ITEA Best Publication of the Year Award for Two-Part Journal Article

ITEA Journal 2009; 30: 401–408

Paths for Peer Behavior Monitoring Among Unmanned Autonomous Systems

Rick Dove
Kennen Technologies, Questa, New Mexico

Aberrant behavior in unmanned weapon systems can be caused by design flaws, system malfunction, malevolent control penetration, and human error. As systems manifest more autonomous operation, unpredictable emergent behaviors add to this list. The near-term expected future calls for swarms of heterogeneous unmanned autonomous systems (UAS) to be employed on a mission. The emergent group behaviors will add new dimensions to testing, posing potentially explosive centralized monitoring and evaluation tasks with large groups. The impossibility of specifying and testing all potential situational conditions is recognized, and the safety of the testing environment itself is of concern. Lessons from social animal life show that peer behavior monitoring has evolved to detect and mitigate aberrant behavior among members, and mitigating that behavior if it is evaluated as intolerable. This article explores a foundation for peer behavior monitoring among UAS under both test and mission conditions.

Key words: Aberrant behavior; emergent social behavior; ethics; parallel pattern recognition; peer monitoring; robots; self-organizing systems-of-systems; soldiers.

Aberrant behavior in mobile unmanned autonomous weapons is likely. Regardless of the degree of autonomous control, aberrant behavior can be caused by design flaws, system malfunction, malevolent control penetration, and human error. In fully autonomous operation, unanticipated emergent behaviors are both likely and desirable in dealing with the infinite possibilities of situational reality. Simulation and test of individual units with these autonomous capabilities have their own sets of challenges and cannot predict how these units will behave in group operations. Individual behavior cannot be ignored as simulation or testing advances to group behavior and poses an explosive centralized monitoring and evaluation task with large groups.

Social animal life exhibits built-in systemic mechanisms for detecting aberrant behavior among its members, and mitigating that behavior if it is evaluated as intolerable. This article identifies a foundation for employing socially attentive monitoring in Unmanned Autonomous System (UAS) predeployment testing, and in perpetual peer evaluation after deployment. The suggested approach was instigated by studies of self-organizing systems-of-systems in a graduate systems-engineering course at Stevens Institute of Technology (Dove 2007) and deemed possible by capabilities of a new pattern detection-engine technology (Dove 2009). The foundation explored in this research was shaped by this targeted technology.

This article reports on part one of a two-part study, identifying a promising behavior detection approach that might benefit from a massively parallel pattern recognition capability. Part two of the study investigates the potential of massively parallel classification technology for leveraging the detection approaches outlined in this article.

In *The Principles of Ethics*, Herbert Spencer reaches into the animal kingdom to support his theories on the origins and enforcements of natural laws within social groups:

> *"There arises such general consciousness of the need for maintaining the limits, that punishments are inflicted on transgressors—not only by aggrieved members of the group, but by the group as a whole. A 'rogue' elephant (always distinguished as unusually malicious) is one which has been expelled from the herd: doubtless because of conduct obnoxious to the rest—probably aggressive. It is said that from a colony of beavers an idler is banished, and thus prevented from profiting by labours in which he does not join:*

> *a statement made credible by the fact that drones, when no longer needed, are killed by worker-bees. The testimonies of observers in different countries show that a flock of crows, after prolonged noise of consultation, will summarily execute an offending member. And an eye-witness affirms that among rooks, a pair which steals the sticks from neighbouring nests has its own nest pulled to pieces by the rest."* (Spencer 1893, 12–13)

Though stories of beaver and rook justice, and anecdotal witness to crow judgment and execution exist, scientific evidence is illusive; nevertheless, the values and varieties of peer judgment constraining and enforcing societal behavior are well known among humans and are studied and observed in animal (Flack et al. 2006) and insect societies (Heinze 2003, Monnin et al. 2002, Ratnieks, Foster, and Wenseleers 2006).

This article suggests that peer evaluation of behavior is necessary and valuable in the domain of autonomous unmanned systems when they are working together as a team on a warfighting mission, and perhaps even more so when these systems are being tested, as they are less likely to be well behaved. The suggestion is prompted by the positioning and planning for an Unmanned Autonomous System Test (UAST) focus area described in a 2008 Broad Area Announcement:

> *"Due to the mobility inherent in all UAS, their close proximity to humans (e.g., soldiers, testers, population centers, etc.) and their capability for unpredictable behavior; a reliable fail-safe system is needed. This effort seeks technologies for all aspects of system safeties as they pertain to UAS, Systems of Systems, and Complex Systems. This includes safe test conduct, testing for top level mishaps, safety fail-safes, truth data assessment for safety, and safeties associated with large numbers of collaborating UAS."* (Office of the Secretary of Defense 2008, 21)

It is also recognized that testing outcomes can have "an almost infinite number of possibilities, depending on UAS cognitive information processing, external stimuli, operational environment, and even possible random events (hardware/software failures, false stimuli, emergent behavior, etc.)." (Office of the Secretary of Defense 2008, 23)

Emergent behavior is later recognized as something less than random:

> *"UAS formation control, swarming, and aggregate intelligent agent behavior are an emergent characteristic of this technology arena. … System behavior, in a multi-agent system, can be difficult to predict and often unexpected system behaviors occur which lead to poor system performance. These unexpected system behaviors result from unforeseen group actions of agent groups and agent-group behavior that is not directly coded by the agent designers."* (Office of the Secretary of Defense 2008, 54–55)

Such unexpected system behaviors can be good as well as bad. In fact, the goal of fully autonomous intelligent behavior is creative problem-solving in situations without precedence. It is unlikely that unleashing a swarm of UASs that are only capable of dealing with well-defined cataloged situations will be effective.

We cannot know the situations that will arise, nor can we directly control how things should play out. Instead, we must recognize and embrace uncertainty within a framework of governance principles that will bound the outcomes within an acceptable space. The principle described in this article classifies behavior as unacceptable based on absolute boundary-infraction recognition, rather than attempts at imperfect reasoning or restrictions to specifically approved behaviors.

UAS will necessarily be tested and fielded in situations that have no precedence in cataloged responses. How will we constrain the outcomes to those we can live with? More to the point, how will we detect unacceptable behavior in time to intervene if unacceptable consequences are the likely outcome?

Dove and Turkington (2009) characterize agile systems as class 1 if they are (operator) reconfigurable and class 2 if they are (systemically) reconfiguring. It can be useful to think of a class-1 UAST system testing class-2 UAS systems. This agile-system class distinction arose from a graduate course in the School of Systems and Enterprises at Stevens Institute of Technology (Dove 2007). The course reviews the literature in various bodies of knowledge relevant to self-organizing systems, and challenges collaborative student analysis to identify recurring and necessary patterns across bodies of knowledge. Five cycles through to date, this investigation is beginning to yield some promising fundamental patterns. One in particular is the genesis of this article's focus: successful social systems often exhibit a pattern of peer behavior-enforcement arising when the stability of the social system is put at risk.

A related body of work led by Ronald Arkin (2007) at Georgia Institute of Technology is concerned with

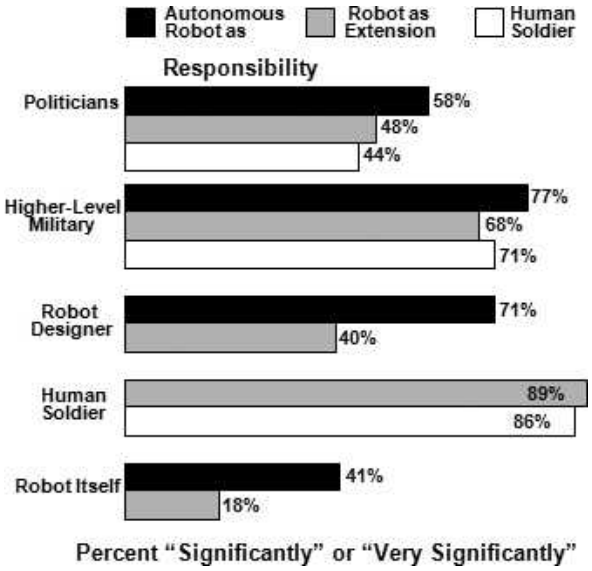

Figure 1. Responsibility for lethal errors by responsible party based on a survey of 430 respondents from demographic groups. The soldier was found to be the most responsible party, and robots the least. (Reproduced with permission from the survey reported in Moshkina and Arkin 2007.)

