



HAL
open science

Probabilistic Semantics

Salvatore Flavio Pileggi

► **To cite this version:**

Salvatore Flavio Pileggi. Probabilistic Semantics. *Procedia Computer Science*, 2016, 80, pp.1834-1845.
10.1016/j.procs.2016.05.472 . hal-01327353

HAL Id: hal-01327353

<https://hal.sorbonne-universite.fr/hal-01327353>

Submitted on 6 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Probabilistic Semantics

Salvatore F. Pileggi

INRIA & UPMC-LIP6, Paris, France
flavio.pileggi@lip6.fr

Abstract

This paper proposes a concise overview of Probabilistic Semantics according to a technology-oriented perspective. Indeed, while the progressive consolidation of Semantic Technology in a wide context and on a large scale is going to be a fact, the non-deterministic character of many problems and environments suggests the rise of additional researches around semantics to integrate the mainstream. Probabilistic extensions and their implications to the current semantic ecosystems are discussed in this paper with an implicit focus on the Web and its evolution. The critical literature review undertaken shows valuable theoretical works, effective applications, evidences of an increasing research interest as the response to real problems, as well as largely unexplored research areas.

Keywords: Semantic Technology, Probabilistic Computing, Knowledge Representation

1 Introduction

During the past years Semantic Technology has been object of an intense research activity that has pushed the materialization of theoretical models in concrete standard technologies to address the enhanced semantic support aimed in the third version of the Web [1]. That support should help machines to "understand" the information on the Web so that they can provide richer discovery, data integration, navigation, and automation of tasks. The need for semantics is an evolving concept, as the Web itself is in a constant evolution in terms of scale and purpose. The recent rise and affirmation of Big Data [2] is a clear example.

On the other hand, the general relevance of uncertainties and the consequent importance of probabilistic models to integrate deterministic approaches are well known, consolidated and universally accepted concepts in Computer Science. The uncertainty can have a different meaning depending on the context, such as unknown, partially known, probable. Semantics have a general purpose. They are definitely not an exception as often they deal with problems with uncertainties. Uncertainties can be implicit in the model itself (e.g. in predictive models [3]), as well as inducted by processes and methods (e.g. in Information Retrieval [4]). Semantics in Computer Science are a challenge themselves [5]. But it is a fact that uncertainties are everywhere (or almost) in computation.

According to a first naive analysis, semantic and uncertain knowledge seem to be somehow mutually exclusive: the former aims at formal specifications to *avoid ambiguities* in the knowledge representation and processing; and the latter reflects the *non-determinism*. This contradiction is evidently only apparent as the non-determinism, if formally modelled, is part of the knowledge. Therefore, a particular line of work focusing on the combination of probabilistic models with description logic (e.g.[6]), known as *Probabilistic Semantic*, has progressively gained importance. That convergence is indeed suggested by practical needs which are also well supported by philosophical considerations: the awareness of the unknown is part of the knowledge.

Several interesting works on probabilistic semantics are currently well known. However, there is also a fundamental lack of integrated discussion on the topic, as well as there are largely unexplored research areas. It seems that, even though recognizing the importance of uncertainties in semantic computing, most researchers are working on the main branch, focusing on the capability of the models and on their computation more than in extensions. This is absolutely understandable considering the current technological trends aimed at the consolidation of the technology. On the other hand, probabilistic extensions would probably extend the scope of *Semantic Technology* and its reliability on a large scale also in presence of non-determinism.

This paper proposes a concise overview of relevant works in the area of probabilistic semantics with a conceptual focus on web-technology. An ideal, exhaustive and rigorous approach would include an overview of the theoretical background and its materialization in concrete technologies. Furthermore, uncertainty is a very complex concept which can assume different meanings depending on the context. Just as an example we can distinguish among probabilistic/possibilistic uncertainty, qualitative/quantitative non-determinism, ambiguity, vagueness, lack, entropy, similarity, approximation and application-specific concepts that can be modelled as an uncertainty in a given context (e.g. trusting). This paper has a more reduced scope and focuses on the "front-end" technology and its relation with the theoretical background in a context of applied semantics, providing a snapshot of the current status of the art. Semantics' key concepts are briefly explained as part of a unique model which assumes semantics on a large-scale in a Web context (*Web Semantics*). This model emphasizes above all on the semantic representation and processing of the information and doesn't deal explicitly with other applications of semantics (e.g. Semantic Web Services [7] or agents [8]). The probabilistic meaning of semantics is discussed both with an overview of the major works in literature and examples where needed and possible.

2 From Semantics to Probabilistic Semantics

Semantic technologies are progressively emerging as a response to new evolving requirements from the Information Society [9]. The consolidation of the key concepts, as well as their progressive application in real systems, has produced a new generation of semantic ecosystems where the information is modeled and processed according to standardized semantic models.

