INTELLIGENT AGENTS AND SYNTHETIC CHARACTERS GROUP

Socio-Emotional Reward Design for Intrinsically Motivated Learning Agents

Thesis Defense

Ph.D. Program in Information Systems and Computer Engineering

Pedro Sequeira

September 18, 2013

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

Emerging Emotions

Socially-Aware Learning Agents

Conclusions

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

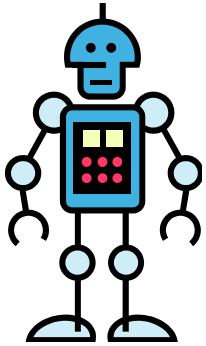
Emerging Emotions

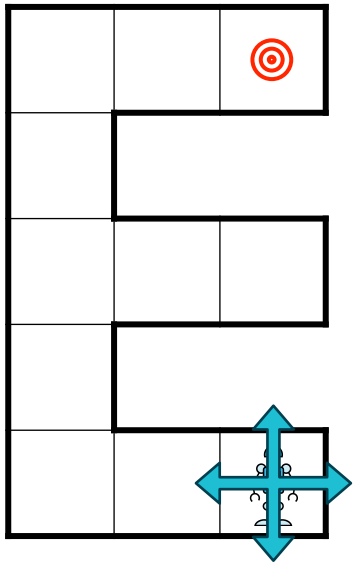
Socially-Aware Learning Agents

Conclusions

Reinforcement Learning (1/2)

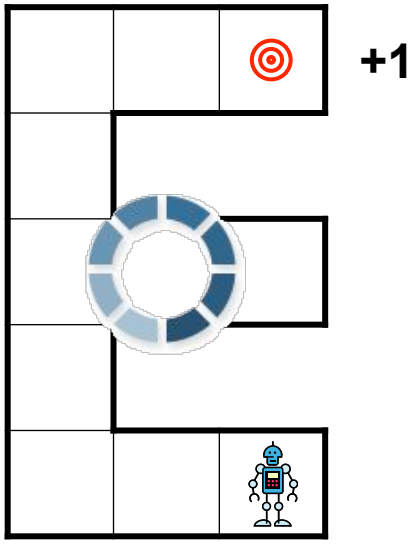
- Simple navigation problem
  - *State:* agent location
  - *Actions:* movement





Reinforcement Learning (2/2)

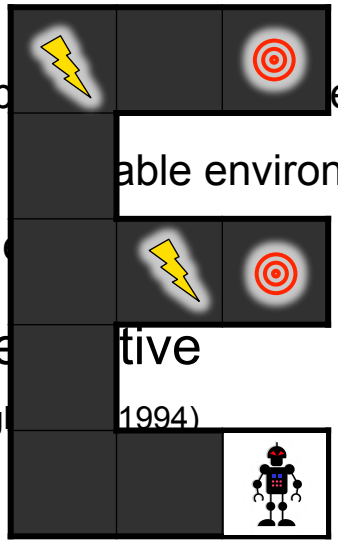
- RL main ideas (Sutton & Barto, 1998; Kaelbling *et al.*, 1996)
  - Objective: maximize the reward throughout time
  - Task: discover which actions maximize reward in each state
    - *e.g.*, using Q-learning (Watkins, 1989)



Challenges in RL

- Assumptions (Sutton & Barto, 1998; Kaelbling *et al.*, 1996)
  - Fully observable environments
  - Infinite visits to all states and actions
  - Stationary environments
- Agent limitations
  - Limited perception and comp es
  - Dynamic, unpredictable and able environment
  - Demand for manual adjustm
- Design assumptions too re tive

  (Littman, 1994; Loch and Singh, 1998; Sing 1994)



Reward Mechanism

- Rewards are fundamental (Sutton & Barto, 1998; Kaelbling *et al.*, 1996)
  - Implicitly defines the agent's *task*
  - Impact on the *learning time*
  - Impact on *what* is learned
- Major challenge (Abbeel and Ng, 2004; Ng and Russell, 2000; Sorg et al., 2010a)
  - Build reward mechanisms so the task is learned efficiently
  - Flexible and robust
  - Enhance agent's *autonomy*

