

Towards a Society of Affect-driven Agents

A.E. Nicholson, I. Zukerman and C.D. Oliver
School of Computer Science and Software Engineering
Monash University
Clayton, VICTORIA 3168, Australia
phone: +61 3 9905-5225 fax: +61 3 9905-5146
{annn,ingrid,cdo}@csse.monash.edu.au

Abstract

We describe a hybrid agent architecture capable of simulating emotional behaviour. Agents have certain fixed personality traits, which influence emotion levels as well as relationships with other agents. Changes in emotions are modelled through propagation in an emotion network. We extend a standard action representation to include emotional preconditions and effects. Results from a test-bed environment with two sample scenarios show how differences in environmental factors and personality affect the social behaviour and emotions of the agents in a group over time.

Introduction

We describe an architecture that combines deliberative and reactive mechanisms for simulating affect-driven agents which operate in agent communities. The deliberative mechanism uses a simple planner which chooses actions to satisfy physiological and emotional needs. Reactive behaviour occurs when extreme emotions trigger an overriding action.

Much of the recent work on social agents focuses on the development of individual believable agents for interaction with human users (Loyall and Bates, 1997; Webber and Badler, 1995; Rich et al., 1994) or for completing domain-specific tasks (Martin and Firby, 1991). In contrast, the focus of this paper is on the exploration of the long-term effects of environmental conditions and personality traits on the interactions between agents in a group. This focus motivates a representation of actions and speech acts at a higher level of granularity than that of the above systems, and also the design of simpler agents, which engage in complex interactions. Our agent architecture, like that described in (Bates et al., 1992; Loyall and Bates, 1997), supports the modeling of emotions and personalities. However, our architecture is in the modular style of (Webber and Badler, 1995; Rich et al., 1994). Our system further differs from that described in (Bates et al., 1992; Loyall and Bates, 1997) in that we explicitly model relationships between agents, and our agents perform actions and speech-acts that achieve emotional effects in other agents as well as themselves.

Agent Architecture

The architecture for our affect-driven agents is shown in Figure 1. The simple interface with the world consists of the perception of external events (by the Perception module) and the execution of actions (by the Action-Execution module). Perceived events lead to an update of an agent's Memory, which uses a predicate-based representation of the states of

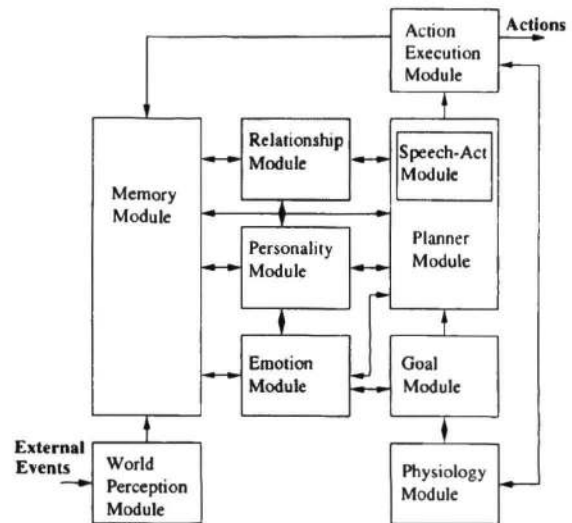


Figure 1: Agent architecture.

the world, e.g., at (x), have (x). The architecture reflects the focus on affect-driven agents with modules for emotions, personality and relationships, which interact with the Planner and Goal modules and the Memory module. We now describe briefly the Physiology, Planning, Goal, Execution and Personality modules, followed by a detailed description of the Emotion, Relationship and Speech-Act modules.

Physiology Module. An agent's physiological state is modelled by a set of physiological needs, such as hunger, tiredness and health. These needs range from 0.0 to 1.0. They increase over time, and are changed by appropriate actions, e.g., an agent eats to reduce its hunger. In the current implementation, a simplified physiological model is used where the needs increase by a constant amount each time step, and appropriate actions reduce them by a constant amount. The health of an agent is a combination of its tiredness and hunger (thirst has not yet been implemented), and can also be directly affected by actions, e.g., it will decrease if the agent is physically hurt (say through fighting another agent); if it goes below a certain value, the agent dies.