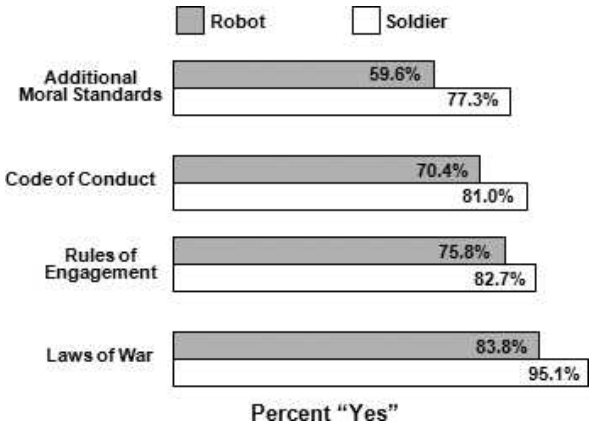

Figure 2. Ethical behavior for soldiers and robots. Applicability of ethical categories is ranked from more concrete and specific to more general and subjective. (Reproduced with permission, from the survey reported in Moshkina and Arkin 2007.)

ethical behavior of UAS used in military operations and recognizes the potential for peer monitoring: "When working in a team of combined human soldiers and autonomous systems, they have the potential capability of independently and objectively monitoring ethical behavior in the battlefield by all parties and reporting infractions that might be observed."

A recent survey investigated opinions about the use of, and responsibilities for, lethal autonomous systems among four demographic groups (Moshkina and Arkin 2007). A total of 430 respondents were distributed demographically as 54% robotics researchers, 30% military, 16% policymakers, and 27% general public. *Figure 1* depicts who the respondents felt was responsible when behavior was unacceptable. Our interest here is in the autonomous devices, not the "robot as extension" case, in which a human directs the unmanned system. Interesting to note: lethal mistakes made by a UAS are blamed on higher-level military, UAS designers, and politicians, in that order.

The survey showed that all four demographic groups want ethical standards for UAS to be considerably higher than those for soldiers. *Figure 2* shows how the groups felt about specific constraints that should be enforced. Monitoring for ethical behavior infractions is a subset of what must be monitoring for safe and secure behavior overall.

Two sections follow that discuss social peer-behavior monitoring; first in terms of temporal relationships and then in terms of spatial relationships. Temporal behavior monitoring compares the temporal event sequence of the intended tactical plan against the actual sequence of events, on an agent-by-agent basis. Spatial behavior monitoring compares the spatial aspects of the intended plan against actual trajectories on an agent-by-agent basis.

Temporal behavior leverage

Our fundamental interest is in the ability to detect and evaluate certain aspects of the behavior of team members as they work toward a common goal. This common goal may encompass a set of tasks that are not necessarily shared by all team members but are nevertheless a part of the activities pursuing common goal achievement. Task plans for achieving this common goal will have constraints. For instance, the end may not justify any possible means. Achievement may also have constraints on team member behavior (e.g., team members are expected to work toward team goals according to an established coordination plan).

A team is defined as a collection of members (agents) working toward a common goal. Working together implies some form of activity coordination. Coordination comes in a range of forms from centralized planning and micro-direction of the agents at the one extreme, to mindless local-reaction agent-behaviors resulting in emergent swarm effects at the other extreme. Our interest is in neither extreme, but rather with autonomous agents that possess and employ both self-awareness and social awareness of other team members and their behaviors as they jointly pursue a mission. Agents will have a sense of team and a sense of mission, and use this information to detect when another agent is clearly not behaving in the team's interest.

In a broader socially aware sense, the team's interest includes the team's image among outsiders—a weap-

ons-toting team member gone rogue can impair the team's long-term existence likelihood. This is a new behavior focus not seen in prior research of Multi-Agent Systems (MAS).

Malone and Crowston (1994) in a broad interdisciplinary survey of the "*emerging research area, sometimes called coordination theory,*" define coordination as "*managing dependencies between activities.*" Noting the onslaught of the electronically connected world, they proposed that new forms of organizing and new forms of coordination structures would emerge. They also observed that different disciplines were already exploring domain-specific coordination concepts, and that there was now value to be gained in finding domain-independent underlying principles of coordination. Their stated intent was to help kick-start the development of a *theory of coordination* by illuminating these cross-discipline similarities, noting that "*It is not enough just to believe that different systems are similar, we also need an intellectual framework for 'transporting' concepts and results back and forth between the different kinds of systems.*" This idea is germane to the present discussion as this article suggests that coordination concepts of social systems inform how we deal with aberrant behavior in UAS.

About the same time as Malone and Crowston pulled together their survey, Jennings (1993) modeled coordinated agent communities on a foundation of commitments and conventions. Jennings defines commitments as mutually agreed upon plans of action, and conventions as the means for monitoring commitments under changing circumstances. He goes on to suggest that all coordination mechanisms can be seen as joint commitments and their related social monitoring conventions. Jennings acknowledges that it is infeasible in any community of reasonable complexity for total monitoring to occur, due to communication bandwidth and processing time. In our own human experience we see this to be true in teamwork, where some awareness of other team-member activity is naturally maintained, but any attempt at totally detailed and continuous monitoring knowledge is impossibly overloading and counterproductive.

Important to our monitoring interests, Jennings shows why the behavior (alone) of a collection of agents as seen by an outside observer is insufficient to determine if coordination is present, and concludes that "*coordination is best studied by examining the mental state of the individual agents.*" He then goes on to say: "*The exact make up of this mental state is still the subject of much debate, however there is an emerging consensus on the fact that it contains beliefs, desires, goals and commitments (intentions).*"

Jennings raised issues that are addressed in Gal Kaminka's Ph.D. thesis (2000) and related publications (Kaminka and Tambe 1997, 1998, 2000). Kaminka pursued what Jennings dubbed the *social conventions* aspect and developed a "mental state" representation based on goal hierarchies presumably shared by a team of agents—recognizing that some agents may have tasks different than others and some may choose to achieve a common task differently than others. Notably his work features primary examples of unmanned autonomous (aerial) systems, where individual UAS (agents) monitor and recognize when a member of the team doesn't behave as mutually agreed to in the plan. Kaminka's approach also enables an agent to detect self-failure often but not always.

Kaminka credits inspiration to Leon Festinger's seminal work on *social comparison theory* (Festinger 1954), which is founded on the hypothesis that humans have a drive to evaluate their own opinions and abilities and will employ a social comparison process if this evaluation cannot be tested directly.

Kaminka (1997) kicks off this path of work by proposing an approach to failure detection which he called unique to a multi-agent setting: the key idea being that agents observe each other and use that information to inform themselves about the situation and about the appropriateness of the behaviors of self and others. Basically each agent evaluates its own behavior by observing that of others, and comparing those observations with its own behavior, beliefs, goals, and plans.

Kaminka's early tack had Festinger's self-centered focus: agents used cues from others to evaluate their own fitness. Subsequently, his investigations broadened to both a centralized agent that could monitor team and other-agent behavior, and multiple agents monitoring team and other-agent behaviors.

Kaminka makes the case in his thesis for distributed monitoring and detection, showing that a centralized monitor using his algorithms does as well as can be done, whereas multiple monitor/detectors among socially aware agents do best, as they can exploit their own local state as part of the information. He shows that the centralized approach can provide either sound (no false positives) or complete (no false negatives) results, whereas the decentralized approach provides both sound and complete results—meaning no incorrect detections and no missed detections. He also shows that this can be accomplished without any one agent monitoring all the agents, and without all the agents having this monitoring capability.

Recent work is getting even closer to the detection of threatening aberrant behavior. Avrahami-Zilber-

brand and Kaminka (2007) extend the social comparison concept into the detection of suspicious behavior by an agent. The general approach is to monitor a large group of agents and note that one or some agents are behaving decidedly different than expected. Two types of suspicious behavior recognition are employed: explicit and implicit. Explicit recognition classifies behavior as suspicious if it reflects a reference pattern known to be suspicious. Implicit recognition classifies behavior as suspicious if it does not conform to cataloged reference patterns of "normal" behavior. This work is part of a more general interest in dynamic tracking of multi-agent teams.

Sviatoslav Braynov (Braynov 2004, Braynov and Jadliwala 2004) has investigated the use of coordination graphs built from filtered action data to recognize coordinated behaviors among multiple agents maliciously working toward an undesirable goal. This is done by an aberrant behavior detector examining forensic data, with suggestions that real-time log-data examination might recognize a coordinated attack in early stages of setup and initiate counteraction. This approach may be useful for identifying the agents, actions, and situational conditions that participate in the manifestation of an emergent behavior. Proactively, such emergent behaviors that are determined to be undesirable could thereafter become recognizable patterns that generate increasing states of concern as manifestation of the conditions increases.

