Gonzalez, C.; Ben-Asher, N.; Oltramari, A.; Lebiere, C. (2014). Cognition and Technology. In Kott, C., Wang, A. & R. Erbacher (eds.), *Cyber defense and situational awareness.* ISBN 978-3-319-11390-6. Springer International Publishing Switzerland 2014. DOI 10.1007/978-3-319-11391-3

Cognition and Technology

Cleotilde Gonzalez, Noam Ben-Asher, Alessandro Oltramari, and Christian Lebiere

1 Introduction

As the previous chapters emphasized, the human cognition – and the technology necessary to support it – are central to Cyber Situational Awareness. Therefore, this chapter focuses on challenges and approaches to integration of information technology and computational representations of human situation awareness. To illustrate these aspects of CSA, the chapter uses the process of intrusion detection as a key example. We argue that effective development of technologies and processes that produce CAS in a way properly aligned with human cognition calls for cognitive models – dynamic and adaptable computational representations of the cognitive structures and mechanisms involved in developing SA and processing information for decision making. While visualization and machine learning are often seen among the key approaches to enhancing CSA, we point out a number of limitations in their current state of development and applications to CSA. The current knowledge gaps in our understanding of cognitive demands in CSA include the lack of a theoretical model of cyber SA within a cognitive architecture; the decision gap, representing learning, experience and dynamic decision making in the cyberspace; and the semantic gap, addressing the construction of a common language and a set of basic concepts about which the security community can develop a shared understanding.

Far from being downgraded to interconnected computer technologies that constitute its physical substratum, cyberspace can be seen as a communication infrastructure built by humans to access and share information in real-time by means of a variety of interfaces and languages. In this regard, "cyberspace is defined as much by the cognitive realm as by the physical or digital" (Singer and Friedman 2014). The centrality of cognition in the cyber world is clearly illustrated in the process of detection, where a human analyst (i.e., a defender) is responsible for protecting client networks from illegal intrusions and hostile activity (i.e., cyber attack) that would jeopardize the integrity of its information and infrastructure. The detection process may be seen as analogous to the Data-Information-Knowledge-Wisdom (DIKW) hierarchical model that is central for information and knowledge management (Rowley 2007). In the DIKW model, often depicted as a pyramid, a hierarchical process is proposed where data is transformed into information, information into knowledge, and knowledge into wisdom.

Figure 1 illustrates this process for Detection. The existence of multiple and diverse sensors result in a large amount of network activity data. Cyber security tools (e.g., Intrusion Detection Systems, IDS) are meant to organize and structure network activity to make it relevant, meaningful and useful to support traffic monitoring and to minimize the damage that an attack can cause. Cyber security technologies provide ways to facilitate and protect an analyst from the cognitive challenges that the cyber world presents. For example, it does so by reducing, filtering and organizing large amounts of network events and by preprocessing events to help reduce the

information workload of the human analyst. These technologies would help in improving the analyst's Situation Awareness (SA): an accurate perception of the elements in the network within a volume of time and space, the comprehension of their meaning, and the projection of their future status (Endsley 1988). However, SA is rarely integrated into technology that would combine information with understanding and capability. Although there are multiple critical technologies to support an analyst in intrusion detection, they are often static and do not adapt to the analyst's state of mind and SA. Furthermore, SA is not an end in itself but rather the means by which analysts can make informed decisions in these complex, fast moving situations. SA is a precondition to make accurate intrusion detection decisions.

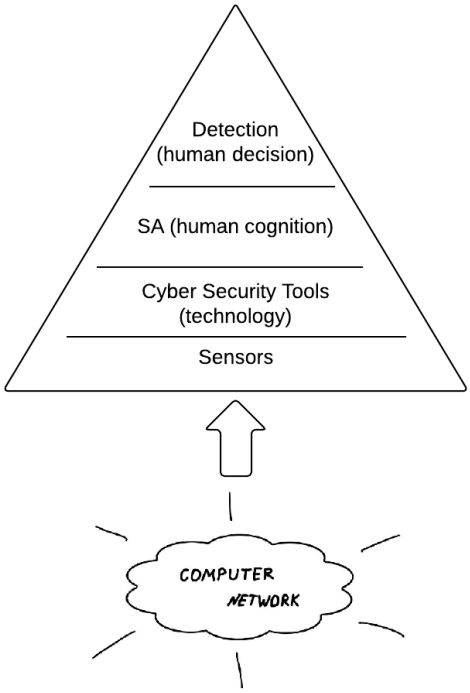

Fig. 1 The Detection Process

To properly design dynamic, adaptive technologies that support the detection process, one needs a strong, quantitative, validated model of the human cognitive processes. Otherwise, the result is often a system that works at counter-purposes with the human user, such as the infamous Microsoft paperclip that constantly changed the ordering of information in menus in a futile attempt to optimize physical movements at the greater cognitive cost to the user of constantly having to relearn a new interface.

Cognitive models are dynamic and adaptable computational representations of the cognitive structures and mechanisms involved in developing SA and processing information for decision making. Cognitive modeling technologies have been developed in the context of the cognitive sciences, which rely on theories of mind that allow for the construction of generative

models to be eventually tested against behavioral, physiological, and neural data. The advantage of cognitive models[1] resides in their ability to dynamically learn from experience, to adjust to new inputs, environments, and tasks in similar ways as humans do, and to predict performance in situations that haven't been encountered and for which data is not yet available. In this regard, cognitive models differ from purely statistical approaches, such as machine learning, that are often capable of evaluating only stable, long-term sequential dependencies from existing data but fail to account for the dynamics of human cognition, including learning processes and short-term sequential dependencies (Lebiere et al. 2003; West and Lebiere 2001).

Cognitive models are often built within a cognitive architecture. Cognitive architectures are computational representations of unified theories of cognition (Newell 1990). They represent the invariant mechanisms and structures of cognition, as implemented in the human brain. For example, the well-known ACT-R architecture (Anderson and Lebiere 1998; Anderson et al. 2004), discussed later, is organized as a distributed framework of modules, each devoted to processing a particular kind of information that is integrated and coordinated through a centralized production system module, which may represent the SA and decision making processes. A cognitive model of SA and decision making should represent the perception, comprehension, and projection status of the human mind, which are the pre-conditions to choice and decision making (Gonzalez et al. 2006). However, to build a cognitive model of cyber SA, more research on the particular cognitive challenges involved in the cyber world is needed.

Research on *cyber* SA is relatively new (Jajodia et al. 2010), and it will require large amounts of collaborative work to determine how much of what is known of SA in the physical world is applicable to the cyber world. The dynamics in the cyber environment do not follow the laws of physics and are not subject to physical constraints. For example, a cyber attack does not utilize physical weapons (a gun, a knife, a bomb) that we can see, touch, or hear and for which we have good established mental models. Cyber attacks use digital weapons that are mostly indiscernible at the human level and for which we often do not have strong intuitions. A cyber attack is not limited by geography and political boundaries. In contrast to physical wars, attacks can be highly distributed, meaning that the attacker can initiate the attack from multiple places at the same time and the same cyber attack can hit multiple targets at once (Singer and Friedman 2014). Furthermore, cyberspace is highly dynamic and it is also a distributed system, "one in which the failure of a computer you didn't even know existed can render your own computer unusable" (Lamport 1987). Thus, the traditional SA triad of perception, comprehension, and projection may have very different meanings in the cyber arena.

This chapter aims at outlining current knowledge gaps in our understanding of cognitive demands in the cyber world; and to present challenges that cognitive architectures and computational approaches face in order to represent and support SA and decision making in the cyber security domain. In what follows, we discuss some particular challenges for obtaining SA and achieving optimal decision making in the cyber world. The gaps identified and discussed in the subsequent sections are: the *cognitive* gap, namely defining a theoretical model of cyber SA within a cognitive architecture; the *decision* gap, representing learning, experience and dynamic decision making in the cyberspace; the *semantic* gap, addressing the construction of a common language and a set of basic concepts about which the security community can develop a shared understanding; the *adversarial* gap, developing ways to represent adversarial behavior; and the

[1] Note that the distinction between 'model' and 'agent' when dealing with cognitive architectures is a blurred one. In general, an agent can be conceived as a cognitive model that dynamically interacts with the environment.

network gap, scaling up models of human behavior to complex networks and cyber warfare representations. Next, we discuss existing technology developed to support the analyst and recent cognitive models of cyber SA and decision making from which new research may derive.

