SeDAn: a Plausible Reasoning Approach for Semantics-based Data Analytics in Healthcare

Hossein Mohammadhassanzadeh[1], Samina Raza Abidi[2], Mohammad Salman Shah[1], Mehdi Karamollahi[3] and Syed Sibte Raza Abidi [1]

[1] NICHE Research Group, Faculty of Computer Science, Dalhousie University, Canada
{hassanzadeh, raza.abidi}@dal.ca, msalmanshah1987@gmail.com

[2] Medical Informatics, Faculty of Medicine, Dalhousie University, Canada
samina.abidi@dal.ca

[3] Faculty of Computer Science, The University of Rome La Sapienza, Italy
karamollahi.1771946@studenti.uniroma1.it

Abstract. Plausible Reasoning (PR) is an inferencing mechanism to derive solutions when dealing with incomplete knowledge. When developing data-driven models for clinical decision support, the completeness of the data is always a consideration. PR provides a practical approach to extend the knowledge-base of a clinical decision support system by abstracting plausible assertions from heath data. Implementation of plausible reasoning relies on fine-grained knowledge of how different concepts are semantically related. The Semantic Web provides formalisms to semantically represent knowledge at various levels of expressivity, and to reason over the knowledge to perform semantic analytics based on healthcare data. This paper proposes a SEmantics-based Data ANalytics framework (SeDan) to investigate the potential of implementing plausible reasoning using the Semantic Web technologies. In particular, we will evaluate the efficacy of the proposed framework in healthcare to perform effective semantic analytics using partial health data to make better decisions in disease diagnosis and long-term care. We demonstrate the efficacy of SeDan by answering medical queries posed by BioASQ challenges using Disease ontology, DrugBank and Semantic MEDLINE databases.

Keywords: Plausible Reasoning, OWL, Semantic Analytics, Semantic Web Reasoning.

1 Introduction

The massive volume of diverse data from clinical practice, healthcare and biomedical research is an opportunity for medical big-data analytics. However, due to the intrinsic nature of data that may be incomplete and inaccurate, the interpretation of data and its associations might be a serious challenge [1]. In this regard, innovative methods, algorithms and tools are needed to facilitate knowledge representation, exchange and reasoning, which is understandable for both human and machine [2].

In applying data analytics to real health applications, especially with large and complex datasets, the patient data is typically sparse and incomplete. To deal with missing

D. Impedovo and G. Pirlo (Eds.), Workshop on Artificial Intelligence with Application in Health, Bari, Italy, November 14, 2017.

data, two approaches exist: (i) removing the objects (entities) or features with incomplete data, and (ii) data filling based on experts' experience, fuzzy, or Bayesian models for a best guess estimation [3]. The former solution considerably reduces the size of data, and the latter needs expert's input, calculating statistical associations between data, or requires probability distribution that may not be always available. While Plausible Reasoning (PR) is an alternative reasoning method that derives solutions when complete data is lacking or non-existent.

Comparing to clinical decision-making process, physicians, basically, observe the available knowledge to make a diagnosis or order a treatment. If the existing knowledge is not sufficient, then the physicians leverage their own tacit knowledge to discover the correlations within existing medical data, draw new relationships and infer the missing knowledge [1]. Plausible reasoning follows the physicians' thinking process to generate new hypothesis. Plausible reasoning does not conform to strict logical formalisms; but it provides a mechanism to infer new knowledge, albeit a weaker inference, especially when working with the Open World Assumption (OWA) [4]. For such cases, PR can infer new and missing relationships by leveraging how different concepts are semantically interrelated [5].

The Semantic Web (SW) framework provides logic-based formalisms to semantically represent knowledge at various levels of expressivity. The SW also offers effective built-in support for deduction-based reasoning, including Description Logic (DL) reasoning and rule-based languages, that conform to the OWA. The results, therefore, are demonstrative and consistent with the knowledge. Despite the great potential of the SW technologies in different domains, including healthcare, there is currently a lack of support for representing and reasoning with uncertainty and incompleteness in the SW framework, which is an irresolvable part of our daily life [6], [7]. This shortcoming limits the use of the SW-based approaches in clinical decision support systems that require efficient handling of incompleteness [8].

