Summary

Recent events have probably lead us to wonder why people make decisions that seem to be irrational, and that go against any easily understandable logic. The fact that these decisions are emotionally driven often explains what, at first glance, does not have a plausible explanation. Evidence has been found that proves that emotions and other affective characteristics guide decisions beyond a purely rational deliberation. Understanding the way emotions take place, the way emotions change, and/or the way emotions influence behavior, has traditionally been a concern of several fields including psychology and neurology. Moreover, other sciences such as behavioral economics, artificial intelligence, and in general, all sciences that aim to understand, explain, or simulate human behavior, acknowledge the important role of affective characteristics in this task.

Specifically, artificial intelligence uses psychological findings in order to create agents that simulate human behavior. Nevertheless, individual research efforts in modeling affective characteristics are often overlapped, short of integration, and they lack of a common conceptual system. This deprives individual researches of the exchange and cooperation's inherent benefits, and makes the task of computationally simulating affective characteristics more difficult. Although much individual effort has been put in classifying, formalizing and modeling emotions and emotion theories on some fields, recognized researchers of emotions' and affective processes' modeling report that a common formal language, an informal conceptual system, and a general purpose affective agent architecture will greatly improve the interdisciplinary exchange and the intradisciplinary coordination.

The research literature proposes a wide amount of affective models that deal with some of: relationship between emotions and cognition, relationship between emotions and behavior, emotions and their evolutionary account, emotions for appraising situations, emotion regulation, etc. These models are useful tools for addressing particular emotion-related issues. Furthermore, computational approaches that are based on particular psychological theories have also been proposed. They often address domain specific issues starting from a specific psychological theory. In such solutions, the absence of a common conceptual system and/or platform, makes difficult the feedback between psychological theories and computational approaches.

This thesis systematizes and formalizes affect-related theories, what can benefit the interdisciplinary exchange, the intradisciplinary coordination, and hence, allows the improvement of involved disciplines. Specifically this thesis makes the following contributions: (1) a theoretical framework that includes the main processes and concepts that a model of an affective agent with practical reasoning should have; (2) a general-purpose affective agent architecture that shares the concepts of the proposed theoretical framework; (3) an implementation-independent formal language for designing affective agents that have the proposed architecture; and (4) a specific agent language for implementing affective agents which is an extension of a BDI language.

Some studies with human participants have helped to validate the contributions of this thesis. They include classical games of game theory, and an study with 300 participants, which have provided the necessary information to evaluate the contributions. The validation has been performed in three directions: determine whether the proposed computational approach represents better the human behavior than traditional computational approaches; determine whether this approach allows to improve psychological theories used by default; and determine whether the proposed affective agents' behavior is closer to human behavior than the behavior of a purely rational agent.