A simplified *Semantic Ecosystem* is represented by the concept map in fig. 1. In this model, the raw, atomic and unorganized piece of information is the *data*; heterogeneous data can be composed to represent a most consistent *information* which is, therefore, a simple aggregation or composition of atomic data; the processing of data and/or information produces *knowledge*, which differs from information as it cannot be obtained by obvious operations or transformations (such as merging or aggregation) but only through a proper computing or inference in a given context. The border between information and knowledge often result a bit fuzzy in practice,

where it could be just theoretical or, more properly, context-specific. All these three progressive concepts are represented according to a formal semantic model (*Semantic Representation* [10]), even though the representation of the knowledge cannot be always formalized due to its intrinsic complexity (for instance when the understanding of results requires a domain specific interpretation or contextual meaning). The very first goal of the semantic representation is the support of enhanced models for interoperability to support machine-to-machine interaction. Indeed, *Semantic Interoperability* [11] is the ability of computer systems to exchange data with unambiguous, shared meaning. It wants to overcome common interoperability models, based on simple syntactic equivalences (*Basic Interoperability*) as simple exchange of messages between systems or on extensions that provide a common structural definition for message fields (*Functional Interoperability*). Functional Interoperability evidently extends the basic model. Semantics enforce further extensions enabling links among formalized data structures.

The semantic representation commonly refers to the process of annotating resources with semantic metadata (*Semantic Annotation and Indexing* [12]) and to the use of rich data models called *Ontology* (e.g.[13]), understood as the formal specification of a conceptualization. In practice, an Ontology represents, according to a standard language, the concepts of a given domain, as well as the relations among those concepts. The last generation model for Ontology [14] is based on the capability of defining concepts and properties explicitly and on the inference of further concepts (like additional layers of knowledge) from the explicit information. *Query* on semantic rich data models requires additional languages (e.g.[15]) that make the knowledge accessible according to a smart and direct philosophy. Semantics are so important to assure the effectiveness and the scalability of the next generation information systems that the evolution of the whole web model is currently driven by semantic technologies. They should lead from the current consolidated understanding of the Web (Web 2.0 or *Social Web* [16]) to a real semantic-oriented model (Web 3.0 or *Semantic Web* [1]), along multiple understandings of intermediate models (Web 2.x [17]) such as Internet of Services or Semantic Social Web, as well as domain specific semantic webs (e.g. *Web of Data* [18] and *Semantic Sensor Web* [19]).

Semantic Computing aims at the achievement of some knowledge by processing or analysing semantic data or information. The most relevant model for computing semantic data is the *Semantic Reasoning*, which provides capabilities to infer logical consequences from a set of asserted facts or axioms with a clear similitude with human reasoning (e.g. in [20]). Further paradigms may be defined and are progressively emerging (e.g. *Semantic Learning*) as the consequence of the increasing popularity of the *Semantic Technology*.

Once a simplified view of a semantic ecosystem has been proposed, the key question is: *can it be a probabilistic ecosystem?* The answer is positive and, as explained in the next sessions, it can be potentially a very consistent probabilistic ecosystem.

3 Probabilistic Semantics

As evident is fig.1, the real central concept in a semantic ecosystem is the semantic representation, which is the simplest concept in the ecosystem but also the most ambiguous. The centrality of a given concept has a context-specific understanding. In this case, the expressiveness and the features of the semantic representation determine, directly or indirectly, the capabilities of the whole ecosystem.

Assuming this key concept in its probabilistic meaning (*Probabilistic Semantic Representation*) redefines completely the scope of a semantic ecosystem as a *Probabilistic Semantic Ecosystem*, which is able to model and represent uncertainties as part of the knowledge according to a probabilistic approach to the non-determinism. Following this approach, the probability