Planner and Goal Modules. Agents' goals are mainly basic *maintenance* goals relating to physiological and emotional needs (rather than goals of achievement). They use a simple action representation for their actions and apply a basic backward-chaining planning procedure.

Maintenance goals are prioritized according to a fixed utility based on how much the current need is below a specified

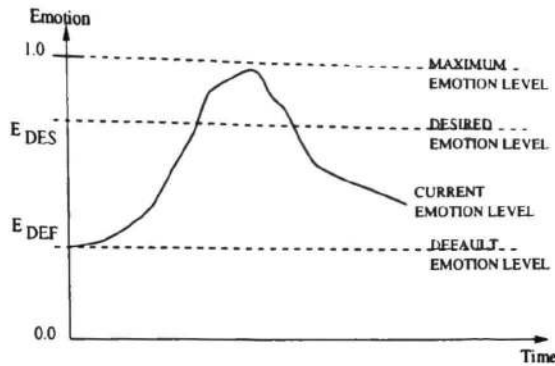


Figure 2: Emotion model.

threshold. Goals of maintenance based on emotional needs (§ *Emotion Module*) are given a lower priority than physiological goals.

Execution Module. This module is given a plan to be executed. This plan may be interrupted by situations that require an immediate response (this is a form of Brooks' subsumption architecture (Brooks, 1986)), e.g., an agent being scared and having to run away. In this case, our simple agents replace the current plan with the action(s) that address the more urgent need. The implementation of a stack architecture, which supports the resumption of the current plan after urgent events have been handled, will be undertaken in the future.

Personality Module. Each agent has a set of personality attributes. We have chosen to represent these attributes by means of the "Big-Five Factor Structure", namely introversion/extroversion, pleasantness or agreeableness, conscientiousness or dependability, emotional stability, and intellect or sophistication (Goldberg, 1990). Numerical values for these factors are in the $[0,1]$ range, with 0.5 representing "normal".

Emotion Module

Emotion Model. The emotion model is based on that proposed in (Nicholson and Dutta, 1997). We model a number of emotion categories (Goleman, 1996), e.g., fear, happiness, sadness, anger and love. The value for each category is in the $[0,1]$ range. For each emotion, an agent has a desired emotion value, E_{DES} , which may be considered a goal of maintenance, and a default emotion value, E_{DEF} . These are functions of the agent's personality profile. At each time step, if there are no external events to influence emotions, the current emotion value will move towards the default value (with a slight random perturbation). We also use an agent's emotional stability (from its personality profile) to determine the degree to which its emotions are affected by an emotional influence (Figure 2 illustrates variations in emotion over time). In the future, we will also take into account the relation between this emotional influence and the agent's pleasantness, e.g., an event which causes anger or a related emotion is likely to have a strong effect on agents who tend to be angry.

Emotion Network. An agent's emotions are represented by a network, with each node E_i representing a separate emotion (Figure 3) (Ortony et al., 1988). Each node has a set of *targets*, $\{T_1, T_2, \dots, T_n\}$, which represent the entities at which that emotion is directed (emotions such as happiness and surprise may be directed at oneself). This set of targets is dy-

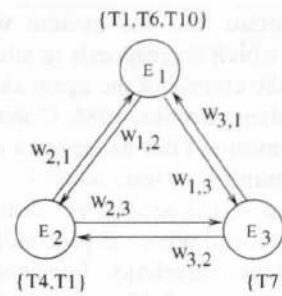


Figure 3: Emotion network with target lists.

namic during a run of the simulation and depends on who the agent has interacted with. Thus, node E_i contains n values, $E_i(T_1), \dots, E_i(T_n)$, which represent the level of emotion associated with the various targets.

