Building an Ontology of Cyber Security

Alessandro Oltramari and Lorrie Faith Cranor
CyLab, Carnegie Mellon University
Pittsburgh, USA

Robert J. Walls and Patrick McDaniel
Department of Computer Science
Pennsylvania State University
University Park, USA

***Abstract*—Situation awareness depends on a reliable perception of the environment and comprehension of its semantic structures. In this respect, cyberspace presents a unique challenge to the situation awareness of users and analysts, since it is a unique combination of human and machine elements, whose complex interactions occur in a global communication network. Accordingly, we outline the underpinnings of an ontology of secure operations in cyberspace, presenting the ontology framework and providing two modeling examples. We make the case for adopting a rigorous semantic model of cyber security to overcome the current limits of the state of the art.**

Keywords— cyber security, ontology, situation awareness, ontology patterns.

I. Introduction

As disclosed by a recent report[1], there has been half a billion cyber security breaches in the first semester of 2014, matching the record set across the entire precedent year. In general, this alarming trend should not surprise when we consider that the bedrock of the Internet is a technological infrastructure built almost 35 years ago for trusted military communications and not for data exchange in the wild (see [1], p.58). The picture gets even worse when considering that the ability to grasp the risk and threats associated with computer networks is averagely poor: recent surveys have actually shown that 65% of the victims of intrusion and information theft in the private sector are notified by third parties and that the detection process usually takes up to 13 months (e.g., see [2], p.10).

Though not exhaustive, such rough statistics at least suggest that if the inadequacy of the technological infrastucture is a key aspect to explain the vulnerabilities of networked computer systems, the *human factor* also plays a central role. As proposed in [3], to improve situation awareness of users and security operators, a shift of focus from system to environment level is highly necessary when modeling cyber scenarios: to this end, a full-fledged science of cyber security needs to be founded, whose core tenet is *cognizing* the cyberspace as a hybrid framework of interaction between humans and computers, where security and privacy policies play a crucial role. As stated by [4], this *cognizance* depends on both a reliable perception of the elements of the environment and, most importantly for our work, on the explicit representation of their semantics. Accordingly, the current article presents the underpinnings of an ontology of secure cyber operations: by and large, the concepts and the relationships that structure this semantic model are peculiar to the domain. That is, notions that are suitable for representing security in the physical world cannot be directly transferred to the cyber environment (e.g., "attack attribution" [5]). We build upon existing ontologies, expanding them to support novel use cases as needed[2]. Our goal is to use the proposed ontology as basis for improving the situation awareness of cyber defenders, allowing them to make optimal operational decisions in every state of the environment.

The rest of the paper is organized as follows: Section II makes the case for the adoption of ontologies in the cyber security realm; Section III outlines the structure of 'CRATELO', a Three Level Ontology for the Cyber Security Research Alliance program funded by ARL[3], and describes two simple cyber scenarios modeled by means of our approach; finally, Section IV draws preliminary conclusions and outlines an agenda for future research.

II. Related Work

Every science is concerned with distinct objects and strives to build rigorous models of the phenomena involving them [6]: accordingly, the objects of a science of cyber security correspond to the attributes of (and the relations between) network of computer devices, security policies, and the tools and techniques of cyber attack and cyber defense [7]. Therefore, inasmuch as ontologies are formal models of a domain, building ontologies of the aforementioned attributes and relations is critical for the transformation of cyber security into a science.

In 2010, the DoD sponsored a study to examine the theory and practice of cyber security, and evaluate whether there are underlying fundamental principles that would make it possible to adopt a more scientific approach. The study team concluded that the most important requirement would be "the construction of a common language and a set of basic concepts about which the security community can develop a shared understanding. A common language and agreed-upon experimental protocols will facilitate the testing of hypotheses and validation of concepts" [8]. The need for controlled vocabularies and ontologies to make progress toward a science of cyber security is recognized in [9] and [10] as well. In this domain, ontologies would include the classification of cyber attacks, cyber incidents, and malicious and impacted software

[1] https://www.riskbasedsecurity.com/reports/2014-MidYearDataBreachQuickView.pdf

[2] For instance, exploiting material available in this portal: http://militaryontology.com/cyber-security-ontology.html

[3] http://cra.psu.edu/

programs. From our point of view, where the human component of cyber security is also essential, the analysis needs to be expanded to the different roles that attackers, users, defenders and policies play in the context of cyber security, the different tasks that the members of a team are assigned to by the team leader, and the knowledge, skills and abilities needed to fulfill them.

There has been little work on ontologies for cyber security and cyber warfare. Within a broader paper, there is a brief discussion of an ontology for DDoS attacks [11] and a general ontology for cyber warfare is discussed in [12]. To the best of our knowledge, Obrst and colleagues [13] provide the most comprehensive description of a cyber ontology architecture, whose vision has actually inspired the work presented in this paper (the scale of the project and its difficulties are also discussed by Dipert in [10]). By and large, efforts that have been made toward developing ontologies of cyber security, even when expressed in OWL, RDF or other XML-based formats, typically do not utilize existing military domain or middle-level ontologies such UCORE-SL[4]. With regard to human users and human computer interaction, the most important step in understanding a complex new domain involves producing accessible terminological definitions and classifications of entities and phenomena, as stressed in [9]. Discussions of cyber warfare and cyber security often begin with the difficulties created by misused terminology (such as characterizing cyber espionage as an attack): in this regard, the Joint Chiefs of Staff created a list of cyber term definitions that has been further developed and improved in a classified version[5]. None of these definitions, however, are structured as an ontology. Likewise, various agencies and corporations (NIST[6], MITRE[7], Verizon[8]) have formulated enumerations of types of malware, vulnerabilities, and exploitations. In particular MITRE, which has been very active in this field, maintains two dictionaries, namely CVE (Common Vulnerabilities and Exposures [9]) and CWE (Common Weakness Enumeration[10]), a classification of attack patterns (CAPEC - Common Attack Pattern Enumeration and Classification[11]), and an XML-structured language to represent cyber threat information (STIX - Structure Threat Information Expression [12]). Regardless of the essential value of these resources, without a "shared semantics" the sprawling definitions they contain are hard to maintain and port into machine-readable formats.

