Toward Affective Cognitive Robots for Human-Robot Interaction

M. Scheutz, P. Schermerhorn, C. Middendorff, J. Kramer, D. Anderson, and A. Dingler
Artificial Intelligence and Robotics Laboratory
Department of Computer Science and Engineering
University of Notre Dame, Notre Dame, IN 46556, USA
{mscheutz,pscherm1,cmidden1,jkramer3,danders,adingler}@cse.nd.edu

Abstract

We present an architecture for complex affective robots for human-robot interaction. After describing our rationale for using affect as a means of "architectural integration", we give a quick conceptual example of how affect can play an organizational role in a complex agent and then describe our proposed affective architecture, its functionality and implementation, and results from experimental evaluations on an autonomous robot.

Introduction

Complex robots that have to interact with people in real-time using natural language pose several design challenges, from the functional definition of the overall robotic architecture, to the integration of different modules that operate at different temporal scales using different representations, to the implementation in parallel hardware that has to be fault-tolerant and allow for high reactivity. Work on high-level cognitive agent architectures has traditionally focused on reasoning and planning in simulated agents, thus largely ignoring real-world, real-time constraints of embodied agents, while work in AI robotics has mainly been centered around lower level perceptual-motor couplings and behaviors, often purposefully ignoring higher-level deliberative functions. Different from other "hybrid architectures" (Arkin & Balch 1997; Gat 1991; Simmons *et al.* 1997) we believe that for realistic human-robot interactions an affective architecture is needed which (1) effectively integrates affect perception and display with architecture-internal affect processing to produce realistic, believable behaviors, and (2) uses "affective control" to coordinate the many, very diverse components of the architecture.

In this paper, we present our proposal for such an architecture. Specifically, we give an overview of its functional organization, outline the role of affect in the integration of various subsystems, describe some of the implemented components relevant to affect processing, and provide results from preliminary experiments with a robot in the context of a robotic waiter scenario (Maxwell *et al.* 1999)).

Affective Architectures for Complex Robots

Affect is deeply intertwined with cognitive processing in humans. *Negative affect*, for example, can bias problem solving strategies in humans towards local, bottom-up processing, whereas *positive affect* leads in many case to global, top-down approaches (Bless, Schwarz, & Wieland 1996). Moreover, humans often seem to rely on *affective memory* (Blaney 1986) to evaluate a situation quickly instead of performing a longer, more complex cognitive evaluation (Kahneman, Wakker, & Sarin 1997) as *affective evaluations* seem to encode implicit knowledge about the likelihood of occurrence of a positive or negative future event (Clore, Gasper, & Conway 2001)). Finally, affect is crucially involved in *social control* ranging from signaling emotional states (e.g., pain) through facial expressions and gestures (Ekman 1993) to perceptions of affective states that cause approval or disapproval of one's own or another agents' actions (relative to given norms), which can trigger corrective responses (e.g., guilt).

Although affect (and in particular, emotion) has been investigated to varying degrees ever since the beginning of AI (Pfeifer 1988), only a few projects currently attempt integrate affect in complex simulated agents (Marsella & Gratch 2002; Hudlicka 2004). And while the utility of artificial emotions for human-robot interaction has been investigated in non-cognitive robots (Breazeal 2002), we are not aware of any other project that investigates *the possible roles of affect in architectures for complex cognitive robots*.

Architectural Requirements for Affect

Affect might have several functional roles in agent architectures (Scheutz 2004). Some have argued that affect serves the purpose of integration and management of multiple processes and is thus required for the effective functioning of an autonomous system (Ortony, Norman, & Revelle forthcoming). Specifically, affect allows for motivational signals originating not from changes in the external environment detected *via* sensors, but from components within the architecture itself (e.g., from deliberative subsystems). Such signals can then influence various other parts of the architecture and modify goal management, action selection, and learning.