Changes in the level of an emotion may be caused by the actions of other agents or the agent itself (details are given later in this section). Emotions are not independent and changes in one emotion may trigger changes in related emotions, e.g., an increase in happiness normally causes a reduction in sadness or anger. To represent how emotions influence one another, and to provide a mechanism for propagating emotional changes, nodes are connected by directed links. Each directed link from node E_i to node E_j has a (positive or negative) weight, $W_{i,j}$, that represents the influence of emotion i on emotion j . We use a form of activation with decay (Anderson, 1983), spreading from the emotion which has changed, $E_i(T_j)$, to propagate emotional change. Activation through links with heavier weights receive priority, and the spreading activation process ceases when an attempt is made to update any node for the second time (i.e., feedback cycles are not allowed).

Emotion Influencing Behaviour. Emotion influences behaviour in two ways: (1) by changing the actions available to the planner to choose from, through the use of emotional preconditions; and (2) by triggering an emotional action response when extreme emotions are experienced.

Prohibiting or enabling particular actions depending on emotions is achieved by adding emotional preconditions to the action description. We call these *E-Preconditions*, to distinguish them from ordinary physical preconditions, which we call *P-Preconditions*. The E-preconditions are satisfied when the level of an emotion exceeds a particular threshold (this is similar to the stimulus conditions described in (Green and Carberry, 1992)). This causes an action to become a candidate to be considered by the planner. The *run-away(x)* action illustrates these preconditions; the P-precondition is that the agent is at location x , while the E-precondition is that the agent's fear exceeds the threshold value 0.7.

Action name: run-away (x)
 P-Preconditions: at (x)
 E-Preconditions: $E_{fear} > 0.7$
 Effects: \neg at (x)

The addition of emotional preconditions means that an agent's emotional levels determine which goals can be satisfied. For example, if eat has the E-precondition $E_{fear} < 0.8$, an agent will be unable to achieve its *decrease-hunger* goal if it is very afraid.

We also incorporate into our system what we call an *emotional trigger*, which corresponds to situations where an agent's emotions take control of the agent and override previously constructed plans (Brooks, 1986; Goleman, 1996). This happens when an emotion i has exceeded a trigger threshold, th_i , for a certain number of time steps T . In this case, the Planner chooses one of the actions which have emotion i in their E-preconditions and whose E-precondition threshold is greater than the trigger threshold. For example, if fear has a trigger value greater than 0.95 and an agent's fear goes to 1.0, then the Planner chooses one of the actions that have fear above threshold 0.95 in their E-preconditions (say *run-away* and *strike*). The rationale for this policy is that if an event is so extreme that it warrants a purely emotional response, then extreme responses, i.e., those which require this high threshold to be activated, are appropriate. Other policies for handling emotional triggers will be considered in the future.

Behaviour Influencing Emotion. The action representation also includes the possible effects of an agent's behaviour on its own emotions and those of others. For example, consider a situation where agent A is angry at agent B ($E_{anger}(A, B) = 0.94$), and insults B as a result of the activation of A 's *insult(x)* action:

Action name: *insult(x)*
 P-Preconditions: *near-to(x)*
 E-Preconditions: $E_{anger}(A, x) > 0.9$
 Effects: $E_{anger}(A, x) \downarrow$, $E_{shame}(x, x) \uparrow$,
 $E_{anger}(x, A) \uparrow$

The *Effects* field of this action description indicates that as a result of the insult, A 's anger at B will decrease ($E_{anger}(A, x) \downarrow$), B 's anger at A will increase ($E_{anger}(x, A) \uparrow$) and B 's shame – directed at itself – will increase ($E_{shame}(x, x) \uparrow$). In the current implementation, the degree to which an action changes an agent's emotional level is a function of the agent's emotional stability only. We envisage extending this to include the relationship between the agent performing the action and the recipient of the action.

