

Probabilistic Approaches in Ontologies: Joining Semantics and Uncertainty for AUV Persistent Autonomy

F. Maurelli, Z. A. Saigol, G. Papadimitriou, T. Larkworthy, V. De Carolis, D.M. Lane

Ocean Systems Laboratory
School of Engineering & Physical Sciences
Heriot-Watt University
EH14 4AS Edinburgh
Email: f.maurelli@hw.ac.uk

I. INTRODUCTION

In the ocean technology research community, interest is growing in robots able to operate persistently and autonomously in dynamic, partially-unknown and partially-observable environments. As part of the European project FP7 PANDORA, a key focus for our research group goes towards persistent autonomy for Autonomous Underwater Vehicles (AUVs), which means the capability for the robot to robustly operate in those challenging environments. One key enabler for extended run-times is the ability to recognise and adapt behaviour to unexpected conditions. In order to achieve this high-level goal, robots need to have an understanding of the surrounding world and to maintain geometric and semantic descriptions of it, and to update these descriptions according to new information coming from sensors and from reasoning over the knowledge base.

Most of the state-of-the-art techniques for knowledge representation using ontologies do not consider uncertainty at all. This is a major shortcoming, especially in the robotics field, where uncertainty is the norm rather than the exception. Sensor readings are noisy, both on the robot state and on the robot perception of reality. Matching a sensor reading with a specific semantic concept naturally involves taking uncertainty and probability into account. In addition to the noise in the sensor measurements, a partial view of an object might not be enough to associate it with a specific concept without any doubt.

The evolution of the world, in response to specific actions, can also be non-deterministic and actions can transform the world (and the AUV) into different states with different probabilities. Therefore there is a strong need to incorporate a Bayesian framework in the standard OWL web ontology language, in order to represent and propagate probabilities that quantify the extent to which semantic statements in the world model are correct.

The first part of the paper will therefore focus on probabilistic approaches in ontologies, on the way probability is represented and on the inference processes used to augment

the knowledge of the world. The focus will then move to the architectural choices which link the ontology part with the other AUV modules, such as the planning and the navigation systems.

II. RELATED WORK

Work on probabilistic ontologies (POs) builds on a large prior body of work in AI which seeks to combine logical representations of the world with probability. The goal is to benefit from both the expressive power of first-order logic (FOL) to describe the world, and the robustness and inference capabilities facilitated by probabilistic representations. Early approaches simply added probabilities into FOL entities [8], but this proved difficult to reason with. Later, decidable subsets of FOL were used, for example by Heinsohn *et al.* [9], built on Description Logics (the basis for OWL-DL) by allowing statistical properties of the domain to be captured (e.g. 95% of birds can fly). An early project to use Bayesian Networks was P-CLASSIC [10], a probabilistic extension of CLASSIC [1], which was one of the earliest formal ontology systems used in AI. A brief review of this proto-probabilistic-ontology work is given in [6], who note that all of the work cited here is only capable of expressing probabilities for entities (TBox, in DL terminology), rather than instances or objects in a particular world. Reviews of some recent work with this goal are given in Larik *et al.* [12] and Predoiu *et al.* [15]. Although [13] are strongly in favour of an approach based on Bayesian Network, the computational load is a disadvantage for this approach, which gives its best only in case of needed complex inference processes. One key strand of work is by Yun Peng's group at University of Maryland Baltimore County, who have developed the BayesOWL framework [6], [7]. BayesOWL generates a Bayesian Network from an OWL ontology together with conditional probability tables for the classes in the ontology. It only considers the TBox, i.e. the classes in the ontology, and not instances from the ABox. Each class from the TBox (including classes defined by restrictions) is mapped into a node in the graph, and directed links between nodes are generated according to a set of rules. The CPTs for

each node are supplied separately, although the BayesOWL software tool [18] allows these tables to also be encoded in OWL. The final DAG is associated with a joint probability distribution for the input ontology, and allows queries such as finding the degree of similarity between two classes, or finding the most similar class to some queried concept. An important application of such queries is in mapping concepts from one ontology to another, which, as mentioned previously, is a very active area of research in the ontology field.

