TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

DIPLOMARBEIT

Affective Acting: An Appraisal-based Architecture for Agents as Actors

ausgeführt am

Institut für Medizinische Kybernetik und Artificial Intelligence
der Medizinischen Universität Wien

sowie am

Österreichischen Forschungsinstitut
für Artificial Intelligence

unter Anleitung von

o. Univ.-Prof. Ing. Dr. Robert Trappl

und

Dipl.-Ing. Dr. Paolo Petta

als verantwortlich mitwirkendem Universitätsassistenten

durch

Stefan Rank

Sonnenuhrgasse 1
A-1060 Wien

Wien, Februar 2004

Zusammenfassung

Emotion spielt eine zentrale Rolle für fesselndes Drama. Konflikte zwischen den Charakteren in einem Stück und die Emotionen, die bei der Lösung beteiligt sind, sind die Bausteine einer dramatischen Struktur, einer Handlung. Diese Prämisse führt zu der Annahme, dass eine Gruppe von Charakteren ausreicht, um dramatische Strukturen nur durch ihre Interaktion in einer konfliktgeladenen Umgebung zu erzeugen, falls sie mittels einer Simulation des Prozesses, den die Appraisal Theory of Emotion beschreibt, gesteuert werden. Das sollte auch ohne eine zentrale Einheit, die die Charaktere kontrolliert, möglich sein. Der Einsatz einer appraisal-basierten Architektur wird demnach als Schlüssel zur Konstruktion von emotional und dramatisch glaubwürdigen Charakteren für interaktives Drama gesehen.

Diese Diplomarbeit präsentiert das ActAffAct Projekt: ein Versuch, diese Ideen anhand einer Implementierung auf die Probe zu stellen, um Erfahrungen bezüglich Machbarkeit und Komplexität zu sammeln. Als Grundlage für die Implementierung der Appraisal Komponenten, die den psychologischen Modellen des Appraisal Prozesses folgt, dient eine belief-desire-intention Architektur für Software Agenten. Für die Konstruktion der Umgebung und für den internen Aufbau der synthetischen Akteure werden wichtige Aspekte verschiedener Dramentheorien berücksichtigt.

Ziele, Standards und Präferenzen, sowie emotionale Ausdrucksarten und sogenannte "Coping" Aktivitäten sind nicht nur zentrale Punkte der Emotionstheorien, sie liefern auch die notwendigen kausalen Verknüpfungen, die eine dramatische Struktur von einer einfachen Handlungssequenz unterscheidet. Mit ActAffAct gelingt die Erzeugung solcher Strukturen, allerdings nur für eine stark eingeschränkte Bedeutung des Begriffs Drama. Dennoch liefert es eine vielversprechende Forschungsperspektive.

Abstract

Emotion plays a central role in engaging drama. The conflicts between the characters in a play and the emotions involved in resolving them are the constituents of a dramatic structure, a plot. This premise leads to the assumption that a cast of characters driven by a simulation of the process described by the appraisal theory of emotion might realize dramatic structures by simply interacting in an environment that is prone to conflict. This should even be possible without any directing entity that controls the characters. Using an appraisal-based architecture is thus seen as a key to construct emotionally and dramatically believable characters for interactive drama.

This thesis presents the ActAffAct project, an effort to test these ideas with an implementation of such characters in order to gain experiences about the feasibility and the involved complexity. A belief-desire-intention architecture for software agents is extended by an appraisal component following a psychological model of the appraisal process. Key features of different theories of drama are considered during the construction of the environment and the internal setup of the synthetic actors.

Goals, standards and preferences as well as emotional expressions and coping activities are not only central aspects of theories of emotion but can also provide a dramatic structure with the causal relationships needed to set it apart from a mere sequence of actions. ActAffAct succeeds to create such structures, if only for a rather limited meaning of drama. It nevertheless shows a promising direction for further work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Imagine taking part in an unfolding story without being limited to the role of the audience as with theatre, film or television. If the lead characters of the narrative do not act as you would, you just modify their motivations to your liking.

Imagine that the responses and actions of other characters do not follow utterly predictable patterns as with most modern computer games. On the contrary, the characters you encounter respond believably and consistently with the actions of others and develop over the course of the story, pursuing their own goals, trying to solve their conflicts and thereby advancing the story, turning it into compelling drama.

Today's animated characters are capable of showing emotional responses in real-time, a necessary precondition for using them to convey dramatic meaning, but what about the engines driving their behaviour and integrating it into a plot? How can we tell a story that functions according to dramatic principles without overly restricting the freedom that a programmable medium like a computer offers? Foreseeing every possible move is not feasible in story settings of more than basic complexity.

The present project, named ActAffAct, researches a bottom-up approach at imitating emotional characters that interact in a story-world with the goal of achieving a plot-like structure while limiting the needed top-down directing control. The ideal level of top-down control would be none at all, resulting in the emergence of plot from the characters interaction.

This chapter will state further reasons why it would be a good idea to build interactive story-worlds, it will enumerate other projects that tried to do something similar (1.2) and describe how ActAffAct, an abbreviation for Acting Affective affecting Acting, attempts to do something new (1.3).

Chapter 2 presents the different theoretical aspects that played a role in the conception of the project, and chapter 3 describes how these are implemented and influence the results. These in turn are discussed in chapter 4, followed by the conclusion in chapter 5.

1.1 Motivation

This work describes the effort to build a control software that reacts and acts appropriately in a simulated environment according to its dramatic role in it: emotionally and dramatically believable software agents.

Drama is the art that deals with a refined version of emotional interaction between individuals. Drama theories describe plays, film scripts and stories in general as the changing constellations of emotional characters (see 2.1). These ideas provide a starting point and success criteria for the creation of dramatic story-worlds. The goal is to simulate characters whose behaviour is coherent in the context of drama. Explicit directions, however, given by a "director" component, should not be necessary. The driving vision for this proof of concept is that the dramatic structure should emerge from the very interaction of the characters, because their behaviour is generated by an emotional system. Simplified formalizations of theories of drama can be used to test the project's results. The final product should bring us closer to software agents that "act in character" without being directed explicitly. The stories that arise should be motivated by emotional constellations, not by plot directions. Based on the observation that emotions are a prominent part of analyzing and writing successful drama, we ask whether the psychological understanding of the functions and functioning of emotion can be used to generate interactions that resemble drama.

Synthetic characters with this ability could be put to use in entertainment and e-learning software. Synthetic tutors or life-like pedagogical agents are used to facilitate playful learning. The "AI" in computer games controls the actions of so called NPCs (non-player characters). Both could be seen as synthetic actors that follow a script but nevertheless have to react to a changing environment in order to show appropriate and believable behaviour. A truly adaptive engine for these applications could reduce the feeling that the interactions are "scripted", i.e. predictable and possibly boring.

Emotion in humans influences their behaviour. The expression of emotion, be it consciously or unwillingly, conveys important information to others. Emotional processes are important for reactive and adaptive actions in a dynamic environment (see 2.2). Synthetic actors with the ability to express emotions and to understand emotions expressed by others would be able to use this information in the planning of their actions. Using an imitation of the human emotion system to influence their actions could provide them with flexibility and adaptivity in an environment that is complex and unpredictable; for example in an environment in which autonomous actors strive for possibly conflicting goals, leading to a struggle that needs to be resolved; i.e. in a dramatic environment.

Interacting with a human user is probably the most unpredictable interaction that a software system can take part in. Software with an "understanding" of emotion could possibly enhance this interaction, not only in the gaming and tutoring area; whether such a software should also have a human-like representation (an avatar) in other applications, is a highly debated subject. But as long as the interaction is more predictable it will lead to better usability [Cooper, 1999], and it will be more predictable, if the user can maintain an intentional stance towards the software artifact [Dennett, 1987]: Humans ascribe intentions and beliefs to others, which allows them to understand their actions in retrospect and, given their limited processing capacity, it is a "reasonable" way for predictions of their future behaviour. A strategy that works well with other humans, but also every object that is complex enough, like most software, is treated this way [Reeves and Nass, 1996]. A suitable basis for software that uses and understands emotions and that allows (and possibly uses) the intentional stance are software agents.

Agents are a very active research topic and the term is often exploited as a buzz-word as there is no consensus in the agent community about what exactly an agent is (see 2.3). But even if we use a weak concept of agency, if we assume that agents are independent entities that can decide about their actions, that not only react but also "pro-act", and interact with each other or maybe even with humans, then an environment including such agents will be highly dynamic. The actions of such agents, however, should be understandable for humans as well as for other agents if they are operating in the context of a multi agent system.

Software agents in general could benefit from the properties that an emotional system promises, it may be the decisive factor in a dynamic environment. Not only the interaction with users is inherently dynamic and unpredictable, also the environment of independent software agents in computer networks, up to a global scale like the Internet, might be easier to master with this advantage. Implementing emotional processes in agents might therefore be useful in areas different from interactive drama.

Using such agents to achieve interactive experiences with emotional impact opens possibilities for art in new media [Murray, 1997, Manovich, 2001, Wardrip-Fruin and Montfort, 2003]. And the implementation of such a system can also be seen as part of the Cognitive Science effort to understand the human mind using agents as a model. It is very likely that emotions are an integral part of human intelligence; integrating simplified versions in working models can further our understanding. Systems that are "situated" only in a virtual, simulated environment might not be perfect for this, but they can nevertheless provide insights into the applicability of emotion theories.

The role of emotion for believability, as a means of conflict resolution in the control structure of agents, and as (one) driving force in drama connects all these goals and it leads to a great diversity of research.

1.2 State of the Art, Related Research Projects

The following section will list a number of projects that are related to the one presented here. All of them deal either with the computational treatment of emotion or with interactive fiction or with both, but they differ significantly in their dates, aims, and assumptions.

On the theoretical end of the spectrum, research efforts aim at flexible architectures for agents that incorporate the understanding and possibly the imitation of emotions. Some projects try to implement a whole framework and tools that can be used in applications that would benefit from knowledge about emotions or from the incorporation of emotional processes. Other projects are oriented towards applications in the pedagogical domain or in the area of interactive fiction.

Different projects naturally start from different theoretical assumptions. In this case, this includes the ideas about emotions, drama, and agents. Furthermore, practical applications might always differ in terms of the technology used.

The following "historical" section will mention very influential projects from the beginnings of this inter-disciplinary field. Afterwards, I will relate my project to more recent efforts, vaguely ordered by the projects' aims.

1.2.1 "Historical" Projects

The computational treatment of emotions has a long history, but only a small number of projects can be discussed here; the following, already completed ones will only be mentioned: BORIS, its extension OpEd and DAYDREAMER [Dyer, 1987]; the first two are systems for the understanding of emotions in written text, the latter manages emotional states that influence its internal behaviour. THUNDER performs ethical reasoning about the motives in textual stories [Reeves, 1991]. SIES tries to map situations described by their features to emotions retrospectively [Rollenhagen and Dalkvist, 1989]. ACRES is a program that explores the idea of action selection by emotion for operation in an uncertain environment [Frijda and Swagerman, 1987]. WILL, a successor of ACRES, is described below, along with the Affective Reasoner, the Oz project and the Virtual Theatre project.

Affective Reasoner The Affective Reasoning project is a collection of different undertakings to explore the possibilities of computational emotion. The dissertation [Elliott, 1992] presents the Affective Reasoner (AR) itself, an "emotion engine" based on the OCC model [Ortony et al., 1988] (see 2.2.2), and its use in the simulation of emotional responses of taxi drivers. The virtual Taxi-World is inhabited by agents that have different personalities and try to fulfill goals like earn money, avoid getting robbed, avoid getting speeding tickets, and so on. Following the OCC model, there are 24 different emotion types, their intensity is determined by several different variables, and emotions that were elicited can be expressed through different channels, ranging from involuntary somatic reactions — like turning pale — to highly conscious plans like taking revenge. The agents also use so-called Concerns-Of-Others databases to model other agents they encounter for up to twice removed reasoning. Possible response actions are selected out of "action databases" influenced by the personality of the agent. Newer incarnations of AR agents can interact with users in realtime using multi-modal perceptive and expressive channels including speech-recognition, text-to-speech, morphed schematic faces and music [Elliott, 1997a]. The Affective Reasoner is a very sophisticated work, used not only for running virtual worlds: it was also used to analyze emotional structures in story-telling [Elliott, 1997b, Elliott et al., 1998].

A recent cleaning up of the OCC-model [Ortony, 2003] that tries to alleviate the complexity of emotion imitation was a starting point for the project presented in this thesis. ActAffAct is therefore very similar to the AR, although less ambitious and less general. ActAffAct tries to avoid reifying discrete emotions and it is aimed at the real-time generation of dramatic structure only, not at the analysis of stories.

WILL [Moffat and Frijda, 1995] describes the architecture WILL, developed as a successor to ACRES to overcome some of its shortcomings. The project focuses on the architectural perspective and is based on the appraisal theory of emotion of Frijda (see 2.2). The model consists of separate modules running in parallel on a shared database of facts. All facts the system knows about are regularly checked for relevance against the concerns of the system. Relevant ones are then subjected to secondary appraisal, in the terms of Frijda, possibly leading to a change in action readiness, i.e., the probability of specific actions changes.

In [Moffat, 1997] an implementation of the model for the purpose of playing the prisoner's dilemma is presented. Emotion and personality are seen as closely related, emotions being described as temporal inconsistencies, personality as temporal consistency.

The domain of the prisoners dilemma is a restricted test-bed for an emotional agent, the only expressive actions to choose from are one of the two possible moves at each turn and verbal (i.e. written) statements that do not affect other agents.

The Oz-project The aim of the Oz-project at Carnegie-Mellon University was to facilitate the creation of interactive fiction [Bates, 1992] including synthetic characters. Intelligent emotional agents that live in virtual worlds should act as believable characters. The underlying system used to generate emotions called Em [Reilly and Bates, 1992] is also based on a subset of the OCC-model and integrated into the overall architecture called Tok. The Em system includes standards and attitudes and generates discrete emotions when sensing events that conflict or conform with these.

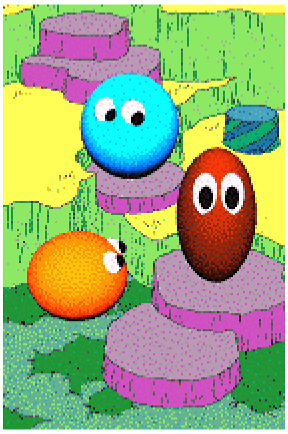

Figure 1.1: The Woggles

Oz was implemented in two ways, as a realtime animation version and as a text-based one similar to classic computer games. The latter gave rise to Lyotard, a simulated house cat. An early manifestation of the animated Oz were the Woggles in the Edge of Intention [Bates et al., 1993], ball-like creatures that display emotions by their morphing movements and do not act according to a predefined narrative (Fig. 1.1). The Woggles have been a major influence for this thesis, although they focus on engaging interaction with characters and not on dramatic structure. [Bates, 1994] stresses the importance of appropriately timed and clearly expressed emotion, a prerequisite to turn software agents in virtual worlds into *believable agents*, a term coined by Bates, alluding to the notion of a believable character used in the art of writing. Such believable characters give the “illusion of life” which allows the “suspension of disbelief" for the audience, the ultimate goal of Hollywood-style cinema. An ongoing work related to Oz that focuses more on interactive drama is Façade, described in the next section. The original Oz-project has been discontinued as of December 2002, Zoesis is a related spin-off company.

The Virtual Theatre project The Virtual Theater project at Stanford University aims to create a virtual environment for users to take part in a theatre production, with the computer fulfilling all the missing roles [1]. The current focus lies on individual synthetic characters that can be directed, but also act according to their personality and in a way consistent with their emotions, moods, and roles. Several sub-projects have explored different applications of this idea.

The Virtual Theater project used the original Woggles from the Oz-project to implement Animated Puppets, a system where the user can control one Woggle using a simple interface, controlling its mood and giving simple directions for actions and utterances, while another one "improvises" suitable reactions. This production focused on an engaging interaction for children [Hayes-Roth et al., 1995].

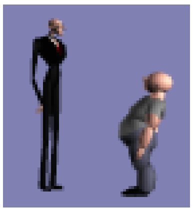

Figure 1.2: Master and Servant

Other sub-projects study personality [Rousseau and Hayes-Roth, 1997] and modes of interaction with synthetic characters [Maldonado et al., 1997]. The Master/Servant scenario (Fig. 1.2) shows the interaction between two autonomous agents that exemplify the reversal of roles, an often used pattern in improvisational acting where the status of the characters change, the servant overpowering the master and taking its place [Hayes-Roth et al., 1997].

[1]Further infos at http://www.ksl.stanford.edu/projects/cait/index.html
(This and all further weblinks were last checked on February 12th, 2004)

1.2.2 Recent Work

There a very different projects that involve emotional agents or dramatically believable agents. Focussing on architectures and frameworks often involves a more principled approach to the treatment of emotion, while the focus on pedagogical applications or interactive fiction is rather product-oriented, although theoretical considerations and their application can and should not be separated. I will nevertheless start with research with a tendency towards theory and the implementation of general frameworks.

Architectures and Frameworks

The Cognition and Affect project Aaron Sloman leads the very broad research effort at the University of Birmingham, entitled the Cognition and Affect project [Sloman, 1997]. Their aim is not to search for a specific architecture that implements a specific capability comparable to a human one. They want to map the space of possible agent designs and their applications ("design space" and "niche space" in terms alluding to evolutionary biology), always considering as broad an architecture as possible, even if it may remain shallow. This approach might lead to insights not only about how our minds work, but also why our minds work exactly like they do, whether there would be other functional possibilities, and how they relate to each other[2].

As a sort of testbed for some of the ideas, they developed the minder-scenario, a simple world for agents to live in. One is the minder or nursemaid, whose task is to look after "baby"-robots in a limited area, keeping them out of trouble until they are mature enough to leave this nursery. Sources of trouble are e.g. ditches the babies could fall into or their innate need to be recharged regularly.

The architecture developed in Birmingham uses a layered approach and specifically tackles the issues of dealing with restricted processing resources and other constraints any physically existing information processing architecture is necessarily subjected to. Some research focused on the relations of such an architecture to high level emotions like grief, originally seen as a perturbation resulting in a partial loss of control [Wright et al., 1995]. Furthermore,

[2]More infos at http://www.cs.bham.ac.uk/~axs/cog_affect/COGAFF-PROJECT.html

Sloman addresses the question how a well-funded conceptualization of emotion could result out of such architectural constraints [Sloman, 2001].

Affective Computing at MIT The Affective Computing division at the MIT Media Laboratory explores possibilities to enhance computing by including emotional information. They include communicating affective information from and to the computer, sensing, and recognizing affective reactions in humans also by using new interfaces.