This drawback of SW has led to several approaches [6] introducing probabilistic variants and fuzzy extensions [9] to the Web Ontology Language (OWL) to deal with vague information. Such probabilistic/fuzzy OWL extensions improve the capability of SW reasoners in dealing with uncertainty. However, they are only applicable to cases where the truth of facts has some degree of ambiguity (qualitative uncertainty), not the cases where uncertainty is result of lack of knowledge (quantitative uncertainty) [10].

In this regard, [1] implemented a multi-strategy reasoning framework, including deductive, inductive and analogical reasoning, within the SW framework. They leveraged ontological knowledge to increase the expressivity and accuracy of plausible reasoning methods. They showed that implementing plausible reasoning methods can extend the coverage of an incomplete KB, and exploiting enriched OWL ontologies can significantly increase the accuracy of the results. However, there is still a lack of non-deductive reasoning support in the logic layer of the SW. Current study, aims to introduce plausible reasoning as one non-deductive approach targeting the logic layer of the SW.

In this research, we propose the concept of semantic analytics as the analysis of semantically annotated data, i.e., data represented in Resource Description Framework (RDF) to infer new knowledge, whilst adhering to the SW's OWA about knowledge incompleteness, by using expressive semantics and semantics relevant reasoning methods [11]. We believe that RDF Schema and OWL, expressing additional semantics on top of RDF, is one way to achieve semantic analytics. There are a number of ways that new facts can be inferred when we have complete knowledge, however within the OWA we need to account for incomplete knowledge that may lead to non-deductive reasoning, and plausible reasoning is one such reasoning approach.

This research aims to investigate the potential of implementing plausible reasoning within the SW, targeting a semantic analytics framework for health data analytics, especially when working with large health datasets. In line with this objective, we aim to: (i) introduce additional markups (plausible extension to OWL) that extend OWL semantics to better capture and represent plausible semantics, (ii) develop a semantic analytics framework using query-rewriting algorithm to discover new associations between underlying domain-specific data, (iii) evaluate framework using health data.

2 Plausible Reasoning

Plausible Reasoning, which is non-demonstrative, ampliative and non-monotonic, is a weak inference approach that identifies the associations between the question and the knowledge retrieved from memory and draw the line of inference based on those associations. Plausible reasoning performs inferencing by using a set of frequently recurring patterns that do not occur in formal logic [12]. A plausible reasoning stack is introduced in [1]. The stack is comprised of a set of plausible patterns and 3 plausible reasoning mechanisms that use the patterns to infer new rules and facts. [1] also classifies plausible patterns into 3 groups (Table 1): hierarchy-based patterns, order-based and hybrid.

Table 1. - Plausible Patterns [1]

Plausible Pattern	Description
Generalization [a]	Passing from a given set of objects to a larger set that contains the given set.
Specialization [a]	Passing from a given set of objects to a smaller set that is contained in the given one.
Interpolation [b]	Creating a new relation from observation space X to conclusion space Y, where $x_i \in X$ is not mapped to any y $\in Y$ (unknown relation), but other relations from $x_h, x_j (\neq x_i)$ to Y and $x_h < x_i < x_j$ are known.
A Fortiori [b]	An inference from a proposition with high degree of confidence to a less confident proposition that is not clearly specified but is implicit in the first one.
Similarity/ Dissimilarity [c]	Moving between any two comparable nodes (siblings) in the concept hierarchy.

[a] Hierarchy-based patterns, [b] Order-based patterns, [c] Hybrid patterns

Hierarchy-based patterns move between the nodes in hierarchical structure, from parent to child or vice versa, to perform a hierarchical plausible inference. Order-based patterns leverage measurable properties (partial order) to compare concepts regarding their size, order, location, ranking, etc. and infer new pieces of knowledge. However, hybrid patterns will be performed using both hierarchical relations and partial order of concepts to infer a plausible answer; they probe hierarchy and move between any two comparable nodes, or consider the concepts that are analogues regarding some measurable properties. The utility of ordered-based patterns within inductive and analogical reasoning has been studied and approved [1]. While, current study investigates the efficiency of all the plausible patterns together, either working alone or in a combination with other patterns. Definition 1 provides a formal notation for PR.

