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

ITEA Best Publication of the Year Award

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

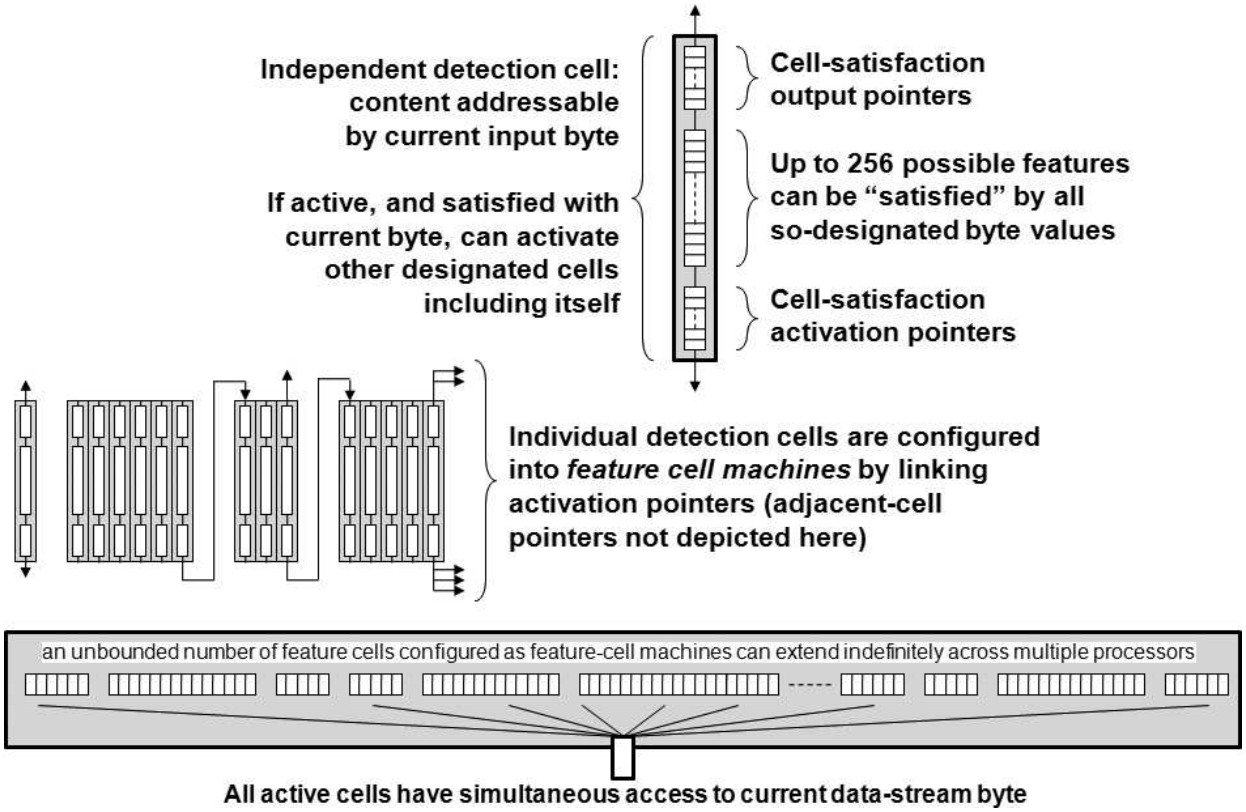

Figure 1. Configurable feature cells and feature cell machines.