Synthesizing affect in machines is one of their research interests that "stands somewhat apart from the rest" according to their website[3]. One of their projects attempts to incorporate affective signals into the gaming engine Quake2 (see the description of Haunt2 in section 1.2.4), another one tries to create an Affective Learning Companion. This agent is intended to assist children in their learning efforts by recognizing their affective states and responding accordingly, encouraging them, posing questions, pointing, etc. The emotions are modelled as points in a five dimensional space specifically tailored to the learning situation [Kort et al., 2001].

Figure 1.3: Sheep|Dog: Trial by Eire

The C4 brain The Synthetic Characters Group at the MIT Media Laboratory has developed an architecture for a synthetic "brain" called C4. Two interactive installations have been implemented using C4: "Sheep|Dog: Trial by Eire" (Fig. 1.3) and "Clicker" [Isla et al., 2001, Burke et al., 2001], both starring a virtual dog the user may train using vocal commands. Their focus lies on the imitation of animal behaviour, which reduces complexity compared to modeling

[3]http://affect.media.mit.edu/AC_research/synthesizing.html

a human mind. (The same holds for generating behaviour typical of dramatic stock-characters.)

The C4 cognitive architecture consists of a perceptive subsystem that receives only input the creature could realistically get from its subjective point of view. It includes a history of percepts and the expectations about future ones, thus enabling the system to be surprised about novelty in its surroundings, which is also one of the key appraisal criteria in emotion theories. At the end of the processing stages lies a complex navigation and motor component. The system also features a reinforcement learning system for which the Clicker installation, allowing to train a virtual dog like a real one, is a showcase.

High level emotions like pride, hatred, or gratitude are not used, as the architecture is mainly intended to model a dog's mind, and, of course, because the rest of the architecture provides big enough a challenge already.

"(void*) : a Cast of Characters" was another installation of the Synthetic Characters group, presented at SIGGRAPH99. It featured creatures the user could dance with using a simple interface. The different characters reacted according to their personality and explicitly modeled emotional states. The presented system also included a separate camera and music creature that dynamically directed the virtual camera and the musical accompaniment respectively [Tomlinson, 1999].

SAFIRA Supporting Affective Interactions for Real-time Applications[4] was a European research project with the ambitious goal to provide a whole framework for affective interactions. All stages in the information processing during the interaction with an intelligent agent shall be enriched by providing tools to integrate affective information and processes. Knowledge acquisition, representation, reasoning, planning, communication and expression are listed as the addressed components that should in the end lead to more believable interactions.

The questions raised in the present work are just a humble subset of those of the SAFIRA project, although the latter is focused on the problems of conveying affective information in human-computer interaction, and not between agent and agent in a dramatic context.

[4]Further infos at http://gaips.inesc.pt/safira

The Influencing Machine [Sengers et al., 2002], FantasyA (a game using SenToy an "affective input device" [Paiva et al., 2002]) and James the Butler (a personal sales assistant) are programs that demonstrate some of the possibilities of the SAFIRA toolkit [Paiva et al., 2001].

Cathexis The Humanoid Robotics Group at the MIT houses research about robots with human qualities. As far as emotions are concerned, these robots mainly try to either recognize them in humans or to imitate them in speech and gestures. The Cathexis model [Velásquez, 1998], however, captures the idea of reaching higher level cognitive emotions by basing them on "affect programs", a term that represents evolutionarily successful stimulus-response patterns. These are to correspond to the so-called basic emotions, a disputed concept in emotions psychology.

Cathexis and its application in Yuppie, an emotional pet robot (a Yamaha puppy), therefore focuses on basic control circuits for specific emotions, which is in contrast to the higher level emotion processes that this thesis tries to capture.

TABASCO [Staller and Petta, 1998, Petta, 2003] present a Tractable Appraisal-Based Architecture for Situated Cognizers, an architecture for situated agents that aims at integrating the emotion process, introducing all the benefits that emotion offers to agent engineering. It uses a layered approach and considers several variations of the appraisal theory of emotion (see 2.2). The architecture has been implemented e.g. in an interactive exhibit: The Invisible Person [Petta, 1999].

TABASCO is of special importance to this thesis as it was a main inspiration for it, its influence showing clearly throughout this work.

1.2.3 Pedagogical Applications

Several projects explore the use of computers for pedagogical purposes. Most of the time when software agents are involved, they are intended to be synthetic tutors facilitating and maybe enhancing the learning process. In most of these applications the agent takes the role of a pedagogical expert or at least uses knowledge from pedagogical expertise. In rare cases, however, technology is used in a simulation of the domain about which the user may learn by interacting with it.

Figure 1.4: Teatrix

The Support and Guidance architecture (SAGA) is an example that was applied in **Teatrix** [Machado et al., 2001], a virtual story creation environment for children (Fig. 1.4). The dramatic theory used to formalize the process of story creation is mainly based on the ideas from Propp's analysis of the Russian folktale [Propp, 1968]. In Teatrix, children control the actions of several characters and a director component keeps track of the flow of the story and, if unexpected actions are taken, will question the child for its motives. One can easily imagine software agents taking on the roles of missing story characters.

This idea is realised in **Carmen's Bright IDEAS** [Marsella et al., 2000], an interactive pedagogical drama designed to help mothers of pediatric cancer patients. The characters in the story are autonomous agents, and two other agents — a director and a cinematographer — manage the flow of the story. The agent models employ emotional appraisal as a key component to influence action selection using a two-step appraisal process (see 2.2). The interaction with the user is accomplished through predefined dialog choices.

Another project along these lines is the **Mission Rehearsal Exercise**, which however uses a virtual 3D environment. It is funded by the U.S. army and intended to teach decision-making skills to soldiers [Hill et al., 2003]. The "virtual humans" in this project also model emotion by using appraisal and coping mechanisms implemented in SOAR, a reasoning system that has also been used in the virtual tutor Steve [Rickel and Johnson, 1997].

1.2.4 Interactive Drama and Virtual Worlds

The following projects try to achieve dramatic interaction with agents in a virtual world.

SimHuman and DIVA These projects at the University of Piraeus are trying to situate autonomous agents based on a BDI architecture into a virtual 3D-environment. Distributed Intelligent Virtual Environments is an approach to combine the research in Distributed Artificial Intelligence and Virtual Environments [Vosinakis et al., 1999]. As an example they designed a VRML maze for an agent to escape from [Panayiotopoulos et al., 1999].

SimHuman is an effort to model realistic human-like agents that can populate these virtual worlds [Vosinakis and Panayiotopoulos, 2001]. It is concerned with physically based modeling of the 3D avatars, the agent architecture used for planning is a BDI-model (see 2.3.1). An example using the classic blocks world task (stacking blocks in a specific order) was implemented; emotional assessment of the situation, however, is not used.

SCREAM Scripting Emotion-based Agent Minds is a system that was developed to allow for synthetic emotional characters in web-based interactions[5] [Prendinger and Ishizuka, 2002, Prendinger et al., 2002]. It incorporates ideas from appraisal theory and affective communication to yield easily customizable interface characters. Investigations into the role of social status in such communication scenarios are also considered.

Figure 1.5: “Akko-chan’s Got a Secret” using SCREAM

[5] Details at http://www.miv.t.u-tokyo.ac.jp/~helmut/agents.html

Helmut Prendinger has implemented different scenarios to demonstrate the capabilities of the system. The coffee shop features an employee that reacts differently to a customer request and to questioning by his boss, depending on his mood and the tone of the questions. Interaction with the system is accomplished through selection out of a choice of predefined responses. Another scenario features a black jack table in a casino, while a third one is inspired by a Manga comic (Fig. 1.5). The most recent scenario deals with corporate decision making and is driven by a story that is influenced through the user's decisions.

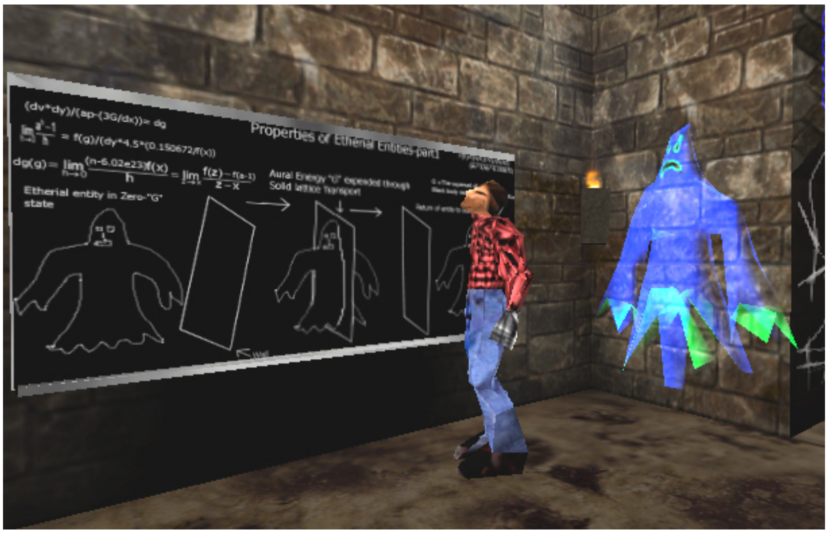

Figure 1.6: Haunt2

Haunt2 [Laird et al., 2002] describes this game that is realized as a "mod" for the Unreal game engine (Fig. 1.6). The SOAR agent architecture is used to create believable characters that populate the game world, as in the predecessor project, Quakebot [Laird, 2001]. The player takes the role of a ghost that is trapped in a house populated by several characters who are unaware of its existence. The goal of the game is to escape from the house, and as ghosts cannot interact easily with their environment the user must resort to influencing the other characters.

The main objective of the project is to learn about the sensing and acting abilities of the characters and their simulated physical drives, and to try to use them for one's own purposes, by inducing emotional reactions according to their personality.

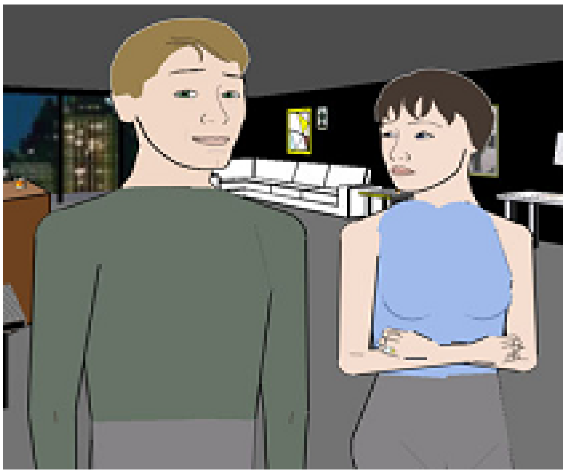

Figure 1.7: Façade

Façade Façade[6] tries to achieve an interactive drama in a realtime 3D world (Fig. 1.7). According to [Mateas and Stern, 2002] it will be publicly released in 2004. The user will experience a narrative from the first person perspective, taking the part of a longtime friend of a married couple that invited her for dinner, an occasion where problems unfold. The goal of the project is a very rich interaction not only through dialog but also through embodied interaction.

The general aim of the project is similar to that of this thesis: to create dramatically interesting worlds populated by emotional agents. But there is a significant difference in the method used: In Façade, there is a separate component in the system that monitors the conformance to a theory of drama, i.e. a drama-manager that influences all agents to achieve a coherent structure of the narrative according to dramaturgy. The drama is managed as a "bag of beats" that structures the plot events to achieve a dramatic curve that builds up tension until a breaking point is reached, and tries to find a resolution at the end (see 2.1).

The ActAffAct project in contrast tries to achieve a simpler, but similar effect without central control.

[6]Details at http://www.interactivestory.net/#facade

Interactive Storytelling A project at the University of Teesside is aimed at creating dynamic narratives capable of reacting to the interaction with a user. [Cavazza et al., 2002] describes a system that uses real-time hierarchical planning techniques to simulate the roles of actors. Their interaction is the basis to generate a storyline, whilst the user can interfere at any time. An interference might change the plans of the characters, thereby changing the course and outcome of the narrative.

This approach is very similar to the one explored in this project, the main difference lies in the use of an emotional process as an integral part of the planning system that drives the actors.

1.3 Outline of ActAffAct

The project that this thesis presents is called ActAffAct. Its aim is to create believable agents that interact in an environment ripe for conflict. It is a test of the idea that the interaction between simple agents, capable of experiencing simple emotions, is sufficient to achieve simple drama. The hypothesis is that this holds true in a (simulated) world loaded with possibilities for conflict, and that there is then no need for a "director". Such characters that imitate emotion using an appraisal process, and act appropriately in a dramatic environment, would be perfect story-telling tools in a programmable medium.

ActAffAct is a system for experimenting with multiple agents in an environment that is conducive to dramatic conflict. The environment is visualized using 2D graphics and basic animations. At the beginning of chapter 3 I will briefly discuss the system surrounding the agents. The agents themselves use plans for task-directed behaviour, but they are mainly driven by emotion-related plans and actions resulting from appraisal of their perceptions. Emotions of the agents are not explicitly represented, an appraisal component evaluates the relevance of perceived events as an ongoing process.

Stories What kind of stories can be expected with such a system and how to evaluate whether an interaction counts as drama? The goal I want to achieve are only "minimal" plots, stereotypical cliché stories enacted by so-called stock-characters. Section 2.1 presents a glimpse at the wide field of drama theory and highlights the ideas underlying this system: What is the minimal plot of a story, how does it relate to the characters in it, and why is emotion important or even necessary for it? It also presents an approach how to evaluate the "story-ness" of a series of interactions.

The section about the implementation (3.2.3) presents the basic setup of the ActAffAct system: the four prototypical dramatic characters, one Hero, one Villain, one Victim, and one Mentor agent, modeled after archetypes of storytelling. What obstacles and conflicts might they deal with?

Emotions The relevant details of appraisal theory are presented in section 2.2. A lot of different but still related theoretical treatments of emotion exist, some of them are more suitable for an implementation in a computer system. This section goes into detail about the ideas that had to be specified more exactly to be amenable for implementation.

Even more detail about the actual implementation can be found in section 3.4: What methods were used to determine relevancy of the perceptions, and how appraisals are mapped to action tendencies; what simple learning strategies can be used to adapt the agents during their interactions; what other relevant decisions were made during development, in a more ad-hoc manner.

Agents Section 2.3 explains the special features of software agents. Agents are used only as a vehicle for creating synthetic actors, the specialized field is considerably richer. This section specifically presents the relevant ideas that could benefit from an implementation of emotions. It also traces the roots of the specific agent architecture that this project is based on, an instantiation of the so-called Belief-Desire-Intention (BDI) architectures.

How the agents in ActAffAct are built is shown in section 2.3.2, and more details follow in chapter 3. The roles of specific plans, facts and goals in the model are presented, and the differences between the four agent types are explained in further detail.

A further enhancement and natural extension of the architecture would be to design a satisfying interaction that enables a user to influence one or more of the agents directly, this would shift the focus of the project from dramatic structure towards engaging interaction. Different levels of control could lead to completely different styles of interaction. This is certainly a desirable direction for future work, but for the sake of simplicity, the complexities of a good user interface were excluded from this project.

Chapter 2

From Theory...

This chapter will explain the theoretical work that influenced the ActAffAct project. The effort to use computing technology to implement a psychological theory in order to yield a narrative is obviously an interdisciplinary one. The first of the different fields that are relevant is drama theory. Later on in this chapter I will present how psychology's view of emotion can be fitted to fulfill a driving force in narrative. And the final section will deal with a suitable foundation in computer science to build on, agent theory, an area which might benefit from the application of psychological ideas.

2.1 Drama Theory

There are different types of theories about drama depending on what purpose they serve. The main function of most drama theories is to analyze already existing drama in retrospect, only few of them try to provide guidance in writing successful or good drama (which is not necessarily the same). Both of these are important for this work, as on the one hand I will try to generate something like a story automatically. On the other hand I also need a means to evaluate the success rate and thus a measure of the "story-ness" of the product.

We might also learn from looking at drama theory how to choose a sensible cast of characters. What types of characters will be needed and what kind of actions must the environment they will be in support in order to make dramatic interaction possible at all. What is a minimal version of dramatic interaction? The result will probably not be interesting by itself but it should be sufficient for testing whether the behaviour of the participating characters can be called dramatically appropriate, or, more daringly still, believable.

I will therefore start to give a broad overview of some of the theories that use a more formalized system of analysis, as those will influence the way of testing my results. This is followed by an account of guides for authoring a story, which use a kind of generative approach. The next section will point out how relevant emotions are in all of these theories, which leads to a possible way of describing a generated story in emotional terms. An example of such a minimal story will be given in a very simple toy-world, showing the dreadful restrictions and therefore lowering too high an expectation about the outcome of this project. General problems about the interpretation of a story by human observers will be mentioned, too.

2.1.1 Analysis and Formalization

Theories that are used for the analysis of stories can help to understand the internal structure of a narrative. Using a suitable decomposition of stories they could also be used for a top-down generative approach using combinatorics; the focus here is, however, the possibility of a bottom-up approach.

An example of a rigid structural analysis of stories is the "Morphology of the Folktale" [Propp, 1968], which focuses on Russian folklore. It is an attempt to derive a suitable classification scheme for tales based on their structural properties. According to Propp the majority of folktales in his study are constructed from so called plot moves taken from a sequential list: not all of them are present in every story, but their relative order is always the same.

In folktales, there are only about 30 of these plot particles which are termed "functions of the dramatis personae". A function is seen as the action of a character in terms of its significance for the rest of the tales plot [Propp, 1968, p. 21].

A tale is analyzed as starting from an initial situation α, connecting a selection of functions acted out by one or more of the cast until it ends (usually in a wedding). All of these functions represent categories of possible actions: Propp gives every one of them a symbol, thereby reducing Russian folktales to a series of Greek and Roman letters (except for two of the functions, "departure" and "return" of the hero, which are important ones that almost no tale misses. They have been assigned an upward and a downward arrow).

Examples of these functions are the "violation" of an "interdiction" shortly after the beginning of a tale, "trickery" — usually executed by the villain — and "punishment" at the end of the story (usually of the villain). They range from rather specific, like "provision or receipt of a magical agent", to very general categories, like "difficult task". The latter function for example could be fulfilled by milking a herd of wild mares, riddle guessing, or winning a game of hide and seek. The former bears a significant resemblance to the plot moves of Campbell (see below), as several others of the functions do.

Some of these functions can be grouped according to by whom they are normally executed. Specific members of the dramatis personae tend to fulfill, not surprisingly, specific functions. The classes of characters in Propp's analysis include a villain and a hero, the princess and her father, a donor, a helper, a false hero, and a dispatcher.

Several other books about drama lay emphasis on the structural analysis. [Pfister, 2001] is a very complex guide handbook on literary treatments of drama focussing on theater. It contains sections about the cast of a play and possible constellations of characters, furthermore an extensive section about technical terms (e.g. the German terms Handlung, Geschichte und Geschehen translatable as plot, story and action) and about the possibilities of presenting the plot either directly or through dialog. It is aimed at descriptive analysis, but is far too complex for evaluating plots of the kind we might generate.