Definition 1: Let $\mathcal{K}$ be a knowledge base including terminological constructs $\mathcal{T}$ and incomplete assertional knowledge $\mathcal{A}$ ($\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$) and Q a query. A plausible reasoner $\mathrm{plbRes}(Q, \mathcal{T}, \mathcal{A})$ returns a set of solutions for Q:

$$\mathrm{P}_{Q}^{\mathcal{K}} = \{ \langle plbAns , \{\pi_1, \ldots, \pi_n\} \rangle \mid \pi \in \Pi, 1 \leq \mathrm{n} \}$$

Which *plbAns* is a plausibly inferred solution, Π is the set of plausible patterns, and $\{\pi_1, \ldots, \pi_n\}$ demonstrates the plausible pattern(s) involved in the reasoning process.

3 Query Rewriting within the Semantic Web

Query Rewriting (QR) algorithms use ontological constructs to transform a given query to an expanded version that extracts both explicit (what a KB knows) and implicit (what it assumes) knowledge from the data [13], [14]. Therefore, QR can be used as a technique to implement plausible patterns and solve queries over an incomplete KB. Within the SW framework, OWL 2 QL profile provides a query rewriting mechanism to query data through an ontology. OWL 2 QL is underpinned by DL-Lite family of description logics. The OWA made in DLs makes OWL 2 QL suitable to work with incomplete knowledge in the SW scenarios [14], [15]. Independence from data and support of other variants of DL-Lite have made QL a suitable approach to Ontology Based Data Access (OBDA) in large RDF stores with different levels of expressivity.

However, the DL-Lite underlying OWL 2 QL roughly describes the allowed operators, which limits their expressivity when it comes to the domains with uncertainty and incompleteness [16]. The axioms within QL support variety of inferences in OBDA, but it may not cover all the plausible semantics. The goal of our work is to introduce a plausible extension to OWL QL to support plausible relations and properties.

4 SeDan: Semantics based Data Analytics Framework

To achieve the semantic analytics, we propose a framework (Fig. 1) that implements a plausible reasoner to infer new knowledge from RDF knowledge bases. This reasoner develops plausible reasoning patterns by manipulating the underlying graph directly with SPARQL query rewriting using OWL DL constructs.

The proposed framework mainly includes three modules: knowledge sources, plausible reasoner and user interface. *Knowledge sources* provide terminological constructs to be consumed during the reasoning process, and assertional knowledge to be used to evaluate the extended query. The *plausible reasoner* (discussed more in the following section) delivers semantics analytics by running a query rewriting algorithm to perform plausible patterns and infer a set of so-called certain solutions. The system accepts the query with a list of desired plausible patterns via the *user interface*, and in return, delivers the plausible answer(s) and their justifications.

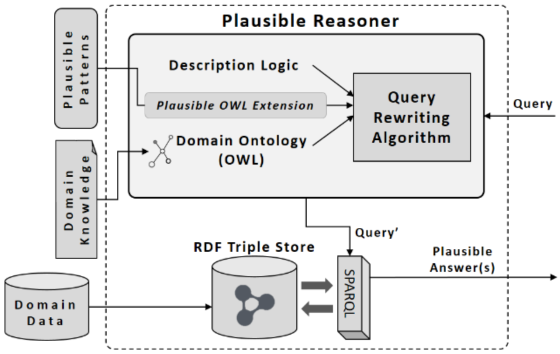

Fig. 1. Proposed semantics based data analytics framework

4.1 Plausible Extension to OWL

Standard reasoning capabilities within OWL QL profile support various types of ontology-based inference – *rdfs:subClassOf* represents hierarchical relations and *owl:sameAs* conducts similarity. However, QL does not support all the semantics required in plausible patterns, like partial order or context. Therefore, it is needed to consider how we can extend OWL in the cases that it has not enough expressivity. In this regard, plausible extension to OWL includes defining new classes, followed by defining new properties that use new (and existing) classes to express new relations. Table 2 demonstrates a subset of the proposed extension to OWL.

