Discovering Human Desired Norms of Interaction Improves Social Adeptness in Agents

Henry Hexmoor and Srinivas Battula

Department of Computer Science & Computer Engineering
University of Arkansas
Fayetteville, AR 72701
Tel : 501 575 2420 Fax : 501 575 5339
{hexmoor, sbattul}@uark.edu

Abstract

A methodology for agents to acquire human desired norms for social interaction is discussed. To improve social adeptness, agents need to recognize social influences of their social actions and to act accordingly. An implemented system with a number of Unmanned Combat Aerial Vehicles (UCAVs) either under autonomous agent control or a human remote pilot control is presented. We illustrate interpersonal norms that agents discover from offering "help" to a human-controlled plane as well as asking for "help" from the human-controlled plane and in both cases receiving simple feedback.

Introduction

We envision agents to adjust their understanding of a human's preferences for interaction so that their interaction is more cognitively appealing to the human user. Since in general it is difficult to extract the human-desired characteristics of interaction and often these desired methods don't fit a mathematical model, we wish for agents to use simple human feedback to learn normative interactions that are preferred. Therefore, we are developing a methodology to capture normative interactions favored by human users. Whereas norms are specific patterns of interactions, social attitudes such as autonomy are more enduring. Elsewhere, as part of building socially adept agents, we have explored how agents might learn human-desired levels for such attitudes (Hexmoor 2001a and 2001b, Hexmoor, Holmback and Duncan 2001).

Norms provide guides to behavior. They encode conditions under which to perform social actions. However, they are flexible and defeasible. Norms have been developed to enhance the reliability of communication, to make an agent's actions predictable and verifiable, to facilitate coordination of agent actions, to enhance social stability, and for other purposes (Conte, Falcone and Sartor 1999). Recently, norms have been reported useful for incorporating emotions and emotional actions (Moldt and Sheve 2001). We will not consider emotional implications in this paper. In (Carberry and Schroeder 2001) we find a report on recognizing attitudes and understating motivations in dialogues. In contrast, our work is limited to very simple human input in the form of a number that reflects satisfaction or dissatisfaction. We are also not considering dialogue but rapid interactions in a dynamic environment.

Eliciting user preferences requires modeling the user. Trust and understanding between user and agent can be increased by allowing the agent to get further input from the user about his preferences and desires (Fleming and Cohen 1999). Preference modeling is being widely used in building automated travel schedules. For instance, Greg Linden's Automated Travel assistant starts with minimal information about the user's preferences, and preferences are inferred incrementally by analyzing the feedback given by the user (Linden, Hanks and Lesh 1997). Performance in the system is improved by learning from the user feedback as well as giving more appropriate information in subsequent iterations. Our work is similar to Pattie Maes's learning interface agent architecture in that agents act primarily by observing their users. As in our system, our agents collect a case history of scenarios. She has devised a case retrieval mechanisms based on the distance between the current state and each of the past situations the system has stored in its memory using a weighted sum of several relevant features. We have tested our methodology in a testbed of a team of Unmanned Combat Aerial Vehicles (UCAVs). This testbed was also developed to explore teamwork and explicit reasoning about roles (Hexmoor and Zhang 2002). In our implemented testbed three or more fighter aircraft agents have the mission to deliver a bomb over a remote designated site. There is a one to one relationship between agents and planes. Artificial agents control all the planes except one, which is controlled by a human operator. The human operator controls his/her plane in the field along with the other planes, and will have similar visual and auditory sensing as well as similar flight maneuvering capabilities. The system is implemented in Java. We simulate Surface to Air Missile sites (SAMs), which are randomly generated each time when the program starts running.
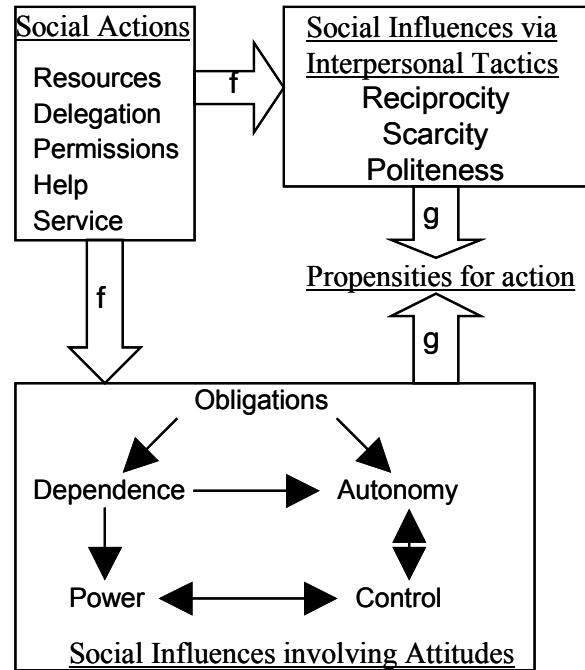
Figure 1 Social actions and influences; values and norms are not shown in the figure