In summary, the cited works in this article bring the concepts of social awareness into play with good effect for detecting behaviors not in keeping with team goals, agent tasks, and coordination plans. Separate paths by Kaminka and Braynov are beginning, respectively, to attack computational scaling issues and the detection of coordinated alien activity within groups.

Spatial behavior leverage

Stephan Intille (1999) opened an interesting path that explored *visual recognition of multi-agent action*. His work analyzed films of American football games and identified the plays being made according to visual analysis of the trajectories of the players, matching the offense player trajectory's against the team's playbook patterns. There is a considerable difference between idealized chalk-board play patterns *(Figure 3)* and actual game-time trajectories given the unpredictability of the 11 defensive-team players, as well as the infinite variety of trajectories the 11 offensive-team players may take in reaction to defensive play. Yet he built a system that could recognize appropriate single-agent and multi-agent actions in this domain under "noisy" trajectory data of player and ball movements.

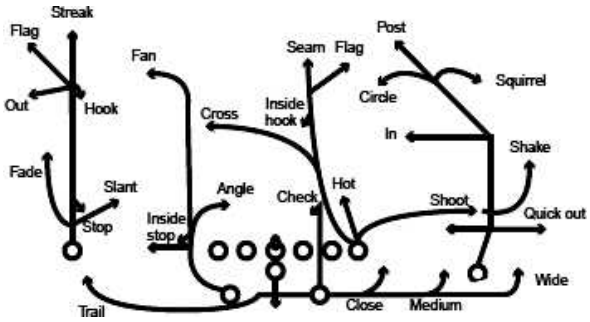

Figure 3. The different types of pass patterns that receivers can run constrained by the rules and nature of the game. (Reproduced with permission, from Intille 1999.)

"Recognizing Planned, Multiperson Action" (Intille 2001) is a mature digestible presentation of his work covering approximately 6 years. He has focused on plan recognition, attempting to identify the play by classifying the observed actions, movements, and spatial relationships of the players. Intille notes certain aspects of American football and the nature of its team interaction that shape the recognition approach:

- Infrequent collaborative replanning—though agents adjust their playing to fit the real-time situation, the intended play coordinates the general action.
- Agent based goal description—each agent has a goal (e.g., catch pass, block, receive handoff, etc.) for any given play. The system attempts to identify the goals the agents are pursuing based on spatial and temporal relationships of agents and their trajectories.
- Existing taxonomies—a common, fairly universal terminology exists among coaches and fans for describing all low-level agent actions (e.g., blocking, running through a hole) and higher level aggregated actions (e.g., executing a specific play)—within the boundaries of game constraints and experience. This forms a succinct and closed domain language. *Figure 3* shows some of that common terminology and the nature of its reference.
- Purposeful behavior—every agent is expected to contribute to the play's objective, nothing happens without a reason.
- No statistical databases—large statistical databases describing much of what has transpired in football action do not exist. A recognition algorithm cannot be based on that type of resource. Instead, a linguistic description of plays is provided by a domain expert (coach).

- Structured but uncertain—each offensive play begins as a highly structured coordination plan. The defensive agents rarely cooperate, so a great deal of variation exists in individual agent movements, individual agent goal achievement, and overall trajectory maps.

This model has potential for describing joint UAS maneuver patterns and detecting when an agent is not contributing as planned. Though a great deal of latitude is expected in the execution of a maneuver pattern, general characteristics should prevail and indicate an individual UAS not working on team behalf. Judgment of cause and severity for out-of-scope behavior is a separate issue not dealt with here.

Concluding remarks

This article reported on the first part of a two-part study, instigated by a technology for massively parallel pattern recognition and studies in self-organizing systems-of-systems. This first part identified a research base for aberrant behavior detection in multi-agent systems that might benefit from a massively parallel pattern recognition capability. The second part of the study, to be published shortly, will indicate how the detection approaches outlined in this article might be implemented advantageously in a massively parallel classification technology.

The temporal and spatial discussions in prior sections are complementary, each having a potential role in a socially attentive solution platform. It is not suggested that the research referenced here is the only way to approach the problem effectively, or that what was presented is completely sufficient, but rather that this basis appears promising as a foundation for a solution path worth exploring.

It is likely that the future of UAS is pervasive employment in human society, regardless of purpose, warfighting or otherwise. Such "things" will need to be socialized, as do the children of all species. Simple behavior safeguards will not be sufficient. Right or wrong, ready or not, we will expect these things to exhibit respect for life and property, ethics, self-control, and peer-policing capabilities approaching our own. To the extent that they don't, we will object to their presence.

In a test environment, especially in early years, as well as later with the presence of legacy units, such detection mechanisms are not likely to be present on board. The physical location of these mechanisms is not necessarily important, provided suitable sensor data are available. Under testing conditions, the testing arena is likely bounded and populated with various observer installations and mobile facilities. Suitable sensors located in these facilities, and perhaps sent from transmitters on board UAS, can provide the raw data feeds. As was shown by Kaminka (2000), it is not necessary to have a one-to-one ratio of monitors to agents in order to ensure high detection accuracy. Thus, multiple such mechanisms might be located in testing and observation facilities in suitable proximity to the testing arena. Alternatively, special units could be deployed among the UAS under test much as many field sports employ referees on the field.

In the end, such mechanisms also belong on board as an integral part of every UAS, as UAS will operate outside of ready observation and are subject to attrition by enemy destruction. In such live cases, aberrant behavior must be detected and evaluated to sense control penetration by the enemy as well as malfunction that threatens the mission or might provide a disabled UAS to the enemy for later recovery.

This concept of a socially aware team of autonomous agents has application well outside the UAS and UAST focus of this discussion. For instance, socially aware security agents can be employed as a community watch among networked groups of computers or sensors, keeping watch on each other. For another instance, Braynov (2004) is investigating ways in which coordination graphs can be employed in the recognition of coordinated attacks by groups of autonomous agents working toward a common goal. The platform suggested here for UAST can merge with Braynov's work to pursue application in intrusion detection areas. ❑

Rick Dove is a founding partner of Kennen Technologies, an adjunct professor in the School of Systems and Enterprises at Stevens Institute of Technology, and Chairman of Paradigm Shift International. He is author of Response Ability—The Language, Structure, and Culture of the Agile Enterprise*; and* Value Propositioning—Perception and Misperception in Decision Making. *He is an active researcher and consultant in agile and self-organizing systems. He has 40 years experience at start-up, turn-around, and interim executive management. He holds a bachelor's degree in electrical engineering from Carnegie Mellon University and has done graduate work in Computer Science at the University of California, Berkeley. E-mail: rick.dove@kennentech.com*

References

Arkin, Ronald C. 2007. Governing lethal behavior: Embedding ethics in a hybrid deliberative/reactive robot architecture. Technical Report GIT-GVU-07-

11. Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, Atlanta, Georgia. http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf (accessed May 15, 2009).

Avrahami-Zilberbrand, Dorit, and Gal A. Kaminka. 2007. Towards dynamic tracking of multi-agent teams: An initial report. In *Proceedings of the AAAI Workshop on Plan, Activity, and Intent Recognition*, July 23, Vancouver, Canada, 17–22. Menlo Park, California: AAAI Press. www.cs.biu.ac.il/~galk/Publications/Papers/pair07.pdf (accessed May 15, 2009).

Braynov, Sviatoslav. 2004. A model for detecting malicious groups. *3rd European Conference on Information Warfare and Security*, June 28–29, London, UK, 21–30. Reading, UK: Academic Conferences International.

Braynov, Sviatoslav, and Murtuza Jadliwala. 2004. Detecting malicious groups of agents. In *Proceedings the First IEEE Symposium on Multi-Agent Security and Survivability*, August 29–31, Philadelphia, Pennsylvania, 90–99. Piscataway, New Jersey: IEEE.

Dove, Rick. 2007. *Self-organizing agile systems and enterprises: Architectural patterns enabling self-organizing systems-of-systems* (course binder for SDOE-683). Hoboken, New Jersey: Stevens Institute of Technology, School of Systems and Enterprises.

Dove, Rick. 2009. Pattern recognition without tradeoffs: Low-cost scalable accuracy independent of speed. In *Proceedings Cybersecurity Applications and Technologies Conference for Homeland Security (CATCH)*, March 3–4, Washington, D.C., 255–260. Piscataway, New Jersey: IEEE. www.kennentech.com/Pubs/2009-PatternRecognitionWithoutTradeoffs.pdf (accessed May 15, 2009).