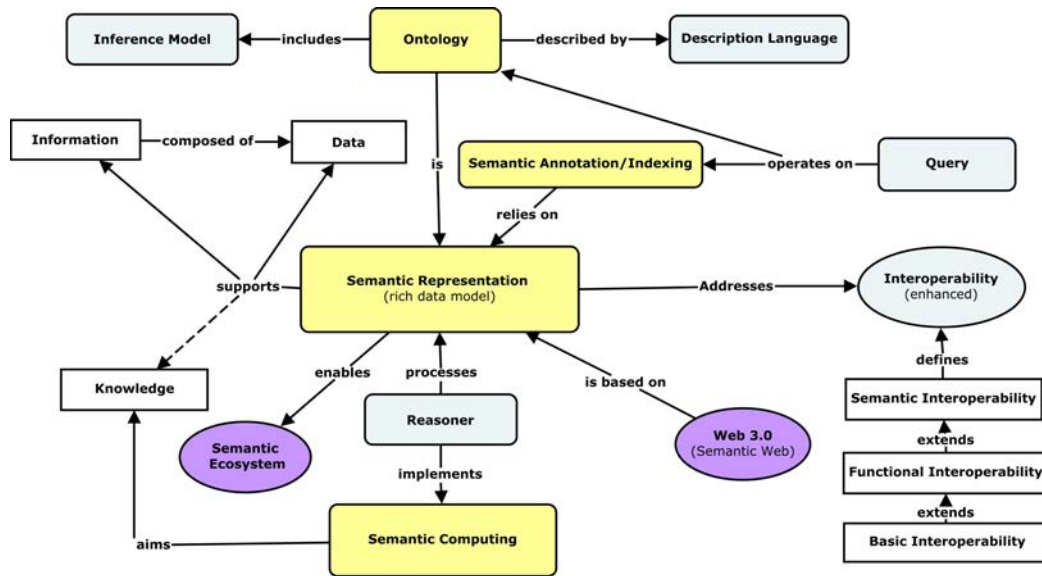


Figure 1: A simplified concept map representing a Semantic Ecosystem.

is considered as an additional dimension for semantics. Therefore, probabilistic semantics are implicitly understood as an extension of common semantics to integrate further capabilities in terms of representation and computing.

Ad-hoc solutions are very interesting but they are often hard to apply as a generic solution. As well as approaches different from the extension/integration that do not guarantee compliance with the existing technology seem somehow hard to be proposed at this stage of the web evolution. Indeed the mainstream philosophy is pushing towards the standardization to convert local semantics into *Web Semantics*, making hard the affirmation of completely new approaches.

However, the last generation of *Semantic Technology* is extremely flexible and intrinsically allows variants and evolutions, providing an interesting research contest. As one of many possible examples, OWL 2 EL [21] is used by large-scale ontologies from practical applications to improve the reasoning performance in common problems such as classification and instance checking.

3.1 Probabilistic Semantic Representation: the theoretical background

The models for probabilistic semantic representation are, somehow, the theoretical background around technologies as they define the logic to represent the knowledge according to a probabilistic approach. Knowledge representation languages invariably reflect a trade-off between expressiveness and tractability. The model of uncertainty adds a further complexity.

There are different approaches from different areas, such as probabilistic description logics [22, 23] and statistical relational languages [24]. While the former have mainly contributed at a theoretical level, statistical relational languages have been applied in a number of practical applications, such as data integration through ontology matching [25]. Fuzzy logic [26] has also been extensively explored [27] producing a number of significant results in the field of semantic technologies (e.g. [28, 29, 30]). However, the probability theory seems more suitable to make

prediction about events from a state of partial knowledge.

Evidences suggest that the compromise between expressivity and tractability chosen by description logics is a particularly successful one. But, as a variant of first-order logic, description logics have severe limitations to express uncertainties. The integration of original models with probabilistic extensions seems to be the most effective approach. It is the case of P-CLASSIC [22] which extends the original model by establishing *Bayesian networks* [31] to represent uncertainty. The resulting inference procedure (*probabilistic subsumption* [22]) is the natural consequence of combining representation logic with Bayesian networks: it computes the probability that a random individual in a class A is also a member of class B. The extension of classical description logic for the management of vague/imprecise knowledge is deeply discussed in [23]. Some of the logics currently adopted in key semantic languages have been extended to address probabilistic knowledge: for example *SHOQ(D)* is extended by *P-SHOQ(D)* [32].

3.2 Probabilistic Semantic Annotation

Semantic annotation and indexing are processes based on the capability to extend common content with semantic meta-data to relate pieces of knowledge each other. As a case study, this section considers the probabilistic extensions of the well know RDF [33]. The *Resource Description Framework* (RDF) is a paradigm for knowledge representation that has rapidly emerged as a data model for representing information about World Wide Web resources. Despite its enormous popularity, RDF is not able to express probabilistic knowledge.

There are different theoretical approaches, such as [34] where the author briefly discusses a method for representing probabilistic relations in RDF by using Bayesian networks. Common query mechanisms can be adapted to work efficiently on probabilistic RDF data graphs [35].

