Affective Agent Architectures

Matthias Scheutz

Artificial Intelligence and Robotics Laboratory

Department of Computer Science and Engineering

University of Notre Dame

Notre Dame, IN 46556, USA

mscheutz@cse.nd.edu

http://www.nd.edu/~airolab/

Overview

- What are "affective agent architectures" and why should we care about them?
- The role of affect in "complex agents"
- Challenges for "affective AI"
- Some results from experiments with simple affective agents
- Concluding thoughts
- References

university of notre dame
computer science and engineering

Affective Agent Architectures

- *Affective agent architecure* := a control architecture of an agent that has *components*, which in connection with other *internal* and *external* (i.e., *environmental*) *states* can instantiate *affective states*
- *Affective state* :=$_{\text{appr.}}$ a *positively* or *negatively valenced*, *teleological* state that the agent does or does not desire to a varying degree
- NOTE: most (if not all) "affect concepts" are *cluster concepts* (and may thus, as a cluster, not have any feature in common; no necessary and sufficient cond.)

Affect-why should we care?

- Because it may not be possible to design complex agents without affect
- Because it may be beneficial to integrate affect (e.g., "computationally cheaper" for certain tasks)
- Because complex agents may not be able to interact effectively with human otherwise
- Because there are no agents without affect in nature
- Because we want to know if we can do it
- Because the game industry wants them

The Foundational Character of Affect

- Simple organisms have no deliberative capabilities, but they are "affective" (e.g., they have simple control states that give rise to attractive-aversive behavior, "fight-or-flight" behavior)
- Complex organisms have affective states at their base and on top a complex deliberative system (which often is used to control the affective system!)
- Affect seems to be used for internal and external control!

Different Roles of Affect I

- Control function for *immediate actions* (e.g., fear triggers a run-away or freeze behavior)
- Control function for *change in short-term and long-term behavioral disposition* (e.g., anxiety leads to increased alertness, but possibly to depression and loss of interest long-term)
- Control function for change in *problem solving* (moods or "negative affect" can change between global and local processing, e.g., top-down vs. bottom up, Bless et al. 1996, Gasper & Clore 2002)

Different Roles of Affect II

- Control function in *decision making*:
  - use *affective memory* (i.e., past affective appraisal of an object, agent or event) instead of longer, more complex cognitive re-evaluation (e.g., Kahneman 1997)
  - use *affective evaluation* as an implicit measure of the likelihood of the occurrence of a positive or negative future event (e.g., implicit knowledge about events may be represented as such and thus not be directly accessible to cognitive processes; see also, affective disorders, Damasio 1994)

Different Roles of Affect III

- Control function for *social behavior*:
  - signalling behavioral dispositions is a beneficial mechansim to coordinate groups as it allows for the prediction of individual behavior (e.g., by indicating “pain”, “pleasure”, “fear”, “anger”, etc.)
  - affective approval or disapproval of own or other agents' actions (relative to norms) can trigger corrective response (e.g., “shame”, “guilt”, “pride”, “awe”, “contempt”, etc.)
  - cultural changes of innately aversive stimuli (e.g., “liking the burn of red chili peppers”, Rozin 1990)

lab
university of notre dame
computer science and engineering

Possible Roles of Affect in Agent Architectures

- *Action selection and behavior arbitration* (e.g., pick the "affectively preferred" behavior)
- *Decision making* (e.g., for choices under time pressure, "tie breaker", substitute for lack of knowledge)
- *Learning* (e.g., affective evaluations as Q values)
- *Integration* (e.g., control flow, resource management)
- *Goal processing* (e.g., arrangement and prioritization)
- *Coordinated behavior* (e.g., acting efficiently in unpredictable multi-agent enviornments)

lab
university of notre dame
computer science and engineering

Callenges for "affective AI"