[Martinez and Scheffel, 2002] is another theoretical book, that looks at the general process of telling stories (albeit in the form of literature). The general questions posed are how something is narrated and what is told. The "how" part is again tailored to descriptive analysis and categorization and not applicable to this project, as the goal of ActAffAct is to generate a plot without varying the mode of telling it. The "what" part again tries to define important terms for describing a story, it also stresses that the plot of a story is defined not only by a sequence of events but mainly by the *causal relationship* between these events. They cite an illustrative example, originating from the English novelist and essayist E.M. Foster.

> "The king died, and then the queen died", is a story.
> "The king died, and then the queen died of grief", is a plot.

Please note that the causal connection between the two events is an *emotional reaction*.

Martinez and Scheffel also mention the resemblances of narrative theory with ideas from other disciplines, like cognitive psychology and anthropology.

The monomyth [Campbell, 1968] is a very influential book that describes every existing story as the derivation of an archetypal monomyth. Though it can be doubted that really every successful story must adhere to the principles he lists as defining the steps of a hero's journey, many a book and screenplay has been shown to work according to this monomyth.[1] Hollywood producers are said to use it extensively as an indication of the quality of a new screenplay, [Vogler, 1996] is a newer take of this basic story specifically targeted at screenwriters, and the software Dramatica (see 2.1.2) also owes a lot of its concepts to Campbell.

The basic premise of the monomyth is that every story can be seen as the journey of a hero, from her ordinary surroundings into an unknown and challenging world. It can either be a physical one, searching for a special item for example, or an inward journey of the mind and heart. Any good story needs the hero to grow and change, quoting Vogler:

> [...] a journey from one way of being to the next: from despair to hope, weakness to strength, folly to wisdom, love to hate, and back again...

Note again the abundance of emotional terms.

This journey is the story of the archetypal hero, during which she meets several other significant types of characters, which need not be embodied as actually different characters, but there are ones that mainly fulfill one of these functions (for an example see below about Dramatica). To give you an impression how this monomyth can be described, here is a short summary of the prototypical hero's journey, adapted from [Vogler, 1996]:

> The hero starts in her common-day world (her status-quo), from where she is called to adventure as a problem or lack is spotted.

[1] Star Wars being one of the well-known examples.
See e.g. http://www.nasm.edu/StarWars/

> After first refusing the call, she meets her helper and eventually crosses the threshold into the unknown world, by overcoming the threshold guardian. She is then subjected to trials and tests ending in a supreme ordeal (the central crisis of the story). Be it passing or failing the ordeal, afterwards the hero is changed and she can take the road back home. This is complicated by a last climactic difficulty, and finally she returns into the ordinary world, changed by the events and possibly bringing along the remedy for the initial problem.

While this is a far cry from a formalized structure, it can nevertheless help at a qualitative evaluation of generated stories. [Campbell, 1968] is analytical in its nature, while [Vogler, 1996] tries to harness the ideas in the form of a guide for writers. And there are many more of this sort.

Section 2.2.3 in this thesis mentions the core relational themes of the appraisal process [Lazarus, 1991] which resemble concise descriptions of the elements of the monomyth.

2.1.2 Writing Guides

A lot of the practical aids for the novelist, the playwright, the screenwriter are top-down in their nature, starting from a basic structure and refining its constituent parts step-by-step. The basic structure is normally one out of a limited number of basic plots. Apart from [Vogler, 1996], there is the example of [Tobias, 1993] with 20 master plots, and the line of ancestors may start with Aristotle's two basic plots of tragedy and comedy. Another famous basic number is 36, attributed to Carlo Gozzi and taken up by Georges Polti in [Polti, 1977], who claims that these plots derive from the 36 different emotions a human can experience, without listing them explicitly. There is no structure in his enumeration, the plots are defined by two or three main characters and a description of their relationships and conflicts, as for example plot number 19, "Slaying of a Relative Unrecognized: The Slayer, the Unrecognized Victim".

Some of the writing guides, however, start from a detailed view of the sine-qua-non parts of a good story, most notably a promising character constellation. [Cowden et al., 2000] enlist several archetypal characteristics of heroes and heroines, each one with its special qualities and flaws, with its special way of thinking and feeling. For them it is the characters who define the plot.

> [...] the action in a story, from the inciting incident to the epilogue, must be inevitable in light of the personalities of the characters. [Cowden et al., 2000, p.97]

Their categorization scheme, however, does not lend itself to an implementation. Its clear-cut distinctions of archetypes are not easily applicable, as heroes in typical stories are often evolving from one class to another or are, as they call it, layered archetypes to begin with. We thus might need to dig deeper.

[Seger, 1990] is another book that focuses on the importance of the cast of characters for a story. For her, the starting point of writing a good story is knowing the backstory of its characters, everything that happened before the narrated events, which is only hinted in the actual story. These defining characteristics of the cast, their psychological profile, and the relationships between them are enough groundwork to build an enticing story.

Characters with a rather rigid pre-defined profile, so called stock characters, are the principal building block used in 16th century Commedia dell'Arte. Each participating actor only chose one of the well-known roles and they all agreed on a basic theme and a suitable ending for the play. This kind of theatre relied heavily on the improvisational talent of the participants, but nevertheless there were rather rigid rules of combining the possible bits of comedy into a coherent whole [Duchartre, 1966].

A very influential book for writers is [Egri, 1946]. Its premise is that a good story needs three basic building blocks: a premise, characters, and conflict. It takes at least two opposing characters, protagonist and antagonist, to get a narrative going. Their conflict is the source for the main action and the society of characters advances it, the "characters are plotting their own play". The conflict arises normally when personalities with different affective traits are put into opposition by their environment (including other characters).

But what about premise? Egri states that for a writer it is essential not to begin with a set of characters or a conflict-prone environment; the starting point needs to be a premise, what the whole narrative should be about, what message it should convey. Examples are "Great love defies even death" (e.g. Romeo and Juliet, West Side Story), "Jealousy destroys oneself and the object of one's love" (Othello). This premise is the guide in choosing characters and an environment for conflict. A well-formulated premise always consists of a character trait; an indication of conflict; and the suggested end point of the narrative, as for example "Dishonesty (character trait) leads to (conflict) exposure (the end)", the premise of many a comedy plot. For Egri, the premise trumps emotion and situation.

> No emotion ever made, or ever will make, a good play if we do not know what kind of forces set emotion going. Emotion, to be sure, is as necessary to a play as barking is to a dog. [Egri, 1946, pp.6f]

An environment could be ripe for conflicts that deal with several possible premises, but according to Egri a kind of director would also be needed to influence the characters, or as an alternative all of them need to have knowledge of the premise that should be proved by the story.

The commercial software Dramatica Pro[2] is a software for authors that helps to plan a narrative in a very structured way. It incorporates and enlarges many of the ideas mentioned so far and includes a voluminous theory section which is a good example of how more complex and convoluted the theory of narrative can get once a vast variety of working stories is to be subsumed. The character is again a central concept, and an explicit theme is used that should lead to a plot, detailing the conflict between the characters.

According to this system of character types, there are eight archetypal characters (Protagonist, Antagonist, Reason, Emotion, Sidekick, Skeptic, Guardian, and Contagonist). They can be seen as pairs of characters: some of them drive the action, others are more passive. Any two but for those that form an opposing pair, can be combined as one actual member of the cast. These characters are seen as separating the different motivations in a human, which are then explicitly dealt with in a narrative. The following tables (2.1, 2.2) show the character types of Dramatica as well as possibly corresponding terms from Campbell/Vogler (although the mapping is not as simple as shown here), and some examples from well-known films.

[2] http://www.dramatica.com

Protagonist (Hero)	⇔	Antagonist
Guardian (Mentor)	⇔	Contagonist (Threshold Guardian)
Reason (Herald)	⇔	Emotion (Shadow)
Sidekick	⇔	Skeptic (Shapeshifter, Trickster)

Table 2.1: Dramatica's eight archetypal characters as opposing pairs (terms from Campbell or Vogler in parentheses)

Dramatica	Star Wars	The Matrix
Protagonist	Luke	Neo
Guardian	Obi-Wan Kenobi	Morpheus
Reason	Princess Leia	Trinity
Skeptic	Han Solo	Cypher
Antagonist	The Empire	Agent Smith

Table 2.2: Examples of clear-cut archetypes in film

What do all these theoretical ideas about narrative imply for the possibility of a generated story? What characters and what environments are needed to make a new kind of storytelling possible, in the computer as a new medium for storytelling, as described in [Murray, 1997]? Ideally, authorship would consist just of providing an environment in which to participate, and creating characters with which to interact. What are the requirements for a minimal version of this ambitious idea?

2.1.3 Minimal Requirements for a Story

In all of the theories reviewed, emotional terms featured prominently when describing the structure of a narrative. Emotions are used for the causal links between events in a story. Could it be sufficient to describe a story as a succession of emotional reactions? Mateas and Stern use the "beat" as the basic unit for stories [Mateas and Stern, 2000]. The main purpose of a beat is to change *dramatic values* that describe the situation between the characters, character traits or emotional relationships. Action that does not change a value does not qualify as beat.

Drama is the interaction of multiple characters, and can be seen as governed by intentions and driven by emotions. If you had a system that imitated several emotional characters reacting autonomously to each others' actions and the events in their environment, it could be enough to provide a sufficiently interesting cast of characters, put them in an environment ripe for conflict, and watch the story unfold from there.

The naive premise of ActAffAct is that such a system can be implemented (as a radically simplified sketch) as a multi-agent system. Egri's concept of premise that guides the whole narrative, however, will be excluded from the reasoning, on the one hand because it might be enough to provide such a diversified environment that several premises could be seen as starting points of the interaction, on the other hand, because the proposed plots to be achieved here are of such a simple and limited structure that discerning premises would not make sense. ('Love overcomes hate' should suffice.)

What kind of a character constellation should be used as a starting point? I chose to use four players: a hero (protagonist), a villain (antagonist), a mentor, and a victim. The goal is for them to act believably in their environment and to choose courses of action that qualify as dramatic. To make conflict possible, the characters must be guided by their intentions and at least two of them need to have opposing goals (with respect to the shared environment).

What does it mean for the interacting characters to act believably in a story context? Their actions should be understandable in the light of their history in this environment and what we already know about them. Hints about the motives, that originate in a character's back-story, are very varied and often subtle in stories, and as the goal here is only to generate suitable courses of action, these problems will be ignored. The aim is to have several (i.e. four) autonomous characters interact with each other so that their reactions to each other can be interpreted as part of a storyline. The environment and the actions feasible therein should be as simple as possible, examples might be stealing objects from somebody, using a rope to bind somebody, or giving a key as a reward for solving a simple puzzle.

So, imagine a simple world with four characters, two of which have opposing goals, like getting a third one to love them (and only them, otherwise there would be no conflict). The world is filled with simple objects, to be used for good or bad purposes, like flowers, ropes or swords. What could be a minimal story-line for such a world?

> Hero falls in love with Victim and wants to give her a flower. This fails, because Villain has bound her with a rope, so Hero has to solve Mentor's puzzle of stacking blocks, in order to get the keys to the treasure from him. He then bribes Villain with the treasure and saves Victim by cutting her free with his sword. She falls in love instantly, of course.

2.1.4 "Story-ness"

Suppose we generated a plot, how to evaluate if it is any good? If evaluation was to be done automatically, the bottom-line is we could not easily achieve that. Building a system that could do that would be a major endeavour in itself, the result of which could possibly be used for a generative approach. But as this is a proof of concept project, a qualitative check of the output should suffice.

The story-ness of a plot is a highly subjective quality anyway, depending on either the taste of the viewer or the particular theory used for analysis. A final test of generated stories would be to use the system in an interactive way and to question the audience, but this project concentrates on implementation issues (and such evaluations include numerous issues in their own right).

2.2 Emotional Agents and Appraisal Theory

What can be implemented in computational characters to create the impression of emotional behaviour and believable reactions? This section will give a short overview of current psychological models of emotion and the emotion process with an emphasis on those parts that can be implemented in computational systems. The last part will summarize the eclectic features used in the realization of the ActAffAct system.

2.2.1 Action Leads to Emotion Leads to Action

> *Felix qui potuit rerum cognoscere causas*
>
> Virgilius

Emotion is a way to reason about the subjective significance and the causes of events and actions we observe, it is possible to make sense of others by using simplified models of them to explain their actions. This is also the way in which emotions are used as explanatory parts of drama theory. But what process in an individual is causing emotion and the actions that are attributed to emotions?

Most words that are used in emotion research have a vast range of connotations stemming from "folk-psychology", a term used in psychology and philosophy of mind to denote the common understanding of mental states (or commonsense psychology). The usage of words like "belief", "desire", "pain" or "anger" in natural language points to the concepts and possibly a whole theory used by us to explain behaviour of other people; others interpret them as a proof that no corresponding mental states exist ([Davies and Stone, 1995, Goldman, 1993] for a glimpse of the controversy). Nevertheless, these words probably only scratch the surface of the underlying processes responsible for the experience and display of emotions that allows us to use these terms for reasoning. One of the motivations of emotions research is to clarify the nature of the responsible structures and processes.

Processes involving emotion are now considered to be an important, if not downright essential, part of the control structures of boundedly rational situated beings. They allow them to react adaptively in a dynamic, changing, and highly complex environment. They are highly flexible (compared to reflexes e.g.) and still do not necessarily require conscious or voluntary reasoning.

Physiological reaction	arousal, mostly uncontrollable
Motor expression	e.g. facial expression, reflects internal state
Subjective feeling	motivation, influence on cognitive processes

Table 2.3: Observable indicators of emotions (response triad)

Psychology has several theoretical ideas about this process in humans (see [Lewis and Haviland, 1993, Dalgleish and Power, 1999, Davidson et al., 2003] and with a different focus [Martin and Clore, 2001, Forgas, 2000]). The one family I will elaborate upon is the cognitive appraisal theories of emotion. Here, a basic function of the emotional in a system is seen as the permanent assessment of the surroundings according to the entities' goals, intentions and standards. Its situation in the environment is checked with every perception for its relevance to the individual. Several criteria are evaluated. When a perception is found to be of (e.g., negative or positive) importance, this does, however, not lead to a response directly. Another level of indirection is used. Appraisal triggers first a motivation for possible further action, a so-called action tendency.

For an implementation a very detailed theory of this process is needed. There is some consensus about what criteria are needed for the appraisal process (e.g. Scherer's "stimulation evaluation checks" [Scherer, 2001]), but less on how such checks are exactly executed. Opinions differ also on the influence of action tendencies.

Out of the several works on appraisal [Arnold, 1960, Frijda, 1986, Ortony et al., 1988, Frijda, 1993, Scherer et al., 2001], a very influential one for implementation and application has been [Ortony et al., 1988]. Examples of such deployed systems include [Bates et al., 1993, Elliott, 1994, Elliott et al., 1997, Martinho and Paiva, 1999, André et al., 1999, Bartneck, 2002, Krenn, 2003].

The next section will give a short overview of the revised version of the OCC-model published in [Ortony, 2003].

2.2.2 The Ortony2003 Model

Ortony classifies the types of emotional reactions on the one hand, and the types of emotional responses on the other. The first scheme includes what can be subject to appraisal and according to what it might be appraised. A basic distinction is made between positive and negative reactions to a situation. Apart from that, the object of the appraisal can either be an event, an action (by somebody) or an object itself. Events might be relevant to an individual's goals, to the standards it tries to uphold, or to its tastes. Table 2.4 shows the types of reactions an appraisal can result in.

	Positive Reactions	*Negative Reactions*	
Event	...because something good happened (joy, happy-for, gloating)	...because something bad happened (distress, sorry-for, envy)	**Goal**
	...about the possibility of something good happening (hope)	...about the possibility of something bad happening (fear)	
	...because a feared bad thing did not happen (relief)	...because a hoped-for good thing did not happen (disappointment, sadness)	
Action	...about a self-initiated praiseworthy act (pride, gratification = pride + joy)	...about a self-initiated blameworthy act (shame, remorse = shame + distress)	**Standard**
	...about an other-initiated praiseworthy act (admiration, gratitude = admiration + joy)	...about an other-initiated blameworthy act (reproach, anger = reproach + distress)	
Object	...because one finds something appealing or attractive (love, liking)	...because one finds someone/-thing unappealing (hate, dislike)	**Taste**

e.g. Victim is relieved that Hero defused the bomb

Table 2.4: Valenced Reactions (from [Ortony, 2003, p.194])

To be able to process its situation in such a manner, the individual needs to have goals, standards and tastes. To make its reactions believable, i.e. consistent, these values have to be relatively stable and coherent. Goals of an individual directly affect its actions, humans are said to have a hierarchy of goals dependent on each other, with long-term higher-level goals on top, and transient subgoals beneath. Standards and norms are responsible for moral value in the individual's world-view, and tastes or preferences give certain values to single important objects of perception. All these perceptions can also be directed inward, perceiving bodily changes or cognitive processes.

The emotion instances given in parentheses in table 2.4 are only examples that fall under the respective category, they are by no means the definitive members of their group: they are just emotion-words used in everyday language to explain reactions that might belong to this category.

Each of the emotion types is associated with a varied group of reactions or behaviours. The appraisal of an event, action, or object as belonging to one of these types triggers an inclination to use one of these behaviours, a response

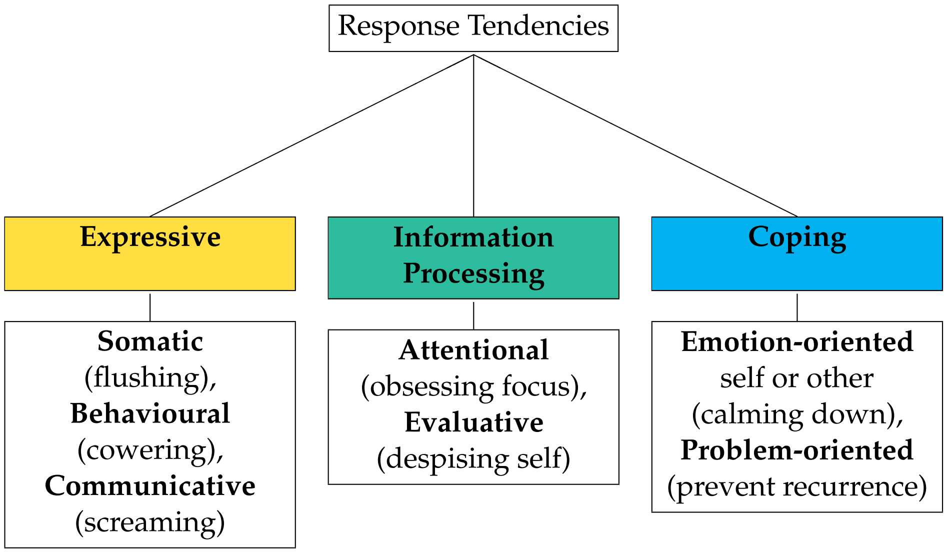

Figure 2.1: Emotion Response Tendencies (from [Ortony, 2003, p.198])

tendency, as Ortony calls it.

