

Emotions for Artificial Agents

Jan Škoda

Faculty of Mathematics and Physics, Charles University in Prague

22th November 2016



Introduction

- "A New Approach to Modeling Emotions and Their Use on a Decision-Making System for Artificial Agents"
- Simulation of simple autonomous agents in an environment
- Emotions and learning
- Experimental evaluation and comparison of performance with and without emotions

Layout

- Problem introduction
- Motivation
- Modelling agent with artificial emotions
- Testing the agent in simulation
- Conclusion and some philosophy

Motivation

- Necessary for human–robot interaction
 - Needed to mimic human behaviour
- Can improve other (non)cognitive functions

Quote

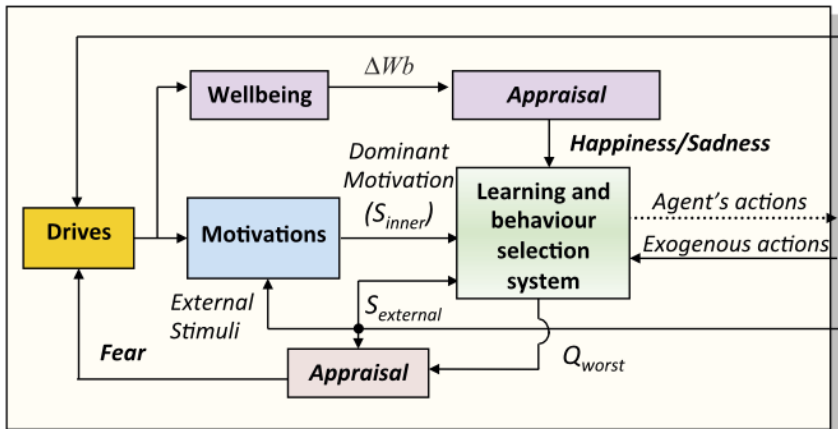
”Emotions are necessary for the survival of the individual and the species. Therefore, all organisms on earth need emotional systems... Thus, a robot designed to survive in the world – would require an equivalent system, one that instills urgency to its actions and decisions.”

A. Kelley

Approach

- Decision-making system based on drives, motivations and emotions
- Emotions: Fear, Sadness, Happiness
- Goal: maximalization of well-being, a function of drives
- Reinforcement learning of actions in different situations
- Sadness and Happiness serve as the reinforcement function
- No inherited/innate knowledge

Agent model



Drives and well-being

- Well-being function:
- $Wb = Wb_{ideal} - \sum_i \alpha_i \cdot D_i$
- α_i weight/importance of the individual drives
- D_i value of the individual drives
- $\Delta Wb > 0 \rightarrow$ Happiness
- $\Delta Wb < 0 \rightarrow$ Sadness

Drives and well-being

- Well-being function:
- $Wb = Wb_{ideal} - \sum_i \alpha_i \cdot D_i$
- α_i weight/importance of the individual drives
- D_i value of the individual drives
- $\Delta Wb > 0 \rightarrow$ Happiness
- $\Delta Wb < 0 \rightarrow$ Sadness

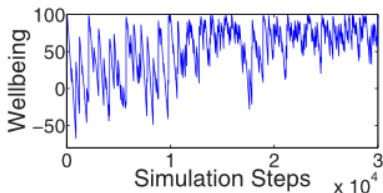
Appraisal of Fear

- What is the worst state, that can happen if the agent will chose actions to maximize utility, but the other agents/change will chose actions to minimize it?
- Fear is a function of utility of that state

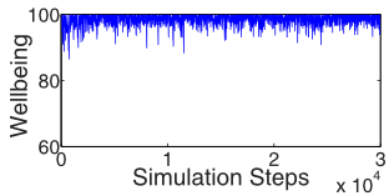
Simulation

- Drives: Thirst, Hunger, Loneliness, Weakness, Fear
- Actions: Eat, Drink, Take Medicine, Get*, GoTo*
- External agent actions: Steal, Give, Greet, Kick, Nothing

Advantage of emotions



(a) Wellbeing as the reinforcement function



(b) Happiness and sadness as the reinforcement function

Conclusion

- Agent with emotions > Agent w/o emotions?
- Emotions ~ States ~ Memory
- Which emotions are "real" ?

Conclusion

- Agent with emotions $>$ Agent w/o emotions?
- Emotions \sim States \sim Memory
- Which emotions are "real" ?

Conclusion

- Agent with emotions $>$ Agent w/o emotions?
- Emotions \sim States \sim Memory
- Which emotions are "real"?

Thank you for your attention.
Questions?