BayesOWL converts ontologies to Bayesian networks, as opposed to more flexible extensions such as Markov Logic Networks (MLNs) or Multi-Entity Bayesian Networks (MEBNs). An example of using ontologies backed by an MLN is given in [11], who create a tool called *Probabilistic Profile Analysis Ontology*. This tool is effective at analysing web user profiles based not only on similarity of interests, but on which profiles are the most influential within a given interest, although the authors note that current MLN reasoners need significant optimisation to scale to datasets of tens of thousands of users.

PR-OWL [13], [5] is an ongoing research work aimed at extending the OWL web ontology language so it can represent POs. It is a probabilistic extension to OWL that provides a framework for authoring POs, and is based on MEBNs. PR-OWL essentially specifies, in standard OWL, classes that can be used to represent the components of an MEBN: random variables, MFragments, MTheories, and associated probability distributions. There is also a graphical interface called UnBBayes [2] that supports both creating PR-OWL ontologies and performing inference with the corresponding MEBN. A significant drawback with the original PR-OWL system is that, as pointed out in [15], there is no way of explicitly linking the concepts in a domain ontology with the probabilities and DAG structures in a PR-OWL ontology. This issue is addressed in PR-OWL 2 [3], [4], where the PR-OWL-random-variable to OWL-property linkage proposed by [14] is adopted. This link consists of mapping MEBN predicates and functions to OWL properties, and special care is taken to ensure multi-argument predicates can be correctly mapped into OWL, which only allows logical triples. To further increase compatibility with existing OWL ontologies, PR-OWL 2 also allows MEBN random variables to have OWL datatypes.

While both BayesOWL and PR-OWL use a structured network for probabilistic reasoning, PR-OWL is an “upper ontology” so provides a foundation that ontologies must be built on top of. BayesOWL on the other hand can be used with pre-existing ontologies, by simply specifying the conditional membership probabilities for each class in the TBox. However, BayesOWL is not appropriate for domains where probabilities must be associated with instances in the ontology, for example where there is a requirement to associate an uncertainty with sensor readings. This effectively rules out BayesOWL for the Pandora project. Further, the ability to work with existing ontologies is much enhanced in PR-OWL 2. However, the chief advantage of network-based approaches is their ability to represent conditional probabilities for entities and perform

inference using these, as opposed to storing an absolute probability for each entity. For domains where sophisticated inference is not required, and the key purpose of probabilistic information is to represent uncertainty in the world state, it may be clearer to simply add probability annotations. Additionally, the annotation approach makes any novel PO developed much less tied to a particular PO tool.

III. UNCERTAINTY AND WORLD BELIEF

In the past section different ways to model probability into ontologies have been analysed. However, an important issue is to define which kind of uncertainty the robot deals with, and at which level. Only after this analysis, it is possible to decide which structures are best suitable to represent the needs of the Pandora project. This section therefore focuses on different levels of uncertainty, briefly describing different situations where a probabilistic approach can be beneficial to correctly estimate the reality and to quantify the extent to which statements are mirroring the reality.

A. Instances

The first type of uncertainty that is presented deals with the presence of a physical instance in the world. Given the sensor readings, the robot can build a probabilistic belief of the surrounding world. Due to noise, partial views, occlusions, *etc.* the sensor readings might not refer uniquely to a specific object in the world. The robot might believe that a certain object is present in the world, with a certain probability, after processing the sensor data. A typical example for this kind of uncertainty is represented by:

valve(V1) exists with 80% probability

This statement means that the robot has processed the sensor data, arriving to a conclusion that the object in the field of view is a valve, with 80% of probability.

In this case, the uncertainty is given only by the uncertainty in the sensor processing and therefore in the robot belief of the world. The world itself might not be uncertain at all. The valve is either present or not, although its belief is uncertain.

