

# Affective Revision

César F. Pimentel and Maria R. Cravo

Instituto Superior Técnico, Dep. Eng. Informática, GIA,  
Av. Rovisco Pais, 1049-001 Lisboa — Portugal  
`cesar.pimentel@dei.ist.utl.pt`, `mrcravo@gia.ist.utl.pt`

**Abstract.** Moods and emotions influence human reasoning, most of the time in a positive way. One aspect of reasoning is the revision of beliefs, i.e., how to change a set of beliefs in order to incorporate new information that conflicts with the existing beliefs. We incorporate two influences of affective states on belief maintenance identified by psychologists, in a AI belief revision operation. On one hand, we present an alternative operation to conventional Belief Revision, Affective Revision, that determines the preference between new and old information based on the mood of the agent revising its beliefs. On the other, we show how beliefs can be automatically ordered, in terms of resistance to change, based on (among other aspects) the influence of emotion anticipations on the strength of beliefs.

## 1 Introduction

The influence of emotions in reasoning has been recognized by philosophers since Aristotle. Generally, this influence was believed to be negative, i.e., emotions were seen as an impairment to rational thought.

Given this state of affairs, it is no wonder that AI researchers did not think of emotions as a useful component of AI systems. Exceptions to this are the views of Simon and Minsky: Simon [17] says that “a general theory of reasoning and problem solving must incorporate the influences of motivation and emotion”; and Minsky [11] says that “the question is not whether intelligent machines can have emotions, but whether machines can be intelligent without emotions”.

In the last years, work done in neuroscience and psychology has shown that, more often than not, emotions are essential to rational reasoning, decision making and social interactions.

As a consequence, AI researchers started to show a growing interest in the possible benefits that the study of emotions can bring to AI systems. This led to the birth of a new area known as *Affective Computing*, defined by Rosalind Picard as the study of “computing that relates to, arises from or deliberately influences emotions” [15].

In this paper we focus on the benefits of considering the influences of affective states in a particular aspect of AI reasoning, belief revision. The aim of belief revision theories is to define how a rational agent should change its set of beliefs, in order to incorporate a new belief, which is inconsistent with this

set. Psychologists have by now identified several ways in which affective states in general, i.e., moods and emotions, do influence the maintenance of beliefs in people. In this paper we model the incorporation of some of these influences in a belief revision theory.

This paper is organized as follows. In the next section we describe the main aspects of the work we built upon: belief revision theories developed in AI, and influences of affective states on the maintenance of beliefs (as identified by psychologists). In this section, we also briefly mention a system of emotion generation that will be needed for the present work. In Section 3, we show how two such influences can be modelled in AI. In Section 4, we present an illustrative example. Finally, in Section 5 we present some conclusions and point directions in which this work will be continued.

## 2 Background

### 2.1 Belief Revision and Permissive Belief Revision

An essential aspect of commonsense reasoning is the ability to revise one's beliefs, i.e., change one's beliefs when a new belief is acquired, that is not consistent with the existing beliefs.

In AI, *belief revision* theories decide which of the previous belief(s) should be abandoned in order to incorporate the new belief, and keep the set of beliefs consistent. All belief revision theories try to keep the change as small as possible, according to what is called the *minimal change principle* [10]. The reason for this principle is that beliefs are valuable, and we do not easily give them up. However, this principle is not enough to determine, in a unique way, the change to be made, and so belief revision theories assume the existence of an order among beliefs, which states that some beliefs are less valuable than others, and should more easily be abandoned.

A number of belief revision theories have been developed since the seminal work of Alchourrón, Gärdenfors and Makinson [1]. These theories assume that beliefs are represented by formulas of the language  $\mathcal{L}$  of some logic, and represent the revision of a set of beliefs  $\beta$  with a belief  $F$  by  $(\beta * F)$ . This represents the new set of beliefs and must be such that: 1) It contains  $F$ ; 2) It is consistent, unless of course,  $F$  is a contradiction. To ensure that the result is a unique set of beliefs these theories either assume the existence of a total order among beliefs, or allow for a partial order, and abandon all beliefs whose relative value is not known, thus abandoning more beliefs than necessary.

In this paper, we consider a particular belief revision theory [2], based on a non-monotonic extension of First Order Logic. For the purpose at hand, it is enough to say that this theory allows for any number of partial orders among beliefs, and that its result is a set of sets of beliefs, instead of a unique set of beliefs. For instance, suppose that the initial set of beliefs  $\beta$  contains the formulas  $A$  and  $A \rightarrow B$ , and that nothing is known about the relative value of these formulas. Then  $(\beta * \neg B) = \{\{A, \neg B\}, \{A \rightarrow B, \neg B\}\}$ . Note that if a single result is required, then we can not do better than  $\{\neg B\}$ , unless an arbitrary