Mereotopological Analysis of Formal Concepts in Security Ontologies

Gonzalo A. Aranda-Corral and Joaquín Borrego-Díaz

Abstract. In this paper an analysis of security ontologies, using an mereotopological interpretation of the relationship amongst their classes, based on the entailment in the ontology, is presented. The analysis is carried out by means of a graphical tool (called *Paella*) that implements such an interpretation and it can suggest the potential debugging of anomalies. The analysis also suggests how to interpret the representational anomalies.

1 Introduction

The envisioned Semantic Web (SW) [2] aims to turn the information of the current web into knowledge for solving the informational chaos inherent with the current WWW, by providing trustworthy processing of the information. Its progressive introduction -mostly by institutions and companies- will represent a fundamental change in the understanding of information on the Internet, and, more importantly, it will change the management of digital information about consumers, governments, users, etc. Nevertheless, from the social and security point of view, certain risks exist with these improvements. One of these is the globalization of digital information, which was not considered at the beginning of the current WWW. Another risk could be the deficient transformation/management of the information. In the case of Semantic Web, it presents the opportunity to evaluate and reflect upon them, from different perspectives: computational, epistemological, logical, trustworthy, etc.

Gonzalo A. Aranda-Corral
Department of Information Technology, Universidad de Huelva,
Crta. Palos de La Frontera s/n. 21819 Palos de La Frontera, Spain
e-mail: Gonzalo.Aranda@dti.uhu.es

Joaquín Borrego-Díaz
Department of Computer Science and Artificial Intelligence,
Universidad de Sevilla, Avda. Reina Mercedes s/n. 41012 Sevilla, Spain
e-mail: jborrego@us.es

Ontologies are formal theories of Knowledge that bridge different resources, solving the interoperability problem at a formal level in SW. Therefore, Knowledge processing needs of a sound understanding of ontologies. However, then formal languages proposed for the representation of ontologies can not be used by non-experts in ontologies, for example, experts in information security. Thus evaluation process is a key stage in Knowledge Engineering (KE) applied to security. In fact, The introduction of SW technologies can produce a ”semantic divide” which can not be avoided by implementing standard technologies. There are some *formal semantic literacy* tasks, which are necessary to produce technologies that make the formal aspects invisible to the user. Other cases where the problem appears are where the SW is applied in Web 2.0 and *Open Data Strategies (ODS)*. The use of SW techniques within the Web 2.0, needs paradigms where ontologies are transparent (i.e. , Freebase[1] or OpenCalais service[2]) as well as tools to represent the ontology in a user friendly way. This would be beneficial to users who attempt to understand the knowledge that companies or institutions hold about them, and how it is interpreted. Likewise, the adoption of Web 2.0 strategies in institutions and companies also need such technologies. A closely related issue which has recently emerged is the adoption of *Open Data Strategies* (ODS) by institutions. ODS causes citizens to be interested in the data stored about them. ODS combined with SW tools can provide a best representation of the data through interoperability. Therefore, we can think about this and we can understand how we are viewed as a consumer/user. An image that emerges from the knowledge that companies possess about us, through personal data and the ontology used. Misinterpretation of information (by the user or the system) is an evident danger.

The aim of this paper is to show how to analyse the robustness of security ontologies using automated and visual reasoning -implemented in a semantic tool- as well as to detect some kind of representational anomalies. From this analysis, we try to enhance the conceptual information processing of security ontologies by means of a Semantic Information Representation tool. The tool lets non expert users -on ontological engineering- both to understand and to debug the relationships among critical security concepts in a logical and trustworthy method.

The structure of the paper is as follows. The next section considers the role of ontological engineering issues in security ontologies. In Sect. 3 we present the formal principles of ontology visualization based on reasoning services that represent the logical basis of the Paella tool, described in Sect. 4. Sect. 5 is devoted to discussing the main results of the analysis of a set of security ontologies. The paper ends with some remarks on future work.

2 Knowledge Representation in Security Ontologies

The ontology-based approach enables the definition of the security concepts and their dependencies in a comprensible way both for humans, and software agents[12].

[1] www.freebase.com

[2] http://www.opencalais.com

In Security Information Technology -and from a KE perspective- the development of trustworthy ontologies is imperative [13]. Beside consistency and complexity, the absence of representational anomalies (see Sect. 2.1) is mandatory, because information management in Security affects citizens' civil rights [9]. Two challenges are present.

