Computational approaches to affective processes: evolutionary and neural perspectives

den Dulk, P.

Citation for published version (APA):

den Dulk, P. (2002). Computational approaches to affective processes: evolutionary and neural perspectives
Chapter 1

Introduction

Computational modeling of emotion and affective processes has been relatively neglected in the cognitive sciences. This could be due to the view that biology, in which affective processes seem to be embedded, is largely irrelevant to human information processing. Another reason for this neglect might be that from a subjective point of view the deep experience of emotions and consciousness does not seem accessible to ‘reductionist’ methods like computer simulation (Nagel, 1974). On the other hand the nature of affective processes – often considered lower cognitive processes and more interwoven with bodily reactions (Damasio, 1994) – is also often seen to make them less suitable for the symbolic nature of many modeling languages.

In this thesis a number of affective processes will be studied by means of computational modeling. The extreme diversity and richness of affective phenomena, ranging from social and cultural, via cognitive, to physiological and neural processes, makes any model vulnerable to the criticism that a certain perspective has been left out. A bottom-up approach was chosen here, starting from simple, biologically inspired, and less conceptually laden elements, in an attempt to see how far this approach leads in modeling higher level processes. Such biologically inspired paradigms might be better suited for modeling affective processes, because they stay closer to the physiology of the brain, and do not revert to symbolic abstractions which might not be appropriate for some aspects of affective processes. Several techniques will be employed in this thesis. Models of mental processes will be built with artificial neural networks, which are abstractions of biological neural networks (Rumelhart, Hinton, & Williams, 1986; Rumelhart & McClelland, 1986). Evolutionary processes will be simulated computationally with the aid of genetic algorithms (Holland, 1975; Michalewicz, 1999). In the final simulations a combination of genetic algorithms and neural networks techniques will be presented.

All projects in this thesis are computer simulations, and mostly constitute models of data gathered in laboratory experiments (on humans or animals). Perhaps the most common view of how modeling should proceed is that first experimental data need to be collected, and subsequently a model is constructed which can suggest an explanation of how these data came about. There is no general agreement, however, on the precise function of computer models (Webb, 2001), and not all studies in this thesis are typical examples of this view. Some of the studies may, in a strict sense, not be called models at all, because they do not have the aim to simulate experimental data. Some studies may function as a proof of principle (Chapter 2),
whereas others explore different solutions to a specific task (as in Chapter 6). The function of modeling can differ from project to project, and in this thesis even differs from chapter to chapter. It, therefore, seems appropriate to first elaborate on the function of modeling, and how it is treated in this project.

Neural network paradigms, as used in Chapters 3 to 6, have often been criticized for its focus on implementation instead of abstraction (e.g., Broadbent, 1985). Many would agree that it should be a goal of science to develop good abstract descriptions, which integrate a large number of empirical phenomena. Such descriptions need to include aspects that are relevant, and ignore aspects that are not. Models are, by definition, imperfect and always neglect some phenomena that are not deemed central to the issue at hand. What is relevant and what not, however, is not a priori clear. Often it only becomes clear during the scientific endeavor which abstractions are most helpful in understanding the problem. Another complication is that what should be considered a good abstraction depends on the question the researcher is trying to answer. Different researchers may have different questions, and thus need different abstractions.

To illustrate this with an example, suppose the inner workings of a car are studied, without any prior knowledge of the underlying mechanics or engines in general. Researchers can have different objectives in such a project. One research objective could be to come up with a complete understanding of how to control a car, i.e., learn how to drive. This enables us to use it as a means for transport. Another research objective could be to gather knowledge on how to repair the car in case of a breakdown. The benefit is obvious. A third objective could be to gather understanding about the general mechanism that enables transport of a vehicle without the need for human or animal strength. Such knowledge could enable us to develop similar devices. These three different perspectives will consider different aspects relevant or irrelevant. Those who are interested in repair of the car should study the engine in great detail. For those who would want to learn how to drive the car, however, everything that is happening under the hood of the car is not important, and should be considered irrelevant data. What they need to know is how to handle the gears, pedal and break. This is less important for the third objective, understanding the engine itself. Here it would be a research breakthrough, if one could come up with a basic fuel driven vehicle that simply moves forward at constant speed. Even if it is completely lacking the gears, break and pedal, which are so crucial for a potential driver. Thus, deciding what is relevant or not depends on the objectives of a research project. There is no such thing as 'the abstraction' with respect to the object of research, only with respect to the research question.

In developing the models presented in this thesis choices have to be made with respect to what is included and what is not. Those choices could best be understood on the basis of Dennett’s (1994) concept of reverse engineering. This is “the interpretation of an already existing artifact by an analysis of the design considerations that must have governed its creation” (Dennett, 1994). The term reverse engineering refers to a strategy in the industry, where one company buys the product of another company to find out how it works, with the purpose to produce a similar device. The engineers will test the device, will benchmark it, will take it apart and try to understand the function of every component. The emphasis in reverse engineering lies on the functional principles that govern the workings of the machine and not on the hardware performing these functions. The aim is to discover the abstractions that were made by the designers. Dennett (1994) compares this strategy to the functionalist approach in cognitive science, where the analysis of behavior
Introduction

proceeds by decomposing it into all functional requirements (Marr, 1982; Newell, Yost, Laird, Rosenbloom, & Altmann, 1990). In Dennett’s view what makes a neuroscientist a cognitive scientist is the acceptance of this project of reverse engineering.

