Towards an Ontology of Trust for Situational Understanding

Owain Carpanini and Federico Cerutti

Cardiff University, U.K.

Abstract. In this paper we propose a computational methodology for assessing the impact of trust associated to sources of information in situational understanding activities—i.e. relating relevant information and form logical conclusions, as well as identifying gaps in information in order to answer a given query. Often trust in the source of information serves as a proxy for evaluating the quality of the information itself, especially in the cases of information overhead. We show how our computational methodology support human analysts in situational understanding by drawing conclusions from defaults, as well as highlighting issues that demand further investigation.

Keywords: computational models of trust, situational understanding, uncertainty

1 Introduction

Individuals and organisations have access to a rich and diverse source of information that can be exploited for situational understanding activities—i.e. relating relevant information and form logical conclusions, as well as identifying gaps in information in order to answer a given query. However, an open and enduring problem resides in managing the impact of trust measurements in such activities.

We propose a computational methodology for assessing the impact of trust in situational understanding to support human analysts with a sound ontology and the ability to reason with partial information. While the aspect of reasoning with partial information shares similarities with other approaches to qualitative decision making [5], to our knowledge in this paper we propose the first ontology for trust in situational understanding. We ground our preliminary investigation in a case study on the Wakefield case on the alleged links between vaccination and autism that will help us describing our desiderata for supporting human analysts (Section 2). We then propose a methodology satisfying those desiderata in Section 3, and we illustrate it developing further our case-study. As this preliminary work is part of an ongoing investigation, we will extensively discuss future directions in Section 4.

2 Motivational Scenario and Desiderata

In [12] (now retracted), Wakefield et al. suggest a link between some vaccinations and autism. This paper triggered extensive studies that resulted in a retraction notice [10]

that states: "it has become clear that several elements of the 1998 paper by Wakefield et al are incorrect."

Let us suppose that an analyst needs to answer the question "Do vaccinations cause autism?" (situation understanding task). For simplicity let us suppose they have access only to [12] and [10]: both are scientific papers discussing a medical topic, thus the analyst might assume that both of them can be highly trusted.

Our first desideratum is therefore (**Des1**) that trust has to be context-dependent: we can trust The Lancet on medical issues, but we should trust the Daily Mail on celebrities lifestyle.

Moreover, the analyst can see that [10] disputes [12]. Such a conflict might be automatically resolved if one of the sources is more trustworthy than the other: this requires (second desideratum, **Des2**) that trust needs to be expressed through an ordinal scale, i.e. it should be possible to determine whether, given the same context, a source of information is more trustworthy than another.

## 3   An Ontology of Trust for Situational Understanding

3.1   What is an Ontology?

An ontology comprises two components [1]: the vocabulary (TBox) and the assertions about individuals (ABox). The vocabulary consists of concepts, which denote sets of individuals; and roles, which denote binary relationships between individuals.

Elementary descriptions are atomic concepts and atomic roles. Complex descriptions can be built from them inductively with concept constructors. In abstract notation, we use the letters A and B for atomic concepts, the letter R for atomic roles, and the letters C and D for concept descriptions. Concept descriptions in $\mathcal{ALI}+$ are formed according to the following syntax rule: $C, D \rightarrow A$ (atomic concept) $| \top$ (universal concept) $| \perp$ (bottom concept) $| C \sqcap D$ (intersection) $| \forall R.C$ (value restriction, or range) $| \exists R.\top$ (limited existential quantification, or domain). R^- denotes the inverse role; roles can be transitive and symmetric. The syntax of $\mathcal{ALI}+$ can clearly be presented using the OWL 2 Web Ontology Language.[1] Due to space constraints we omit the formal description of semantics, that are given—as usual—by means of an interpretation. For ease of representation, in the following we will use a graph-based syntax, where nodes are either atomic concepts or individuals (identified by ♦), and edges are labelled with one of the following labels: *is-a* (representing sub-concept relation), *type* (membership assertion), or R (roles).

Moreover, as presented in [8], ORL is a language for expressing Horn clause rules that extends the OWL language in a syntactically and semantically coherent manner. A rule has the form *antecedent* → *consequent*, where both *antecedent* and *consequent* are conjunctions of atoms written $a_1, \ldots, a_n$. Variables are prefixed with a question mark— e.g. ?x. The model-theoretic semantics for ORL is an extension of the semantics given by an interpretation. A rule is satisfied by an interpretation iff every binding—mapping to elements of the domain—that satisfies the antecedent also satisfies the consequent. An interpretation satisfies an ontology iff it satisfies every axiom (including rules) and fact in the ontology [8].

