Go world champion (2017)

SC2LE



- Starcraft 2 API
- PySC2
- Starcraft 2 Binary (the game itself)

MINI GAMES

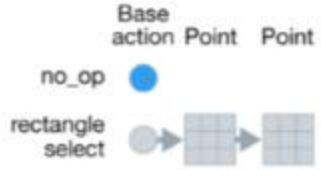
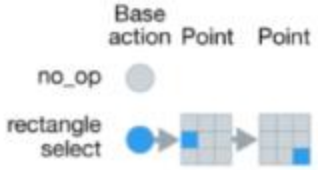
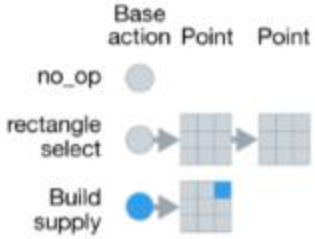
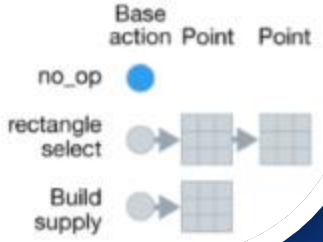
- Small scenarios of the full-game
- Testing a subset of actions
  - Movement
  - Collecting resources
  - Building units
  - …

OBSERVATIONS

- Feature layers
- Two sets:
  - Minimap layers
  - Window layers

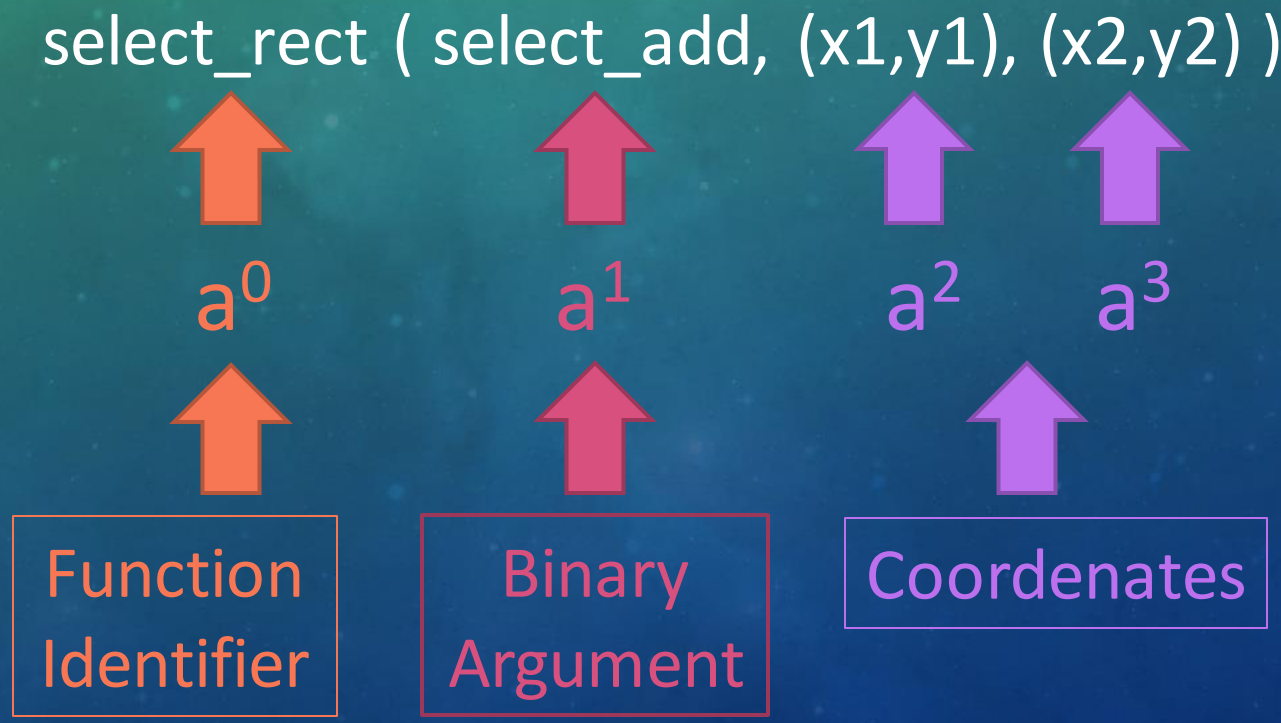
ACTIONS




Human Actions	IDLE	Left_Click_Hold (p1) Release (p2)	Press B + S Left_Click (p3)	IDLE
Agent Actions	no_op	select_rect(p1, p2)	build_supply(p3)	no_op
Available Actions				