A more comprehensive contribution is represented by pRDF [36]. A pRDF instance (fig. 3) extends common RDF triples (fig. 2) with a further dimension which allows the unconditioned probability distributions over a set of possible values of an RDF triple. As an example of the strong limitations of RDF and of the usefulness of probabilistic extensions, we can consider a common phrase: *the complications of the flu are, among others, acute bronchitis (70%) and pneumonia (15%)*. The human understanding of this affirmation easily detects an uncertain knowledge related to the implicitly defined concept of *other*. This uncertainty integrates the explicit probabilistic knowledge expressed by the statement. There is, indeed, an intuitive conclusion deduced by the fact that the sum of the probabilities is lower than 100%. The concept of *other* is hard to model in the context of semantic environments due to its intrinsic ambiguity. Furthermore, its potentially extensive use would introduce significant uncertainties at an application level. Standard RDF only allows an incomplete and, therefore, wrong knowledge representation (fig. 2) which is limited to the deterministic knowledge. Indeed, the RDF triple (*Flu, hasComplication, AcuteBronchitis,Pneumonia*) is commonly understood like *Flu has two complications, acute bronchitis and pneumonia*, that is far away from the original semantic. The quadruple (*Flu, hasComplication, AcuteBronchitis,Pneumonia,.7, .15*) can be intuitively read as: *whenever the Flu condition has complications, there is a 70% probability that the complication is AcuteBronchitis and a 15% probability that the complication is Pneumonia* (fig. 3). That is still somehow vague due to an implicit lack of knowledge that can generate ambiguities in the interpretation. But it is much closer to the original semantic.

A more detailed analysis of pRDF suggests also the definition of a simplified model. Indeed, the previous pRDF statement is really composed of two different atomic probabilistic statements, (*Flu, hasComplication, AcuteBronchitis, .7*) and (*Flu, hasComplication, Pneumonia, .15*), merged in an unique piece of knowledge or *assertion*. A simplified approach associates an

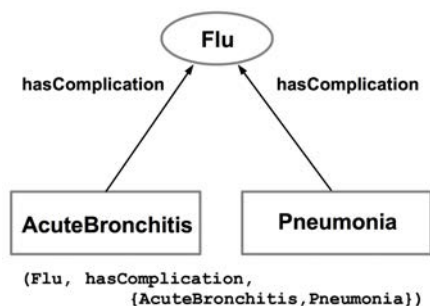


Figure 2: RDF.

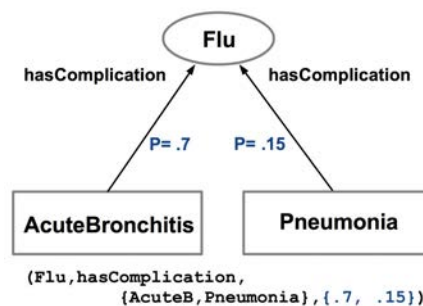


Figure 3: pRDF.

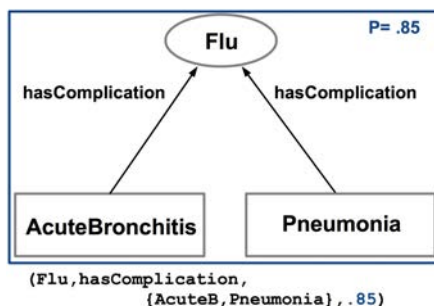


Figure 4: Simplified pRDF.

unique probability to the whole assertion (fig. 4). From this brief analysis two different models for probabilistic extension of RDF-like annotation can be defined as follows:

Def. A *fully-probabilistic semantic annotation* assumes that all the concepts of the n-tuple representing the assertion are related each others with a given probability (e.g. pRDF).

Def. A *simplified probabilistic semantic annotation* associates an unique probability to the whole assertion, even composed of multiple tuples (e.g. the simplified pRDF model previously described).

The two models are equivalent only if assertions are atomic statements. In general, the former is more expressive than the latter. Different approaches to model and compute the knowledge could privilege the full or the simplified model.

3.3 Probabilistic Ontology

Ontology Web Language (OWL) [37] consists of three increasingly-expressive variants known as OWL Lite (from *SHIF(D)* logic [38]), OWL DL (from *SHOIN(D)* logic [39]), and OWL Full. The first only allows simple statements, the second implements a consistent inference model and the last is a free mixing of OWL with RDF that does not guarantee completeness and decidability [40]. OWL 2 polishes OWL and is backward compatible with it. Without altering expressiveness, semantics and complexity, OWL 2 makes some patterns easier to write allowing

more efficient processing in reasoners.

Assuming the most expressive decidable version of OWL (OWL DL basically), an Ontology is composed of a set C of *concepts* (or *classes*), a set P of *properties* and a set I of *individuals*. Inside an Ontology whichever statement is an *assertion*. Assertions establish hierarchies among concepts, define individuals, associate individuals with classes, as well as they set the properties of individuals, establishing relations among concepts. Furthermore, a set R of inference rules allows the inference of the knowledge; therefore, individuals may be associated through those rules to concepts without explicit assertion. According to an individual-centric perspective [4], at a theoretical level, two different kind of probabilistic models for OWL Ontologies may be observed. They are defined as follows:

Def. An Ontology has a *non-deterministic population* if assertions are non-deterministic, meaning they are associated with a given probability (fig. 5).

Def. An Ontology has a *non-deterministic schema* if each inference rule is associated with a given probability (fig. 6).