2 Challenges of the Cyber World and Implications for Human Cognition

In contrast to the physical world, there are many distinct cognitive challenges that a decision maker confronts in the cyber world. First, the amount of data available to the analyst is unusually large and highly diverse. This is due to the relatively inexpensive ways of collecting data (network activity) and to the number and diversity of possible data sources (each network node or piece of equipment can serve as a sensor).

Second, cyber attacks can take many forms, and each form might target different parts or services in the network. As such, an attack might be represented only in one data source or in combinations of several data sources, but not in all the data sources at the same time and in the same manner. Thus, the analyst needs to expend more effort in searching and diagnosing information to achieve the comprehension level of SA.

Third, the cyber world involves rapid and constant change. In normal day-to-day operation, changes like the maintenance of network equipment, the addition of sub-networks, and changes in services or users may be legitimate operations; however, they may also resemble signs of an attack. Furthermore, changes in network behaviors can be abrupt, drastic, and caused by both internal and external factors. For example, a sudden spike in network activity on a retailer network can be caused by an approaching holiday (external), the retailer having a sale (internal), or a cyber attack.

Fourth, the cyber SA of an analyst highly depends on the information coming from sensors (network monitoring equipment, logs, etc.). The analyst needs to constantly determine his level of trust in the sensors and whether to rely on the information coming from them; as it is not possible to directly evaluate the sensors' reliability. For example, an attacker may first compromise sensors to deceive the analyst about the status of the network before and during the attack.

Fifth, cyber attacks are adversarial digital ways of determining who gets power, wealth, and resources. Thus, beyond the SA of one individual, defenders (analysts and end users) in the cyber world need to be aware of cyber attackers. Attackers have one important advantage over defenders: they know their target and decide who, when, and how to attack. Defenders face many difficulties in identifying the origin, attribution, and goal of these attacks. In the cyber world, it becomes very difficult to determine the identity, organizational affiliation, and nationality of those sitting behind a computer with malicious intentions. Furthermore, the defender monitors the network, identifies threats, and repairs each and any vulnerability, while the attacker needs to find a single vulnerability that can be exploited. This simplified view highlights the asymmetric relationship between the defender's SA and the attacker's SA. Cyber SA for a defender, thus, must involve awareness of the attackers' SA and intentions. This is a concept that is not currently well-known in the SA literature. A good amount of research has been devoted to the concept of *Shared SA*, a requirement to perform well in teams and achieve coordination and collaboration among team members (e.g., Gorman et al. 2006; Saner et al. 2009). Shared SA represents the "degree to which team members possess the same SA on shared SA requirements" (Endsley and Jones 2001, p. 48). While the information requirements by one individual that overlap among

members of a group are essential elements for shared SA in friendly situations (Saner et al. 2009), the disparity, conflict, and disagreement of information needed to successfully deceive defenders and attackers is one of the most important weapons of agents involved in a cyber war. Thus, a concept of *Adversarial* SA needs to be developed to enhance the theory and models of theory of mind in cyber settings.

In summary, given the challenges of the cyber world and their implications for human cognition outlined above, it is clear that the development of cognitive models and computational approaches to represent and support cyber SA and decision making of the analyst are only in their infancy. In the next section, we review some existing technologies aimed at representing and supporting cyber SA and the detection of cyber attacks. We also introduce the ACT-R cognitive architecture and cognitive models aimed at representing processes involved in cyber defense. In these descriptions, we highlight the current knowledge and outline how cognitive architectures and models can be used to address these gaps.

3 Technology for Supporting an Analyst in Intrusion Detection

A cyber analyst is mainly responsible for reviewing logs from various security tools and network traffic analyzers; they compile information and report incidents based on the intrusions that are detected. Given the cognitive challenges discussed above (e.g., large amounts of raw data collected by network sensors; variable speeds and workloads of events; and complex interrelationships of various elements of a network), the analyst's ability to grasp pieces of information as a coherent *whole* diminishes when dealing with a cyber environment. An important technology that helps support cyber SA and human decision making in the detection of threats and cyber attacks is the Intrusion Detection System (IDS). IDS are relatively well-established technology, and they are widely used in different settings to automatically analyze packets for signs of possible incidents and to highlight those to the human analyst. A comprehensive review of the IDS-based methodologies and technologies that are more commonly used for intrusion detection and prevention are presented by Bernardi and colleagues (2014). IDS and their derivatives are mostly rule-based systems that require knowledge of the vulnerabilities in the networks. Snort (http://www.snort.org/) is probably the most well-known IDS: it is an open source software with millions of users, and it is considered a standard capable of performing packet sniffing and real-time traffic analysis. Snort rules are supported by an active community that improves the rules and the tool's capabilities. Other open source tools such as Bro (https://www.bro.org/) offer faster network capabilities and have also increased in popularity. Bro was developed as a research platform for intrusion detection and is commonly used by the research community.

A main challenge for the analyst is that the IDS generates a large number of false alarms, from which an analyst must identify real threats. IDSs may be used in conjunction with many other tools that help human detection. Of particular interest is the development of correlation models and the estimation of relationships between suspicious events flagged by the IDS, which may help humans detect patterns, the paths of attacks, and the attackers' intentions. Attack graphs have also been widely used to highlight alert correlations and to improve the prediction of the attackers' intentions. These attack graphs highlight the dependencies between network components and known vulnerabilities, and they may be important in providing an analyst with improved SA regarding the possible attack propagation within the network. Combining attack

graphs with dependency graphs, which capture dependencies among assets in the network, can provide the analyst with a more informed decision making process (Albanese et al. 2011).

Another way to support the analyst' cyber SA is with computational assistance tools that filter and visualize data and help prevent "cognitive overload" (Etoty et al. 2013). By and large, as Erbacher (2012) has recently pointed out, the vast majority of these state-of-the art assistance tools are targeted at network analysts with the common function of correlating cyber events within a network topology and facilitating the interpretation of low-level events (where an "anomaly" is essentially a cyber event that violates some pre-defined constraints and deviates from previously observed patterns). This kind of tool (e.g., VisAlert: http://www.visalert.com/, NVisionIP: Lakkaraju et al. 2004, etc.) leverages machine learning and information fusion techniques to extrapolate meaningful structures for the cyber analyst, but they are not designed to either provide a high-level representation of the data (which would include notions like risk management, agility handling tasks, etc.) or to factor into play the distinctive cognitive elements in genuine SA, such as perception, attention, memory, experience, reasoning capabilities, expectations, confidence, performance, etc. Hence, the aim of most existing visualizations tools is to make the data more accessible to the analyst and alleviate some of the effort of the perception phase. Such tools provide less support to the comprehension and projection phases of cyber SA. Furthermore, numerous pitfalls of visualizations can bias the analyst's SA and should be carefully considered when visualizing network data (Tufte and Graves-Morris 1983). For example, visualizations can highlight some data attributes and can lead to over-consideration of these attributes in the decision process while directing less attention to other relevant attributes.

When huge amounts of network traffic need to be analyzed, Machine Learning (ML) methods can provide a means to instantiate IDS processes (Chauhan et al. 2011; Harshna and Kaur 2013). In general, ML techniques are split into two large groups, namely "classification" and "clustering": the former aims at minimizing the number of false positives (normal events mistakenly classified as attacks) and false negatives (undetected attacks) by using labeled data sets as training examples; the objective of the latter is to extract clusters of similar patterns from a dataset, thus *de facto* creating multiple data subsets differentiated by some suitable distance measure. The main advantage of clustering is that it does not involve any training phase, which conversely makes classification more effective for a dataset where training data are available, but classification is less reusable across scenarios and less adaptive to novel situations. Among the ML classification techniques used for intrusion detection, we find Inductive Rule Generation (e.g., the Ripper system; Cohen 1995), Genetic Algorithms, Fuzzy Logics, Neural Networks, Immunological-based techniques, and Support Vector Machines. Concerning ML clustering techniques, statistical methods based on Bayes estimators and Markov models represent the most complex frameworks of analysis, where patterns can be computed in a variable time-scale and in a per-host or per-service scale. Overall, ML tools can be very efficient in handling large amounts of data and can provide meaningful insights regarding the state of a network. However, they rely on complex algorithms and intensive computational processes when detecting threats. Eventually, the analyst is provided with a recommendation without the ability to understand the details of the processes that generated that recommendation. Without the ability to acquire the appropriate level of SA, this can expose the analysts to various biases related to trust in automation and eventually harm the comprehension and projection levels of SA.

