Assessing Conflicts in Ontologies

ANUSCH DAEMI and JACQUES CALMET
IAKS Calmet
Universität Karlsruhe (TH)
Am Fasanengarten 5, 76131 Karlsruhe
GERMANY

Abstract: - We present a method to assess the possible conflict of beliefs in ontologies. Those beliefs represent for example the assessment of an ontology in regard to some goal, i.e. hydrologists may assess a flooding ontology only with maximum flood protection in mind where a politician considers more monetary issues. These different beliefs lead to different interpretations or semantics of the ontology which in turn results in a conflict between them. The method we present for measuring this conflict is based upon the concept of entropy and relative entropy.

Key-Words: - Ontology, Distance, Entropy, Relative Entropy, Conflict, Belief, Knowledge Assessment

1 Introduction

In the upcoming E-Society, which will also be a knowledge society, ontologies will play an important role, partly due to their ability to formally represent knowledge. The concept of ontology, defined for example by Gruber in [1], is nowadays routinely found as a suitable representation of knowledge, especially due to the Semantic Web [2]. Usually, it is defined as a kind of taxonomy or even more crudely as a list of words describing domain specific knowledge. A proper definition of ontology though implies that the knowledge is structured, which is also consistent with the usual concept first introduced by philosophers. In knowledge engineering structured knowledge is a well-understood concept that means the existence of a well-defined data structure to represent knowledge. A common use of ontology covers for example simple hierarchies in the case of taxonomies and there are various portal solutions available for their development [3].

In this paper we add to the concept of ontology the notions of belief [4] and conflict. If an ontology is assessed by different people they may assign different beliefs to the concepts of the ontology, which are based on their interpretation of the view on the world represented by this ontology. This means that one and the same ontology can have different semantics resulting from the different beliefs assigned to the concepts of the ontology. Hence, the conflict of beliefs results in a conflict of interpretation and thus semantics of the ontology. We propose to measure this conflict with the notion of relative entropy, which is by the way consistent with the use of entropy in information theory. The result of the computation of relative entropy between the different beliefs of the concept gives us some information about the compatibility of the interpretation and thus semantics of the ontology. For example a small conflict means that the beliefs and thus the interpretation are similar where a significant conflict connotes a difference of semantics in the knowledge represented by the ontology. It should be noted that this approach differs from classification algorithms used in numerical statistics or numerical taxonomy [5] and from distance measures used in computational linguistics.

In the field of numerical classification there is, according to [6], an old distinction between classifications based on type i.e. with cognitive contents, and on pure statistics. However, the well-known "dynamical clouds" method of Diday [7] has been seen as sitting somewhere between these two features. While classification amounts to summarize in a meaningful way sets of data and to identify the structural links within these sets, our goal is to resolve conflicts in ontology. We base our method on distances generated by relative entropy, indeed a concept of statistics, but there are so many different distances that have been used in statistical data analysis, several not based upon entropy, that this does not constitute a link to classification. Even though we look for structures in ontology, we do not aim at clustering ontologies along these structures. The attempt in the machine learning community to develop robust classifier-learning methods has lead to converting numerical classification into text classification [8]. Our work leads as a byproduct to the investigation in the opposite direction. From an ontology we deduce a structure defining the context in which this ontology is defined. From this context we deduce some sort of semantic characterization of the ontology that is specific to the view of the world enclosed in this ontology. Then, we compare and resolve these different views.

In the domain of computational linguistice there are two major approaches to measure similarity, one using vector space models (VSM) [9] (for an application of VSM models see i.e. [10]) and the other using frequencies of words [11]. A well known measure originating from the frequency approach is [12], which introduced the concept of entropy in this domain. The basic idea of this distance measure is based on the assumption that the more information (in the entropic sense) two concepts have in common, the more similar they are. The information shared by two concepts is indicated by the information content of the concepts that subsume them in a given taxonomy.

The paper is structured as follows. In section 2 we will give, for completeness sake, a brief overview of the well known concept of entropy and the principles of entropy to which relative entropy belongs. Section 3 details the way how we can identify conflicts of beliefs and thus conflicts of interpretation and semantics in an ontology. Further, an interpretation of some results of the application of relative entropy to ontologies is given. Section 4 offers an idea of a first and very simple implementation of this approach, illustrated with an example from the domain of flooding. We conclude the paper in section 5 and give an outlook on further research.