There are mostly involuntary expressive reactions on the one side of the spectrum, changes in the style of internal information processing and attention, and coping activities that include complex planned actions on the other end. Figure 2.1 shows the three categories of response tendencies and related subgroups (examples in parentheses).

Frijda defines an emotion as a change of action tendency (e.g. "approach" or "avoidance"), a change in the readiness to alter one's relation to external or internal objects. Ortony's response tendency can therefore be seen as realizing a change of action tendency on the one hand, when they trigger plans and alter internal states, and on the other hand as the result of a change of action tendency, when they deploy expressive reactions.

One theorized purpose of expressive response tendencies is to reveal the elicitation of an emotion directly to others that are watching, at a subconscious level (e.g. [Reisenzein, 2001]).

Effects of information processing can either result in redirecting attention or changing general evaluations about other agents or values that the individual has. Coping is a very wide category (and a prominent term in emotion research): Problem-oriented coping is the most direct approach to solving the cause of an emotional reaction, trying to use a plan that brings the situation under control, thus improving a bad situation or realizing or extending a good one. Emotion-oriented coping is more involved, it is directed either at oneself or another participating individual, by trying to change oneself or someone else, the individual tries to change or stabilize the situation indirectly.

Ortony also elaborates on models of personality as an important part of modeling a consistent and therefore believable character. Several models of personality exist, most of them rely on the characterization of people along a small number of dimensions, e.g. the big five traits [Pervin, 1994].

A proposal for an approach to implement a mechanism in that way is to start with only few parameters as for example the distinction of promotion and prevention focus. This could influence the choice of coping strategies directly. Promotion focus would tend to choose plans that strive for pleasure, whereas prevention would first try to prevent harm.

2.2.3 The Process of Appraisal

It is a big step from a qualitative theory about appraisal to an actual implementation, therefore I will now summarize what features of the emotion process I will try to include in this project, going into further detail. From an engineering standpoint, appraisal is under-specified, a lot of the steps from a perception to an action are unclear. There also must be an underlying model of the perceiving entity, it must be capable of remembering its goals, standards and preferences, must be able to act and perceive, to expect and anticipate.

The basic action-reaction loop starts with the perception of the environment. Every perception passes through appraisal. This might lead to an entry in an Appraisal Registry, which in turn causes several response tendencies to be considered by the Action part of the model. This in turn effects the environment in a hopefully beneficial way.

This is a vast simplification for the purpose of modeling; the processes that supposedly happen in a human agent do not justify a clear-cut distinction between perception and response. It could be described more aptly as a sensorimotor circuit where sensations and actions happen in a close feedback loop. Sensations are actively sought for, caused by the stimulus of ongoing action that needs input. The roles of cause and effect are constantly shifting; sensations are attained only by action and one's own actions are sensed as well. They get their specific interpretation only through the integration into a larger context [Dewey, 1896, Pfeifer and Scheier, 1994].

I assume that the simplification is valid in this case, because in the proposed model actions do have direct access to all current perceptions to actively modify their parameters, and the actions themselves are processed as perceptions. The implementation of the model, however, does not include the concept of an attention focus, all changes in the environment are perceived almost instantly, thus reducing the complexity of sensorimotor coordination on the one hand and on the other greatly depriving the model (for a very detailed proposal of an architecture that incorporates appraisal see [Petta, 2003]).

I will now try to flesh out the steps with a simple specification. The first issue will be the appraisal criteria, the Stimulus Evaluation Checks that each perception passes. Several theories exist, but there seems to be a consensus on at least the criteria in the following list:

- relevance, goal significance, focus
- standards compliance, blameworthiness
- intrinsic pleasantness, valence, appealingness
- novelty, unexpectedness, suddenness, familiarity
- responsibility
- coping potential

[Lazarus, 1991, Smith and Lazarus, 1993] analyze the elicitation of emotions on two levels: on the "molecular" level, appraisals are divided into components corresponding to the dimensions of stimulus evaluation checks. When they are combined on the "molar" level, they are characterized by so called core relational themes, which incidentally bear a striking resemblance to distilled versions of plot elements in the theory of narrative. The core relational theme for jealousy for example is "resenting a third party for loss or threat to another's affection", the one for compassion reads "being moved by another's suffering and wanting to help".

The first three out of the above list of evaluation criteria can be seen as corresponding directly to the three necessary value types in an agent: goals, standards, and preferences. The latter three are applicable to every appraisal.

Novelty modifies the intensity of an appraisal: if the perception it deals with has not been expected, the intensity of the appraisal is higher, as opposed to for example one's own actions.

Responsibility of an action attributes causality to an agent, which is directly usable in directing coping plans or expressive actions, and for changing evaluative processing about this agent.

Coping potential represents the possible degree of control about the situation. It could be higher for perceptions that can easily be answered with several possible coping plans or for those where there is already a suitable plan in execution; when there is no plan available, the coping potential would be low. The value representing the degree of control could also modulate the intensity of the appraisal, as with novelty.

But first we need a basic intensity value for an appraisal. The first three evaluation criteria can be seen as corresponding to the three possible objects of appraisal as differentiated by [Ortony, 2003] (event, action or object). They may be implemented as quantitative evaluations that provide a basis for an intensity value. This, however, only provides an initial account of an appraisal intensity neglecting the dynamics of intensity. The laws of hedonic asymmetry in [Frijda, 1986] point to a fundamental difference in the changes of positive and negative intensities, respectively; furthermore repeated sensation of the same type of perception should lead to a decreasing intensity. In the proposed model the intensity values simply decrease over time when there is no further perception reinforcing the same appraisal. A more detailed account of the intensity values in the implementation of the ActAffAct system will be given in the sections 3.3.2 and 3.4.1.

Relevance, Conformance, Preference First, consider the case where a perception should be matched against the goals present in a system to determine its *relevance*. Let us assume a simple case where there exist goals to either try to prevent a specific action from happening or to try to encourage (or, as a special case, perform) this action. If a perception stating the completion of this action is encountered, a goal encouraging it would lead to a positive relevance for this goal, a prevention goal to a negative one. It would be the other way round with a perception that indicated the failure of an action. And if there are also perceptions that indicate the possibility of an action happening in the near future, i.e. someone is trying to perform the action, all of the goal-relevant appraisals in table 2.4 (page 36) are possible.

It can be argued that determining what counts as failure of an action in the real world can be rather complicated, even more so for the prospect of something happening. In a simulated world, it is easy to work around this problem. It would of course be much more challenging to solve this more realistically. Another point of improvement would be to incorporate pre- and post-conditions of goals into the matching process. A perception that corresponds to a precondition for a goal would also be conducive to this goal, preventing a precondition would be obstructive. But this would mean a more complicated reasoning process involved in appraisal, although it should still be possible without relinquishing the idea of appraisal as a basic low-level function.

Perceptions about the success of an action can be matched directly against an agent's standards to yield standards *conformance*. The standards should attribute a moral value to a specific action or more realistically a group of actions. The value of the *responsibility* check is used to differentiate between one's own actions and those initiated by others. The *preference* for agents or specific objects in the environment can also lead to an appraisal about one's feeling about them. This of course must be preceded by a change of this value.

To cut-down on the number of appraisals that get registered, the calculated values of relevance, conformance, or preference must exceed a certain significance threshold.

Response Tendencies If a threshold is exceeded, an appraisal can have a threefold effect, corresponding to the distinction of response tendencies in figure 2.1. If the appraisal is intense enough, there should be an immediate expressive reaction, revealing it to others. Information processing effects would be changing the values about the relationship to the responsible person or involved objects in proportion to the intensity, this would be the main way of changing preferences for others (cf. [Staller and Petta, 2001]).

The most important effect is the installment of goals for coping activities, which could in turn lead to the execution of a suitable plan and the spawning of new goals.

The underlying model is responsible for the pursuit of goals, scheduling of concurrent plans and sorting out the conflicts between action tendencies.

2.3 Applying Appraisal to Agents

What model of an agent lends itself to the incorporation of appraisal? This section will look at the theory of autonomous agents and discuss agent architectures.

There is no undisputed definition of the term "computational agent" in the scientific field, but [Wooldridge and Jennings, 1995] name four agent characteristics that are significant for agents: reactivity, social ability, autonomy and proactiveness. The incorporation of emotion promises to contribute to all of these characteristics. A response that is both fast and adapted to the situation as interpreted by the agent would even approve on the definition of *reactivity*, and emotional expressions also add to the information that is communicated between agents. This mostly involuntary communication of emotional states is also a contribution to the *social abilities* of agents, as well as the consideration of standards (or social norms) in the elicitation of emotions and the possibility of long-term relationships through the longevity of preference values for other agents. Goals, one of the necessary elements for the process of appraisal, also provide a basis for *autonomy*, and the emotion process as a whole is directed at selecting and steering action. Checking external as well as internal perceptions for their relevance to the agents concerns also leads it to taking the initiative, which is a paraphrase for *pro-activeness*.

Appraisal could lead to added flexibility in the responses of an agent and the responses should be adapted to its environment in a more consistent manner. This is also the most basic requirement for believability. Emotional expressive reactions have, as mentioned above, also the effect of revealing inner states to others, thereby potentially adding a lot to the social relationships between agents. Emotions have an important role in judging the relative importance of things to do next, appraisals with a high intensity improve the probability of associated actions to be executed, and these actions might spawn others, hopefully compatible with the already existing goals of the agent.

For the actual implementation, I chose an existing implementation of an agent model, that already included many of the necessary features: JAM, the Java Agent Model [Huber, 2001b]. JAM belongs to the family of belief-desire-intention (BDI) agent models. The next section will deal with JAM and BDI models in general, I will then detail the modifications that ActAffAct introduces to JAM.

2.3.1 JAM as a BDI-model

JAM is an implementation of an agent model in Java[3].

The Java Agent Model is an Architecture that can be characterized as a hybrid intelligent agent architecture, whose roots lie in the Procedural Reasoning System (PRS, [Georgeff and Ingrand, 1989, Ingrand et al., 1996, Georgeff et al., 1999]), which in turn is based on ideas about resource bounded practical reasoning [Bratman et al., 1988]. It provides its own Plan representation language, meta-level and utility-based reasoning, goal-driven and event-driven behaviour. An agent is defined as a collection of its beliefs (facts it believes to be true), its desires (goals), and its capabilities (specified as plans in a JAM specific language with definable primitive actions at the lowest level). Capabilities scheduled to achieve a goal form the third name-giving component, intentions. Because of the powerful plan language, JAM agents can be very flexible, if possibly at the price of high complexity. There is no special support or model for the processes of perception and action, but any Java code can be used for the purpose. Nor is there special support for inter-agent communication. JAM supports both top-down goal-based as well as bottom-up data-driven reasoning; goals and plans (intentions) are selected for execution based on utility values, meta-level plans can be used to override these utilities.

Research on JAM has also focused on a definition of autonomy in BDI agent architectures. The proposed view on autonomy is as the level of separation between external influences like perceptions about other agents and the internal

[3] It was written by Marcus J. Huber and can be downloaded freely from http://www.marcush.net, the website of his company Intelligent Reasoning Systems, for non-profit use. A manual [Huber, 2001b] and several papers [Huber, 1999b, Huber, 1999a] provide documentation.

structures in the agent, its goals, but also its beliefs, intentions and capabilities ([Huber, 1999a, Huber, 2001a]).

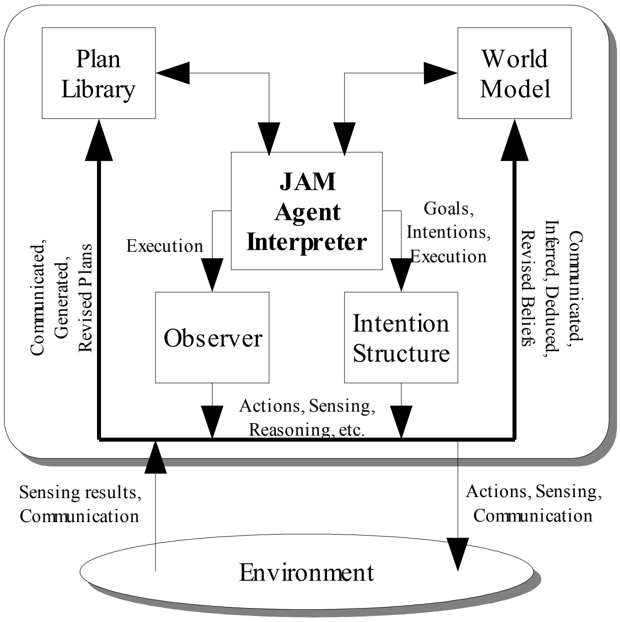

Figure 2.2: The JAM Agent Architecture (from [Huber, 2001b, fig.2])

Figure 2.2 shows the main components of JAM The central part of the Agent is the Interpreter, which encapsulates the cyclic execution model. A cycle consists of first executing the Observer, a simple JAM plan that might be used to update data from outside the agent regularly. After that, an Applicable Plan List (APL) is calculated and the most suitable plan at the moment is chosen (via utilities or meta-level reasoning) and added to the Intention Structure. Finally, the intention with the currently highest utility that is not blocked is executed, which means one step in its plan is taken. This stepwise execution realizes the actual behaviour.

The central knowledge item of JAM is the *Relation*, a propositional data representation with a name and an arbitrary number of arguments. The arguments can be of any type (String, Real or Integer value and Java object references), no type-checking is performed.

```
relation-name  argument1 argument2 ... argumentN ;
```

The World Model is a collection of relations that can be indexed by relation name or queried by relation name and one or more fixed values for the positional arguments from plans. The Intention Structure starts as a list of top-level goals that consist of the goal type (a relation that specifies the desired state), and a utility (that defaults to zero). The goal type can be one of ACHIEVE, PERFORM or MAINTAIN (there is a fourth, QUERY, that is just a synonym for ACHIEVE). The first two differ in the success semantics; the third is not removed upon completion, but checks for its success relation regularly. Intended plans can spawn subgoals or explicitly post new top-level goals. The utility of top-level goals is used to choose intentions for execution.

The Plan Library holds all the capabilities the agent has, they are identified by the goal they are applicable to. Plans can be given utilities, but these are used mainly to prioritize the list of plans that are applicable to one goal. Plans have a precondition, which is checked before a plan can be intended; a context, which must evaluate to true at every cycle of execution; and failure and effect fields, which are executed if the plan fails or succeeds, respectively. Several constructs are provided for the procedural body of a plan, including conditional and looping structures, contributing to the overall flexibility of JAM.

The focus of JAM is on the reasoning capabilities, there are no constraints as to what environment an agent can be embedded in, therefore there is also no special support.

What is necessary to adapt a JAM agent to a conflict-laden environment and how to enhance it with appraisal? The next section deals with the elements to be added to JAM and ways to integrate it with an environment.

2.3.2 ActAffAct's Agents

JAM is a general purpose and very extensible platform which must be tailored for specific use. Most features of JAM are very useful for the ActAffAct agents, the only ability not used at all is meta-reasoning as it is optional, and using utility values proves sufficient for the prioritization of task directed and emotion related behaviour. Several things need to be added, most notably the appraisal step.

Goals are an integral part of JAM that drive the execution cycle. Appraisal also needs standards and preferences of an agent; these can be represented as special relations in the world model of the agent. A separate appraisal component has to be added, that will affect the other components as described below. ActAffAct agents are intended to be used in (pseudo-)parallel multi-threaded environments, JAM however is strictly single-threaded with no thread-safe mechanisms. On the one hand the internals of an agent should somehow be visualized in the multi-threaded interface, on the other hand one dedicated JAM agent is responsible for the environment logic, and as multiple agents are acting on the environment concurrently, the parts that could be interacted with need to be thread-safe.

Perception can be achieved through a simple Java routine called directly from the observer plan. It uses a direct Java object reference, so there is no further complication as through another communication channel. The perceptions themselves are represented directly as relations (i.e., the native knowledge unit of the agent). This constitutes a further (see 2.2.3) drastic simplification compared to the complexity of sensorimotor activities, but it reduces the complications of implementation. Requests for actions are done in the same way, so each ActAffAct agent has a reference to the Environment object it uses for perception and action; the particularities and synchronization issues involved will be described in the implementation section. To facilitate verification and debugging, an agent should be able to easily and cleanly pause execution, a feature missing in JAM.

An idea taken from [Moffat, 1997] is the notion of auto-boredom, the reduction of the utility of a goal if it cannot be processed successfully. So for example if no applicable plan is found, the goal's utility is decreased by a certain percentage and the interpreter skips to the next goal in queue. The main benefit is to decrease the utility of goals that have been posted by a coping activity but that the agent cannot solve with the capabilities available to it. Coping goals are added by the appraisal step described below.

Each perception is interpreted in terms of indicating a request for an action, its success, or its failure, and these interpretations are subjected to appraisal and matched with all goals about agent behaviours and with all standards. Matching is done in the same way as the matching of relations in JAM, with the addition of a simple placeholder mechanism that allows goals and standards to encompass a range of behaviours instead of a specific one by allowing any value for certain arguments.

Relevances and Conformances are calculated by taking into account the goal utility or the “moral” value of the standard respectively. Afterwards, all preferences of the agent are inspected. If a result exceeds the relevant threshold, the appraisal’s effects are executed.

As a direct consequence, an appraisal can lead to a change in the agent’s preferences. Agents that are responsible for positively appraised actions, for example, rise in the esteem of the appraising agent. If the appraisal is intense enough, another immediate effect is the posting of an impulse goal that leads to an expressive action suitable for the type of appraisal, as long as there is no appraisal with even higher intensity, it is guaranteed to be executed, thereby revealing the internal state to the other agents. The third consequence of an appraisal is the posting of a coping goal which can lead to the intention of coping activities, if the agent has a suitable capability for the situation.