- *What kinds of affective states are useful and how can they be defined?* (don't wait for psychologists!)
- *How can they be integrated into agents architectures?* (e.g., what are the architectural requirements)
- *For what kinds of tasks are they beneficial, and are there tasks for which they are necessary?*
- *Do we need "embodied agents" for affect?*
- *Do we want "affective agents"?* (e.g., McCarthy 1995)
- *How do/can we know when we have them?*

university of notre dame
computer science and engineering

Experiments comparing Affective and Deliberative Agents

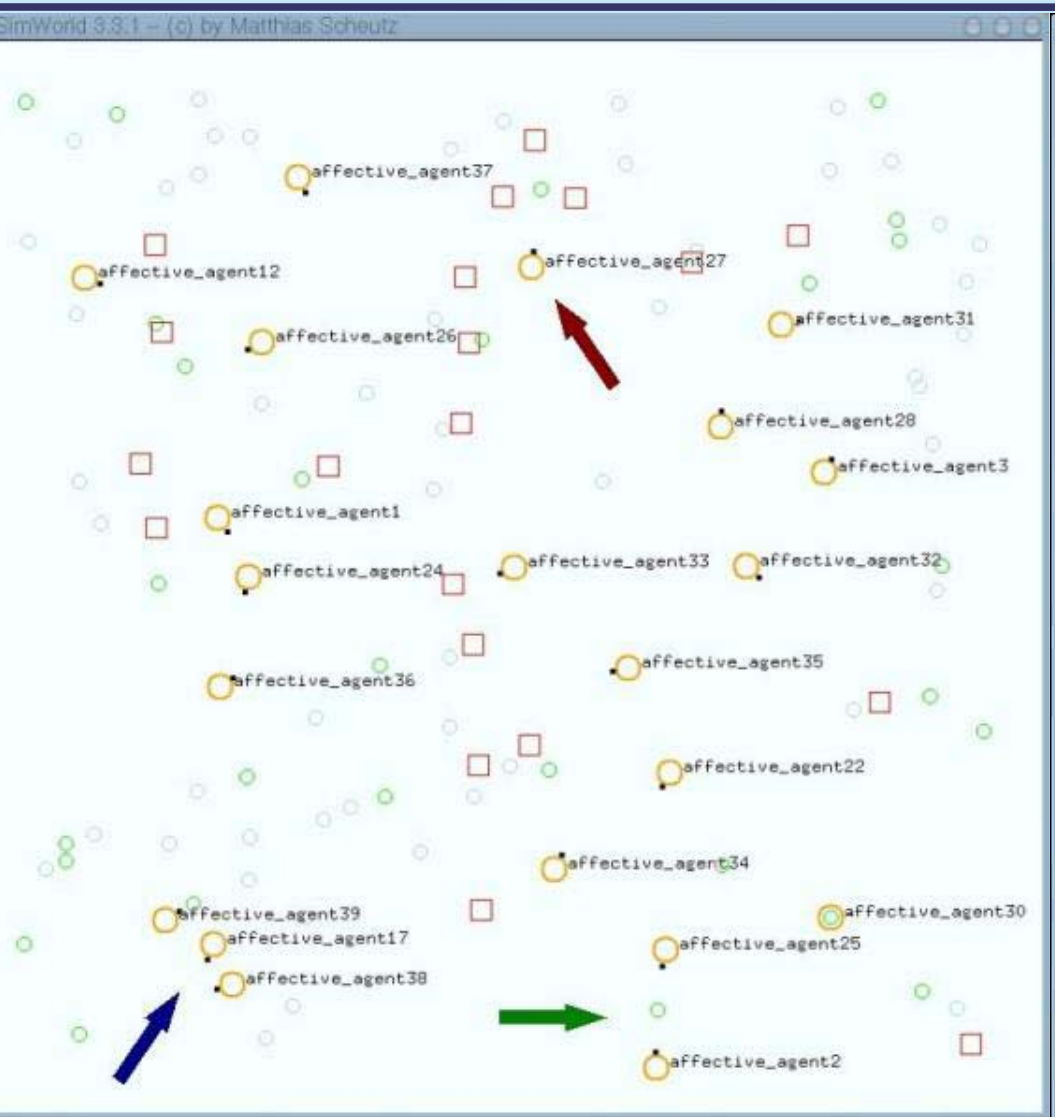
- *Idea of experimental comparison:* start simulation with different distributions of different kinds of agents in different environments
- The average number of surviving agents after a predetermined period is a *fitness measure* that can be used to assess *the relative advantage* of various architectural features of agents and their trade-offs
- *Caveat:* these kinds of experiments do not entail statements about architectural features *in general*
- (see Scheutz 2001, Scheutz and Schermerhorn 2002)