In ontological terms, the uncertainty is related only to the physical instantiation of a physical specific entity in the *ABox* of the ontology, assuming that the ontology does not represent the world, but the belief on the world. In this case, the *TBox* has no uncertainty at all.

B. Relations

Remaining in the field of uncertainty about world belief, this type of uncertainty models the confusion in the belief of a certain relation among physical instances. The robot is not sure whether a certain relation among physical instances is true or not. Through sensor processing, it can formulate a probabilistic belief. An example for this kind of uncertainty is represented by:

panel(Panel1) is connected with pipe(Pipe1) with 74% probability

Like the previous case, the uncertainty is given only by the uncertainty in the sensor processing and therefore in the robot

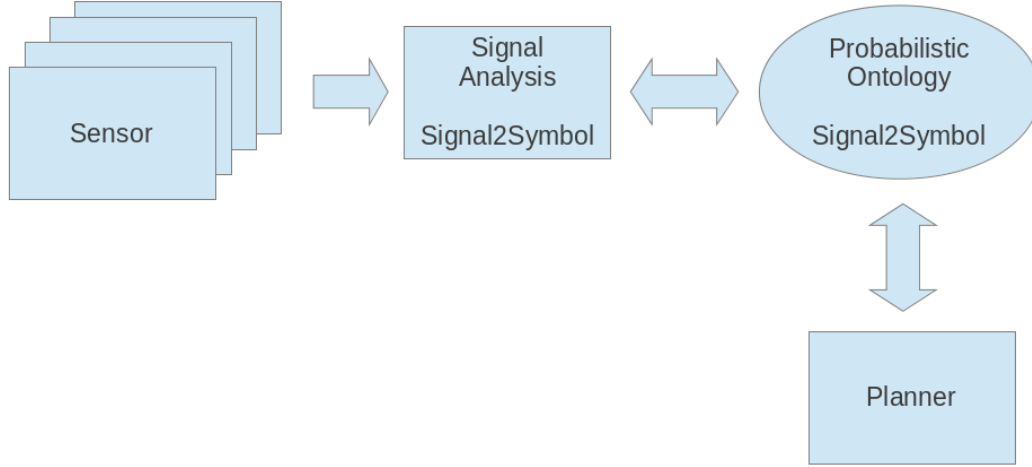


Fig. 1. The full flow from sensor data to the ontology.

belief of the world. The world itself might not be uncertain at all. The panel is either connected with the pipe or not, although its belief is uncertain.

In ontological terms, the uncertainty is again related only to the physical instantiation of a physical specific relation in the *ABox* of the ontology, assuming that the ontology does not represent the world, but the belief on the world. In this case, the *TBox* has no uncertainty at all.

C. Inferred Relations and Concepts

This section now analyses the uncertainty in the robot belief given by reasoning on the knowledge base. The uncertainty is linked on the relations that can be inferred reasoning over the knowledge base. An example for this kind of uncertainty is represented by:

Red light on a panel means: "Panel in use mode", with 99% probability "Panel in test mode", with 0.9% probability "Panel is off, but the light switch is broken", with 0.1% probability

Different from the previous sections, in this case the uncertainty is given by the probabilistic inference process and not by the limitation given by perception. It is to be noticed that again the uncertainty is only related to the belief of the world, not on the world itself. In the example drawn above, the state of the panel is uniquely determined in the world, but its belief is probabilistic, as the information gathered by the robot are not enough to discriminate among the different options. Unlike the previous cases, here the uncertainty is directly related to a probabilistic rule in the *TBox*, applied to physical instances in the *ABox*.

D. Evolving World

In all three cases analysed in the above sections, the world is static and well defined. The uncertainty is related to the belief on the world only. Belief that can be affected by uncertainty in the existence of physical instances, in the relations among them and in the inference process, in order to arrive to a bigger

and more complete knowledge base on the belief of the world. This section will now focus on the uncertainty linked to the evolution of the world. Given a set of actions, for example, the result is very often not deterministic (or *deterministic with hidden variables*).