Later in this paper we will see simulator screens. As the human operator requests help (which is a social action), agents respond to this request and the human will rate their offer of help. This is the main form of normative behavior we will examine.

In the remainder of this paper we will first discuss social actions, influences, and our methodology for capturing norms. We then briefly describe our testbed and example of specific norms we have captured for a single social action of offering and asking for "help." We will then describe our plans for agents' use of norms in their future interaction. We will end by some concluding remarks.

Social Actions, Social Influence, and Norm Discovery

Whereas *physical actions* predominantly produce physical change, and *speech actions* predominantly produce epistemic change, *social actions* predominantly produce influence. Social action might be an action by one agent toward another, a mutual action of multiple agents, a bilateral action by multiple agents, or a group action. For an example, consider actions about Help. The actions are performed to aid others in their task and specific actions might be to provide, to withhold, to request, or to reject help.

As shown in Figure 1, social actions promote social influences. Since social factors are inter-related, initial social influences due to social actions might produce secondary and indirect influences. Actions of agents are guided by social values and norms, which affect the relationships among social influences. Therefore, social actions of one agent will influence the other agent in the context of prevailing norms and values as well as by the strength of the social action.

In this paper, influences are normative (utilitarian) influences in social psychology in that agents reason about gains and losses from interpersonal interactions (Kassin 2001). Influences might involve overt interpersonal tactics such as reciprocity, scarcity, and politeness. Overt influences are immediate and deliberate as in interactions described in FaintPop (Ohguro et al. 2001). Reciprocity is when one agent returns an act by another or in effect *pays* for an act. An agent might reason about the reciprocity norm and perform a social act (e.g., help) based on an expected propensity in repayment. Scarcity is the norm that short supply produces a demand. An agent might use that norm and deny or hide services or resources. This is commonly used in theories of persuasion (Larson 1998). Politeness as an interpersonal tactic is to get another agent to yield to another agent.

Influences might also be indirect and of indefinite duration. One type of indirect influence is via changes of attitudes. This is shown as the box in the lower part of Figure 1. These are perceived changes in social relationships that affect an agent's ties. The Figure shows our focus on Autonomy, Dependence, Obligations, Control, and Power and salient relationships we see among them. Later in this section we will discuss interdependencies among attitudes. Let's refer to the set of influences as I. Let's refer to the set of social actions as A. We define function f that maps the agent's current beliefs B, a set of currently active values V,

a set N of currently active norms, and a set of social actions A to a set of influences I:

$$f: B \times V \times N \times A \rightarrow I.$$

An example is the assigning of homework by a teacher to a student. Following shared norms governing relationships among teachers and students, assigning homework produce the influence for the student to adopt the obligation to carry out the homework.

Agents use influences that result from social actions they experience in their action selection. In addition to social influences, action selection accounts for means end analysis and rationality principles that are governed by the agent's endogenous sources. How action selection is affected by social influences is a complex issue that is beyond our current scope and is denoted as function g in Figure 2. Agents can project such a propensity for action in deciding to perform a social action. The reasoning might also include a chain effect where one agent produces an influence in another, which in turn produce an effect in another and so on. An agent can intend such a proliferation of influences and intentionally start such a chain reaction. This in fact is commonplace in a team setting.

