SOVEREIGN: A Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation System

William Gnadt[1], Stephen Grossberg[2]

Department of Cognitive and Neural Systems
Center for Adaptive Systems
Boston University
677 Beacon Street, Boston, MA, 02215
wgnadt@cns.bu.edu, steve@bu.edu

Abstract

Both animals and mobile robots, or animats, need adaptive control systems to guide their movements through a novel environment. Such control systems need reactive mechanisms for exploration, and learned plans to efficiently reach goal objects once the environment is familiar. How reactive and planned behaviors interact together in real time, and are released at the appropriate times, during autonomous navigation remains a major unsolved problem. This work presents an end-to-end model to address this problem, named SOVEREIGN: A Self-Organizing, Vision, Expectation, Recognition, Emotion, Intelligent, Goal-oriented Navigation system. The model comprises several interacting subsystems, governed by systems of nonlinear differential equations. As the animat explores the environment, a vision module processes visual inputs using networks that are sensitive to visual form and motion. Targets processed within the visual form system are categorized by real-time incremental learning. Simultaneously, visual target position is computed with respect to the animat's body. Estimates of target position activate a motor system to initiate approach movements toward the target. Motion cues from animat locomotion can elicit orienting head or camera movements to bring a new target into view. Approach and orienting movements are alternately performed during animat navigation. Cumulative estimates of each movement, based on both visual and proprioceptive cues, are stored within a motor working memory. Sensory cues are stored in a parallel sensory working memory. These working memories trigger learning of sensory and motor sequence chunks, which together control planned movements. Effective chunk combinations are selectively enhanced via reinforcement learning when the animat is rewarded. The planning chunks effect a gradual transition from reactive to planned behavior. The model can read-out different motor sequences under different motivational states and learns more efficient paths to rewarded goals as exploration proceeds. Several volitional signals automatically gate the interactions between model subsystems at appropriate times. A 3-D visual simulation environment reproduces the animat's sensory experiences as it moves through a simplified spatial environment. The SOVEREIGN model exhibits robust goal-oriented learning of sequential motor behaviors. Its biomimetic structure explicates a number of brain processes which are involved in spatial navigation.

Overview

Goal-oriented movements by animals in the natural world are necessary to obtain essential items for survival, such as food, water, shelter and mates, as well as to avoid predators. These movements are generated to satisfy internal drives, such as hunger, thirst, pain, sex and fear. To quell these drives, animals learn to approach specific objects or locations that have previously led to satisfaction of their needs while avoiding contact with obstacles and other objects that have resulted in punishment. This is known as *approach-avoidance* behavior. Within this paradigm, learning occurs via an *action-perception-cognition-emotion* cycle in which an action and its consequences elicit sensory cues that are associated with them. The success of the action hereby affects the likelihood that the same action will be repeated in the future. For example, an animal's previous positive experiences, say with food or water, may elicit approach movements and eating or drinking behaviors at a later time. On the other hand, negative experiences, such as an electric shock, may cause the animal to retreat.

The acquisition of spatial behavior has been extensively studied in rats. It is initially characterized by *reactive* movements during an exploratory phase. These movements may appear to be locally random, as the animal orients toward and approaches many local stimuli.

