EVENT TEMPLATE HIERARCHIES AS MEANS FOR HUMAN-AUTOMATION COLLABORATION IN SECURITY SURVEILLANCE

David Woods, Stephanie McNee
Cognitive Systems Engineering Laboratory
The Ohio State University

James W. Davis, Alexander Morison
Computer Science and Engineering
The Ohio State University

Patrick Maughan
Security and Fire Prevention
The Ohio State University

Klaus Christoffersen
CS Designs, Hamilton, Canada

Advances in remote sensing systems provide human monitors access to more data. The current challenge is to help extract relevant patterns and direct the attention of human monitors and human problem holders to the changing security picture with respect to acute situations (a normal market scene turns into an ethnic confrontation) or longer term trends (seeing new patterns of 'typical' behavior to avoid false alarms). Security surveillance monitoring can be advanced through new event recognition capability of autonomous monitors and by effectively coupling these sensor/algorithm systems to human monitors and problem holders. To meet these challenges in security surveillance, we have developed an event-sensitive architecture where machine agents provide event-based information to human monitors and problem holders and are re-directable given contextual information. The key innovation is a context-based *hierarchical event template structure* which can be used to integrate data over a distributed sensor system.

Challenges in surveillance and sensors

As people mill in a outdoor marketplace, can peacekeeping forces anticipate signs of tensions rising and notice when factions begin to coalesce and square off (or similarly, in a parking lot or street after a game, can police separate 'normal' post-game activities from the beginnings of mob violence)? A lone figure enters a large parking lot and roams from lane to lane; in the homogenous backdrop of the lot, he cannot tell where he parked much earlier that day. Obscured in the stream of pedestrians after a football game, four figures on opposite sides of a courtyard identify a victim and begin to close in. In a rural landscape near a U.S. base, a robed figure is walking along with a flock of goats—is he caring for the flock or are his movements unrelated? Is he carrying a light load of a meal and water in a sack or is there a heavier load of arms obscured under his robes?

In monitoring and surveillance, these kinds of scenarios represent the difficulties in discriminating unusual and threatening behavior against the backdrop of the ranges of 'typical' behaviors of people in that environment. What is informative in surveillance includes recognizing activities, knowing typical and threat patterns of behavior, inferring intent, seeing through attempts to mask threats with 'normal' behavior.

Advances in remote sensing systems provide human monitors access to more data. Simply providing more and different sensors over wider areas does not in itself help extract informative patterns from the increasing sea of low level data. Security surveillance can be advanced by organizing monitors around event patterns to recognize changing security issues, e.g., a normal market scene begins to turn into an ethnic confrontation, or noticing new patterns of 'typical' behavior to avoid false alarms.

While there has been rapid advances in enabling technology islands, there are important challenges for taking advantage of sensor advances in future surveillance systems. These include how to model complex events against changing background of activity and at different time scales, how to integrate inputs from data sources that vary in modality and in level of autonomy, how to cull out unimportant data and avoid the debilitating consequences of false alarms, how to combine multiple information sources into a coherent picture for human monitors, how to direct the attention of human monitors to 'interesting' behaviors or changes in behavior, and how to provide means for human monitors or problem holders to re-direct and interact with partially autonomous surveillance systems.

Event-Sensitive, Collaborative Autonomy

Previous work has shown that event patterns are strongly context-sensitive which is a major barrier for fully automated event recognition. Cooperative human-computer solutions that greatly amplify the natural ability of knowledgeable human monitors to perceive and reason in terms of events provide a more promising approach.

An interdisciplinary team is developing a collaborative autonomy architecture where machine agents provide event-based information to human monitors and problem holders and are re-directable given contextual information. The key innovation is a context-based *hierarchical event template structure* which can be used to integrate data over a distributed sensor system.

Building the hierarchical event structure begins with the identification, modeling, and processing of events (Warren and Shaw, 1985; Woods, 1995b; Zacks et al., 2001). The formal methodology developed by Christoffersen, Blike and Woods (2003) is used to discover events in surveillance (based on previous work by Newston, 1973; Zacks et al., 2001). This method (a) can identify patterns that are important but may not be available from interviews and other knowledge elicitation techniques, (b) is time-based so that it provides information about the actual temporal flow of events, (c) reveals the background of expectations that influence what is recognized as an event, given all of the other changes which are or may be going on, (d) is sensitive to difference between experienced and inexperienced practitioners, (e) can discriminate complex events that are defined by more than just data features (e.g., the absence of an expected change as an event).

What are events?

For our purposes, an event is a meaningful pattern of change (or lack thereof) for an observer in a given environment. Events are defined in terms of a contrast with expectations (cf. Teigen and Keren, 2003).