While the class of affective states found in nature is comprised of many different kinds of states and processes (sensations, feelings, emotions, moods, etc.), we will focus on complex motivational and emotional states (such as "desiring to win a grant" or "worrying about whether the proposal can be completed in time"), which can be caused by a combination of perceptions and processes internal to the agent (e.g., results of complex deliberations about the utility of trying to achieve a particular goal compared to alternatives). Complex emotions, for example, may include any of the following, based on (Beaudoin & Sloman 1993; Ortony, Clore, & Collins 1988) and others:

1. an elicitor (e.g., *the grant proposal*)
2. an eliciting condition (e.g., *the possibility of (1) not being completed by the deadline*)
3. criteria for the evaluation of (2) based on various factors such as beliefs, goals, norms, standards, tastes, attitudes, etc. (e.g., *completing (1) by the deadline is crucial to research career*)
4. an evaluation of (2) in terms of (3) (e.g., *(2) is undesirable*)
5. possible causes for (2) (e.g., *deadline approaching rapidly, work progressing too slowly, etc.*)
6. a hedonic attitude towards (2) (e.g., *displeasure*)
7. a measure of the urgency to act on (1) given (2) (e.g., *urgent*)
8. a set of strategies to cope with (2) (e.g., *cancel meetings, focus attention on (1), etc.*)
9. a set of motivations to be instantiated based on (8) (e.g., *being able to continue one's research, being able to fund students, etc.*)
10. a set of emotions to be instantiated based on (4) through (8) (e.g., *distress*)
11. the selected motivation (if any) based on (4) through (8) (e.g., *being able to continue research*)

Consequently, complex representational and processing mechanisms (e.g., pattern matching and rule instantiation) are required for architectures to be able to represent possible or impossible, likely or unlikely future or hypothetical states and thus support complex motivations and emotions.

We have defined an architecture that allows for the representation and instantiation of complex motivational and emotional states in an effort to study the utility of these states in architecture internal processes (in addition to being able to detect affect in others and to express it).

A Partial Affective Architecture

Figure 1 depicts a partial view of the functional organization of the proposed affective architecture for complex robots.[1] Columns separate sensors, perceptual, central, and action processing components, and effectors. All boxes depict autonomous computing components that can operate in parallel and communicate via several types of communication links (in the figure, only links that are part of affective processing pathways are shown).[2] Labels of components denote the functional role of the component in the overall system. The colored/grayish components are part of the affect system, which is used to control the robot's actions and drive the behavior of the robot over time by providing internal feedback about the success or failure of an action.

We will now illustrate the architecture-internal roles of affect (i.e., motivations and emotions) in a concrete example taken from a (fictitious) robotic waiter scenario in a bar, which specifically demonstrates the roles of *joy*, *pride*, *gratitude*, and *relief* as they interact with *goal management*, *memory*, and *social control*.

Example: An Envisioned Interaction between a Patron and a Robotic Waiter

Suppose the bartender was late fixing the drink ordered by patron X. The waiter W has noticed that (by examining the *urgency slot* in the representation of its goal to serve X's drink) and instantiated a *fear state* about the possibility of being yelled at by an angry X. Among W's current goals are to keep patrons happy, to deliver the drink to X, and to scold the bartender for being late with the drink preparation. This goal is caused by the fear state and might result in a "shift blame" strategy, if yelled at by X. W has a positive attitude towards happy people in general, and consequently a positive attitude towards X. W's affect evaluation system uses this as a norm to thank people when they make compliments. W also has an extended notion of "self" that includes the bartender (e.g., as a result of their both being part of the same institution). Currently, W is located next to the X (after having picked up the drink from the bartender, although without having had a chance to scold him), and has just said, "Here's your Manhattan, Sir, sorry for the delay." X has taken the drink, sipped it, smiled, and said, "No problem, you guys make the best Manhattans!"