Let's consider a variation in function f where an agent performs a social action. In our simulation we considered social action for an agent to offer help and five variations for asking for help. The human user senses a social influence and provides feedback between -10.0 (to indicate disapproval) to $+10.0$ (to indicate approval). We keep norms but omit values for simplicity. So now we have:

$$B_{shared} \times N_{human} \times A_{agent} \rightarrow I_{human} \times Feedback_{human}$$

B_{shared} is the situation (e.g., plane and SAM locations) known to both agent and human. The agent can consider the human feedback as the reward. The agent who performed the social action concludes rules of the following form:

$$B_{shared} \rightarrow A_{human\text{-}desired} \times Desired\text{-}preference.$$

This rule is an inferred norm. $A_{human\text{-}desired}$ is social action the agent believes human wishes for it to perform (i.e, A_{agent}). Initially, desired-preference is unknown and can be randomly assigned for pairs of behavior and action. But as feedback is provided the agent can compute desirability by considering past values of desirability and the amount of reward:

Desired-preference = Feedback + (λ * Desired-preference)

This formulation of desirability is similar to update function in Q learning. λ is usually set to low levels such as 0.25. This does not simplify our method to Q learning. We consider the learned rule to be the agent's perception of a norm of its interaction with the human. This type of norm is interpersonal norm as opposed to community governed. In this methodology we simply suggest recording

rule-like norms of interaction and have not applied machine-learning techniques (For application of machine learning in user modeling see (Webb, Pazzani and Billsus 2001)).

Other variations to our methodology are possible. Consider the norms are known (i.e., shared) instead of inferring them. In one variation, given prevailing situations (i.e., shared beliefs) to an agent, it can predict human social actions. . If there are competing norms possible, it is useful to guess which norms is being followed. The human feedback (i.e., social influences) along with human social action and current situation (i.e., shared beliefs) can be used to in another variation where an agent can guess which norm matches best. We will explore these variations in our future work.

Testbed and Norms

Figure 2 shows the main simulator screen with one agent-controlled plane flying close together followed by a human-controlled plane. The screen depicts a mountainous terrain, SAMs (Surface to Air Missiles) placed at predetermined positions and a set of agents (planes).
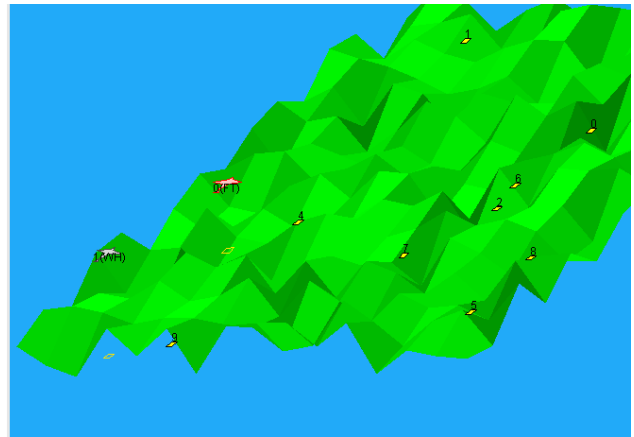

Figure 2 Testbed Screen

A timer is started as soon as the planes take off from the base to reach the Destination. Each agent has a *visible region* beyond which the agent can't see. In addition, each SAM has a visible region. The SAMs attack the agents flying within their visible region and the agents attack the SAMs within their visible region. The agent experiences a highest hit probability when it first enters the visible region of a SAM with the amount of 0.05 and it increases by 0.001 with each increment of the simulated time. So if an agent stays in the visible region of a SAM for 10 cycles the probability of the agent being hit will be: Hit probability = 0.05 + 0.001 * 10 = 0.06. Suppose the agent is in the visible region of two SAMs A and B at the same time the agent's hit probability will be the union of both the hit probabilities since hit probabilities in SAMs A and B are disjoint.
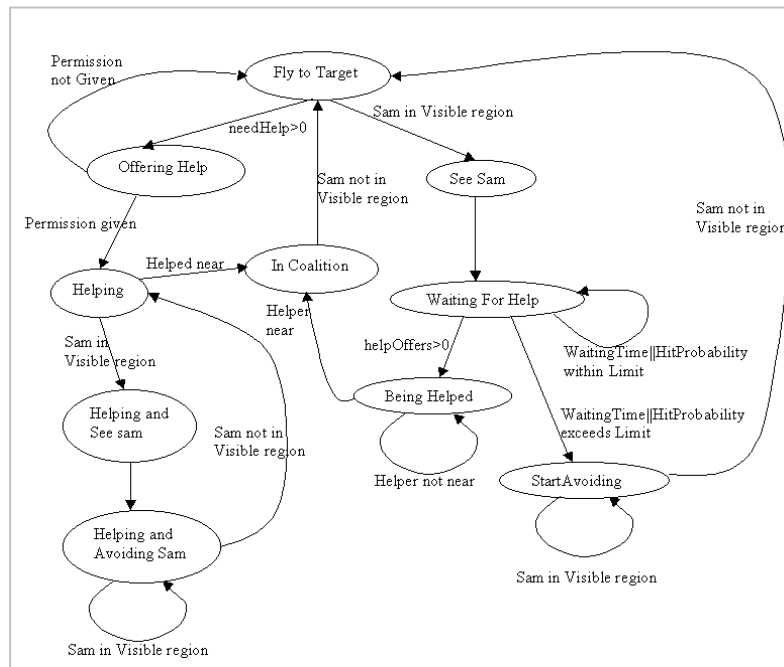
Figure 3 State diagram

