Emergence of Emotional Appraisal Signals in Reinforcement Learning Agents

Pedro Sequeira
Northeastern Universtity
202 WVH 360 Huntington Avenue, 02115 MA, USA
psequeira@ccs.neu.edu

***Abstract*—The positive impact of emotions in decision-making has long been established in both natural and artificial agents. In the perspective of appraisal theories, emotions complement perceptual information, coloring our sensations and guiding our decision-making. However, when designing autonomous agents, is *emotional appraisal* the best complement to their perceptions? Mechanisms investigated in the affective sciences provide support for this hypothesis in biological agents. In this paper, we look for similar support in artificial systems. We adopt the intrinsically motivated reinforcement learning framework to investigate different sources of information that can guide decision-making in learning agents, and an evolutionary approach based on genetic programming to identify a small set of such sources that have the largest impact on the performance of the agent in different tasks, as measured by an external evaluation signal. We then show that these sources of information: (i) are applicable in a wider range of environments than those where the agents evolved; (ii) exhibit interesting correspondences to emotional appraisal-like signals previously proposed in the literature. The results of the study thus point towards our departing hypothesis that the appraisal process might indeed provide essential information to complement perceptual capabilities and thus guide decision-making.**

I. Introduction

Emotions indirectly drive behaviors that lead individuals to achieve goals and satisfy needs. Studies show that damage to regions of the brain identified as responsible for emotional processing affects the ability to properly learn aversive stimuli, plan courses of action, *i.e.*, to take advantageous decisions [1]. Appraisal theories of emotions suggest that they arise from evaluations of specific aspects of the individual's relationship with their environment, providing an adaptive response mechanism to situations occurring therein [2].

In artificial systems, the area of affective computing (AC) has investigated the impact of emotional processing capabilities in the development of autonomous agents, often based in appraisal theories of emotions. Appraisal-inspired mechanisms were shown to improve the performance of artificial agents [3], [4]. Within these mechanisms, *appraisal signals* "translate" information about the history of interaction of the agent with its environment that aid decision-making and focus behavior towards dealing with the situation being evaluated [5]. Although one of the driving motivations for the use of emotional appraisal-based agent architectures is the creation of "better agents", one fundamental question remains mostly unaddressed in the literature: *in the search for information that may complement an agent's perceptions, is the emotional appraisal process the best mechanism to provide such information?*

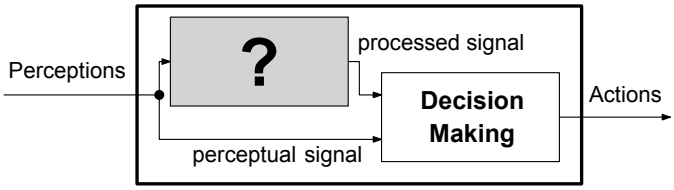

Figure 1. How to complement decision-making?

In this paper we contribute to this question, providing empirical evidence that appraisal-like signals may arise as natural candidates when looking for sources of information to complement an agent's perceptual capabilities. Our approach to the question can be conceptualized as depicted in Fig. 1—we want to analyze which kind of perceptual information processing mechanism best complements an agent's decision-making in a given set of tasks.[1]

II. Identification of Sources

In our study, we rely on *intrinsically motivated reinforcement learning* (IMRL) that provides a principled manner to integrate multiple sources of information in the process of learning and decision-making of artificial agents [7]. These complementary sources of information endow the agent with a richer repertoire of behaviors that may successfully overcome agent limitations, *e.g.*, [4], [7]. As such, IMRL is a framework naturally suited to our investigation.

Fig. 2 depicts the evolutionary procedure based on genetic programming (GP) [8]. We depart from an initial population of agents, each endowed with a certain reward function that is encoded as a genetic program combining information about different aspects of the agent's interaction with its environment, *e.g.*, the number of visits to some state, the value associated with some action or the perceived distance to some estimated goal state. Each agent/reward function is evaluated in a set of environments of interest. For experimental purposes we tested our IMRL agents in grid-world foraging scenarios, where fitness is measured as the total amount of preys captured. The perceptual limitations of the agent in the different environments pose challenges that directly impact

[1] A full version of this manuscript can be found in [6].

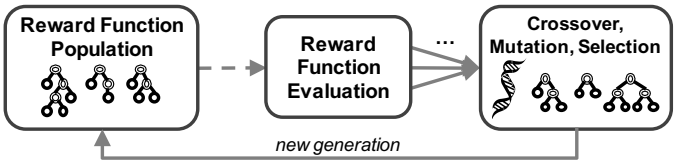

Figure 2. Evolving reward functions with GP.