2 Entropy and Relative Entropy

The concept of entropy originates in physics through the second law of thermodynamics [13]. Another important area where entropy plays a central role is statistical mechanics, which is due to the work of Maxwell, Boltzmann, Gibbs [14] and others. Entropy has then been increasingly popular in computer science and information theory, particularly through the paper of Shannon [15].

Maxwell, Boltzmann and Gibbs extended the notion of entropy from thermodynamics into the domain of statistical mechanics, where *macrostates* and *microstates* play an important role. The definition of entropy by Boltzmann [14] is

$$-k \sum_i P_i \log P_i$$

where the P_i are the probabilities that particle i will be in a given microstate, and all the P_i are evaluated for the same macrostate; k is the famous Boltzmann constant. Therefore, entropy in statistical mechanics denotes the uncertainty about which state the system is in.

Entropy, as defined in the work of Shannon [15], represents the *information content* of a message or, from the point of view of the receiver, the uncertainty about the message the sender produced, prior to its reception. It is defined as

$$-\sum_i p(i) \log p(i).$$

$p(i)$ is the probability of receiving message i and Shannon has shown that $\log p(i)$ is the only function that satisfies all requirements to measure information. Some years later Renyi proved the uniqueness of $\log$ as a measure of information in [16].

Entropy is nowadays heavily used by physicists for problems ranging from quantum computing to black hole physics, for an overview see [17]. We restrict our investigations of entropy to its application in the context of knowledge. In this first attempt we use methods that are similar to those already investigated in linguistics, social network analysis [18] and information theory. For an example of the application of mutual information (an information theoretic distance measure) to ontologies see [19].

2.1 Relative entropy

Relative entropy $D(\mathbf{p}||\mathbf{q})$, also known as minimum cross entropy, directed divergence or Kullback-Leibler distance, belongs to the Ali-Silvey class of information-theoretic distance measures [20]. It was first introduced by Kullback and Leibler [21] as a measure to discriminate between probability distributions

$$\mathbf{p} = p_1, \ldots, p_n \text{ and } \mathbf{q} = q_1, \ldots, q_n$$

of a system, where **p** denotes assumed distributions of a system and **q** is the real, usually unknown, distribution. The formal definition of this distance (only by analogy, since it violates the axiom of symmetry) is:

$$D(\mathbf{p}||\mathbf{q}) = \sum_i p_i \log \frac{p_i}{q_i}$$

This concept has been extended by Shore and Johnson [22] into the domain of information theory, where it is known as the principle of minimum cross-entropy in contrast to the well known principle of maximum entropy of Jaynes [23] which is nowadays heavily used in non-monotonic reasoning [24]. The interpretation of Shore and Johnson is that relative entropy applies when a prior distribution **q**, that estimates an unknown distribution, is known in addition to other constraints imposed on the unknown distribution. Among those distributions **p** that satisfy the constraints, the one with the least relative entropy should be chosen.

Similar in this vein is the use of relative entropy in probabilistic inference (see for example [25]), where **q** is the prior probability and **p** the posterior probability. The distribution to choose for an updating of probabilities is the one with the minimum distance.

We propose to add another interpretation to relative entropy for ontologies in the context of knowledge: It should represent the *conflict* of different beliefs, and therefore the conflict between different interpretations, placed in an ontology. This is also consistent with one of the classical interpretations of entropy in information theory [26].

3 Semantic Conflicts in Ontologies

As stated in the introduction, an ontology is structured knowledge and thus able to represent a view, on a particular domain of the world, of a person or group of persons. An ontology consists of concepts, maybe with attributes, and relations between those concepts. Usually there are at least two kinds of relations: *is-a* relations representing an inheritance hierarchy and relations representing the interaction between concepts. These relations depict a model of the world represented by the ontology and how it works. We take as example the domain of flooding: An ontology in this domain may describe flooding causes and mitigation incentives. The *is-a* hierarchy may consist of urban and rural areas, concrete, water and forest areas linked to urban areas and so on (see figure 1). The relations between the concepts in turn are describing, for example, causal effects like: urban areas sealed with concrete result in fast runoff and a decrease in the sealed area reduces the risk of flooding.