III. A Three-Level Ontology for the Cyber-Security Research Alliance

Top-level ontologies capture generic characteristics of world entities, such as spatial and temporal dimensions, morphology (e.g., parts, edges, sides), qualities (e.g., color, volume, electric charge), etc.; because of their inherent generality, they are not suited to model contextual aspects. Nevertheless, it's good practice to describe the fine-grained concepts that constitute a *domain-level* ontology in terms of foundational (or *top-level*) categories, adding core (or *middle-level*) notions to fill contingent conceptual gaps. For instance, an ontology of mineralogy should include notions like "basaltic rock", "texture" and "metamorphic reaction". In order to describe the meaning of those specific concepts, high-level categories such that "object", "quality" and "process" must be employed; the ontology should also define an intermediate notion like "metamorphism", which is common across domains (biology, chemistry, computer science, architecture, etc.), to explain how the different phases, end products, and features of metamorphic reactions are bound together.

Our ontology of cyber security makes no exceptions to the tripartite layering described above: in particular, CRATELO is an ontological framework constituted of a domain ontology of cyber operations (OSCO), designed on the basis of DOLCE top ontology extended with a security-related middle-level ontology (SECCO). The three levels of CRATELO (schematized in figure 1) currently include 223 classes and 131 relationships (divided into 116 object properties and 15 datatype properties) and encoded in OWL-DL. The expressivity of the ontology is SRIQ, a decidable extension of the description logic SHIN (see [14] for more details).

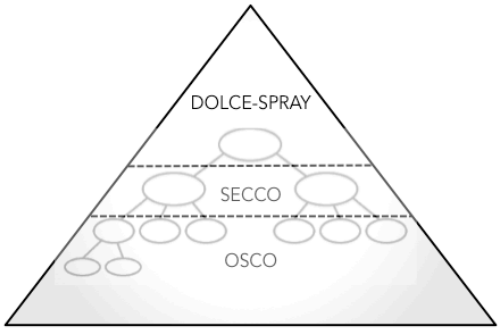

Figure 1: The schematics of CRATELO

A. Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)

DOLCE is part of a library of foundational ontologies for the Semantic Web developed under the WonderWeb EU project[13]. As reflected in the acronym, DOLCE holds a cognitive bias, i.e., aiming at capturing the conceptual primitives underlying natural language and commonsense reasoning [15]. In order to reduce the complexity of the axiomatisation, in the current work we adopt DOLCE-SPRAY[14], a simplified version of DOLCE [16].

The root of the hierarchy of DOLCE-SPRAY is ENTITY, which is defined as the class of anything that is identifiable as an object of experience or thought. The first relevant distinction is among CONCRETE ENTITY, i.e., whose instances

[4] http://www.slideshare.net/BarrySmith3/universal-core-semantic-layer-ucoresl
[5] http://publicintelligence.net/dod-joint-cyber-terms/
[6] http://www.nist.gov/
[7] http://www.mitre.org/
[8] http://www.verizon.com/
[9] https://cve.mitre.org/
[10] http://cwe.mitre.org/
[11] https://capec.mitre.org/
[12] https://stix.mitre.org/language/version1.1.1/
[13] http://wonderweb.man.ac.uk/
[14] Categories are indicated in small caps; relationships in italics. Mutiple individuals instantiating the same category are denoted by adding an 's' to the category name (e.g., REQUIREMENTs). Presenting the axiomatisation of DOLCE-SPRAY is out of scope in this paper.