Research Statement

*"Design **reward mechanisms for RL** agents that are able to*

***alleviate** their inherent perceptual limitations and*

*make them operate in a **wide variety of domains***

*without the explicit **intervention of others** or relying on*

***expert or domain knowledge** about a particular task."*

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

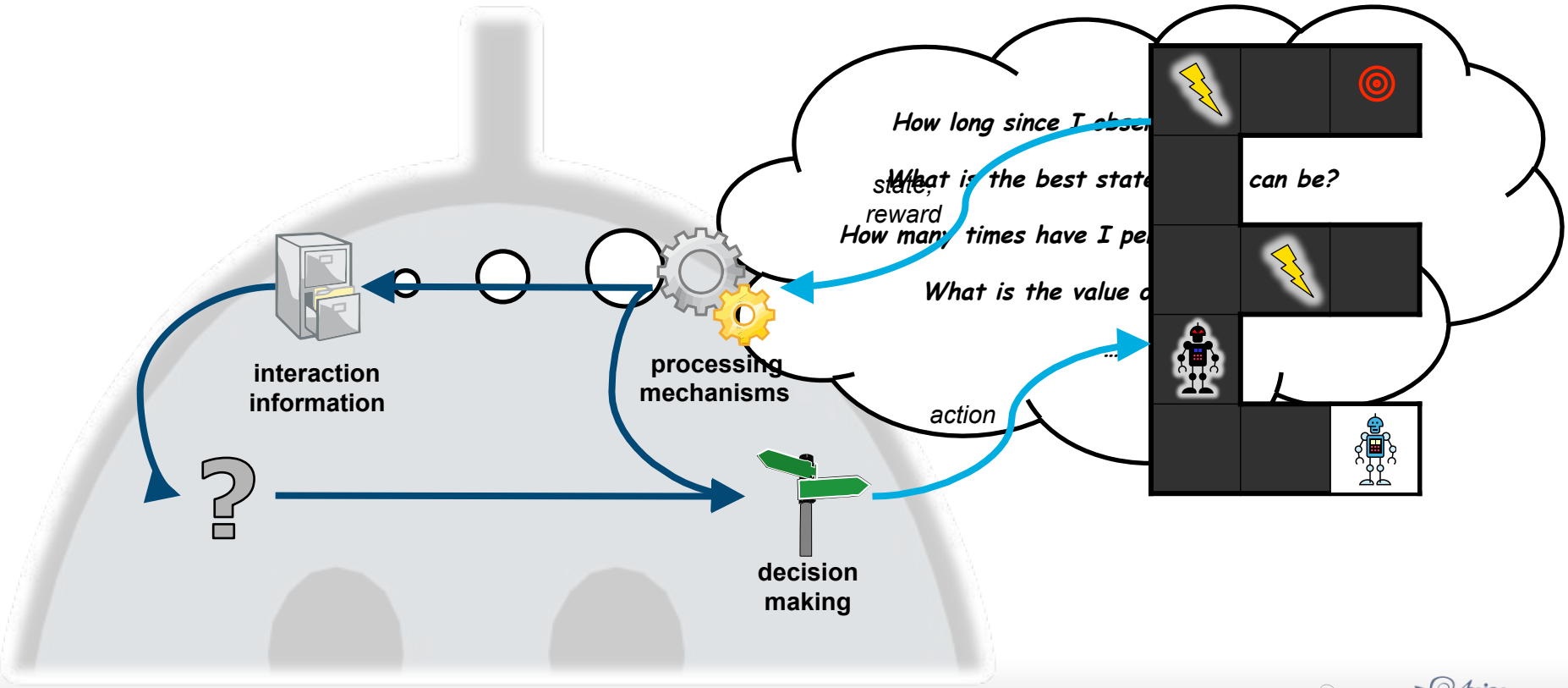
Emerging Emotions

Socially-Aware Learning Agents

Conclusions

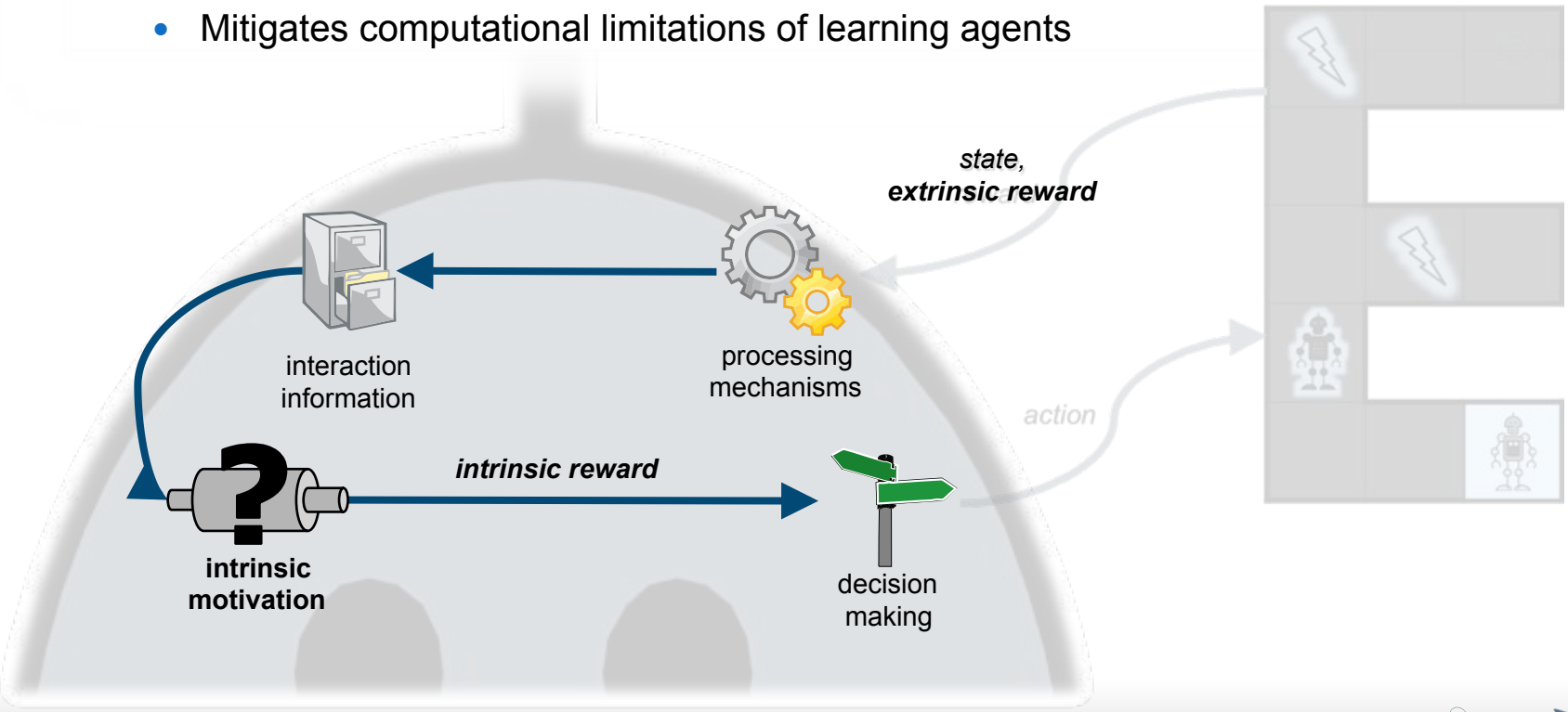
Sources of Information

- Agent has access to several sources of information
  - Perception, reasoning, learning, etc.



Intrinsic Motivation

- Intrinsically Motivated RL (Singh *et al.* 2009, 2010; Sorg *et al.* 2010)
  - Agent learns with intrinsic rewards
  - Fitness: measures performance
  - Mitigates computational limitations of learning agents



Inspiration

- Parallel with biological organisms
  - Limited perception and resources
  - Dynamic, unpredictable and unreliable environment
- Natural motivational mechanisms
  - Shaped by evolution
  - Provide adaptive advantages
- Social mechanisms
  - Cooperation in inherently competitive environments

Hypothesis

*"We focus on the **role of emotions** and also on the*

*way **individuals interact and cooperate** with each other as a social group*

*to **design more flexible and robust reward mechanisms** that*

***enhance the autonomy** of RL agents in*

both single and multiagent settings."

Contributions

- Emotion-based Intrinsic Rewards
  - Role of emotions in decision-making
  - 4 emotion-based domain-independent reward features
- Emerging Emotions
  - Emergence of useful sources of information
  - Discuss relation with emotions
- Socially-Aware Learning Agents
  - Extend IMRL to multiagent scenarios
  - Socially-aware behaviors