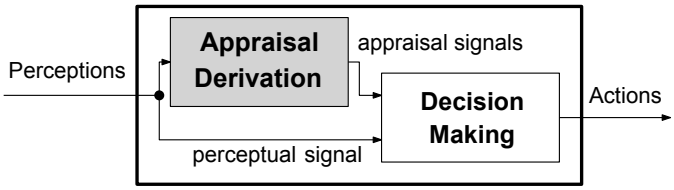

Figure 3. Emotional appraisal-based agent.

its ability to capture preys and, consequently, its fitness. In each generation, the fittest agents are selected and the reward structure of the other agents is modified by means of the GP operations of mutation and crossover. Upon convergence, the set of agents able to attain the largest fitness in the desired scenarios was identified. The analysis of the corresponding reward structures provided the information about which signals are potentially most *useful* to complement the perceptual capabilities of our IMRL agents. In that regard, the results of our experiments revealed the emergence of the following informative signals:

- **Fitness**, corresponding to the estimate of the extrinsic reward received after executing some action in a state;
- **Value**, assessing the estimated long-term extrinsic reward of executing an action after observing some state;
- **Advantage**, evaluating how good it is to execute an action when in some state, relative to the estimated best action;
- **Predictability**, measuring the probability of transitioning to some state after executing some action;
- **Novelty**, providing a (negative) measure of how novel a state is given the agent's past observations.

III. Validation of Emerged Sources

In order to assess the general-purpose of the emerged sources of information, *i.e.*, whether they can guide the decision process of learning agents in general, we tested their applicability in a different set of scenarios. For that purpose we created a set of environments inspired by the game of Pac-Man, which are considerably larger and more complex than the foraging scenarios in which the reward signals emerged. We again resorted to the IMRL framework to define a space of reward functions as a linear combination of the 5 emerged signals. By using a basic uniform-sampling optimization procedure, the results of this second experiment showed that indeed the signals are more informative than the extrinsic rewards while mitigating the agent's inherent perceptual limitations when playing Pac-Man.

IV. Discussion & Conclusions

After showing the applicability and general-purpose of the emerged sources of information, we then compared the kinds of information processed by each reward signal with the nature of evaluations made by humans according to appraisal theories of emotions. We based our analysis in the "major appraisal dimensions" identified in [2], namely *novelty*, *intrinsic pleasantness*, *motivational bases*, *power and coping* and *social dimensions*. Each dimension defines an appraisal "theme", *i.e.*, the criteria and kind of information that is used to perform a certain appraisal, according to particular aspects of the subject-environment relationship. The results of our analysis revealed that the identified sources of information exhibit dynamics and properties that can be related to the way natural agents evaluate their environment according to appraisal theories of emotions.

In conclusion, our experimental study highlighted the usefulness of appraisal-like processes in identifying different aspects of a task—the sources of information—that complement an agent's perceptual capabilities in the pursuit for more reliable artificial decision-makers, as depicted in Fig. 3. Notably, the emerged features result from reward functions that provided optimal performance as measured by a fitness measure that has nothing to do with emotions or appraisal. Moreover, these features, much like appraisal processing in nature, also proved to be useful in different environments. We also contributed for research within IMRL by emerging 4 domain-independent reward signals, relying on common information used by most reinforcement learning (RL) algorithms, that could be applied in different scenarios with distinct purposes. As a consequence, the emerged signals provide a general-purpose reward mechanism for artificial RL agents while reducing the need for designers to hand-code reward functions for specific scenarios.

Acknowledgments

This work was done in collaboration with Francisco S. Melo and Ana Paiva, both with INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, Portugal.

References

[1] A. Bechara, H. Damasio, and A. Damasio, "Emotion, decision making and the orbitofrontal cortex," *Cerebral Cortex*, vol. 10, no. 3, pp. 295–307, 2000.

[2] P. Ellsworth and K. Scherer, "Appraisal processes in emotion," in *Handbook of the Affective Sciences*. Oxford University Press, 2003.

[3] T. Rumbell, J. Barnden, S. Denham, and T. Wennekers, "Emotions in autonomous agents: Comparative analysis of mechanisms and functions," *Autonomous Agents and Multiagent Syst.*, vol. 25, no. 1, pp. 1–45, 2011.

[4] P. Sequeira, F. S. Melo, and A. Paiva, "Learning by appraising: an emotion-based approach to intrinsic reward design," *Adaptive Behavior*, vol. 22, no. 5, pp. 330–349, 2014.

[5] S. Marsella, J. Gratch, and P. Petta, "Computational models of emotion," in *Blueprint for Affective Computing*. Oxford University Press, 2010.

[6] P. Sequeira, F. S. Melo, and A. Paiva, "Emergence of Emotional Appraisal Signals in Reinforcement Learning Agents," *Autonomous Agents and Multi-Agent Systems*, vol. 29, no. 4, pp. 537–568, 2014.

[7] S. Singh, R. Lewis, A. Barto, and J. Sorg, "Intrinsically motivated reinforcement learning: An evolutionary perspective," *IEEE Trans. Autonomous Mental Development*, vol. 2, no. 2, pp. 70–82, 2010.

[8] J. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.