Active Perception in Navigation of Partially Observable Grid Worlds

Paul A. Crook & Gillian Hayes

paulc@dai.ed.ac.uk

Institute of Perception, Action and Behaviour

School of Informatics

University of Edinburgh

http://www.dai.ed.ac.uk/homes/paulc

Breakdown Of Talk

- Introduction
- Review of Active Perception Work
- Developments in the Reinforcement Learning of Partially Observable Markovian Decision Processes (POMDPs)
- Experimental Results from Our Approach

Introduction

- What are Partially Observable Environments?

Introduction

- What are Partially Observable Environments?
  - An environment where the observation received by an agent may not uniquely identify state.

Introduction

- What are Partially Observable Environments?
  - An environment where the observation received by an agent may not uniquely identify state.
- What is Active Perception?

Introduction

- What are Partially Observable Environments?
  - An environment where the observation received by an agent may not uniquely identify state.
- What is Active Perception?
  - The ability of an agent to direct its perception of the environment.

Introduction

- Why of Interest?

Introduction

- Why of Interest?
  - Consider a simple, reactive, model-free agent that is situated and embodied in its environment.

Introduction

- Why of Interest?
  - Consider a simple, reactive, model-free agent that is situated and embodied in its environment.
  - Such an agent will probably find its environment to be partially observable.

Introduction

- Why of Interest?
  - Consider a simple, reactive, model-free agent that is situated and embodied in its environment.
  - Such an agent will probably find its environment to be partially observable.
  - The responses of a such an agent equipped with fixed perception is limited to the number of states it can distinguish.

Introduction

- Why of Interest? (continued)

Introduction

- Why of Interest? (continued)
  - A similar agent with active perception is more flexible. Provided it can find some distinction between states that the fixed agent regarded as identical it can then vary its response.

Introduction

- Why of Interest? (continued)
  - A similar agent with active perception is more flexible. Provided it can find some distinction between states that the fixed agent regarded as identical it can then vary its response.
  - I'm interested in how difficult it is for such an agent to learn to utilise the flexibility afforded it by active perception.

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	**Actions**	**States**
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI’91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI’92)]**	Generalises; through storing statistics	No Separation	

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	Actions	States
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI'91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI'92)]**	Generalises; through storing statistics	No Separation	

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	Actions	States
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI'91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI'92)]**	Generalises; through storing statistics	No Separation	

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	**Actions**	**States**
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI’91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI’92)]**	Generalises; through storing statistics	No Separation	

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	Actions	States
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI'91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI'92)]**	Generalises; through storing statistics	No Separation	

Previous Approaches

Previous Approaches in Learning Active Perception with Reinforcement Learning:

	Applicable Environments	Actions	States
Whitehead (PhD 1992, Rochester) **Ming Tan (ICAI'91)**	Limited to Deterministic Environments	Perceptual Actions treated separately	Learns Model Relating Observations to States
McCallum (PhD 1996, Rochester) **[Chrisman (AAAI'92)]**	Generalises; through storing statistics	No Separation	

Developments in RL of POMDPs

- Identification of “optimal memoryless policies” (Littman, SAB’94).