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

Emerging Emotions

Socially-Aware Learning Agents

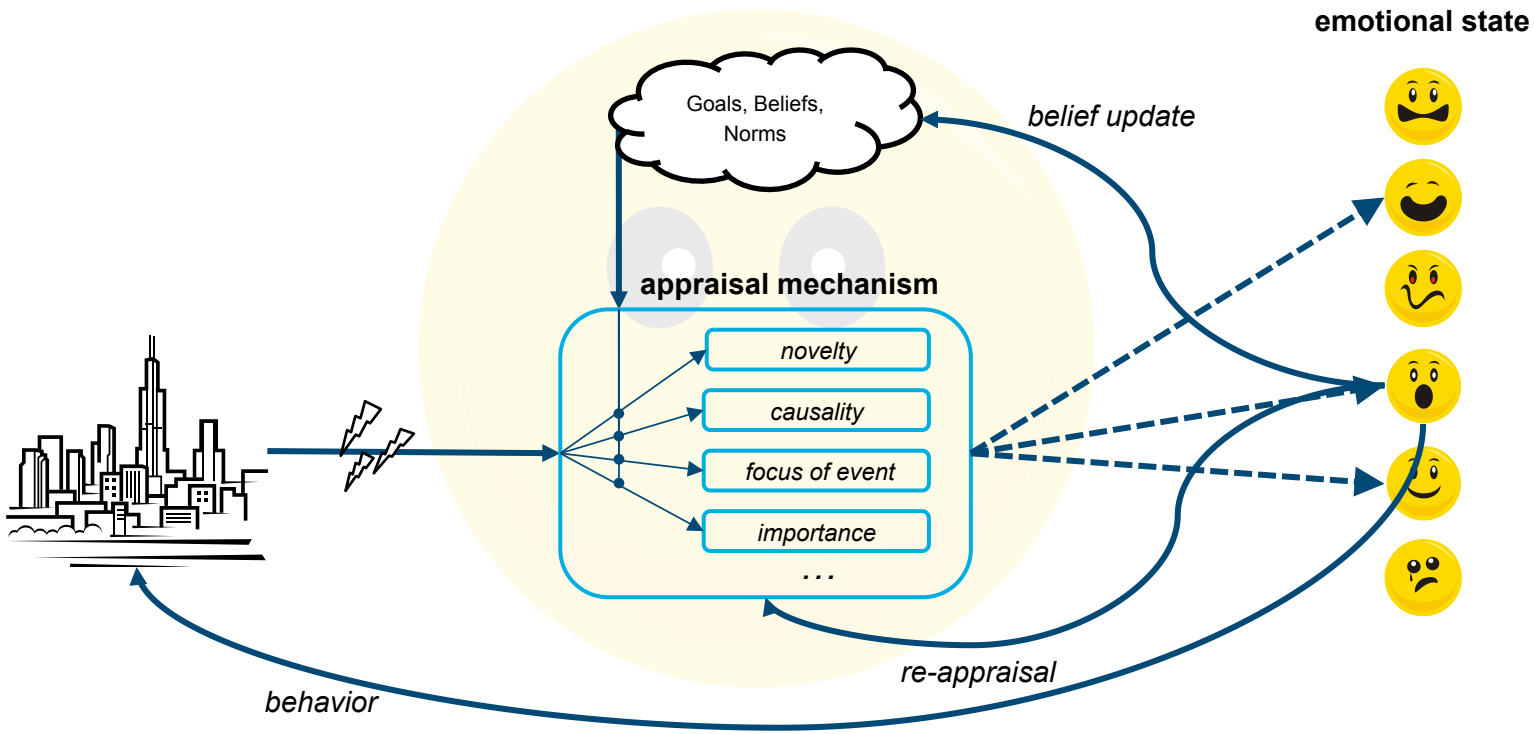
Conclusions

The Role of Emotions

- Studies show that emotions:
  (Cardinal *et al.* 2002, Dawkins 2000, Phelps & LeDoux 2005)
  - Basic and ancient survival mechanism
  - Beneficial adaptive mechanism for decision-making
  - Elicit physiological signals
- Absence of emotions
  (Bechara *et al.* 2000, Damasio 1994, LeDoux 2000)
  - Impairs taking advantageous decisions
- *How to provide emotion-based motivation?*