Dennett (1994), however, warns against what he calls the top-down approach to reverse engineering. Assuming that the necessary functions can be specified on the assumption of optimal execution, could lead to over-idealization of the design problem. The evolutionary process does not design systems like engineers do with the goal of the project in mind from the start. Also evolution does not seem to start with similar abstractions as designers do. Although most researchers know perfectly well that evolution does not function this way, they nevertheless assume the method is applicable. Biological solutions are, however, fundamentally different from artificial ones. When human engineers design something (forward engineering), they must guard against a notorious problem: unforeseen side effects. The only practical way to guard against unforeseen side effects is to design the subsystems to have relatively impenetrable boundaries that coincide with the boundaries of the function the creator had in mind. The designer attempts to insulate the subsystems from each other, and insists on an overall design in which each subsystem has a single, well-defined function within the whole. Evolutionary ‘design’ does not have such a protection against unwanted side effects and the system will often not fix it, but compensate for it, or in some cases even utilize it (Gould, 1991). Biological systems are, thus, likely to be composed of multi-function, multi-effect elements, quite incompatible with an engineer’s ideas of design. As an alternative, Dennett proposes what he calls bottom-up reverse engineering, which starts with the specification of small elements, and from there tries to move towards a description of the behavior of the larger ensembles. As examples of this approach Dennett (1994) mentions artificial life and neural networks, both approaches are part of this thesis.

It should be stressed that the ‘bottom’ in the bottom up approach does not refer to the lowest possible level of physiology or physics. Rather, the bottom level should be seen as a lower conceptual level where the simple elements have no strong prior conceptual interpretation. Understanding of the system needs not be restricted to this low level of description. Those elements can be combined to form more complex assemblies, which may again be understood at a more abstract conceptual level. Why not start out at such a higher conceptual level right away? The problem is that we do not know in advance what kind of abstract concepts are the most appropriate for the system under investigation. By starting out with lower level elements that do not have a strong a-priori conceptual interpretation it is possible to come up with other, perhaps new, higher level abstractions that would have been missed otherwise.

On the other hand, the simplified elements which form the basis of much of the research in this thesis (neurons or nodes), are also abstractions of a lower level physiology, and are not completely free from interpretation. Many details are not included because they are assumed to be irrelevant. With respect to this issue Webb (2001) has made a similar point as the one discussed above ‘The existence of an attractive formalism might end up imposing its structure on the problem so that alternative, possibly better, interpretations are missed’ (Webb, 2001). For example, in this thesis the complex chemical processes which are involved in the firing of a neuron are simplified to a single number indicating a node’s activation. Whether this representation is sufficient is not clear, and we need to look at lower levels of neural network research and modeling to find out if this captures all important aspects (Segev, 1992). In our view, different levels of research can coexist next to each other,
without any one level being the proper level of research (Bakker & Den Dulk, 1999), and each level can learn from lower and higher levels to improve its own.

The danger of abstracting away critical aspects at a lower level seems particularly relevant to the area of affective processes. Although symbolic representations are probably also involved in emotional processes (Damasio, 1994; Frijda, 1986; Lazarus, 1991), at least in part many 'sub-symbolic' processes, such as biological (Oatley & Johnson-Laird, 1987; 1995), neural (Panksepp, 1982) and physiological (Cacioppo & Berntson, 1994; Lang, 1995) phenomena, are involved as well. Therefore, it seems more appropriate to start building a model for affective processes with components at a lower level of representation. If at a later stage symbolic representations would become necessary they could be constructed from lower level components. For some of the studies presented neural networks seem appropriate for more obvious reasons. In the study on conditioning of Chapter 4, the original data applied to neural pathways in the brain. Rather a direct link can be made between the neural pathways studied in rats, and the bundles of neural connections simulated in these models. Neural network models of psychological processes, like affective priming in Chapter 5, possibly also offer the opportunity to make more direct connections with the accumulating number of neuro-imaging studies in this area.

Although the different Chapters of this thesis are related they should not be considered an integral study of a single subject. All chapters are about affective processes, and all chapters discuss biologically inspired computational models. Different affective processes are studied, however, with different computational models. The thesis is divided into two main parts: Part I: individual differences (Chapter 2 and 3), and Part II, dual route models of affective processes (Chapter 4, 5 and 6). The contents of the five subsequent Chapters are summarized below.

Chapter 2 is a study of altruistic behavior in groups. The study will not simulate the way altruistic behavior itself functions within individuals, but demonstrates under which circumstances altruistic tendencies may develop. It also demonstrates that different groups emerge within the population, which differ in both altruistic and migrating behavior. Individual differences thus arise as adaptive solutions in this specific environment.

Chapter 3 consists of two parts. The first part discusses a neural network model of individual differences with respect to anxiety applied to the Stroop phenomenon. The second part discusses a model on the differences between high and low anxious groups in performance on an experiment on attentional deployment. From an evolutionary standpoint it can also be argued that it is adaptive to have groups of different anxiety levels coexist in a single population.

Chapter 4 introduces the neurobiological model central to this thesis, the dual route model of LeDoux (1996). In this chapter an artificial neural network is constructed of the LeDoux model within the same connectionist framework (i.e., of competitive networks) as in Chapter 3. It constitutes a replication of a model by Armony et al. (1995) implemented in a different paradigm and extended with a number of new simulations.

In Chapter 5 a dual route model is again presented, but now applied to simulate experimental results in the area of affective priming (Murphy & Zajonc, 1993). The stronger suboptimal- than optimal priming pattern which is characteristic for affective processing is reproduced, and the conditions under which it can be obtained are investigated computationally.
In Chapter 6 the evolutionary origin of the dual route architecture is investigated in an artificial life type of simulation. Simulated creatures have to survive in an environment requiring both fast reactions and accuracy. By means of a genetic algorithm the dual route organization is optimized. Conclusions are drawn with respect to the adaptive value of the neural architecture as has been postulated by LeDoux (1996). The evolutionary justification by LeDoux of the dual-route model is supported by this computational study.