Relationship Module

The relationship between agents influences their behaviour in each others' presence. The Relationship module monitors an agent's opinions of other agents. An opinion may be based on personal interactions or on what an agent has heard about other agents. At present, only the former has been implemented.

A relationship value ranges from 0.0 to 1.0. In the current implementation, the relationship between two agents depends on their initial reaction to each other and their interaction history. The initial reaction may be based on a number of factors, including the agents' personality profiles, their current emotional state and the other agent's physical appearance. However, at present only the first of these factors has been implemented in the computation of the initial reaction. This is done by the following formula.

$$R(B, A)^{t_0} = R(A, B)^{t_0} = \frac{1}{N} \sqrt{\sum_i^N (P(i, A) - P(i, B))^2},$$

where t_0 is the first time-step when the relationship is being monitored, $P(i, x)$ is factor i in the personality profile of agent x , and N is the total number of profile factors. This formula is based on the notion that agents with similar personality traits are likely to be positively disposed towards each other. Thus, it implicitly assumes that the initial relationships between agents are symmetrical. This simplifying assumption will be relaxed in the future through the introduction of additional factors in the calculation of a value for the initial reaction. As stated above, the effect of the initial reaction is moderated with the passage of time by the emotional influences which result from the interactions between agents. An agent's relationship with another agent is changed by how this other agent affects its emotions; if agent B performs an action that moves agent A 's emotions towards (or away from) the desired level, A 's relationship with B , $R(A, B)$, will improve (or deteriorate). This improvement (or deterioration) is a function of the difference between agent A 's emotional levels before B performed the action in question and A 's emotional levels after.

Speech-Act Module

Speech-acts are a particular subgroup of an agent's possible actions and are handled by the Speech-act module, within the Planner module. We use a standard speech-act classification (propose, reject, accept, request-clarification, query, answer) plus additional speech-acts which directly affect emotions (share-feeling, insult, scream, praise). These emotion related speech-acts are represented using action descriptions (§ *Emotion Module*).

Recall that the planner chooses actions depending on the current goal priorities. If an agent's physiological needs are currently satisfied, the planner will attempt to satisfy emotional needs. This may be done either through a physical action, e.g., kissing, or through the performance of a speech-act, which at this stage is unspecified. The Speech-act module then determines the actual speech-act, its addressee(s), and when it should be uttered.

Selection of a Speech-act. The speech-acts that are available at each point in time, i.e., those that may be reasonably uttered after other actions or speech-acts, are controlled by a state-transition model (Stein and Thiel, 1993). For example, the *accept* and *reject* speech-acts may be chosen only if an existing proposal is yet to be resolved. In addition, after a *propose* speech-act, likely candidates are *accept*, *reject* and *request-clarification*, but an agent may generate other speech-acts if it is highly compelled to do so, e.g., it will execute a *share-feeling* speech-act, even after a proposal, if its emotional state exceeds the emotional trigger threshold (§ *Emotion Module*). In the current implementation the choice of speech-act depends on factors that pertain to an agent's emotional state, rather than on objective attributes of the situation. For example, an agent is likely to *reject* a particular proposal if its relationship with the agent who made this proposal is negative, regardless of the merit of the proposal.¹ The following factors affect an agent's

¹A module for the evaluation of proposals will be implemented in the future, and its operation integrated with the affect-related modules.

choice of a speech-act among the available ones: (1) the agent's personality profile, (2) the number of agents in the group, (3) the conversational history of the group, (4) the suspected emotional state of the other agents in the group (not yet implemented), (5) the relationships between the agents in the group, and (6) the agent's goals. Normally, only a subset of these factors influences the selection of a particular speech-act. For instance, at present the selection of an insult speech-act is based on an agent's anger level, while the selection of a praise speech-act is a function of the agent's pleasantness. The selection of an accept or reject speech-act depends on an agent's personality profile (whether it is agreeable), its relationship with the agent who uttered the proposal, its emotional state (e.g., how scared it is) and the responses of the other agents in the group so far (e.g., whether most have agreed with the proposal). If upon completion of the speech-act selection process there is more than one possible candidate, e.g., *reject* or *insult*, a random selection is made.