ACTIONS

- 300 action-function identifiers
- 13 possible types of arguments



RL – LEARNING ALGORITHM

- A deep neural network with parameters θ defines a policy π_θ
- At time step t with a state s_t, the agent selects an action a_t with $\pi_\theta(a_t|s_t)$ probability
- Then, it receives a r_t reward
- In order to learn θ, A3C is used

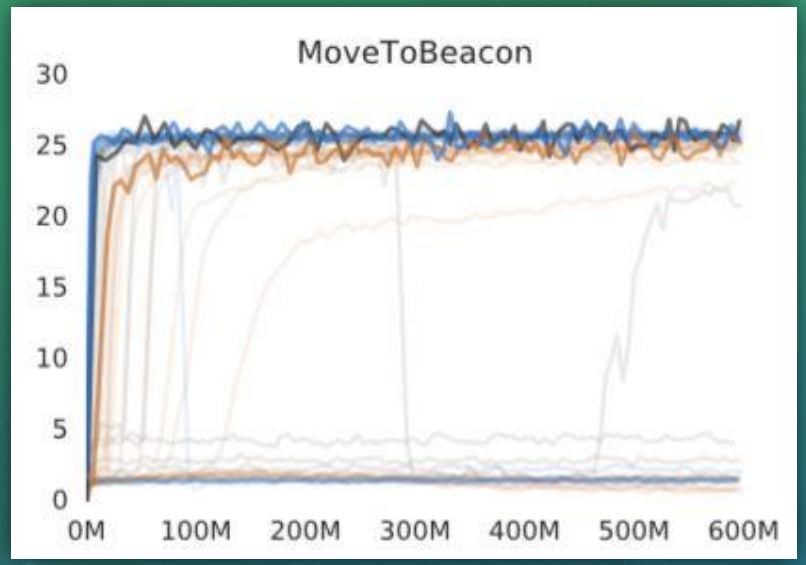
$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