The technology to support the analyst in intrusion detection is critical to the analyst's acquisition of cyber SA and decision making. But in order to create adaptable technology that accounts for the analyst's mode of thinking, the analyst's cognitive processes and limitations

ultimately need to be represented in this technology. Next, we discuss the ACT-R cognitive architecture and the instance-based learning theory (IBLT) (Gonzalez et al. 2003), a theory of decisions from experience in dynamic tasks, which has recently been used to create cognitive models of the intrusion detection process.

4 ACT-R Cognitive Architecture

Cognitive architectures are computational instantiations of unified theories of cognition (Newell 1990). They represent the invariant mechanisms and structures of cognition, as implemented in the human brain. The ACT-R architecture (Anderson and Lebiere 1998; Anderson et al. 2004) is organized as a set of modules, each devoted to processing a particular kind of information that is integrated and coordinated through a centralized production system module (see Figure 2). Each module is assumed to access and deposit information into a buffer associated with the module, and the central production system can only respond to the contents of the buffers, not the internal encapsulated processing of the modules. Each module and associated buffer has been correlated with activation in particular brain locations (Anderson 2007). The visual module and buffer keep track of objects and locations in the visual field. The manual module and buffer are associated with control of the hands. The declarative module and retrieval buffer are associated with the retrieval of information from long-term declarative memory. The goal buffer keeps track of the goals and the internal state of the system in problem solving, while the imaginal buffer (not pictured) keeps track of problem information. Finally, the procedural module is charged with coordinating the activity of other modules by directing the flow of information between them. That module, implemented as a production system, includes components to pattern matching against buffer contents, to select a single production rule to fire at one time, and to trigger activity in various modules by directing information into their buffer.

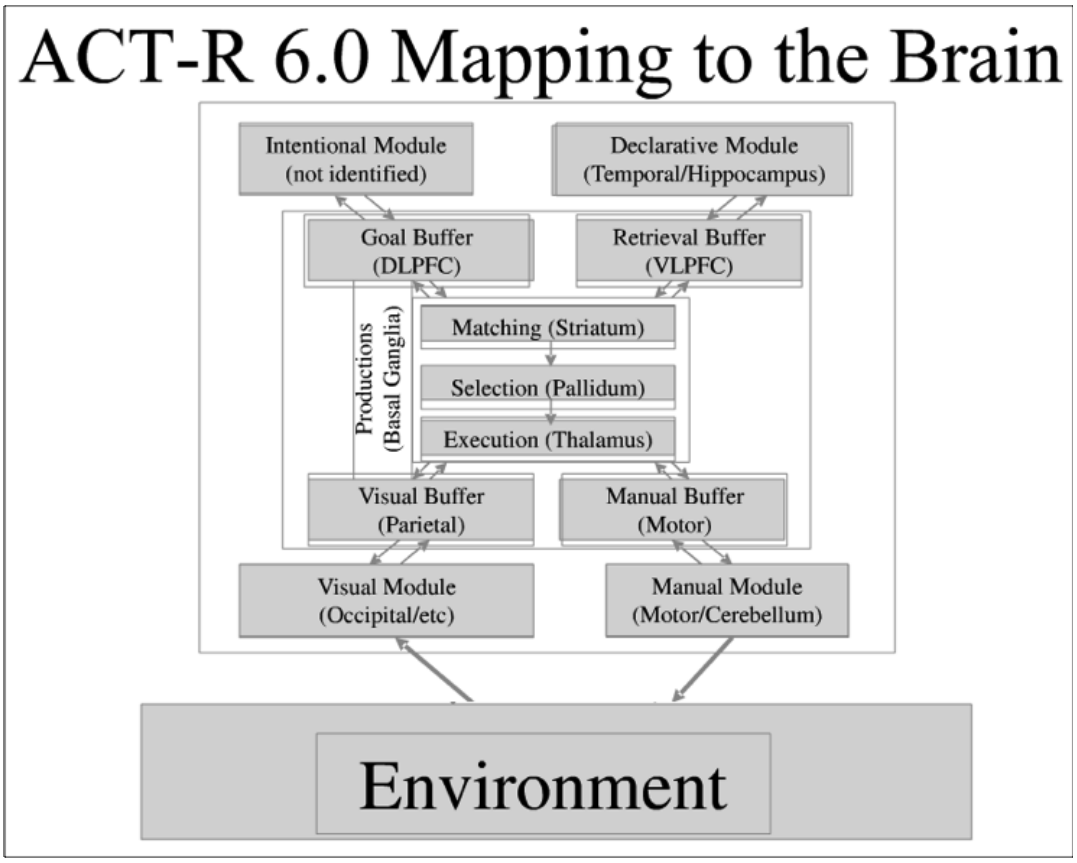

Fig. 2 ACT-R is a production system architecture with multiple modules corresponding to different kinds of perception, action, and cognitive information stores. Modules have been identified with specific brain regions. A central procedural module synchronizes information flow to and from the other modules

The declarative module and procedural module, respectively, store and retrieve information that corresponds to declarative knowledge and procedural knowledge. Procedural knowledge consists of the implicit skills that we display in our behavior, generally without any conscious awareness. Production rules represent procedural knowledge in the form of the strategies and heuristics used to manipulate that information and achieve problem solving. They specify procedures that represent and apply cognitive skill in the current context to retrieve and modify information in the buffers and transfer it to other modules. While those procedures could specify expert solutions to the problem, it is generally assumed that achieving expert levels of performance requires up to thousands of hours of experience in the most complex domains. Instead, a common hypothesis in modeling task performance is to assume that individuals rely on direct recognition or recall of relevant experience from declarative memory to guide their solution or, failing that, resort to very general problem-solving heuristics. This compute-vs-retrieve process is a common design pattern used to structure ACT-R models (Taatgen et al. 2006). For instance, it would apply in cyber security when a novice analyst would painstakingly apply a procedure to make a judgment about a new intrusion, while an expert analyst would simply recognize the pattern and make a snap judgment.

Declarative knowledge is the kind of explicit knowledge that a person can attend to, reflect upon, and usually articulate in some way (e.g., by declaring it verbally or by gesture).

Declarative knowledge in ACT-R is represented formally in terms of chunks that are structured sets of slots and associated values, which can in turn be other chunks, thus enabling the creation of complex hierarchical representations. The chunks in the declarative memory module correspond to episodic and semantic knowledge that stores the long-term experience of the model. A chunk typically integrates information available in a common context at a particular point in time in a single representational structure. Chunks are retrieved from long-term declarative memory by an activation process (see Table 1 for detailed equations) that reflects the statistics of the environment (Anderson 1993). Each chunk has a base-level activation that reflects its recency and frequency of occurrence, which accounts for the power laws of practice and forgetting that are pervasive in human behavior. Activation spreads from the current focus of attention, including goal and imaginal buffers, through associations among chunks in declarative memory to account for phenomena such as associative priming, in which the context plays an implicit role in our access to information. These associations are built up from experience and they reflect how chunks co-occur in cognitive processing. The spread of activation from one cognitive structure to another is determined by combining the weight of attentional focus from the originating cognitive structure with its associative strength to the other structure. Chunks are compared to the desired pattern specified in the retrieval buffer by using a partial matching mechanism that subtracts its degree of mismatch to the desired pattern from the activation, referred to as similarity. This is done additively for each component of the pattern and corresponding chunk value, weighted by a mismatch penalty factor. This ability to match to imperfect information allows us to deal with changing, approximate, and probabilistic environments. Finally, noise is added to chunk activations to make retrieval a probabilistic process governed by a Boltzmann (softmax) distribution, accounting for the probabilistic nature of human cognition. While the most active chunk is usually retrieved, a blending process (Lebiere 1999) can also be applied; which returns a derived output encoding the consensus value reflecting the similarities between the values of the content of all chunks, weighted by their retrieval probabilities as determined by their activations and partial-matching scores. This blending process is often used to provide a constrained way of making decisions in continuous domains as proposed in IBLT (Gonzalez 2013; Gonzalez and Dutt 2011; Gonzalez et al. 2003), which is described next.