On the one hand, data management with logical trust (consistency, model theoretic properties, etc.) is related to the need for extending or revising ontologies. This task is, from a company's point of view, dangerous and expensive: since every change in ontology could affect the overall knowledge about the organization. It is also hard to automate, because some criteria for revision cannot be fully formalized [1]. Moreover, security experts would be responsible for reviewing these tasks and usually they are not ontologists.

On the other hand, the sound understanding of concepts, properties and axioms of an ontology is not only a Ontological Engieering challenge. It includes the problem of understanding the structure of concepts to anticipate potential failures, and that requires the combined work of Knowledge engineers security experts. It is more important to visualize the relationships, the internal structure of concepts for recommending changes. Important features to analyse are consistency, compliance with current Security Standards, and fidelity to the intended model. The latter is about the sound representation of some concepts, this means, it the specification represent de intentions of security experts and there is not axioms or properties clearly incompatible with real concepts.

2.1 Representability of Security Issues

Security Ontologies have usually been built on security information resources. Since these kind of resources have not designed to fit ontological structures, several deficiencies of representation arise. In [6], the autors detect several representational problems when aim to enrich a security ontology with Information security which are easily comparable:

1. No concept for some kind of vulnerabilities
2. Vague connections between threats and controls
3. No relationships between threats
4. Inconsistent granularity of information
5. Redundancy and overlapping of information

Note that problems (2) (4) and (5) have mereological nature (that is, they deal with extensional interpretations of concepts/classes as identities and no elements are considered). In fig 2, a NRL ontology is drawed and some of this threats emerge, such as Redundancy and Overlapping, etc. Upper Ontologies as SUMO, DOLCE or OPEN-CyC use mereology for representing (and reasoning) abstract concepts.

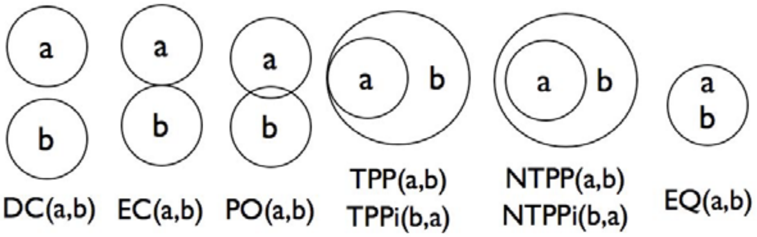


Fig. 1 The relations of RCC8

3 Ontology Visualization Based on Reasoning Services

Thinking in a visual representation, two principles can be considered. First, in order to relate two classes, the picture does not necessarily have to represent elements within them. That is, is a mereological representation in nature. Second, it does not have a prefixed shape, place or size for classes in the representation. That is, is a topological representation in nature. Combining both principles, it has to consider mereotopological representation and reasoning.

The rationale behind the semantic tool used is to use a logical interpretation of Qualitative Spatial Reasoning (QSR) as a logical basis for representing and reasoning with the classes of an ontology. The selected theory is the well known *Region Connection Calculus* (RCC) [5], a mereotopological approach to QSR; it describes topological features of spatial relationships. It has been used in several subfields of AI as well as in SW [7]. In RCC, the ground relation is the *connection*, $C(x,y)$, with intended meaning: *"the topological closures of* x *and* y *intersect"*. The basic axioms of RCC are $\forall x[C(x,x)]$ and $\forall x,y[C(x,y) \rightarrow C(y,x)]$ and a set of definitions of main spatial relations jointly with another set of auxiliary axioms (see [5]). The theory provides a robust specification of QSR.

The set of binary relations formed by the eight jointly exhaustive and pairwise disjoint (JEPD) relations given in figure 1 is denoted by RCC8. This set is thought as a calculus for Constraint Satisfaction Problems (CSP) (see e.g. [11]). Another interesting calculus is RCC5, based on the set $\{DR, PO, PP, PPi, EQ\}$. Roughly speaking, the main difference between RCC5 and RCC8 is that the latter one allows the representation of knowledge that depends on topological frontiers. It has been empirically constated that RCC8 is more adequate than RCC5 as a tool for representing topological relations discriminated by humans [8]. The cognitive impact of this distinction on the spatial representation of a concept has been discussed in [3].