$$P(A \text{ and } B) = P(A) + P(B) - P(A) * P(B)$$

Agents can be in any of the Following nine States: *FlytoTarget, SeeSam, WaitingforHelp, OfferingHelp, BeingHelped, StartAvoiding Helping, HelpingandAvoidingSam,InCoalition.* The state diagram in Figure 3 shows the states and their transitions.

An agent asks for Help when its hit probability exceeds a certain limit[1], and enters into *WaitingforHelp* state. Agents that are in the *FlytoTarget* state offer help to agents needing Help. Agents *WaitingforHelp* receive the help offered by other agents and select the best offer. It then issues permission to that agent and enters into *BeingHelped* state. The agent that received permission to help changes its state to *Helping*.The helping agent moves towards the agent it is helping. When the helper and the helped agents come close to each other, they enter into the coalition state. They remain in the coalition state until each of the planes is out of the SAM's visible region. When they move out of the SAMs visible region, they break the coalition and each agent enters *FlytoTarget* state.

In addition to the transition of states explained above there are some more transitions. The agent being helped can't indefinitely wait for the helping agent to arrive. So if the hit probability exceeds a certain level the agent stops waiting and enters into *StartAvoiding* state. The helping agent is informed of this transition and it changes its state to *FlytoTarget* state. Figure 4 shows the control panel for the human control of one of the planes. The human pilot can speed up or down, gain or lose elevation, turn right or left, launch missile, land, ask for help, respond to help (by issuing "I am coming"), and rate offers of help by providing feedback. Another screen (shown in Figure 5) displays the rate of hit probabilities for an agent. This display contains a 2 dimensional plot where the y-axis shows the hit prob abilities in the range $0.0 - 1.0$. The x-axis gives the cycle number of the simulator. The dropdown list below the plot is used to select the agent for which the hit probability is plotted.

Consider a scenario in which the human pilot asks and receives help. When the human-controlled plane asks for help other agents respond. In each scenario the human gives a feed back indicating satisfaction with the offer of help. The feedback ranges from -10 to 10. The feedback along with the situation under which this help was offered is captured for later use. Figure 5 shows a few scenarios and rewards (feedbacks). Each scenario is basis for a norm. Beliefs in a norm are (a) distance, (b) position, and (c) state of the agent when it offered help. Since the agents don't yet use these scenarios, feedback is not converted into desirability values. Figure 5a shows one scenario with positive feedback of 6 and one with negative feedback of -4. Figure 5b shows two negative scenarios with feedbacks of -1 and -7. For an alternate scenario consider when an agent-controlled pilot's choice of phrase when it asks for help from the human pilot. Natural language phrase for help request such as "may day" or "Can I get some assistance" partly reflect the intensity of need for help. There is also subjectivity among humans in what phrase fits the situation best.

---

[1] Currently, this value is at 0.2. However, this value can be adjusted.

Figure 4 Human control panel

This makes it difficult to hard-code a match between help phrases and situations an agent may encounter. Our agents learn this for each human interaction. In our simulation we allow an agent five options at different intensity levels: "Mayday", "I want help", "I need help", "Can I get some assistance", and "I may need Help".