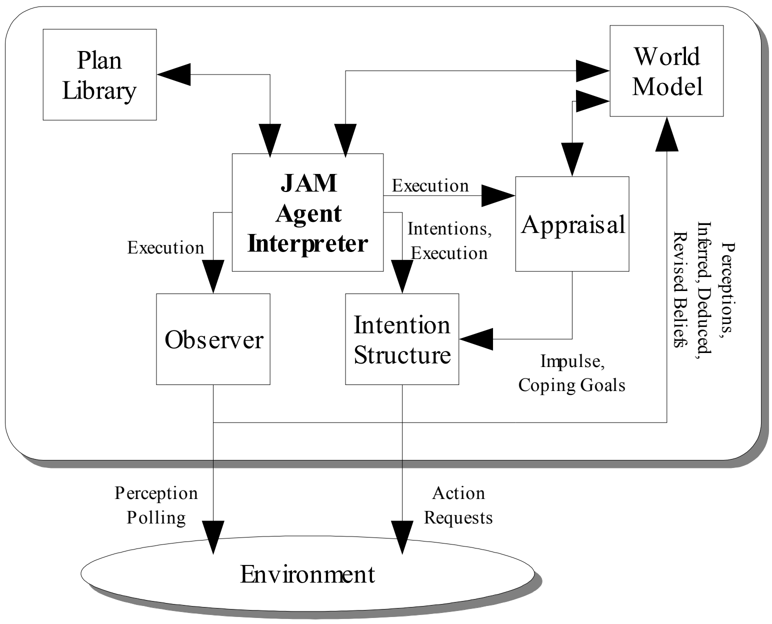

Figure 2.3: The ActAffAct Agent Architecture

The utilities of the coping and impulse goals are proportional to the intensity of the appraisal, and any update to the perception also updates the utility of coping goals. Furthermore, with every cycle all the active appraisals are checked for execution of an associated intention, and if there is none, their recency is decreased and stale ones are removed.

Figure 2.3 gives an impression of the intended differences in components and internal communication compared to the original JAM architecture and the restrictions in flexibility. The actual implementation and its use in the environment of the ActAffAct application will be detailed in the next chapter, including the changes in the execution cycle of the agent.

Chapter 3

...to Practice (Implementation)

This chapter will provide a detailed view of the implementation starting with the application as a whole and ending at the level of the appraisal process. The first section describes the viewer application used to visualize the agents in action, the initial setup of a stage for the agents is described. Apart from the graphical representation the stage also points to the definition of the environment and the agents that inhabit it. The objects that can be found in the environment and the possible actions are described in the following section, including the interaction of the environment and the acting agents. The internals of the four agents are the topic of the third section, changes applied to the JAM architecture and the facts and plans used during the execution are explained. The last section goes into further detail about the appraisal process.

3.1 An AAAViewer for an AAAStage

To view ActAffAct agents in action, requires an initial situation to put them in, and their actions need to be staged.

3.1.1 Actors Need a Stage

A constellation of agents is described in a stage file. XML (eXtensible Markup Language) was chosen as an easy-to-parse format for this simple configuration file. It consists of an entry for the file that contains the visual representation of the environment and one for the definition of the JAM agent that is responsible for the environmental dynamics. There are also a visual file and a definition for each actor[1], and there can be a list of several jam-files that contain common plans or facts for all actors. Furthermore a directory for sounds can be specified and the level of detail at which the viewer application should produce logs can be set.

Sounds are used simply as an indication of an action taking place. The contents of the JAM files and how these JAM agents interact will be discussed below. The graphical representations are specified as SVG (Scalable Vector Graphics[2]) files.

SVG Batik The Batik toolkit is used as the rendering engine for the SVG graphics. It is a Java-based toolkit for display, generation, and manipulation of SVG images developed as a part of the Apache XML project[3].

SVG was chosen as an open way of providing a simple animated display of the interaction between the agents. Although Batik is not (yet) a graphics library that is suited for fast and dynamic animation it provides a flexible environment for simple high-quality 2D graphics that can be manipulated through a DOM interface. The loss in performance can be tolerated for a proof-of-concept implementation. The Batik thread that handles the display is one of the three main threads of the viewer application.

Multi-threading Alongside the Batik thread, a JAM agent handles the environment, and one additional thread is started for each actor. All of these run in parallel, and the only control exerted from outside is starting and possibly pausing them. The actors request the environment to implement actions and poll it for perceptions (which includes the actual outcomes of their own actions), the environment in turn starts and monitors animations executed by the graphics engine (see figure 3.1)[4]. The states involved in taking an action, from request to completion, and the corresponding facts are discussed in section 3.2.

[1] I will use the term actor only for the autonomously acting agents using the appraisal-based architecture.

[2] http://www.w3.org/TR/SVG/

[3] http://xml.apache.org/

[4] The diagrams in this chapter are roughly conforming to the UML (Unified Modeling Language) standard.

The communication between the agents is very simple. Actor agents periodically try to perceive anything new in the Environment, nothing is hidden from them. Everything that is in the World Model of the Environment agent (the state of the environment) is perceived by the Actor agents.

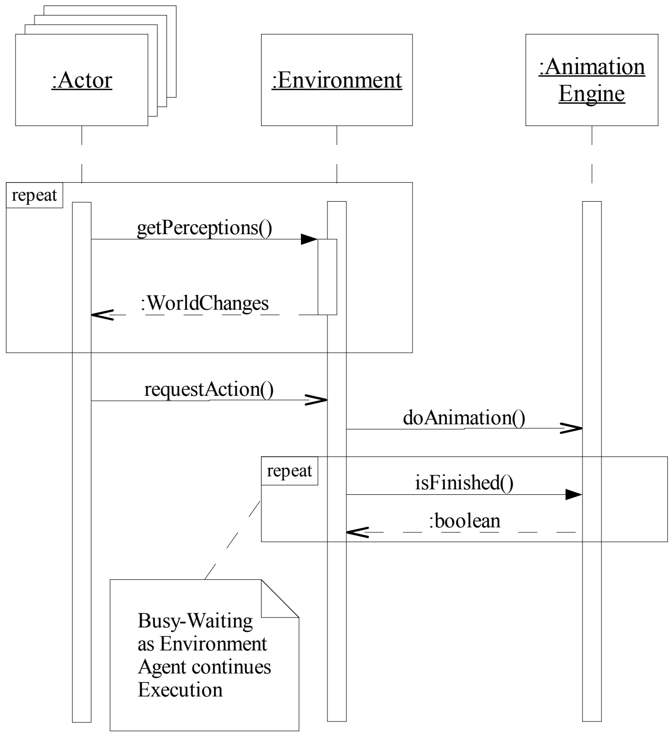

Figure 3.1: Sequence Diagram: Interaction of the Stage Threads

The threads implementing the actual agents are given lower thread priority (actually minimal priority in the coarsely grained Java thread model), so that the already slow SVG graphic updates do not feel even less responsive because of repetitive agent computations.

3.1.2 AAAViewer Application

The simple viewing application is used to show the graphical representation of agent interactions and a little bit of the agent internals. There are tree-like displays for the IntentionStructure of each agent and a display of its WorldModel as a list of JAM facts. Furthermore, the facts constituting the current state of the environment are shown, providing an overview of the current intentions and beliefs of the actors.

The stage created for this project is an extremely simple world inhabited by four actors. It does not try to be a work of art, but only a proof of concept, and its expressive power is therefore intentionally limited, not to say crippled. The setting can be described as four small children in a backyard playing with more or less dangerous toys, which they use to affect each other (I call it the "Bad Boys (and a Girl) in the Backyard" scenario). A snapshot of the stage displayed in the AAAViewer application is shown in figure 3.2.

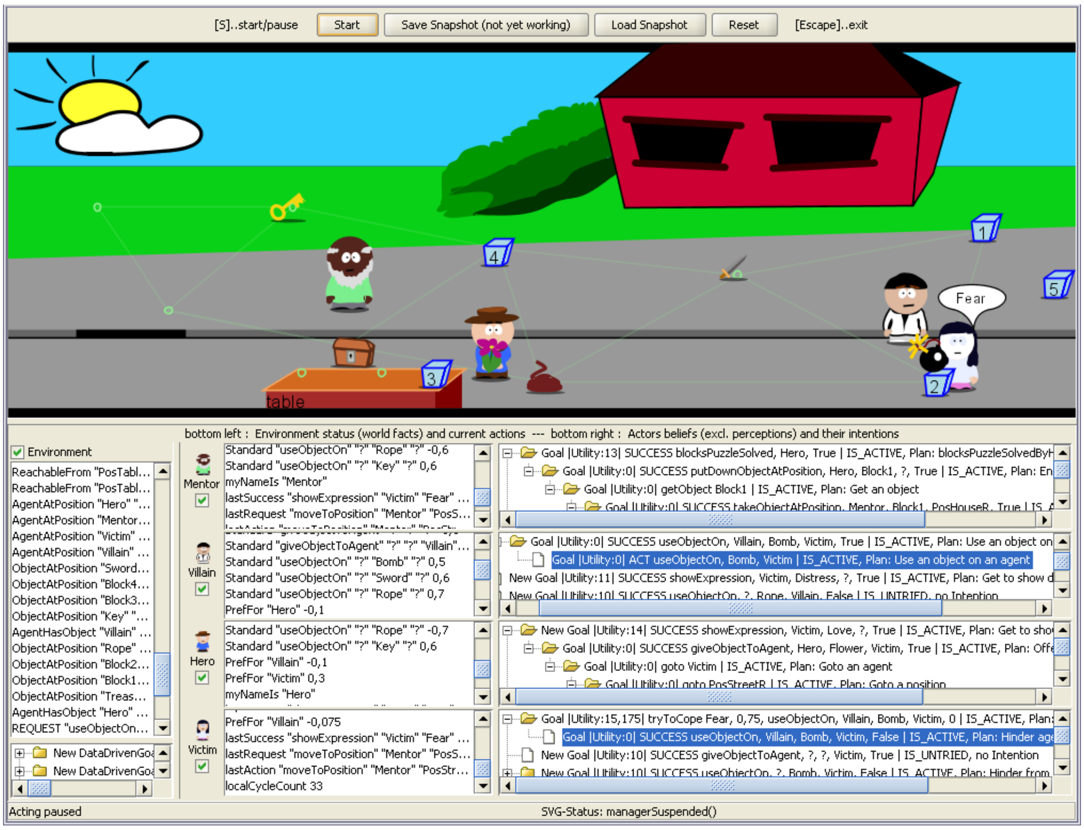

Figure 3.2: The AAAViewer application

3.2 The Environment as Agent

The environment is a simpler JAM-Interpreter than the other agents, it uses mainly CONCLUDE plans, which are triggered by facts that indicate an actor's action request. These plans try to perform the actions for the requester, effecting the necessary changes in the state of the world if possible. The dynamic features of the Environment, its characteristics, are contained in these plans that can use the full arsenal of control structures offered by JAM. The CONCLUDE plans, however, have no PRECONDITION and no CONTEXT, as they should be executed immediately and fail-fast, i.e. if it is impossible to comply to an request, that should be stated promptly. Synchronizing concurrent action requests is the main concern of these plans.

3.2.1 Objects of Interest

> *May you live in interesting times.*
>
> Chinese curse

Apart from the actors, several objects can be manipulated in the environment (cf. Fig. 3.2). Table 3.1 enumerates the possible activities that the objects in the ActAffAct world might be used for.

Bomb	activate or deactivate
Rope	bind someone
Flower	give it as a present
Key	unlock the treasure
Sword	loosen the rope or threaten someone
Block[1-3]	stack them
Treasure	give it as a present

Table 3.1: Objects in the Environment

These objects and activities were chosen with conflict in mind and to provide the possibility of solving a puzzle. For most of the actions it should also be obvious whether they are beneficial of detrimental for the target. Thus, the objects can be used to influence how others think of an actor. The villain of the story would be recognizable as the one who threatens and hinders others. The "good guy" on the other hand might help others solve the puzzle and use items to gain acceptance by others. The main puzzle of this small world is the stacking of blocks in the right order, this quest is however not explicitly posed by any of the actors, it is implicit in the circumstances. Such a problem that needs solving would be part of the beginning of a minimal storyline.

3.2.2 World State

> *I. Die Welt ist alles, was der Fall ist.*
>
> Ludwig Wittgenstein

Apart from encapsulating the dynamic aspects (the “physics”), the environment is a container for facts that describe the current state of the world. It carries out actions requested by the actors, which are also tracked by facts in the world model, including facts about the success, or, if something prevented it, the failure of actions.

static		
IsAgent	*agent* ;	
IsObject	*object* ;	
IsPosition	*position* ;	
IsAgentPosition	*position* ;	
PathTo	*position1 position2* ;	(reflexive)
ReachableFrom	*position1 position2* ;	
dynamic		
AgentAtPosition	*agent position* ;	
ObjectAtPosition	*object position* ;	
ObjectOnObject	*object1 object2* ;	
AgentHasObject	*agent object* ;	
IsOpen/IsActive/IsBound	*object* ;	
action-related		
REQUEST	*action actor [args...]* ;	
SUCCESS	*action actor [args...] True/False* ;	
LOCKED	*[type] actor/object* ;	

Table 3.2: Environment Facts

The facts that are visible to onlooking actors can be divided into static and dynamic information about the environment, the former group comprising facts that name the agents and objects that are present and facts about the (discrete) positions and paths in the environment. The dynamic aspects include the current whereabouts of objects and agents, state-related information about objects, and action-related facts that might be seen as a separate category. Table 3.2 summarizes the possible relations that describe the current state of the environment, i.e. the actor-independent “ontology” of the domain.

Some of the named positions in the environment, those representing a table, are only suitable for objects. Consequentially there is no path to these positions that an actor could walk, but they are reachable from some of the agent's positions. Alterations of objects are indicated by special facts in the case of Treasure, Rope, and Bomb. Mutual exclusion for sensitive plan regions is enforced through the use of the LOCKED fact.

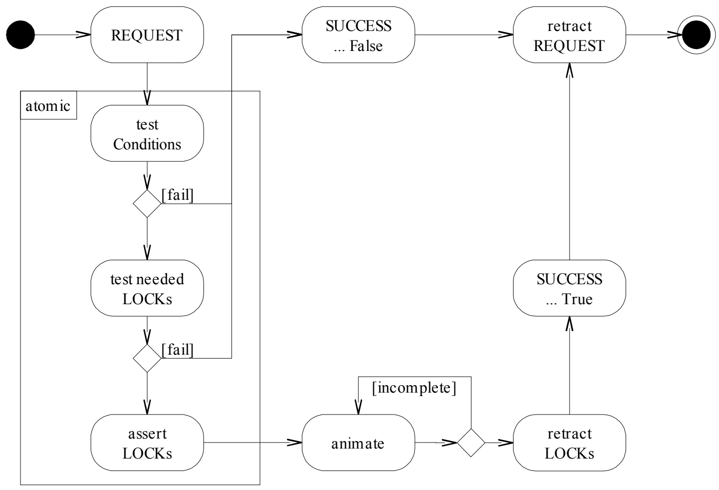

Figure 3.3: Environment activities for an Action REQUEST

Action Dynamics When an agent decides to act on the environment, it issues a REQUEST to the environment which is added as a fact to the environment's world model (and can thus be perceived by all agents). This REQUEST is guaranteed to be followed by a corresponding SUCCESS fact (also perceived by all agents) indicating with a boolean argument, whether the action could be completed. This SUCCESS fact remains part of the world model until the same agent tries exactly the same action again. Figure 3.3 shows as a simple activity diagram how the environment treats an action request. These actions are the building blocks that the actors can use for their behaviour in the environment.

3.2.3 Changing the World

The actions that are possible in the environment are moving around, using objects and moving them around, and, last but not least, expressive actions.

The latter ones substitute all the delicate ways in which people express their emotional state, voluntarily or not, so that others can react to them. Expressive actions are represented graphically mainly (and inadequately) as speech bubbles. Table 3.3 lists the reqests available to the actors. The “useObjectOn” action takes on radically different meanings depending on the arguments.

```
REQUEST  "moveToPosition" actor position ;
REQUEST  "putDownObjectAtPosition" actor object position ;
REQUEST  "takeObjectAtPosition" actor object position ;
REQUEST  "putObjectOntoObject" actor object object2 ;
REQUEST  "takeObjectFromObject" actor object object2 ;
REQUEST  "stealObjectFromAgent" actor object actor2 ;
REQUEST  "giveObjectToAgent" actor object actor2 ;
REQUEST  "useObjectOn" actor object object2/actor2 ;
REQUEST  "showExpression" actor expression ;
```

Table 3.3: Action REQUESTs

Initial Setup Any combinatorially possible “status quo” can be set up for the four actors in this environment. When the stage is loaded, the actors and objects are randomly positioned in the environment, but for the purpose of testing the variability of the agent’s responses in the same situation, specific initial setups are also used.

3.3 Stage Folk

This section describes the different actors that are “on stage”. It also describes the types of internal facts that are used by the agents, and the interplay between facts, goals, and plans, as well as where the appraisal process comes into the equation, which is then described in detail in the next section.

3.3.1 The Actors

The cast consists of four graphically distinct agents that use the same underlying architecture. The choice of names for the four different agents is influenced by the notion of character archetypes. There are no architectural differences, but while one actor is in principle not limited to a particular role in a story, their initial beliefs and goals differ, thus suggesting a particular role.

The Cast In order to achieve minimal storylines that are at least intelligible, the characters have to be discernible not only by their behaviour but also by their appearance. By choosing very abstract names that nevertheless imply a bit about their role in a dramatic interaction, I hope to further the inclination of the viewer to interpret all actions as dramatically significant.

Hero the lead character, the one who should substantially change.

Villain the antagonist, trying to disrupt the hero's action.

Victim the possibly suffering one, in need of help.

Mentor leads the hero's way and lends a helping hand.

The victim is portrayed as female because cliché plots ask for that (no sexism intended, as the result should only caricature a storyline).

There is a thin line between encouraging the viewer to look for drama and avoiding the exploitation of our innate urge to "make sense" of what we see. Humans are very motivated to interpret even unsensible things as coherent, which needs to be considered when the results are evaluated. This project is, however, not intended for a broad audience and thus it can safely be left to the individual viewer to decide whether action sequences are sensible or not.

To get a plot started, at least one of the characters needs a task-oriented motivation to begin with. To foster drama, two characters with opposing ambitions should be present. The goal in setting up the initial situation is seeding it with potential conflicts, which means conflicting interests and goals.

The initial concerns (top-level goals of the intention structure) of the four agents are shown in table 3.4, along with their initial preferences and a small selection of their standards. All of the actors have a permanent goal, their main "concern", that they keep on their agenda throughout their lifetime. This "task-oriented" goal generates behaviours, and its subgoals are used to evaluate the perceptions of the actor as conducive to or obstructive of that goal. The standards are also used as set points for this evaluation and include the implicit goal to uphold the standard. The meaning of the actions mentioned in the standards will be explained in the following section 3.3.2, the question marks act as simple placeholders subsuming all possible values.

The actors are not limited to one concern and a more complex environment would ask for several conflicting concerns in every agent from the beginning. In this toy environment single concerns lead to a much more understandable initial cause of action. Further top-level goals are introduced when the actors express emotions and try to cope with important appraisals, so the possibly fruitful complexity of inner-agent conflicts is not avoided at this level.

To decide about when to do what, the actors need to use their knowledge about the current situation encoded in the facts in their world model.

3.3.2 Actor Knowledge

JAM uses relations to store the data an agent knows about in its personal world model. The ActAffAct actors are concurrently trying to modify the state of a shared environment, and they are monitoring it regularly for changes. But for their reasoning and for the appraisal process, further types of knowledge facts are needed in the agent architecture.