Addressee(s) of a Speech-act. When there are groups of agents, it is difficult to speak to everyone equally. This is modeled by a factor, which we call *conversational attention*, that indicates how much attention an agent is giving to each other agent in the group when generating a speech-act. This attention, which ranges from 0.0 to 1.0, is a combination of the suspected relevance of the speech-act to the other agent and the relationship with that agent; the details of how this attention is determined vary for the different speech-acts and their different instantiations.

When to Perform a Speech-act. Once a speech-act has been chosen and the addressees determined, the Speech-act module decides when the speech-act is to be performed. An agent has a *speech-act activation level*, ranging from 0.0 to 1.0, which is used as part of a trigger mechanism, in the same way as emotions can trigger actions. Each agent also has an activation level threshold, which is fixed based on its personality. A speech-act is triggered when its speech-act activation level exceeds its activation threshold. The speech-act activation level for an agent is a function of the following factors: (1) the number of agents in the conversation, (2) the agent's level of introversion-extroversion, (3) the average relationship value between the agent and all other agents in the group, (4) the time elapsed since the last speech act, and (5) a "relevance" value, between 0.0 and 1.0, which reflects how well the chosen speech-act fits within the current conversation. For example, the speech-act activation level of an introverted agent will be low when a large number of agents is participating in the conversation. In addition, talking to enemies will reduce the speech-act activation level, while talking to friends will increase it. The level of activation also increases with the passage of time, since participating in the conversation becomes more compelling. Finally, the relevance value will be low if the agent has just asked a question or made a proposal, or if the chosen speech-act is now inappropriate given other agents' recent speech-acts. In the current implementation, the relevance value is obtained from a simple lookup table.

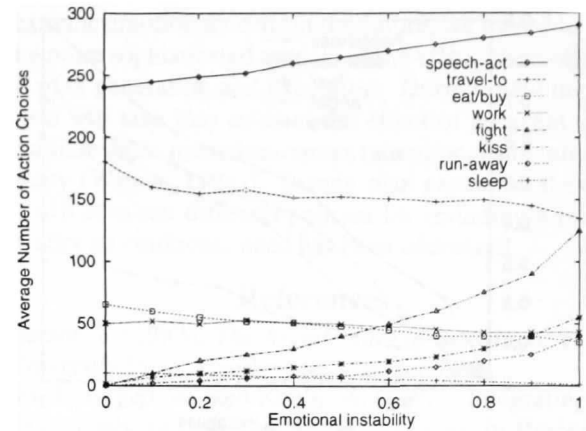


Figure 4: Average choice of actions as the average emotional instability increases (32 agents, after 1000 time steps).

Experimental Results

In the current implementation of the system, agents have the following goals: *decrease-hunger* and *sleep*, which involve reducing the physiological need levels, and *satisfy-emotions*, which involves moving emotions towards their desired levels. Agents have the following actions available to them: *work*, *buy*, *sleep*, *fight*, *kiss*, *run-away*, *travel-to* and *eat*, which are specified in the modified action representation, and a generic *speech-act*, with the exact speech-act determined by the Speech-act module. Two scenarios have been developed to test the system.

Scenario 1. Large Population, General Living. This scenario involves a population of agents (32 for the results given in this section), who are given a home where they can sleep and a workplace where they can work. They know about the location of shops where they can buy food. An agent's physiological needs are satisfied by sleeping or eating. Its emotional needs are satisfied by interacting with other agents. Personality traits, relationships and emotions are modelled as described in previous sections. This scenario is executed over discrete time intervals (1000 time cycles for the results given in this section). Priorities of goals vary over time, as physiological needs and emotions vary. The physiological values used for these experiments are as follows.