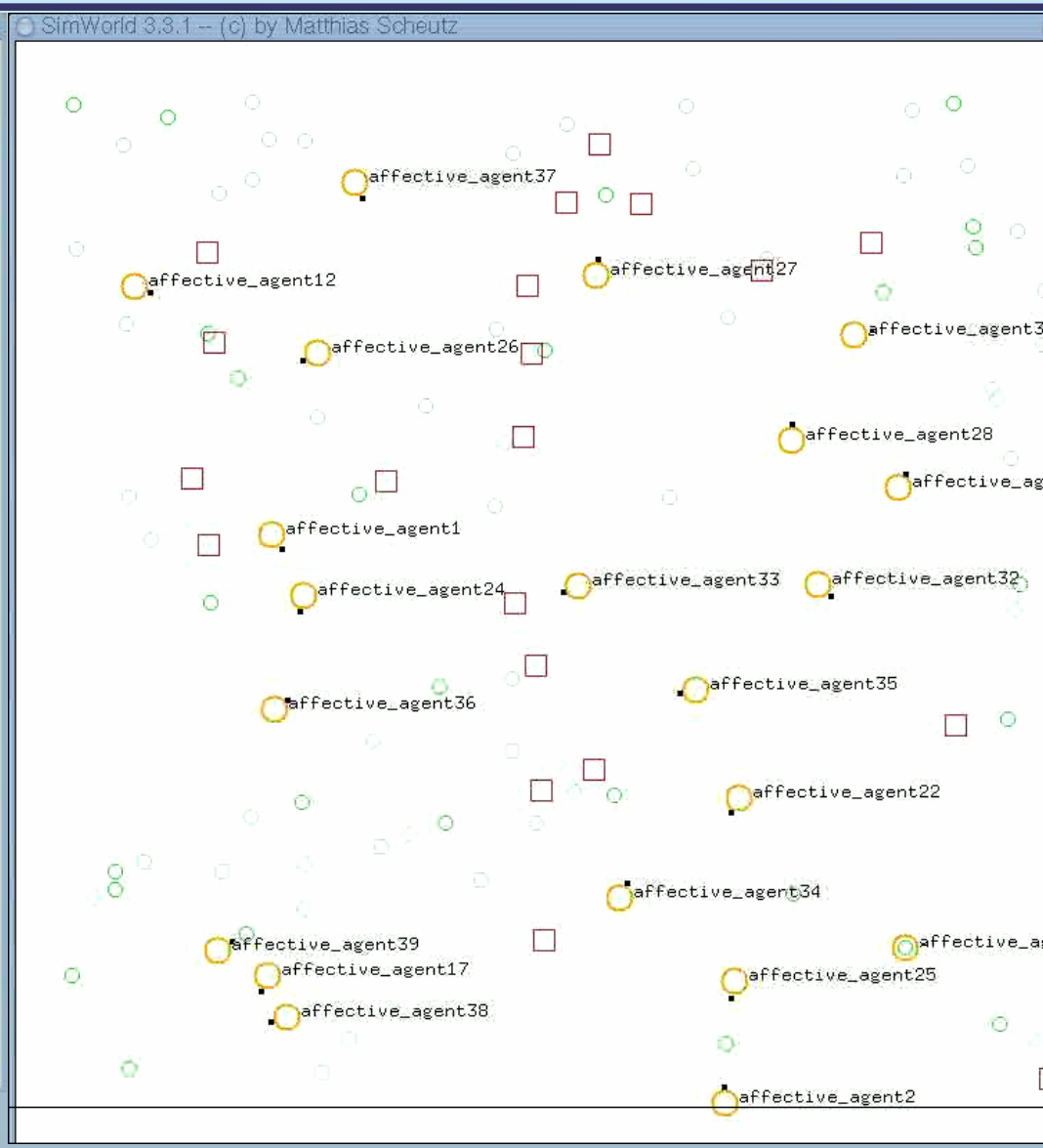
nlab
university of notre dame
computer science and engineering