In addition to the facts that are used to maintain the state of the environment and that are also passed as perceptions to the actors (see table 3.2), the actors interpret these perceptions, translating them into facts that seem relevant in their particular context. To accomplish appraisal of actions, actors have beliefs about standards (“moral” values attributed to certain actions) and about their preferences for other objects and agents.

The Facts

The state of the environment is perceived by every actor regularly in its entirety. So all of these facts are part of every actor’s beliefs, and they are marked as perceptions. Although an actor can build up further knowledge through reasoning, there already is a process of actively interpreting perceptions when they reach the actor. This interpretation process should be an evolving subsystem, but in this simple setting a static set of rules is used. The concept of a focus of attention has for now been neglected, but a primitive form is realized by these interpretations.

 Hero	**MAINTAIN** *beLoved :UTILITY 12 ;* (tries to get the most promising candidate to love him) **Standard** *? giveObjectTo Flower ? 0.7 ;* **Standard** *? giveObjectTo ? Hero 0.7 ;* **Standard** *? solveBlocksPuzzle 0.9 ;* **Standard** *? defuseBomb 0.6 ;* **Standard** *? threatenActor ? -0.8 ;* ... **PrefFor** *Victim 0.5 ;* **PrefFor** *Villain -0.1 ;*
 Villain	**MAINTAIN** *beMean :UTILITY 12 ;* (generates subgoals intended to harm promising victims) **Standard** *? giveObjectTo ? ? -0.55 ;* **Standard** *? stealObjectFrom ? ? 0.6 ;* **Standard** *? stealObjectFrom ? Villain -0.9 ;* ... **PrefFor** *Hero -0.25 ;* **PrefFor** *Victim -0.25 ;*
 Victim	**MAINTAIN** *beActive :UTILITY 8 ;* (generates simple random behaviour suitable in her situation) **Standard** *? giveObjectTo Flower Victim 0.8 ;* **Standard** *? giveObjectTo ? Victim 0.7 ;* **Standard** *? solveBlocksPuzzle 0.9 ;* ...
 Mentor	**MAINTAIN** *beHelpful :UTILITY 12 ;* (help the people liked, mainly with solving the blocks puzzle) **Standard** *? solveBlocksPuzzle 0.9 ;* **Standard** *? stealObjectFrom ? Mentor -0.6 ;* ... **PrefFor** *Hero 0.2 ;*

Table 3.4: Initial concerns of the agents

Interpretations

> *There are no facts, only interpretations.*
>
> Friedrich Nietzsche

These rules have been implemented as simple JAM plans that get executed for each asserted or retracted perception. The parser for JAM files has been extended to include a third type of plan: in addition to plans that react to goals and those that react to newly added facts, called CONCLUDE-plans, there now exists a third type that reacts to a change of the perceptions in the agent's WorldModel: PERCEIVE plans. In contrast to the former plans that are executed step-by-step, with a whole cycle in between, these new plans are executed once in their entirety during the observer procedure. This is the main difference to CONCLUDE plans, as the PERCEIVE plans are guaranteed to be executed immediately and fully whenever a perception is asserted or retracted, they are not influenced by plan and goal utilities. Their main use is to translate the perceptions of the agent, that correspond directly to the facts of the environment (see table 3.2), into interpreted facts that are meaningful only to this specific agent. The different types of interpreted facts are listed in table 3.5.

These interpretations represent the immediate surroundings of an agent and its subjective knowledge about intentions and behaviours of others. The more complex case of twice-removed reasoning which involves assumptions about the knowledge and assumptions of other agents about third parties is not used. Many plans that the actors execute use these facts as conditions, but they can also resort to the original perceptions, which is necessary on the lower level of the plan hierarchy.

The interpretations about the intentions and completed behaviours of others are the basis for further appraisal. These facts are matched against goals and standards as described in section 3.4.

Standards A behaviour may conform to standards or may violate them. Of course, behaviours can be completely unrelated to any specific standard. In the evaluation of an action, standards are used to attribute a "moral" value to an action. In theory this can also include a simulation of the possible effects

static
myNameIs *agent* ;

dynamic
myObject *object* ;
ObjectReachable *object* ;
AgentNearMe *agent* ;
AgentNextToMe *agent* ;

action-related
AgentWantsTo *agent behaviour [args...]* ;
AgentDid *agent behaviour [args...]* ;
AgentFailedTo *agent behaviour [args...]* ;

originating from the agent's own plans
IWantTo *behaviour [args...]* ;
IDid *behaviour [args...]* ;
IFailedTo *behaviour [args...]* ;

Table 3.5: Interpreted Facts

a behaviour might have, but in ActAffAct simple pattern matching is used. Only some specific behaviours have an intrinsic negative or positive value for some of the actors, although of course every behaviour can be assigned any value, and they could be dynamically altered through the course of execution. The standards are represented as relations of the following form:

Standard *action args..* *real-value(-1..+1)* ;

Every argument can be substituted by a question mark, which is used as a placeholder to make simple generalizations possible. Some of the actually used standards in the scenario have been shown in table 3.4.

Preferences The preferences of an actor are facts that store a numerical value for every other actor and object it has encountered so far. They might change with every new appraisal of a situation that involves the object or actor, and they significantly influence the choice of an actor's activities.

PrefFor *object/actor real-value(-1..+1)* ;

To change the preferences of other actors about oneself is the main goal of the Hero character and can therefore be seen as a success condition for a simple storyline. If it succeeds to induce a high preference for it in another actor, i.e. if another one “loves” it, its top concern is fulfilled and a significant dramatic change has been completed.

The Plans

The plans that an actor uses to achieve its goals are organized in a hierarchy. The lowest layer contains all those plans that try to act on the environment, i.e., that post action requests. These are used to build action packages which are employed by the behaviour level. This latter level is special as it is also used during appraisal, and its plans use timed control structures. Mediating between the top-level goals or concerns of an agent and the behaviours is the level of activities. In the top layers variability of behaviour and the selection of suitable targets are the main issues, while at the bottom the focus lies in the robust execution of the behaviour that was decided upon.

ACT Plans They correspond directly to the actions that are possible in the environment (see table 3.3), their sole purpose is to translate the required action from a personalized format to the REQUEST that is understood by the environment. All ACT plans use busy-waiting to check the result of their request, i.e., they wait for the corresponding SUCCESS fact.

At this level the plans only use the fact myNameIs and the perceptions IsObject, IsAgent, and IsPosition, that correspond to a naming service for the domain to ensure that the request is sensible.

Action packages Simple ACT plan actions can be combined into action packages. These action package plans do the real work of agent behaviours. To achieve that, they need the actual environmental facts, the agent’s perceptions, to request actions. Action packages use contextual conditions of JAM plans to ensure a fast reaction to changing circumstances (e.g., an attempt to put down an object is only sensible as long as you still own it).

The action packages that are used in the ActAffAct scenario are listed in table 3.6, they translate the desired result into one or several action requests, i.e., ACT plans, or even none at all if the desired effect is already attained.

```
goto  actor/object/position ;
wander   ;
fleeFrom  actor ;
clearHands   ;
getObject  object ;
takeObject  object ;
putDownObject  object ;
putBlockOnTable  object ;
putBlockOnBlock  object1 object2 ;
clearObject  object ;
stealObjectFrom  object actor ;
giveObjectTo  object actor ;
useObjectOn  object thing ;
wait  seconds ;
```

Table 3.6: Actor Action Packages

How can these action packages be connected to the concerns of an actor? Let us reverse the direction of description and start from the top level.

Concerns At the highest point in the hierarchy there are the top-level goals that an actor sets out with and those that get posted during execution. These can be seen as the basic and surface concerns of the actor [Frijda, 1986], and if they are permanent they are realized as MAINTAIN goals, i.e. they are never removed and, if there is no fact directly stating that their goal is achieved, they will always be considered for planning. The starting set of concerns is shown in table 3.4, these are the main objectives of the four actors. These concerns trigger plans that instantiate a whole range of subgoals to satisfy the concern.

These subgoals then are the activities that the agent decides to embark on in order to fulfill his need.

Activities The next lower level consists of PERFORM goals that repeatedly try to further the fulfillment of the actor's concerns (see table 3.7). These activities are used in the concern plans, but they can also be called from coping plans as described below. They decide on a suitable behaviour and the specific target of the actions depending on preferences and the current surroundings of the actor.

```
givePresentsToActor  actor ;
stealStuffFromActor  actor ;
impressActor  actor ;
deterActor  actor ;
helpActor  actor ;
hinderActor  actor ;
giveHintsToActor  actor ;
examineObject  object ;
```

Table 3.7: Actor Activities

Behaviours This level of planning includes goals and corresponding plans that try to implement a certain combination of actions on one's own, but on the other hand an actor can also attempt to help another actor or hinder it from accomplishing something. These three general forms are shown in table 3.8. They are used during the appraisal of interpreted facts that guess the current behaviours of the other actors. These interpretations are stored as the **AgentWantsTo**, **AgentDid** and **AgentFailedTo** facts by perception plans.

```
do  myname behaviour arguments ;
help  actor behaviour arguments ;
hinder  actor behaviour arguments ;
```

Table 3.8: Behaviour Goal Types

The behaviours that actors know about can also be seen as the main parts of a story, which would also include their justifications: either a top-level concern or a coping activity initiated by an appraisal. Plans at this level are often subdivided into **different phases** that influence how easily they can be interrupted or replaced by a different plan. Initially, all the necessary preconditions are evaluated and if the behaviour is possible as far as the actor knows, then it is repeatedly tried over a fixed time-slot. If the plan reaches a certain stage of completeness and only final steps are needed, it can then only be stopped by the repeated failure of the subgoals still pending.

The repeated failure during an earlier stage triggers the reconsideration of the behaviour by the calling activity. Nevertheless, behaviours can be interrupted by more important coping goals that get posted to the intention structure.

A wide variety of plans is needed to yield interesting interactions. The small number of behaviours of ActAffAct actors are summarized in table 3.9.

```
goAwayFrom  actor/object ;
giveObjectTo  object actor ;
stealObjectFrom  object actor ;
bindActor  actor ;
freeActor  actor ;
solveBlocksPuzzle   ;
unlockTreasure   ;
threatenActor  actor ;
defuseBomb   ;
emote  impulse intensity ;
```

Table 3.9: Actor Behaviours

Coping If the appraisal of an interpretation the actor has about another actor calls for a response, a coping goal is posted. The plans that get triggered by these goals have the following general form:

```
tryToCope  type intensity behaviour responsible arg1 arg2 success ;
```

The utility of the goal, that decides whether it is executed or not, depends on the intensity of the appraisal (see section 3.4). The appraisal type can be one of the categories taken from the Ortony2003 model. This coping goal also explicitly includes the reason for the appraisal, i.e., the behaviour and the responsible actor as well as the actual arguments and whether it was successfully completed or not (yet).

The copings actually used in ActAffAct include e.g., the retribution for a behaviour that made someone angry, trying to hinder an actor from achieving what it supposedly wants to do, or cleaning up attempts of hindering if the behaviour the actor wanted to prevent has already happened.

Utility Values for Goals and Plans Task-directed behaviour consists of initial top-level goals: there might also be new top-level goals posted at run-time. As a convention, in ActAffAct these goals have utility values between 10 and 20. Values above 20 are used for immediately showing expressive reactions, values below 10 are achieved by coping goals that were not considered for execution, as these utilities decrease in proportion to the age of the appraised fact.

When there are several plans that are applicable for a certain goal at any level, then plan utilities between -1 and 1 are added to the respective goal utilities before the decision on the currently executing intention is made.

3.3.3 The Agent Execution Cycle

JAM agents use a sequential execution cycle that starts with the observer plan, intended as the place to interact with the outside world regularly. In ActAffAct, the observer is used to do exactly that, every cycle the environment is polled for changes and the response is then translated into agent-internal interpreted facts.

After the completion of the observer, the appraisal process is started, which is detailed in the next section. It evaluates each change in the environment for its relevance to the agent under the current circumstances. Its effects include changes to appraisal relevant facts, like preferences and goals that are posted to the intention structure. The goals are either used to start complex coping plans, or they will trigger an expressive action that shows an emotional reaction.

Then a list of applicable plans is formed, which includes all plans that can be used to achieve a top-level goal or a subgoal that is currently unfulfilled and without an intention. If there are new plans on this Applicable Plan List (APL) the most important one is intended, i.e. the one with the highest utility. The cycle concludes by executing one step in the current top-level plan of highest utility or its current sub-plan. Figure 3.4 gives a simple activity diagram for this cycle.

3.4 Appraisal

In every cycle, new facts in the world model are evaluated to determine their significance for the actor. To check whether a new fact is relevant, it is compared to the possibly matching goals and standards; the numerical values "relevance" and "conformance" are calculated and compared against thresholds. If the thresholds are exceeded, an object is created for this appraisal that immediately imposes its effects on the world model and the intention structure

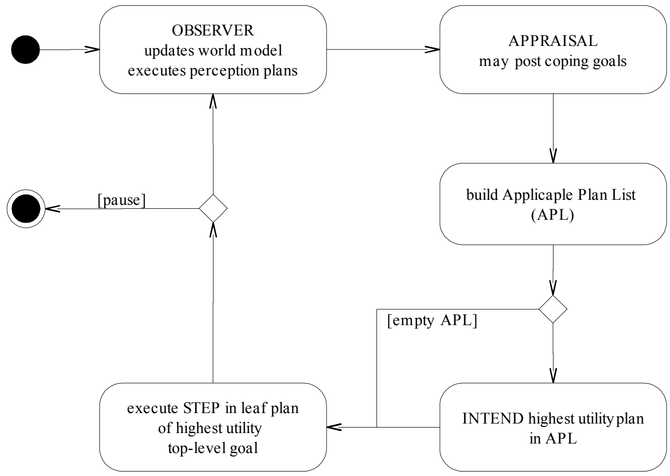

Figure 3.4: Activity Diagram: Agent Execution Cycle

of the agent. It is kept in an appraisal register and will be visited each cycle to update its effects, until it is finally found too old to be of any significance, at which point it is finally removed. If the preference for an object or agent exceeds the corresponding threshold, a similar appraisal object is registered and treated likewise.

3.4.1 Matching

The facts that are considered for appraisal are those that deal with behaviours, they either state that an actor completed a behaviour, that it failed to do so or that it seems to contemplate that behaviour. The facts can therefore be categorized by the success value the fact indicates — attempt, failure or success — and the responsible actor, which can be anyone including the appraising actor.

All the behaviour-related facts are matched with the actor's behaviour goals; matching with standards is only used for successful actions. For each standard, however, there is a virtual goal to uphold this standard, and it is used for the calculation of the relevance of prospects about actions that are relevant to this standard.

Calculation of Relevance, Conformance, Preference How are the numerical values calculated that are needed for algorithmic processing? The part of the relations which indicate the success value are stripped from the facts as well as the type for behaviour goals (do, help or hinder), what remains is the type of behaviour and its arguments including the responsible actor. Bound variables are considered fixed (variables are only present in goal relations), and to allow for crude placeholders, those elements of a relation that contain a question mark only are replaced by variables for this match and can be bound to anything, thus it is possible to construct and match the goal of "preventing anyone from using the sword". If a match between a fact and a goal is found, a non-zero relevance value is calculated as follows (the same procedure is used for matching standard-fact pairs).

The values for relevance, conformance, and preference always range from -1 to 1. To calculate the relevance value of a fact, the utility of the matched goal is used, but as JAM only considers the utilities of top level goals, the intention structure has to be walked to retrieve the significant value. Furthermore, the result needs to be normalized to obtain values in the range of 0 and 1. The same `UTILITY_FACTOR` is used as for the generation of coping goals. The sign of the relevance value is determined by multiplying the success value of the fact and the type of behaviour goal: Prospects and successful behaviours are represented as positive values as opposed to failures, the same holds true for do and help goals as opposed to hindering. For standards, the moral value of the standard is used as conformance measure. To ensure that an appraised fact is at least more important than the goal it matched, a corrective value (0.25) is added, and though not necessary, the same is done for matched standards for consistency.

Relevance = `(successValue * goalType * normalizedGoalUtility)`
`+ CORRECTION`
Conformance = `moralValue + CORRECTION`
Preference = `accumulated preference changes`

3.4.2 Appraisal Objects and their Effects

> *If the facts don't fit the theory, change the facts.*
>
> Albert Einstein

If the value for an appraisal — its relevance, conformance or a preference — exceeds the corresponding threshold, an appraisal object is created and registered. The types of appraisals are labeled following Ortony's categories of valenced reactions (see table 2.4).

The intensity of the appraisal is calculated by combining its value so far with the recency of the appraised fact and a crude approximation of the concept of coping potential, i.e., if there is already a plan in effect that tries to cope with an identical appraised fact, the current one will receive a lower intensity. The recency of any appraised fact will initially be 1, indicating a new fact, but as cycles pass, each update of the appraisal will lower the intensity value according to the age of the fact.

This intensity value is used for the effects of the appraisal. First, another threshold determines whether the intensity is high enough for an expressive reaction. If so, a goal is posted to the intention structure to show an impulse, with all the details of the appraisal to modulate the expression. In ActAffAct, however, this always results in a speech bubble with a font-size proportional to the intensity. This impulse goal is only posted if the agent is not currently showing an impulse with a higher intensity, as the actors cannot mix the display of concurrent emotions.

Second, the preference for the responsible actor is changed in proportion to the intensity and according to the type of the appraisal.

Third, a coping goal is posted with a utility proportional to the intensity, which will possibly lead to a behaviour that is better suited for the current circumstances of the actor.

The goals that are used for coping and showing impulses are subjected to the normal planning procedure of JAM agents and are therefore accessible to each other, enabling an actor to suppress the display of emotions or to fake it, but only if there is a "conscious" effort to do so.

Tweaking the appraisal process Several constants are used during the appraisal that influence the behaviour of actors to a great extent. The most obvious are thresholds that decide whether the intensity of an appraisal is high enough to consider it for stipulating actions. Too low an intensity dooms an appraisal into oblivion, nevertheless, even if it is considered, only the most important ones get to say anything in choosing further actions. The following list gives a short overview of the factors involved in the calculation of intensities and the corresponding utilities for appraisals.

- thresholds for the absolute values of relevance, conformance, and preference.
- a corrective value that determines how much more important coping activities are compared to task-directed behaviour.
- the age of an old world model relation in cycles determines how fast a fact and a corresponding coping activity decay.
- the impulse threshold, intensities below this value do not result in an expressive reaction, but can still instigate coping activities.
- each type of appraisal defines a factor for the change of the preference it affects.

3.4.3 A Comprehensive Example

As an example of appraisal and coping, suppose that the Hero character decided to give a flower to the Victim in order to be liked by her. The Hero's concern caused him to fetch the flower and he moves towards the Victim. As she perceives that, she assumes that he wants to do just that, and therefore an appraisal of the type "hope" is created. That happens because of the Victim's goal to receive presents, which is implicit in her standards. The Victim expresses her hope, and her preference for the Hero rises. If he should fail to meet her expectations, her preference would be decreased again by the ensuing sadness.