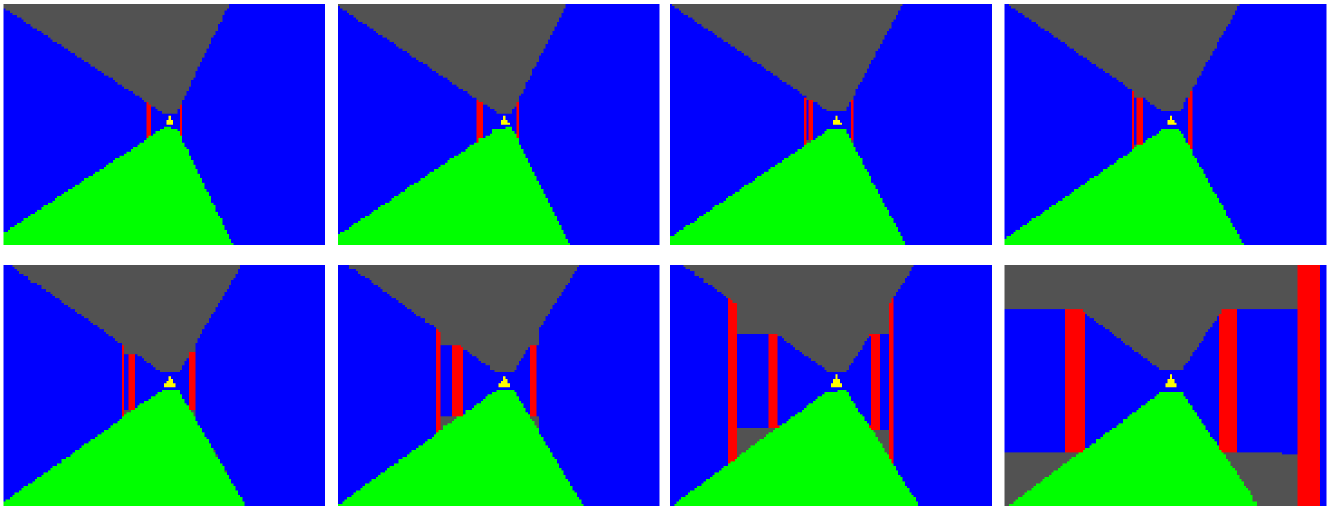

Figure 1: Snapshots from the 3-D virtual reality simulation depict changes in the scene during reactive homing toward the triangle-shaped cue.

During this period, the animal becomes familiar with its surroundings and learns to discriminate between objects likely to yield a reward and objects that are dangerous. When objects are out of direct sensory range, multiple movements may be needed to reach them. To do so efficiently, animals store complex behaviors as a plan.

Planned behavior reduces the need for exhaustive searching, or for trial-and-error efforts, both of which are metabolically expensive, by providing readout of a previously learned solution to a task. Eventually, the exploratory behavior is replaced by more direct trajectories between fixed locations within a familiar environment.

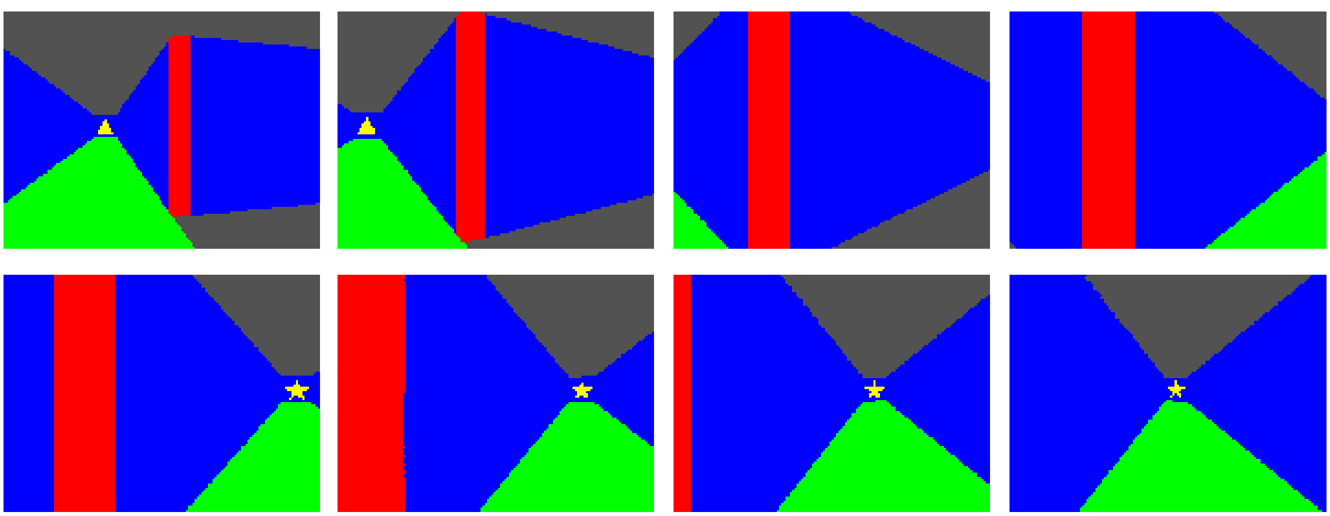

Figure 2: During reactive approach to the triangle cue, visual motion signals trigger a reactive head orienting movement to bring the star cue into view.

This research focuses on one form of animal problem-solving, namely spatial navigation and planning. The following major issues in goal-oriented spatial behavior are helpful in formulating a model of this type of problem-solving:

1. What is the role of sensory input, especially visual input, in both target identification and target position estimation?
2. What coordinate system is used for representing target positions? How is it calibrated?
3. What coordinate system is used for representing body position in space? How is it calibrated? What is the role of vestibular and proprioceptive feedback?
4. Are movement sequences composed of sensory and motor items (cue-command pairs) sufficient for some types of spatial navigation?
5. How are items within planned movement sequences computed and accumulated? How can a memory system be designed to store and recall such sequences, thereby solving the goal paradox?
6. How does reinforcement enable an animal to selectively attend to rewarded sources of sensory input?
7. How does reinforcement interact with the short-term memory representation of experienced sequences in order to differentially activate them?
8. Can latent learning, or learning without reward, be explained?
9. How can a spatial-to-motor coordinate transformation of target position be computed?
10. How can a selection mechanism be constructed for resolving conflicts between visually reactive and planned movement commands?