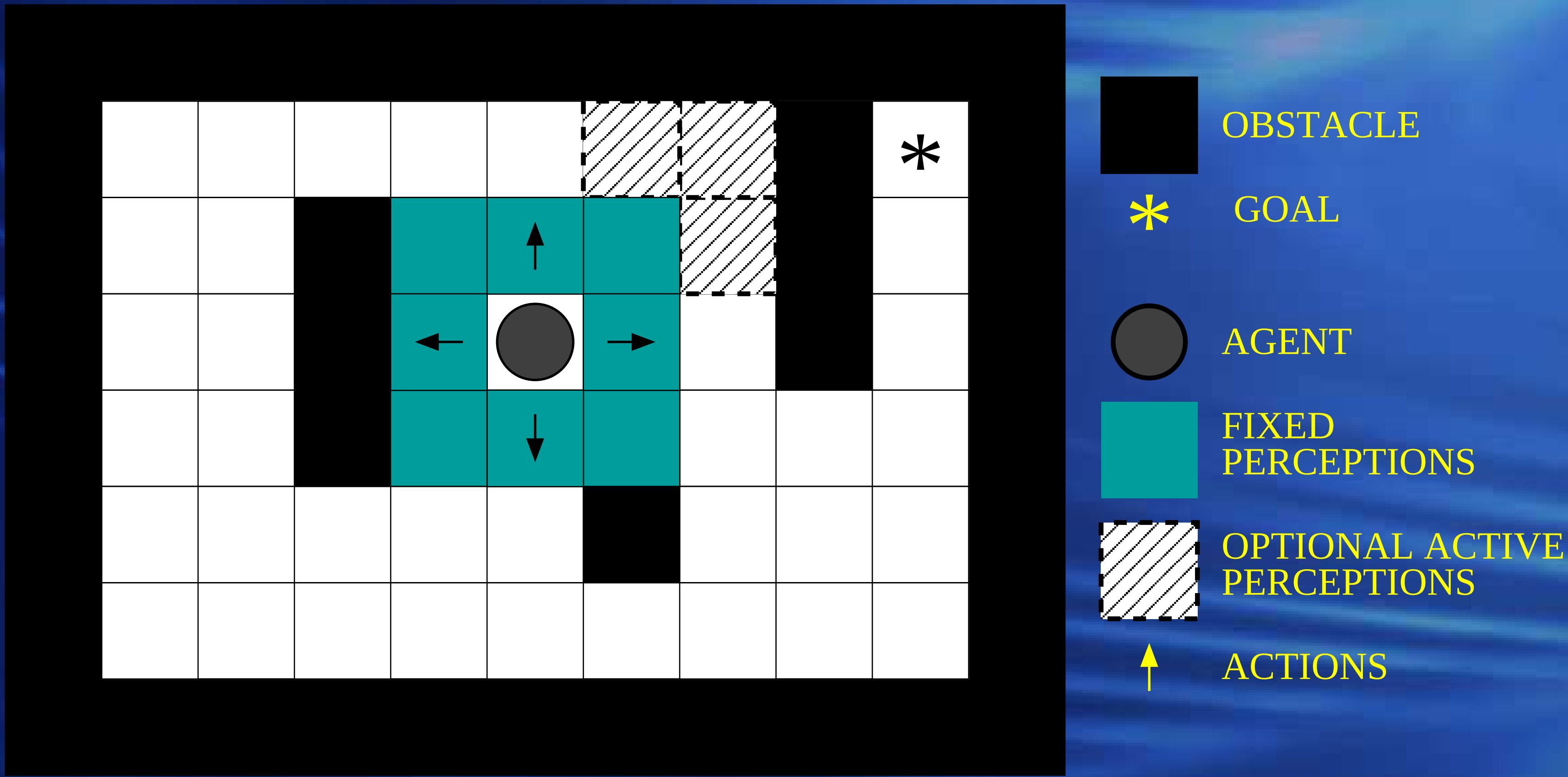
Developments in RL of POMDPs

- Identification of "optimal memoryless policies" (Littman, SAB'94).
- Empirical observation that SARSA(λ) is robust in partially observable environments (Loch & Singh, ICML'98).