Table 1 List of activation mechanisms in the ACT-R Architecture

Mechanism	Equation	Description
Activation	$A_i = B_i + S_i + P_i - \varepsilon_i$	B_i: Base-level activation reflects the recency and frequency of use of chunk *i* S_i: Spreading activation reflects the effect that buffer contents have on the retrieval process P_i: Partial matching reflects the degree to which the chunk matches the request ε_i: Noise value includes both a transient and (optional) permanent component (permanent component not used by the integrated model)
Base-Level	$B_i = \ln(\sum_{j=1}^{n} t_j^{-d}) + \beta_i$	n: The number of presentations for chunk *i* t_j: The time since the j^{th} presentation d: A decay rate (not used by the integrated model) β_i: A constant offset (not used by the integrated model)
Spreading Activation	$S_i = \sum_k \sum_j W_{kj} S_{ji}$	k: Weight of buffers summed over are all of the buffers in the model j: Weight of chunks which are in the slots of the chunk in buffer *k* W_{kj}: Amount of activation from sources *j* in buffer *k* S_{ji}: Strength of association from sources *j* to chunk *i*
	$S_{ji} = S - \ln(fan_{ji})$	S: The maximum associative strength (set at 4 in the model) fan_{ji}: A measure of how many chunks are associated with chunk *j*

Partial Matching	$P_i = \sum_k PM_{ki}$	**P:** Match scale parameter (set at 2) which reflects the weight given to the similarity M_{ki}**:** Similarity between the value *k* in the retrieval specification and the value in the corresponding slot of chunk *i* The default range is from 0 to -1 with 0 being the most similar and -1 being the largest difference
Declarative Retrievals	$P_i = \frac{e^{A_i/s}}{\sum_j e^{A_i/s}}$	P_i**:** The probability that chunk *i* will be recalled A_i**:** Activation strength of chunk *i* $\sum A_j$**:** Activation strength of all of eligible chunks *j* **s:** Chunk activation noise
Blended Retrievals	$V = \min \sum_i P_i \cdot (1 - Sim(V, V_i))^2$	P_i**:** Probability from declarative retrieval Sim_{ij}: Similarity between compromise value *j* and actual value *i*

5 Instance-Based Learning Theory and Cognitive Models

The notion that learners have a general-purpose mechanism whereby situation-decision-utility triplets are stored as chunks and later retrieved to generalize solutions to future decisions originates from *instance-based learning theory (IBLT)* (Gonzalez et al. 2003). IBLT is a theory of decisions from experience in dynamic tasks. A simple cognitive model, derived from IBLT, has recently been proposed for representing individual learning and for reproducing choice behavior in repeated binary choice tasks (Gonzalez and Dutt 2011; Lejarraga et al. 2012). This model has shown to be a robust accounting of the choice and learning process in a large variety of tasks and environmental conditions (for a summary, see Gonzalez 2013). Its greatest strength is that it offers a single learning mechanism to account for behavior observable in multiple paradigms and decision making tasks (for a summary, see Gonzalez 2013). However, Gonzalez and colleagues (2003) argue that the strength of IBLT is the explanations of decision making in complex dynamic situations, such as cyber security. With the aim of scaling up from simple binary choice models to the type of complex dynamic tasks that IBLT intended to explain, Gonzalez and colleagues have used the cognitive model for binary choice to represent the detection process in cyber security.

Dutt, Ahn, and Gonzalez (2011) proposed an IBL model to study cyber SA. The model represented the cognitive processes of a cyber-security analyst who needs to monitor a computer network and detect malicious network events that constitute a simple island-hopping cyber attack. In this model, the memory of a simulated analyst was pre-populated with instances encoding possible network events, including a set of attributes (e.g., IP address, whether the IDS issued an alert, etc.) that define a network event. An instance also included the analyst's decision regarding that specific combination of attributes, meaning whether the analyst decided that the event (i.e., set of attributes and their values) described malicious network activity or not. Finally, an instance also stored the outcome of that decision, indicating whether the event actually represented a malicious network activity or not. Controlling the representation of the analyst's memory provided the ability to manipulate situation awareness by adjusting the amount of instances in memory that represent malicious network activity. For example, the memory of a very selective analyst had 75% malicious instances and 25% non-malicious instances, while a less selective analyst's memory had 25% malicious instances and 75% non-malicious instances. When making a decision about whether a new network event is part of a malicious network activity or not, the model retrieved similar instances from memory according to the cognitive judgment mechanisms. Through the process of judging, the modeled analyst accumulated evidence that can indicate if there is an ongoing cyber attack. The risk tolerance parameter of the

model governed this accumulation process. The number of malicious network events that the model detected was constantly compared to the analyst's risk tolerance, and once the number of malicious events was equal to or higher than the risk tolerance, the modeled analyst declared that there is an ongoing cyber attack. Thus, risk tolerance served as a threshold for evidence accumulation and risk taking.

The results from simulating different cyber analysts demonstrated that both the risk tolerance level and the past experiences of the analyst affect the analyst's cyber SA, with the effect of experiences (in memory) being slightly more impacting than risk tolerance. This work also highlighted the importance of modeling the adversary's behavior, by comparing the influence of impatient and patient attacker strategies on the performance of the defender. Patient attacker strategy and longer delays between the threat incursions on the network can challenge the security analyst and decrease her ability to detect threats. Thus, the cognitive model was capable of capturing the phenomenon that some attack patterns are more challenging than others to the simulated security cyber analyst.

6 Research Gaps for Understanding the Cognitive Demands of the Cyber World

Many advances need to be made in several research directions to make cognitive models useful and effective in representing and supporting the job of a cyber security analyst. Based on the current state of technology discussed above, we identified five gaps in our understandings of the cognitive demands of the cyber world.

6.1 The Cognitive Gap: Mapping Cognitive Architecture Mechanisms to Cyber SA

The general processes of a cognitive architecture such as ACT-R can be mapped systematically onto the concepts of cyber SA, such that the distinct levels of situation awareness can be related to concrete cognitive mechanisms. This mapping does not take the form of a one-to-one correspondence between cyber SA concepts and cognitive modules, but it instead maps those concepts onto modeling idioms that leverage multiple modules using common patterns. The first level of cyber SA corresponds to the processes involved in the direct acquisition of information from the environment. This perception level can be directly associated with the perceptual modules of the ACT-R cognitive architecture, including the visual and aural modules. However, those modules do not operate on their own, but through the direct supervision and control of the procedural module. Attention is a fundamental construct that reconciles the limited processing resources of our cognitive (including perceptual) modules with the considerable demands arising from the open-ended complexity of the external world. Attentional focus is used to decompose complex external scenes, like a complex cyber security display, into simple components that can be processed directly by our perceptual systems.

The typical flow of control for perception in an ACT-R model (e.g., Anderson et al. 2004) proceeds in a top-down manner. While attention can be directed by external events in the environment, effective performance of complex tasks in information-rich environments typical of cyber security requires structured, goal-directed perceptual processing of information. The

first step of perception is therefore a request for a location that matches a specific content condition.[2] This location might already be known if the user is sufficiently familiar with his environment and the environment is stable enough, in which case it will be provided by retrieval from declarative memory. Otherwise, it is directly supplied by a production rule, if sufficient experience has transformed that knowledge into a skill through production compilation. If not, the location will be determined by searching the environment to match the specified condition. Once the location has been obtained, it is supplied to the visual buffer to trigger processing of that area of the visual field in the visual module. This will result in the chunk representing the object recognized at that location to be returned in the same visual buffer. That chunk is then transferred to the imaginal buffer holding the representation of the current situation being elaborated on, which is where the process of comprehension starts. Hence, in the context of cyber SA, this phase correspond to the process through which a cognitive model retrieves and encodes source and destination IP address, protocol type, and other attributes of the network. Comprehension corresponds to the second level of cyber SA, which results in the semantic representation of a perceived situation, a product of the cognitive process known as sensemaking (Klein et al. 2006a). According to Klein et al. (2006b), sensemaking is the process of abstraction that maps concrete situations to the general by using mental representations called frames, which correspond to structured conceptual models of the world. Lebiere et al. (2013) describe how sensemaking is fundamentally compatible with IBLT, and more specifically how frames can be mapped onto the chunk representations of situations used in that process. For instance, in the domain of geospatial intelligence, frames correspond to a pattern of input data, aggregating layers of information from independent sensors and associating them with specific hypotheses. "Comprehension" thus corresponds to the process of gradually aggregating the information from perception into hierarchical chunks implementing integrated frames. In the next section, we argue that "ontologies" can enhance this second level of cyber SA by mapping ACT-R declarative chunks to highly expressive semantic structures that formally specify the conceptual models encapsulated in frames. Going back to cyber SA and detection, in this comprehension phase, IP address obtained during the perception phase are organized into categories that reflect whether it is internal or external to the monitored network. This type of reasoning can also bind an event (e.g., an IDS alert) and the reason that the event occurred (e.g., an IDS rule regarding the maximal number of open connections for a communication protocol), thus generating a hypothesis for the observed behavior that will drive further investigations. The third level of cyber SA corresponds to the process of projection, or the generation of expectations about future states of the system. Those changes in system state can result from the actions of the decision maker, from those of an opponent or teammate, or from other independent parts of the system. Projection is essential in evaluating the effect of potential actions by including feedback from the outcome of past actions. Because many cyber security interactions are fundamentally adversarial, it is essential to also being able to generate expectations of the opponent's future actions, encompassing both independent actions and actions taken in response to one's own decisions. Finally, since the actions of third parties, such as system users, also impact the outcome of security measures, generating expectations of their actions is crucial to projecting future system states and effective system control. From the cyber SA perspective, this phase occurs after perceiving an IDS alert and comprehending that it was generated by a rule that limits the number