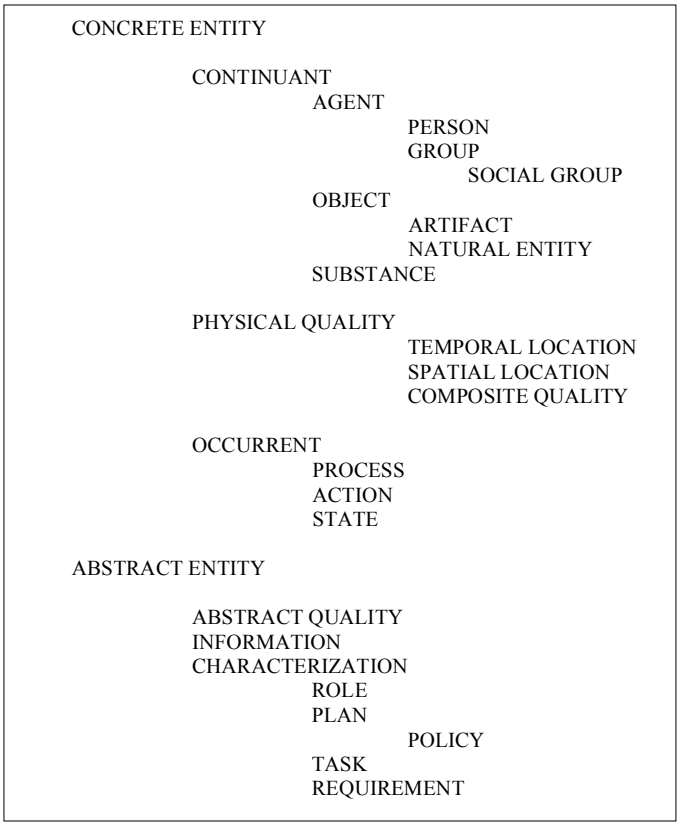

Figure 2: DOLCE-SPRAY backbone taxonomy

are located in definite spatiotemporal regions, and ABSTRACT ENTITY, whose instances don't have inherent spatiotemporal dimensions. CONCRETE ENTITY is further divided into CONTINUANT, OCCURRENT, and QUALITY, respectively entities with inherent spatial parts (e.g., artifacts, animals, substances), entities with inherent temporal parts (e.g., events, actions, states) and entities whose existence depends on their host (for instance 'the color of a flower', 'the duration of a football game', 'the area of a construction site', etc.). DOLCE's basic ontological distinctions are maintained in DOLCE-SPRAY: the substantial differences come from a) merging ABSTRACT and NON–PHYSICAL–ENDURANT categories into DOLCE-SPRAY's ABSTRACT ENTITY and b) by breaking the class QUALITY into PHYSICAL QUALITY and ABSTRACT QUALITY, moving the latter under the branch ABSTRACT ENTITY. Accordingly, the class ABSTRACT QUALITY designates the qualities that don't have any defining spatiotemporal dimension, such as the price of goods, the usefulness of a service, etc. A sibling of ABSTRACT QUALITY under the ABSTRACT ENTITY branch, INFORMATION refers to any content that can be conveyed by some physical OBJECT, from the metal boards used for road signs to the memory location of a Python script. CHARACTERIZATION is defined as a mapping of *n*-uples of individuals to truth-values. Individuals belonging to CHARACTERIZATION can be regarded to as 'reified concepts' (e.g., 'manufactured object'), and the irreflexive, antisymmetric relation *characterizes* associates them with the objects they denote ('a collection of vintage shoes'). Among the relevant sub-types of CHARACTERIZATION we can find: ROLE, i.e., the classification of an entity according to a given context or perspective (e.g., 'instructor'); PLAN, namely the generic description of an action (such as 'the disassembly of a 9mm'); TASK, that is a representation of the specific steps that are needed to execute an ACTION according to a PLAN (e.g., 'removing the magazine', 'pull back the slide'); REQUIREMENT, whose instances can be seen as the conditions that need to be satisfied as part of a PLAN (e.g., 'the weapon must be clear before proceeding'). A specific sub-class of PLAN is POLICY, whose instances need to satisfy specific REQUIREMENTs adopted or proposed by some SOCIAL GROUP (e.g., a government, a party, a no profit association, a private company, etc.). In general, the branch of DOLCE-SPRAY rooted on CHARACTERIZATION distills the extensions introduced in [17]. An overview of DOLCE-SPRAY backbone taxonomy is represented in Figure 2.

B. Security Core Ontology (SECCO)

This section outlines a set of security concepts based on DOLCE-SPRAY primitives.

An entity is a THREAT ϕ for an ASSET α *valued by* a STAKEHOLDER σ and *protected by* a DEFENDER δ, if and only if ϕ is *used by* an ATTACKER κ to *exploit* a VULNERABILITY ϖ of α in an OFFENSIVE_OPERATION τo. To prevent τo, a specific collection of SECURITY_REQUIREMENTs υs need to *be satisfied by* a SECURITY_POLICY π, enforced to protect α. But if τo strikes, δ has to promptly defend α, performing a suitable DEFENSIVE_OPERATION δo to *deploy* a COUNTERMEASURE χ for neutralizing PAYLOAD ψ conveyed by τo[15]. The class OPERATION can be represented as the union of τo and δo: any OPERATION o is carried out on the basis of a MISSION-PLAN λ whose sequence of MISSION_TASKs ξs *are executed in* o[16]. Note that in order to delineate λ in a DEFENSIVE_OPERATION δo, δ would also need to run a RISK-ASSESSMENT μ of the RISK ρ *associated to* ξs (datatype properties can be used to represent ρ as a parameterization of the expected losses, probabilities of attack, etc.)[17]. The formalization below (1-30) represents a basic alignment between SECCO and DOLCE-SPRAY. The relations *isPartOf*, *participates* (and its inverse *hasParticipant*), *isQualityOf*, *characterizes*, *definedIn*, *satisfies hasRole*, *hasRequirement*, are imported from DOLCE-SPRAY. We used self-explanatory abbreviations (e.g., OFF_OP instead of OFFENSIVE_OPERATION) to keep the list compact, when possible. For reasons of space, presenting a comprehensive set of axioms for SECCO is out of scope in this paper.

$$\text{ATTACKER}[18] \sqsubseteq \text{ROLE} \sqcap \forall\, characterizes.\text{AGENT} \quad (1)$$

$$\text{DEFENDER} \sqsubseteq \text{ROLE} \sqcap \forall\, characterizes.\text{AGENT} \quad (2)$$

[15] Both countermeasures and payloads are artifacts of some sort, e.g., an antidote and a poison.

[16] o can be a single ACTION or a complex collection of interconnected actions.

[17] Although risk assessment needs to be done preemptively, continuous monitoring is also required for up-to-date situational awareness.

[18] In our model, instances of ATTACKER, DEFENDER and STAKEHOLDER are not equal to instances of PERSON, GROUP and, in general, AGENT. In this perspective, 'Alessandro' (instance of PERSON) *qua* DEFENDER would correspond to team member 'Alpha1' (instance of DEFENDER). *Qua*-entities have been formally analyzed in [33]. Also, since in different situations a defender may play the role of an attacker (and vice versa), we don't consider the two classes as disjoint.