Dove, Rick, and Garry Turkington. 2009. On how agile systems gracefully migrate across next-generation life cycle boundaries. *Journal of Flexible Systems Management* 10 (1): 17–26.

Festinger, Leon. 1954. A theory of social comparison processes. *Journal of Human Relations* 7 (2): 117–140.

Flack, Jessica C., Michelle Girvan, Frans B. M. de Waal, and David C. Krakauer. 2006. Policing stabilizes construction of social niches in primates. *Nature* 439 (7075): 426–429.

Heinze, Jürgen. 2003. Reproductive conflict in insect societies. In *Advances in the Study of Behavior*, ed. P. Slater and J. Rosenblatt 34: 1–57. New York: Academic Press.

Intille, Stephen S. 1999. Visual recognition of multi-agent action. Ph.D. dissertation, School of Architecture and Planning, Massachusetts Institute of Technology. http://web.media.mit.edu/~intille/papers-files/thesis.pdf (accessed May 15, 2009).

Intille, Stephen S. 2001. Recognizing planned, multiperson action. *Computer Vision and Image Understanding* 81 (3): 414–445. http://web.media.mit.edu/~intille/papers-files/cviu01.pdf (accessed May 15, 2009).

Jennings, Nick. 1993. Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review* 8 (3): 223–250. https://eprints.kfupm.edu.sa/30491/ (accessed May 15, 2009).

Kaminka, Gal A. 2000. Execution monitoring in multi-agent environments. Ph.D. dissertation, University of Southern California. www.isi.edu/soar/galk/Publications/diss-final.ps.gz (accessed May 15, 2009).

Kaminka, G. A., and M. Tambe. 1997. Towards social comparison for failure detection. In *Proceedings Fall Symposium on Socially Intelligent Agents*, November 8–10, Cambridge, Massachusetts, 64–66. Menlo Park, California: American Association for Artificial Intelligence.

Kaminka, G. A., and M. Tambe. 1998. What is wrong with us? Improving robustness through social diagnosis. In *Proceedings of the 15th National Conference on Artificial Intelligence*, July 26–30, Madison, Wisconsin, 97–104. Menlo Park, California: American Association for Artificial Intelligence.

Kaminka, G. A., and M. Tambe. 2000. Robust multi-agent teams via socially-attentive monitoring. *Journal of Artificial Intelligence Research* 12: 105–147. http://teamcore.usc.edu/papers/2000/jair-socmon.pdf (accessed May 15, 2009).

Kaminka, G. A., D. V. Pynadath, and M. Tambe. 2001. Monitoring deployed agent teams. In *Proceedings of the Fifth International Conference on Autonomous Agents*, May 28–June 1, Montreal, Canada, 308–315. New York: ACM. www.cs.biu.ac.il/~galk/Publications/Papers/agents01.pdf (accessed May 15, 2009).

Malone, Thomas, and Keven Crowston. 1994. The interdisciplinary study of coordination. *ACM Computing Surveys* 26 (1): 87–119. www.cs.unicam.it/merelli/Calcolo/malone.pdf (accessed May 15, 2009).

Monnin, T., F. L. W. Ratnieks, G. R. Jones, and R. Beard. 2002. Pretender punishment induced by chemical signaling in a queenless ant. *Nature* 419 (6902): 61–65.

Moshkina, Lilia, and Ronald C. Arkin. 2007. Lethality and autonomous systems: Survey design and results. Technical Report GIT-GVU-07-16. Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, Atlanta, Georgia. www.cc.gatech.edu/ai/robot-lab/online-publications/MoshkinaArkinTechReport2008.pdf (accessed May 15, 2009).

Office of the Secretary of Defense. 2008. *Unmanned and Autonomous System Testing (UAST) Broad Agency Announcement <BAAUAST0002>*. U.S. Department of Defense, Issue Date April 4, 2008.

Ratnieks, F. L., K. R. Foster, and T. Wenseleers. 2006. Conflict resolution in insect societies. *Annual Review of Entomology* 51: 581–608. www.people.fas.harvard.edu/~kfoster/RatnieksetalAnnreventomol2006.pdf (accessed May 15, 2009).

Spencer, Herbert. 1893. Sub-human justice. In *The Principles of Ethics*, Volume 2, 8–16. New York: D. Appleton and Company.

Acknowledgments

This work was supported in part by the U.S. Department of Homeland Security award NBCHC070016. *Figures 1 and 2* are included with permission from Ronald Arkin. *Figure 3* is included with permission from Stephen Intille.

ITEA Best Publication of the Year Award for Two-Part Journal Article

ITEA Journal 2009; 30: 504–512

Methods for Peer Behavior Monitoring Among Unmanned Autonomous Systems

Rick Dove
Kennen Technologies, Questa, New Mexico

As multi-agent weapon systems manifest more autonomous operation and work in teams, unpredictable emergent behaviors will occur in both individual agents and in teams. These group behaviors will add new dimensions to testing, posing potentially explosive centralized monitoring and evaluation tasks with large groups. The impossibility of specifying and testing all potential situational conditions is recognized, and the safety of the testing environment itself is of concern. A prior companion article explored a socially attentive foundation for peer behavior monitoring among Unmanned Autonomous Systems (UAS) under both test and mission conditions. This article explores certain complexity issues of aberrant behavior detection in UAS, draws parallels with human cognition capabilities, and provides a technology foundation for massively parallel behavior-pattern detection.

Key words: Aberrant behavior; control penetration; expertise; parallel recognition capabilities; pattern recognition; peer-peer socially attentive monitoring; unmanned autonomous systems; very-large-scale integrated (VLSI) circuit.

The trend toward increased autonomy in unmanned weapon systems has raised concerns about methods for testing these devices both individually and in tactical group maneuvers (DoD 2008). Increased autonomy is generally enabled and permitted by increased intelligence of the artificial kind in Unmanned Autonomous Systems (UAS). Intelligent systems, be they human or artificial, exhibit behaviors in response to situational conditions. Situational conditions are unpredictable and infinite in potential variety, leading to emergent behaviors at both the individual and group level. For UAS in warfighting, emergent behavior is necessary and desirable when it is appropriate and useful, and potentially a major problem when inappropriate.

Range testing can never duplicate the situational variety that will arise in warfighting, any more than prequalifying the capabilities and performance of Michael Vick as a sports team player was able to avoid later behaviors that reflected poorly on all players by association (subsequent rehabilitation not withstanding). UAS that run amok in any way will reflect poorly on all UAS—eroding necessary public trust.

In a prior companion article (Dove 2009b), Arkin (2007) and Moshkina and Arkin (2007) were cited for identifying the important need of UAS conformance to rules of ethics, rules of war, and related high-level behaviors expected by the public of weapons toting UAS. Infractions can be devastating to continued public acceptance as well as to life and property. Range testing alone cannot assure safety under warfighting conditions. It is suggested that testing for appropriate behavior become a continuous process throughout the operational life of UAS, carried out by peer-peer monitoring.

Peer behavior monitoring occurs naturally and constantly in social animals. Each member of the group evaluates the others for adherence to social norms and threats to social coherence and security. Rogue elephants, for instance, are the result of banishment for unacceptable behavior. Social insects are known to restrain and even kill members of the group that overstep certain social bounds (Monnin et al. 2002; Heinze 2003; Flack et al. 2006; Ratnieks, Foster, and Wenseleers 2006).

Humans monitor the behavior of others in ways more sophisticated and more complex than animals of lesser cognitive capability. The process is often carried out formally as a test for granting new candidates membership in a group. Initial tests are typically for similar values, compatible behaviors, acceptable capabilities, and even for synergy in mission-based groups such as sports teams or Special Forces.

A revealing example of human peer-behavior monitoring and punishment was recently published

in Myers (2008). Professor Myers studied social reactions in on line game play. He played by the rules of the game but not by the cultural rules of the dominant player group. The degree of escalating retaliation as the group turned against him is an interesting study in human peer-behavior policing.

Unlike traditional approaches at sophisticated behavior detection and classification through reasoning, this article suggests an approach inspired by human expertise studies, where it appears that a conclusion is driven by a vast quantity of simultaneously accessible "experience" patterns rather than a compromising sequential search or reasoning process. This approach is both suggested and conceptually possible with new processor architectures offering massively parallel pattern-recognition capabilities. Many of these architectures are inspired by human cortical learning and classification models but may not offer ready post-learning algorithm cloning nor behavior-capability transparency. A single-processor architecture without integrated learning that features massive parallel classification capability for explicit patterns avoids these potential limitations, and will be used to establish a conceptual foundation for peer-peer socially attentive monitoring.