Developments in RL of POMDPs

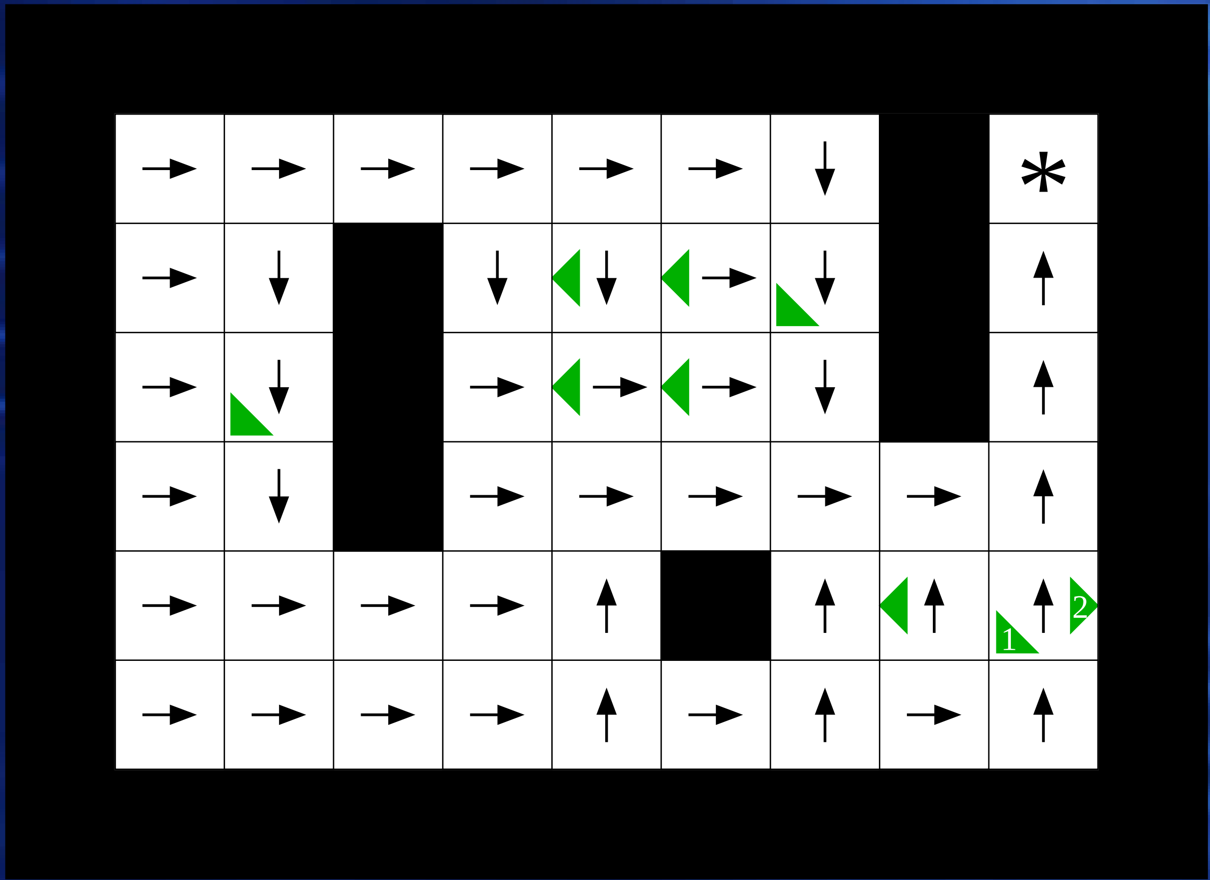
- Identification of "optimal memoryless policies" (Littman, SAB'94).
- Empirical observation that SARSA(λ) is robust in partially observable environments (Loch & Singh, ICML'98).
- Recent work indicates that provided certain conditions are met, then there is a guarantee of convergence to a consistent policy, rather than oscillating between policies (T.J.Perkins ICML'02),
  (though this could be a double edged sword)