Experimental Setup: a 2-Resource Foraging Task

- Experiments conducted in SIMWORLD environment
- Each experiment consists of 20 runs of the simulation (for 10000 update cycles each)
- Obstacles are placed at random locations in the environment
- "Food and water rates" are fixed
- "Procreation age" is set to 250 update cycles
- Other parameters (e.g., food energy, ingestion time, movement energy, etc.) are also fixed in advance

The SIMWOLRD Environment





nd lab
university of notre dame
computer science and engineering

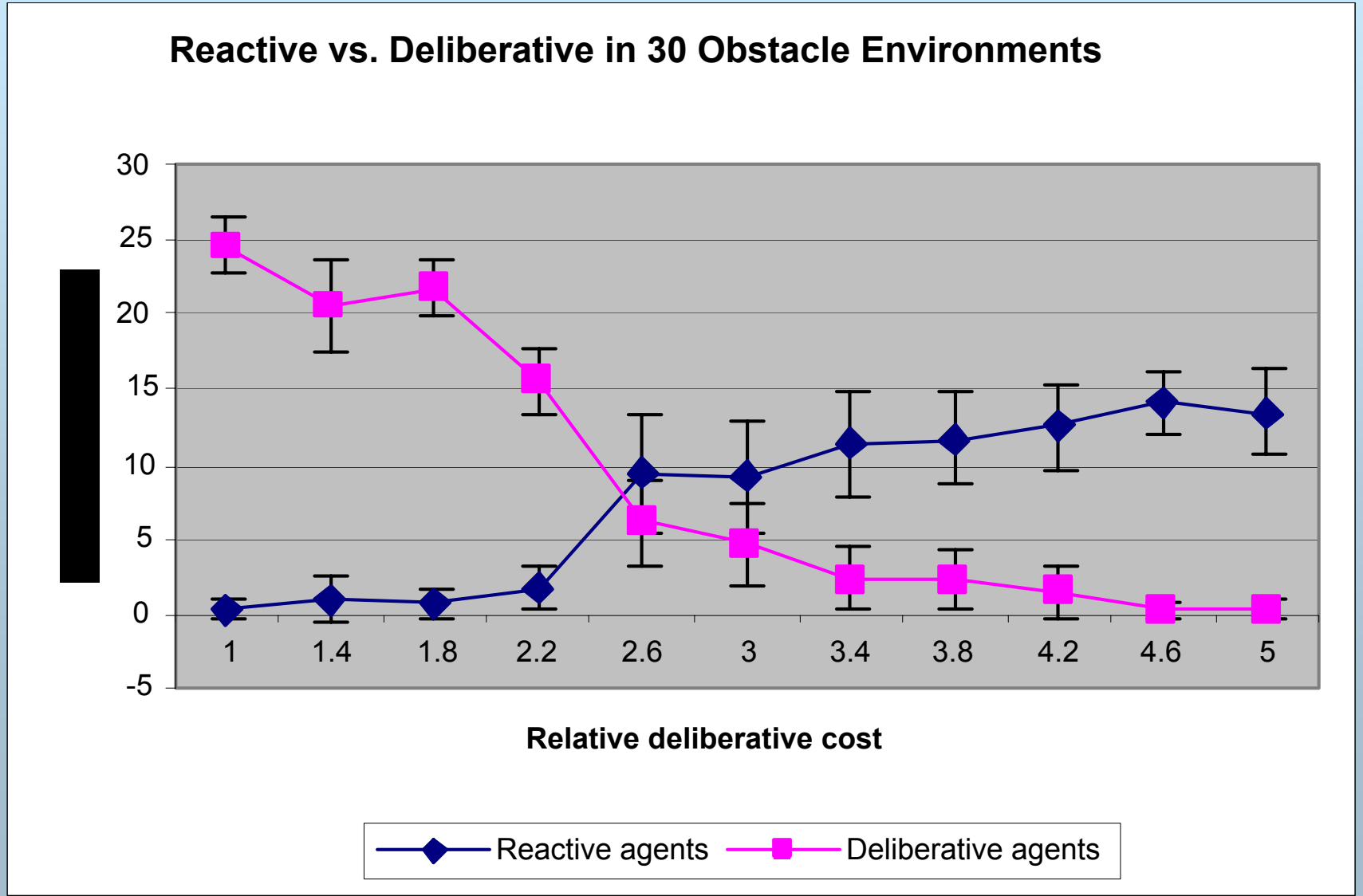
The Reactive and Affective Agents

- *Reactive* (as baseline): schema-based architecture ("greedy search")
- *Affective*: reactive + "affective extension" (i.e., control components to implement emotional control, Scheutz 2001, Scheutz under review)
- Simple fear mechanism ("fear of obstacles" and "fear of other agents" which will temporarily change the behavioral dispositions of an affective agent)
- "Need-based" foraging (through "hunger" and "thirst" states)