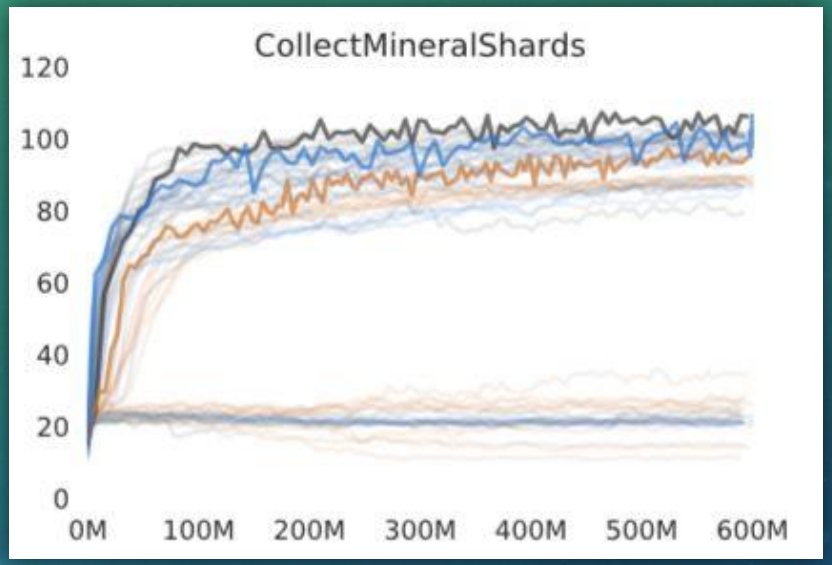
RL – POLICY REPRESENTATION

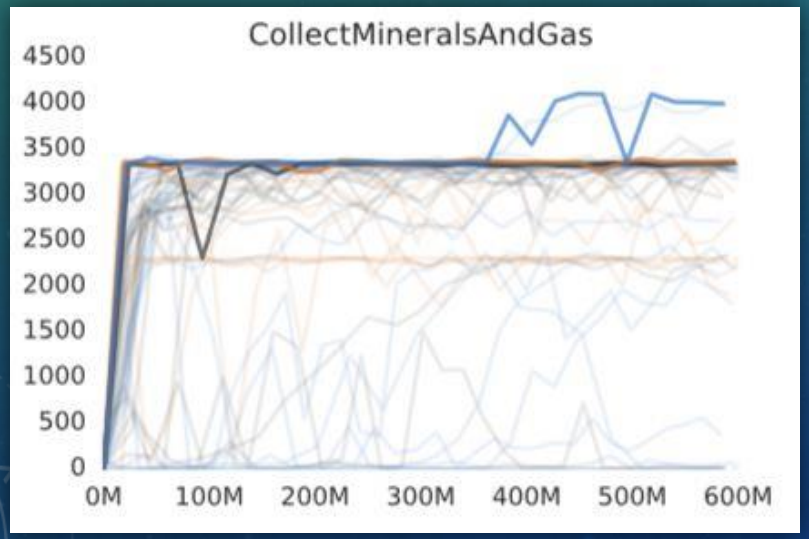
- In each state, millions of possible actions
- The policy is represented in an autoregressive manner, using this chain rule.

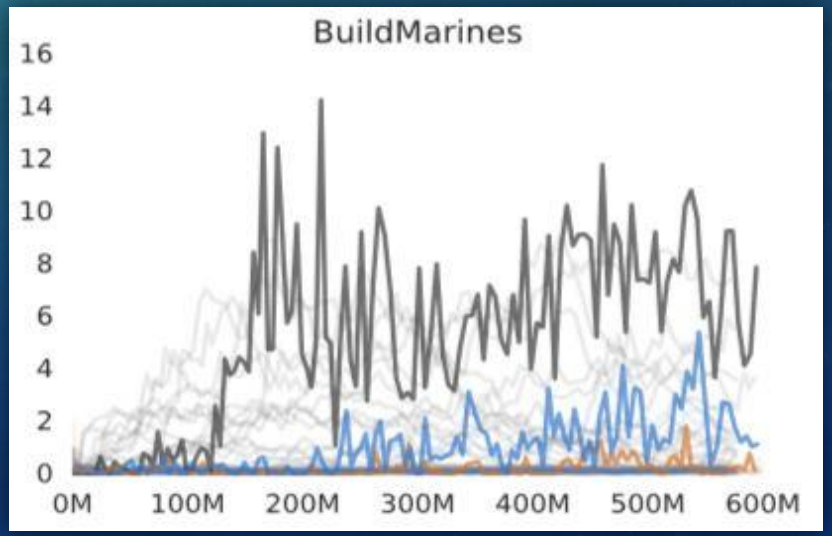
$$\pi(a|s) = \prod_{l=0}^{L} \pi(a^l|a^{<l}, s)$$

MINIGAMES RESULTS









MINIGAMES RESULTS

- Even relatively simple mini-games present interesting challenges
- Better not to talk about the full game results…

Agent	Metric	MoveToBeacon	CollectMineralShards	CollectMineralsAndGas	BuildMarines
DeepMind human player	MEAN	26	133	6880	138
	MAX	28	142	6952	142
StarCraft GrandMaster	MEAN	28	177	7566	133
	MAX	28	179	7566	133
Atari-net	BEST MEAN	25	96	3356	< 1
	MAX	33	131	3505	20
FullyConv	BEST MEAN	26	103	3978	3
	MAX	45	134	4130	42
FullyConv LSTM	BEST MEAN	26	104	3351	6
	MAX	35	137	3995	62

BIBLIOGRAPHY

- Article - “Release of the SC2LE” by DeepMind

  https://deepmind.com/blog/deepmind-and-blizzard-release-starcraft-ii-ai-research-environment/

- Paper - “StarCraft II: A New Challenge for Reinforcement Learning” by DeepMind & Blizzard

  https://deepmind.com/documents/110/sc2le.pdf

- Video - “A guide to DeepMind’s StarCraft AI Environment” by Siraj Raval

  https://www.youtube.com/watch?v=URWXG5jRB-A