Mereotopological Interpretation of concepts of an ontology

Spatial representation is based on the *strong mereotopolical interpretation* [3]:

Two concepts $\mathsf{C}_1, \mathsf{C}_2$ of an ontology Σ are **Σ-connected** if

$$\Sigma \not\models \mathsf{C}_1 \sqcap \mathsf{C}_2 \equiv \bot$$

In logical terms, two concepts are connected if a logical model of Σ exists, where the interpretation of these concepts intersect. That is, there exists a potential situation of use of the ontology where the concepts share elements. The remaining RCC relations can be interpretated by means of their corresponding definition. The strong interpretation works on abstract spatial encodings of Σ. That is, it does not work on a concrete spatial interpretation of concepts, and does not need use individuals of the ontology. Formally, strong interpretation works as a powerfull logic formalism that insures robust representation.

The sound use and understanding of ontologies is based on a agreement between user and ontologist. Unfortunately, several reasons obstruct the agreement. The main one is that end users do not know the logical formalisms behind ontology web languages, so the user could not know hidden principles on which the ontologies are built. It would not help to increase the understanding of technologies involved in SW tools. Anyway, this fact might not be important if he uses amenable technologies for representing/repairing the anomalies found in its own ontology project. Visual encodings are very interesting for such purposes.

End-user preferences on visual representation are well known in other related fields of Information Systems. The spatial metaphor is a powerful tool in human information processing. The user will feel encouraged to repair the anomaly, although some obstacles exist: on the one hand, visual reparation may not be corresponded by a logical reparation of the ontology source. This occurs if there is no a formal semantics for supporting the change; on the other hand, repairs can be logically complex.

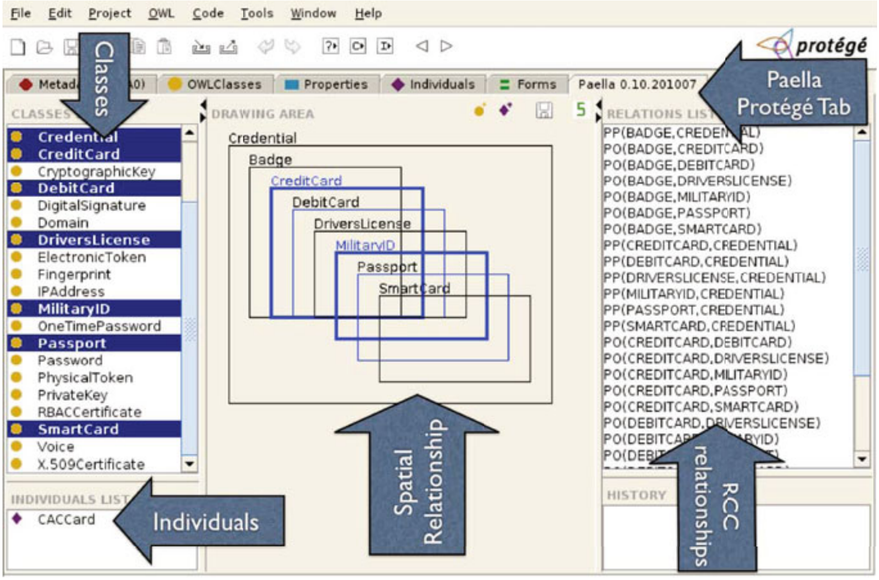
Visual cleaning of ontologies will be important for future end users of ontology debugging systems[10] mainly due to three reasons. Firstly, it allows the user to summarize ontology contents. Secondly, since the user's information is often fuzzily defined, visualization can be used to help the user to get a nice representation. lastly, visualization can therefore help the user to interact with the information space.

4 Paella Tool

Roughly speaking, Paeela is an ontology reviewer through spacial metaphors. Specifically, this tool uses a visual/topological interpretation based on RCC and the logical/mathematical properties of ontologies, where non-expert users can transform ontologies as they see them, but keeping safe the formal properties of the ontology source. Although there are many tools for visual representation (for example Jambalaya), Paella also allows transformations. Therefore, this prototype represents

- A very useful tool for socially appropriate management of formal ontologies;in the past, restricted to expert users only.
- A tool to uncover hidden relationships in the ontology code between concepts, which would discover intentionally hidden relationships or even harmful to personal data, data security,etc, which are referenced to them.