The two models are not equivalent and, a priori, the first is more expressive than the second. As discussed before for RDF, different models of computation may privilege one model or another. Also a *fully-probabilistic* Ontology can be defined as the combination of the two models. The co-existence of a non-deterministic schema and a non-deterministic population defines a sophisticated inference model that combines both kinds of uncertainty.

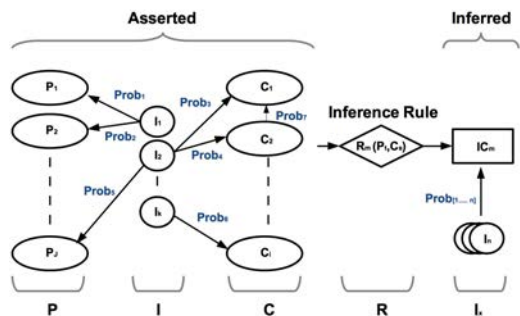


Figure 5: Probabilistic population.

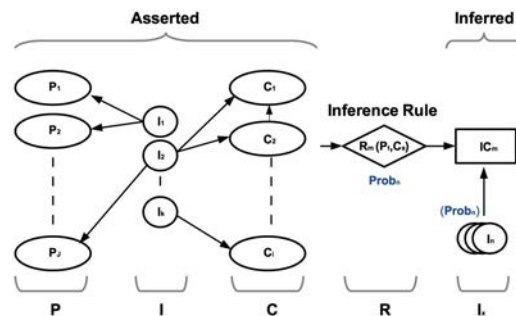


Figure 6: Probabilistic schema.

3.3.1 Probabilistic extensions for OWL

As a *de facto* standard, OWL is definitely the most used language for Ontology development. Several probabilistic extensions for OWL have been proposed during the last years to support uncertain ontology representation, reasoning and mapping. The most common approaches provide inference models based on the Bayesian paradigm (generic conditional dependency between cause/assertion and effect/inferred) or on a joint (independent) probability model. All them propose a convergence between the original logic inference model and statistical inference.

The work proposed in [41] incorporates Bayesian networks to model uncertain knowledge. Bayesian networks are also used in [42] to provide enhanced probabilistic ontology mapping. In

[43] the authors proposes PR-OWL as the response to the ubiquity of uncertainty: a Bayesian framework for probabilistic ontologies and plausible reasoning services determine a deterministic view of a probabilistic world, overcoming in fact one of the major weakness of the current Semantic Technology [43]. While PR-OWL provides OWL constructs for representing *Multi-Entity Bayesian Network* (MEBN) [44] theories, it presents several limitations extensively discussed in [45], including the lack of a proper integration of the MEBN formalism with the logical basis of OWL on the meta level. A formal mapping between random variables defined in PR-OWL and concepts defined in OWL is addressed by PR-OWL 2 [46].

3.3.2 Prolog-like technology

Prolog-like technology [47] is based on *Horn clauses* and has significantly contributed to education and research for a long time. In a second moment, it has been discharged by industry and, consequently, by many education programs. But it has never been dropped by researchers which follow using declarative programming and Prolog-like solutions in many domains and applications (e.g. Artificial Intelligence).

Even though there are examples that seem to prove the contrary (e.g. [48]), prolog-like technology is not considered a priori a Semantic Technology. It has already extensively contributed to the development of the last generation of semantic technologies (e.g. SWRL [49]), as well as it will probably lead also some further evolution.

Languages such as SWRL [49] combines OWL and subsets of *Datalog* [50]. Datalog is syntactically a subset of Prolog often used as a query language for deductive databases. Datalog has recently found new applications in data integration, information extraction, networking, program analysis, security, and cloud computing. SWRL is built on OWL DL and shares its formal semantics. It is more expressive than OWL DL alone but this additional expressivity comes at the expense of decidability. Common reasoners (e.g. Pellet [51]) support SWRL built-ins, enabling the use of SWRL in fact.

Probabilistic extensions of Datalog (*pDatalog*) have been proposed several years ago [52, 53]. They can intuitively contribute to the model of probabilistic semantics, implicitly or explicitly. For example, a simple program in pDatalog ($0.5male(x) \leftarrow person(x)$) may express that a person is with probability of 50% male, or that Tom is a person with probability of 80% ($0.8person(Tom)$). The work reported in [54] adds probabilities and rules to OWL Lite by using Probabilistic Datalog, as well as probabilistic extension of SWRL have been recently proposed [55, 28, 29].

3.4 Probabilistic Semantic Interoperability

The ontological representation (according to the OWL model) can currently be considered the highest layer of sufficient maturity in the theoretical hierarchical structure of the Semantic Web. Once the Ontology layer and the back-end architecture is extended to accommodate the representation of uncertain knowledge, high-level concepts based on the semantic infrastructure, such as Semantic Interoperability and Computation, change consequently. As far as the author knows, *Probabilistic Semantic Interoperability*, an intuitive consequence of the semantic representation, is not explicitly addressed in literature.