[2] This discussion will be focused on visual attention, though the same principles apply to other perceptual modules such as auditory perception.

of open connections. Now, when the number of open connections exceeds the limit, projection is used to evaluate whether this is a temporary benign spike in the demand for a service or if it is an indication for a cyber attack. Making such a decision requires integration of additional information that can be perceived and comprehended explicitly from the environment, like the source IP addresses of the connections, as well as consideration of implicit information like the consequence to the network if the number of open connections will continue to increase.

6.2 The Semantic Gap: Integrating Cognitive Architectures with Ontologies of Cyber Security

In the previously mentioned models, modelers themselves directly specified the semantics of the representation. In order to enable full-fledged reasoning capabilities in cognitive architectures, these systems need to incorporate "re-usable declarative representations that correspond to objects and processes of the world" (McCarthy 1980). Similarly, cognitive architectures must provide a way to represent world entities (Sowa 1984), i.e., an "ontology"[3]. An ontology is a language-dependent cognitive artifact committed to a certain conceptualization of the world by means of a given language[4] (Guarino 1998). Thus, in broad terms, an ontology corresponds to a semantic model of the world (or of a portion of it, i.e., a "domain"): when the model is simply described in natural language, an ontology reduces to a *dictionary*, *thesaurus*, or *terminology*; when the model is expressed as an axiomatic theory (e.g., in first order logic), it is called a *formal ontology*. Ultimately, if logical constraints are encoded into machine-readable formats, formal ontologies take the form of *computational ontologies*, and enter *de facto* in the family of *semantic technologies*, which include search engines, automatic reasoners, knowledge-based platforms, etc. In the context of a cognitive architecture like ACT-R, computational ontologies can extend the semantics of the chunks stored in declarative memory. Although these extensions are not usually required by ACT-R models that perform relatively narrow cognitive tasks, declarative memory should be designed to encompass a rich spectrum of concepts when dealing with decision making in complex scenarios like cyber operations, including classifications of cyber security policies, risks, attacks, system's functionalities, human responsibilities, user's privileges, as well as the mutual connection among them. Widening the scope beyond ACT-R, state of the art work on cognitive architectures has also gone in the direction of mapping ontologies (like Cyc, see Lenat et al. 1985) to declarative memory (see Ball et al. 2004; Best et al. 2010; Edmond 2006). It aims to enhance not only the "capability" of representing the available knowledge of a domain but also the functionality of automatically deriving inferences from it, a feature that would also help to increase the "Comprehension" level in cyber SA. In this regard, the role of ontologies in cognitive architectures is to 1) formally characterize chunks in long-term memory that depict conceptual models of situations (frames) and 2) foster *automaticity* of certain cognitive tasks, "that significantly benefit SA by providing a mechanism for overcoming limited attention" and improve the decision making process.

---

[3] This was the genesis of using the word 'ontology' in AI. Ontology, 'the study of being as such' – as Aristotle named it – originated as a philosophical discipline.

[4] Guarino distinguishes between 'Ontology' as a discipline (with the capital 'o') and 'ontologies' as engineering cognitive artifacts.

There has been little work on ontologies for cyber security and cyber warfare. An ontology of IDS is discussed by Undercoffer, Joshi, and Pinkston (2003); within a broader paper, there is a brief discussion of an ontology for DDoS attacks (Kotenko 2005); and a general ontology for cyber warfare is discussed in D'Amico et al. (2009). Obrst et al. (2012) provides the best sketch of a cyber warfare ontology, and the scale of the project and its difficulties are discussed by Dipert (2013). With regard to human users and human-computer interface, the most important step in understanding a complex new domain involves producing accessible definitions and classifications of entities and phenomena. Mundie (2013) stressed this point when talking about the Jason Report (The MITRE Corporation 2010). Discussions of cyber warfare often begin with the difficulties created by misused terminology (such as characterizing cyber espionage as an "attack"). The Joint Chiefs of Staff created a list of cyber term definitions (Joint Staff Department of Defense 2010) that has been further developed and improved in a classified version. Nevertheless, none of these definitions has been encoded in OWL (Staab and Studer 2003) or in any other computational semantic format, which is a necessary requirement to make them machine-understandable. Likewise, various agencies and corporations (NIST, MITRE, Verizon) have formulated enumerations of types of malware, vulnerabilities, and exploitations, sometimes expressed in XML-based semantics: but without a common vocabulary, their sprawling English descriptions in large, incompatible databases are not directly machine-usable and are nearly impossible to maintain. Efforts that have been made toward developing computational ontologies of cyber security and cyber warfare typically do not work within any standard framework and do not utilize existing military reference ontologies such as UCORE-SL, which define concepts such as the notion of "agent," "organization," "artifact," "weapon," etc.

As a consequence of this general deficiency, one of the first and perhaps most generally useful tasks that will need to be completed to fill the "semantic gap" is to collect definitions of key cyber security concepts that are currently scattered across existing ontologies, controlled vocabularies, doctrines, and other documental resources and to suitably harmonize them in a homogenous computational ontology. As a second step, the capabilities of this cybersecurity ontology will have to be dynamically tested in cognitive models of decision making in cyber operations.

6.3 The Decision Gap: Representing Learning, Experience, and Dynamic Decision Making in the Cyber World

Given the complexity and variability of the cyber environment, there is an ongoing effort to provide decision makers with tools that can support their decision process and provide insights to manage the complex dynamics of the cyber world. To gain and maintain situation awareness, the decision maker is constantly required to make multiple and interdependent decisions in a highly dynamic environment. Dynamic decision making requires an understanding of multiple, interrelated attributes and the ability to anticipate the way that the environment will develop over time. Making the right decision and acting appropriately and in a timely manner can maximize the decision value (Brehmer 1992; Edwards 1962; Gonzalez 2005; Gonzalez et al. 2005).

The modeling of human decision processes in cyber security highlights some important aspects of cyber SA that cognitive models need to account for. For example, pattern recognition under uncertainty represents a defender's attempt to find patterns in the attacker's sequence of

actions in order to predict the attacker's next operation and to provide the best response to it. However, if the attacker is aware of these attempts to detect sequential dependencies, one possible path of action is to constantly change the malicious operations and to exploit the sequential dependencies. Cognitive models in ACT-R (Anderson and Lebiere 1998, 2003) and neural networks (West and Lebiere 2001) are capable of accounting for the human ability to detect sequential dependencies, and they use the perceived sequence to project the next action that an opponent will most likely take in a strategic interaction. Through their natural stochasticity, those models can balance the exploitation of the opponent's patterns with some measure of deception and self-protection by avoiding becoming too predictable themselves. Also, cognitive models such as those derived from ACT-R and IBLT provide the capability to learn from experience and the ability to utilize past experiences in novel decision situations.

Human decision makers use the same cognitive system for a vast array of divergent tasks. The underlying cognitive system represents a highly efficient, multipurpose mechanism that has evolved to be as effective as possible across a wide variety situations and conditions (West et al. 2006). Cognitive architectures share the same flexibility and diversity, and as such can efficiently represent and capture human decision making in cyber security. However, continued efforts are needed to maintain and update the formal representation of the cyber environment that the architectures use. This requirement emphasizes the need for cognitive architectures to develop better and more efficient models of perception and information encoding. For cognitive architectures to serve a meaningful role in future cyber security engagements, two main aspects should be carefully developed: the first is the flexibility of reasoning that underlies human adaptivity and the second is the active and efficient perceptual processes that search, detect, and encode information in a dynamic environment.