Tracing this scenario through the architecture, where colored/dashed links indicate the information flow connected to various affective states (red=*relief*, green=*gratitude*, purple=*pride*, dark blue=*joy at having achieved the goal of delivering the drink*, light blue = *joy at having achieved the goal of making the patron happy*): W's attention is focused on X, whose face becomes the focal point of the scene. The completion of the *deliver drink* goal elicits *joy*. W sees X smile and hears the tone of his exclamation, from which it concludes that X is pleased. This elicits a *joy* reaction, due to facilitating the *keep patrons happy* goal. It is also taken as evidence that X is not angry, which reduces W's level of *fear*. W recognizes X's exclamation as a compliment, which elicits *relief* due to the conclusion that X is not angry and has the added effect of canceling the *scold bartender* goal. Further, *pride* is elicited as a combination of facilitating the *keep patrons happy* goal and being part of the compliment's object. W obsesses further about the attribution of the compliment, determines that the bartender is responsible, and generates a *congratulate bartender* goal, while committing this event to episodic memory (because of its affective salience). Finally, due to the *thank people* norm, W is motivated to say, "Thank you very much, sir. I'll tell the bartender you said so". The whole event will be stored in W's episodic memory for subsequent analysis. In particular, W will be able to

[1] This diagram of the architecture is restricted to visual and auditory sensors and camera motors and speaker as effectors; other sensors such as sonar and laser sensors, or effectors, such as grippers and wheel motors, are not shown. Moreover, several internal components related to navigation and complex action sequencing (e.g., spatial maps, sequencers, etc.) are not depicted either.

[2] The implementation builds on the ADE system available at `http://ade.sourcefore.net/`.

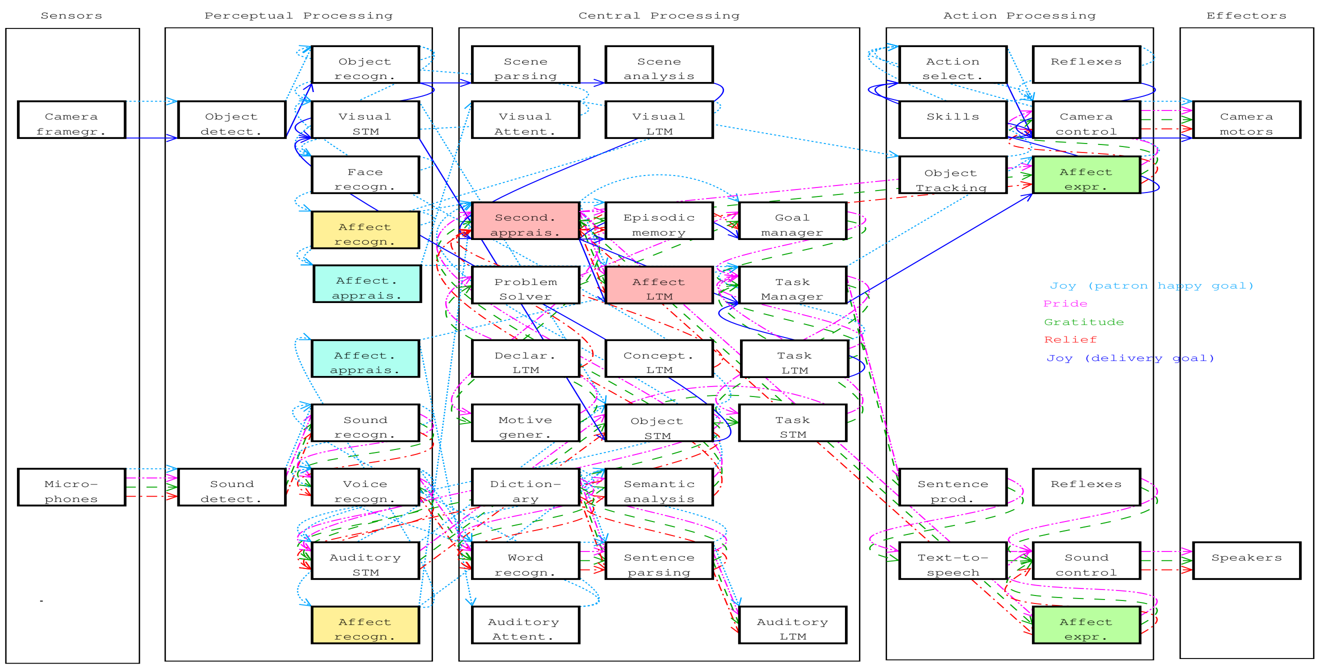

Figure 1: A partial view of the proposed affect architecture for a complex robot. Boxes depict concurrently running components of varying complexity and colored/grayish components indicate affective control mechanisms. The five different arrow types depicted indicated the information flow through the architecture involved in processing five different affective processes.