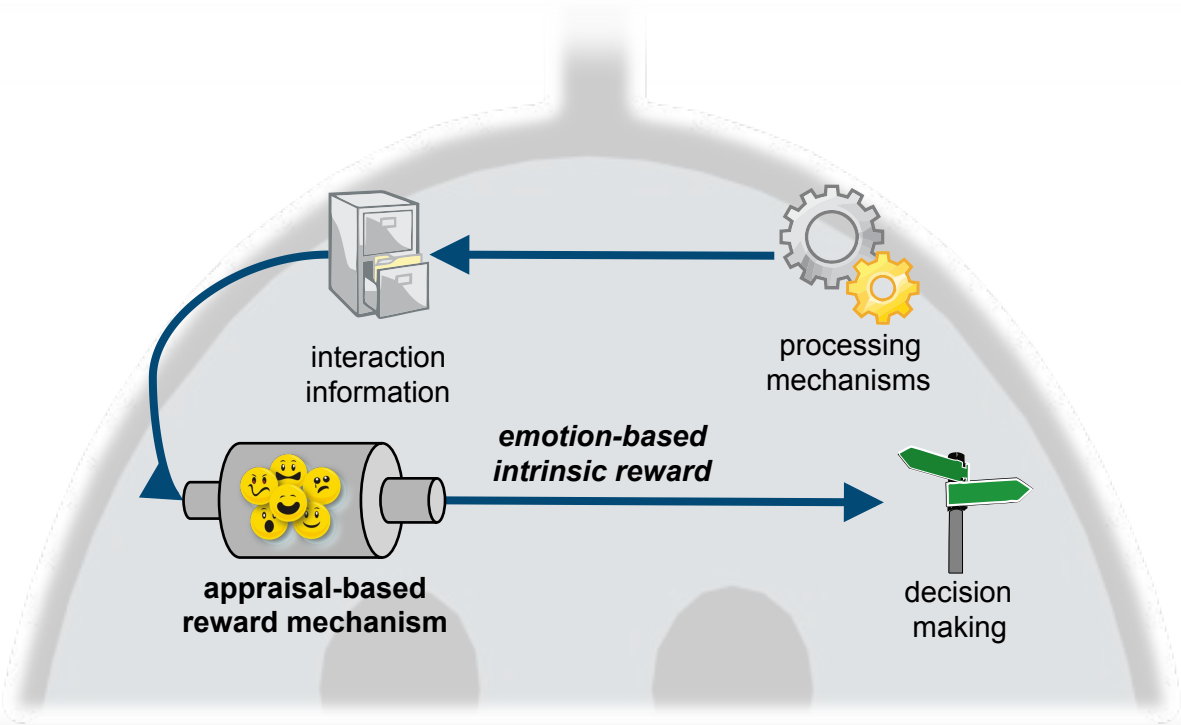
Appraisal Theories of Emotions

- Evaluation through appraisal
  (Frijda and Mesquita 1998, Lazarus 2001, Reisenzein 2009, Roseman 2001, Scherer 2001)



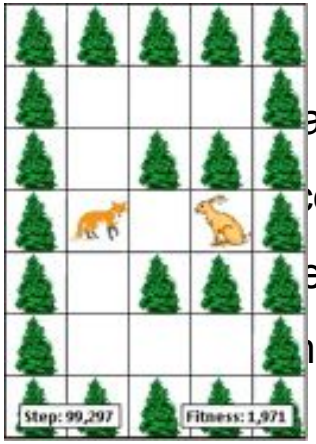
Emotion-based Reward Design

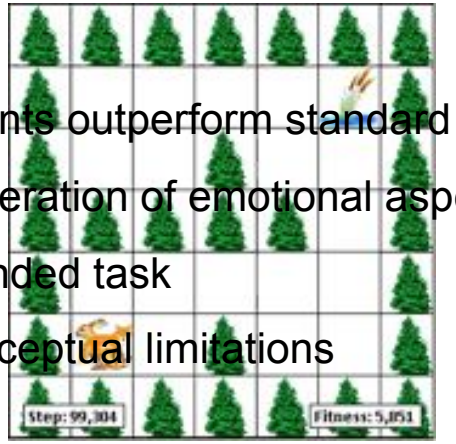
- *Major dimensions* of appraisal (Ellsworth & Scherer 2003, Leventhal & Scherer, 1987)
  - 4 domain-independent reward features
  - Evaluate agent's history of interaction with environment

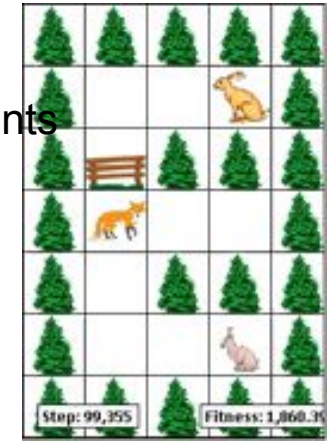


Experiments

- Foraging scenarios
  - Each presents distinct challenge
  - Partially-observable
  - Compare performance emotion-based vs. fitness-based







al agents outperform standard agents

consideration of emotional aspects

e intended task

ne perceptual limitations

Contributions

- 4 emotion-based reward features
  - Novelty, Valence, Goal relevance and Control
  - Domain-independent
  - General-purpose guiding system for RL agents
  - Mitigation of perceptual limitations
  - Departs from previous works within Affective Computing
    - Domain-independent appraisal-based
    - Does not alter RL algorithm
    - Does not focus on a set of basic emotions