The commonly accepted concept of Semantic Interoperability (SI) is based on the indirect relation established by semantically linking contents to formally defined concepts as part of shared *vocabularies* (fig. 7). A probabilistic understanding of SI (fig. 8), even though maintaining the same philosophy and structure of the original model, adopts probabilistic semantic links. Therefore, also the resulting relation of interoperability among contents is probabilistic.

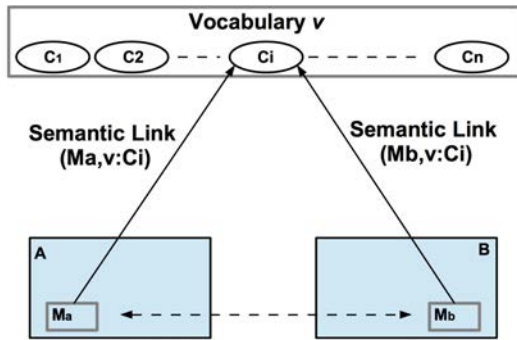


Figure 7: Semantic Interoperability (SI).

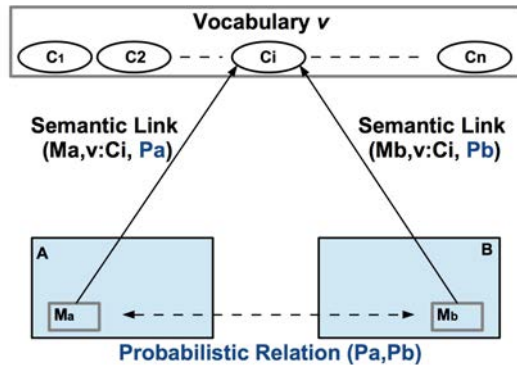


Figure 8: Probabilistic SI.

4 An intuitive comparison

An intuitive analysis is proposed in table 1. The different classes of probabilistic extensions for semantic technologies are informally compared according to simple qualitative metrics and without using formal methods.

Representation logic has provided a solid theoretical background which has not yet been converted into technology, proposing a huge still unexplored potential in terms of application. Statistical relational languages propose the opposite situation as they have been used in practice without providing a real semantic support. They are defining, in fact, a kind of gap between the semantic representation and its probabilistic meaning. Future efforts could be aimed at covering that gap. Finally, solutions based on fuzzy logic are definitely promising but need experimentation in practice to archive some maturity at a technological level.

Focusing on applied semantics, the extensions of the current semantic technology seem the most serious candidate to approach the real world. Due to its simplicity, the technology based on RDF looks the easiest to use; on the other hand, the potential in terms of further evolution seems reduced. The most promising solutions extend OWL. Those models propose enhanced expressiveness with respect to RDF. Their use is not always intuitive but the potential in terms of evolution seems really consistent. The considerations about OWL can be easily extended to the Prolog-like technology, which proposes higher expressiveness at the price of decidability.

5 Conclusions

The fundamental question addressed in the paper (*can semantic ecosystems model and process also probabilistic knowledge?*) has a clearly positive answer. Indeed, the critical literature review undertaken shows valuable theoretical works, effective applications, as well as the evidences of an increasing research interest as the response to real problems.

The critical analysis of the state of the art has also put in evidence largely unexplored research areas, as well as a theoretical gray zone that should provide a kind of integration path to the main Semantic Technology for extensions and variants.

Probabilistic Semantic Representation				
	Theory	Application		
Description Logic	High	Low		
Statistical Relational Lang.	Low	High		
Fuzzy Logic	Medium	Low		
Probabilistic Semantic Technology				
	Expressiveness	Decidable	Usability	Evolution
Probabilistic RDF	Medium\Low	yes	High	Low
Simplified Prob. RDF	Low	yes	High	Low
OWL-like extensions	High\Medium	yes [†]	Medium	High
Probabilistic Population	High\Medium	yes [†]	Medium	High
Probabilistic Schema	Medium	yes [†]	Medium	High
Prolog-like extensions	High	no	Medium	High

[†] Assuming OWL-DL

Table 1: Qualitative comparison.

Acknowledgments

This research is supported in part by European FP7 project 609551 SyncFree (2013–2016).