learn that it is good to wait for the outcome of an action A (like *deliver drink*) before shifting blame to agents causing a delay in A (e.g., scolding the bartender right away), since the result of A might not cause W to be blamed for A being delayed. "Learning" here could mean reducing the strength of the connection between the *fear state*, which instantiates *anger*, and ultimately *blame*, and the urgency parameter of the goal *blame cause of anger*.

Evaluation of the Current System

While being able to achieve a natural interaction as described in the ficticious scenario is still a long-term goal and will unlikely be feasible in the near future in any robotic system, it is possible to test the proposed architecture in a variety of *limited versions* of the above scenario. However, it is important to point out that a *comprehensive quantitative evaluation* of such a robotic system, even though highly desirable in principle, can be very difficult to achieve in practice.[3] Ideally, we would like a highly controlled set of initial conditions, under which the robot has to perform a given task. Based on a set of performance criteria and by carefully manipulating various dimensions (ranging from task, to architectural, to environmental parameters), we could then empirically determine a "performance space", which reflects the advantages and disadvantages of various architectural mechanisms (e.g., affective control) in given situations. For example, one could set up a scenario, where patrons with different levels of frustration make too many requests within a short period of time for the robot to be able to satisfy them all and examine the utility of using affective evaluations of the patrons and prioritization of goals based on affective states to determine the best order in which to satisfy the patrons' requests such that their overall levels of frustrations as measured by special purpose devices (Riseberg *et al.* 1998) will be minimized. We believe that this kind of evaluation will be critical for future robotic systems in order to determine not only their utility, but also the appropriateness and safety in human-robot interactions (both of which will become more and more critical issues as AI technology matures and robots become part of society).

Unfortunately, such quantitative evaluations are not yet possible with current robotic systems, which are still too brittle in their performance (given the low reliability of visual and auditory perceptual processing in noisy environments, the computational limitations on autonomous systems, the burden of real-time processing of multiple information channels, etc.). Hence, we have to restrict ourselves to a *quantitative evaluation*, which nevertheless can verify some of the critical features of the system. In particular, it can demonstrate that (1) the system is capable of successful interaction with people in real-world environments using natural language, (2) it is capable of expressing and using affect as part of the task performance, and (3) it can accomplish the waiter task in a variety of initial conditions.

[3] Even in the case of human waiters general quantitative, agreed-upon criteria for performance evaluation might not exist, even though qualitatively it seems easy to distinguish good from bad waiters based on efficiency in serving, polite behavior with patrons, personality, etc.

Next we will describe the architecture implementation, the experimental setup, the action interpreter, and results from experiments in a simplified waiter task.

Implementation and Experimental Setup

The implementation of the proposed architecture is–almost by necessity–work in progress given the intrinsic complexity of the system.[4] However, several subsystems are functional and their interactions allow the robot to integrate sensory information from sonar, laser, bumper, vision, and auditory sensors, process information at several levels in parallel, and perform several actions: from finding and tracking multiple people in a room, to following trajectories in a room, to producing emotional speech output. For space reasons, we can only briefly touch on some subsystems.