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

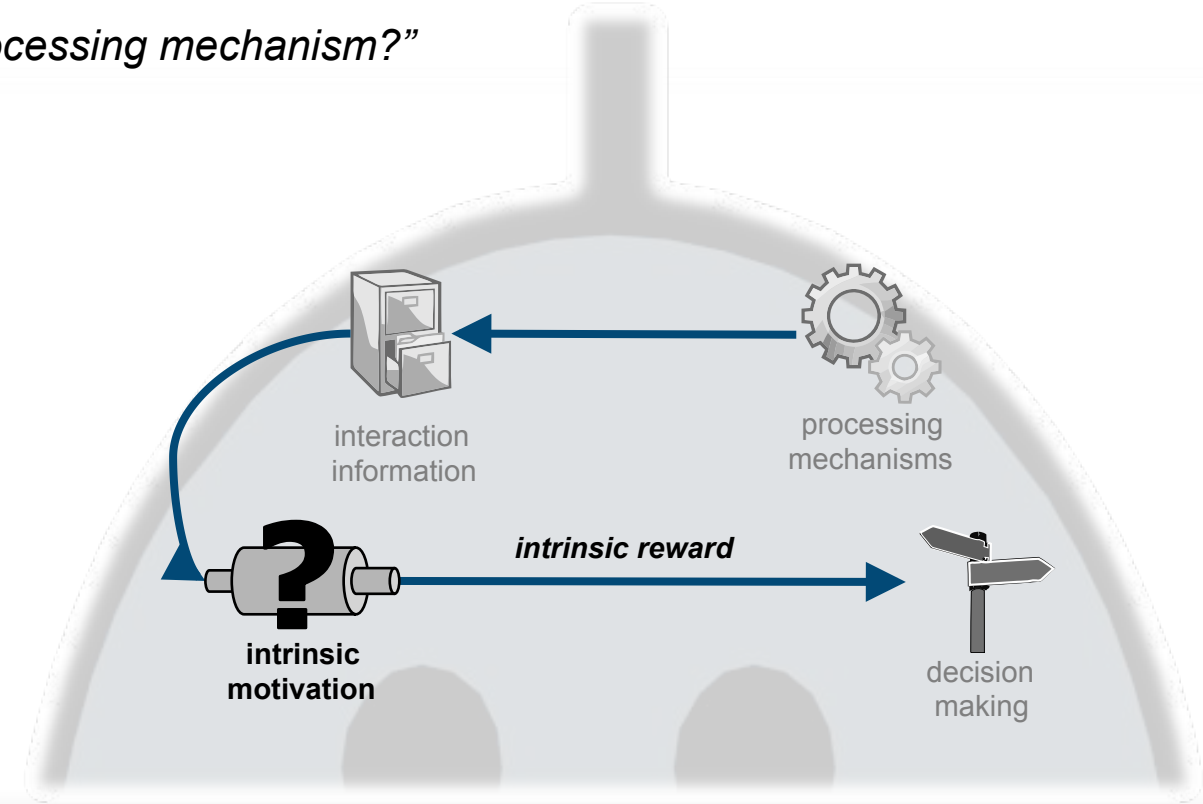
Emerging Emotions

Socially-Aware Learning Agents

Conclusions

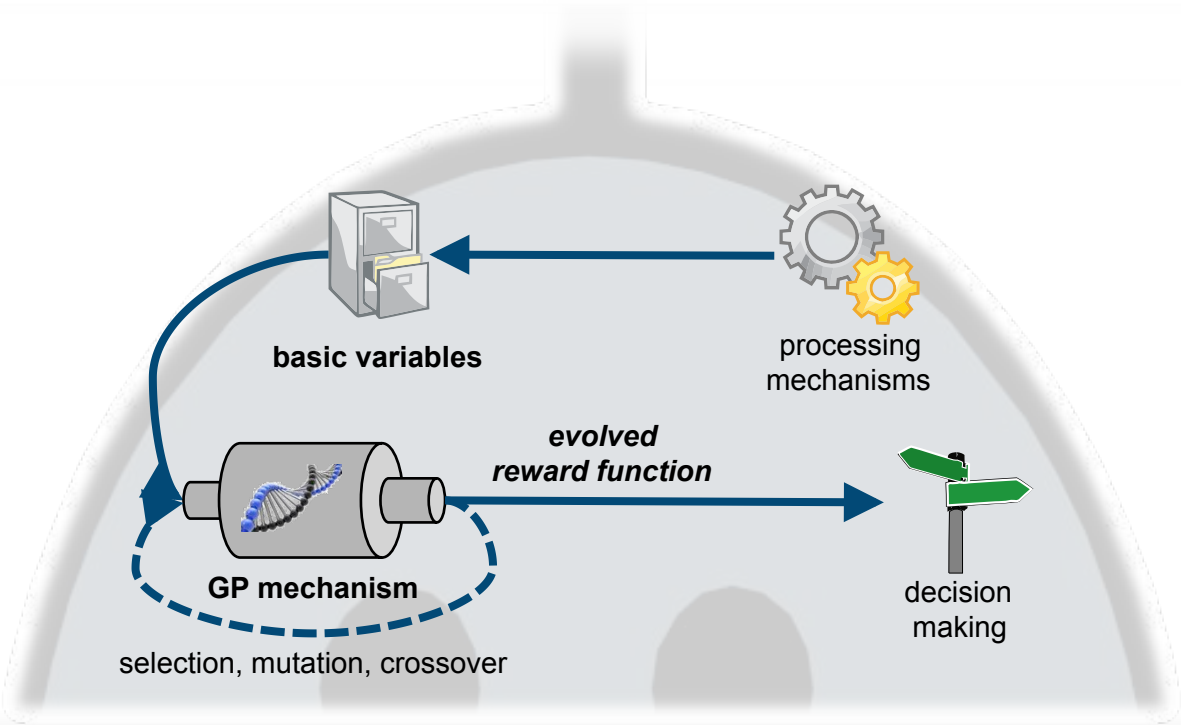
Motivation

- Answer the question:
  - *"Are emotions the best candidate to complement the agents' information processing mechanism?"*



Optimal Sources of Information

- Reward optimization using Genetic Programming (Niekum *et al.* 2010)
  - Population of reward functions
  - Evolved and evaluated according to agent's performance



Foraging Experiments

- Same foraging scenarios as before
  - Observe resulting evolved optimal rewards
  - Discover patterns in the reward functions' expressions
- Results
  - Set of 5 informative signals
  - Fitness, relevance, advantage, prediction, frequency
  - Each signal can be used as an intrinsic reward feature

PacMan Experiments

- Scenarios based on PacMan
  - Objective: validate emerged optimal sources of information
  - Different and much more complex scenarios
  - Limited perception

t outperfo

ed task

eptual lim

Analysis

- Analyze emerged information signals