An example for this kind of uncertainty is represented by: *after switching on the water jet: "the water flow will start", with 99.9% probability "the water flow won't start", with 0.1% probability*

However, the system can also naturally evolve into different states, with different probability, without a specific action from the robot. An example is represented by:

Red light on the underwater system means that after 2 minutes: "the system will be switch off", with 50% probability "the system will be disconnected", with 50% probability

Unlike the previous cases, the uncertainty is given by the inference process of the world evolution. It is not related to perception problems (first two cases), neither on reasoning on the current world (third case), but it is rather focused on the uncertain evolution over time.

IV. PROBABILISTIC ONTOLOGY FRAMEWORK

The general proposed architecture of the system is highlighted in Figure 1. Sensor data are collected and analysed, with a resulting interaction with the ontological representation of the reality, which is then accessed by the Planner system in order to take decisions for the vehicle, based on the world representation. Over the multiple types of uncertainty described in the previous section, the ones related to the instances - and relations among instances - are considered for the purposes of this paper. Analysing the diagram, it is important to discuss the term "*Signal2Symbol*" which is present in two different boxes. An ontological representation is by definition a symbolical representation of high-level concepts. The authors therefore considered that the *Signal2Symbol* module should be positioned in front of the Ontology to query it about high-level knowledge and interact with it. To make a concrete

example, let us consider an environment with several pillar and underwater structures, like the one that will be described in more details in Section V. The high level concepts are for example: *pillar* and *crossbeam*. In this scenario the sensor processing module would run an ATR algorithm (Automatic Target Recognition), to detect pillars and crossbeams and to put them in the knowledge base, if they are not present yet. The *Signal2Symbol* is therefore only in the sensor processing module. The interaction with the knowledge base is related to the update and merge of the information. For example, if subsequent analysis show the presence of a pillar in two very similar location, this information should be merged, without creating in the knowledge base two distinctive pillars overlapping each others. The reasoning on the knowledge base, which is one of the key advantages of an ontological system, compared to other knowledge base framework, happens on the high-level concepts. The probability in the knowledge base is given by the results of the sensor processing module, and is generally represented by the confidence given by the ATR algorithm. However, there is also a different way of proceeding, incorporating the *Signal2Symbol* in the ontology itself, through the reasoning. This would mean to also store basic concepts in the ontology, like for example: *line*, *circle*, *obstacle*, and have ontological rules which put those basic concepts in relation to higher level concepts, like for example *pillar*. In this case the probability in the ontology could be given by probabilistic rules in itself, associating the basic concepts with the higher level ones. For example, two circles found at roughly the same location, but at different depth represent a pillar with 80% probability. Surely both approaches are important as they provide a different perspective. For the purpose of this paper, only the first one has been demonstrated, as the aim is to present a general framework able to address probabilistic ontologies, fully integrated in the vehicle architecture, which could be then extended and improved in each single portion. The second case deserves also a special attention in the probabilistic rules. If the basic concepts are probabilistic, as well as the rules in the ontology itself, the full probabilistic framework might become too complex to be correctly updated during the life of the mission.

Following sensor data analysis, probabilities stored in the ontology are updated real-time according to the sensor processing system. This however leads to an always-changing world, at every new sensor reading. Moreover, most of the planning systems including the one developed for Nessie VII AUV, the HWU vehicle used in this project cannot cope with uncertainty. Probabilistic processing needs to be addressed separately, as considering a full probabilistic framework for everything would quickly become an intractable problem. The system therefore keeps two levels of world modelling: one with uncertainty, regularly updated, and one without any uncertainty, used by the planner. The first world model system is continuously updated. On the other hand, the second system is updated only when something changes radically. In Section V two implementations will be discussed.



Fig. 2. The Underwater Centre, in Fort William.