---

[1] https://www.w3.org/TR/owl2-overview/
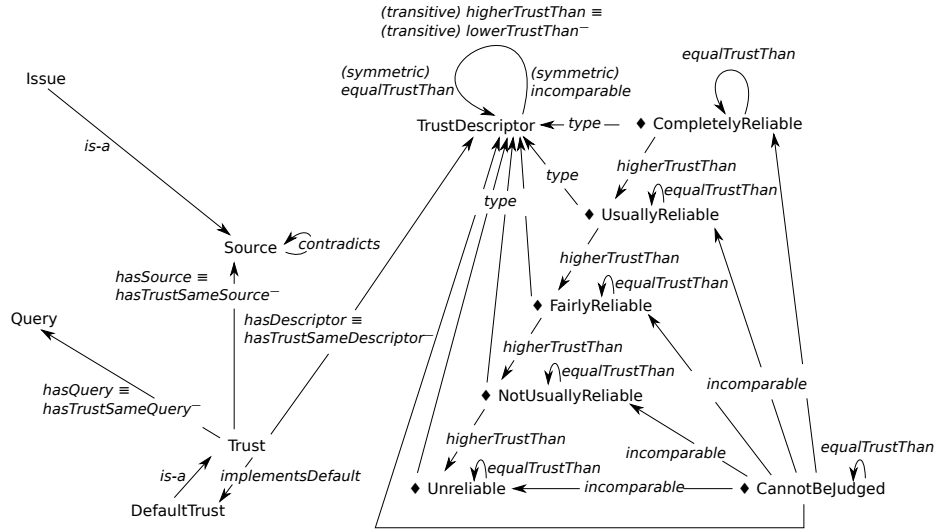
Fig. 1. Graphical representation of SitUTrustOnto: nodes with ♦ are individuals, otherwise atomic concept. Edges are labelled with *is-a* (subclass relations), with *type* (membership assertion); and with atomic roles. For instance, *(transitive) higherTrust* ≡ *(transitive)lowerTrustThan*⁻ describes the role *higherTrust* with domain and range Trust, such as it is the inverse of *lowerTrust*.

R1:	$\left\{\begin{array}{l}\text{Trust(?x), DefaultTrust(?y),}\\\text{implementsDefault(?x,?y), TrustDescriptor(?t),}\\\text{hasDescriptor(?y,?t)}\end{array}\right\}$ → hasDescriptor(?x,?t)
R2:	$\left\{\begin{array}{l}\text{Source(?x), Source(?y), contradicts(?x,?y),}\\\text{Trust(?ax), Trust(?ay),}\\\text{hasTrustSameSource(?x,?ax),}\\\text{hasTrustSameSource(?y,?ay), Query(?q),}\\\text{hasQuery(?ax,?q), hasQuery(?ay,?q), hasDescriptor(?ay,?ty),}\\\text{hasDescriptor(?ax,?tx), equalTrustThan(?tx,?ty)}\end{array}\right\}$ → Issue(?x)
R3:	$\left\{\begin{array}{l}\text{Source(?x), Source(?y), contradicts(?x,?y),}\\\text{Trust(?ax), Trust(?ay),}\\\text{hasTrustSameSource(?x,?ax),}\\\text{hasTrustSameSource(?y,?ay), Query(?q),}\\\text{hasQuery(?ax,?q), hasQuery(?ay,?q), hasDescriptor(?ay,?ty),}\\\text{hasDescriptor(?ax,?tx), lowerTrustThan(?tx,?ty)}\end{array}\right\}$ → Issue(?x)
R4:	{Issue(?x), Source(?y), contradicts(?x,?y)} → Issue(?y)

Table 1. ORL Rules in SitUTrustOnto.

3.2   Our Proposal: SitUTrustOnto

Figure 1, together with Table 1, depicts SitUTrustOnto, our proposed ontology of trust for situational understanding.[2]

   The concept Source contains sources of information, e.g. blog posts, twits, scientific papers.... The concept Query describes the situation that needs to be understood, e.g. specific queries an analyst has to address, such as "Do vaccinations cause autism?"