$$\text{STAKEHOLDER} \sqsubseteq \text{ROLE} \sqcap \forall\, characterizes.\text{AGENT} \quad (3)$$
$$\text{STAKEHOLDER} \sqsubseteq \neg\ (\text{ATTACKER} \sqcup \text{DEFENDER})^{19} \quad (4)$$
$$\text{ASSET} \sqsubseteq \text{ROLE} \sqcap \forall\, characterizes\ (\text{OBJECT} \sqcup \text{INFORMATION}) \quad (5)$$
$$\text{ASSET} \sqsubseteq \neg\ \text{THREAT} \quad (6)$$
$$\text{THREAT} \sqsubseteq \text{ROLE} \sqcap \forall\, characterizes\ (\text{OBJECT} \sqcup \text{INFORMATION}) \quad (7)$$
$$\text{THREAT} \sqsubseteq \neg\ \text{ASSET} \quad (8)$$
$$\text{SEC\_REQ} \sqsubseteq \text{DEF\_REQ} \sqsubseteq \text{REQUIREMENT} \quad (9)$$
$$\text{SECURITY\_POLICY} \sqsubseteq \text{POLICY} \sqcap \forall\, satisfies.\text{SEC\_REQ} \quad (10)$$
$$\text{OFF\_REQ} \sqsubseteq \text{REQUIREMENT} \quad (11)$$
$$\text{OFF\_REQ} \sqsubseteq \neg\ \text{DEF\_REQ} \quad (12)$$
$$\text{DEF\_REQ} \sqsubseteq \neg\ \text{OFF\_REQ} \quad (13)$$
$$\text{OPERATION} \sqsubseteq \text{ACTION} \quad (14)$$
$$\text{DEF\_OP} \sqsubseteq \text{OPERATION} \quad (15)$$
$$\text{OFF\_OP} \sqsubseteq \text{OPERATION} \quad (16)$$
$$\text{OFF\_OP} \sqsubseteq \neg\ \text{DEF\_OP} \quad (17)$$
$$\text{DEF\_OP} \sqsubseteq \neg\ \text{OFF\_OP} \quad (18)$$
$$\text{MISSION\_PLAN} \sqsubseteq \text{PLAN} \quad (19)$$
$$\text{MISSION\_TASK} \sqsubseteq \text{TASK} \sqcap \forall\, isDefinedin.\text{MISSION\_PLAN} \quad (20)$$
$$\text{RISK} \sqsubseteq \text{ABST\_QUALITY} \sqcap \forall\, isQualityOf.\text{MISSION\_TASK} \quad (21)$$
$$\text{RISK\_ASSESSMENT} \sqsubseteq \text{ACTION} \sqcap \exists\, hasParticipant\,.\text{RISK} \quad (22)$$
$$\text{COUNTERMEASURE} \sqsubseteq \text{ARTIFACT} \sqcap \forall participates.\text{DEF\_OP} \quad (23)$$
$$\text{PAYLOAD} \sqsubseteq \text{ARTIFACT} \sqcap \forall\, participates.\text{OFF\_OP} \quad (24)$$
$$\text{VULNERABILITY} \sqsubseteq \text{ABST\_QUALITY} \sqcap \forall\, isQualityOf.\text{ASSET} \quad (25)$$
$$\text{DEF\_OP} \equiv \exists\, hasParticipant.\text{DEFENDER} \sqcap \exists executes.\text{MISSION\_PLAN} \sqcap \exists hasParticipant.\text{COUNTERMEASURE} \sqcap \exists\ hasRequirement.\text{DEF\_REQ} \quad (26)$$
$$\text{OFF\_OP} \equiv \exists\, hasParticipant.\text{ATTACKER} \sqcap \exists executes.\text{MISSION\_PLAN} \sqcap \exists hasParticipant.\text{PAYLOAD} \sqcap \exists\ hasRequirement.\text{OFF\_REQ} \quad (27)$$
$$\text{ATTACKER} \equiv \forall\, exploits.\text{VULNERABILITY} \sqcap \exists uses.\text{THREAT} \quad (28)$$
$$\text{DEFENDER} \equiv \forall\, protects.\text{ASSET} \sqcap \exists uses.\text{COUNTERMEASURE} \quad (29)$$
$$\text{STAKEHOLDER} \equiv \forall\, values.\text{ASSET} \sqcap \exists enforces.\text{SECURITY\_POLICY} \quad (30)$$

SECCO's categories are positioned at a too coarse-level of granularity to capture the details of domain-specific scenarios: properties like THREAT, VULNERABILITY, ATTACK, COUNTERMEASURE, ASSET are orthogonal to different domains and, in virtue of this, they can be predicated of a broad spectrum of things: for instance, infections are a threat to the human body, Stuxnet is a threat to PLCs, the impact of large asteroids on the Earth's surface is a threat to the survival of organic life forms, dictatorship is a threat to civil liberties, and so on and so forth. Though there seems to be a consensus in the literature on the core ontological concepts of security (see [18] and [19]), the minimal set presented here has been occasionally expanded along alternate directions. For instance, Fenz and Ekelhart [20] introduce the concept of 'control', by means of which stakeholders implement suitable countermeasures to mitigate known vulnerabilities of assets[20]. A 'policy', in this context, is defined as a regulatory or organizational form of control (SECCO definition of POLICY is more functionality-centered). Fenz and Ekelhart [20] also outline a taxonomy of assets, distinguishing 'tangible' (e.g., 'wallet') from 'intangible' ones (e.g., 'credit card credentials'), where the former can be furthermore split into 'movable' (e.g., 'car', 'jewelry') and 'unmovable' (e.g., 'house', 'land'). Interestingly enough, Fenz and Ekelhart reify the procedure of assessing a risk into the concept of 'rating', whose attributes can be expressed qualitatively (e.g., in Likert scale – high, medium and low) or quantitatively (measuring the probability of a risk). Avižienis and colleagues present a comprehensive analysis of security where the notion of 'fault' is introduced to denote an interruption of the services delivered by a given system in the environment [21]. A middle-level ontology of security can be possibly extended beyond SECCO: in this respect, the key contribution of this module doesn't rely on the coverage (or 'concept density' – see [22], p. 187) of security primitives but on the formalization driven by a top-level ontology. Our approach has some similarities with the effort described in [23], though Massacci and colleagues were principally concerned with the ontological analysis of a specific software development methodology, Secure Tropos.