6.4 The Adversarial Gap: Representing Adversarial Cyber SA and Decision Making

Cognitive architectures provide rich and flexible modeling environment. Using these architectures, it is possible to generate models that represent the analyst' decision making process and SA, as well as models of the adversary. For each of these models, there is a need to define knowledge base, learning processes, and decision making process. Furthermore, the models of the analyst and the adversary interact within a defined environment (i.e., the cyber world) that dictates a set of possible action each model can choose from. Thus, there is a need to define the possible interactions between multiple cognitive models. In addition to defining the possible interactions, there is a need to define how and what kind of feedback the models would receive regarding the outcomes of their combined decision making processes. Issues concerning delayed feedback and incomplete or imperfect feedback are highly relevant when modeling studying decision making and learning in dynamic systems. Therefore, a comprehensive formal representation that can bring together the analyst, the adversary, and the environment in which they interact is needed. Game theory has been successfully used to capture the essence of complex and dynamic situations that involves two or more agents that interact within a well-defined environment. We posit that combining game theoretical perspective and cognitive modeling can provide a controllable, but still ecological valid, representation of interactions in the cyber world and serve as a potent framework for studying cyber SA.

Game theory has been popularized as a potent approach to characterize and analyze decisions in situations that involve social dilemmas and conflict situations. Stackelberg games

have been used to model and capture the strategies of defenders and attackers in airport security, as well as for optimizing resources allocation in sensitive settings (Pita et al. 2008). Similarly, game theory has been used for decision making in cyber security (Alpcan and Baar 2011; Grossklags et al. 2008; Lye and Wing 2005; Manshaei et al. 2013; Roy et al. 2010). However, most game-theoretic approaches to security hold some limitations and assume either static game models or games with perfect or complete information (Roy et al. 2010). To some extent, these assumptions misrepresent the reality of the network security context where situations are highly dynamic and the decision maker must rely on imperfect and incomplete information. To overcome this, recent studies that apply game theory to security attempt to account for the bounded rationality of human actors, especially human adversaries (Pita et al. 2012). However, this and other game-theoretic approaches still do not fully address the cognitive mechanisms like memory and learning that drive the human decision making processes and can provide a first-principled predictive account of human performance, including both capabilities and suboptimal biases.

Behavioral Game Theory relaxes some of the constraints of Game Theory with the study of human decision makers and how they interact in strategic situations involving more than one decision maker (Camerer 2003). Using Behavioral Game Theory, it is possible to address some of the limitations imposed by game-theoretic approaches and examine how learning from experience and adaptation to the environment influences decision making and risk taking in cyber security (Gonzalez 2013).

As discussed earlier, ACT-R and IBLT have proven to be highly beneficial to studying the interplay between learning and decision making processes of an individual. One ongoing effort aims at scaling up cognitive models to study interactions between two or more decision makers in social conflicts like the Prisoner's Dilemma (Gonzalez et al. in press) and the Chicken Game (Oltramari et al. 2013). However, scaling up models of human cognition and SA to cyber worlds with more than two agents involved is still a challenge (Gonzalez 2013). An important issue for all levels of SA is the availability of information regarding the other entities. Recently, cognitive models have been extended to study how the availability of information and the source of the information influence decision making and learning.

Recent studies examine how the availability of descriptive and experiential information influences interactions in social dilemmas (Martin et al. 2013; Oltramari et al. 2013). The key findings of these studies suggest that information is needed for cooperation, and the lack of information fostered situations in which one decision maker tended to exploit the other. Another relevant finding is related to trust and its role in cooperative behavior, indicating that decision makers dynamically weigh the partner's information based on surprise (i.e., the gap between expectations or projections and the observed outcome). Learning models that incorporate surprise into the decision process and combine both descriptive and experiential information can capture the complex dynamics of iterated interaction between two decision makers in conflict situations (Gonzalez et al. in press; Ben-Asher et al. 2013). Overall, these finding emphasize the interplay between information and cognitive processes in order to achieve SA and finally making a decision.

6.5 The Network Gap: Addressing Complex Networks and Cyber Warfare

Cyber warfare is the extension of the traditional attacker-defender concept that involves multiple units (individual, state-sponsored organizations, or even nations) simultaneously executing offensive and defensive operations through networks of computers. In a cyber war, units can execute attacks against targets in a cooperative and simultaneous manner. Any defending unit can also be attacked by multiple enemies, eventually acting as both attacker and defender at the same time.

The dynamics of a cyber war, which are driven by multiple decision makers making simultaneous decisions, are hard to predict. Achieving and maintaining SA in such an environment is crucial and at the same time challenging. The fact that multiple units operate simultaneously in the environment might imply that a decision maker has to maintain SA in different levels. The decision maker has to perceive, comprehend, and make projections regarding interactions in which the unit itself is involved directly, interactions between other units which do not involve the decision maker directly, and the overall aggregated SA at the environment level. Scaling up cognitive models of SA from the dyad perspective (an analyst and an adversary) to the SA needed in an environment where large networks of units can interact simultaneously requires careful consideration and examination of environmental attributes and their relation to SA. For example, the topology of the network that connects units involved in a cyber conflict has an extensive impact on the availability of information, trust in information, and information propagation.

To support SA and decision making in large scale cyber conflicts, simulations using multiple cognitive models connected in a network can provide predictions and answer what-if questions. Similarly, simulations that combine multiple cognitive models and human decision makers can train humans to acquire and maintain SA in cyber conflicts. Recently, there has been an increasing interest in N-Player models of social conflict that share some similarities with cyber warfare (Kennedy et al. 2010; Hazon et al. 2011). In parallel, there are attempts to study cyber attacks and cyber warfare through multi agent-based modeling (e.g., Kotenko 2005, 2007). However, many of these models use strategic agents and not cognitive models. Such strategic agents are designed to execute an optimal strategy, rather than learn the maximizing strategies from experience; and thus not only fail to replicate SA, human learning, and decision making mechanisms but are fundamentally incapable of coping with fluid, dynamic situations commonly encountered in cyber warfare.

The CyberWar Game (Ben-Asher and Gonzalez 2014) is a multi-player framework that aims to capture some of the characteristics and the dynamics of the environment in cyber warfare and aspects of the decision maker. It is inspired by Hazon et al.'s (2011) N-Player model. Considering important aspects of cyber warfare and conflicts in general, the CyberWar Game introduces two relevant concepts that characterize a player: power and assets. In the context of cyber warfare, power represents the ability to successfully accomplish a goal, which for a defender is to block an attack and for an attacker is to accomplish a malicious goal. Power can be seen as a representation of the robustness of cyber security infrastructure and is likely to be a function of investment in cyber security. An asset is an abstraction of what the defender is trying to protect and what the attacker wants to gain. In general, assets are the motivation for building both defense system and attack systems, and selfish assets maximization is the shared goal of all the decision makers in this environment. Power represents the potential of these systems to achieve this goal.

In this paradigm, as illustrated in Figure 3, several players simultaneously attack each other or defend themselves from attacks. Thus, a player is not assigned to be an attacker or a defender in this game, but it is the players' decision what role they play. Furthermore, this resembles distributed attacks over the network and also incorporates the idea that power can be distributed between multiple goals. A player needs SA and learning processes to identify who might try to attack and who can be a valuable target to attack. For example in Figure 3, Player 1 and Player 3 are likely to attack Player 2 as she is the weakest player. However, if Player 1 invests all her power in the attacking without defending from Player 3, Player 3 can take advantage and attack only Player 1, who has the highest asset's value. The decision of whether or not to attack an opponent is not straightforward, as the player has to incorporate additional aspects like the cost of attack, the cost of defense, the attack severity (i.e., what percentage of opponent assets it is possible to gain when winning an attack), and the effectiveness of defense. Frameworks like the CyberWar Game allow us to examine the role of SA at the operational level (who to attack and from whom to defend), as well as at the tactical and strategic levels (which coalition to join).

