Joy, Distress, Hope, and Fear in Reinforcement Learning

(Extended Abstract)

Elmer Jacobs
Interactive Intelligence, TU Delft
Delft, The Netherlands
elmer.j.jacobs@gmail.com

Joost Broekens
Interactive Intelligence, TU Delft
Delft, The Netherlands
joost.broekens@gmail.com

Catholijn Jonker
Interactive Intelligence, TU Delft
Delft, The Netherlands
C.M.Jonker@tudelft.nl

ABSTRACT

In this paper we present a mapping between joy, distress, hope and fear, and Reinforcement Learning primitives. Joy / distress is a signal that is derived from the RL update signal, while hope/fear is derived from the utility of the current state. Agent-based simulation experiments replicate psychological and behavioral dynamics of emotion including: joy and distress reactions that develop prior to hope and fear; fear extinction; habituation of joy; and, task randomness that increases the intensity of joy and distress. This work distinguishes itself by assessing the dynamics of emotion in an adaptive agent framework - coupling it to the literature on habituation, development, and extinction.

Categories and Subject Descriptors

I.2.6 [**Computing Methodologies/Artificial Intelligence**]: Learning

General Terms

Human Factors

Keywords

Reinforcement Learning, Emotion Dynamics, Affective computing

1. INTRODUCTION

Emotion and reinforcement learning play an important role in shaping behaviour. Emotions are forms of feedback about the value of alternative actions [3, 13] and directly influence action selection, for example through action readiness [7]. Reinforcement Learning (RL) [20] is based on exploration and learning by feedback and relies on a mechanism similar to operant conditioning. The goal for RL is to inform action selection such that it selects actions that optimize expected return. There is neurological support for the idea that animals use RL mechanisms to adapt their behavior [4, 11]. This results in two important similarities between emotion and RL: both influence action selection, and both involve feedback. The link between emotion and RL is supported neurologically by the relation between the orbitofrontal cortex, reward representation, and (subjective) affective value (see [14]).

While most research on computational modeling of emotion is based on cognitive appraisal theory [9], our work is different in that we aim to show a direct mapping between RL primitives and emotions, and assess the validity by replicating psychological findings on emotion dynamics, the latter being an essential difference with [5]. We believe that before affectively labelling a particular RL-based signal, it is essential to investigate if that signal behaves according to what is known in psychology and behavioral science. The extent to which a signal replicates emotion-related dynamics found in humans and animals is a measure for the validity of giving it a particular affective label.

We propose a computational model of joy, distress, hope, and fear instrumented as a mapping between RL primitives and emotion labels. Requirements for this mapping were taken from emotion elicitation literature [12], emotion development[19], and habituation and fear extinction [21, 10]. Using agent-based simulation where an RL-based agent collects rewards in a maze, we show that the emerging emotion dynamics are consistent with this psychological and behavioral literature.

2. MAPPING EMOTIONS

We propose to map RL primitives (e.g., reward, value, update signal) to emotion labels, in particular joy/distress and hope/fear. Such a mapping should honor the fact that emotions develop. In the first months of infancy, children exhibit distress and pleasure [19], followed by joy, sadness and disgust (3 months). This is followed by anger, suprise then fearfulness, usually reported first at 7 or 8 months.

Further, a mapping of RL primitives to emotion should be consistent with habituation and extinction. Habituation is the decrease in intensity of the response to a reinforced stimulus resulting from that stimulus+reinforcer being repeatedly received, while extinction is the decrease in intensity of a response when a previously conditioned stimulus is no longer reinforced [10, 21].

Reward, desirability, unexpectedness and habituation all modulate the intensity of joy [18, 12]. We map joy / distress as follows:

$$J(s_{t-1}, a_{t-1}, s_t) = (r_t + V(s_t) - V(s_{t-1}))(1 - P^{a_{t-1}}_{s_{t-1}s_t}) \quad (1)$$

where J is the joy (or distress, when negative) experienced after the transition from state s_{t-1} to state s_t through action a_{t-1} with V the value and P the probability to end up in

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

state s_t. Joy is calculated before updating $V(s_t)$.

Hope and fear should emerge after joy and distress, should be dependent on the expected joy/distress and likelihood of a future event [12], and should allow fear extinction (e.g, through a mechanism similar to *new learning* [10]). We model the intensity of hope/fear HF as follows:

$$HF(s_t) = V(s_t) \tag{2}$$

3. VALIDATION

We now briefly report on whether the model adheres to several important requirements. For details see [8]. We observed in our agent-based simulation experiments that joy/distress is the first emotion to be observed followed by hope/fear. As mentioned earlier, human emotions have an order in their developent in individuals from simple to complex [19]. We observed joy habituation when the agent was repeatedly presented with the same reinforcement, and fear extinction over time due to a mechanism a mechanism similar to *new learning* [10]. We were unable to confirm if lowered expectation decreases hope and results in a higher intensity for joy/distress [21, 12]. Finally, we were able to confirm that increasing the unexpectedness of results of actions (by modulating task randomness) also increases the intensity of the joy/distress emotion [12, 15].