Table 2. Plausible extension to OWL (PLOWL)

Class Name	Supper Class		On Property	
OrderedProperty	ObjectProperty		-	
Context	Class		hasContext	
PlausiblePattern	Class		inferredViaPattern	
Property Name	**Type**	**Domain**	**Range**	**Inverse Property**
standsBefore	Ordered Property	Entity	Entity	standsAfter
standsAfter	Ordered Property	Entity	Entity	standsBefore
hasContext	Object Property	Entity	Context	-
inferredViaPattern	Object Property	Plausible Answer	Plausible Pattern	-

An *ordered property* is a property to reflect partial order of two entities w.r.t a measurable property (*plowl:Context*). More formally, if *P* is an *plowl:OrderedProperty*, any instance of *P*, like *(x,y)*, implies *x* is bigger, older, etc. than *y* or vice versa. From this, the plausible reasoner would be able to conduct interpolation and a fortiori reasoning. In Table 2, *plowl:standsAfter* and *plowl:standsBefore* are instances of *plowl:OrderedProperty* demonstrates how entities are comparable. Similarly, *plowl:hasContext* indicates the specific context in which the ordered property is meaningful.

4.2 Query Rewriting Algorithm

In this section, we present the proposed QR algorithm (Algorithm 1) that supports plausible reasoning patterns in the SeDan reasoning engine. We makes use of GCLRR algorithm [17] which transforms a query Q into a *Union of Conjunctive Queries (UCQs)* by applying the TBox axioms to the body atoms of the query. UCQ is one of the most common approach for computing a so-called perfect rewriting of a query. A UCQ is a set of conjunctive queries of the same arity and the same query predicate. Algorithm 1 demonstrates the proposed algorithm.

To start the rewriting, the algorithm needs an initial query, a set of preferred plausible patterns to limit the extended query to those patterns, and an ontology based on $DL - Lite_T$ axioms that is semantically enriched with introduced Plausible OWL extension. Starting with the initial query, the algorithm tries to replace the body atom of the query D (step 7), with new atom D'. The atom D' should be (i) semantically related to D ($\exists\alpha \in \mathcal{T}\ \alpha(D, D')$), and (ii) applicable to the preferred plausible patters (step 6). For example, *rdfs:subClassOf* is applicable to generalization, *owl:instanceOf* is used in specialization, *owl:sameAs* conducts (dis)similarity, and *plowl:standsAfter* is applicable to a fortiori and interpolation. The new conjunctive query resulting from replacing an atom will be added to R, the set of conjunctive queries. This algorithm keeps formulating new queries until there is no unique query to be added.

Algorithm 1. The proposed QR algorithm

Input: A query in a triple format, a set of plausible patterns
$\pi \in \Pi: \{GEN, SPEC, SIM, DIS, FORT, INTP\}$, $DL - Lite_T$ TBOx $\mathcal{T}$ enriched with PL-OWL extension
Output: R, a set of rewriting queries.
1: $\mathrm{R} = \{Q\}$;
2: **repeat**
3: **foreach** $query$ Q $\in$ R **do**
4: **foreach** $atom\ D\ in\ Q$ **do**
5: **foreach** $axiom\ \alpha \in \mathcal{T}$ **do**
6: **if** α is applicable to any $\pi \in \Pi$, w. r. t. D
7: $\mathrm{Q}' = \exists D'.\mathrm{Q}(D \rightarrow D') \wedge \alpha(D, D')$;
8: $\mathrm{R} = \mathrm{R} \cup \{Q'\}$;
13: **until** no unique query can be added to R;
14: **return** R;

5 Improving clinical decision support using SeDan

The SeDan framework has the potential to be used for decision-making and problem solving in any domain, which suffer from incomplete knowledge. However, we have focused on healthcare applications for the following reasons:

- Semantic analytics is very relevant to healthcare, as it is predominantly a knowledge-intensive domain. The opportunity to capture and leverage semantics via inference or query processing is vital for supporting both disease diagnosis and long term care (e.g. predictive and preventive diagnosis of chronic diseases) [18].
- A vast amount of health data is available from many diverse automated information systems including Electronic Health Records (EHR), Personal Health Records (PHR), Electronic Medical Records (EMR). Effective semantic analytics of data enables the extraction of potential relationships existing in healthcare data to provide insights that can assist healthcare providers to make better decisions.

To demonstrate the efficacy of SeDan and the feasibility of our query rewriting algorithm, we provide two case studies where we attempt to answer two questions from BioASQ challenges [19] using DrugBank [20], Disease Ontology [21] and Semantic

MEDLINE[1] database [22]. BioASQ challenges are a series of tasks in which participants are asked to respond to a set of questions posed by medical expert. The DrugBank is a bioinformatics and cheminformatics resource that includes detailed drug data. Disease ontology is standardized ontology for human disease, and Semantic MEDLINE database is a repository of 89.2 million sematic triples extracted from PubMed articles. For the sake of simplicity, in the examples below, we only discuss one conjunctive query out of the possible dozens of queries resulting from query rewriting algorithm.

