Biomimetic Emotional Learning Agents

Samuel H. Kenyon

Northeastern University
120 Snell Engineering Center
Boston, MA 02115-5000
samkenyon@computer.org
http://flanneltron.com/bel.html

I am researching and developing a biomimetic emotional software agent system. An agent in this project is an autonomous real-time control system. A central feature is low-level emotional learning, which includes a simple knowledge base that links homeostatic/innate drives to sensory perception states, and a novel sliding-priority drive-motivation mechanism. Learning occurs in two types of training: phylogenetic (innate, genetic) and ontogenetic (lifetime). Currently the agents are tested in environment simulations, which themselves lead to many different types of experiments. These primitive agents are extensible and intended as the bottom layers (or parent classes) of planning-capable AI architectures—which are intended for practical applications not just in software but especially to be ported to embodied situated mobile robots.

The Biomimetic Emotional Learning (BEL) framework is basically reactive and behavioral at the lowest levels, but has the ability to learn and retain the learned facts/rules. BEL is partially inspired by various neurocognitive hypotheses of homeostasis, instincts, and emotion's role in survival, namely its relation to learning, reactive behavior, and higher-level decision-making, especially that of (Damasio 1994), (Damasio 1999).

BEL is also biomimetic in the way agents must be trained in phases analogous to the accumulated learning of a species during evolution, then childlike learning of an individual agent, and then continued learning during an agent's lifetime. Other emotional software agent systems have been researched (Breazeal 2003), (Vlad, Vachkov, and Fukuda 2002), (Vlad and Fukuda 2002), (Moffat and Frijda 2000), (Mochida et al. 1995), but take different approachs than me. There are also various neuromimetic projects that are trying to capture simple animal behavior and robustness by reproducing the morphology, sensors, and the nervous system (Ayers 2002), which is far more esoteric than my approach. I intend to avoid turning this system into some equivalent of prior robot control systems, such as the various three-layer architectures (Gat 1998), thus avoiding at least their limitations. Nor is BEL intended to produce anthropomorphic agents or robots with external emotional expressions—the point is to make low-level high-survivability extensible systems.

This project continues towards my intentions of practical emotional robots first displayed in (Kenyon 2003). I think a control system built from the ground up with emotional learning to develop its knowledge base and ontology, and a clearcut distinction between ontogenetic and phylogenetic training, could outperform current mobile robot systems in long-term, harsh, dynamic environments.

Robotic emotional learning has been attempted in various forms before (Fujita et al. 2001), (Gadanho and Hallam 2001), but the research has hardly been exhausted. The cognitive/learning architecture of these agents is similar to what could be the system of primitive emotions, homeostasis, and emotional learning that is the integral backbone of neurological reactive/reflexive behavior that on average promotes survival. This backbone is integrated with higher level emotions, moods, self-awareness, decision-making, planning, etc. Some parts of the BEL system will diverge from or improve on biomimetic concepts to be better suited for agents/robots. However, to get biological evolutionary advantages the artificial system has to start with a very similar core and acknowledge all the connected internal subsystems, hence this project.

In biological systems, innate survival priorities are related to homeostasis (internal management of critical processes/levels). Homeostasis is often associated with the autonomous nervous system (ANS). These homeostatic and innate survival drives give rise to motivations and actions to satisfy them, and hence stay alive. Besides internal regulation, there are also external actions to maintain internal ranges, e.g., seek shelter, stay out of extreme temperature, find food, eat food, etc.

A BEL agent has variable "sliding" priorities. Each entry in the motivation list (homeostatic) has a current intensity and follows a formula for growth/reduction over time. There can be instant promotion from stimulus—any motivation in the upper set (top three in current version) can be instantly promoted to #1 priority if stimulated by a relevant sensor state. This may be an important part of the core of higher emotions in this architecture.