Type	Initial value	Threshold	Increment
<i>hunger</i>	0.6	0.8	0.0010
<i>tiredness</i>	0.0	0.8	0.0005

Figure 4 shows how the behaviour of the agents varies, in terms of the number of times the different actions were performed, as a function of the emotional instability personality trait.² As the personalities of the agents are made on average more emotionally unstable (for experimental purposes), the physical actions of eating and working take less of a central role, whereas the number of affect-related actions, such as speech-acts, running away and fighting, increases. Figure 5 shows that the more emotionally unstable an agent is.

²In the current implementation, food can be obtained only by buying it, hence eating and buying are performed the same number of times.

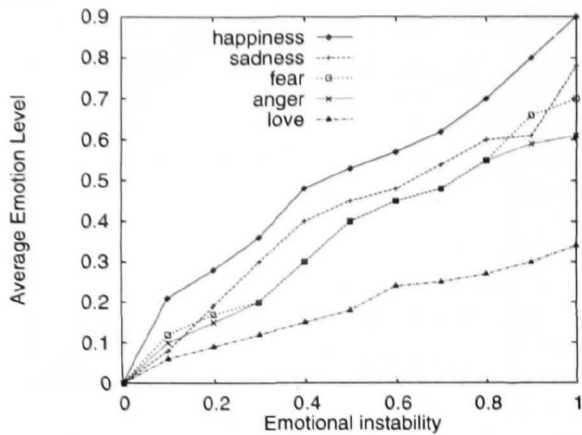


Figure 5: Average emotion levels as the average emotional instability increases (32 agents, after 1000 time steps).

the higher its average emotion level (for happiness, sadness, fear, anger and love). Since the emotion levels of emotionally unstable agents are more volatile, more emotional preconditions are being met or the emotional trigger mechanism is activated as emotional instability rises, thereby affecting agent behaviour. In addition, the more emotional an agent becomes, the more its behaviour affects other agents' emotions, so there is a snow-balling effect. Note that the level of both happiness and sadness rises on average for all agents. This is because for each agent the reduction in one due to an increase in the activation of the other is lower than the increase in activation.

The simulation also shows an increase in the average number of deaths as the emotional instability increases (due to an increase in fighting); and a reduction in the average relationship level as the pleasantness level decreases.

Scenario 2: Small Group, Verbal Interaction. A group of five agents is trying to escape from their place of imprisonment; they have reached a door and found it locked.³ Subsequent physical actions and speech-acts are recorded.

In this experiment all the agents have a medium level (i.e., 0.5) for all five personality traits, and each has 0.75 level relationships with half the other agents (reasonably good friends) and 0.25 with the other half (reasonably bad enemies). With this configuration we vary what we call the *environmental influence* by increasing the values for fear and anger at each time step (simulating negative external events). This constitutes a crude implementation of results from (Lepore et al., 1991), whereby some chronic environmental stressors (e.g., crowding) may increase the impact of social stressors.⁴

In Figure 6 we see that as the negative environmental influences increase, the agents choose more emotionally charged speech-acts; the incidence of *reject*, *insult*, *scream* and *share-feeling* increases, while the incidence of the other speech-acts decreases. The *Average-choice-of-speech-act* axis represents a proportion of the available actions at any time; for example, after a *propose* speech-act, the subsequent speech-acts are primarily *accept* and *reject*, with a few *request-clarification* speech-acts and some

³This scenario was also used in (Nicholson and Dutta, 1997).

⁴In the future we expect to implement a function that links external events with an agent's emotional levels.

