Learning From Multiple Sources

Luís Nunes ..... *luis.nunes@iscte.pt*
Eugénio Oliveira ..... *eco@fe.up.pt*

Summary:
- ***Objective***
- ***Concepts***
- ***Experiments***
- ***Results***

Objective

Use **information from different sources** during the learning process to **improve learning performance**

Environments' characteristics:
- Partially observable and dynamic
- **Several Agents** dealing with similar problems
- Using **different learning algorithms**
- Communicating
- Rewards:
  Immediate/delayed + individual/team

Concepts

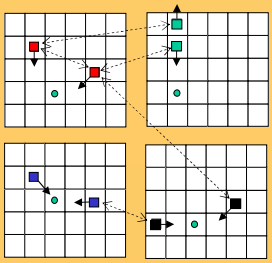
- **Reinforced Learning** (environment feedback)
  - Evolutionary Algorithms (EA)
  - Q-Learning (QL)
  - Genetic-Programming (GP)
- **Supervised Learning** from peers' advice
  - EA+Backprop
  - QL+ Virtual Experience/Imitation/Bonus
  - GP+ID3
- **Heterogeneous vs. Homogeneous** environments
- **Roles** (static)
- **Trust and "Learnability"**
- **Learning Stages**: Exploration, Novice, Intermediate, Expert
  - **Changing: Advice Type / Learning Parameters**

Information exchanged

- **Advice**:
  - Specific/Batch
  - Online/Offline
  - Standing/Multiple advisor(s)
- **Rewards**:
  - Combined validation
  - Specific (for a given state)

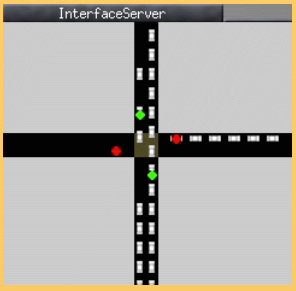
Experiment 1: Pursuit Problem

- 15x15 arena
- Limited vision, spherical world
- Individual and cooperative catch



Experiment 2: Traffic Control

- Traffic Light Control
- Simulation based on real data
- Simplified car movement



Experiment 3: Load Balance

- Based on Whiteson and Stone 2005
- Hidden-state problems
- Constant Load: 500 jobs

Results

- Information exchange **improves performance** (typical) at the expense of communication
- **Improvement over advisors' skill** (problem dependent)
- **Heterogeneity helps low-performance agents** (problem dependent)
- **Stable interaction** of different learning algorithms - inter/intra agent – (typical)
- Most add-ons (trust, roles, learning-stages, etc.) **do not produce significant changes** (problem dependent)

Dead-Ends (... was Future Work last year)

- Storage of information in a common format: what format?
- Combination of advice from several sources: confusion!
- Influence-exchange: even more confusion!
- Team Supervisors: centralization!
- State generalizations in stored experience: difficult for some learning algorithms.
- Dynamic role assignment: too slow.

Still in evaluation:

- **Changing Learning Parameters**
- **Exception Lists**
- **Filters**

The End (hopefully ...)

Thank you for your attention ...