Based on studies of human competence at event recognition in monitoring complex processes, Christoffersen and Woods (2003) have developed a new model that captures the complexities of event patterns. These complexities raise considerable barriers for fully automated event recognition. Instead the results point to the promise of cooperative human-computer solutions that utilize intermediate levels of automated pre-processing to organize and present data in ways that greatly amplify the natural ability of knowledgeable human monitors to perceive and reason in terms of events (Warren and Shaw, 1985).

The event cycle model is illustrated in Figure 1. The event cycle consists of the influence model which establishes expectations about future behavior. Sensed data is analyzed for these expected data relationships. Patterns noticed in the sensed data, because of their deviation from expected observations, are used to modify the model of the current situation and the cycle continues. This model provides a means to structure data about recurrent event patterns—event template hierarchies.

Event Template Matrix

Event patterns are composed of episodes of behavior and each episode is analyzed in terms of a matrix based on the event cycle model. Episodes define the rows of the matrix including recurrent types of *onset* (e.g., the surprise episode), *body* of the event (episode types such as deviation from typicality, response to interventions), and *resolution*. Onsets are precipitators or harbingers of events. The event body consists of the activities that carry the semantic content of the event, and the resolution is the termination of the activity and movement into next event (e.g., resolution in the Deteriorate/Recovery pattern in Table 2 is a return to target values, but the recovery can overshoot the target and lead to a new event sequence).

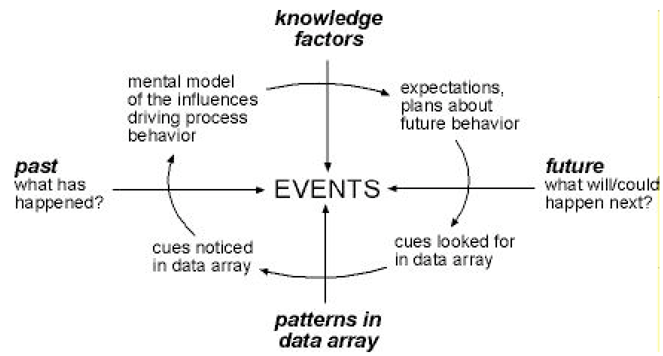

Figure 1. A model of factors underlying the recognition of event patterns in a dynamic telemetry stream (from Christoffersen and Woods, 2003).

Event pattern templates can be likened to frames in which attribute slots and their partial temporal ordering constitute the event template. An event template codifies a sequence of episodes that make up the event type plus the behaviors, feature relationships and contextual information that make up each episode.

Each of the three episodes in a generic event structure is broken down in terms of the 4 regions of the Event Cycle model and the columns of the matrix specify the behavior, relationships and context for each (see Table 1; Tables 2 illustrates the use of the event pattern templates).

Example of Event Pattern Templates

A general example of the Event Pattern Template structure is the Deteriorate/Recover recurrent pattern where the key relation at the outset is the difference between the actual value of a leading indicator P and its target. Since this surprising deviation, or anomaly, risks significant consequences, there is a need for agents to act to bring the process and this indicator back to normal range. The cause of the anomaly is unknown. The episodes that make up the body of the template are characterized by the further movement away from the target—continued deterioration. Expectations focus on the effect of intervention taken to counteract whatever influence is continuing to push P off-target. The next episode is marked by the recognition that P does not respond as expected to interventions – it continues to drift higher. Finally, the resolution episode or "turnaround" corresponds to where P eventually begins to respond appropriately to the intervention and returns towards the target range. This resolution not only concerns the recovery (return to target) but also the possibility that the recovery can overshoot the target and lead to a new event sequence.

Using event discovery methods in a particular domain generates a set of templates that constitutes a model of event patterns that can drive algorithm development and human-computer cooperation design.

Table 1. General Event Pattern Template
Event: NAME

Episode: ONSET	***Behavior***	***Relations***	***Context Knowledge***
features noticed			
model of influences			
expectations			
features looked for			

Table 2. Event template for the first (onset) episode of the Deteriorate/Recovery pattern.

SURPRISE	Behavior	Relations	Context Knowledge
features noticed	P is critically moving off-target;	level/direction of P relative to target range	target range/limits for P in this context; "criticality" of anomaly?
model of influences	presence of unknown influence moving/holding P off-target	Anomaly-disturbances-fault relationships	candidate set of problems based on situation and background information
expectations	intervention urgently required; responsiveness of P to intervention will be diagnostic too	Interventions to fault relationships	affordance set; what interventions are appropriate, possible
features looked for	evidence of intervention; any change in behavior of P in response	level of P relative to itself over time; relative to target over time	

Event Hierarchies

The template structure is naturally multi-level as episodes may be events in their own right with their own episode sequence and feature relationships. This induces an event template hierarchy as higher-level events are generated from lower-level events. While in principle, an event hierarchy has an arbitrary number of levels, we have generally found that a 3 level structure of low-level, mid-level, and high-level events is a very useful heuristic.