Paella uses SWI-Prolog (and the library for CHR for reasoning with constraints) and integrates JAVA-SWI. The tools provide three different spatial interpretations,



Credential
CreditCard
CryptographicKey
DebitCard
DigitalSignature
Domain
DriversLicense
ElectronicToken
Fingerprint
IPAddress
MilitaryID
OneTimePassword
Passport
Password
PhysicalToken
PrivateKey
RBACCertificate
SmartCard
Voice
X.509Certificate

CACCard

```
PP(BADGE,CREDENTIAL)
PO(BADGE,CREDITCARD)
PO(BADGE,DEBITCARD)
PO(BADGE,DRIVERSLICENSE)
PO(BADGE,MILITARYID)
PO(BADGE,PASSPORT)
PO(BADGE,SMARTCARD)
PP(CREDITCARD,CREDENTIAL)
PP(DEBITCARD,CREDENTIAL)
PP(DRIVERSLICENSE,CREDENTIAL)
PP(MILITARYID,CREDENTIAL)
PP(PASSPORT,CREDENTIAL)
PP(SMARTCARD,CREDENTIAL)
PO(CREDITCARD,DEBITCARD)
PO(CREDITCARD,DRIVERSLICENSE)
PO(CREDITCARD,MILITARYID)
PO(CREDITCARD,PASSPORT)
PO(CREDITCARD,SMARTCARD)
PO(DEBITCARD,DRIVERSLICENSE)
PO(DEBITCA[illegible]RYID)
PO(DEB[illegible])
```

Fig. 2 Screenshot of Paella

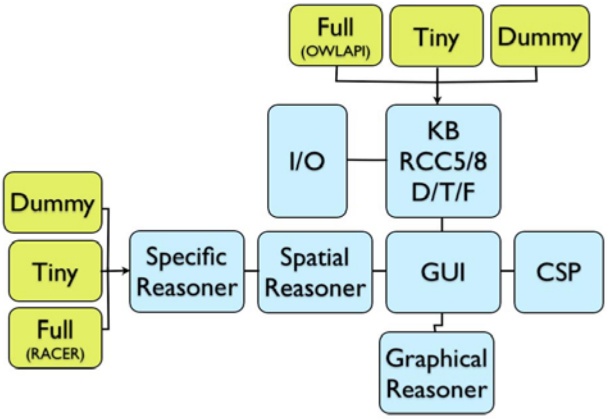


Fig. 3 Architecture of Paella

according to the nature of data which have associated different debugging methods: Dummy Paella, Tiny Paella and Full Paella (see Fig. 3). The latter is the used one for visual analysis of OWL ontologies, and it uses as automated reasoning system RACER[3] for computing the spatial relationship between classes.

[3] http://www.racer-systems.com/products/tools/index.phtml

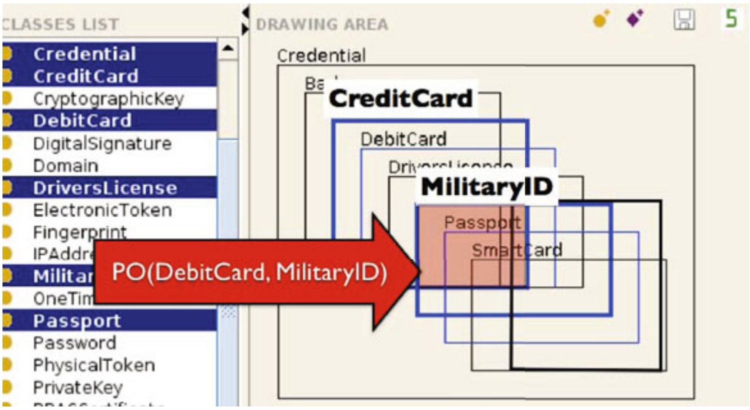


Fig. 4 Representational anomaly in serviceSecurity.owl

5 Analysis with Paella of Ontologies on Security

In order to show how Paella is useful to detect potential anomalies, three security ontologies, with different logical complexity. The selected ontologies are serviceSecurity.owl, SecurityOntology_min.owl[4] and MemoryProtection.owl[5]. Using Paella, the experiments show that several representational anomalies exist. The graphical representation allows one to visualize anomalies of type (2) and (4).

The most common anomaly is the vague relationship between critical concepts. To illustrate a problem of type (2), it is interesting to analyse a specific example in serviceSecurity.owl (part of NRL security ontology). In Fig. 4 a screenshot of Paella is depicted, showing that C*reditCard* partial overlaps to M*ilitaryID* (under Strong Interpretation). That is -by the strong interpretation- this ontology is potentially dangerous if a population of data considers a credit card as military identification for military installations where the access is restricted. Paella also provides graphical movements to make both classes disjoint, translating this spatial configuration of ontology source (that is, it repairs the anomaly). Note that this kind of anomaly does not imply logical inconsistency, only warns of potential non intended models of the ontology. In the example of 4, the reparation in Paella consists in an axiom stating that the classes are disjointed.