The *vision subsystem*, for example, provides information about detected faces using and histogram methods from (Yang, Kriegman, & Ahuja 2002) as well as skin color, color of clothes, and camera angle relative to the robot, which are used in conjunction with information about the distance of an object (coming from sonar and laser sensors) to find and track people in a room (Scheutz, McRaven, & Cserey 2004) and determine some of their salient features such as height (Byers *et al.* 2003) for future identification. The emotion recognition component currently only classifies faces as "happy", "sad" or "neutral".[5]

The *natural language processing subsystem* integrates and extends various existing components (CMU's *SPHINX* and IBM's *ViaVoice* for spoken word recognition, an enhanced version of "thought treasure" (Mueller 1998) for natural language understanding and production, and a modified version of the University of Edinburgh's *Festival* system for speech synthesis). Novel components are employed for affect recognition in voices and an affect expression in spoken language to recognize and express the four emotional states "angry", "frightened", "happy", and "sad".[6]

The *action control subsystem* is based on a novel *action interpreter*, which interprets scripts for natural language understanding augmented by action primitives for the control of actions (see following Subsection). These scripts can be combined in hierarchical and recursive ways, yielding complex behaviors from basic behavioral primitives, which are grounded in basic *skills* (the bottom layer control structures are implemented as motor schemas as in (Arkin & Balch 1997)). Moreover, several spatial maps are used for the representation of locations of the robot, people, and other salient objects in the environment, as well as path planning and high-level navigation.

The architecture is implemented on a fully autonomous Pioneer Peoplebot from ActivMedia (with sonar, pan-tilt-zoom camera, a SICK laser range finder, three sonar rings, two microphones, two speakers, one built-in Linux PC104 board, and two added Dell laptops, one running Linux, the other running Windows XP, connected through a local Ethernet network, with wireless access to the outside world).

Action Processing in the Waiter Task

Figure 2 depicts snapshots from a typical run with the robot in the waiter task. Here the robot first has to find people in the reception area (A) (this is achieved using camera and range finder information). Once a person has been detected, the robot approaches the person and starts a conversation (B) (the figure shows the face and eye-brow tracking for emotion detection), the length of which depends on the robot's "happiness", which, in turn, is affected by the conversation topic and the degree to which the robot is able to find an appropriate semantic representation of the human's utterance. Eventually, the robot asks if the person would like a drink (C), takes the order (if it knows that the drink is available), and moves to the bar area (D), which is stored in an internal map. It detects the bar tender and asks him to get the requested drink (E) (the figure shows the bartender putting the drink on the robot's tray), after which it returns to the person asking the person to take the drink (F). Successful timely task completion leads to a change in the robot's "happiness", which is then reflected in its voice output and propensity to converse more with people. All motions are controlled by the action interpreter based on scripts stored in long-term memory:

```
====serve-drink//serve.V
role01-of=waiter|
role02-of=human|
role03-of=beverage|
role04-of=bar|
timeout-of=600sec|
event01-of=[wander waiter]|
event02-of=[shiftFOA waiter human]|
event03-of=[converse waiter human]|
event04-of=[say-to waiter human [fetch-from waiter bar beverage]]|
event05-of=[move-to waiter bar]|
event06-of=[say-to waiter bar [active-goal waiter
                               [fetch-from waiter bar beverage]]]|
event07-of=[move-to waiter human]|
event08-of=[shiftFOA waiter human]|
event09-of=[say-to waiter human [active-goal-weaker waiter
                             [fetch-from human waiter beverage]]]|
```

"Roles" indicate entities and "events" denote of descriptions of actions and state-of-affairs involving these entities. In the above script all events denote actions (MOVE-TO, SHIFT-FOA for "shift focus of attention", and SAY-TO) are action primitives, while "wander" and "converse" refer to other scripts, leading to a recursive structure of scripts).

A primitive action like MOVE-TO($\langle waiter \rangle, \langle location \rangle$) then has a particular meaning to the robotic system. In this case, the action interpreter passes the action on to the navigation system (after having substituted "self" for $\langle waiter \rangle$), which interprets it as a command to move the robot to the coordinates (x, y) of $\langle location \rangle$ (represented in a discrete, topological map). The high-level navigation system translates the action into commands for the low-level navigation

[4]Most components are implemented in some rudimentary form, even though the majority of the components does not yet update in parallel.