The *is-a* relations between the concepts are also able to represent the strength of a belief a person places in a concept, i.e. how important this concept is in the context of the complete ontology, maybe in regard to some goal. Take the flooding example from above: The ontology was generated by some domain experts (hydrologists) and faithfully describes a model for flooding and mitigation. Then they assign belief values to the concepts in the ontology, where those values show what they belief is important for flood mitigation (the goal), i.e. reduce concrete in some urban areas (decrease the *belief* of the urban areas that are sealed), like parking lots and replace them for example with water-permeable stone. A politician on the other hand may place entire different beliefs on the concepts in the context of flooding and mitigation because he dislikes the idea of replacing concrete in parking lots due to costs and dirty shoes, i.e. he has different goals.

The assignment of the belief values to the concepts can be done either manually by the involved parties or (semi-) automatically through different techniques like data mining, document analysis or graph theoretical measures.

In our approach we propose to measure this conflict of belief via the notion of relative entropy, which can be seen as a generalization of normal entropy. It is more general in the sense, that we can write the usual definition of entropy (ignoring signs)

$$E(\mathbf{p}) = -\sum_i p_i \log p_i$$

also as

$$D(\mathbf{p}||\mathbf{q}) = \sum_i p_i \log \frac{p_i}{q_i}$$

if we set $q_1, \ldots, q_n = 1$, denoting a uniform distribution. Since a uniform distribution of beliefs in an ontology is very unlikely, because it would mean absolute indifference about what is important and not (everything would be equally possible), we choose relative entropy as a measure for conflict. This measure allows arbitrary distributions **q** as long as they comply with the standard axioms for a probability distribution. Thus the relative entropy $D(\mathbf{p}||\mathbf{q})$ describes the conflict of the different interpretations or semantics of the ontology represented by the different beliefs **p** and **q**, for example the different beliefs in flooding incentives by politicians **p** and hydrologists **q**.

To actually measure the conflict between the interpretations we choose a reference belief **q**, that of the domain experts for example, as to which we calculate via relative entropy the conflict in regard to the beliefs **p** of the politicians. The result of this calculation represents the conflict between two different interpretations of the same ontology, in the example above the conflict between maximum flood protection (hydrologist) and monetary issues (politician). This means, we can assess the interpretation of the knowledge represented by the ontology and the corresponding belief values in relation to other interpretations.

3.1 Interpretation

A low value of the relative entropy,

$$D(\mathbf{p}||\mathbf{q}) < 0.5$$

can be interpreted as having similar beliefs **p** in the concepts of the ontology in regard to the reference beliefs **q**, and thus the interpretation and semantics of the ontology are still valid. The actual value of 0.5 is a result of the computations of relative entropy on a simple flooding ontology and is surely different for other ontologies and applications. The minimum of conflict is achieved if the beliefs **p** and **q** are equal, i.e. both interpretations are the same, and hence $D(\mathbf{p}||\mathbf{q}) = 0$. A high value of the relative entropy,

$$D(\mathbf{p}||\mathbf{q}) \geq 0.5$$

denotes a significant conflict, which means that the beliefs **p** in the concepts of the ontology are very different from the reference beliefs **q**. Therefore the interpretation and semantics of the knowledge represented by the ontology differs. One can also say, that the assessment of the knowledge is different.

The difference between maximum and minimum of the relative entropy,

$$Spec(\mathbf{p}||\mathbf{q}) := \max_i\{D(p_i||q_i)\} - \min_i\{D(p_i||q_i)\}$$

called spectrum, provides information about the possibility of a conflict. If the spectrum is very small, a different interpretation cannot occur, but if we have a large spectrum then there are a lot of different interpretations possible. Also the spectrum gives information about how different the interpretations may be.

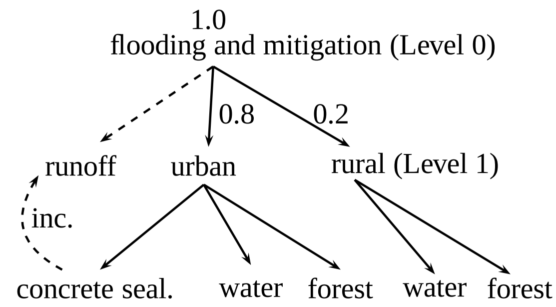

Figure 1: Simple example ontology. Thick lines are *is-a* relations, dashed lines are causal relations.

4 Implementation and a possible application

In a first step, a simple algorithm was devised to calculate the relative entropy between the beliefs of some reference ontology, represented by a simple taxonomy with a tree structure, and all other possible beliefs with accuracy of 0.05. This means that the strength of the beliefs could vary in steps of 0.05, where the minimum of a belief in a concept is 0.05 and the maximum depends on the value of the parent concept and its descendants because of the requirement that the minimum belief is 0.05.