C. Ontologies of Secure Cyber Operations (OSCO)

One of the major cyber security problems for government and corporations is the widespread "operational chaos" experienced by analysts, as Michael Susong has recently called the phenomenon of "having too many alarms (false positives) in a network, not enough trained people to deal with them, and a consequent poor prioritization of risks and countermeasures" [21]. In this regard, the objective of an ontology of cyber security is to shape that chaos into a framework of meaningful and reusable chunks of knowledge, turning the operational disarray into a systematic model by means of which cyber analysts can improve their situation awareness. As mentioned in section 1, the key to this augmented *cognizance* relies on a consistent assessment of the context and on a comprehensive understanding of its elements at the semantic level. But how is a cyber operation usually defined? In a document released in 2010, the Joint Chiefs of Staff describes a "cyberspace operation" as the "employment of cyber capabilities where the primary purpose is to achieve objectives in or through cyberspace. Such operations include computer network operations and activities to operate and defend the Global Information Grid" [24]. Drawing on this broad definition and relying on DOLCE-SPRAY and SECCO, in OSCO we represent a CYBER_OPERATION ψ as an OPERATION *executed by* a CYBER_OPERATOR φ, who can play either the role of DEFENDER in a DEFENSIVE_CYBER-OPERATION or the role of ATTACKER in an OFFENSIVE_CYBER-OPERATION. In the context of cyber security we can also distinguish between those OFFENSIVE_CYBER_OPERATIONs whose MISSION-PLANs satisfy the OFFENSIVE_REQUIREMENT of remaining undetected, and those that don't: we use the class CYBER_EXPLOITATION to the denote the former, and CYBER-ATTACK for the latter. As Lin points out in [5], from a technical viewpoint cyber-attacks and cyber exploitations are very similar: they use the same access paths and focus on the same vulnerabilities. The difference is on the delivery and

[19] Note that δ and σ may or may not coincide: in the second case, the latter needs to delegate the former to act in her behalf. The notion of delegation (and trust) in agent ontologies has been extensively studied by [26], but it's currently not included in CRATELO, as (6) shows.

[20] In cyber security, exploitations of unknown vulnerabilities correspond to the so-called Zero-Day Attacks.

[21] Dr. Micheal Susong is an Intelligence Subject Matter Expert affliated to iSIGHT Partners; he gave an invited talk at Carnegie Mellon University on September 8th, 2014.

execution of the PAYLOAD that must be performed undetectably in CYBER_EXPLOITATIONS (e.g., port scanning or SQL injections). The list of class-inclusions below (33-51) denotes the alignment between OSCO and SECCO categories and some specializations of OSCO domain concepts. For reasons of space we could not include a formal characterization of specific cyber threats and cyber vulnerabilities (comprehensive classifications can be consistently found in military reports, doctrines and academic articles - see [25] [26] [27]).

$$\text{CYBER\_OPERATION} \sqsubseteq \text{OPERATION} \quad (31)$$
$$\text{OFF\_CYBER\_OP} \sqsubseteq \text{CYBER\_OPERATION} \quad (32)$$
$$\text{DEF\_CYBER\_OP} \sqsubseteq \text{CYBER\_OPERATION} \quad (33)$$
$$\text{OFF\_CYBER\_OP} \sqsubseteq \text{OFF\_OP} \quad (34)$$
$$\text{OFF\_CYBER\_REQ} \sqsubseteq \text{OFF\_REQ} \quad (35)$$
$$\text{DEF\_CYBER\_REQ.} \sqsubseteq \text{DEF\_REQ} \quad (36)$$
$$\text{UNDETECTABILITY} \sqsubseteq \text{OFF\_CYBER\_REQ} \quad (37)$$
$$\text{CYBER\_COUNTERMEASURE} \sqsubseteq \text{COUNTERMEASURE} \quad (38)$$
$$\text{CYBER\_ASSET} \sqsubseteq \text{ASSET} \quad (39)$$
$$\text{CYBER\_THREAT} \sqsubseteq \text{THREAT} \quad (40)$$
$$\text{CYBER\_SEC\_REQUIREMENT} \sqsubseteq \text{SEC\_REQUIREMENT} \quad (41)$$
$$\text{CYBER\_SECURITY\_POLICY} \sqsubseteq \text{SECURITY\_POLICY} \quad (42)$$
$$\text{CYBER\_VULNERABILITY} \sqsubseteq \text{VULNERABILITY} \quad (43)$$
$$\text{CYBER\_ATTACKER} \sqsubseteq \text{ATTACKER} \sqcap \forall\, exploits.\text{CYBER\_VULNERABILITY} \sqcap \exists uses.\, \text{CYBER\_THREAT} \quad (44)$$
$$\text{CYBER\_ANALYST} \sqsubseteq \text{DEFENDER} \sqcap \forall\, protects.\text{CYBER\_ASSET} \sqcap \exists uses.\, \text{CYBER\_COUNTERMEASURE} \quad (45)$$
$$\text{CYBER\_STAKEHOLDER} \sqsubseteq \text{STAKEHOLDER} \sqcap \forall\, values.\text{CYBER\_ASSET} \sqcap \exists enforces.\, \text{CYBER\_SECURITY\_POLICY} \quad (46)$$
$$\text{CYBER\_ATTACK} \sqsubseteq \text{OFF\_CYBER\_OP} \sqcap \exists hasParticipant.\, \text{CYBER\_ATTACKER} \sqcap \neg\exists hasRequirement.\, \text{UNDETECTABILITY} \quad (47)$$
$$\text{CYBER\_EXPLOITATION} \sqsubseteq \text{OFF\_CYBER\_OP} \sqcap \exists hasParticipant.\, \text{CYBER\_ATTACKER} \sqcap \exists hasRequirement.\, \text{UNDETECTABILITY} \quad (48)$$
$$\text{DEF\_CYBER\_OP} \sqsubseteq \text{DEF\_OP} \sqcap \exists hasParticipant.\, \text{CYBER\_ANALYST} \sqcap \exists hasRequirement.\, \text{DEF\_CYBER\_REQ} \quad (49)$$

Since the development of a full-scale domain ontology is currently underway within our project, for the sake of this article we will limit ourselves to model only two sample scenarios.

