

Regular Decision Processes

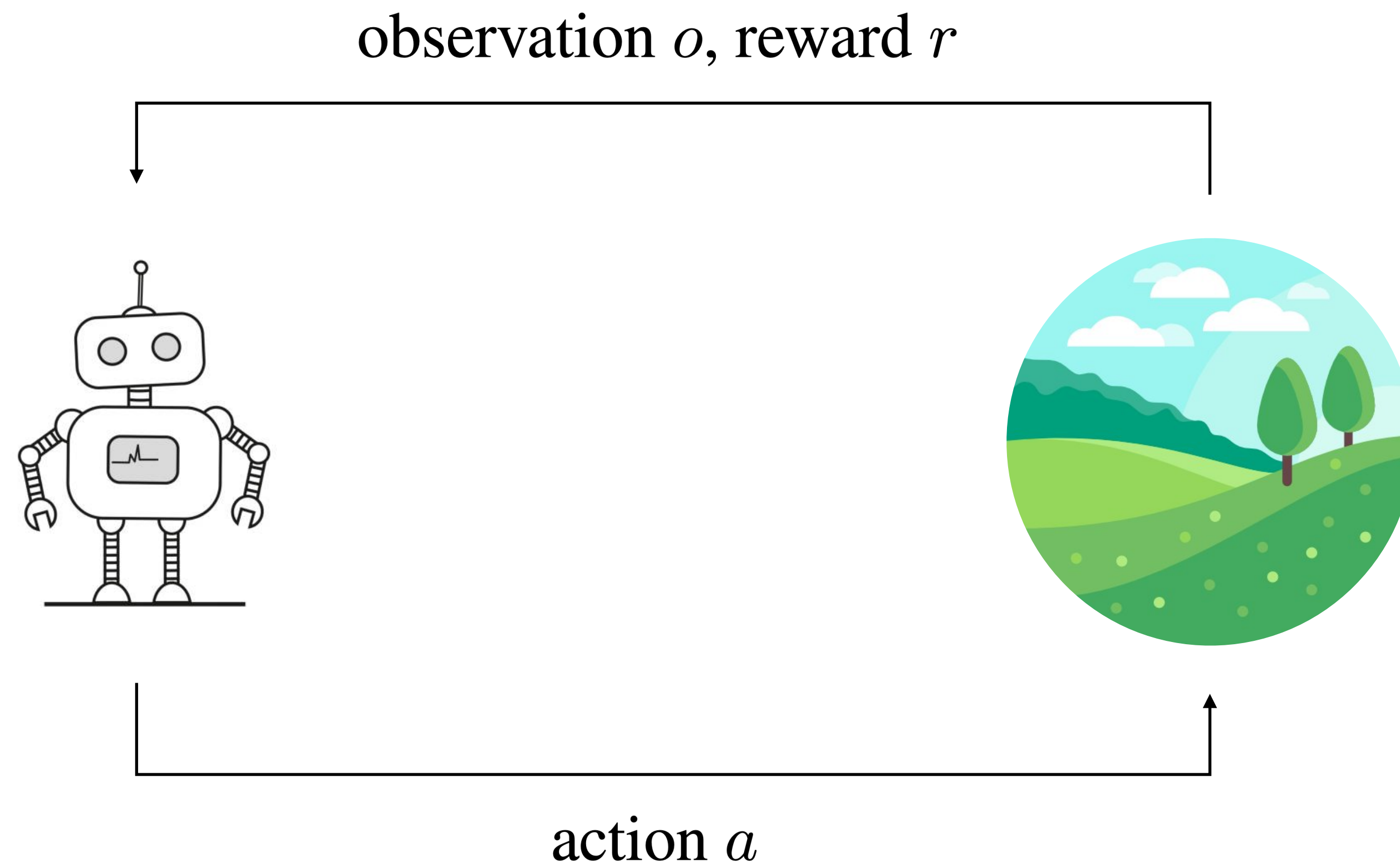
Alessandro Ronca

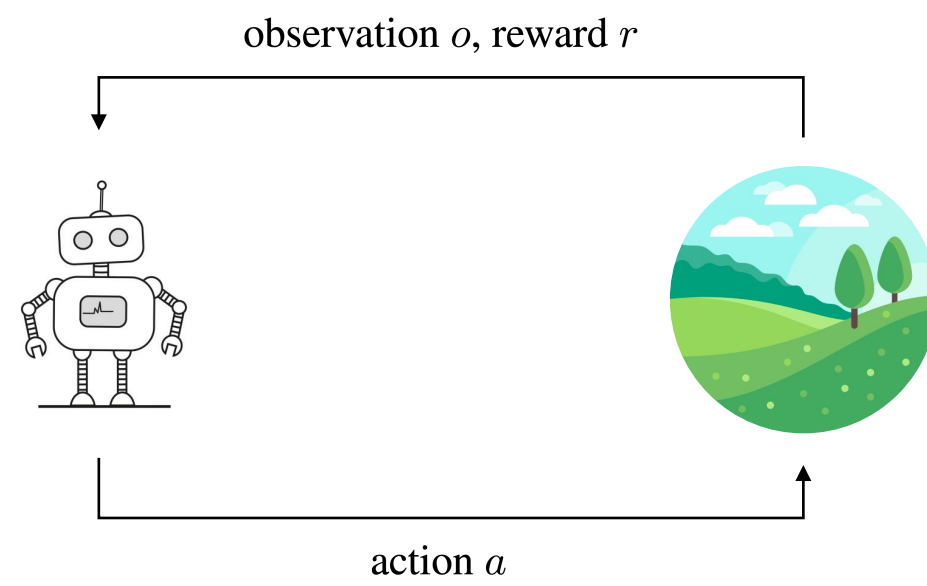
ronca@diag.uniroma1.it

9 April 2021

An agent interacts with the environment

Performing **actions**, receiving **observations** and **rewards**





Non-Markov setting