V. OPERATIONAL RESULTS

In-water trials took place at the Fort William Underwater Centre (Figure 2), in Loch Linnhe, using the Nessie VII AUV. The Underwater Centre is located in the loch itself, with a long bridge connecting it to mainland Scotland. It is therefore an ideal environment for testing underwater cognitive robotics, in presence of man-made structures. The vehicle was driven either by joystick or by waypoint requests through its interface, and the data were consequently post-processed in lab condition. The full architecture was designed and tested to be fully compatible with the vehicle existing architecture, and the tests were done in the same ROS-environment, replaying the data bags (i.e. replaying all sensor data as they were streamed real-time during the mission by the sensors). The ontology has been designed using the tool Protege (<http://protege.stanford.edu/>). In Figure 3, the world model is presented. Blue diamonds represent individuals, whilst the yellow circles represent the taxonomic representation in classes, with their relations. All the concepts are defined in advance, although the vehicle can modify, add and remove instances at run-time. It is important to see the relations among the different concepts. For the implementation of the ontology knowledge base and its interaction with the vehicle's system, two framework were analysed. The first one is JENA - jena.apache.org/. It is a framework written in Java, which provides easy API to interact with an ontology. A message serving module was implemented in order to connect the ROS-based vehicle architecture with the Java-based ontology interaction module.

An initial incomplete world model is given to the vehicle. Experimental results show the ability for the autonomous vehicle to incorporate new semantic information into a probabilistic world model, with a mechanism to provide deterministic information ready to be used by the planner system. Moreover, the vehicle dynamically updates the probability of the semantic concepts which are present in the world model, following sensor data analysis, thus being able not just to insert new entities, but also to remove entities and relations if their probability drops. When queried by the planning system, the ontology module is able to provide a discretisation of the probabilistic ontology, therefore feeding data to the planner in

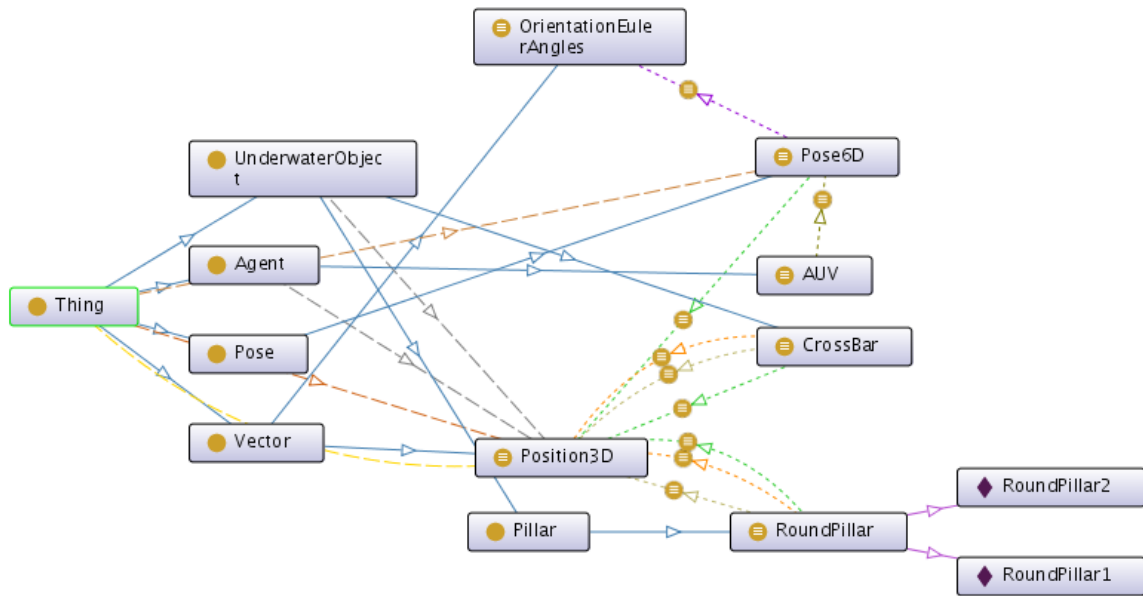


Fig. 3. The world model representation. Blue diamonds represent individuals, whilst the yellow circles represent the taxonomic representation in classes, with their relations.