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

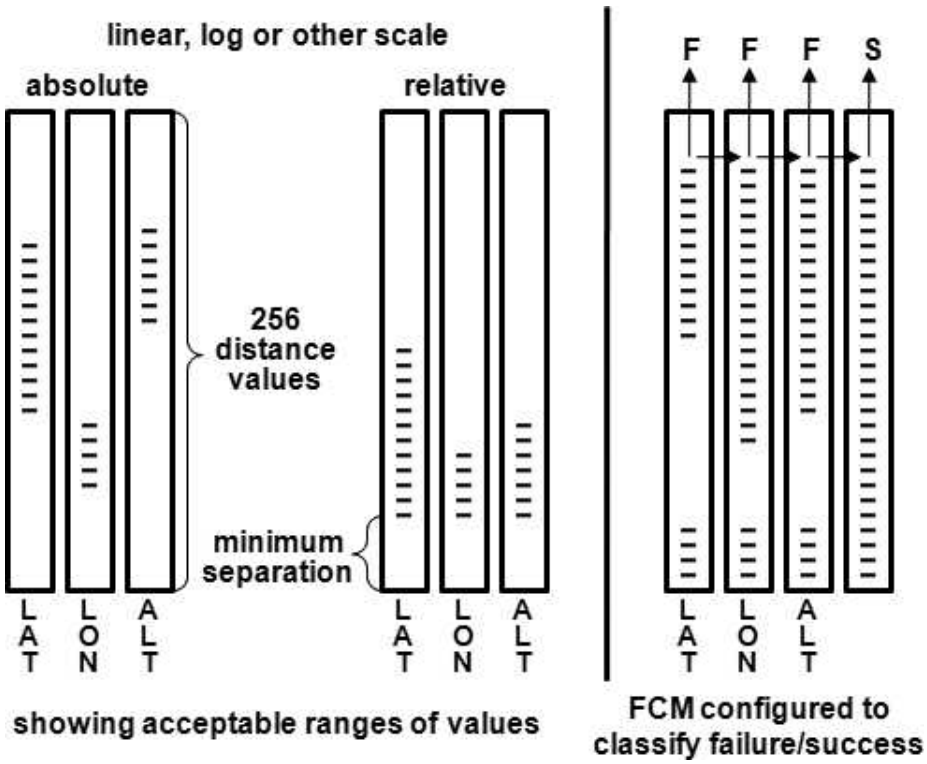

Figure 2. Some possible ways of encoding an envelope of acceptable values for latitude, longitude, and altitude. Multiple detection cells and data-stream bytes could be used for each. Minimum separation ensures two Unmanned Autonomous Systems (UAS) do not get dangerously close. A detection cell only 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

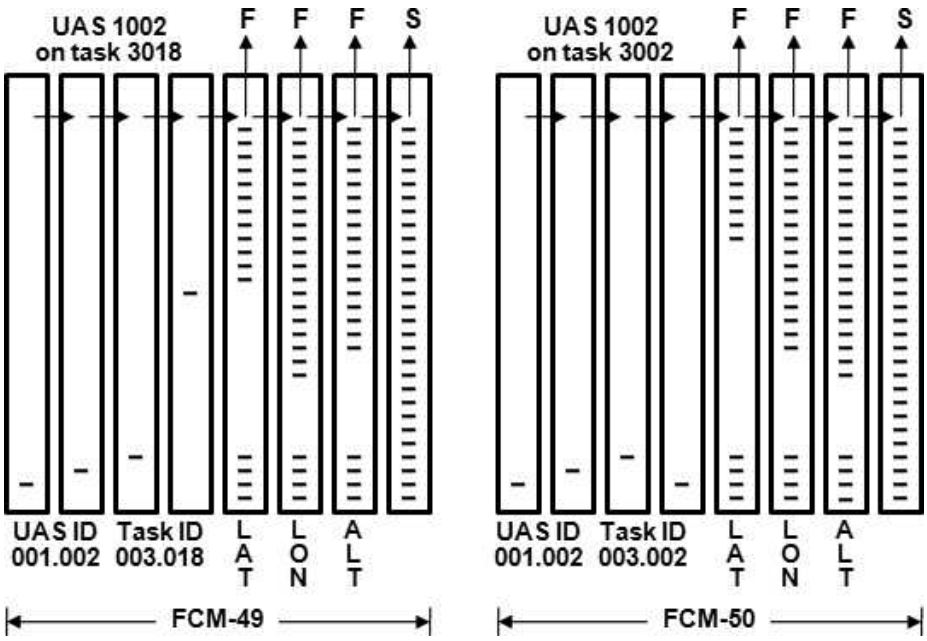

Figure 3. This example employs a packet approach to data-stream packaging. Packets here have a two-part header: the first two cells/bytes signify the Unmanned Autonomous Systems (UAS) associated with the data, and the second two signify the task that this UAS is currently supposed to be executing. Detection 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.