Event-oriented base features are interesting qualities of the sensed data that are directly computable from the sensed data itself, in particular, the data needed to specify *transitions*. Note that this step shifts data representation from elements such as states (e.g., current value) to an initial set of relationships based on transitions. This shift is essential to provide the foundation for event recognition (Thronesbery et al., 1999).

Low-level events are significant relationships across the base features that specify some form of behavior, i.e., change relative to a reference state (e.g., presence of a figure in a restricted area; rapid closing distance between 2 moving figures that stands out against typical relative motion in a market crowd). Note that assessing how the relationship of a transition or sequence of changes conforms to or differs from typicality for that context is at the heart of the shift from data relations to event patterns. In other words, a relationship across relationships—transitions relative to typicality or reference. We have found that algorithms to extract low level events can be accomplished if domain-specific accounts of typical behavior over domain contexts are modeled.

Mid-level and high-level events have an extended temporal structure of episodes that we model by the *onset-body-resolution* triple. Mid-level event triples are composed of low-level events and often represent shorter time scale structures or acute events. High level events generally also specify a context for mid-level events. High level event triples are made up of mid level events and can be used to express longer time scale patterns and support anticipating trends.

For example, consider monitoring of the RMS or remote manipulator system on Space Shuttle (the robot arm). High level events are how the RMS supports flight mission goals—grabbing or releasing a satellite and are composed of transitions in the RMS. The latter then function as mid-level events (e.g., unstowing or stowing the arm) which exist in the context of the plan, e.g. to grab a satellite in need of repair, and which are composed of the transitions that make up stowing (e.g., latching the arm). The latter then function as low level events—latching—in the context of e.g. stowing and composed of the event oriented base features which are various state transitions signaled by telemetry of sensors attached to the arm and related equipment. This example illustrates how the 3 level heuristic is a means to cope with the inherent context sensitivity and multi-level nature of events—what is an event at one level of representation is an element or a context at other levels of representation.

An important property of events is that they exist in relation to changing background activity and changing views of what should or what usually occurs in a context. This is similar to dynamic background update when extracting the foreground in image processing. In order to maintain a sense of this background, a model is dynamically updated for the background activity or what is typical in different contexts (e.g. what is typical goat herding behavior varies in terms of season, threats to the flock, and previous events such as a storm scattering the flock). It is this *background activity model* that sets the context for the events to occur in. We have found repeatedly in evaluating visualization systems on event recognition that a critical test is the ability of human monitors to see the lack of change in behavior when expectations change as a significant event (e.g., Christoffersen et al., 2001).

Event Templates for Security Surveillance

To illustrate the scenario-based event discovery process and to capture the complexity of the relationships in event patterns consider one example from security surveillance—Asset Defense Through Concentric Protective Rings. In this situation, a buffer zone is set up to protect an asset given some organized situation nearby (e.g. people within a stadium attempting to reach a specific location on the stadium field). An outer perimeter is set up to prevent groups or some sub-group from gaining access to the asset (a soft boundary set up based on the groups and their activities and the kind of asset). In case of barrier breakdown, an inner perimeter is also planned to protect the asset and ensure public safety.

A scenario is used to discover event patterns: the situation escalates as the outer boundary is stressed and then overwhelmed. Fallback is subsequently initiated and a secondary boundary is actively set up. The security interventions increase and inadvertently act as a catalyst for further disturbances that increase the threat to the asset and make resolution of the incident more difficult.

In this scenario note that the onset episode is not any given threat against the boundary in itself, rather it is related to how the threats may escalate relative to the security provided. How do security monitors recognize that threats are escalating: as individuals begin to coalesce into groups, as the energy of the crowd grows, as groups begin to move toward entry points to the targeted area.

The boundary is actively maintained (manned by ushers with a few supporting security personnel). As groups continue to push toward key parts of the boundary, contact with ushers occurs repeatedly, stressing the outer perimeter. As resources and attention flow to these points, other sub-groups become more aggressive and begin to probe for other points to gain access to the target area.

Since the boundary has not been breached at this point and there has been no overt aggressive behavior, the focus of ushers is moving the accumulating crowds away from the key entry points. The increasing scale of the surging groups threatens to spill over to the stadium field.

Then the outer perimeter is penetrated simultaneously at multiple locations. The loss of containment generates a significant security response as uniformed security and police personnel respond. The intervention influences the people's behavior as security personnel begin to feel pressed in on and threatened and respond more aggressively to activate the inner perimeter and try to contain the crowd. Tension levels increase dramatically and hostile interactions between the crowd and security begin to occur (e.g., imprecise use of pepper spray instigates more chaotic and aggressive behaviors).