[5]Emotion detection in uncontrolled environments is very difficult and turned out to be too inexact to pick up on subtle emotional expressions and changes. The current system is too unreliable to be of practical use. We recently learned, however, that Jeff Cohen and Takeo Kanade with their groups at CMU have been making progress in tracking action units in faces, which could lead to much more reliable emotion detection and tracking.

[6]A separate paper on the details of these subsystems, which employ a combination of signal processing and neural-network methods, is in preparation.

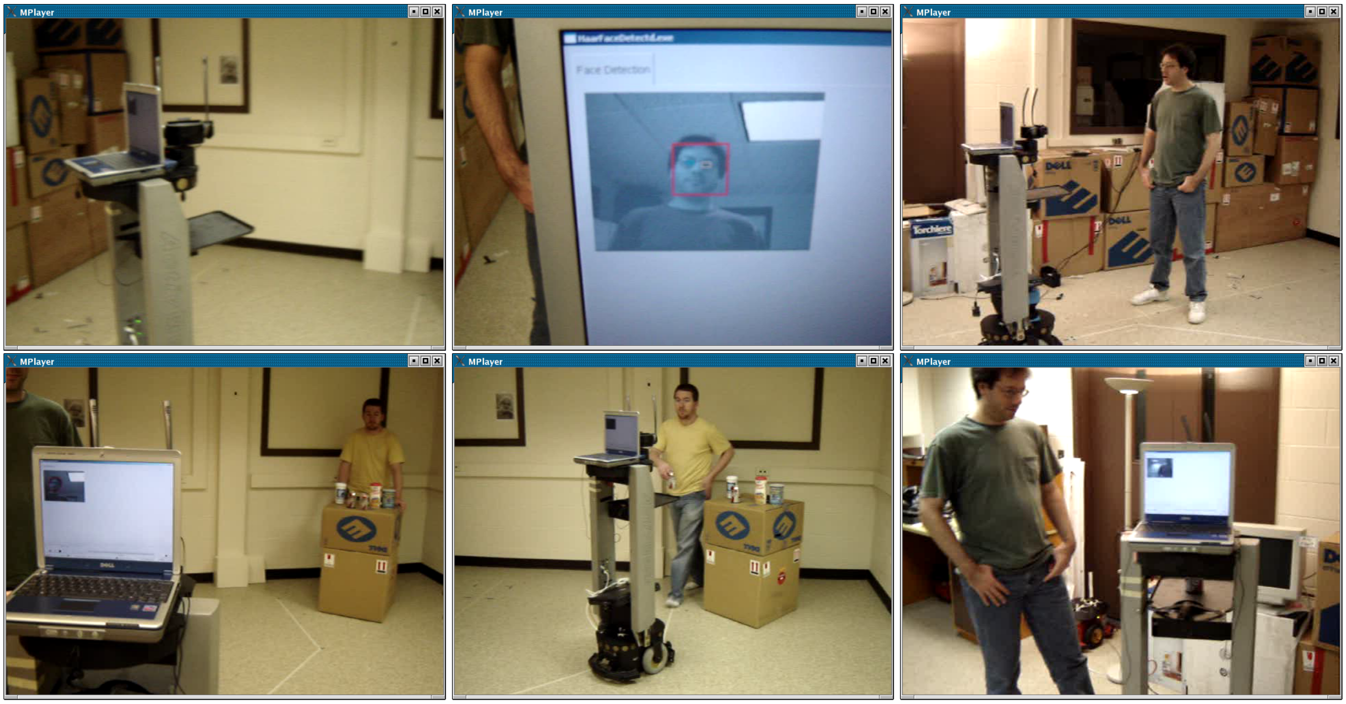

Figure 2: Snapshots (A) through (F) (top left to bottom right) from the robot's performance of a waiter task (see text for details).

system, which eventually causes the robot to move in a particular direction, if possible (e.g., it will not move there if obstacles block the location, although it will attempt to move around obstacles that obstruct the path to the final location).