Furthermore, a coping goal is created in the Victim that leads to the activity of "waiting for the present", i.e. putting down any other objects and standing still for some time for the present to be received: Actions that should help the Hero in completing the behaviour that the Victim thinks he is trying.

Hero moves towards Victim holding the flower

Victim
1. assumes **`AgentWantsTo`** *`Hero giveObjectTo Flower Victim`* `;`
2. matches **`help`** *`? giveObjectTo Flower Victim`* `;`
3. expresses Hope and starts coping by clearing her hands and waiting

If the Hero succeeds in giving the flower to the Victim, appraisals of joy and admiration will follow, because he completed what the Victim expected from him and his actions conformed to her (admittedly strange) standard.

Hero gives flower to Victim

Victim
1. assumes **`AgentDid`** *`Hero giveObjectTo Flower Victim`* `;`
2. matches **`help`** *`Hero giveObjectTo Flower Victim`* `;`
3. matches **`Standard`** *`? giveObjectTo Flower Victim 0.8`* `;`
4. expresses Joy (Admiration is weaker)

If the Hero's intentions were in fact different ones and he would continue to do anything else except giving the flower to the Victim, she would "feel" sad about that and possibly try to cope with it by avoiding the Hero for some time.

Hero goes away

Victim
1. assumes **`AgentFailedTo`** *`Hero giveObjectTo Flower Victim`* `;`
2. matches **`help`** *`Hero giveObjectTo Flower Victim`* `;`
3. expresses Sadness

The next chapter will present episodes including possible forking paths because of different appraisals and coping activities.

Chapter 4

Results

This chapter presents episodes that were observed in the working ActAffAct system. In order to give an impression of the differences resulting from several initial setups of the environment, the first section presents the features of simple episodes that exclude the Villain, who is the main source of conflict. When this character is excluded, the action that takes place loses a lot of potential. Furthermore, it is compared to the sequences that the full cast produces.

The second section contains details of episodes and their relevance for dramatic structures. The question is, whether the resulting sequences can be seen as sensible mini-dramas without using a too impoverished idea of dramatic interaction. Obvious points of possible enhancements are mentioned.

All of the following descriptions are only qualitative, there was no quantitative analysis of the action sequences produced in different runs of the system, as the intention is only to proof the possibility of “emerging” drama. As I will point out at the end of this chapter, more practical work will be needed, as well as a principled approach towards formalizing actions in an environment.

4.1 Varying Initial Setup

The premise of this project is that a cast of autonomously acting characters whose behaviour is driven by emotional considerations will lead to dramatically interesting interactions if there is enough potential for conflict in the initial setup. An obvious question you can ask is what happens to these characters if there is no conflict in their world. As there are no conflicting initial concerns in any one of the agents, this can be accomplished easily in the ActAffAct setup by simply leaving the Villain character out.

The normal setup for ActAffAct is to use all four agents and position them and the objects randomly on the stage.

4.1.1 No Antagonist

If there is nobody that opposes the goals of the Hero, he just goes about his business and finishes his tasks, which consist of getting someone to love him. This normally involves giving the flower to the actor he likes most and trying to impress same actor, for example by solving the blocks puzzle. The latter behaviour is also sometimes triggered by the Mentor who tries to help the Hero by handing over the necessary blocks in the right order to him.

When nobody tries to hinder the Hero from achieving any of these goals, problems only arise from random actions the Victim wants to do, as for example when she tries to pick up a block that the Hero currently needs.

Misunderstandings arise quite frequently because the actors have only positive or neutral preferences for each other that only rise during conflict-free interactions. A high preference for another actor makes it more likely to assume that he wants to give his object as a present, these assumptions are not that frequent when other appraisals of higher importance have to be dealt with. If an expected present is not received, sadness is expressed.

Because of the intense "cooperation" of Mentor and Hero on the solving of the blocks puzzle it often happens that they interact so much in a positive way that the Hero starts to express Love for the Mentor (remember that Love is only a label for the group of affections that express a strong preference for another actor).

A typical sequence for a setup without the Villain could be as follows.

- Hero wants to give the flower to Victim, he picks it up.
- Meanwhile Mentor gets Block2. He wants to give it to Hero in order to stipulate the solving of the blocks puzzle. Block1 is already on the table.
- Hero interprets the approach of Mentor holding Block2 as a possible present, getting a block from somebody is appraised as positive, because

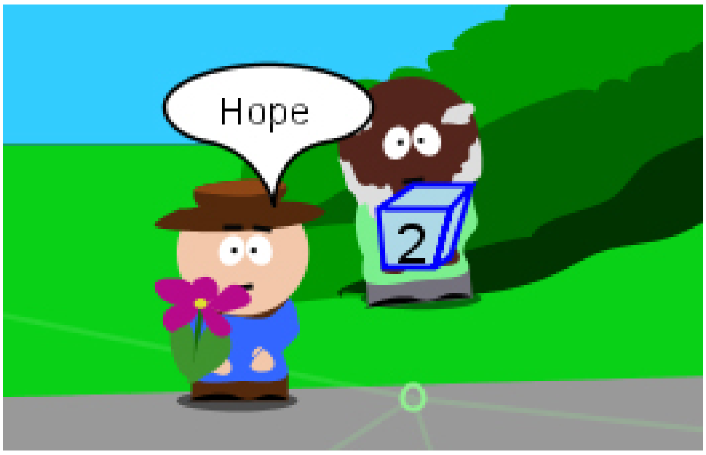

Figure 4.1: Hero expects to get Block2 from Mentor

of his

```
Standard  ? giveObjectTo ? Hero 0.7 ;
```

Because this is more important than his task-oriented behaviour he starts coping with this Hope (see figure 4.1) by clearing his hands and awaiting the present.

Figure 4.2: Hero admires Mentor for giving him Block2

- Mentor succeeds in giving the block to Hero, the latter admires Mentor for his selfless deed (see figure 4.2).
- Hero now decides to start solving the blocks puzzle and puts the block he received onto Block1 on the table.

- In case that the Victim also watches this, i.e., she is near the table, she expresses Hope that the Hero might solve the puzzle. If she was near enough to watch the Hero picking up the flower earlier, she might also be sad about not receiving it as a present.

This continues until the blocks puzzle is solved and Mentor rewards Hero by giving him the key which Hero uses to unlock the treasure. He can then use this item as a present for Victim, which will change her preference for him considerably to the better. This and the successful presenting of the flower, possibly more than once, might even make her love him. It is an easy simulated world after all.

The most prominent expressions of emotion in an episode with this small cast are hope, admiration, pride, and joy. The only negatively valued appraisal is the sadness that results from an unfulfilled — i.e. false — hope. This spectrum is enriched substantially when we introduce the antagonist to the cast.

4.1.2 Full Cast

When the full cast of Mentor, Hero, Victim, and Villain is used, the interaction gets considerably richer and the appropriateness of the sequences for drama can be discussed. With an opposing force, expressions of fear and relief or anger are more common and influence the plans that the actors use. The results depend largely on the sophistication of the plans used for coping. These plans can be tweaked to get better suited responses from actors.

The aim of ActAffAct is to provide a template for organizing plans in intelligent actors that leads to suitable responses combined with the appraisal method of evaluating the actors surroundings. Using behaviour goals and plans that explicitly state the intent of favouring or hindering the actions of other actors directly provides goals that can be used sensibly during coping. This structure is beneficial for the evaluation during appraisal and also for the implementation of coping activities.

When the initial situation is set up randomly, the task-directed behaviour of the actors is also influenced by objects that they can reach at the beginning of planning. This is the first source of variation that does not rely on appraisal.

Successful episode

- Hero wants to give the flower to Victim, he does that and her preference for him rises.
- Meanwhile Mentor gets Block1.
- Villain wanted to use the bomb to threaten Victim, but as he sees that Hero gave her the flower he feels anger and therefore decides to steal the present.
- Villain puts the bomb down again and goes towards Victim.
- While Villain succeeds in stealing the flower, Hero hopes for the currently needed block from Mentor.

Figure 4.3: Villain steals flower, Hero awaits a block

- After that, Villain expresses joy about the success of his devious plan of retribution, all others could watch it and are overwhelmed by anger.
- Victim still feels the need to cope with her fear and flees from Villain (i.e., the intensity of the appraisal is still higher than the utility of her task-directed behaviour).
- Villain on the other hand, after his joy and pride are gone, intends to steal the block that was exchanged between Mentor and Hero.

- Hero now feels fear at the prospect of Villain approaching while he himself holds a valuable object in his hands. He therefore copes with this fear by removing himself from the source of the emotion, i.e., he flees from Villain.
- Villain fails to steal the block, feels sad about it and decides to threaten the Hero with the bomb
- After Hero escaped, he uses the block he got to start solving the blocks puzzle.
- Shortly after Hero puts the block on the table — thus completing the first step of the blocks puzzle — Villain threatens him with the bomb.
- Hero is frightened and runs away, but Mentor and Victim were nearby and saw what happened. Both try to defuse the bomb and Mentor succeeds.
- The result is a stronger "bond of preference" between Hero, Victim and Mentor.
- ...

Villain is forced to use other methods like deterring the Hero with the sword, but most of the time his obstacles are overcome. Finally Hero solves the blocks puzzle and can give the treasure to Victim. In the end Victim expresses love for Hero and often for the helping Mentor too.

Variations Assuming that the Villain succeeded in stealing the block from Hero, he then decides to focus on the Victim and to bind her with the rope. This would lead to the Hero and the Mentor trying to free the Victim again, a very positive deed, that would increase the preference that Victim has for them even faster, leading to an earlier feeling of love. In this case the solving of the blocks puzzle would not be needed, as the main problem would be to free the bound Victim.

In both cases the motivations of the actors are visible in the succession of their planning activities, and although they are inadequately expressed in

ActAffAct, the knowledge could be used to provide the audience with the necessary elements for emotional participation.

The end of this account of an episode is of course a little bit idealized. Because of the limited repertoire of actions that the actors can use, a lot of repetition is encountered until the episode comes to an "ending". This can easily be solved by simply providing a larger set to chose from along with the tendency to use new methods if available.

More important problems arise through the inability to combine simultaneous emotions. Several emotions are manifested at once and they are also expressed, but only the most important one will be coped with, without the knowledge about the other appraisals that happened at almost the same time.

This problem is one candidate for meta-level planning, but several other types of meta information could be included in the ActAffAct planner to enhance the believability of action sequences. Timing of plans and a count of their success rate could be used to implement a simple learning scheme. This and a controllable bias of the actor towards either exploration of new behaviours or exploitation of already known successful ones could yield substantially better plan selection.

Meta-level plans can also be used to replace the crude scheme of utilities, but more explicit meta-data for plans would be of great use. That would include relations between plans like "plan A facilitates plan B" or "plan A prohibits plan B", the same kind of information that can be used for appraisal. If the pre-appraisal step of interpretation leads to the assumption that another actor is executing plan A, one's own knowledge about the relations between plans can be used to determine whether that would be conducive or obstructive for one's own plans.

This is a natural way to make the simple scheme of plans for helping or hindering others more explicit that would also provide valuable information for normal task-directed planning.

But although the methods used in ActAffAct leave much to be desired, it is still possible to see that the use of the appraisal process is beneficial for the implementation of intelligent actors. In short action sequences, the causes for actions and reactions are understandable (i.e. in the terms of the simulated world), and though there were no explicit precautions to ensure a coherent long-term planning of the actors, the effects of appraisal and coping lead the actors to a sensible succession of goals and plans.

4.2 “Story-ness” Revisited

Let us recapitulate which aspects of the requirements for drama are dealt with in ActAffAct and what is needed on the road towards engaging interactive drama.

ActAffAct addresses the initial need for conflict to start the sequences towards the resolving it. Only two actors with conflicting goals suffice to produce a wide range of emotional reactions and diverse reactions. Appraisal and coping provide the actors with causes for their actions, a substantial feature for a plot, and the long-term effects help to achieve a sensible order in the sequence of actions.

Patterns of dramatic functions [Propp, 1968] can be discerned that arise naturally like for example the “violation” of a standard that is followed by the “punishment” of the wrong-doer.

To accomplish his tasks, the Hero is assisted by the Mentor who gives him the necessary objects, not only a correspondence with Propp’s functions, but also with the role of the mentor in the monomyth [Vogler, 1996].

Interpretations like this of the simple actions that are possible in ActAffAct are probably too ambitious, but nevertheless the bottom-line of this project is a positive one. The first signs of the necessary components for believable drama are visible and the methods used are conducive to the efficient planning of emotionally justified actions.

Chapter 5

Conclusion

This last chapter summarizes the experiences gained in this project and, above all, tries to list some of the many possible enhancements that the goal of interactive drama could benefit from.

This project was started to get closer to the goal of dramatically and emotionally believable software agents. Starting from notions in the theory of drama it was assumed that characters drive the plot of a story: If their actions are caused by a simulation of the emotional process, this should be a good basis for construction a dramatic structure. Other preconditions for such a feat would be the potential for conflict in their environment and their initial goals and a diverse repertoire of actions.

In the setting of a very limited simulated world, actors were constructed that, while they are not portrayed in a very believable way, most of the time choose actions in a sequence that can be seen as dramatically appropriate. To achieve this, they are driven mainly by activities that are targeted at coping with the appraisal of a recent perception of their surroundings, thus providing the connected causes or motivations that turn sequences of separate actions into a plot.

To simplify the task at hand a lot of relevant issues were excluded from the project or dealt with only rudimentarily. There is no focus of attention of a single actor at an early stage of perception, as would be needed for a realistic simulation of an actor that is situated in its environment, even if the environment is a virtual one. In the process of interpreting the all-knowing perceptions, the actor selects those that are more important because of his environmental situation. More far-reaching would be an implementation that needs an active dedication of resources for the process of perceiving, simulating in more detail the complexities of sensorimotor coordination.

Apart from the pre-appraisal step of interpreting an actor's perceptions, these interpretations were subjected to the appraisal process, i.e., evaluated according to an actor's current goals, standards, and preferences. Here another step of simplification was done by limiting the process of appraisal to the level of behaviours specified in an easily comparable format. A lot of work remains to do in order to achieve a principled way of specifying behaviour plans for actors and actions in an environment, so that the necessary matching during appraisal can always be accomplished. Further work will be focused on common domain ontologies for the actors including meta-data about the actions available to them and a format for relations between plans, and between plans and actions, that can be used for task-related planning as well as the process of appraisal. Simple restricted simulations of the interpreted behaviours of others could be used to determine possible effects and their implications for one's own concerns. This could also involve a more explicit representation not only of the behaviours that others pursue but also of their activities and concerns, and if necessary an actor should be capable of extending that to the twice-removed case of another actor's assumptions about a third party.

Meta-information could be used in a more explicit form of meta-level planning including a scheme for improving an actor's choices through learning. The information learned during meta-level planning, like the timing and success properties of plans and their relations, should also be usable for coping activities as for example the repeated failure to deal with an appraisal could lead to the dedication of more resources and the shifting of the focus of attention.

Another aspect that was excluded — although it is certainly one of the biggest issues on the road towards believable interactive drama — is the interaction itself, i.e. the interaction between one or more human users and the system. How and at what level of complexity might such an interactive drama be influenced by the user? The obvious but also relatively complex way would be to let users control one of the actors that participate, but the question remains at what level. They might only alter personality parameters or they might control every detail of an actor's actions. Another fascinating possibility that could open up potential further work would be to restrict the users' control to the objects in the environment. They would try to alter the actors' behaviour simply by changing the environment they are situated in, in order to see in what direction this would lead the dramatic structure.

Generally, it seems to be a good idea to move a lot of complexity into the environment. Specifying the environment with the focus of providing a diverse set of actions with nevertheless clear relations between them, that are available to the actors during reasoning, might render the task of appraisal considerably easier.

The results of the ActAffAct project are far from perfect and a lot needs to be done, but while the idea of drama emerging from the interaction of autonomously acting agents seemed impossible only years ago, it now seems to be feasible, at least on a simple level.

Bibliography

The numbers after each bibliography entry indicate the pages where the citations occur.