To control and de-escalate the situation with minimal risks of violence and injury while still protecting the asset requires additional personnel and activities as the scenario continues.

Scenarios such as above are used with security professionals to discover the event boundaries and relationships captured generically in Figure 1 and the event template hierarchy described earlier. Given an event template hierarchy the question shifts to the issue of how to design for collaboration between automated monitoring sensor/algorithm nets and human monitors.

Design for Human-Machine Coordination

Human monitors are ultimately responsible for ascribing significance to a situation to meet security and surveillance goals. As the above scenario illustrates, the background behavior can be so rich and varied that fully automatic methods are guaranteed to fall short of performance criteria (the fundamental brittleness and literal-mindedness of automata; see Woods, 2002). On the one hand, human monitors can be overwhelmed if only raw or base sensor data is thrown at them. Hence the target of the research is to produce a human-machine architecture that dramatically advances the ability of human monitors to find emerging threats in security situations.

The event template structure and event hierarchies provides the basis for a new architecture that enables advances in:
(a) Intelligent Sensor Networks—the increasing the sophistication and diversity of the sensors allows networks to become increasingly attuned to informative events. Event patterns support fusion of the data collected over different sensors with different information, different quality, different speeds, different coverage and at different distances,
(b) Reconfigurable, Adaptive Sensor Networks—when sensors are seen as active resources, adaptation can be seen as reorganization of the sensor network/algorithms in terms of events of interest which will allow human monitors to direct system focus,
(c) Intelligent Alerting—when sensor system outputs are organized relevant to event patterns, machine alerts and prompts to areas, times and patterns of interest can be used to *redirect focus* of human monitors and problem holders (Woods, 1995a).
(d) Dynamic Field of Interest for Collaboration—a joint *focus* or *field of interest* is a critical base concept for any architecture that supports bi-directional refocusing across time scales (Woods and Elias, 1988). A smart sensor network also has a field of interest defined spatially, defined in how it adapts autonomously to e.g. sensor failures, but also defined in terms of its knowledge of event patterns including models of typicality or reference conditions.

Acknowledgements

The ideas articulated are based on discussions with R. Parent and R. Machiraju on the role of event patterns in advancing research on intelligent sensor networks.

This work has been carried out in part through participation in the Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0009. Support is provided by the National Science Foundation under IIS-0428249 Division of Information, Robotics & Intelligent Systems.

References

Christoffersen, K., Woods, D. D. and Blike, G. T. (2001). Extracting Event Patterns From Telemetry Data. Proceedings of the Human Factors and Ergonomics Society 45th annual meeting. 8-12 October, Minneapolis, MN.

Christoffersen, K., Blike, G. T., and Woods, D. D. (2003). Discovering the Events Expert Practitioners Find Meaningful in Dynamic Data Streams. Cognitive Systems Engineering Laboratory, Institute for Ergonomics, The Ohio State University, Columbus OH. Prepared for Advanced Decision Architectures CTA of Army Research Laboratory.

Christoffersen, K. and Woods, D. D. (2003). Making Sense of Change: Extracting Events From Dynamic Process Data. Cognitive Systems Engineering Laboratory, Institute for Ergonomics, The Ohio State University, Columbus OH. Prepared for Advanced Decision Architectures CTA of Army Research Laboratory.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, 28, 28-38.

Newtson, D. and Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, 12, 436-450.

Thronesbery, C. G., Christoffersen, K., and Malin, J. T. (1999). Situation-oriented displays of Space Shuttle data. In Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, (pp. 284-288). Houston, TX: HFES.

Warren. W. H. and Shaw, R. E. (1985). Events and encounters as units of analysis for ecological psychology. In W. H Warren and R. E. Shaw (Eds.), Persistence and change: Proceedings of the first international conference on event perception (pp. 1-27). Hillsdale, NJ: Erlbaum.

Woods, D. D. (1995a). The alarm problem and directed attention in dynamic fault management. *Ergonomics*, 38(11), 2371—2393.

Woods, D. D. (1995b). Towards a theoretical base for representation design in the computer medium: ecological perception and aiding human cognition. In J. Flach, P. Hancock, J. Caird, and K. Vicente (Eds.), *Global perspectives on the ecology of human-machine systems* (vol.1; pp. 157-188). Hillsdale, NJ: Erlbaum.

Woods, D.D. (2002). Laws that Govern Cognitive Work. Proceedings of the 24th Annual Meeting of the Cognitive Science Society [Plenary Address].

Woods, D. D. and Elias, G. (1988). Significance messages: an integral display concept. In Proceedings of the Human Factors Society, 32nd Annual Meeting, Santa Monica, CA.

Zacks, J. M., Braver, T. S, Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L. and Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6), 651-655.

Zacks, J. M., Tversky, B. and Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130(1), 29-58.