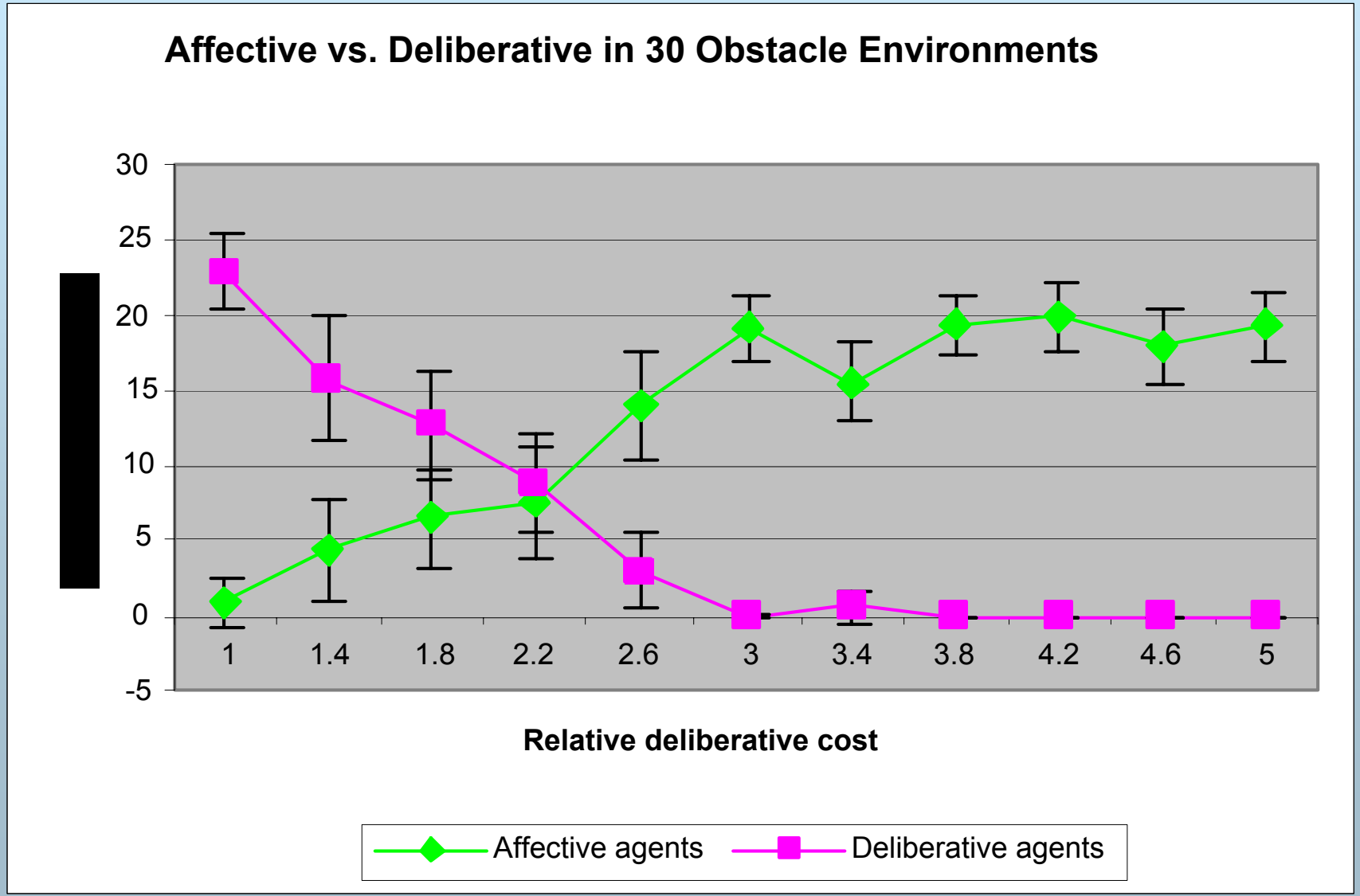
The Deliberative Agents

- (Unlimited) memory component to store location of objects in the environment
- Update mechanism for relative positions of stored entities to adjust for movements
- A* planner to compute optimal paths to resources
- Coherence mechanism to check whether locations of objects agree with perceived locations or whether goal item disappeared (which triggers re-planning)
- Replanning also triggered by closer goal item