   The concept TrustDescriptor contains trust descriptors: given our interest in supporting human analysts and not to replace them, we chose to rely on the *admiralty rating* [9] that identifies the following five descriptors listed in decreasing order: *Completely Reliable*, *Usually Reliable*, *Fairly Reliable*, *Not Usually Reliable*, *Unreliable*; plus a sixth, incomparable, one, namely *Cannot Be Judged*. Figure 1 shows the six individuals belonging to TrustDescriptor, and their relationships expressed by *equalTrustThan*—identifying that two descriptors are equivalent; *higherTrustThan* and *lowerTrustThan*—expressing ordering, and thus satisfying (**Des2**); and *incomparable*—when two descriptors cannot be compared.

   The concept Trust describes the relationship between a source of information, a query, and a trust descriptor, thus providing the context in which to assess the trust in a source of information for a given query. Please note that there is no role linking directly Source to TrustDescriptor, nor Source to Query, thus satisfying (**Des1**). Moreover, DefaultTrust is a sub-concept of Trust that provides *default* trust accounts between some types of queries and some sources of information. For instance, we might want to express that scientific papers addressing medical queries generally have high trust, and thus we can create a specific individual belonging to DefaultTrust. This means that when a new piece of information is added to the ontology, we can exploit defaults for assessing its trustworthiness using the rule R1 from Table 1 (see Section 3.3 for a complete example).

   Finally, since different Sources can contradict each other, Issue is populated by the means of rules R2, R3, and R4 of Table 1. This is the case where two sources of information ♦ a and ♦ b, related to the same query, contradict each other, i.e. ♦ a *contradicts* ♦ b, and either they have the same level of trust (*equalTrustThan*), cf. R2 in Table 1; or the trust in ♦ a is *lowerTrustThan* ♦ b, cf. R3 in Table 1. This is based on the assumption that if ♦ a *contradicts* ♦ b, but ♦ a is more trustworthy than ♦ b, an analyst would accept ♦ a and discard ♦ b. We agree that this is not necessary the case, and further comments are outlined in Section 4. Finally, if a source of information also belong to Issue, also the sources it is in conflict with belong to Issue (cf. R4 in Table 1.

3.3   SitUTrustOnto and our Case Study

Figure 2 depicts (in blue) SitUTrustOnto enriched with elements of the scenario discussed in Section 2. The query ♦ Do vaccinations cause autism? is a member of Vaccination, a sub-concept of Medicine, sub-concept of Query. The Wakefield et al. [12] paper ♦ https://goo.gl/83pRSA and the retraction notice [10] ♦ https://goo.gl/XpFQgK both are ScientificPapers serving as Source, and they are linked to the query through two individuals belonging to Trust, ♦ trust83pRSA and ♦ trustXpFQgK respectively.

---

[2] Available at https://goo.gl/DLq6dz
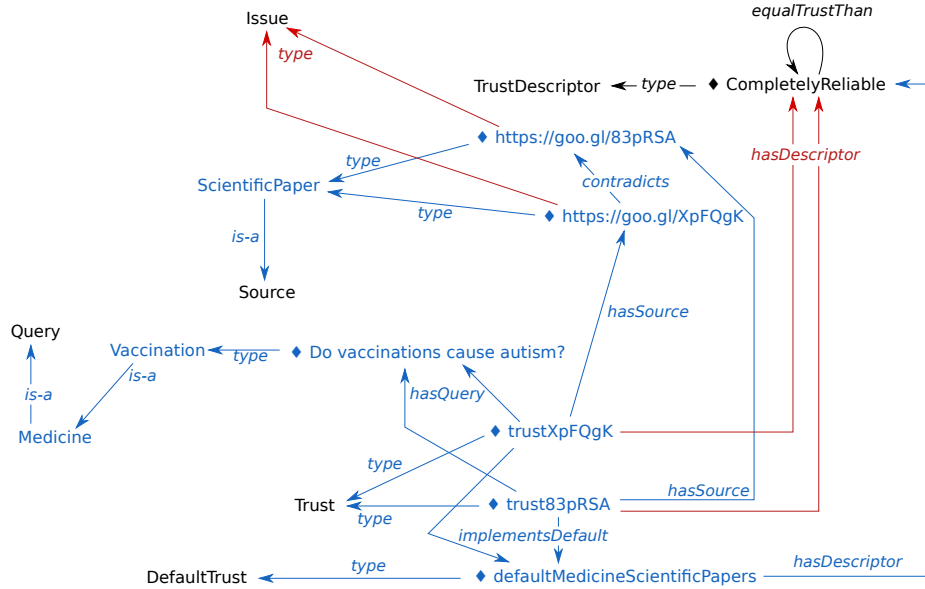
Fig. 2. SitUTrustOnto extended to include elements of the use case (in blue), and (some of) the inferred relationships (in red).