Experiment



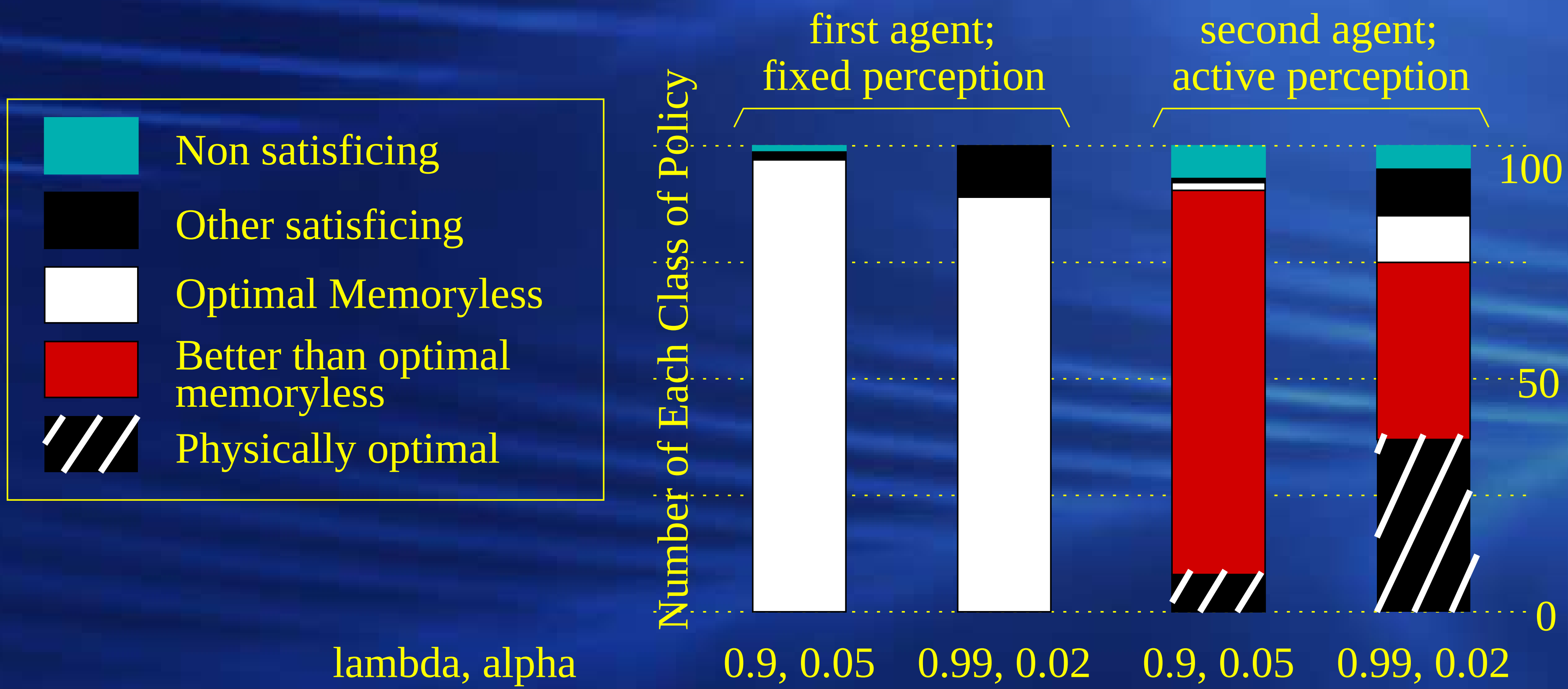
*
OBSTACLE
GOAL
AGENT
FIXED PERCEPTIONS
OPTIONAL ACTIVE PERCEPTIONS
ACTIONS

Results: A Typical Learnt Policy



Results: Policies Learnt

Policies classified based on Total Physical Actions (*i.e.* Perceptual Actions are not counted).



Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.

Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.
- Future Work:

Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.
- Future Work:
  - Why not more physically optimal policies?

Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.
- Future Work:
  - Why not more physically optimal policies?
  - Look at the effects of Exploration.

Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.
- Future Work:
  - Why not more physically optimal policies?
  - Look at the effects of Exploration.
  - Try other environments; especially stochastic environments and ultimately use a real robot.

Conclusions & Future Work

- Main Conclusion: demonstration that a simple agent can learn to use active perception to learn physically optimal solutions.
- Future Work:
  - Why not more physically optimal policies?
  - Look at the effects of Exploration.
  - Try other environments; especially stochastic environments and ultimately use a real robot.

Welcome feedback & suggestions