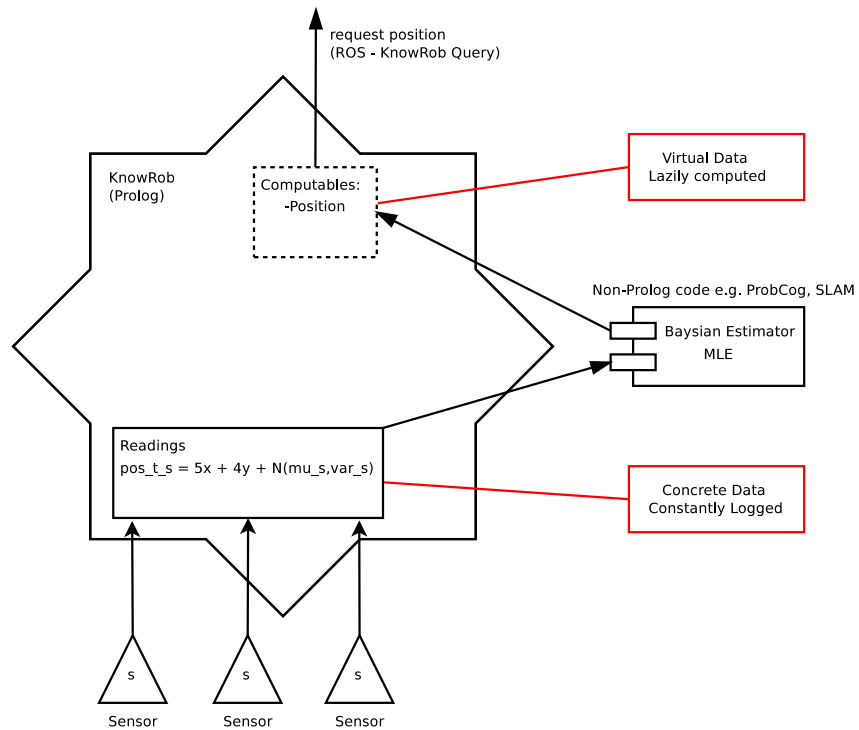


Fig. 4. Proposed architecture, based around KnowRob as the ontology storage and logical processing framework.

a way it can use them.

The second system that has been tested is KnowRob [16], [17]. Figure 4 represents the related architecture. Firstly developed for indoor robotics, adaptations were made to make it fit the challenges of the present research. With built-in reasoner, based on Prolog, KnowRob is able to export services in ROS (Robotic Operating System) and therefore communicate with the other modules in the vehicle. On the implementation point of view, the deterministic ontology is actually represented by *computables*, which provides - when queried - deterministic values to the planner system.

VI. CONCLUSIONS

This work starts from the need of incorporating uncertainty into ontological representations, given that its presence is an inevitable fact for any real field experience. Applications in the underwater domain are very relevant, especially in the oil&gas field. A vehicle able to maintain such a level of cognition, incorporating uncertainty and semantics, is more reliable, can address different situations and safely operate on underwater infrastructures. The paper presents ways to incorporate a Bayesian framework into traditional owl-based ontologies, and analyses its implications for the overall system. A novel solution is proposed, in order to dynamically update the belief on the world model, providing at the same time means for the other modules - especially the planner - to access discretised values. The full architecture and proposed system have been integrated in Nessie VII AUV and results from post-processing of mission data are presented.