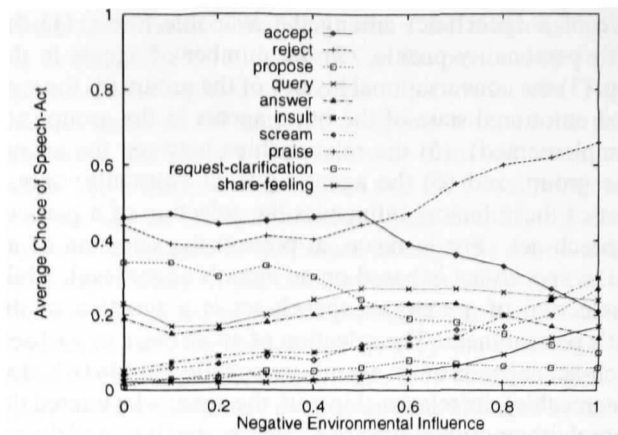


Figure 6: Speech-act choices as negative environmental influences increase (medium level agent archetypes).

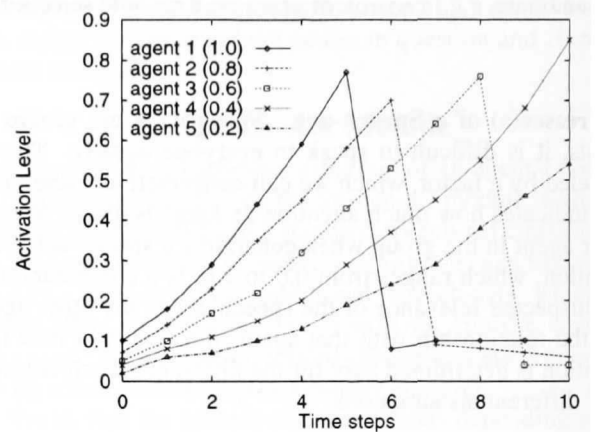


Figure 7: Activation levels of five agents, with extroversion/introversion values of 0.2, 0.4, 0.6, 0.8 and 1.0.

chance of an “interruption” with an unrelated speech-act. The average number of acceptances and rejections exceeds the number of proposals because a single proposal may be accepted or rejected by more than one agent.

Turn taking in conversation emerges from the agents' speech-act activation levels (§ *Speech-Act Module*). Agents with similar personality profiles normally try to activate at similar times, which in turn results in contention for a speaking turn. Figure 7 shows the activation levels over time for five agents who have varying extroversion/introversion values but similar speech-act activation thresholds. Agent 1, which is the most extroverted, breaks its activation threshold first and makes a proposal, after which it waits for a response (indicated by a reduced activation level). The order of the speech-act activations for the remaining agents corresponds, as expected, to their extroversion levels. We also found that the amount of contention for a conversational turn rises as the extroversion levels of the agents rise, and as the number of agents in the conversation increases.

Finally, Figure 8 shows how the average ratio of acceptances and rejections of proposals changes as the average relationship level increases (for an introverted and emotionally unstable group of agents).

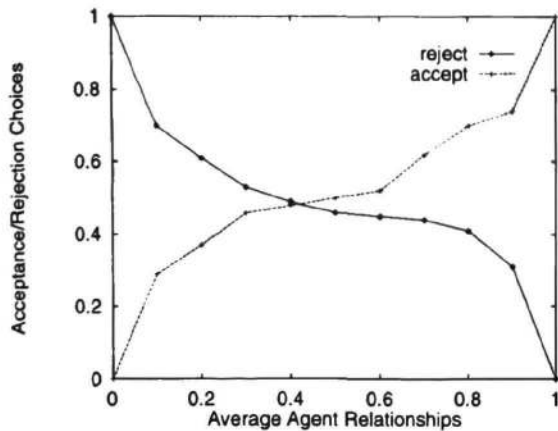


Figure 8: Average acceptances and rejections of proposals as the relationship among agents improves.

Conclusions and Future Work

We have described a modular agent architecture for the simulation of emotional behaviour. This is done by adding emotional preconditions and effects to a standard representation of actions, and using limited spreading-activation to model affective impact. Personality traits and relationships with other agents also influence emotional levels. Results from two sample scenarios illustrate how complex interactions can be achieved with comparatively simple constituent modules. This suggests that our system provides a suitable platform for simulating psychological and social theories.