As an example look at figure 1, where we have as root concept *flooding and mitigation*, as its child concepts regarding *is-a* relations *rural* and *urban* areas. As children of *urban* areas *concrete sealing*, *forest*, *water* and so on. The belief in our root concept would be 1 (important for our view on the world is now only *flooding and mitigation*), *urban* areas would be rather important (belief is 0.8) and we largely disregard *rural* areas (0.2). The minimum that could be assigned to *concrete*, *forest* or *water* would be 0.05, though they must sum up to the belief of the parent node, i.e. *urban* areas $= 0.8$. That is, the beliefs in the children of a concept should not exceed the belief in their parent. This represents the constraint, that the concepts are neither independent of each other, as it would be the case if the *is-a* relations of each concept would sum up to one. Nor do we require a complete assessment of all states of the world, i.e. every concepts in the ontology must sum up to one, as this implies, that all the concepts are disjunct. Hence, the maximum belief of *concrete*, *forest* or *water* in this example is 0.7, because we have as requirement that the other two have as minimum belief 0.05.

Zero was not chosen as a minimum for the beliefs p_i, because this would result in an infinite relative entropy and thus infinite conflict between the interpretations:

$$\lim_{p_i \to 0} \log \frac{p_i}{q_i} = \infty$$

This is understandable, because if we remove some concept of our ontology (we do not belief in it any longer), then a change of our model, that represents how the world is shaped and works, occurres and it may not be appropriate to compare it to the old ontology for this specific part anymore.

Another constraint for the assignment of the beliefs is that every belief p_i on one level of the ontology must sum up to one, i.e. the sum of *urban* and *rural* in level 1 must be equal to one. This is necessary due to the additivity requirement of the distributions $\mathbf{p}$ and $\mathbf{q}$ for the calculation of relative entropy.

Further it is a constructive algorithm, because it assigns to every concept its beliefs value p_i, subject to the restrictions mentioned above, and calculates the relative entropy in regard to the reference beliefs q_i. With this information it is possible to identify sets of beliefs, where the conflict is minimal, $0 \leq D(\mathbf{p}||\mathbf{q}) < 0.5$, or maximal, $D(\mathbf{p}||\mathbf{q}) \geq 0.5$. If sets representing minimal or maximal conflict are available, one can also deduce how to adjust the beliefs of an ontology with large conflict to achieve a belief structure with lesser conflict potential.

Another advantage of this method is its effectiveness, because the calculation only requires the computations of sums, fractions and logarithms. This is not the case if we would use other information-theoretic distance measures, for example the Chernoff distance [27], which requires computation of maxima (a non-trivial optimization problem). The same problem is encountered if we would use the principle of maximum entropy (and minimize the results to get the minimum conflict) as it is done for example in non-monotonic reasoning [28]

One possible application for this method could be a fast assessment of the status quo regarding other, more desirable states which are modeled by an ontology. For example, the status quo in flood mitigation is modeled by an ontology and given to experts in this field. They make the appropriate changes with maximum flood protection in mind and then the conflict between both ontologies is calculated, which provides an average value about the changes that will occur. This value can help for example decision makers in assessing the consequences of the proposed changes very fast, i.e. in case of emergency.

5 Conclusion

We have presented a new application of the notion of relative entropy as a measure of conflict between different beliefs placed in an ontology, which is consistent with the interpretation of entropy in information theory. The measure is able to identify similar interpretations and thus semantics of an ontology or radically different ones, depending on the value of the relative entropy. This may not be the case if we would compare just the belief values, because they may be misleading in the case where a rather large difference occurs between the beliefs, but the parent concept of those beliefs is not an *important* one, i.e. has a low belief value.

Further research has to be done regarding the exact numerical value, when a belief is seen as similar or different to a reference belief, even though this may be application dependent. The performance of the algorithm computing the relative entropy for all possible beliefs can also be improved because some computations are redundant. Another interesting aspect is how a change of belief in the concepts affects the causal relations in the ontology. For example, if we change the belief in *urban* areas that are sealed with *concrete* in figure 1, we want to asses how this affects the causal implication (dashed lines) to the *runoff*, which is strongly related to knowledge assessment.

References

[1] Gruber T. Towards principles for the design of ontologies used for knowledge sharing. *Knowledge Acquisition*, 1993, pp. 199 – 220.

