Affect and Learning: a computational analysis

Broekens, D.J.

8

Summary and Conclusion

Here we give a concise overview of the results presented in the different chapters of this thesis, and relate these to each other.

8.1 Affect, Mood and Information Processing

Action selection has been defined as the problem of continuously deciding what next action to select in order to optimize survival (Tyrell, 1993). In a Reinforcement Learning (RL) context, action selection is the process of selecting the next action from a set of actions proposed by a model of interaction with the world, such that the model can both be *learned* by means of interaction and be *used to optimize* received reward. In our case the RL mechanism is a model-based Reinforcement Learning method (see Kaelbling et al., 1996). An agent can select actions in a variety of ways, such as greedy (take the best proposed action) or random (take any action). This is an important issue in robot learning: when should the action-selection mechanism explore versus exploit. We have shown (Chapter 3) that this action-selection trade-off can be partly controlled by artificial affect, when artificial affect is defined as a measure that keeps track of how well the agent is performing compared to what the agent is used to. If the agent performs well, artificial affect is positive and action selection can be greedy, reflecting the relation “good performance—keep doing what you do” (exploit). If the agent performs badly, artificial affect is negative and action selection must be more random, reflecting the relation “bad performance—try new stuff” (explore). By doing so, we have shown that computational modeling can give insights into the possible relations between affect and learning on a meta-level. Artificial affect can be used to control learning parameters.

The type of agents used in the experiments just mentioned is reactive. Agents do not have an abstraction of thought. They behave using an input-output mapping: states go in; actions come out (via the value function that learns value-state-action mappings and the action-selection mechanism that subsequently selects one action from the set of proposed actions). However, thoughts, like actions, have to be selected in some way too. We have shown that thought-selection can also be controlled up to a certain extent by artificial affect (Chapter 4). Thought in this case is interpreted as internal simulation of behavior (Hesslow, 2002). The agent can internally simulate potential interactions with the world. In our experiments, simulation is bounded to one imaginary step ahead, however multiple possibilities exist for that one step (different actions are possible, and different states could result from those actions). The agent has to choose between selecting only several good thoughts (the agent is in a good mood and thinks greedy) or a lot of diverse thoughts including thoughts that are evaluated as bad (bad mood, the agent thinks “explorative”). Again, there is a trade-off between exploration (internally simulate all potential next interactions) and exploitation (internally simulate the option that is perceived as best). Note that in our computational studies, the agent can not really explore mentally, as it can only think of things it has already encountered, explaining the quotes around exploration.

We have shown that internal simulation in this sense is beneficial to the learning performance of an agent, and that artificial affect can be used to control thought selection. Internal simulation of all possible options results, in all experiments, in the best learning performance when compared to no simulation, some simulation, or affectively controlled simulation. However, internal simulation of all options every step is a waste of effort. Sometimes, the learned world model and value function are very good; simulation is actually not needed and the agent can just use a purely reactive mode of operation. In other cases, the learned model is bad; the agent should try to look ahead in a broad sense in order to predict possible consequences of its actions. Artificial affect can control this trade-off. When positive artificial affect is coupled to less, but greedy, internal simulation and negative artificial affect is couple to more, “explorative”, internal simulation, the resulting amount of internal simulation that is needed for a learning performance comparable to one resulting from simulation of all options every step is reduced. To be more precise, coupling artificial affect to internal simulation, in the way just mentioned, enables a learning agent to have about the same learning performance gain compared to an agent that simulates all possible interactions every step, but using considerably less internal simulation. This means that agents that “feel good” can think ahead in the narrow sense freeing mental resources for other things, while agents that “feel bad” should think ahead in a broad sense fully using mental resources to plan ahead. This is compatible with the psychological literature on human mood, as discussed in Chapter 3 and 4.

An interesting issue that has not been discussed yet is that the most beneficial relation between artificial affect and action selection on the one hand, and artificial affect and simulation selection on the other is the same, i.e., positive relates to narrow and negative relates to broad. This is important for the Simulation Hypothesis (Hesslow, 2002). One of the cornerstones of this hypothesis is evolutionary continuity (Hesslow, 2002). It must be possible to move, in the evolutionary process, from agents that act reactively to agents that think and act. Our finding that the direction of the most beneficial relation between artificial affect and thought selection and artificial affect and action selection is the same is an indication that at the level of behavior modulation this continuity exists. However, one has to be very careful with such conclusions, as computational models are complex, large structures containing many choices. We return to this point, made in Chapter 7, in Section 8.3.

In the studies reported upon in Chapter 3 and 4, we have used a definition of affect that relates to the positiveness versus negativeness of *mood*. It is a long-term signal originating from the relative success of the agent. As such, we have used artificial affect as a meta-level signal: artificial affect is used to control learning parameters, not as reward. However, the latter is certainly possible (Chapter 6 and Section 8.2).

8.2 Affect, Emotion and Reinforcement

When affect is related to the positiveness versus negativeness of a situation in a short-term, object/situation-directed sense, it relates more to *emotion* than mood. As such, artificial affect can be used to tell the learning agent something about the current situation, instead of about its general situation. Further, affect can be elicited by external factors, such as communicated emotional expressions, instead of originate from the agent itself. We have taken this approach in Chapter 6, and we have shown that communicated affect can help learn an artificial agent. More precise, a human observer reacts affectively (by means of emotion recognition from facial expressions) to a simulated robot while that robot learns. This affective reaction is translated to a positive or negative reward. The reward is used by the robot in addition to the rewards it gets from interacting with the world. This interaction helps the robot to learn a grid-world task.