References

- [1] Tim Berners-Lee, James Hendler, Ora Lassila, et al. The semantic web. *Scientific american*, 284(5):28–37, 2001.
- [2] Andrew McAfee, Erik Brynjolfsson, Thomas H Davenport, DJ Patil, and Dominic Barton. Big data. *The management revolution. Harvard Bus Rev*, 90(10):61–67, 2012.
- [3] C Soize. A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive models in structural dynamics. *Journal of sound and vibration*, 288(3):623–652, 2005.
- [4] Salvatore F Pileggi. An individual-centric probabilistic extension for owl: Modelling the uncertainty. *Procedia Computer Science*, 51:1742–1751, 2015.
- [5] Nigel Shadbolt, Wendy Hall, and Tim Berners-Lee. The semantic web revisited. *Intelligent Systems, IEEE*, 21(3):96–101, 2006.
- [6] Ian Horrocks et al. Daml+oil: A description logic for the semantic web. *IEEE Data Eng. Bull.*, 25(1):4–9, 2002.
- [7] Sheila A McIlraith, Tran Cao Son, and Honglei Zeng. Semantic web services. *IEEE intelligent systems*, 16(2):46–53, 2001.
- [8] James Hendler. Agents and the semantic web. *IEEE Intelligent systems*, (2):30–37, 2001.
- [9] Manuel Castells. *The rise of the network society: The information age: Economy, society, and culture*, volume 1. John Wiley & Sons, 2011.
- [10] Thomas L Griffiths, Mark Steyvers, and Joshua B Tenenbaum. Topics in semantic representation. *Psychological review*, 114(2):211, 2007.

- [11] Aris M Ouksel and Amit Sheth. Semantic interoperability in global information systems. *ACM Sigmod Record*, 28(1):5–12, 1999.
- [12] Atanas Kiryakov, Borislav Popov, Ivan Terziev, Dimitar Manov, and Damyan Ognyanoff. Semantic annotation, indexing, and retrieval. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2(1):49–79, 2004.
- [13] Xiao Hang Wang and et al. Ontology based context modeling and reasoning using owl. In *Pervasive Computing and Communications Workshops, 2004*.
- [14] Deborah L McGuinness, Frank Van Harmelen, et al. Owl web ontology language overview. *W3C recommendation*, 10(10):2004, 2004.
- [15] Jorge Pérez, Marcelo Arenas, and Claudio Gutierrez. Semantics and complexity of sparql. In *International semantic web conference*, volume 4273, pages 30–43. Springer, 2006.
- [16] Tim O'reilly. What is web 2.0: Design patterns and business models for the next generation of software. *Communications & strategies*, (1):17, 2007.
- [17] Salvatore F. Pileggi, Carlos Fernandez-Llatas, and Vicente Traver. When the social meets the semantic: Social semantic web or web 2.5. *Future Internet*, 4(3):852–864, 2012.
- [18] Christian Bizer. The emerging web of linked data. *Intelligent Systems, IEEE*, 24(5):87–92, 2009.
- [19] Salvatore F Pileggi, Carlos E Palau, and Manuel Esteve. Building semantic sensor web: Knowledge and interoperability. In *SSW*, pages 15–22, 2010.
- [20] Nicola Henze, Peter Dolog, and Wolfgang Nejdl. Reasoning and ontologies for personalized e-learning in the semantic web. *Journal of Educational Technology & Society*, 7(4):82–97, 2004.
- [21] F. Baader, S. Brandt, and C. Lutz. Pushing the envelope. In *19th International Joint Conference on Artificial Intelligence (IJCAI 2005)*, Morgan-Kaufmann Publishers, Edinburgh, UK.
- [22] Daphne Koller, Alon Levy, and Avi Pfeffer. P-classic: a tractable probabilistic description logic. *AAAI/IAAI*, 1997:390–397, 1997.
- [23] Thomas Lukasiewicz and Umberto Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):291–308, 2008.
- [24] Lise Getoor and Lilyana Mihalkova. Learning statistical models from relational data. In *2011 ACM SIGMOD International Conference on Management of data*.
- [25] Mathias Niepert, Christian Meilicke, and Heiner Stuckenschmidt. A probabilistic-logical framework for ontology matching. In *AAAI*, 2010.
- [26] George Klir and Bo Yuan. *Fuzzy sets and fuzzy logic*, volume 4. Prentice Hall New Jersey, 1995.
- [27] Thomas Lukasiewicz and Umberto Straccia. Description logic programs under probabilistic uncertainty and fuzzy vagueness. *International Journal of Approximate Reasoning*, 50(6):837–853, 2009.
- [28] Jeff Z Pan, Giorgos Stoilos, Giorgos Stamou, Vassilis Tzouvaras, and Ian Horrocks. f-swrl: A fuzzy extension of swrl. In *Journal on Data Semantics VI*, pages 28–46. Springer, 2006.
- [29] Xing Wang, Zongmin M Ma, Li Yan, and Xiangfu Meng. Vague-swrl: a fuzzy extension of swrl. In *Web Reasoning and Rule Systems*, pages 232–233. Springer, 2008.
- [30] Mauro Mazzieri, Aldo Franco Dragoni, and UPD Marche. A fuzzy semantics for semantic web languages. In *ISWC-URSW*, pages 12–22, 2005.
- [31] David Heckerman, Dan Geiger, and David M Chickering. Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning*, 20(3):197–243, 1995.
- [32] Rosalba Giugno and Thomas Lukasiewicz. P-shoq (d): A probabilistic extension of shoq (d) for probabilistic ontologies in the semantic web. In *JELIA*, volume 2, pages 86–97. Springer, 2002.
- [33] Frank Manola, Eric Miller, Brian McBride, et al. Rdf primer. *W3C recommendation*, 10(1-107):6, 2004.
- [34] Yoshio Fukushige. Representing probabilistic relations in rdf. In *ISWC-URSW*, pages 106–107.