We also describe the assumption that when dealing with a medical topic, a scientific paper is ♦ CompletelyReliable with ♦ defalutMedicineScientificPaper belonging to DefaultTrust, and linked to the trust assessments ♦ trust83pRSA and ♦ trustXpFQgK through the role *implementsDefault*.

As a result of automated reasoning (in red in Figure 2), rule R1 is fired twice leading to assess both ♦ https://goo.gl/83pRSA and ♦ https://goo.gl/XpFQgK as ♦ CompletelyReliable. Because of that, and because ♦ https://goo.gl/XpFQgK *contradicts* ♦ https://goo.gl/83pRSA, they both belong to Issue (rules R2 and R4) thus flagging to the analyst the need for further investigation on these two sources of information.

4   Conclusion

In this paper we proposed a computational methodology for evaluating trust associated to sources of information in situational understanding. In particular, we showed how our computational methodology supports situational understanding by drawing conclusions from defaults, as well as highlighting issues due to conflicts between sources of information that demand further investigation to be solved. For instance, in our case-study, the assumption that scientific papers share the highest level of trust when considered in the context of a scientific enquiry proved to be debateable.

This is a first investigation towards a support system for trust management in situational understanding. As part of future work we plan to evaluate techniques for automatic evaluation of trust: we will investigate how topic modelling—i.e. automatically

identifying relevant topics in a written document, for instance using the Latent Dirichlet Allocation (LDA) [3]—and similarities of sources—e.g. articles in The Lancet are more similar to articles in the New England Journal of Medicine rather than to articles in the Daily Mail—can help suggesting trust measures for new pieces of information. Moreover, we will also investigate how to automatically identify problems with default assumptions, i.e. questioning whether there is enough evidence suggesting that a default assignment needs to be revisited. This will be part of a large empirical evaluation.

The notion of conflict in this preliminary paper is quite rudimentary: as part of future investigations we will exploit argumentation systems [2] and the Argument Interchange Format ontology [4] for their ability to reason about different types of conflicts. Moreover, following [11] where an argumentation system for supporting intelligence analysis is proposed, we will also investigate the relationship between trust and provenance of information [7].

Finally, as answers to situational understanding queries often require to fuse pieces of information into a single, coherent document, we will assess how to evaluate the trust of fused documents by building on qualitative decision under uncertainty [6].

References

1. Baader, F., Nutt, W.: Basic Description Logics. In: The Description Logic Handbook, pp. 43–95. Cambridge University Press (2003)
2. Besnard, P., Hunter, A.: Constructing argument graphs with deductive arguments: a tutorial. Argument & Computation 5(1), 5–30 (2014)
3. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research 3, 993–1022 (2003)
4. Chesnevar, C.I., McGinnis, J., Modgil, S., Rahwan, I., Reed, C., Simari, G.R., South, M., Vreeswijk, G.A.W., Willmot, S.: Towards an argument interchange format. The Knowledge Engineering Review 21(04), 293 (2006)
5. Dubois, D., Fargier, H., Prade, H., Perny, P.: Qualitative decision theory: From savage's axioms to nonmonotonic reasoning. Journal of the ACM 49(4), 455–495 (2002)
6. Dubois, D., Prade, H., Rico, A.: Residuated variants of Sugeno integrals: Towards new weighting schemes for qualitative aggregation methods. Information Sciences 329, 765–781 (2016)
7. Hartig, O., Zhao, J.: Using web data provenance for quality assessment. In: First International Conference on Semantic Web in Provenance Management. pp. 29–34. CEUR-WS.org (2009)
8. Horrocks, I., Patel-Schneider, P.F.: A proposal for an OWL rules language. In: Proceedings of WWW '04. pp. 723–731 (2004)
9. Prunckun, H.: Handbook of Scientific Methods of Inquiry for Intelligence Analysis. The Scarecrow Press (2010)
10. The Editors of The Lancet: Retraction–Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. The Lancet 375(9713), 445 (2010)
11. Toniolo, A., Norman, T.J., Etuk, A., Cerutti, F., Ouyang, R.W., Srivastava, M., Oren, N., Dropps, T., Allen, J.A., Sullivan, P.: Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis. In: Proceedings of AAMAS 2015. pp. 781–789 (2015)
12. Wakefield, A., Murch, S., Anthony, A., Linnell, J., Casson, D., Malik, M., Berelowitz, M., Dhillon, A., Thomson, M., Harvey, P., Valentine, A., Davies, S., Walker-Smith, J.: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. The Lancet 351(9103), 637–641 (1998)