Detection-complexity leverage

Progress in pattern recognition has come in the form of trying harder with more elaborate recognition algorithms, pattern-tuned special-purpose processors, multicore processors and clustered servers, multiple graphic processors, and massively parallel supercomputers. All of these approaches continue to make tradeoffs among the same forces in tension: accuracy, time, and cost.

Biological capability is the benchmark for pattern recognition. Machines, like people, cannot recognize situations of which they have no prior knowledge. A healthy person over a lifetime builds up a wealth of experience patterns, stored in memory, adding details and variations as repeated exposure reveals new levels of nuance. How biological entities achieve this remains as conjecture, but it is clear that patterns are developed, retained, and applied in the necessary and constant sense-making of everyday life.

Klein (1998) suggests his Recognition Primed Decision model to explain how humans make decisions without apparent deliberation or reasoning. Well known for his studies of professional firefighters making appropriate choices for situation response almost immediately, he describes the Recognition Primed Decision model as one that uses intuition (pattern recognition) to qualify the first viable action, without conscious weighing and decision making.

Research indicates that human expertise (extreme domain-specific sense-making) is strongly related to meaningful pattern quantity. According to an interview with Nobel Prize winner Herb Simon (Ross 1998: 98–104), people considered truly expert in a domain (e.g., chess masters, medical diagnosticians) are thought unable to achieve that level until they've accumulated some 200,000 to a million meaningful patterns, requiring some 20,000 hours of purposeful focused pattern development. The accuracy of their sense-making is a function of the breadth and depth of their pattern catalog. Of interest, in biological entities, the accumulation of large expert-level pattern quantities does not manifest as slower recognition time. All patterns seem to be considered simultaneously for decisive action. There is no search and evaluation activity evident.

On the contrary, automated systems, regardless of how they obtain and represent learned reference patterns, execute time-consuming sequential steps to sort through pattern libraries and perform statistical feature mathematics. This is the nature of the computing mechanisms and recognition algorithms generally employed in this service.

Ross (2006) talks about the expert mind, and Herb Simon presents a "chunking" explanation for how chess masters can manage and manipulate a vast storehouse of patterns. Ross ties this chunking discussion into the common understanding that the human mind seems limited by seven plus-or-minus two elements in working memory: "By packing hierarchies of information into chunks," Simon argued, "chess masters could get around this limitation, because by using this method, they could access five to nine chunks rather than the same number of smaller details."

Psychologist George Miller (1956) wrote "The magical number seven plus or minus two" that provided the underpinning for Simon's suggestion. Miller's article is a great and rare reading pleasure as well as a rich storehouse of information, far beyond the simple seven-digit limitation to which common reference has reduced it. Of importance, Miller's concept of chunking into hierarchical levels of patterns-of-patterns appears highly relevant in attempting to build pattern-recognition algorithms that exhibit capabilities seen in humans. Subsequent research (Cowan 2001) carries this study of chunks and limits further and makes a case for the number four plus-or-minus one as a more likely limit.

Similar chunked-hierarchy architecture is reported by researchers at MIT (Serre, Oliva, and Poggio 2007). Serre's doctoral dissertation (Serre 2006) describes

> *"a quantitative model that accounts for the circuits and computations of the feedforward path of the ventral stream of visual cortex," and claims "that this may be the first time that a*

> *neurobiological model faithful to the physiology and the anatomy of visual cortex . . . achieves performance close to that of humans in a categorization task involving complex natural images." (Serre 2006)*

Though Serre's work is focused on image recognition, it is inspirational in its fit with the platform developed in this article and will surely guide subsequent steps in this investigation.

This section will close by noting a tie to the discussion in Dove (2009b) of multi-agent trajectory behavior recognition. Hockey legend Wayne Gretsky is renowned for his field sense (Kahn 2007)—knowing where his teammates are without looking and knowing where the puck will be next. Though what sensory mechanisms are involved may be illusive for now, this is expert pattern recognition involving the trajectories of bodies and objects in motion, rather than static chessboard configurations or medical diagnostic symptoms. Intille's American football-play identification from visual image pattern recognition did not have the vast quantity of patterns associated with expertise (Intille 1999, 2001), nor did it have to respond in real time; but Intille's work can offer initial guidance on how an artificial mechanism might represent tactical choreography patterns far in excess of a football playbook.

Technology leverage

Brain circuitry understanding and models of parallel pattern-recognition algorithms with brain-like results at MIT (Riesenhuber and Poggio 1999), at the San Diego Neuroscience Institute (McKinstry, Edelman, and Krichmar 2006), and at Numenta (George 2008) are already being fabricated as experimental Very-Large-Scale Integrated (VLSI) circuits at Stanford (Merolla and Boahen 2006) and at the Ecole Polytechnique in Lausanne (Schemmel et al. 2006). These VLSI chips combine analog and digital circuitry to emulate simple models of neuron/synapse circuitry, and integrate pattern learning with pattern detection.

The integrated nature of learning before recognition may make rapid cloning of the information in these chips difficult, and the nature of the learned patterns may be difficult to verify for behavior boundaries. For example, Jeff Krichmar, a senior fellow of the San Diego Neuroscience Institute, said in a recent interview:

> *"Put a couple of my robots inside a maze, let them run it a few times, and what each of those robots learns will be different. Those differences are magnified into behavior pretty quickly." (Kotler 2009)*

These chips promise remarkable capabilities, but they may also raise some problems for weapons toting UAS test and verification.

On the other hand, conventional stored-program sequential-instruction processors pressed into massive pattern recognition service are severely constrained by trade-offs among speed, cost, and accuracy.

A new VLSI pattern-detection processor architecture, shown partially in *Figure 1*, does not contain integrated learning, can be unambiguously loaded with detection patterns, decouples the speed/accuracy trade-off, and renders the cost/accuracy trade-off negligible (Dove 2009a). The architecture features massively parallel, dynamically configurable Feature Cell Machines (FCMs), which simultaneously process the same data stream. Low-cost VLSI fabrication, unbounded scalability, and high-speed constant-rate throughput independent of pattern number and complexity break current trade space constraints.

This decoupling of the speed/accuracy trade-off constraints enables new possibilities for employing pattern recognition. In particular, the massive quantity of simple patterns associated with expert performance can be investigated as an alternative to time-consuming accuracy-compromising computational heuristics. In one sense it sounds like a brute force approach: enumerating all possible patterns of interest, rather than developing an elegant heuristic. On the other hand, the biological benchmark appears to use this massive-pattern-quantity approach; while "elegant" approaches are made necessary by the nature of the computational mechanisms employed—not the problem in need of a solution—and they extract a cost in both accuracy and time that can be avoided.

The processor architecture and how it eliminates these trade-offs is explained in Dove 2009a. Currently an emulator is used for investigating parallel algorithm development, with field programmable gate array (FPGA) processor prototypes employed for large data streams while VLSI design is in process.

Some understanding of the processor architecture is necessary. Referring to *Figure 1*, a partial view of the processor's architectural concept shows massively replicated detection cells. A quarter to half million such cells on a single VLSI chip appears possible for early generation silicon. These cells are independent units, with four dynamically configurable elements to consider:

1. an activation status,
2. a 256-element feature-vector designating all byte values of interest,
3. a set of pointers to other cells that will be activated if this cell is "satisfied," and