[André et al., 1999] André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T. (1999). Integrating models of personality and emotions into lifelike characters. In Paiva, A. and Martinho, C., editors, *Affect in Interactions: Towards a New Generation of Interfaces, Workshop Notes, Annual Conference of the EC I3 Programme*, Siena, Italy. AC'99. 35

[Arnold, 1960] Arnold, M. B. (1960). *Emotion and Personality*, volume I and II. Columbia University Press, New York. 35

[Bartneck, 2002] Bartneck, C. (2002). Integrating the OCC model of emotions in embodied characters. In *Workshop Notes of Virtual Conversational Characters: Applications, Methods, and Research Challenges*, Melbourne, Australia. Human Factors 2002 Conference. 35

[Bates, 1992] Bates, J. (1992). Virtual reality, art and entertainment. *PRESENCE: The Journal of Teleoperators and Virtual Environments, MIT Press*, 1(1):133–138. 11

[Bates, 1994] Bates, J. (1994). The role of emotion in believable agents. Technical Report CMU-CS-94-136, School of Computer Science Carnegie Mellon University. 11

[Bates et al., 1993] Bates, J., Altucher, J., Hauptman, A., Kantrowitz, M., Loyall, A., Murakami, K., Olbrich, P., Popovic, Z., Reilly, W., Sengers, P., Welch, W., Weyrauch, P., and Witkin, A. (1993). Edge of intention. In Linehan, T., editor, *Visual Proceedings*, Computer Graphics Annual Conference Series, pages 113–114, New York. ACM. 11, 35

[Bratman et al., 1988] Bratman, M. E., Israel, D. J., and Pollack, M. E. (1988). Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4(4):349–355. 43

[Burke et al., 2001] Burke, R., Isla, D., Downie, M., Ivanov, Y., and Blumberg, B. (2001). Creaturesmarts: The art and architecture of a virtual brain. In *Proceedings of the Game Developers Conference*, pages 147–166, San Jose, CA. 14

[Campbell, 1968] Campbell, J. (1968). *The Hero with a Thousand Faces*. Number XVII in Bollingen Series. Princeton University Press, second edition. 27, 28

[Cavazza et al., 2002] Cavazza, M., Charles, F., and Mead, S. J. (2002). Character-based interactive storytelling. *IEEE Intelligent Systems*, special issue on AI in Interactive Entertainment:17–24. 21

[Cooper, 1999] Cooper, A. (1999). *The Inmates Are Running The Asylum*. SAMS publishing. 7

[Cowden et al., 2000] Cowden, T. D., LaFever, C., and Viders, S. (2000). *The Complete Writer's Guide to Heroes & Heroines: Sixteen Master Archetypes*. Lone Eagle Publishing Company. 28, 29

[Dalgleish and Power, 1999] Dalgleish, T. and Power, M. J., editors (1999). *Handbook of Cognition and Emotion*. Wiley, Chichester/London/New York. 35

[Davidson et al., 2003] Davidson, R. J., Scherer, K. R., and Goldsmith, H. H., editors (2003). *Handbook of Affective Sciences*. Oxford Press. 35

[Davies and Stone, 1995] Davies, M. and Stone, T., editors (1995). *Folk Psychology*. Blackwell, Oxford. 34

[Dennett, 1987] Dennett, D. C. (1987). *The Intentional Stance*. MIT Press, Cambridge MA. 7

[Dewey, 1896] Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, 3:357–370. reprinted in 1981. 39

[Duchartre, 1966] Duchartre, P. (1966). *The Italian Comedy*. Dover Publications, New York. 29

[Dyer, 1987] Dyer, M. (1987). Emotions and their computations: Three computer models. In *Cognition and Emotion*, volume 1, pages 323–347. 9

[Egri, 1946] Egri, L. (1946). *The Art of Dramatic Writing*. Touchstone Book, New York. 29, 30

[Elliott, 1992] Elliott, C. (1992). *The Affective Reasoner: A Process model of emotions in a multi-agent system*. PhD thesis, Northwestern University, Evanston, Illinois. 9

[Elliott, 1994] Elliott, C. (1994). Research problems in the use of a shallow artificial intelligence model of personality and emotion. In *Proceedings of the 12th National Conference on Artificial Intelligence*, pages 9–15, Cambridge/Menlo Park. AAAI Press/MIT Press. 35

[Elliott, 1997a] Elliott, C. (1997a). Hunting for the holy grail with "emotionally intelligent" virtual actors. *ACM Intelligence*, 1(1). 10

[Elliott, 1997b] Elliott, C. (1997b). I picked up catapia and other stories: A multimodal approach to expressivity for "emotionally intelligent" agents. In Johnson, W. L. and Hayes-Roth, B., editors, *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, pages 451–457, New York. ACM Press. 10

[Elliott et al., 1998] Elliott, C., Brzezinski, J., Sheth, S., and Salvatoriello, R. (1998). Story-morphing in the affective reasoning paradigm: generating stories semi-automatically for use with "emotionally intelligent" multimedia agents. In Sycara, K. P. and Wooldridge, M., editors, *Proceedings of the second international conference on Autonomous agents (Agents98) St.Paul, MN, USA*, pages 181–188, New York. ACM Press. 10

[Elliott et al., 1997] Elliott, C., Lester, J. C., and Rickel, J. (1997). Integrating affective computing into animated tutoring agents. In Andre, E., editor, *Proc. of IJCAI-97 Workshop "Animated Interface Agents: Making Them Intelligent"*, Nagoya, Japan. 35

[Forgas, 2000] Forgas, J. P., editor (2000). *Feeling and Thinking: The Role of Affect in Social Cognition*. Studies in Emotion and Social Cognition Second Series. Cambridge University Press. 35

[Frijda, 1986] Frijda, N. (1986). *The Emotions*. Cambridge University Press, Paris, editions de la maison des sciences de l'homme edition. 35, 40, 63

[Frijda, 1993] Frijda, N., editor (1993). *Appraisal and Beyond: The Issue of Cognitive Determinants of Emotion*, volume 7 of *Special issue of Cognition & Emotion*. Lawrence Erlbaum Associates, Hove, UK. 35

[Frijda and Swagerman, 1987] Frijda, N. and Swagerman, J. (1987). Can computers feel? theory and design of an emotional system. In *Cognition and Emotion*, volume 1, pages 235–258. 9

[Georgeff et al., 1999] Georgeff, M., Pell, B., Pollack, M., Tambe, M., and Wooldridge, M. (1999). The belief-desire-intention model of agency. In Müller, J., Singh, M. P., and Rao, A. S., editors, *Proceedings of the 5th International Workshop on Intelligent Agents V : Agent Theories, Architectures, and Languages (ATAL-98)*, volume 1555, pages 1–10. Springer-Verlag: Heidelberg, Germany. 43

[Georgeff and Ingrand, 1989] Georgeff, M. P. and Ingrand, F. F. (1989). Decision-making in an embedded reasoning system. In *Proc. of the 11th IJ-CAI*, pages 972–978, Detroit, MI. 43

[Goldman, 1993] Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences*, 16:pp.15–28. 34

[Hayes-Roth et al., 1995] Hayes-Roth, B., Brownston, L., and Sincoff, E. (1995). Directed improvisation by computer characters. Technical Report KSL-95-04, Knowledge Systems Laboratory, Department of Computer Science, Stanford University, Stanford, CA. 12

[Hayes-Roth et al., 1997] Hayes-Roth, B., Gent, R., and Huber, D. (1997). Acting in character. In Trappl, R. and Petta, P., editors, *Creating Personalities for Synthetic Actors*, number 1195 in LNAI, pages 92–112. Springer. 12

[Hill et al., 2003] Hill, R. W., Gratch, J., Marsella, S., Rickel, J., Swartout, W., and Traum, D. (2003). Virtual humans in the mission rehearsal exercise system. *KI Journal*, (4/03). 17

[Huber, 1999a] Huber, M. J. (1999a). Considerations for flexible autonomy within BDI intelligent agent architectures. In *Working Notes of the AAAI Spring Symposium on Agents with Adjustable Autonomy*, pages 65–72, Stanford, CA. 43, 44

[Huber, 1999b] Huber, M. J. (1999b). JAM: a BDI-theoretic mobile agent architecture. In *Proceedings of the third annual conference on Autonomous Agents*, pages 236–243. ACM Press. 43

[Huber, 2001a] Huber, M. J. (2001a). Agent autonomy: Social integrity and social independence. Technical Report IRS-01-01, Intelligent Reasoning Systems. 44

[Huber, 2001b] Huber, M. J. (2001b). *JAM Agents in a Nutshell*. Intelligent Reasoning Systems, Oceanside, CA. 3, 43, 44

[Ingrand et al., 1996] Ingrand, F. F., Chatila, R., Alami, R., and Robert, F. (1996). PRS: A high level supervision and control language for autonomous mobile robots. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA'96)*, volume LAAS Report, Minneapolis, MN, USA. 43

[Isla et al., 2001] Isla, D., Burke, R., Downie, M., and Blumberg, B. (2001). A layered brain architecture for synthetic creatures. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence*. Morgan Kaufmann. 14

[Kort et al., 2001] Kort, B., Reilly, R., and Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy building a learning companion. In *International Conference on Advanced Learning Technologies 2001*. 14

[Krenn, 2003] Krenn, B. (2003). The NECA project: Net environments for embodied emotional conversational agents. In *Proc. of Workshop on emotionally rich virtual worlds with emotion synthesis at the 8th International Conference on 3D Web Technology (Web3D)*, St.Malo, France. 35

[Laird, 2001] Laird, J. E. (2001). Using a computer game to develop advanced AI. *Computer*, pages 70–75. 19

[Laird et al., 2002] Laird, J. E., Assanie, M., Bachelor, B., Benninghoff, N., Enam, S., Jones, B., Kerfoot, A., Lauver, C., Magerko, B., Sheiman, J., Stokes, D., and Wallace, S. (2002). A test bed for developing intelligent synthetic characters. In *AAAI Spring Symposium on AI and Interactive Entertainment, March 25-27*. 19

[Lazarus, 1991] Lazarus, R. (1991). *Emotion and Adaptation*. Oxford University Press, London/Oxford/New York. 28, 39

[Lewis and Haviland, 1993] Lewis, M. and Haviland, J., editors (1993). *Handbook of Emotions*. Guilford Press, New York/London. 35

[Machado et al., 2001] Machado, I., Brna, P., and Paiva, A. (2001). Learning by playing - supporting and guiding story-creation activities. In *International Conference on AI in Education*. IOS Press. 17

[Maldonado et al., 1997] Maldonado, H., Picard, A., Doyle, P., and Hayes-Roth, B. (1997). Tigrito: A multi-mode interactive improvisational agent. Technical Report KSL-97-08, Knowledge Systems Laboratory, Department of Computer Science, Stanford University, Stanford, CA. 12

[Manovich, 2001] Manovich, L. (2001). *The Language of New Media*. MIT Press, Cambridge, Mass. 8

[Marsella et al., 2000] Marsella, S. C., Johnson, W. L., and LaBore, C. (2000). Interactive pedagogical drama. In *Proceedings of the 4th International Conference on Autonomous Agents*, pages 301–308. 17

[Martin and Clore, 2001] Martin, L. L. and Clore, G. L., editors (2001). *Theories of Mood and Cognition - A User's Guidebook*. Lawrence Erlbaum Associates, Mahwah, New Jersey. 35

[Martinez and Scheffel, 2002] Martinez, M. and Scheffel, M. (2002). *Einführung in die Erzähltheorie*. C.H.Beck Studium. C.H.Beck, München, third edition. 26

[Martinho and Paiva, 1999] Martinho, C. and Paiva, A. (1999). "underwater love": Building tristao and isolda's personalities. In Wooldridge, M. and Veloso, M., editors, *Artificial Intelligence Today*, pages pp.269–296. Springer, Berlin/Heidelberg/New York/Tokyo. 35

[Mateas and Stern, 2000] Mateas, M. and Stern, A. (2000). Towards integrating plot and character for interactive drama. In *Socially Intelligent Agents: The Human in the Loop*, Menlo Park, CA. AAAI Fall Symposium, AAAI Press. 31

[Mateas and Stern, 2002] Mateas, M. and Stern, A. (2002). Architecture, authorial idioms and early observations of the interactive drama façade. Technical Report CMU-CS-02-198, School of Computer Science, Carnegie Mellon University. 20

[Moffat, 1997] Moffat, D. (1997). Personality parameters and programs. In und P. Petta, R. T., editor, *Creating personalities for synthetic actors*, pages 120–165. Springer, Berlin,. 10, 46

[Moffat and Frijda, 1995] Moffat, D. and Frijda, N. H. (1995). Where there's a Will there's an Agent. In Wooldridge, M. J. and Jennings, N. R., editors, *Intelligent Agents: ECAI-94 Workshop on Agent Theories, Architectures, and Languages*, pages 245–260. Springer, Berlin,. 10

[Murray, 1997] Murray, J. H. (1997). *Hamlet on the Holodeck - The Future of Narrative in Cyberspace*. Free Press, New York. reprint by MIT Press 2000. 8, 31

[Ortony, 2003] Ortony, A. (2003). On making believable emotional agents believable. In Trappl, R. and Petta, P., editors, *Emotions in humans and artifacts*. MIT Press, Cambridge, MA. 3, 4, 10, 35, 36, 37, 40

[Ortony et al., 1988] Ortony, A., Clore, G. L., and Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press. 9, 35

[Paiva et al., 2002] Paiva, A., Andersson, G., Höök, K., Mourão, D., Costa, M., and Martinho, C. (2002). SenToy in FantasyA: Designing an affective sympathetic interface to a computer game. *Personal and Ubiquitous Computing Journal*, 6(5-6). 16

[Paiva et al., 2001] Paiva, A., André, E., Arafa, Y., Botelho, L., Costa, M., Figueiredo, P., Gebhard, P., Höök, K., Mamdani, A., Martinho, C., Mourão, D., Petta, P., Sengers, P., and Vala, M. (2001). SAFIRA - supporting affective interactions in real-time applications. *CAST - Living in mixed realities, Special Issue of netzpannung.org/journal*. 16

[Panayiotopoulos et al., 1999] Panayiotopoulos, T., Katsirelos, G., Vosinakis, S., and Kousidou, S. (1999). An intelligent agent framework in VRML worlds. In Tzafestas, S., editor, *Advances in Intelligent Systems : Concepts, Tools and Applications*, chapter 3, pages 33–43. Kluwer Academic Publishers, Netherlands. 18

[Pervin, 1994] Pervin, L. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, 5(2):103–113. 38

[Petta, 1999] Petta, P. (1999). Principled generation of expressive behaviour in an interactive exhibit. In Velasquez, J. D., editor, *Workshop: "Emotion-Based Agent Architectures" (EBAA'99), Saturday, May 1, 1999*, pages 94–98, Seattle, WA, USA. Third International Conference on Autonomous Agents (Agents'99). 16

[Petta, 2003] Petta, P. (2003). The role of emotions in a tractable architecture for situated cognizers. In Trappl, R. and Petta, P., editors, *Emotions in humans and artifacts*. MIT Press, Cambridge, MA. 16, 39

[Pfeifer and Scheier, 1994] Pfeifer, R. and Scheier, C. (1994). From perception to action: The right direction? In Gaussier, P. and Nicoud, J.-D., editors, *From Perception to Action*, pages 1–11, Los Alamitos, CA. IEEE Computer Society Press. 39

[Pfister, 2001] Pfister, M. (2001). *Das Drama - Theorie und Analyse*. Wilheml Fink Verlag UTB, München, 11. edition. 26

[Polti, 1977] Polti, G. (1977). *The 36 Dramatic Situations (Trente-six Situations Dramatique 1895)*. The Writer, Boston. 28

[Prendinger et al., 2002] Prendinger, H., Descamps, S., and Ishizuka, M. (2002). Scripting affective communication with life-like characters in web-based interaction systems. *Applied Artificial Intelligence*, 16(7–8):519–553. 18

[Prendinger and Ishizuka, 2002] Prendinger, H. and Ishizuka, M. (2002). SCREAM: SCRipting Emotion-based Agent Minds. In *Proceedings First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-02)*. 18

[Propp, 1968] Propp, V. (1968). *Morphology of the Folktale*. University of Texas Press, Austin, second edition. 17, 25, 79

[Reeves and Nass, 1996] Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. The University of Chicago Press. 7

[Reeves, 1991] Reeves, J. F. (1991). Computational morality: A process model of belief conflict and resolution for story understanding. Technical Report UCLA-AI-91-05, UCLA Artificial Intelligence Laboratory. 9

[Reilly and Bates, 1992] Reilly, W. S. and Bates, J. (1992). Building emotional agents. Technical Report CMU-CS-92-143, School of Computer Science Carnegie Mellon University. 11

[Reisenzein, 2001] Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions. In Scherer, K., Schorr, A., and Johnstone, T., editors, *Appraisal Processes in Emotion: Theory, Methods, Research*, pages 187–201. Oxford University Press, Oxford New York. 37

[Rickel and Johnson, 1997] Rickel, J. and Johnson, W. (1997). Steve: An animated pedagogical agent for procedural training in virtual environments. In *Proceedings of Animated Interface Agents: Making Them Intelligent*, pages 71–76. 17

[Rollenhagen and Dalkvist, 1989] Rollenhagen, C. and Dalkvist, J. (1989). Cognitive contents in emotion: A content analysis of retrospective reports of emotional situations. Technical report, Department of Psychology, University of Stockholm. 9

[Rousseau and Hayes-Roth, 1997] Rousseau, D. and Hayes-Roth, B. (1997). Improvisational synthetic actors with flexible personalities. Technical Report KSL-97-10, Knowledge Systems Laboratory, Department of Computer Science, Stanford University, Stanford, CA. 12

[Scherer et al., 2001] Scherer, K., Schorr, A., and Johnstone, T., editors (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press, Oxford/New York. 35

[Scherer, 2001] Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In Scherer, K. R., Schorr, A., and Johnstone, T., editors, *Appraisal Processes in Emotion: Theory, Methods, Research*, pages 92–120. Oxford University Press. 35

[Seger, 1990] Seger, L. (1990). *Von der Figur zum Charakter*. Alexander Verlag, Berlin. translation of Creating Unforgettable Characters. 29

[Sengers et al., 2002] Sengers, P., Liesendahl, R., Magar, W., Seibert, C., Mueller, B., Joachims, T., Geng, W., Martensson, P., and Hook, K. (2002). The enigmatics of affect. In *Proceedings of the Conference on Designing Interactive Systems (DIS)*. 16

[Sloman, 1997] Sloman, A. (1997). What sort of control system is able to have a personality? In Trappl, R. and Petta, P., editors, *Creating Personalities for Synthetic Actors: Towards Autonomous Personality Agents*, Lecture notes in AI, pages 166–208. Springer. 13

[Sloman, 2001] Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive Processing*, 2(1):177–198. 14

[Smith and Lazarus, 1993] Smith, C. and Lazarus, R. (1993). Appraisal components, core relational themes, and the emotions. In Frijda, N. H., editor, *Appraisal and Beyond: The Issue of Cognitive Determinants of Emotion (Cognition and Emotion)*, volume 7, pages 233–270. 39

[Staller and Petta, 1998] Staller, A. and Petta, P. (1998). Towards a tractable appraisal-based architecture for situated cognizers. Technical Report TR-98-24, Austrian Research Institute for Artificial Intelligence, Wien. 16

[Staller and Petta, 2001] Staller, A. and Petta, P. (2001). Introducing emotions into the computational study of social norms: A first evaluation. *Journal of Artificial Societies and Social Simulation*, 4(1). 41

[Tobias, 1993] Tobias, R. B. (1993). *20 Master Plots (and How to Build Them)*. Writer's Digest Books, Cincinnatti, Ohio. 28

[Tomlinson, 1999] Tomlinson, W. M. (1999). Interactivity and emotion through cinematography. Master's thesis, Massachusetts Institute of Technology. 15

[Velásquez, 1998] Velásquez, J. (1998). A computational framework for emotion-based control. In *Proceedings of the Grounding Emotions in Adaptive Systems workshop*, Zurich, Switzerland. SAB 98. 16

[Vogler, 1996] Vogler, C. (1996). *The Writer's Journey: Mythic Structure for Storytellers and Screenwriters*. Boxtree, London. 27, 28, 79

[Vosinakis et al., 1999] Vosinakis, S., Anastassakis, G., and Panayiotopoulos, T. (1999). DIVA: Distributed Intelligent Virtual Agents. In *Virtual Agents 99 workshop on Intelligent Virtual Agents*. University of Salford, UK. 18

[Vosinakis and Panayiotopoulos, 2001] Vosinakis, S. and Panayiotopoulos, T. (2001). SimHuman: A platform for real-time virtual agents with planning capabilities. In Antonio, A. d., Aylett, R., and Ballin, D., editors, *Intelligent*

Virtual Agents, number 2190 in LNAI, pages 210–223. Springer-Verlag, Berlin Heidelberg. 18

[Wardrip-Fruin and Montfort, 2003] Wardrip-Fruin, N. and Montfort, N., editors (2003). *The New Media Reader*. MIT Press, Cambridge and London. 8

[Wooldridge and Jennings, 1995] Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, 10(2):115–152. 42

[Wright et al., 1995] Wright, I. P., Sloman, A., and Beaudoin, L. (1995). Towards a design-based analysis of emotional episodes. Technical Report CSRP-95-11, University of Birmingham, School of Computer Science. 13

Index