Affect and Learning: a computational analysis

Broekens, D.J.

2

Artificial Affect

In the Context of Reinforcement Learning

In this chapter we present the rationale for the concept of emotion used in the studies reported upon in Chapter 3 and 4, that is, positive and negative affect. We first review different findings on the interplay between emotion and cognition, after which we describe several ways in which affect influences learning, the main phenomenon investigated computationally in this thesis. Finally we introduce a measure for artificial affect, and argue for its validity in the context of Reinforcement Learning.

2.1 Emotion and Behavior Regulation

Emotion influences thought and behavior. For example, at the neurological level, malfunction of certain brain areas not only destroys or diminishes the capacity to have (or express) certain emotions, but also has a similar effect on the capacity to make sound decisions (Damasio, 1994) as well as on the capacity to learn new behavior (Berridge, 2003). These findings indicate that these brain areas are linked to emotions as well as "classical" cognitive and instrumental learning phenomena.

Emotion is related to the regulation of adaptive behavior and to information processing. Emotions can be defined as states elicited by rewards and punishments (Rolls, 2000). Behavioral evidence suggests that the ability to have sensations of pleasure and pain is strongly connected to basic mechanisms of learning and decision-making (Berridge, 2003; Cohen & Blum, 2002). These studies directly relate emotion to Reinforcement Learning. Behavioral neuroscience teaches us that positive emotions reinforce behavior while negative emotions extinguish behavior, so at this level of information processing one type of emotional regulation of behavior has already been established, i.e., approach (rewarded behavior) versus avoidance (punished behavior).

At the level of cognition, emotion plays a role in the regulation of the amount of information processing. For instance, Scherer (2001) argues that emotion is related to the continuous checking of the environment for important stimuli. More resources are allocated to further evaluate the implications of an event, only if the stimulus appears important enough. Furthermore, in the work of Forgas (2000) the relation between emotion and information processing strategy is made explicit: the influence of mood on thinking depends on the information processing strategy used.

Emotion also regulates behavior of others. Obvious in human development, expression (and subsequent recognition) of emotion is important to communicate (dis)approval of the actions of others. This is typically important in parent-child relations. Parents use emotional expression to guide behavior of infants. Emotional interaction is essential for learning. Striking examples are children with an autistic spectrum disorder, typically characterized by a restricted repertoire of behaviors and interests, as well as social and communicative impairments such as difficulty in joint attention, difficulty recognizing and expressing emotion, and lacking of a social smile (for review see Charman & Baird, 2002). Apparently, children suffering from this disorder have both a difficulty in building up a large set of complex behaviors *and* a difficulty understanding emotional expressions and giving the correct social responses to these. This disorder provides a clear example of the interplay between learning behaviors and the ability to process emotional cues

To summarize, emotion can be produced by low-level mechanisms of reward and punishment, and can influence information processing. As affect is a useful abstraction of emotion (see Section 1.4), these aspects inspired us to study (1) how artificial affect can result from an artificial adaptive agent's reinforcement signal, and (2) subsequently influence information processing in a way compatible with the psychological literature on affect and learning. In the next section we present some of the psychological findings related to the latter. In Section 2.3 we introduce the measure of artificial affect we have used in the studies reported upon in Chapter 3 and 4.

2.2 Learning is Influenced by Positive and Negative Affect

The influence of affect on learning is typically studied with the following psychological experiment. Take two groups, one control group and one experimental condition group. Induce affect (positive or negative) into the subjects belonging to the experimental condition group by showing them unanticipated pleasant images or giving them small unanticipated rewards, or violent, ugly images and punishment if negative affect is to be induced in the subject. Measure the subjects' affect. Let the two groups do a cognitive task. Finally, compare the performance results between both groups. If the experimental condition group performs better, affect induction (positive or negative change in a subject's affect due to, e.g., presented images) is assumed to be responsible for this effect, ergo; affect influences the execution of the cognitive task.

Some studies find that non-positive affect enhances learning. For instance, Rose, Futterweit and Jankowski (1999) found that when babies aged 7 - 9 months were measured on an attention and learning task, neutral affect correlated with faster learning. Attention mediated this influence. Neutral affect related to more diverse attention, i.e., the babies' attention was "exploratory", and both neutral affect and diverse attention related to faster learning. Positive affect resulted in the opposite of neutral affect (i.e., slower learning and "less exploratory" attention). This relation suggests that positive affect relates to exploitation and neutral affect relates to exploration. Additionally, Hecker and von Meiser (2005) suggest that attention is more evenly spread when in a negative mood. This could indicate that negative affect is related to exploration.

Interestingly, other studies suggest an inverse relation. For instance, Dreisbach and Goschke (2004) found that mild increases in positive affect related to more flexible behavior but also to more distractible behavior. The authors used an attention task, in which human subjects had to switch between two different "button press" tasks. In such tasks a subject has to repeatedly press a button A or a button B based on some criteria in a complex stimulus. After some trials, the task is switched, by changing several stimulus characteristics. The authors measured the average reaction time of the subjects' button-press just before and just after the task switch. The authors found that increased positive but not neutral or increased negative affect relates to decreased task switch cost, as measured by the difference between pre-switch reaction time and post-switch reaction time. So, it seems that in this study positive affect facilitated a form of exploration, as it helped to remove the bias towards solving the old task when the new task had to be solved instead.