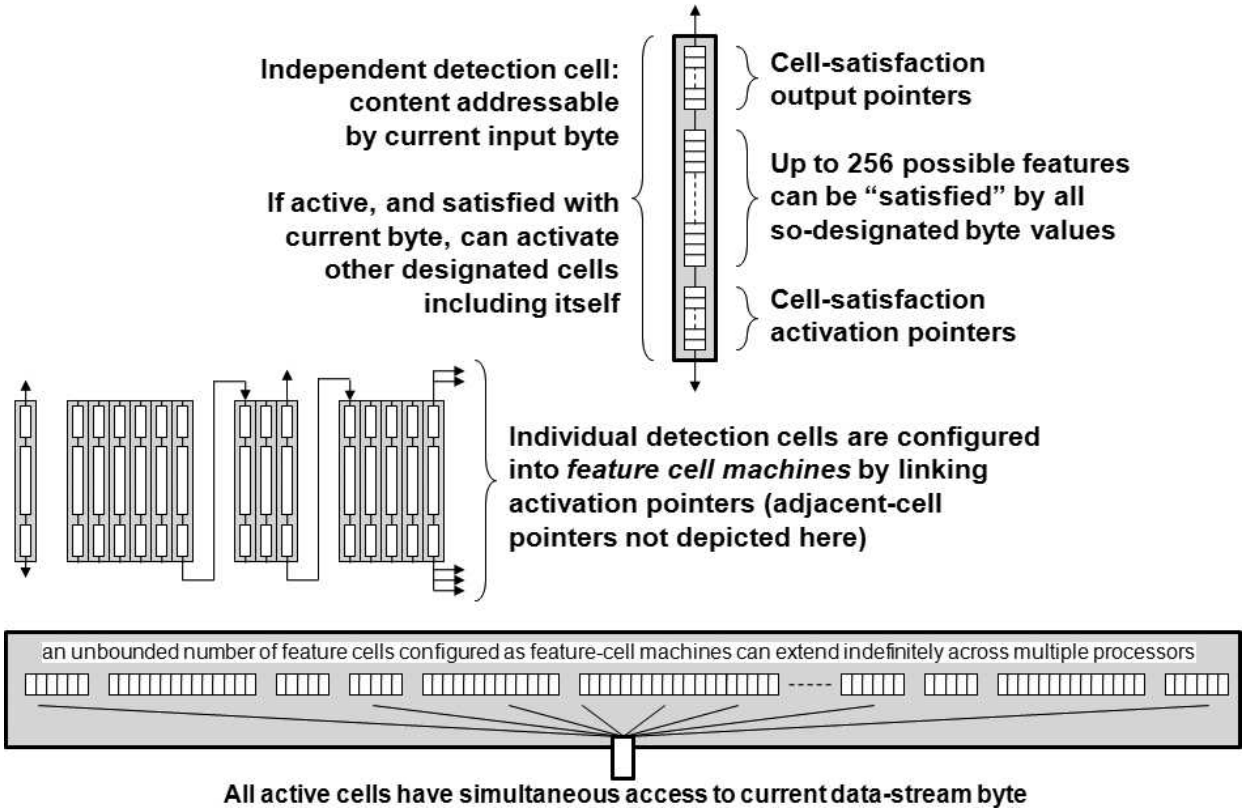

Figure 1. Configurable feature cells and feature cell machines.

4. a set of pointers to output transforms that can logically combine designated cell-satisfactions into buffered output and reset designated processor status.

In operation, an external controller feeds data stream bytes in sequence to a current-byte register in the processor. The presentation of each new byte triggers the beginning of a detection cycle. The current byte acts as an index into the feature vector for all active cells simultaneously. If a cell is active and the current byte value is one of interest, as designated in the feature vector, it is said to be "satisfied." That satisfaction will activate (for the next cycle) all other cells according to the satisfied cell's activation pointers and will cause designated output transformations to occur according to this cell's output pointers. A cell's activation pointers may include one that reactivates itself, as cell activation is effective for a single cycle only. Note that a cell can respond to any number of data-stream byte values, which enables value-based as well as syntactic feature-based classification.

Multiple processors can be employed in parallel and serial arrangements to increase throughput speed and/or reference-pattern capacity. Interleaving packet-based data streams, for instance, across multiple processors can be used to increase throughput speed. Presenting the data stream "current" character to multiple processors simultaneously can be used for unbounded reference-pattern scalability.

With this architecture, groups of detection cells can be configured into FCMs, similar to finite state machines, by setting activation pointers to pass activation successively through a group of successively "satisfied" cells. One cell may activate many other cells, so that multiple pattern branches may become simultaneously active. Any number of such FCMs may be configured within the total cells available within a processor. Typically such FCMs are created to detect (classify) specific patterns or subpatterns of interest.

The next section will provide some simple examples that could be useful in aberrant behavior detection methods for the concepts presented earlier.

Classification techniques

Pattern recognition has two distinct approaches, and a third that blends the two. For a full treatment see Jain, Duin, and Mao (2000).

- Statistical Approach—In this approach an unknown entity or situation is characterized by a set number of features and measured values for each

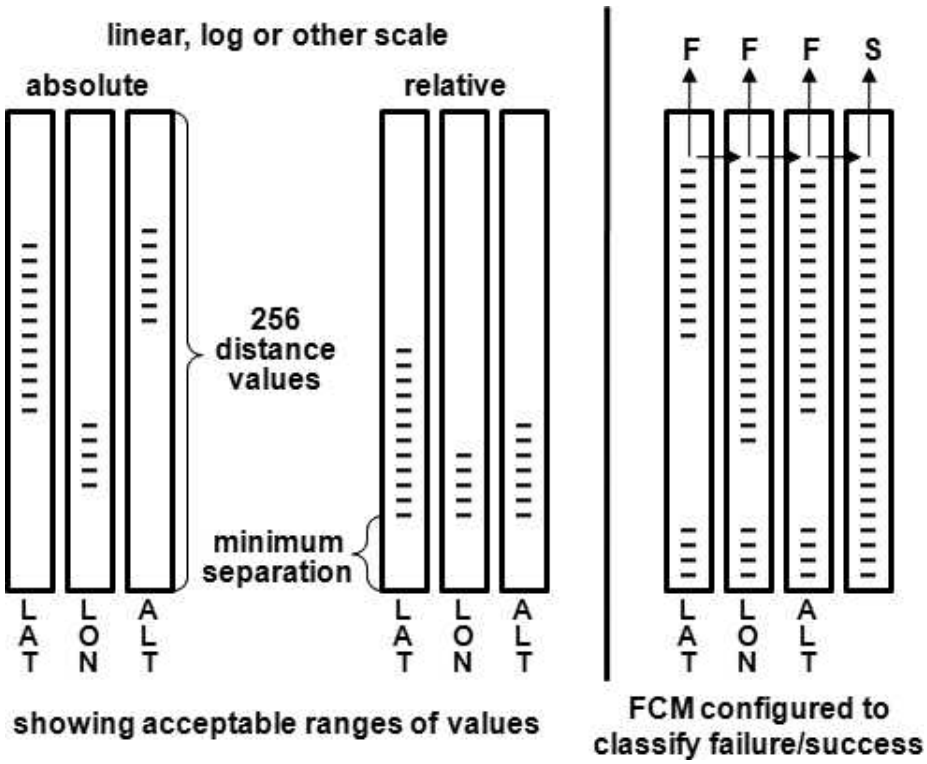

Figure 2. Some possible ways of encoding an envelope of acceptable values for latitude, longitude, and altitude. Multiple detection cells and data-stream bytes could be used for each. Minimum separation ensures two Unmanned Autonomous Systems (UAS) do not get dangerously close. A detection cell only has one satisfaction line, so the Feature Cell Machine (FCM) on the right must satisfy on failure with the complement of the permissible range. The final cell matches on anything and signals success.

feature (a person characterized by height and weight; a danger characterized by velocity, distance, and heading). Mathematically the features become dimensions in a multidimensional space, and the values for each of those features then place an unknown entity or situation at a point in that multidimensional space. Regions of the space are associated probabilistically with pattern classification (man or woman; dangerous or suspicious or benign).

- Syntactic Approach—This approach is structural in nature and generally hierarchical, where patterns are composed of subpatterns, which are in turn composed of subpatterns, with the lowest level subpatterns being simple recognizable primitives. In syntactic pattern recognition, a formal analogy can be drawn between the structure of patterns and the syntax of a language. Language parsing is common usage for this approach, but other patterns such as waveforms and multi-agent trajectory paths, which can be constructed from primitive structural components, lend themselves to syntactic recognition. Syntactic patterns are composed of primitives that follow rules about how they may be combined in relation to each other. Using the common linguistic metaphor, these rules form a grammar of allowable pattern structures.
- Augmented Grammars—This approach combines the two above, which may be done in a variety of ways to suit the raw sensor data, the computational resources being employed, the difficulty of feature extraction, and the speed vs. accuracy trade-offs dictated by the application. In a general sense, augmented grammars have syntactic elements and semantic elements, mixing structural relationships and feature values.

The processor described here is well suited to the syntactic approach, appears highly promising for augmented grammar approaches, and has utility for some statistical approaches.

A few general basic techniques will be shown to give some idea of how detection cells can be organized as FCMs, and how such FCMs can be organized to discriminate syntactic structure, feature values, and pattern groupings.