The main conclusion to be drawn from this study is that affective interaction facilitates robot learning: we have quantitatively shown that a simulated robot learns quicker with social reward than without social reward. However, for this beneficial effect to last, the robot has to learn an additional social reward function that predicts, based on world-state input, the social reward given by the human. If not, the agent simply forgets the social reward when the human stops giving it. This is not a problem if the task has been learned completely, because now the agent already has an optimal model. However it is a problem if the agent is left over to itself after a short social training period. The latter situation is the more plausible and more desirable one. It is more efficient (the human has to observe the robot less often), and it is better related to parent-child interaction: children are not monitored all the time, but in phases. In a non-monitored phase, the child has to try to find out for itself what to do with the guidance given during a monitored phase.

8.3 Formal Models and Computational Limitations

It is interesting to note that affect defined as the positiveness versus the negativeness of a situation (e.g., Gasper & Clore, 2002) is actually a very useful abstraction in the context of Reinforcement Learning. It can be used in many ways, as has been shown in this thesis. However, we have to be careful, again, about conclusions drawn from computational experiments, specifically related to the meaning of the modeled concepts. It can be argued that the way we model affect is quite limited, which is most certainly the case considering the wide variety of emotions and moods that exist in humans. In relation to Reinforcement Learning this definition (and our derived definition of artificial affect) might be adequate, but this does not mean that we have modeled affect in its full glory, or that we can conclude anything about affect in general. Therefore, our psychology-related claims and conclusions have to be interpreted in the context of Reinforcement Learning and instrumental conditioning. Our conclusions are about existence proofs of relations, for they appear beneficial to artificial agents that learn based on different computational models of instrumental conditioning (the versions of RL used in Chapters 3, 4 and 6). As such, they are relevant to experimental psychology. Experimental psychology has difficulties explaining the mechanisms behind relations. In this context, the mechanisms presented in this research are potential candidates that support relations between affect and learning found in the psychological literature. The conclusions should not be carried further than that.

Concrete computer science related results include the control of learning parameters in artificial learning methods by means of abstractions of concepts borrowed from psychology. More specific, artificial affect has successfully been used to control exploration versus exploitation, and affect has been used as reinforcement in an interactive learning setup with a human in the loop. It is very well possible to use affect in a broader sense than the one studied in this thesis. For example, it is interesting to research how affect can be used to control the search through a solution space, as this is also a process of exploration (random jumps, multiple start positions) versus exploitation (hill-climbing). Further, *arousal*, the part of affect that defines the activity or action readiness of the organism—a part we have ignored completely in this thesis—can be modeled and then used to control other parameters. These parameters could be related to the amount of energy available to the agent. Such parameters include the likelihood of acting in the first place and the depth of the thought process.

As mentioned in the previous paragraphs, computational models are limited in their ability to conclude about natural phenomena. This issue has been dealt with related to emotion modeling in Chapter 7. We have shown that it is useful, in fact critical, to use formal models of emotion at an architectural level to advance emotion theory. The analysis has been focused on cognitive appraisal theory, explaining emotions as a result of the subjective evaluation of events in the context of beliefs, desires, and intentions of an agent (being natural or artificial). Our analysis showed that with the formal notation we developed it becomes easier to evaluate whether unexpected behavior resulting from a computational model is due to errors in the computational model or errors in the theory. This is an important issue, as computational models of emotion tend to get very complex and are inspired by many psychological theories (see, e.g., the impressive agent models by Gratch and Marsella, 2004 or Baars and Franklin, 2003). We have further shown that the formal notation can be used to integrate different cognitive appraisal theories, an important issue in the advancement of appraisal theory (Wehrle & Scherer, 2001).

A very valid argument that could be put forward at this point is that we haven't formally described the affect-learning relations studied in Chapter 3 to 6, and as such can not really draw strong conclusions from these studies. We can say two things about this.

First, we did not formally represent the relations studied, and it would be interesting to find out if this is possible using the formalism developed in Chapter 7. However, as argued in Chapter 7 and (Broekens & DeGroot, 2006), emotion psychologist have to also formally annotate the data resulting from, and proposed mechanisms derived from emotion studies. Without this annotation, the computer model can not be evaluated other than in ways done in this thesis or in the work by many other modelers. So, formal modeling by computer scientists is only half the solution, and in this case, half a solution is no solution as there is nothing formal to compare the computer scientist's formal model with. More importantly, the formalism proposed in Chapter 7 is targeted towards cognitive appraisal theory, which is not used as underlying theory for the research in Chapter 3 to 6. We have taken this direction because the number of computational models of emotion based on cognitive appraisal theory is vast, and consequently a formalism targeted at this family of models and theories could have a larger impact.

Second, we can definitely draw conclusions related to psychology from our studies, given that we extensively argued why we modeled affect in the ways we did, as well as how we used it to influence learning. Further, our conclusions should be interpreted as *mechanism existence proofs* than can inspire psychological research, just as psychological research has inspired the modeling work in this thesis. Research can be done in many ways; sometimes the conclusions are clear-cut logical results, sometimes they are hypotheses made plausible. Our conclusions regarding computational results, such as better learning performance, fall into the first category: whatever the underlying mechanisms are or are not based upon, the result is objectively measurable. Conclusions related to the psychological implications of the studies presented in this thesis fall into the second category: given the computational results, the relations and mechanisms we have modeled become more plausible psychologically, although never an exclusive truth.