A description of solutions to these problems comprises the core of this project. Each element of the SOVEREIGN model contributes to showing how a freely moving and self-organizing spatial navigator, or *animat*, achieves robust learning and recall of sensory and motor sequences.

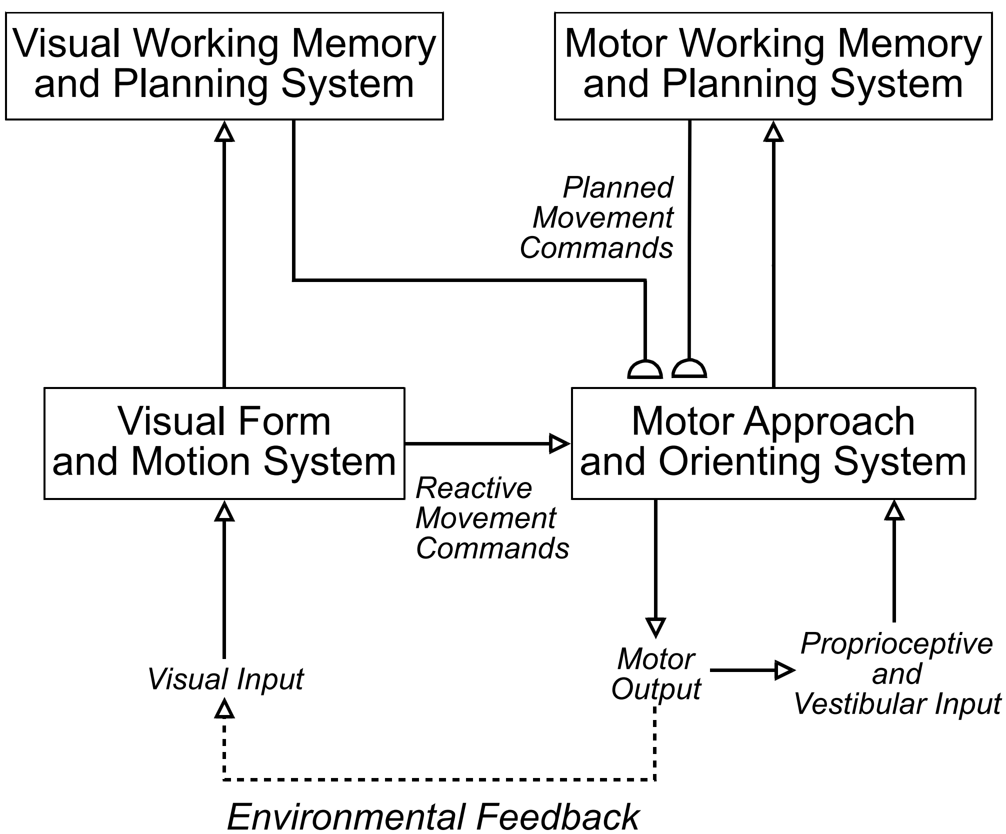

Figure 3: The SOVEREIGN flow diagram showing the convergence of reactive and planned head and body movement commands within the *Motor Approach and Orienting System*.

Model

Three key problems to which this work proposes solutions are: How can an animat learn to balance between reactive and planned behaviors in a task-appropriate way? How can previously learned plans be read-out in the correct spatial contexts and at the correct times, given that navigation may proceed in novel ways on each of an animat's movement trials? How can these problems be solved by an animat control system that uses real-time vision, incremental learned recognition, autonomous exploration, working memory storage, learned planning, reinforcement learning, and action as it navigates a novel environment? In the present study, all of these processes are modeled for an animat that moves within the relatively simple spatial format of a plus-maze. Much of the difficulty arises because of the need to understand how such an animat can self-organize behaviors that can efficiently acquire rewards, even though it may navigate the maze in different ways, including different speeds and directions of movement, on successive learning trials. In this study, a trial starts after placing the animal in the maze, at the end of one arm. The goal location, in one of the other three arms, is baited with a cue that the animal finds rewarding. By shrouding the top of the maze, only route-based visual and motor cues can be used for navigation. Only one visual cue is assumed to be visible at a time, at the end of each maze arm, from any location within the maze.

A 3-D virtual reality simulation was created which can reproduce the visual experience of an animat autonomously navigating in such a plus-maze. A sequence of images from the 3-D simulation during reactive approach toward a visual cue appears in Figure 1. As the animat approaches the choice point, a competitive struggle occurs between the salience of form cues and visual motion signals. Depending upon the balance between previous recognition and reinforcement learning associated with these form cues and motion cues that are sensitive to navigational variations that change from trial to trial, these motion cues may or may not trigger a reactive head orienting movement to the right or left at a choice point, revealing another cue at the end of an adjacent corridor. The sequence of visual scenes that are processed during a typical head orienting sequence is illustrated in Figure 2. Such alternating of approach and orienting movements are characteristic of the animat's exploration of a novel environment.