Feature value discrimination

One likely classification of undesirable behavior might be a UAS that is not where it is expected to be. Perhaps it

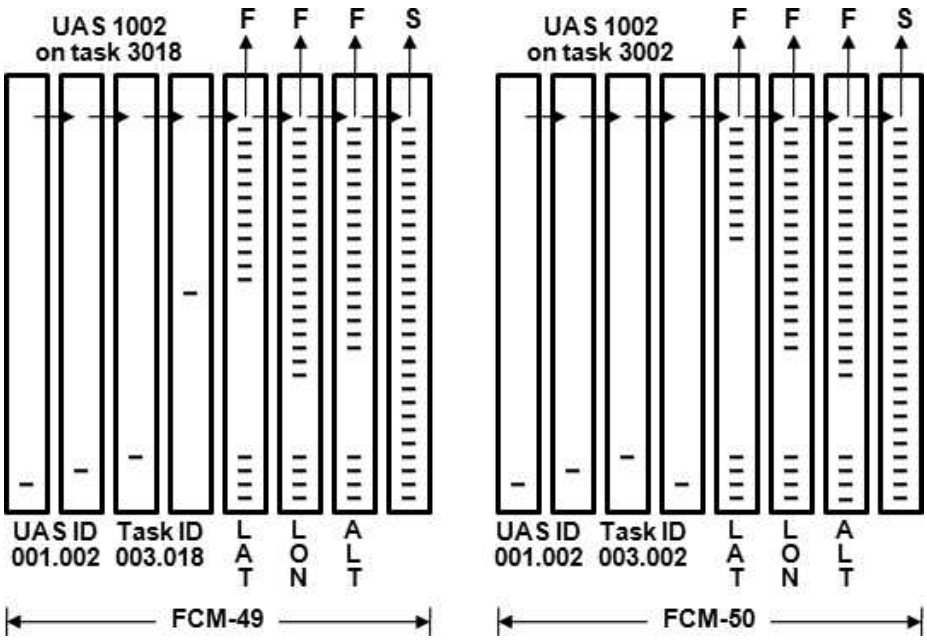

Figure 3. This example employs a packet approach to data-stream packaging. Packets here have a two-part header: the first two cells/bytes signify the Unmanned Autonomous Systems (UAS) associated with the data, and the second two signify the task that this UAS is currently supposed to be executing. Detection of the header activates the correct positioning envelop for that UAS on that task. Two Feature Cell Machines (FCMs) are shown for the same UAS on two different tasks 49 and 50.

has developed a mobility malfunction, missed a cue signaling a new task, been incapacitated by the enemy, or redirected by an unauthorized command. If a team of UAS is coordinated in accordance with a specific plan, each agent is expected to maneuver within some absolute or relative location envelope. This envelope may be narrow during travel to a target area, larger during engagement, and different among some members of the team when sub-groups are deployed on separate tasks. The example in *Figure 2* shows two ways to encode the location of a UAS by three global positioning system coordinates: latitude, longitude, and altitude. The absolute case can detect a UAS that wanders outside of its expected travel envelope for a given leg of a journey. The relative scale might be in relationship to a monitoring team-member during some phase or task of planned teamwork.

Figure 3 shows how the location envelope for two different tasks assigned to the same UAS at different times during a mission might be configured. Imagine thousands or tens of thousands of such reference patterns all prepared to classify incoming data as acceptable or not, all localized to multiple specific UAS. The example FCM could be expanded to include additional behavior data for a specific UAS in a specific task, or separate FCMs could handle separate behaviors and be associated appropriately with data-stream packet headers.

A final technique example is offered to indicate one of the ways this processor could be used to weight different features or subpatterns within a total pattern. *Figure 4* indicates some of the higher level aggregation and output capabilities of the processor conceptually (Dove 2009a). In this example, the down counters are employed to give different weights to different features of a pattern. A down counter can be initialized to some value when a configuration load is sent to the processor. Output pointers associate specific down counters with specific FCM satisfaction lines. *Figure 4* depicts four possible classifications for a large number of features, where one of those features carries a weight of 2, another carries a weight of 3, and the rest each carry a weight of 1.

The examples shown are all simplistic and not indicative of the range of possibilities. They were chosen to show some specific techniques that broaden the purely syntactic applications readily associated with this reconfigurable replicated detection cell architecture.

Concluding remarks

The leverage discussions in this and the prior article (Dove 2009b) are complementary, each having a potential role in a total solution platform. The starting point in the prior article was social attentiveness in certain biological systems that monitor and enforce peer behavior. The end point was a conceptual example of technology that can approach the pattern capacities and speeds of biological systems in bounded subdomains of behavior interest.

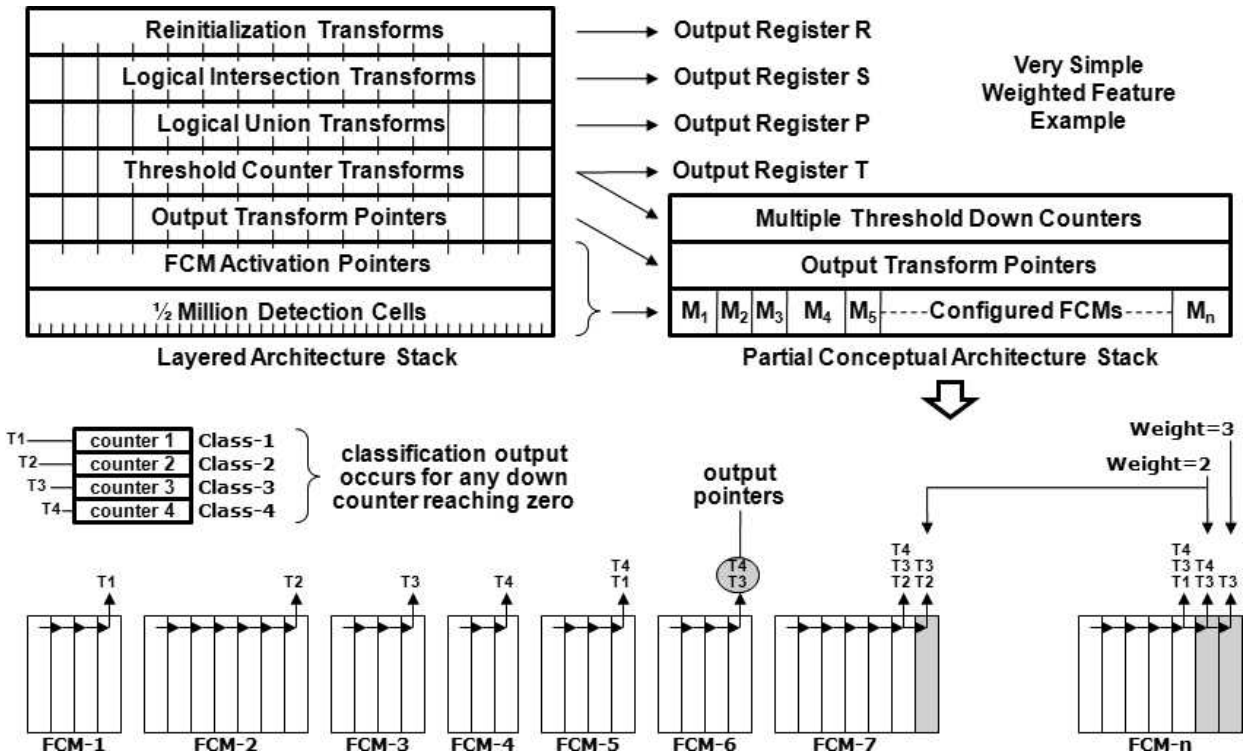

Figure 4. This conceptual depiction of additional features of the processor's architecture, at the top, places the Feature Cell Machines (FCMs) and their association with down counters in context. The bottom depiction shows a large number of multi-element FCMs, which are used to determine a classification. Classification-1, for instance, may have its down counter initialized to 3, while Classification-4 may have its down counter initialized to 4. Thus, Classification 4 can occur in multiple ways, whereas Classification-1 must have one each of the three specifically designated features.

The work reported here attempted to find sufficient connective concepts between the two end points that would warrant a next stage investigation. It is not suggested that the connecting leverage points discussed here and in the prior article are the only way to approach the problem effectively, nor that they are completely sufficient, but rather that they appear promising as a foundation for a solution path worth exploring.

This is preliminary work that sought and constructed a basis from which to investigate algorithms that can detect safety- and security-threatening behavior among UAS working in teams; where speed, accuracy, and breadth of comprehension are key performance factors. The work suggests that a promising basis exists in combining recognition of social behavior and trajectory behavior with a technology that can manage a vast quantity of stored reference patterns structured and accessed in a feedforward chunked hierarchy. Of importance, that technology must employ parallel recognition capabilities that eliminate any need for time-consuming search or sequential algorithms.