  - Compare the “kinds” of evaluation
  - According to appraisal theories of emotion
- Results
  - Informative signals provide similar evaluation
  - Share structural and dynamical properties

Contributions

- Information-processing reward mechanism
  - Emerged by means of genetic programming
  - Domain-independent
  - Mitigation of perceptual limitations
  - Relation with natural agents and emotions
    - Dynamic and structural connections with appraisal dimensions
    - Adaptive mechanism that allows for higher fitness
    - Reinforce the role of emotions in agents adaptation

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

Emerging Emotions

Socially-Aware Learning Agents

Conclusions

Motivation

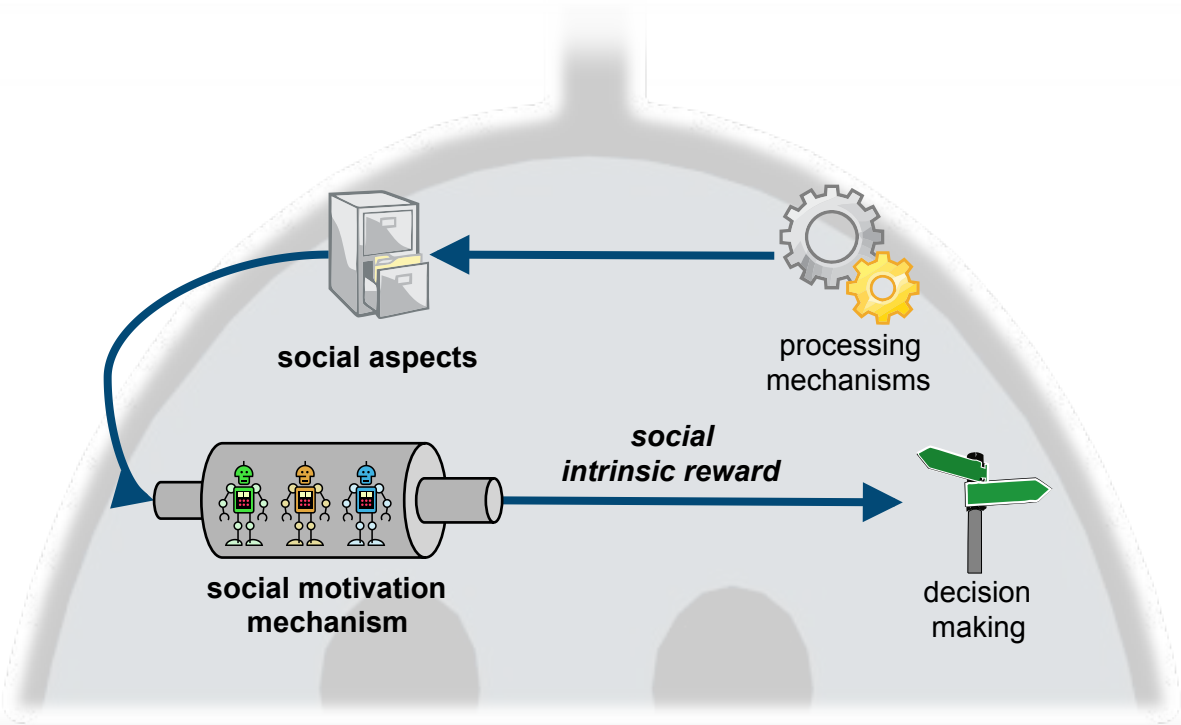
- Extend previous research into multiagent settings
  - Shared environment
  - Each agent has its own goals
    - May be conflicting
  - Achieve cooperation
- Inspiration from social mechanisms
  - Living in group augments survival chances
  - Still there is competition for resources and power
  - Communicate intentions and evaluate each others actions