The animat's control system is split into a number of subsystems shown in the flow diagram of Figure 3. The primary input is via the visual system. Visual cues are selected and categorized via on-line incremental learning within the *Visual Form and Motion System*. Target position information in the form of reactive movement commands is then relayed to the *Motor Approach and Orienting System*, where it directs head orienting movements and body approach movements. Both visual cues and proprioceptive motor cues help the animat to determine where it is in the maze, with the proprioceptive cues capable of guiding its navigation in a familiar environment in the dark. The *Visual Working Memory and Planning System* temporarily stores and categorizes, or chunks, sequences of categorized objects and learns the corresponding approach-orient motor commands via top-down learning. Chunks are learned that are sensitive to sequences of variable lengths. This working memory operates in parallel with a *Motor Working Memory and Planning System* that temporarily stores and categorizes, or chunks, sequences of motor commands. Together these parallel Visual and Motor working memories can disambiguate decisions that only one of them would find ambiguous. When the animat receives reward, the active chunks are associated with active drives and actions. The animat is capable of learning different plans to satisfy different motivational goals. After such learning occurs, when the animat sees a familiar sensory cue under a prescribed motivational state, it can recall a motivationally-compatible plan to reach the site of previous reward. Repeated, random exploration of the environment effects a gradual transition from reactive to more efficient, planned control that leads the animat to its various motivated goals.

SOVEREIGN represents a self-consistent system synthesis of biologically-derived neural networks that have been mathematically and computationally characterized elsewhere. These include Form and Motion visual preprocessing (Grossberg, Mingolla, and Viswanathan, 2001; Raizada and Grossberg, 2003), ART fast incremental learning classifiers (Carpenter, et. al., 1992), STORE working memories (Bradski, Carpenter, and Grossberg, 1994), Masking Field self-similar multiple-scale planning fields (Cohen and Grossberg, 1986, 1987; Grossberg and Myers, 2000), CogEM cognitive-emotional circuits for reinforcement learning (Grossberg and Merrill, 1992, 1996; Grossberg, 2000), Spectral Timing circuits for adaptively timed learning (Grossberg and Merrill, 1992; Fiala, Grossberg, and Bullock, 1996), and volitional (GO) and endogenous (ERG) gain control modulation for exploratory behavior (Gaudiano and Grossberg, 1991; Pribe, Grossberg, and Cohen, 1997). We are not aware of any other autonomous agent that has yet integrated this range of self-organizing biological competences.

Acknowledgments

[1]Supported in part by the Advanced Research Projects Agency (N00014-92-J-4015), the Air Force Office of Scientific Research (F49620-92-J-0225), the National Science Foundation (IRI 90-24877), the Office of Naval Research (N00014-91-J-4100, N00014-92-J-1309, and N00014-95-1-0657), and Pacific Sierra Research (PSR 91-6075-2).

[2]Supported in part by the Air Force Office of Scientific Research (F49620-01-1-0397), the National Science Foundation (SBE-0354378), and the Office of Naval Research (N00014-01-1-0624).

References

Bradski, G., Carpenter, G.A., and Grossberg, S. 1994. STORE working memory networks for storage and recall of arbitrary temporal sequences. *Biological Cybernetics* 71, 469-480.

Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H., and Rosen, D.B. 1992. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks* 3, 698-713.

Cohen, M.A. and Grossberg, S. 1986. Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short-term memory. *Human Neurobiology* 5, 1-22.

Cohen, M.A. and Grossberg, S. 1987. Masking fields: A massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Applied Optics* 26, 1866-1891.

Fiala, J.C., Bullock, D., Grossberg, S. 1996. Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye blink response. *Journal of Neuroscience* 16, 3760-3774.

Gaudiano P. and Grossberg S. 1991. Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks* 4, 147-183.

Grossberg, S. 2000. The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences* 4, 233-246.

Grossberg, S., Mingolla, E. and Viswanathan, L. 2001. Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research* 41, 2521-2553.

Grossberg, S. and Merrill, J.W.L. 1992. A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research* 1, 3-38.

Grossberg, S. and Merrill, J.W.L. l996. The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience* 8, 257-277.

Grossberg, S. and Myers, C.W. 2000. The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review* 107, 735-767.

Pribe, C., Grossberg, S., and Cohen, M.A. 1997. Neural control of interlimb oscillations, II: Biped and quadruped gaits and bifurcations. *Biological Cybernetics* 77, 141-152.

Raizada, R. and Grossberg, S. 2003. Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex* 13, 100-113.