6 Final Remarks and Future Work

The analysis of ontologies used in critical information systems, as Security Ontologies, must be performed using formal methods that insure their safety in potential

[4] With logical complexity AL and ALCHOIQ(D) respectively. Both from NRL ontology, `http://chacs.nrl.navy.mil/projects/4SEA/ontology.html`

[5] A SHOIN(D) ontology, `http://www.ida.liu.se/~iislab/projects/secont/`

uses. In this paper a tool for carrying out the analysis is presented. It is a visual semantic tool that insures logical compatibility between spatial representation and ontology specification by means of an interpretation of RCC as meta-ontology [3]. The tool allows security experts revise ontologies without the need to be ontological engineers. The future work is driven to analyse how such revisions could be suggested from data, using formal concepts as *cognitive entropy*. [4].

Acknowledgements. Partially supported by TIN2009-09492 project of Spanish Ministry of Science and Innovation, cofinanced with FEDER founds.

References

1. Alonso-Jimenez, J.A., Borrego-Diaz, J., Chavez-Gonzalez, A.M., Martin-Mateos, F.J.: Foundational Challenges in Automated Semantic Web Data and Ontology Cleaning. IEEE Intelligent Systems 21(1), 42–52 (2006)
2. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Sci. American, Singapore (May 2001)
3. Borrego-Díaz, J., Chávez-González, A.M.: Visual Ontology Cleaning: Cognitive Principles and Applicability. In: Sure, Y., Domingue, J. (eds.) ESWC 2006. LNCS, vol. 4011, pp. 317–331. Springer, Heidelberg (2006)
4. Borrego-Díaz, J., Chávez-González, A.M.: Using Cognitive Entropy to Manage Uncertain Concepts in Formal Ontologies. In: da Costa, P.C.G., d'Amato, C., Fanizzi, N., Laskey, K.B., Laskey, K.J., Lukasiewicz, T., Nickles, M., Pool, M. (eds.) URSW 2005 - 2007. LNCS (LNAI), vol. 5327, pp. 315–329. Springer, Heidelberg (2008)
5. Cohn, A.G., Bennett, B., Gooday, J.M., Gotts, N.M.: Representing and Reasoning with Qualitative Spatial Relations about Regions. In: Stock, O. (ed.) Spatial and Temporal Reasoning. Kluwer, Dordrecth (1997)
6. Fenz, S., Ekelhart, A.: Formalizing information security knowledge. In: Proc. 4th Int. Symp. on Inf., Computer, and Comm. Security ASIACCS 2009, pp. 183–194. ACM, New York (2009)
7. Grütter, R., Scharrenbach, T., Bauer-Messmer, B.: Improving an RCC-Derived Geospatial Approximation by OWL Axioms. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) ISWC 2008. LNCS, vol. 5318, pp. 293–306. Springer, Heidelberg (2008)
8. Knauff, M., Rauh, R., Renz, J.: A Cognitive Assessment of Topological Spatial Relations: Results from an Empirical Investigation. In: Frank, A.U. (ed.) COSIT 1997. LNCS, vol. 1329, pp. 193–206. Springer, Heidelberg (1997)
9. Miller, R.R.: Information management in the aftermath of 9/11. Comm. ACM 45(9), 31–33 (2002)
10. Murtagh, F., Taskaya, T., Contreras, P., Mothe, J., Englmeier, K.: Interactive Visual Interfaces: A Survey. Artificial Intelligence Review 19, 263–283 (2003)
11. Renz, J.: Qualitative Spatial Reasoning with Topological Information. In: Renz, J. (ed.) Qualitative Spatial Reasoning with Topological Information. LNCS (LNAI), vol. 2293, p. 31. Springer, Heidelberg (2002)
12. Pereira, T., Santos, H.: An Ontology Based Approach to Information Security. Communications in Computer and Information Science 46(part 2), 183–192 (2009)
13. Smith, S.W., Spafford, E.H.: Grand Challenges in Information Security: Process and Output. IEEE Security and Privacy 2(1), 69–71 (2004)