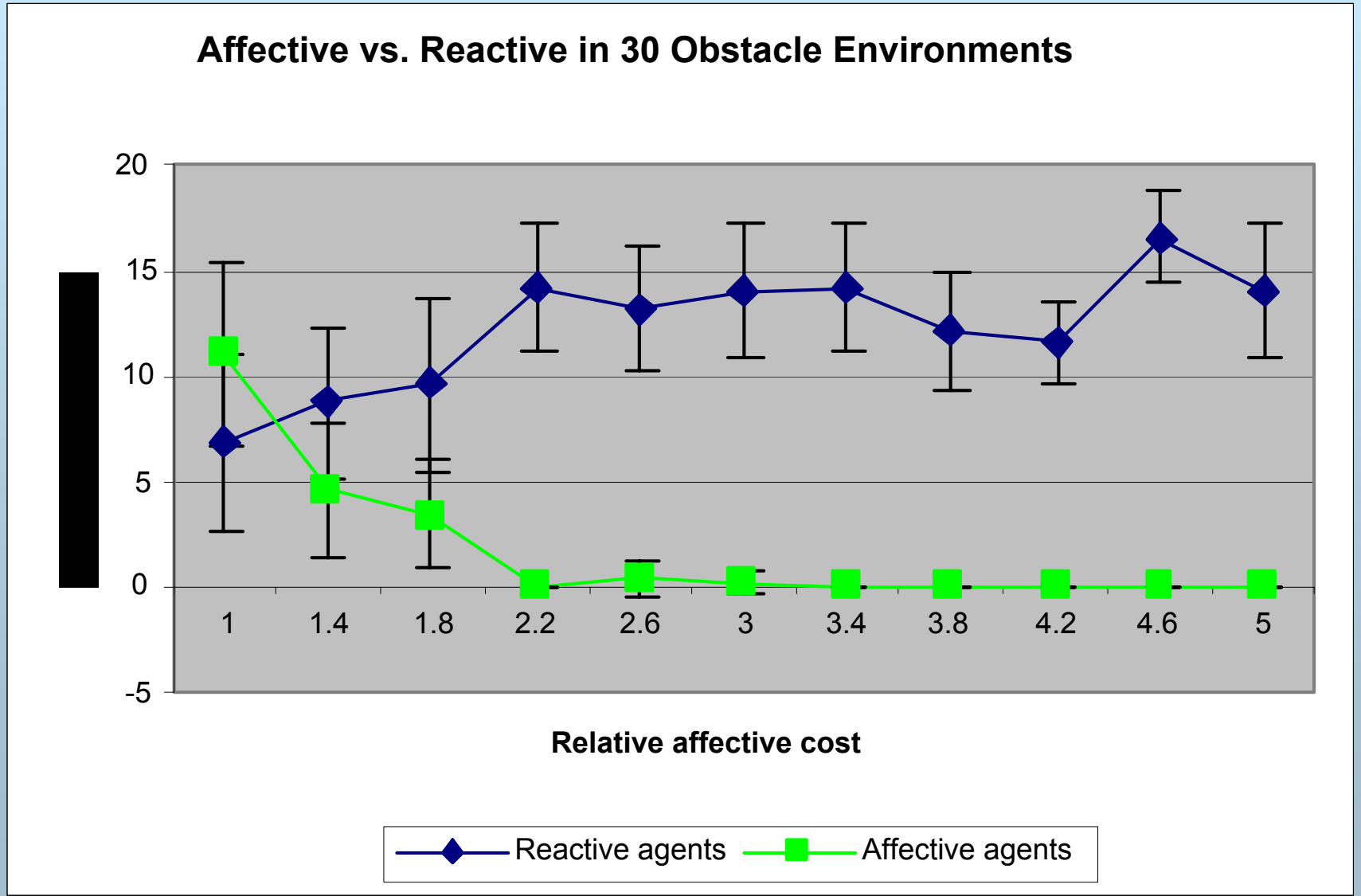
Results Experiment 1



Results Experiment 2



Results Experiment 3



Analysis of the Results

- Fitness ordering: deliberative > affective > reactive
- *Break-even points* in terms of relative cost:
  - deliberative = 3.4 * reactive cost
  - deliberative = 2.5 * affective cost
  - affective = 1.2 * reactive cost
- *But*: reactive and affective computational cost is much lower than deliberative cost (at least by a factor of 100)
- *And*: additional affective cost <20% of reactive

Concluding thoughts

- Affective control seems to be efficient for simple agents-what about complex ones?
- How can affect be utilized to improve cognition?
- How can we categorize “affect” in a way that allows for integration into (existing) agent architectures (e.g., SOAR, ACT-R, and others)?
- How can we implement and test affective mechanisms in complex agents? (e.g., what sorts of tasks would be appriopriate and challenging)

Some References to our Work on Affective Agents

- Scheutz, Matthias (2002) "Agents With or Without Emotions?". In *Proceedings of FLAIRS'02*, AAAI Press.
- Scheutz, M. and Schermerhorn, P. (2002) "Steps Towards a Systematic Investigation of Possible Evolutionary Trajectories from Reactive to Deliberativ Control Systems". In *Proceedings of Alife 8*, MIT Press.
- Scheutz, M. and Sloman, A. (2001) "Affect and Agent Control: Experiments with Simple Affective States". In *Proceedings of IAT-01*, World Sci. Publisher.
- Scheutz, Matthias (2001) "The Evolution of Simple Affective States in Multi-Agent Environments". In *Proceedings of AAAI Fall Symposium'01*, AAAI Press
- See also `http://www.nd.edu/~mscheutz/publications/`