[2] Berners-Lee T., Hendler J., and Lassila O. The Semantic Web. *Scientific American*, 05 2001.

[3] Zhao G., Verlinden R., and Meersman R. Ontology Development Portal. In *Proceedings of 3rd WSEAS Int.Conf. on Artificial Intelligence, Knowledge Engineering, Data Bases (AIKED 2004)*, 2004.

[4] Bacchus F., Grove A., Halpern J.Y., and Koller D. From statistical knowledge bases to degrees of belief. *Artificial Intelligence*, vol. 87, 1996, pp. 75 – 143.

[5] Agarwala R., Bafna V., Farach M., Paterson M., and Thorup M. On the Approximability of Numerical Taxonomy (Fitting Distances by

Tree Metrics). *SIAM Journal on Computing*, vol. 28(3), 1999, pp. 1073 – 1085.

[6] Halphen E. La notion de vraissemblance. Essai sur les fondements du Calcul des Probabilités et de la Statistique Mathématique. In *I.U.S.P Publications*, vol. 4, 1955.

[7] Celeux G., Diday E., and Govaert G. *Classification automatique des donnees*. Dunod, 1989.

[8] Macskassy S.A., Hirsh H., Banerjee A., and Dayanik A.A. *Converting Numerical Classfication into Text Classification*. Preprint Rutgers Univ., June 2002.

[9] Jones K.S. and Willett P. (eds.). *Readings in Information Retrieval*. Morgan Kaufmann, 1997.

[10] Karimi S., Moeini A., and Hejazi M.R. Information Mining Based on Fusing Results of Multi-Perspective Cluster-based Summarizations. In *Proceedings of 3rd WSEAS Int.Conf. on Artificial Intelligence, Knowledge Engineering, Data Bases (AIKED 2004)*, 2004.

[11] Hotho A., Maedche A., and Staab S. Ontology-based Text Document Clustering. *Künstliche Intelligenz*, 04 2002, pp. 48 – 54.

[12] Resnik P. Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In *Proceedings of IJCAI*, 1995.

[13] Maxwell J.C. *Theory of Heat*. Dover; Reprint, 2001.

[14] Tolman R.C. *The principles of statistical mechanics*. Dover, 1979.

[15] Shannon C.E. A mathematical theory of communication. *Bell System Technical Journal*, vol. 27, July and October 1948, pp. 379 – 423, 623 – 656.

[16] Renyi A. *Probability theory*. Springer Verlag, 1970.

[17] Bashkirov A.G. On the Renyi entropy, Boltzmann Principle, Levy and Power-law distributions and Renyi parameter. *eprint arXiv:cond-mat/0211685*, 2003.

[18] Wasserman S. and Faust K. *Social network analysis: Methods and applications, Structural analysis in the social sciences*. Cambridge University Press, Cambridge, 1994.

[19] Calmet J. and Daemi A. From Entropy to Ontology. In Trappl R. (ed.), *Cybernetics and systems 2004 - AT2AI-4: From Agent Theory to Agent Implementation*, vol. 2, 2004.

[20] Ali S.M. and Silvey S.D. A General Class of Coefficients of Divergence of One Distribution from Another. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 28(1), 1966, pp. 131 – 142.

[21] Kullback S. and Leibler R. On Information and Sufficiency. *Annal Math. Stat.*, vol. 22, 1951, pp. 79 – 86.

[22] Shore J. and Johnson R. Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy. *IEEE Trans. on Information Theory*, vol. 26, 1980, pp. 26 – 37.

[23] Jaynes E.T. Where do we stand on maximum entropy? In Levine R. and Tribus M. (eds.), *The Maximum Entropy Formalism*. MIT Press, 1979.

[24] Goldszmidt M., Morris P., and Pearl J. A Maximum Entropy Approach to Nonmonotonic Reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15(3), March 1993.

[25] Caticha A. Relative Entropy and Inductive Inference. *eprint arXiv:abs-physics/0311093*, 2003.

[26] Klir G.J. and Wierman M.J. *Uncertainty-Based Information*. Studies in fuzziness and soft computing. Springer Verlag, 1998.

[27] Chernoff H. Measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annal Math. Stat.*, vol. 23, 1952, pp. 493 – 507.

[28] Lukasiewicz T. Weak Nonmonotonic Probabilistic Logics. In *9th International Conference on Principles of Knowledge Representation and Reasoning (KR2004)*. AAAI Press, 2004.