ACKNOWLEDGMENT

The authors would like to thank all the other members of the Ocean Systems Laboratory at Heriot-Watt University. The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 - Challenge 2 - Cognitive Systems, Interaction, Robotics - under grant agreement No 288273 - PANDORA

REFERENCES

- [1] Ronald J. Brachman, Deborah L. McGuinness, Peter F. Patel-Schneider, Lori Alperin Resnick, and Alex Borgida. Living with classic: When and how to use a kl-one-like language. In John F. Sowa, editor, *Principles of Semantic Networks: Explorations in the representation of knowledge*, Morgan Kaufmann Series in Representation and Reasoning, pages 401–456. Morgan-Kaufmann, San Mateo, California, 1991.
- [2] Rommel Carvalho, Kathryn Laskey, Paulo Costa, Marcelo Ladeira, Lacio Santos, and Shou Matsumoto. UnBBayes: Modeling uncertainty for plausible reasoning in the semantic web. In *Semantic Web*. InTech, 2010.
- [3] Rommel N. Carvalho, Kathryn B. Laskey, and Paulo Cesar G. Costa. PR-OWL 2.0 – bridging the gap to OWL semantics. In *6th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW)*, 2010.
- [4] Rommel Novaes Carvalho. *Probabilistic Ontology: Representation and Modeling Methodology*. PhD thesis, George Mason University, 2011.
- [5] Paulo Cesar G. Costa. *Bayesian Semantics for the Semantic Web*. PhD thesis, School of Information Technology and Engineering, George Mason University., 2005.
- [6] Zhongli Ding and Yun Peng. A probabilistic extension to ontology language OWL. In *Proceedings of the 37th Hawaii International Conference on System Sciences*, 2004.
- [7] Zhongli Ding, Yun Peng, and Rong Pan. BayesOWL: Uncertainty modeling in semantic web ontologies. In Zongmin Ma, editor, *Soft Computing in Ontologies and Semantic Web*, volume 204 of *Studies in Fuzziness and Soft Computing*, pages 3–29. Springer Berlin Heidelberg, 2006.
- [8] Joseph Y. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46(3):311–350, 1990.
- [9] Jochen Heinsohn. Probabilistic description logics. In *Proceedings of the Tenth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-94)*, pages 311–318, San Francisco, CA, 1994. Morgan Kaufmann.
- [10] Daphne Koller, Alon Levy, and Avi Pfeffer. P-CLASSIC: A tractable probabilistic description logic. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, pages 390–397, 1997.
- [11] Pawel Kozak and Karsten Tolle. Analyzing web profiles using probabilistic ontologies. In *Proceeding of 3rd International Conference on Web Science*, 2011.
- [12] Asma S. Larik and Sajjad Haider. Efforts to blend ontology with bayesian networks: An overview. In *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on*, volume 2, pages V2–598 –V2–602, aug. 2010.
- [13] K.B. Laskey, P. Costa, and T. Janssen. Probabilistic ontologies for knowledge fusion. In *Information Fusion, 2008 11th International Conference on*, pages 1 –8, 30 2008-july 3 2008.
- [14] David Poole, Clinton Smyth, and Rita Sharma. Semantic science: Ontologies, data and probabilistic theories. In Paulo Cesar G. Costa, Claudia d’Amato, Nicola Fanizzi, Kathryn B. Laskey, Kenneth J. Laskey, Thomas Lukasiewicz, Matthias Nickles, and Michael Pool, editors, *Uncertainty Reasoning for the Semantic Web I*, volume 5327 of *Lecture Notes in Computer Science*, pages 26–40. Springer Berlin Heidelberg, 2008.
- [15] Livia Predoiu and Heiner Stuckenschmidt. Probabilistic models for the semantic web: A survey. In *The Semantic Web for Knowledge and Data Management: Technologies and Practices*. Information Science Reference, Hershey, PA, USA, 2008.
- [16] M. Tenorth and M. Beetz. KnowRob – knowledge processing for autonomous personal robots. In *IROS’09 Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, pages 4261–4266, 2009.
- [17] Moritz Tenorth. *Knowledge Processing for Autonomous Robots*. PhD thesis, Technische Universität München, 2011.
- [18] Shenyong Zhang, Yi Sun, Yun Peng, and Xiaopu Wang. BayesOWL: A prototype system for uncertainty in semantic web. In *Proceedings of the 2009 International Conference on Artificial Intelligence, ICAI 2009*, volume 2, pages 678–684, Las Vegas, NV, United states, 2009.