4. DISCUSSION

We conclude that our model is a plausible RL-based instrumentation for joy/distress and hope/fear. Our results support the idea that the function of emotion is to provide a complex feedback signal for an organism to adapt its behavior. We show this feedback signal can be operationalized for RL agents. This is important for several reasons. First, RL-based models can help understand the relation between emotion and adaptation in animals. The function of emotions is to provide complex feedback signals aimed at informing the agent about the current state of affairs during learning and adaptation [6, 13, 1]. What do such signals look like in an adaptive agent? If we can operationalize such signals for RL agents, a popular computational model for reward-based learning in animals [4, 11], we can computationally tie emotion to adaptation. Second, the emotional state might be used to increase adaptive potential of artificial agents [16, 17]. Third, from a human-robot interaction point of view the emotional signal can be expressed to a human observer. If this signal is grounded in the learning mechanism of the agent [2] it could help interpret the learning process of the agent or robot. However, we are aware of the difficulties of labeling RL-based signals as particular emotions, and we feel that in general a more structured approach is needed to develop scenarios (tasks/learning approach/RL parameters) to test for the plausibility of affective labeling of RL-based signals.

5. REFERENCES

[1] Joost Broekens, Stacy Marsella, and Tibor Bosse. Challenges in computational modeling of affective processes. *IEEE Transactions on Affective Computing*, 4(3), 2013.

[2] L. Canamero. Emotion understanding from the perspective of autonomous robots research. *Neural networks*, 18(4):445–455, 2005.

[3] A. R. Damasio. *Descartes' Error: emotion reason and the human brain*. Penguin Putnam, 1996.

[4] Peter Dayan and Bernard W. Balleine. Reward, motivation, and reinforcement learning. *Neuron*, 36(2):285–298, 2002.

[5] Magy Seif El-Nasr, John Yen, and Thomas R Ioerger. Flame: fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent systems*, 3(3):219–257, 2000.

[6] N. H. Frijda. *Emotions and action*, page 158âĂŞ173. Cambridge University Press, 2004.

[7] N.H. Frijda, P. Kuipers, and E. Ter Schure. Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2):212, 1989.

[8] Elmer Jacobs, Joost Broekens, and Catholijn Jonker. Emergent dynamics of joy, distress, hope and fear in reinforcement learning agents. In *Adaptive Learning Agents workshop at AAMAS2014*, 2014.

[9] Stacy Marsella, Jonathan Gratch, and Paolo Petta. Computational models of emotion. *K. r. Scherer, t. BÃd'nziger and e. roesch (eds.), A blueprint for affective computing*, pages 21–45, 2010.

[10] K. M. Myers and M. Davis. Mechanisms of fear extinction. *Mol Psychiatry*, 12(2):120–150, 2006.

[11] John P. O'Doherty. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology*, 14(6):769–776, 2004.

[12] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[13] Edmund T. Rolls. Precis of the brain and emotion. *Behavioral and Brain Sciences*, 20:177–234, 2000.

[14] Edmund T. Rolls and Fabian Grabenhorst. The orbitofrontal cortex and beyond: From affect to decision-making. *Progress in Neurobiology*, 86(3):216–244, 2008.

[15] K.R. Scherer. Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research*, 92:120, 2001.

[16] N. Schweighofer and K. Doya. Meta-learning in reinforcement learning. *Neural Networks*, 16(1):5–9, 2003.

[17] Pedro Sequeira, FranciscoS Melo, and Ana Paiva. *Emotion-Based Intrinsic Motivation for Reinforcement Learning Agents*, volume 6974 of *Lecture Notes in Computer Science*, chapter 36, pages 326–336. Springer Berlin Heidelberg, 2011.

[18] JC Sprott. Dynamical models of happiness. *Nonlinear Dynamics, Psychology, and Life Sciences*, 9(1):23–36, 2005.

[19] L Alan Sroufe. *Emotional development: The organization of emotional life in the early years*. Cambridge University Press, 1997.

[20] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. Cambridge Univ Press, 1998.

[21] Ruut Veenhoven. Is happiness relative? *Social Indicators Research*, 24(1):1–34, 1991.