1) Example 1:RETRIEVE_FILE_SECURELY

Figure 3 represents CRATELO's classes and relationships used to model the *Retrieve File Securely* scenario. For issues of visualization, the diagram covers only the most salient notions involved in this cyber operation. In order to retrieve a file without exposing a computer system – and possibly an entire network – to cyber threats, some specific security requirements need to be fulfilled while carrying out that operation. In particular, as it is also the case for other kinds of CYBER-OPERATION, RETRIEVE-FILE-SECURELY must occur over a secure channel of a network, from authenticated computer(s) and through authorized server(s). By and large, abiding to these security requirements while executing the mission-tasks should lead to mission accomplishment. The composite RETRIEVE-FILE-SECURELY-TASK can be further divided into simpler temporally-structured and logically-connected subtasks. Accordingly, a request for a file can be sent to an authenticated server only after locating the desired file in the network; the inspection of the file can trivially occur only once the file has been obtained; and so on and so forth. In CRATELO we can express these basic temporal constraints by means of the foundational layer: in fact, DOLCE includes an adaptation of Allen's axioms [28], which are considered as a powerful logical theory for temporal representation and reasoning (the formalization of these axioms has also been maintained in DOLCE-SPRAY). Moreover, if malware is detected, the file must be removed from the host: the deployment of this preventive countermeasure aims at avoiding a disruption of the isolated computer node and a cyber attack to the network it belongs to. This countermeasure can be expressed as a conditional rule formalized in CRATELO by using an additional modeling apparatus, i.e., the Semantic Web Rule Language (SWRL)[22], which extends OWL-DL axioms. By including rule-based mechanisms in CRATELO we also comply with the core requisites described in [13] of a full-fledged cyber ontology architecture.

As the example exposes, one of the key design principles underlying CRATELO is to separate the temporal dynamics of cyber operations from the abstract generalizations used to describe them, i.e., plans, tasks, requirements. This approach consents to model a cyber operation as an ontology pattern grounded on the top level dyad ACTION-CHARACTERIZATION, unfolded by the middle-level tetrad OPERATION-MISSION_PLAN-MISSION_TASK-SEC_REQUIREMENT, and specified by CYBER_OPERATION-CYBER_MISSION_PLAN-CYBER_MISSION_TASK-CYBER_SECURITY_REQUIREMENT. In recent years, 'ontology patterns' have become an important instrument for conceptual modeling [29]: the rationale, as our work suggests, is to identify some minimal knowledge structures within an ontology to be used for modeling a problem (in this regard, the ontology remains the reference framework whereby the pattern can be expanded). This methodology is also ideal from a reasoning standpoint. For instance, in [30] the authors state that "mission activities are tasks focused on answering mission questions" (where the latter can be seen as partially overlapping the notion of security requirement): but an ontology that fails to discriminate 'activities' from 'tasks' would likely be affected in its inference capabilities, in the degree that reasoning over tasks that have not been executed yet – i.e., that are not activities – would not be supported. It's not difficult to imagine the circumstances where this limit can become a serious drawback for a cyber analyst: mental simulation is commonly adopted by humans to foresee the outcomes of an action before performing it [31], and a semantic framework where mission activities and tasks are conceptually viewed as the same entity precludes that, and might eventually result into pervasive logical inconsistencies (if the ambiguity is not somehow reduced). On the contrary, an ontology-pattern based on CRATELO allows to specify cyber operations at a sufficient level of conceptual granularity.

[22] http://www.w3.org/Submission/SWRL/

2) Example 2: INTRUSION_DETECTION

In a simplified scenario where an SQL injection attack is launched, a defensive cyber operation of INTRUSION_DETECTION can be divided into three essential sub-actions (and corresponding tasks): 1) block the IP address of the attacker; 2) to escalate the level of response; 3) to block all external connections and 4) redirect the incoming traffic to a honeypot for further inspection. Who can perform these actions? In the real world, cyber analysts with different responsibilities and privileges usually form a response team: for instance, we can indicate with L1, L2 and L3 the incremental levels of expertise of cyber analysts. Accordingly, 1) would only be performed by L1 analysts; 2) can only be performed by L1 analysts toward L2 analysts or by L2 toward L3; 3) can only be executed by L2 analysts and 4) only by L3. As a matter of fact, gauging which action fits better the situation is not a one-shot decision, but rather a multi-stage evaluation process where the situational awareness of cyber analysts frequently changes Also, each of those sub-actions has incremental costs and inversely proportional risks: for instance, if blocking all the connections to a web server eliminates the risks of a reiterated attack, suspending the network traffic has a severe impact on the system functionality (e.g., no data access for authorized third parties): escalation, in this context, is an effective means to prevent risk mismanagement. Although this simplified scenario gives only a partial account of the actions that actual analysts have at their disposal, using an ontology of cyber security like CRATELO to model intrusion detection can clearly represent a mean to improve situational awareness and fill the *semantic gap* [32] in our understanding of the cognitive demands in the cyber world. Figure 4 presents a partial view of CRATELO categories and relations used for intrusion detection.

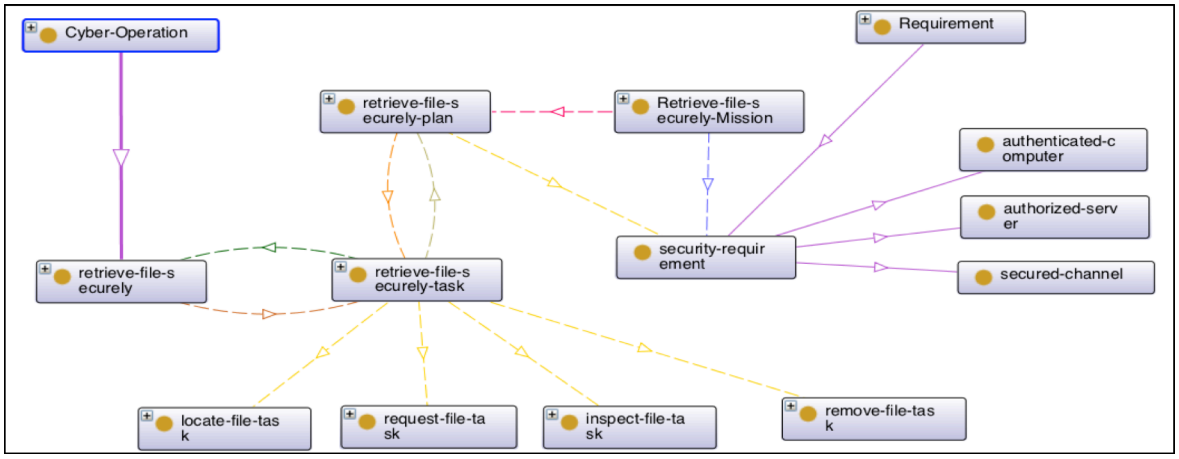

Figure 3 – A visualization of the RETRIEVE-FILE-SECURELY cyber operation modeled in CRATELO. Legend of the arc types: 'has subclass' (purple); 'is executed in' (green); 'executes' (brown); 'has part' (yellow); 'defines task' (orange); 'is defined in task' (ochre); 'satisfies (all)' (fuchsia); 'satisfies (some)' (electric blue).