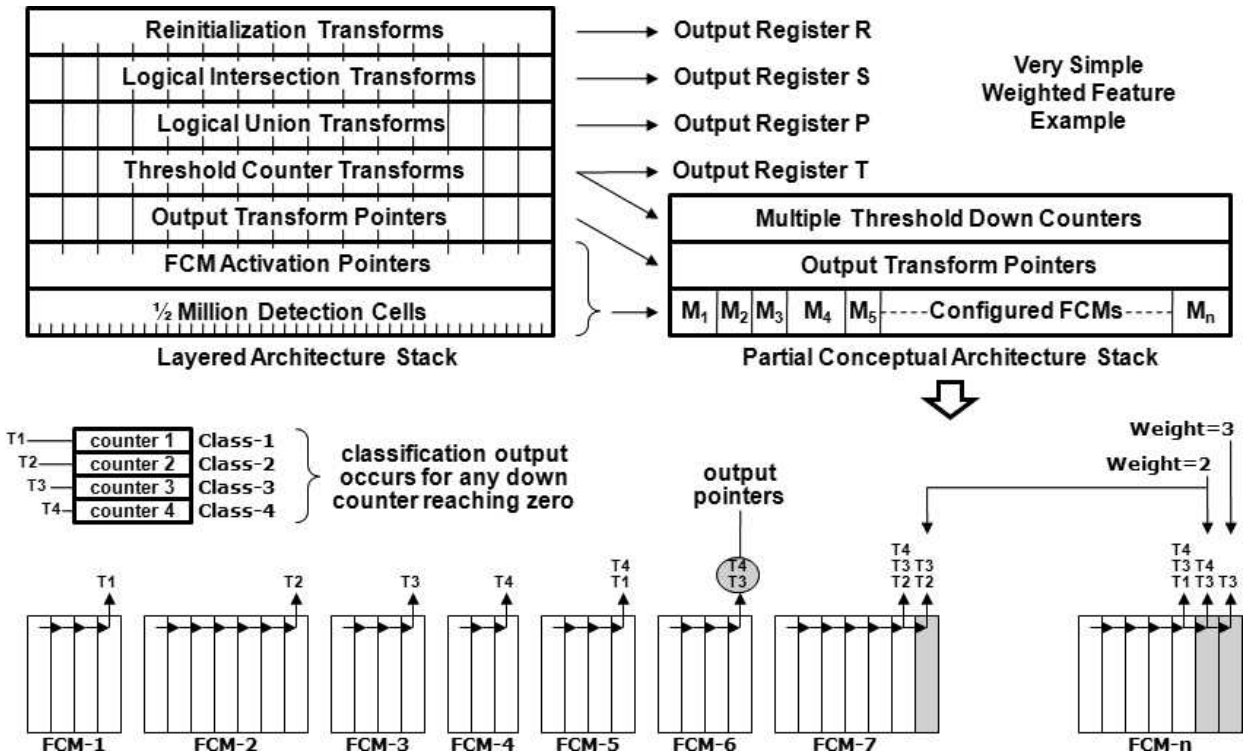

Figure 4. This conceptual depiction of additional features of the processor's architecture, at the top, places the Feature Cell Machines (FCMs) and their association with down counters in context. The bottom depiction shows a large number of multi-element FCMs, which are used to determine a classification. 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.