Cooperation in Nature

- Cooperation in competitive contexts
  (Axelrod 1984, de Waal 2008, Dörner 1999, Falk & Fischbacher 2006, Hamilton 1964, Trivers 1971)
  - Need for *affiliation*
    - Altruistic behaviors despite momentary losses
  - Legitimacy signals
    - Signal socially-aware behaviors
  - Reciprocation mechanism
    - Evaluates “kindness” of others’ actions

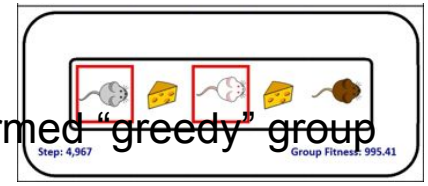
Social Intrinsic Motivation

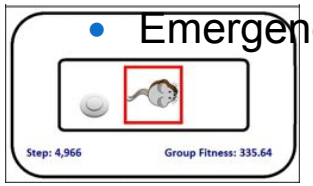
- Limited resource sharing scenarios
  - Internal and external social rewards
  - Evaluate appropriateness of behaviors towards social group

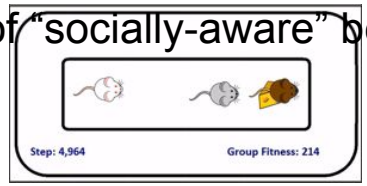


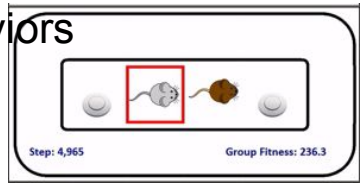
Social Experiments

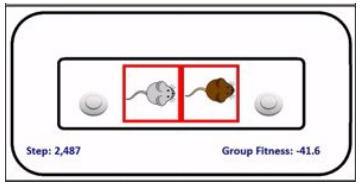
- Limited resource and mutual dependency scenarios
  - 2 or 3 agents interact in the same environment
  - Limited perception
  - Agents learn and act individually, fitness is measured of the social group
- Results
  - Socially motivated agents outperformed "greedy" group
  - Mostly homogeneous populations
  - Emergence of "socially-aware" behaviors











Contributions

- Extend IMRL to multiagent scenarios
  - Social intrinsic motivation
    - Based on affiliation, social signaling and reciprocity
  - Emergence of “socially aware” behaviors
    - Trade-off immediate gains for future collaboration
    - Learned behaviors benefit whole group
  - Accordance with how cooperation thrives in nature
    - Existence of signaling mechanism
    - Reciprocation opportunities
    - Future interactions

Outline

Problem

Approach

Emotion-based Intrinsic Rewards

Emerging Emotions

Socially-Aware Learning Agents

Conclusions

Hypothesis

*"We focus on the **role of emotions** and also on*

*the way **individuals interact and cooperate** with each other as a social group*

*to **design more flexible and robust reward mechanisms** that*

***enhance the autonomy** of RL agents in*

both single and multiagent settings."

Implications (1/2)

- RL and IMRL
  - Alleviate perceptual limitations and modeling effort
  - More autonomous, robust and flexible mechanisms
  - Novel approach for multiagent IMRL
- Affective Computing
  - Show the importance of emotion-based reward design
  - Independent of algorithm, not focused on set of basic emotions
  - Parallel with natural organisms
  - Importance of emotion-related information

Implications (2/2)

- Multiagent Systems
  - Relation with evolutionary game theory
    - Emergence of cooperation with relatedness and reciprocation
  - Signaling mechanism according to internal social standards
  - Emergence of cooperation by means of
    - Social pressures
    - Pure altruism

The End