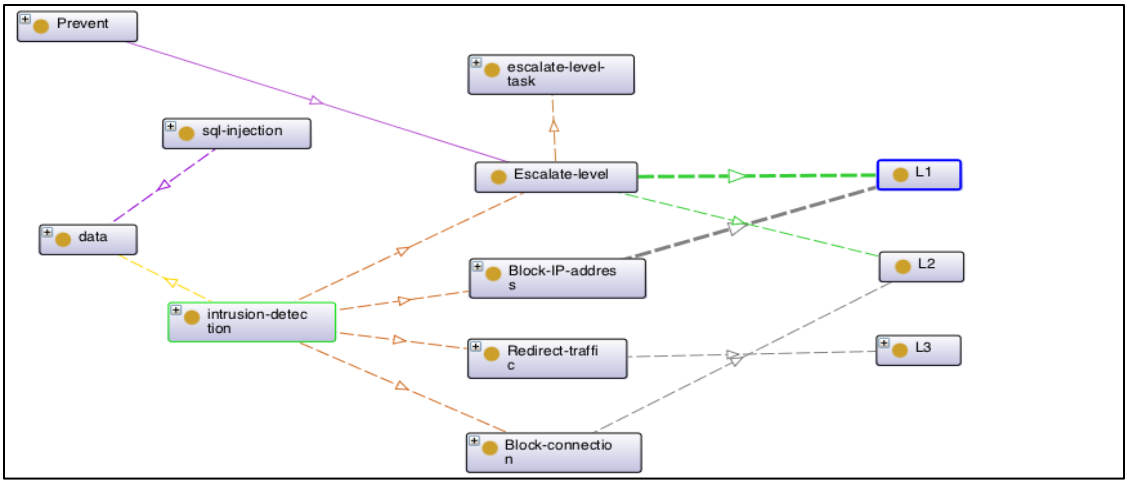

Figure 4 – A subset of actions that can be performed in a cyber operation of INTRUSION_DETECTION. This diagram shows some of the interdependencies between classes of actions and levels of expertise of cyber analysts. Legend of the arc types: 'has subclass' (solid purple); 'targets' (dotted purple); 'defend' (yellow); 'has part' (brown); 'executes task' (light brown); 'involves (only) agent' (gray); 'involves (only disjunction)' (green).[23]

[23] Figure 3-4 were generated and exported using Ontograf (http://protegewiki.stanford.edu/wiki/OntoGraf), a visualization plug-in for Protégé. Even within the same ontology, Ontograf automatically assigns different colors to arcs when a new figure is created: this explains mismatch of colors between the two figures.

IV. Conclusions And Future Work

Notwithstanding the proliferation of taxonomies, dictionaries, glossaries, and terminologies of the cyber landscape, building a comprehensive model of this domain remains a major objective for the community of reference, that includes government agencies, private organizations, researchers and intelligence professionals. There are multiple reasons behind the discrepancy between demand and supply of semantic models of cyber security. Although we cannot thoroughly address this topic here, we are firmly convinced that a great part of the problem is the lack of balance between the 'vertical' and the 'horizontal' directions of the effort. From one side, state of the art consists of several classifications of the domain, as argued in Section II: these efforts typically yield rich catalogs of cyber attacks, exploits and vulnerabilities. On the other side, a rigorous conceptual analysis of the entities and relationships that are encompassed by different cyber scenarios would also be needed, but little work has been done on this horizontal dimension (if we exclude the ongoing MITRE initiative described by Leo Obrst and colleagues in [13]). In this paper we placed ourselves on the second perspective: instead of presenting "yet another" catalog of cyber notions, an endeavor that remains however of undisputable relevance, we decided to explore in depth the semantic space of operations. Our investigation addresses cyber operations as complex entities where the human factor is as important as the technological spectrum: our ontological analysis is grounded on a bedrock of foundational concepts and reaches the domain of cyber operations through an intermediate layer where core notions are defined.

Future work will focus on the following research steps:

- extending SECCO with an ontology of risk;
- populating OSCO with a large set of cyber operations documented in the literature and learned from real-world case studies;
- designing and customizing a methodology for ontology validation based on "competency questions" submitted to domain experts (along to what has been proposed in [20]);
- running cyber warfare simulations within military exercises, collecting data to be modeled with CRATELO;
- studying ontology mappings beteween CRATELO and other semantic models (e.g., MITRE's Cyber Ontology Architecture), ensuring interoperability and reusability of the resource.

We are aware of the challenges ahead of us in pursuing this research agenda, which would usually be very difficult to implement. Nevertheless, we're also persuaded that, in the broad vision framed by the ARL Cyber Security Collaborative Research Alliance, what we have described illustrates a realistic work plan and a necessary step toward the foundation of a science of cyber security.

Acknowledgments

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-13-2-0045 (ARL Cyber Security CRA). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] Yannakogeorgos, P. and Lowther, A. B. "The Prospects of Cyber Deterrence: American Sponsorships of Global Norms," in *Conflict and Cooperation in Cyberspace.*: Taylor&Francis, 2013, pp. 49-77.

[2] L. Mattice, "Taming the "21st Century's Wild West" of Cyberspace?," in *Conflict and Cooperation in Cyberspace.*: Taylor&Francis, 2013, pp. 9-12.

[3] McDaniel, P., Rivera, B., Swami, A. "Toward a Science of Secure Environments," *Security and Privacy*, vol. 12, no. 4, pp. 68-70, July/August 2014.

[4] Endsley, M.R. "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors*, vol. 37, no. 1, pp. 32-64, 1995.