The purpose of training an agent is to develop an innate knowledge base to a point where the agent has a chance of survival. Eventually I would like to further segment the ontogenetic training into "growing pains" phases, in which the agent must survive several microcosms before getting to a final "real world" at which point it is theoretically very well prepared. Some of the experiments to do are:

- What parameters make the agent live the longest in a given environment?
- Extreme, rapid, environment changes.
- Longevity with no training.
- Longevity with just phylogenetic training (allowance of *x* fatal mistakes).
- Longevity with just ontogenetic "infant" training (sandbox environment in which there are less ways to kill itself).
- Longevity with both trainings, and usefulness of learning after training for survival and goal-attainment.
- Extreme noise/errors.

Besides training various versions of agents in various simulations, future extensions/development will require the ability to assign goals among other things. Future research may include:

- Simulated lesions.
- Out-of-body agent experience, reincarnating agents, swapping "brains".
- Evolve innate emotion maps with genetic algorithms?
- Multiple agents, intentionality of objects in the world.
- Can the emotional learning systems be used for swarm intelligence?
- Sub-agents, i.e., agents within agents.
- Time shifts, perception of time, survival traits—useful for robots?
- Learning how to learn.

This framework, which at the moment consists of a growing agent class family and compatible simulation environment (all written in C++), is intended to be extended both through derivations and revisions. Some of the extensions require a lot more experimentation, such as determining the distinctions between this system's homeostasis-emotional architecture versus "background" and "primary" emotions in order to fully exploit the usefulness of low-level biological emotion systems. Related issues include a rudimentary self-awareness and development of self-analyzing mechanisms.

For practical embodied robots, the framework will need intentionality/goals assigned by outside observers, extensible to long-term missions. This will be developed first with internal plans, quick searches through consequences *k* levels deep to find the best immediate actions in reference to several competing drives/plans, and also high-level tactics.

References

Ayers, J. 2002. A Conservative Biomimetic Control Architecture for Autonomous Underwater Robots. In *Neurotechnology for Biomimetic Robots,* Ayers, J., Davis, J.L., and Rudolph, A. eds. 241-259. Cambridge, MA: MIT Press.

Breazeal, C. 2003. Cognitive Modeling for Biomimetic Robots. In *Biologically Inspired Intelligent Robots*, Bar-Cohen, Y. and Breazeal, C. eds. 253-283. Bellingham, WA: SPIE Press.

Damasio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

Damasio, A. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. San Diego: Harcourt.

Fujita, M., Hasegawa, R., Takagi, C., Yokono, J., and Shimomura, H. 2001. An Autonomous Robot that Eats Information Via Interaction with Humans and Environment. In *Proceedings 2001 IEEE 10th International Workshop on Robot and Human Interactive Communication*.

Gat, E. 1998. On Three-Layer Architectures. In *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems,* Kortenkamp, D., Bonasso, R.P., and Murphy, R. eds. Menlo Park, Calif.: AAAI Press.

Gadanho, S.C., and Hallam, J. 2001. Robot Learning Driven by Emotions. *Adaptive Behavior,* 9(1).

Kenyon, S. 2003. The Need for Emotional Architectures in Practical Robots. Draft. Available: http://flanneltron.com/papers/emotional_arch_robot_v1.pdf

Mochida, T., Ishiguro, A., Aoki, T., and Uchikawa, Y. 1995. Behavior Arbitration for Autonomous Mobile Robots Using Emotion Mechanisms. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3: 516-521.

Moffat, D.C. and Frijda, N.H. 2000. Functional Models of Emotion. In *Affective Minds*, Hatano, G., Okada, N., and Tanabe, H. eds. 59-68. Amsterdam: Elsevier.

Vlad, O.P. and Fukuda, T. 2002. Model Based Emotional Status Simulation. In *Proceedings 2002 IEEE Int. Workshop on Robot and Human Interactive Communication*.

Vlad, O.P., Vachkov, G., and Fukuda, T. 2002. Fuzzy Emotion Interpolation System for Emotional Autonomous Agents. In *Proceedings Society of Instrument and Control Engineers Annual Conference 2002*.