Social comparison theory guides us to a comparison of an agent's behavior pattern against behaviors of others on the team, against mission plans, against defined patterns of normal behavior, and against defined patterns of aberrant behavior. Trajectory behavior classification could be considered a special subset of social behavior detection for loosely choreographed teamwork.

Expertise theory, if it can be called that, guides us to a need for an extremely large number of reference patterns that can be simultaneously compared relative to a dynamic situation, eliminating time for sequential evaluation and reasoning steps, and eliminating much of the otherwise selective monitoring and pattern simplification that increases uncertainty.

This work was prompted by the growing concern for testing methods that can keep pace with the growing intelligence of UAS. It is suggested that a very different approach is required, one that never stops, one that carries the test environment into the operational environment, on board every UAS eventually.

There is precedence for and experience with this approach: the training and vetting of Special Forces operatives. One point to note is that we appear comfortable moving operatives into field status after some "testing" period, even though we know they will face situations that have not been tested. Another point to note is that these operatives on mission are always

being evaluated by their peers, who rely on the integrity of each and every member of a team.

UAS will operate outside of ready observation and are subject to attrition by enemy destruction. In such live cases, aberrant behavior must be detected and evaluated to sense control penetration by the enemy, as well as malfunction, that threatens the mission or might provide a disabled UAS to the enemy for recovery post-mission.

Continued study will investigate appropriate classes of behavior for monitoring, the nature of expert-level detection capability, and suitable pattern representations for whatever technology is employed. ❑

Rick Dove is a founding partner of Kennen Technologies, an adjunct professor in the School of Systems and Enterprises at Stevens Institute of Technology, and chairman of Paradigm Shift International. He is author of Response Ability—The Language, Structure, and Culture of the Agile Enterprise; *and* Value Propositioning—Perception and Misperception in Decision Making. *He is an active researcher and consultant in agile and self-organizing systems. He has 40 years' experience at start-up, turn-around, and interim executive management. He holds a bachelor of science degree in electrical engineering from Carnegie Mellon University and has done graduate work in computer science at U.C. Berkeley. E-mail: rick.dove@kennentech.com*

References

Arkin, R. C. 2007. *Governing lethal behavior: embedding ethics in a hybrid deliberative/reactive robot architecture*. Technical Report GIT-GVU-07-11. Atlanta, Georgia: Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf (accessed May 15, 2009).

Cowan, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24 (1): 87–185. http://www.bbsonline.org/documents/a/00/00/04/46/bbs00000446-00/bbs.cowan.html (accessed May 15, 2009).

DoD (Department of Defense) Office of the Secretary of Defense. 2008. *Unmanned and Autonomous System Testing (UAST) Broad Agency Announcement <BAAUAST0002>*. April 4, 2008. Washington, D.C.: DoD.

Dove, R. 2009a. Pattern recognition without tradeoffs: Low-cost scalable accuracy independent of speed. In *Proceedings Cybersecurity Applications and Technologies Conference for Homeland Security (CATCH)*, March 3–4, Washington, D.C., 255–260. Piscataway, NJ: IEEE. http://www.kennentech.com/Pubs/2009-PatternRecognitionWithoutTradeoffs.pdf (accessed May, 15, 2009).

Dove, R. 2009b. Paths for peer behavior monitoring among unmanned autonomous systems. *The ITEA Journal* 30 (3): 401–408.

Flack, J. C., M. Girvan, F. B. M. de Waal, and D. C. Krakauer. 2006. Policing stabilizes construction of social niches in primates. *Nature* 439 (7075): 426–429.

George, D. 2008. *How the Brain Might Work: A Hierarchical and Temporal Model for Learning and Recognition*, Ph.D. dissertation, Department of Electrical Engineering, Stanford University, June. http://www.numenta.com/for-developers/education/DileepThesis.pdf (accessed May 15, 2009).

Heinze, J. 2003. Reproductive conflict in insect societies. In *Advances in the Study of Behavior*, ed. P. Slater and J. Rosenblatt, 1–57. New York: Academic Press.

Intille, S. S. Month 1999. Visual recognition of multi-agent action. Ph.D. dissertation, MIT. http://web.media.mit.edu/~intille/papers-files/thesis.pdf (accessed May 15, 2009).

Intille, S. S. 2001. Recognizing planned, multiperson action. *Computer Vision and Image Understanding* 81 (3): 414–445. http://web.media.mit.edu/~intille/papers-files/cviu01.pdf (accessed May 15, 2009).

Jain, A. K., R. P. W. Duin, and J. Mao. 2000. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (1): 4–37.

Kahn, J. 2007. Wayne Gretzky-style 'field sense' may be teachable. *Wired Magazine*, May 22. http://www.wired.com/science/discoveries/magazine/15-06/ff_mindgames# (accessed May 15, 2009).

Klein, G. 1998. *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kotler, S. 2009. Here Come the Neurobots. *h+ Magazine*, June 5. http://www.hplusmagazine.com/articles/ai/here-come-neurobots (accessed July 15, 2009).

McKinstry, J. L., G. M. Edelman, and J. L. Krichmar. 2006. A cerebellar model for predictive motor control tested in a brain-based device. *PNAS* 103 (9): 3387–3392. http://www.pnas.org/content/103/9/3387.full.pdf+html (accessed May 15, 2009).

Merolla, P. A., and K. Boahen. 2006. Dynamic computation in a recurrent network of heterogeneous silicon neurons. In *Proceedings IEEE International Symposium on Circuits and Systems*, May 21–24, Kos, Greece, 4539–4542. http://www.stanford.edu/group/brainsinsilicon/pdf/ISCAS2006_merolla.pdf (accessed May 15, 2009).

Miller, G. A. 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63 (2): 81–97. www.musanim.com/miller1956/ (accessed May 15, 2009).

Monnin, T., F. L. W. Ratnieks, G. R. Jones, and R. Beard. 2002. Pretender punishment induced by chemical signaling in a queenless ant. *Nature* 419 (6902): 61–65.

Moshkina, L., and R. C. Arkin. 2007. *Lethality and autonomous systems: Survey design and results*. Technical Report GIT-GVU-07-16. Atlanta, Georgia: Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology. http://www.cc.gatech.edu/ai/robot-lab/online-publications/MoshkinaArkinTechReport2008.pdf (accessed May 15, 2009).

Myers, D. 2008. Play and punishment: The sad and curious case of Twixt. In *Proceedings of The [Player] Conference*, August 26–29, Copenhagen, Denmark.

Ratnieks, F. L., K. R. Foster, and T. Wenseleers. 2006. Conflict resolution in insect societies. *Annual Review of Entomology* 51: 581–608. http://www.people.fas.harvard.edu/~kfoster/RatnieksetalAnnreventomol2006.pdf (accessed May 15, 2009).

Riesenhuber, M., and T. Poggio. 1999. Hierarchical models of object recognition in cortex. *Nature Neuroscience* 2 (11): 1019–1025. http://cbcl.mit.edu/projects/cbcl/publications/ps/nn99.pdf (accessed May 15, 2009).

Ross, P. 1998. As I see it: Flash of genius. *Forbes*. November 16. http://www.forbes.com//forbes/1998/1116/6211098a.html (accessed May 15, 2009).

Ross, P. 2006. The expert mind. *Scientific American* 295 (2): 64–71.

Schemmel, J., A. Grubl, K. Meier, and E. Mueller. 2006. Implementing synaptic plasticity in a VLSI spiking neural network model. In *Proceedings International Joint Conference on Neural Networks*, July 16–21, Vancouver, BC, Canada, 1–6. http://www.kip.uni-heidelberg.de/Veroeffentlichungen/download.cgi/4620/ps/1774.pdf (accessed May 15, 2009).

Serre, T. June 2006. Learning a dictionary of shape-components in visual cortex: Comparison with neurons, humans and machines. Ph.D. dissertation, Massachusetts Institute of Technology. http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2006-028.pdf (accessed May 15, 2009).

Serre, T., A. Oliva, and T. Poggio. 2007. A feedforward architecture accounts for rapid categorization. *Proceedings National Academy of Sciences* 104 (15): 6424–6429. http://cvcl.mit.edu/Papers/SerreOlivaPoggioPNAS07.pdf (accessed May 15, 2009).

Acknowledgments

This work was supported in part by the U.S. Department of Homeland Security award NBCHC070016. Curt Harris invented the processor architecture.