Note that the same script can also be used for "understanding" actions of others without having to carry them out, because the "serve-drink script" is not defined in an *indexical way* (i.e., assuming that the robot is the waiter), but rather defined generically so that the action interpreter can use the present context to determine who should assume the waiter role. If based on the context an entity other than the robot should be the waiter, then this will only lead to a new binding of the symbol "waiter" in the present context in the robot's working memory, whereas a binding of "waiter" to "self" will cause the robot to perform an action.

In addition to action primitives or references to other scripts, scripts can also contain descriptions for failure recovery actions, which the action interpreter can use to detect failures in action sequencing. For example, suppose the robot was about to the deliver a drink, but the patron has already left. In that case, the SERVE-DRINK-TO⟨*patron*⟩ goal cannot be met, hence the script SERVE-DRINK fails, and the recovery action might be to bring the drink back to the bar. As a result, the emotional state of "self" can be adjusted, which consequently allows for an additional context-based adjustment of goals, preferences, attitudes, and ultimately behavior. Again note that the emotional adjustments prescribed as part of the scripts are also essential for understanding the internal makeup of other agents (e.g., if the robot needs to understand what it was like for another waiter to serve a drink in vain). To allow for such adjustments of internal states, scripts can include conditional branching on evaluations of internal states (e.g., emotional states) of any entity, including self (to the extent that they are accessible). For example, observing the waiter failing at a task and knowing that failure causes anger or sadness (depending on the prior state of the agent), it is possible to infer that the waiter must be angry or sad. The action interpreter can also perform actions on such representations of internal states (e.g., modify parameters of complex emotions).[7]

Performance

Currently, the most critical performance bottleneck is the natural language processing system, given that all experimental runs in the waiter task have essentially the same outcome: *if* the robot is able to process the human request correctly (i.e., to properly recognize spoken words, parse the sentence, and build an appropriate semantic representation, which happens in about 1 out 5 runs), then it will successfully complete the task (i.e., process the request, move to the bar without collisions, get the desired drink, and return with

[7]The current implementation of the action interpreter is still somewhat impoverished, as variables for other scripts have not been implemented yet. E.g., it is not possible to add "variable actions" to scripts such "pick any script that satisfies preconditions X_i and execute it", which would cause the action interpreter to search through its scripts and match them against the preconditions X_i. Also, the current implementation only supports detection of failures, but not "recursive" attempts to recover from them ("recursive" as recovery actions might themselves fail and might thus lead to recovery from recovery, etc.).

it to the patron).[8]

Conclusion

We have presented an architecture for complex affective robots, where affective signals serve several functional roles in the overall architecture in addition to improving user interaction. A large part of the architecture has been implemented on an autonomous robot and its intended behavior has been verified over several preliminary experimental evaluations in the waiter task.

Future work on the implementation of the architecture will focus on (1) improving the language subsystem–currently the main performance bottleneck–by employing more robust voice recognition and statistical parsing components, and (2) the action interpreter by allowing for different forms of learning via special "learning scripts" that can monitor actions, record failures and successes, and use the internal affective states of the robot to adjust parameters in existing scripts (such as the numeric parameters in emotional representations) or to create new scripts based on the successful completion of tasks for which no single script exists (cp. to "chunking" in SOAR). This will eventually allow more rigorous evaluations based on systematic experiments varying architectural parameters that can assess the utility of affective control processes in a quantitative way.

Over the next few months, we expect the robotic implementation of the architecture to mature to a point where it will be ready for demonstrations of the waiter task outside our lab (in particular, we plan to demonstrate it at the upcoming 2005 AAAI robot competition).

References

Arkin, R. C., and Balch, T. R. 1997. AuRA: principles and practice in review. *JETAI* 9(2-3):175–189.

Beaudoin, L., and Sloman, A. 1993. A study of motive processing and attention. In Sloman, A.; Hogg, D.; Humphreys, G.; Partridge, D.; and Ramsay, A., eds., *Prospects for Artificial Intelligence*. Amsterdam: IOS Press. 229–238.