5.1 Example 1: Migalastat treats Fabry Disease?

In this case study, when the question "*Is Migalastat used for treatment of Fabry Disease?*" (BioASQ challenge, Task 5b) is posed to the SemMedDB, the traditional approach returns a response 'No' as it cannot find any matching triple. The initial SPARQL syntax of the question can be written as below:

```
Initial SPARQL query:
@PREFIX sem: <https://skr3.nlm.nih.gov/SemMed#>
   ASK { "Migalastat" sem:treats  "Fabry Disease" }
```

```
Answer:
No
```

Code 1. Initial query answering if Migalastat treats Fabry Disease

By posing the failed query to SeDan framework, it uses ontological semantics to conduct the query rewriting. The QR algorithm explores the domain ontology to find any hierarchical/ordered relationships that matches any or a combination of the subject, object, and predicate of the triple in the question. Then, the query transformation would be performed by replacing the new atom with the matching atom in the triple.

Regarding the failed query (Code 1), and using the DrugBank ontology, we know *Migalastat is an alpha-Galactosidase* (DrugBank: DB05018). Based on generalization pattern, QR algorithm replace the subject of the triple (*Migalastat*) with its super class (*alpha-Galactosidase*) with this logic that *"if a category of drugs can treat a disease, then any subclass or instance of that category would treat the disease as well".* Using the relevant ontology axiom, transformation t (Fig. 2) can be conducted:

$$\texttt{("Migalastat", sem:treats, "Fabry Disease")} \xrightarrow{\text{t: Migalastat db:isa alpha−Galactosidase}} \texttt{("alpha-Galactosidase", sem:treats, "Fabry Disease")}$$

Fig. 2. Rewritten triple using the ontology axiom

Considering the transformation above, the initial SPARQL query could be written as below (Code 2). By posing this new query over SemMedDB, we will get a 'Yes' answer, as the database contains the matching triple. As seen in Code 2, the plausible answer 'Yes' is accompanied by the plausible pattern, generalization, that is involved in the QR. It shows which plausible patterns has lead to this plausible answer.

```
Rewritten SPARQL query:
PREFIX sem:     <https://skr3.nlm.nih.gov/SemMed#>
   ASK
   { "alpha-Galactosidase" sem:treats "Fabry Disease" }
```

```
Plausible Answer:
(Yes, {GEN})
```

Code 2. Rewritten query answering if Migalastat treats Fabry Disease

1 https://skr3.nlm.nih.gov/SemMedDB/index.html

5.2 Example 2: Herceptin treats Prostate Cancer?

In this case study, we are asking another Yes/No question, *"Is Herceptin of potential use in the treatment of prostate cancer?"* (BioASQ challenge, Task 2b), over the SemMedDB. Making use of the existing triples in the database, there is no matching triple unifying the question. Consequently, the answer will be 'No'. The initial SPARQL syntax of the question is as bellow:

Initial SPARQL query: @PREFIX sem: <https://skr3.nlm.nih.gov/SemMed#> ASK { "Herceptin" sem:treats "Prostate cancer" }	Answer: No

Code 3. Initial query answering if Migalastat treats Fabry Disease

Utilizing Disease ontology axioms (DOID:10286) and existing triples in SemMedDB, we know:

```
("Herceptin", sem:treats, "Malignant neoplasms")            (1)
("Malignant neoplasms", sem:occurs_in, "Prostate carcinoma") (2)
("Prostate carcinoma", do:isa, "Prostate cancer")           (3)
```

Fig. 3. Rewritten triple using the ontology axiom

In the triples above, the *treats* predicate (Fig. 3.1) shows a disease (*malignant neoplasms*) that could be treated by *Herceptin*. The *occurs_in* relationship (Fig. 3.2) characterizes the *"order"* of occurrence of two phenomena, in this case two phases of a disease: *malignant neoplasms* and *prostate carcinoma*. The *is_a* relationship (Fig. 3.3) represents a hierarchical relationship between two diseases, *prostate carcinoma* and *prostate cancer*. Using the semantics above, QR algorithm exploits specialization pattern and a fortiori pattern, to transform the initial query to the expanded query below:

Rewritten SPARQL query: PREFIX do: < http://disease-ontology.org/term#> PREFIX sem: <https://skr3.nlm.nih.gov/SemMed#> ASK { "Herceptin" sem:treats " Malignant neoplasms". "Malignant neoplasms", sem:occurs_in, "Prostate carcinoma". "Prostate carcinoma", do:isa, "Prostate cancer"}	Plausible Answer: (Yes, {SPEC, AFORT})

Code 4. Rewritten query answering if Migalastat treats Fabry Disease

By posing the new query over SemMedDB, we will get a plausible positive answer that is inferred via both specialization and a fortiori patterns. The inference above means: *Herceptin* could *treat prostate cancer*, as *Herceptin* could *treat malignant neoplasms* that is an *earlier* phase (ordered relationship) of *prostate carcinoma*, which *is a type of* (hierarchical relationship) *prostate cancer*. In other words, *Herceptin* could *plausibly treat prostate cancer* as it is administered to some prior phases of the disease.

6 Discussion

Medical experts can make plausible conclusions as they know semantics and understand the relationships between the concepts. They also utilize plausible patterns to draw tentative associations that are currently missing. So, case studies above and similar inferences might seem straightforward to the practitioners making clinical decisions. However, examples above showed even with a large database like SemMedDB (with over 89 million predicates from all of PubMed citations), conventional clinical reasoning engines cannot guarantee an answer. This drawback is due to the lack of support for handling uncertainty resulting from missing associations between data attributes.

Despite the strict logical formalisms in traditional reasonings, case studies above showed PR, as a weak form of inference, can infer new knowledge by exploiting semantics between data. In the first case study, the QR algorithm replaced the subject of the triple in the question by its parent in the hierarchy to conduct a generalization pattern, with this logic that *"when something is true about a set of objects, it might be true for any subset of it"*. In the second case study, the plausible answer is the result of combination of specialization and a fortiori. The rationale behind specialization pattern contrasts with generalization: *"when something is true about a class/entity, it might be true about its super class (parent) as well"*. However, a fortiori pattern, as an ordered-based pattern, conducts the query transformation based on the belief that *"if something is true about a stage of a phenomena, then it might be true for any stages after that"*.

The efficiency of SeDan depends on (i) the collaboration between the plausible patterns, like how human thought process works, and (ii) the ontological constructs that conduct the plausible patterns. A well-designed QR algorithm addresses the first issue. However, the enrichment, validity and variety of semantic annotations and relationships of the ontologies that QR algorithm uses to rewrite a query would be a challenge.

7 Conclusions and Future Work

Healthcare is a knowledge-intensive domain, which typically suffers from incomplete data. To extend the knowledge coverage of medical knowledge-bases and enhance patient health outcomes, machines are required to (i) capture and understand semantics and relationships between data attributes, and (ii) leverage those semantics to extract potential relationships existing in healthcare data (EHR, PHR, EMR, etc.)

To this aim, we introduced the SeDan framework that supports automated clinical decision support via semantics-based data analytics. The plausible reasoner integrates plausible patterns with fine-grained biomedical ontologies. The reasoner infers plausible solution(s) by transforming an initial query with no answer to an augmented union conjunctive of queries. This flexible mechanism extends SPARQL queries with the hope to overcome the existing gap in the medical knowledge bases.

From the theory development perspective, Sedan implements plausible patterns using OWL constructs and SPARQL to provide principled means to represent and reason with incompleteness. Our proposed plausible extension to OWL provides full-fledge support to implement plausible patterns within the SW. From an applied perspective, due to the flexible graph-based data format capable of incorporating new relations, support for rich semantics and automatic DL-based reasoning, the SW technologies provide excellent support for PR to draw semantic inferences from large datasets.

Future work consists of studying the efficiency of SeDan in answering the questions from the latest BioAsk task using Disease ontology, DrugBank and Semantic MEDLINE databases. This is the first step to verify the competency of SeDan in answering to real-world medical questions before using it in a real clinical environment.

Improving the performance of the QR algorithm (i.e., *reduction* phase) to guarantee computational completeness and decidability of the reasoner will be the next step.

Acknowledgment: This research is supported by a NSERC Discovery Grant.