[5] Lin, H. "Escalation Dynamics and Conflict Termination in Cyberspace," *Strategic Studies Quarterly*, vol. 6, no. 3, pp. 46-70, Fall 2012.

[6] Bunge, M. *Causality and Modern Science*. New York: Dover Publications, 1979.

[7] Kott, A."Towards Fundamental Science of Cyber Security," in *Network Science and Cybersecurity*, R. E. Pino, Ed. New York, 2014, vol. 55.

[8] The MITRE Corporation, "Science of Cyber-Security," The MITRE Corporation, McLean, VA, Technical 2010.

[9] Mundie, D. A. and McIntire, D. M. "The MAL: A Malware Analysis Lexicon," CERT® Program - Carnegie Mellon University , Technical 2013.

[10] Dipert, R. "The Essential Features of an Ontology for Cyberwarfare," in *Conflict and Cooperation in Cyberspace - The Challenge to National Security*, Panayotis A Yannakogeorgos and A. B. Lowther, Eds.: Taylor & Francis, 2013, pp. 35-48.

[11] Kotenko, I. "Agent-Based modeling and simulation of cyber-warfare between malefactors and security agents in internet ," in *19th European Conference on Modeling and Simulation*, 2005.

[12] 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. [Online]. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA517410

[13] Obrst, L., Chase, P., & Markeloff, R. "Developing an ontology of the cyber security domain," in *Seventh International Conference on Semantic Technologies for*

Intelligence, Defense, and Security, 2012, pp. 49-56).

[14] Horrocks, I., Kutz, O., Sattler, U. "The Irresistible SRIQ ," in *OWLED '05 - "OWL: Experiences and Directions"*, vol. 188, Galway, 2005.

[15] Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, Schneider, L. A. "The WonderWeb Library of Foundational Ontologies and the DOLCE ontology," Laboratory For Applied Ontology, ISTC-CNR, Technical Report 2002.

[16] Vetere G., Jezek E., Chiari I., Zanzotto F.M., Nissim M., Gangemi A. Oltramari A., "Senso Comune: A Collaborative Knowledge Resource for Italian," in *The People's Web Meets NLP: Collaboratively Constructed Language Resources.*: Springer Verlag, 2013, pp. 45-67.

[17] Gangemi, A., Mika, P. "Understanding the Semantic Web through Descriptions and Situations," in *On The Move to Meaningful Internet Systems - Lecture Notes in Computer Science*. Berlin-Heidelberg: Springer, 2003, vol. 2888, pp. 689-706.

[18] Salinesi, C., Wattiau, I., A. Souag, "Ontologies for Security Requirements: A Literature Survey and Classification," in *Advanced Information Systems Engineering Workshops*, vol. 112, 2012, pp. 61-69.

[19] Schumacher, M. "Toward a Security Core Ontology," in *Security Engineering with Patterns*. Berling-Heidelberg: Springer-Verlag, 2003, pp. 87-96.

[20] Fenz, S., Ekelhart, A. "Formalizing Information Security Knowledge," in *the International Symposium on Information, Computer, and Communications Security (ASIACCS '09)*, New York, pp. 183-194.

[21] Avižienis, A., Laprie, J., Randell, B., Landwehr, C. "Basic Concepts and Taxonomy of Dependable and Secure Computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11-33, January-March 2004.

[22] Prévot, L., Borgo, S., Oltramari, A. "Interfacing Ontologies and Lexical Resources," in *Ontology and the Lexicon - A Natural Language Perspective*, C.R., Calzolari, N., Gangemi, A., Oltramari, A., Prévot, L. Huang, Ed. New York, USA: Cambridge University Press, 2010, pp. 185-200.

[23] Massacci, F., Mylopoulos, J., Paci, F., Thein, T.T., Yijun, Y. "An Extended Ontology for Security Requirements". In *CAiSE 2011 International Workshops*, vol. 83, London, 2011, pp. 622-636.

[24] Joint Staff Department of Defense. Joint Terminology for Cyber Operations. [Online]. http://afri.au.af.mil/cyber/Docs/panel1/Cyber_Lexicon.pdf

[25] Joint Chiefs of Staff, "The National Military Strategy for Cyberspace Operations," Department of Defense, 2006. [Online]. http://www.dod.mil/pubs/foi/joint_staff/jointStaff_jointOperations/07-F-2105doc1.pdf

[26] Air Force Doctrine Document, "Cyberspace Operations,".

[27] Simmons, C. B., Shiva, S. G., Bedi, H., Dasgupta "AVOIDIT: A Cyber Attack Taxonomy," in *9th Annual Symposium on Information Assurance (ASIA)*, Albany, NY, 2014, pp. 2-12.

[28] Allen, J.F. "An interval based representation of temporal knowledge," in *7th International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 1, Vancouver, 1983, pp. 221–226.

[29] Gangemi, A. and Presutti, V. "Ontology design patterns," in *Handbook on Ontologies*.: Springer , 2009, pp. 221-244.

[30] Morris, T.I., Mayron, L.M., Smith, W.B., Knepper, M.M., Reg, I., Fox, K.L. "A perceptually-relevant model-based cyber threat prediction method for enterprise mission assurance," in *IEEE Multi-disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, Miami Beach, 2011, pp. 60-65.

[31] Taylor, S.E., Pham L.B., Rivkin I.D., Armor D.A. "Harnessing the imagination. Mental simulation, self-regulation, and coping.," *American Psychologist* , vol. 53, no. 4, pp. 429-439, Apr 1998.

[32] Gonzalez, C., Ben-Asher, N., Oltramari, A., Lebiere, C. "Cognitive Models of Cyber Situation Awareness and Decision Making," in *Cyber Defense and Situational Awareness*, A., Wang, C., Erbacher, R. Kott, Ed.: Springer, 2014, vol. 62.

[33] Masolo, C., Guizzardi, G., Vieu, L., Bottazzi, E., Ferrario, R. "Relational Roles and Qua Individuals". In *AAAI Fall Symposium on Roles, an Interdisciplinary Perspective*, Virginia, USA. 2005.