Five different situations are considered and the Human controlled maps these situations with the appropriate requests.

The agents may be in any of the following situations (i.e, beliefs):

Situation 1 – Agent doesn't see SAM and its not attacked
Situation 2 -- Agent doesn't see SAM but attacked
Situation 3 – Agent sees SAM and being attacked
Situation 4 – Agent sees SAM and attacked closely
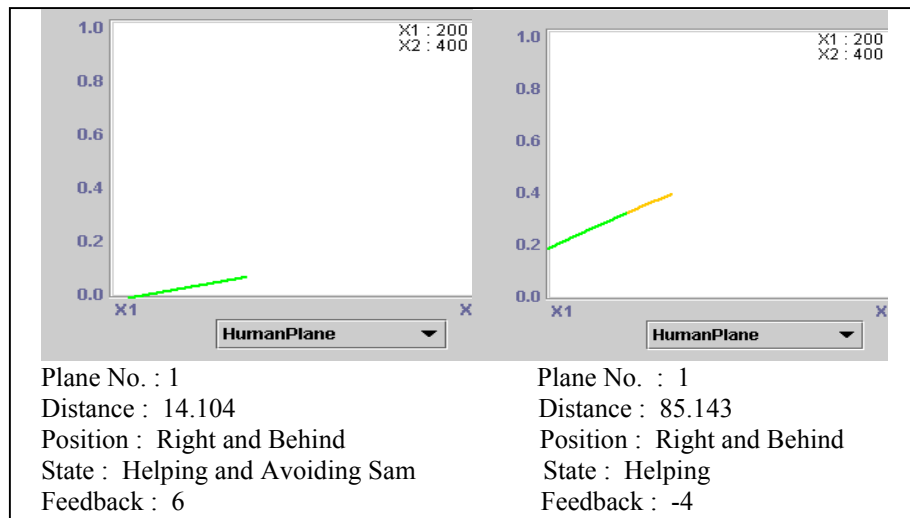Situation 5 – Agent attacked very closely.

Plane No. : 1
Distance : 14.104
Position : Right and Behind
State : Helping and Avoiding Sam
Feedback : 6

Plane No. : 1
Distance : 85.143
Position : Right and Behind
State : Helping
Feedback : -4

Figure 5a a pair of norms for "offer help"

Plane No. : 1
Distance : 16.675
Position : Right and Front
State : Helping And Avoiding Sam
Feedback : -1

Plane No. : 1
Distance : 37.015
Position : Right and Front
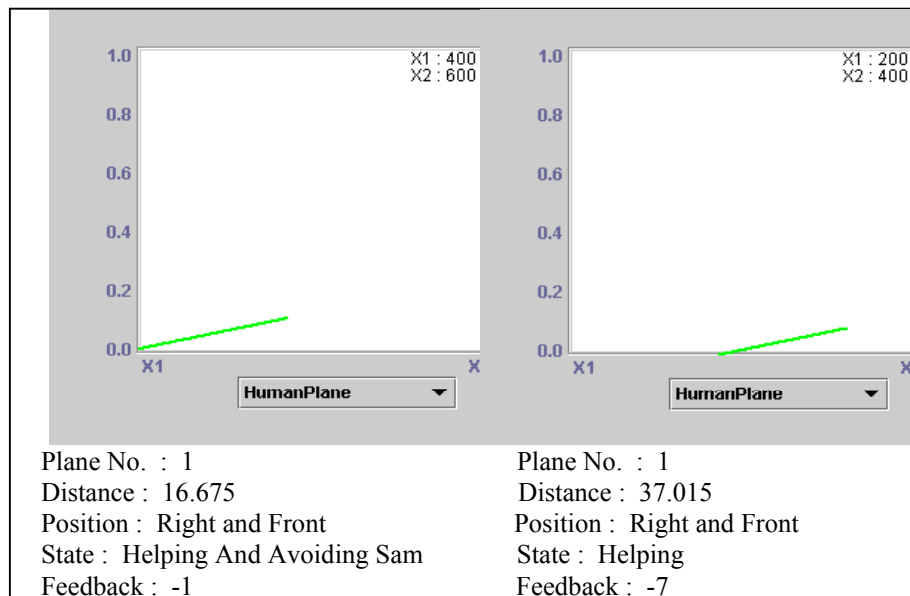State : Helping
Feedback : -7

Figure 5b another pair of norms for "offer help"

The following table shows a set of human feedback values issued during a run.