Blaney, P. H. 1986. Affect and memory: A review. *Psychological Bulletin* 99(2):229–246.

Bless, H.; Schwarz, N.; and Wieland, R. 1996. Mood and the impact of category membership and individuating information. *European Journal of Social Psychology* 26:935–959.

Breazeal, C. L. 2002. *Designing Sociable Robots*. MIT Press.

Byers, Z.; Dixon, M.; Goodier, K.; Grimm, C. M.; and Smart, W. D. 2003. An autonomous robot photographer. In *Proceedings of IROS 2003*.

[8] All other subfunctions such as finding people and approaching them, detecting their faces, producing natural language sentences with emotional expression, prioritizing goals based on affective value, avoiding obstacles, navigating through the environment, etc. work properly given enough time (e.g., a face might not be recognized right away, or getting around an obstacle on the way to the bar will take time.

Clore, G.; Gasper, K.; and Conway, H. 2001. Affect as information. In Forgas, J., ed., *Handbook of affect and social cognition*. Mahwah, NJ: Erlbaum. 121–144.

Ekman, P. 1993. Facial expression and emotion. *American Psychologist* 48(4):384–392.

Gat, E. 1991. Integrating planning and reacting in a heterogeneous asynchronous architecture for mobile robots. *SIGART Bulletin 2* 70–74.

Hudlicka, E. 2004. Beyond cognition: Modeling emotion in cognitive architectures. In *Proceedings of 6th International Conference on Cognitive Modeling*, 118–123.

Kahneman, D.; Wakker, P.; and Sarin, R. 1997. Back to bentham? explorations of experienced utility. *Quarterly Journal of Economics* 112:375–405.

Marsella, S., and Gratch, J. 2002. Modeling the influence of emotion on belief for virtual training simulations. In *Proceedings of the 11th Conference on Computer-Generated Forces and Behavior Representation*.

Maxwell, B. A.; Meeden, L. A.; Addo, N.; Brown, L.; Dickson, P.; Ng, J.; Olshfski, S.; Silk, E.; and Wales, J. 1999. Alfred: The robot waiter who remembers you. In *Proceedings of AAAI Workshop on Robotics*.

Mueller, E. T. 1998. *Natural language processing with ThoughtTreasure*. New York: Signiform.

Ortony, A.; Clore, G.; and Collins, A. 1988. *The Cognitive Structure of the Emotions*. New York: Cambridge University Press.

Ortony, A.; Norman, D.; and Revelle, W. forthcoming. Effective functioning: A three level model of affect, motivation, cognition, and behavior. In Fellous, J., and Arbib, M., eds., *Who needs emotions? The Brain Meets the Machine*. New York: Oxford University Press.

Pfeifer, R. 1988. Artificial intelligence models of emotion. In Hamilton, V.; Bower, G. H.; and Frijda, N. H., eds., *Cognitive Perspectives on Emotion and Motivation, volume 44 of Series D: Behavioural and Social Sciences*. Netherlands: Kluwer Academic Publishers. 287–320.

Riseberg, J.; Klein, J.; Fernandez, R.; and Picard, R. 1998. Frustrating the user on purpose: Using biosignals in a pilot study to detect the user's emotional state. In *Conference on Human Factors in Computing Systems CHI '98*, 227–228.

Scheutz, M.; McRaven, J.; and Cserey, G. 2004. Fast, reliable, adaptive, bimodal people tracking for indoor environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Scheutz, M. 2004. Useful roles of emotions in artificial agents: A case study from artifici al life. In *Proceedings of AAAI 2004*.

Simmons, R.; Goodwin, R.; Haigh, K.; Koenig, S.; and Sullivan, J. 1997. A layered architecture for office delivery robots. In *First International Conference on Autonomous Agents*, 235 – 242.

Yang, M.-H.; Kriegman, D. J.; and Ahuja, N. 2002. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(1):34–58.