history: $h = o_1 o_2 o_3 \dots o_n$

next observation: $h, a \mapsto P(O)$

next reward: $h, a \mapsto P(R)$



Automaton

- Given (LTLf/LDLf)
- Learned

finite state space: $S = \{s_1, \dots, s_m\}$

deterministic function: $h \mapsto s_i$



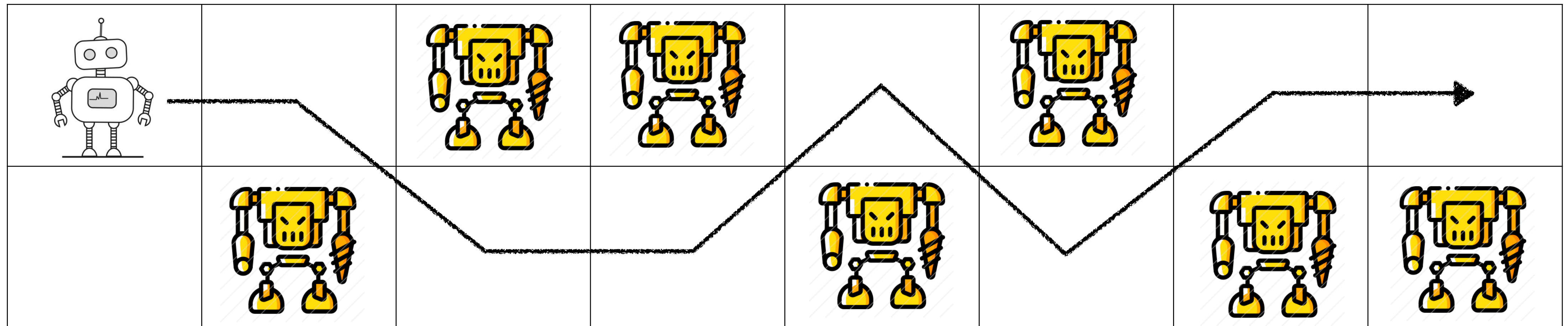
Markov setting (i.e., MDP)