	Mayday	I want help	I need help	I may need help	Can I get help
Situation1	-8.0	-6.0	-2.0	2.0	7.0
Situation2	-7.0	-2.0	2.0	7.0	3.0
Situation3	-1.0	2.0	6.0	1.0	-5.0
Situation4	2.0	8.0	3.0	-1.0	-6.0
Situation5	9.0	6.0	0	-2.0	-8.0

Via human feedback, the agent learns to select the appropriate phrase for the given situation. For instance, in situation 1, "can I get help was most preferred.

Conclusion

To increase social adeptness in agents interacting with human users, we have developed a methodology for agents to acquire human desired norms for social interaction. It is in general difficult for humans to express their preferences but they know when they like or dislike a situation when they are presented with a situation. We have implemented a methodology to automatically construct preferred interpersonal norms of interaction. This is demonstrated in a system with a number of UCAVs either under autonomous agent control or a human remote pilot control. We illustrate norms that agents discover from offering help as well as selecting an appropriate phrase to ask for help and receiving simple human feedback. In future work, we will empirically show usefulness of such norms about help. We will also pursue variations we discussed to our methodology.

Acknowledgement

References

Carberry, S., and Schroeder, L. 2001. Recognizing and conveying attitude and its underlying motivation, In 2nd Workshop on Attitude, Personality and Emotions in User-Adapted Interaction, *User Modeling* 2001, Sonthofen, Germany.

Conte, R., Falcone, R., and G. Sartor, 1999. Agents and Norms: How to fill the gap. *AI and Law, special issue on Agents and Norms.*

Fleming, M., and Cohen, R. 1999. User modeling in the Design of Interactive Interface Agents, In *Proceedings of The Seventh International Conference on User Modeling*, pages 67-76.

Hexmoor, H., Holmback, H., and Duncan, L. 2001. Detecting, Expressing, and Harmonizing Autonomy in Communication Between Social Agents, *AAAI spring symposium on Robust Autonomy*, Stanford, AAAI press.

Hexmoor, H. 2001a. Stages of Autonomy Determination, *IEEE Transactions on Man, Machine, and cybernetics*- Part C (SMC-C), Vol. 31, No. 4, November 2001.

Hexmoor, H. eds. 2001b. A Cognitive Model of Situated Autonomy, In *Advances in Artificial Intelligence*, Springer LNAI2112 -pages 325-334, Kowalczk, Wai Loke, Reed, and William .

Hexmoor, H., and Zhang, X., 2002. Socially Intelligent Combat Air Simulator, In proceedings of The Seventh Pacific Rim International Conference on Artificial Intelligence (PRICAI-2002), Tokyo, Japan.

Kassin, S. 2001. *Psychology*, Third Edition, Prentice-Hall

Larson, C. U. 1998. *Persuasion: Reception and Responsibility,* 9th edition. Boston: Wadsworth.

Linden, G., Hanks, S., Lesh, N. 1997. Interactive Assessment of User Preference Models:The Automated Travel Assistant, In Anthony Jameson, Cécile Paris, and Carlo Tasso (Eds.), *User Modeling: Proceedings of the Sixth International Conference, UM97*. Vienna, New York: Springer Wien New York.

Ohguro, T., Kuwabara, K., Owada, T., and Shirai, Y. 2001. FaintPop: In touch with the social relationships, In International Workshop on Social Intelligence Design, The 15th Annual Conference of JSAI, Japan.

Moldt, D., von Scheve, C. 2001. Emotions and Multimodal Interface-Agents: A Sociological View, In H. Oberquelle, R. Oppermann, J. Krause, (Eds): *Mensch & Computer 2001*. Tagungsband der 1. fachübergreifenden Konferenz. Stuttgart: Teubner Verlag, 287-295.

Webb, G., Pazzani, M.J., and Billsus, D. 2001. *Machine Learning for User Modeling*, User Modeling and User-Adapted Interaction 11: 19-29, Kluwer Academic Publishers, the Netherlands.