The behaviour of our agents is influenced by emotions; all actions whose P- and E-preconditions are satisfied become candidates for the planner. Currently, the choice between candidate actions is made randomly. In the future, we plan to select actions probabilistically based on: (1) the extent to which the emotion levels in the E-preconditions exceed the thresholds, and (2) some personality traits. In addition, at present relationships are taken into account in the selection of only certain speech-acts, such as *accept* or *reject*. In the future, we envisage a more general way of incorporating relationships in the action-selection process, e.g., by having relationship levels affect the probability of choosing actions. Unlike the research described in (Bates et al., 1992), where agents simply have attitudes towards objects (whether physical or other agents), our explicit model of relationships between agents will allow more complex relationship models, e.g., the relationship of an agent *A* with another agent *B* may be influenced by their relationships with other agents.

Both the main five personality factors and the emotion categories could be extended to more detailed sub-categories, e.g., anxiety or concern rather than fear (as done in (Bates et al., 1992)). However, the advantages from the resultant additional distinctions in agent behaviour must be weighed against the increase in computational complexity.

At present, agents have a simple model for the effect of emotions on plan generation and execution. Plan generation is currently affected by means of the E-precondition, which modifies the candidate actions available to a planner; plan execution is affected by the emotional trigger "override" feature, which replaces a current plan with a plan that addresses

an extreme emotional need. In the future, we intend to investigate more sophisticated models of the effect of emotions on both plan generation and execution. During planning these models will take into account the effect of an agent's emotional state on its mental resources (attentional and inferential capacity (Walker, 1996)). During plan execution these models will consider different policies for resuming a previous plan after an emotional need has been addressed.

References

- Anderson, J. (1983). *The Architecture of Cognition*. Harvard University Press, Cambridge, Massachusetts.
- Bates, J., Loyall, A., and Reilly, W. (1992). Integrating reactivity, goals, and emotion in a broad agent. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pages 696–701, Bloomington, Indiana.
- Brooks, R. (1986). A robust layered control system for a mobile robot. *Robotics and Automation*, 2:14–23.
- Goldberg, L. R. (1990). An alternative 'Description of Personality': The Big-Five factor structure. *Social Psychology*, 59:1216–1229.
- Goleman, D. (1996). *Emotional Intelligence*. Bloomsbury, London.
- Green, N. and Carberry, S. (1992). Conversational implicatures in indirect replies. In *Proceedings of the Thirtieth Annual Meeting of the Association for Computational Linguistics*, pages 64–71, Newark, Delaware.
- Lepore, S. J., Evans, G. W., and Palsane, M. (1991). Social hassles and psychological health in the context of chronic crowding. *Health and Social Behavior*, 32(4):357–366.
- Loyall, A. B. and Bates, J. (1997). Personality-rich agents that use language. In *Proceedings of the International Conference on Autonomous Agents*, Marina del Rey, CA.
- Martin, C. E. and Firby, R. J. (1991). Generating natural language expectations from a reactive execution system. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, pages 811–815, Chicago, Illinois.
- Nicholson, A. E. and Dutta, A. (1997). Intelligent agents for an interactive multi-media game. In *Proceedings of the International Conference on Computational Intelligence and Multimedia Applications*, pages 76–80.
- Ortony, A., Clore, A., and Collins, G. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, England.
- Rich, C., Waters, R. C., Schabes, Y., Freeman, W. T., Torrance, M. C., Golding, A. R., and Roth, M. (1994). An animated on-line community with artificial agents. *IEEE Multimedia*, 1(4):32–42.
- Stein, A. and Thiel, U. (1993). A conversational model of multimodal interaction in information systems. In *AAAI93 – Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 283–288, Washington, D.C.
- Walker, M. (1996). The effect of resource limits and task complexity on collaborative planning in dialogue. *Artificial Intelligence*, 85:181–243.
- Webber, B. and Badler, N. (1995). Animation through reactions, transition nets and plans. In *Proceedings of the International Workshop on Human Interface Technology*, Aizu, Japan.