next observation: $s_i, a \mapsto P(S)$

next reward: $s_i, a \mapsto P(R)$

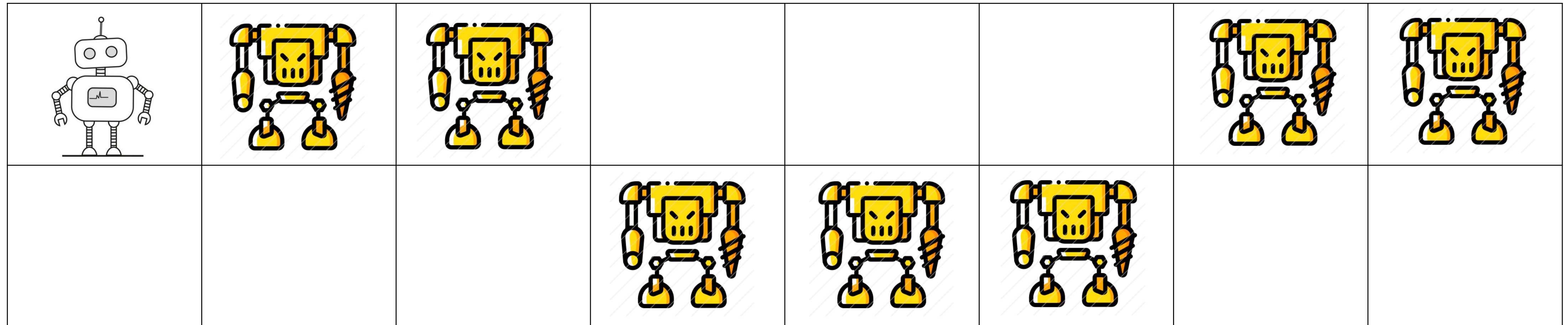
Example of an RDP

An agent that has to avoid enemies across a corridor



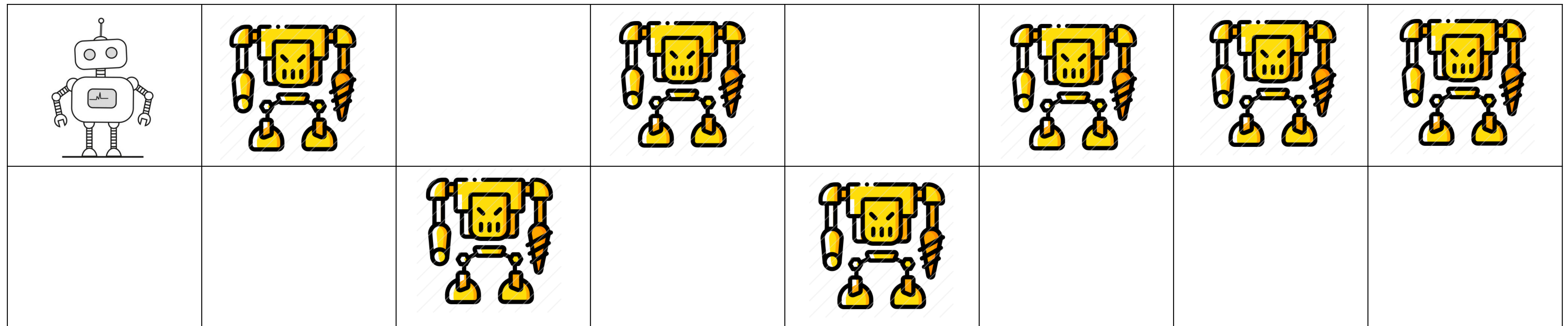
Example of an RDP

An agent that has to avoid enemies across a corridor



Example of an RDP

An agent that has to avoid enemies across a corridor



Example of an RDP

An agent that has to avoid enemies across a corridor

	???	???	???	???	???	???	???
	???	???	???	???	???	???	???

Example of an RDP

An agent that has to avoid enemies across a corridor



Example of an RDP

An agent that has to avoid enemies across a corridor



Enemies change strategy every time the agent hits an enemy! \longrightarrow RDP

References

Ronen I. Brafman, Giuseppe De Giacomo:
Regular Decision Processes: A Model for Non-Markovian Domains. IJCAI 2019: 5516-5522

Eden Abadi, Ronen I. Brafman:
Learning and Solving Regular Decision Processes. IJCAI 2020: 1948-1954

Alessandro Ronca, Giuseppe De Giacomo:
Efficient PAC Reinforcement Learning in Regular Decision Processes. Under review.
(write to me at ronca@diag.uniroma1.it to obtain a personal copy)