- Citeseer, 2005.
- [35] Xiang Lian and Lei Chen. Efficient query answering in probabilistic rdf graphs. In *2011 ACM SIGMOD International Conference on Management of data*.
 - [36] Octavian Udrea, VS Subrahmanian, and Zoran Majki. Probabilistic rdf. In *2006 IEEE International Conference on Information Reuse and Integration*.
 - [37] Deborah L McGuinness, Frank Van Harmelen, et al. Owl web ontology language overview. *W3C recommendation*, 10(10):2004, 2004.
 - [38] Ian Horrocks and Peter F Patel-Schneider. Reducing owl entailment to description logic satisfiability. In *The Semantic Web-ISWC 2003*, pages 17–29. Springer, 2003.
 - [39] Ian Horrocks and Ulrike Sattler. A tableau decision procedure for $\mathcal{S}HOIQ$. *Journal of Automated Reasoning*, 39(3):249–276, 2007.
 - [40] Riccardo Rosati. On the decidability and complexity of integrating ontologies and rules. *Web Semantics: Science, Services and Agents on the World Wide Web*, 3(1):61–73, 2005.
 - [41] Zhongli Ding and Yun Peng. A probabilistic extension to ontology language owl. In *37th Annual Hawaii international conference on System Sciences*, 2004.
 - [42] Prasenjit Mitra, Natasha F Noy, and Anuj Rattan Jaiswal. Omen: A probabilistic ontology mapping tool. In *The Semantic Web-ISWC 2005*, pages 537–547. Springer, 2005.
 - [43] Paulo Cesar G da Costa, Kathryn B Laskey, and Kenneth J Laskey. Pr-owl: A bayesian ontology language for the semantic web. In *ISWC-URSW*, pages 23–33, 2005.
 - [44] Kathryn Blackmond Laskey. Mebn: A language for first-order bayesian knowledge bases. *Artificial intelligence*, 172(2):140–178, 2008.
 - [45] Livia Predoiu and Heiner Stuckenschmidt. Probabilistic extensions of semantic web languages-a survey. *The Semantic Web for Knowledge and Data Management: Technologies and Practices. Idea Group Inc*, 2008.
 - [46] Rommel N Carvalho, Kathryn B Laskey, and Paulo CG Costa. Pr-owl 2.0—bridging the gap to owl semantics. In *Uncertainty Reasoning for the Semantic Web II*, pages 1–18. Springer, 2013.
 - [47] William Clocksin and Christopher S Mellish. *Programming in PROLOG*. Springer Science & Business Media, 2003.
 - [48] Francesco M Donini, Maurizio Lenzerini, Daniele Nardi, and Andrea Schaerf. Al-log: Integrating datalog and description logics. *Journal of Intelligent Information Systems*, 10(3):227–252, 1998.
 - [49] Ian Horrocks, Peter F Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosf, Mike Dean, et al. Swrl: A semantic web rule language combining owl and ruleml. *W3C Member submission*, 21:79, 2004.
 - [50] Stefano Ceri, Georg Gottlob, and Letizia Tanca. What you always wanted to know about datalog (and never dared to ask). *Knowledge and Data Engineering, IEEE Transactions on*, 1(1):146–166, 1989.
 - [51] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. Pellet: A practical owl-dl reasoner. *Web Semantics: science, services and agents on the World Wide Web*, 5(2):51–53, 2007.
 - [52] Norbert Fuhr. Probabilistic datalogic logic for powerful retrieval methods. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 282–290. ACM, 1995.
 - [53] Norbert Fuhr. Probabilistic datalog: Implementing logical information retrieval for advanced applications. *Journal of the American Society for Information Science*, 51(2):95–110, 2000.
 - [54] Henrik Nottelmann and Norbert Fuhr. Adding probabilities and rules to owl lite subsets based on probabilistic datalog. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14(01):17–41, 2006.
 - [55] Yu Liu, Shihong Chen, Shuoming Li, and Yunhua Wang. Bayes-swrl: A probabilistic extension of swrl. In *International Conference on Computational Intelligence and Security (CIS)*, 2013.