References

[1] H. Mohammadhassanzadeh, W. Van Woensel, S. R. Abidi, and S. S. R. Abidi, “Semantics-based plausible reasoning to extend the knowledge coverage of medical knowledge bases for improved clinical decision support,” *BioData Min.*, vol. 10, no. 1, p. 7, 2017.

[2] A. Holzinger and I. Jurisica, “Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions,” *Interact. Knowl. Discov. data Min. Biomed. informatics*, pp. 1–18, 2014.

[3] R. Almeida, U. Kaymak, and J. Sousa, “A new approach to dealing with missing values in data-driven fuzzy modeling,” *Fuzzy Syst. (FUZZ), 2010 IEEE Int. Conf.*, 2010.

[4] D. Walton, C. W. Tindale, and T. F. Gordon, “Applying Recent Argumentation Methods to Some Ancient Examples of Plausible Reasoning,” *Argumentation*, vol. 28, no. 1, pp. 85–119, Nov. 2014.

[5] A. Collins and R. Michalski, “The logic of plausible reasoning: A core theory,” *Cogn. Sci.*, vol. 13, pp. 1–49, 1989.

[6] D. Ausín, D. López-de-Ipina, and F. Castanedo, “A probabilistic OWL reasoner for intelligent environments,” in *Proceedings of the 10th International Conference on Uncertainty Reasoning for the Semantic Web*, 2014, vol. 1259, pp. 1–12.

[7] N. Al Haider, S. Abidi, W. Van Woensel, and S. S. Abidi, “Integrating existing large scale medical laboratory data into the semantic web framework,” in *Big Data (Big Data), 2014 IEEE International Conference on*, 2014.

[8] T. Berners-Lee, M. Fischetti, and M. F. By-Dertouzos, *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. HarperInformation, 2000.

[9] G. Stoilos, T. Venetis, and G. Stamou, “A Fuzzy Extension to the OWL 2 RL Ontology Language,” *Comput. J.*, vol. 58, no. 11, pp. 2956–2971, 2015.

[10] P. Han, W. Klein, and N. Arora, “Varieties of uncertainty in health care a conceptual taxonomy,” *Med. Decis. Mak.*, 2011.

[11] M. Dimartino, A. Calì, and A. Poulovassilis, “Query Rewriting under Linear EL Knowledge Bases,” in *International Conference on Web Reasoning and Rule Systems*, 2016, pp. 61–67.

[12] M. Virvou and K. Kabassi, “Adapting the human plausible reasoning theory to a graphical user interface,” *IEEE Trans. Syst. Man, Cybern. A Syst. Humans*, vol. 34, no. 4, pp. 546–563, 2004.

[13] H. Pérez-Urbina and E. Rodrıguez-Dıaz, “Evaluation of query rewriting approaches for OWL 2,” *Proc. of SSWS+ HPCSW*, 2012.

[14] S. Grimm and B. Motik, “Closed World Reasoning in the Semantic Web through Epistemic Operators,” *OWLED*, 2005.

[15] M. Bienvenu, “Ontology-Mediated Query Answering: Harnessing Knowledge to Get More From Data,” in *nternational Joint Conference on Artificial Intelligence*, 2016.

[16] T. Lukasiewicz and U. Straccia, “Managing uncertainty and vagueness in description logics for the semantic web,” *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 6, no. 4, 2008.

[17] H. Pérez-Urbina, B. Motik, and I. Horrocks, “A Comparison of Query Rewriting Techniques for DL-lite,” *Descr. Logics*, 2009.

[18] O. Mohammed, “Semantic web system for differential diagnosis recommendations,” Lakehead University, 2012.

[19] “The BioASQ Challenge.” [Online]. Available: http://www.bioasq.org/.

[20] V. Law, C. Knox, Y. Djoumbou, and T. Jewison, “DrugBank 4.0: shedding new light on drug metabolism,” *Nucleic Acids Res.*, vol. 42, no. D1, pp. D1091–D1097, 2013.

[21] W. Kibbe, C. Arze, V. Felix, and E. Mitraka, “Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data,” *Nucleic Acids Res.*, vol. 43, no. D1, pp. D1071–D1078, 2014.

[22] T. Rindflesch, H. Kilicoglu, and M. Fiszman, “Semantic MEDLINE: An advanced information management application for biomedicine,” *Inf. Serv. Use*, vol. 31, no. 1–2, pp. 15–21, 2011.