Combined, these results suggest that both positive and negative affective states can help learning but perhaps at different phases during the process, a point explicitly made by Craig et al. (2004). Chapter 3 and 4 of this thesis address exactly this issue. We use simulated adaptive agents to study the influence of artificial affect on learning performance, by controlling several learning parameters. In Chapter 3 artificial affect controls the amount of exploration versus exploitation used by the agent: affect controls the greediness of the action-selection mechanisms. In Chapter 4 artificial affect controls the greediness of its thoughts. In the latter study, an agent can internally simulate a number of interactions before actually executing these. Internal simulation can increase or decrease the likelihood of choosing a particular action, as it biases the value of the next actions (much like a person who imagines the potential results of a certain action, and who decides not to do it because of the imagined consequences). Some of these anticipated possibilities seem good, others do not. Affect is used to control the extent to which the selection of these simulated interactions is biased towards simulating only the positive ones (narrow, greedy "optimistic" thoughts) or towards simulating all anticipated possibilities (broad, evenly distributed thoughts).

2.3 Artificial Affect

To model the influence of affect on learning, we first need to model affect in a psychologically plausible way. Our agent learns based on Reinforcement Learning, so at every step it receives some reward *r*. Here we explain how our agent's artificial affect is linked to this reward *r*.

Two issues regarding affect induction are particularly important. First, in studies that measure the influence of affect on cognition, affect relates more to long-term mood than to short-term emotion. Affect is usually induced before or during the experiment aiming at a continued, moderate effect instead of short-lived intense emotion-like effect (Dreisbach & Goschke, 2004; Forgas, 2000; Rose et al., 1998). Second, the method of affect induction (explained earlier) is compatible with the method used for the administration of reward in Reinforcement Learning. Affect is usually induced by giving subjects small *unanticipated* rewards (Ashby et al., 1999; Custers & Aarts, 2005). The fact that these rewards are unanticipated is important, as the reinforcement signal in RL only exists if there is a difference between predicted and received reward. Predicted rewards thus have the same effect as no reward. It seems that reward and affect follow the same rule: *if it's predicted it isn't important*.

The formula we use for artificial affect is:

$$e_p = (r_{star} - (r_{ltar} - f\sigma_{ltar}))/2f\sigma_{ltar} \tag{2.1}$$

Here, e_p is the measure for affect. If e_p=0, we assume this means negative[1] affect, if e_p=1 we assume this means positive affect. The short-term running-average reinforcement signal, r_{star}, with *star* defining the window size in steps, is the quicker-changing average based on the agent's reward, *r*, as unit of measurement at every step. The long-term running-average reinforcement signal, r_{ltar}, with *ltar*

[1] Low and high values of e_p should not be interpreted as depressed and elated respectively. We assume that we model moderate levels of positive and negative affect, as induced by typical psychological affect-induction studies. Clinical depression and elatedness have different influences on behavior that are out of scope, and are too complex for our current modeling approach.

again defining the window size in steps, is the slower-changing average taking r_{star} as unit of measurement every step. The standard deviation of r_{star} over that same long-term period *ltar* is denoted by σ_{ltar}, and *f* is a multiplication factor defining the sensibility of the measure.

Obviously, artificial affect behaves differently for different values of *f*, *ltar* and *star*. In general, for r_{ltar} to be a good estimate of what the agent is "used to", *ltar* must be considerably larger than *star*. In the studies presented we have varied *ltar*, *star* and *f* across a wide range of values.

Our measure for artificial affect reflects the two issues mentioned above. First, r_{star} uses reinforcement signal averages, reflecting the continued effect of affect induction related to mood not emotion. Second, our measure compares the first average r_{star} with the second longer-term average r_{ltar}. As the first, short-term average, reacts quicker to changes in the reward signal than the second, long-term average, a comparison between the two yields a measure for how well the agent is doing compared to what it is used to (cf. Schweighofer & Doya, 2003). If the environment and the agent's behavior in that environment do not change, e_p converges to a neutral value of 0.5. This reflects the fact that anticipated rewards do not influence affect.

By defining artificial affect purely in terms of rewards and punishments, one could argue that we interpret affect in a too narrow sense, thereby hollowing out the concept. We do not agree. Our meaning of artificial affect is still the same as the meaning of affect: it defines the goodness/badness of a situation for the agent. Further, it is quite compatible with certain theories of emotion (e.g., Rolls, 2000) that emphasize that emotion is fundamentally grounded in (the deprivation/expectancy of) reward and punishment. Finally, as rewards and punishments define what behavior an artificial agent should pursue and avoid, reinforcement *is* the definition of good and bad for such agents. We therefore believe our measure for artificial affect is firmly grounded.