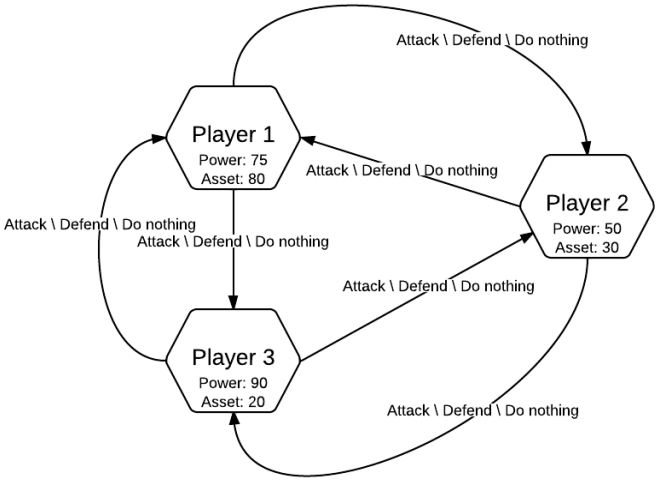

Fig. 3 General description of a CyberWar Game

7 Summary

Human cognition is central to our understanding of the challenges of the cyber world. Cyber security is an extremely complex domain that stretches and challenges many of our theories and concepts of situation awareness and decision making. Current theories of SA have been developed for the physical world, and research is needed to determine whether and how much of what we currently know is applicable or useful for cyber security. The process of detection (protecting networks against illegal intrusions) illustrates the challenges involved in cyber security and the need for integration of information technology and computational representations of human situation awareness. Cognitive models are dynamic and adaptable computational representations of the cognitive structures and mechanisms involved in developing SA and processing information for decision making. Cognitive models differ from

purely statistical approaches, such as machine learning, that are often capable of evaluating only stable, long-term sequential dependencies from existing data but fail to account for the dynamics of human cognition, including learning processes. An important technology that helps support cyber SA and human decision making is the Intrusion Detection System (IDS). Other assistance tools are targeted at network analysts with the common function of correlating cyber events within a network topology and facilitating the interpretation of low-level events. The aim of most existing visualizations tools is to make the data more accessible to the analyst and alleviate some of the effort of the perception phase. Such tools provide less support to the comprehension and projection phases of cyber SA. Machine Learning (ML) methods can provide a means to instantiate IDS processes and are often divided in two large groups, namely "classification" and "clustering." Eventually, the analyst is provided with a recommendation without the ability to understand the details of the processes that generated that recommendation. Without the ability to acquire the appropriate level of SA, this can expose the analysts to various biases related to trust in automation and eventually harm the comprehension and projection levels of SA. In order to create adaptable technology that accounts for the analyst's mode of thinking, the analyst's cognitive processes and limitations must be represented in a cognitive model. Cognitive models are often built within a cognitive architecture. Cognitive architectures are computational representations of unified theories of cognition and the ACT-R architecture is an example. IBLT is a theory of decisions from experience in dynamic tasks; the strength of IBLT is the explanations of decision making in complex dynamic situations, such as cyber security.
An IBL model to study cyber SA represented the cognitive processes of a cyber-security analyst who needs to monitor a computer network and detect malicious network events that constitute a simple island-hopping cyber attack. When making a decision about whether a new network event is part of a malicious network activity or not, the model retrieved similar instances from memory according to the cognitive judgment mechanisms. The model illustrates how both the risk tolerance level and the past experiences of the analyst affect the analyst's cyber SA. The current knowledge gaps in our understanding of cognitive demands in the cyber world are: the cognitive gap, namely defining a theoretical model of cyber SA within a cognitive architecture; the decision gap, representing learning, experience and dynamic decision making in the cyberspace; the semantic gap, addressing the construction of a common language and a set of basic concepts about which the security community can develop a shared understanding; the adversarial gap, developing ways to represent adversarial behavior; and the network gap, scaling up models of human behavior to complex networks and cyber conflict representations. Together, the descriptions of these gaps present a roadmap for new research and development of cognitive-aware technologies that would support the analyst's cyber SA and decision making process.

References

Albanese M, Jajodia S, Pugliese A, Subrahmanian VS (2011) Scalable analysis of attack scenarios. In: Atluri V, Diaz C (eds.) Lecture notes in computer science, vol. 6879. Springer-Verlag, Berlin, p 415-433

Alpcan T, Basar T (2011) Network security: A decision and game-theoretic approach. Cambridge University Press, New York

Anderson JR (1993) Rules of the mind. Lawrence Erlbaum Associates, Hillsdale, NJ

Anderson JR (2007) How can the human mind occur in the physical universe? Oxford University Press, Oxford

Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y (2004) An integrated theory of the mind. Psych Rev 111(4):1036-1060

Anderson JR, Lebiere C (1998) The atomic components of thought. Lawrence Erlbaum Associates, Hillsdale

Anderson JR, Lebiere C (2003) The Newell test for a theory of cognition. Behav Brain Sci 26(5):587-639

Ball J, Rodgers S, Gluck K (2004) Integrating ACT-R and Cyc in a large-scale model of language comprehension for use in intelligent agents. In: Proceedings of the nineteenth national conference on artificial intelligence. AAAI Press, Menlo Park, p 19-25

Ben-Asher N, Dutt V, Gonzalez C (2013). Accounting for integration of descriptive and experiential information in a repeated prisoner's dilemma using an instance-based learning model. In: Kennedy B, Reitter D, Amant RS (eds) Proceedings of the 22nd annual conference on behavior representation in modeling and simulation. BRIMS Society, Ottawa

Ben-Asher N, Gonzalez C (2014) The CyberWar Game: A behavioral modeling of cyber warfare (in preparation)

Bernardi P, McLaughlin K, Yang Y, Sezer S (2014) Intrusion detection systems for critical infrastructure. In: Pathan A-SK (ed) The state of the art in intrusion prevention and detection. CRC Press, Boca Raton, p 115-138

Best BJ, Gerhart N, Lebiere C (2010) Extracting the ontological structure of OpenCyc for reuse and portability of cognitive models. In: Proceedings of the 19th conference on behavior representation in modeling and simulation. Curran Associates, Red Hook, p 90-96

Brehmer B (1992) Dynamic decision making: Human control of complex systems. Acta Psychol 81(3):211-241

Camerer CF (2003) Behavioral game theory: Experiments in strategic interaction. Princeton University Press, Princeton

Chauhan A, Mishra G, Kumar G (2011) Survey on data mining techniques in intrusion detection. Int J Sci Eng Res 2(7):2-4

Cohen, WW (1995) Fast effective rule induction. In: Proceedings of the 12th international conference on machine learning. Morgan Kaufmann, Lake Taho

D'Amico A, Buchanan L, Goodall J, Walczak P (2009) Mission impact of cyber events: Scenarios and ontology to express the relationship between cyber assets. Available online. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA517410

Dipert R (2013) The essential features of an ontology for cyber warfare. In: Lowther A, Yannakogeorgos P (eds) Conflict and cooperation in cyberspace: The challenge to national security. Taylor & Francis, Boca Raton, p 35-48

Dutt V, Ahn Y-S, Gonzalez C (2011) Cyber situation awareness: Modeling the security analyst in a cyber-attack scenario through instance-based learning. In: Li Y. (ed) Lecture notes in computer science, vol. 6818. Springer-Verlag, Berlin, p 281-293

Edwards W (1962). Dynamic decision theory and probabilistic information processing. Hum Factors 4(2):59-73

Emond B (2006) WN-LEXICAL: An ACT-R module built from the WordNet lexical database. In: Fum D, Del Missier F, Stocco A (eds) Proceedings of the seventh international conference on cognitive modeling, University of Trieste, Trieste, 5-8 April 2006

Endsley MR (1988) Design and evaluation for situation awareness enhancement. Hum Fac Erg Soc P 32(2):97-101

Endsley MR, Jones WM (2001) A model of inter- and intrateam situation awareness: Implications for design, training and measurement. In: McNeese M, Salas E, Endsley MR (eds) New trends in cooperative activities: Understanding system dynamics in complex environments. HFES, Santa Monica, p 46-67

Erbacher RF (2012) Visualization design for immediate high-level situational assessment. In: Proceedings of the ninth international symposium on visualization for cyber security. ACM, New York, p 17-24

Etoty RE, Erbacher RF, Garneau C (2014) Evaluation of the presentation of network data via visualization tools for network analysis. Technical Report #ARL-TR-6865, Army Research Lab, Adelphi MD, 20783

Gonzalez C (2005) Decision support for real-time dynamic decision making tasks. Organ Behav Hum Dec 96(2):142-154

Gonzalez C (2013). The boundaries of Instance-based Learning Theory for explaining decisions from experience. In: Pammi VS, Srinivasan N (eds) Progress in brain research, vol. 202. Elsevier, Amsterdam, p 73-98

Gonzalez C, Ben-Asher N, Martin JM, Dutt V (2014) A cognitive model of dynamic cooperation with varied interdependency information. Cog Sci (in press)

Gonzalez C, Dutt V (2011). Instance-based learning: Integrating decisions from experience in sampling and repeated choice paradigms. Psychol Rev 118(4):523-551

Gonzalez C, Juarez O, Endsley MR, Jones DG (2006). Cognitive models of situation awareness: Automatic evaluation of situation awareness in graphic interfaces. In: Proceedings of the fifteenth conference on behavior representation in modeling and simulation. Simulation Interoperability Standards Organization, Baltimore, p 45-54

Gonzalez C, Lerch JF, Lebiere C (2003) Instance-based learning in dynamic decision making. Cog Sci 27(4):591-635

Gonzalez C, Vanyukov P, Martin MK (2005) The use of microworlds to study dynamic decision making. Comput Hum Behav 21(2):273-286

Gorman JC, Cooke NJ, Winner JL (2006) Measuring team situation awareness in decentralized command and control environments. Ergonomics 49(12-13):1312-1325

Grossklags J, Christin N, Chuang J (2008) Secure or insure? A game-theoretic analysis of information security games. In: Proceedings of the 17th international conference on world wide web. ACM, New York, p 209-218

Guarino N (1998) Formal ontology and information systems. In: Guarino N (ed) Formal ontology in information systems. IOS Press, Amsterdam, p 3-15

Harshna, Kaur N (2013) Survey paper on data mining techniques of intrusion detection. Int J Sci Eng Technol Res 2(4):799-802

Hazon N, Chakraborty N, Sycara K (2011) Game theoretic modeling and computational analysis of n-player conflicts over resources. In: Proceedings of the 2011 IEEE international conference on privacy, security, risk and trust and IEEE international conference on social computing. Conference Publishing Services, Los Alamitos, p 380-387

Jajodia S, Liu P, Swarup V, Wang C (2010) Cyber situational awareness: Issues and research. Springer, New York

Joint Staff Department of Defense (2010). Joint terminology for cyber operations. Available online. http://publicintelligence.net/dod-joint-cyber-terms/

Kennedy WG, Hailegiorgis AB, Rouleau M, Bassett JK, Coletti M, Balan GC, Gulden T (2010) An agent-based model of conflict in East Africa and the effect of watering holes. In: Proceedings of the 19th conference on behavior representation in modeling and simulation. Curran Associates, Red Hook, p 112-119

Klein G, Moon B, Hoffman RR (2006a) Making sense of sensemaking 1: Alternative perspectives. IEEE Intell Syst 21(4):70-73

Klein G, Moon B, Hoffman RR (2006b) Making sense of sensemaking 2: A macrocognitive model. IEEE Intell Syst 21(5):88-92

Kotenko I (2005) Agent-based modeling and simulation of cyber-warfare between malefactors and security agents in internet. In: Merkuryev Y, Zobel R, Kerckhoffs E (eds) Proceedings of 19th European conference on modeling and simulation, Riga Technical University, Riga, 1-4 June 2005

Kotenko I (2007) Multi-agent modelling and simulation of cyber-attacks and cyber-defense for homeland security. In: Proceedings of the 4th IEEE workshop on intelligent data acquisition and advanced computing systems: technology and applications. IEEE, Los Alamitos, p 614-619

Lakkaraju K, Yurcik W, Lee AJ (2004) NVisionIP: NetFlow visualizations of system state for security situational awareness. In: Proceedings of the 2004 ACM workshop on visualization and data mining for computer security. ACM, New York, p 65-72

Lebiere C (1999) The dynamics of cognition: An ACT-R model of cognitive arithmetic. Kognitionswissenschaft 8(1):5-19

Lebiere C, Pirolli P, Thomson R, Paik J, Rutledge-Taylor M, Staszewski J, Anderson JR (2013) A functional model of sensemaking in a neurocognitive architecture. Comp Intell Neurosci 2013: 921695.

Lebiere C, Gray R, Salvucci D, West R (2003) Choice and learning under uncertainty: A case study in baseball batting. In Alterman R, Kirsch D (eds) Proceedings of the 25th annual conference of the cognitive science society. Lawrence Erlbaum Associates, Boston, p 704-709

Lejarraga T, Dutt V, Gonzalez C (2012) Instance-based learning: A general model of repeated binary choice. J Behav Decis Making 25(2):143-153

Lenat DB, Prakash M, Shepherd M (1985). CYC: Using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks. Artif Intell 6(4):65-85

Lye K-W, Wing JM (2005). Game strategies in network security. Int J Inf Secur 4(1-2):71-86

Manshaei MH, Zhu Q, Alpcan T, Bacsar T, Hubaux JP (2013) Game theory meets network security and privacy. ACM Comput Surv 45(3):25

Martin JM, Gonzalez C, Juvina I, Lebiere C (2013) A description-experience gap in social interactions: Information about interdependence and its effects on cooperation. J Behav Decis Making (in press)

McCarthy J (1980) Circumscription – A form of non-monotonic reasoning. Artif Intell 13(1-2):27–39

The MITRE Corporation (2010) Science of cyber-security. The MITRE Corporation, McLean, VA, Technical Report.

Mundie D (2013) How ontologies can help build a science of cyber security. Available online. http://www.cert.org/blogs/insider_threat/2013/03/how_ontologies_can_help_build_a_science_of_cybersecurity.html

Newell A (1990) Unified theories of cognition. Harvard University Press, Cambridge

Obrst L, Chase P, Markeloff R (2012) Developing an ontology of the cyber security domain. In: Costa PCG, Laskey KB (eds) Proceedings of the seventh international conference on semantic technologies for intelligence, defense, and security, George Mason University, Fairfax, 23-26 October 2012

Oltramari A, Lebiere C, Ben-Asher N, Juvina I, Gonzalez C (2013) Modeling strategic dynamics under alternative information conditions. In: West RL, Stewart TC (eds) Proceedings of the 12th international conference on cognitive modeling. ICCM, p 390-395

Pita J, Jain M, Marecki J, Ordóñez F, Portway C, Tambe M, Western C, Paruchuri P, Kraus S (2008) Deployed ARMOR protection: The application of a game theoretic model for security at the Los Angeles International Airport. In: Proceedings of the 7th international joint conference on autonomous agents and multiagent systems: industrial track, p 125-132

Pita J, John R, Maheswaran R, Tambe M, Yang R, Kraus S (2012) A robust approach to addressing human adversaries in security games. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems. International Foundation for Autonomous Agents and Multiagent Systems, Richland, p 1297-1298

Rowley J (2007) The wisdom hierarchy: representations of the DIKW hierarchy. J Inf Sci 33(2):163-180

Roy S, Ellis C, Shiva S, Dasgupta D, Shandilya V, Wu Q (2010) A survey of game theory as applied to network security. In: Sprague RH Jr. (ed) Proceedings of the 43rd Hawaii international conference on system sciences. IEEE: Los Alamitos

Saner LD, Bolstad CA, Gonzalez C, Cuevas HM (2009) Measuring and predicting shared situation awareness in teams. J Cog Eng Decis Making 3(3):280-308

Singer PW, Friedman A (2014) *Cybersecurity and cyberwar: What everyone needs to know*. Oxford University Press, New York

Sowa JF (1984) *Conceptual structures: Information processing in mind and machine*. Addison Wesley, Reading

Staab S, Studer R (2003) *Handbook on ontologies*. Springer-Verlag, Berlin

Taatgen N, Lebiere C, Anderson JR (2006) Modeling paradigms in ACT-R. In: Sun R (ed) Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge University Press, New York, p 29-52

Tufte ER, Graves-Morris PR (1983) The visual display of quantitative information, vol. 2. Graphics Press, Cheshire

Undercoffer J, Joshi A, Pinkston J (2003) Modeling computer attacks: An ontology for intrusion detection. In: Vigna G, Kruegel C (eds) Lecture notes in computer science, vol. 2820. Springer-Verlag, Berlin, p 113-135

West RL, Lebiere C (2001) Simple games as dynamic, coupled systems: Randomness and other emergent properties. J Cog Syst Res 1(4):221-239

West RL, Lebiere C, Bothell DJ (2006) Cognitive architecture, game playing, and human evolution. In: Sun R (ed) Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge University Press, New York, p 103-123