Bridging between Sensor Measurements and Symbolic Ontologies through Conceptual Spaces

Stefan Dietze, John Domingue
Knowledge Media Institute,
The Open University,
MK7 6AA, Milton Keynes, UK
{s.dietze, j.b.domingue}@open.ac.uk

Abstract. The increasing availability of sensor data through a variety of sensor-driven devices raises the need to exploit the data observed by sensors with the help of formally specified knowledge representations, such as the ones provided by the Semantic Web. In order to facilitate such a Semantic Sensor Web, the challenge is to bridge between symbolic knowledge representations and the measured data collected by sensors. In particular, one needs to map a given set of arbitrary sensor data to a particular set of symbolic knowledge representations, e.g. ontology instances. This task is particularly challenging due to the potential infinite variety of possible sensor measurements. Conceptual Spaces (CS) provide a means to represent knowledge in geometrical vector spaces in order to enable computation of similarities between knowledge entities by means of distance metrics. We propose an ontology for CS which allows to refine symbolic concepts as CS and to ground instances to so-called prototypical members described by vectors. By computing similarities in terms of spatial distances between a given set of sensor measurements and a finite set of prototypical members, the most similar instance can be identified. In that, we provide a means to bridge between the real-world as observed by sensors and symbolic representations. We also propose an initial implementation utilizing our approach for measurement-based Semantic Web Service discovery.

Keywords: Sensor Data, Conceptual Spaces, Semantic Sensor Web, Vector Spaces.

1 Introduction

Current and next generation wireless communication technologies will encourage widespread use of well-connected *sensor-driven devices* which in fact produce *sensor data* by observing and measuring real-world environments. This has already lead to standardisation efforts aiming at facilitating the so-called *Sensor Web*, such as the ones by the Sensor Web Enablement Working Group[1] of the Open Geospatial Consortium (OGC)[2]. The increasing availability of sensor data raises the need to merge such data with formally specified knowledge representations, such as the ones

---

[1] http://www.opengeospatial.org/projects/groups/sensorweb
[2] http://www.opengeospatial.org/

provided by *Semantic Web (SW)* standards such as OWL [22] or RDF [23]. However, whereas sensor data usually relies on measurements of perceptual characteristics to describe real-world phenomena, ontological knowledge presentations represent real-world entities through *symbols*. The symbolic approach – i.e. describing symbols by using other symbols, without a grounding in perceptual dimensions of the real world – leads to the so-called *symbol grounding* problem [2] and does not entail meaningfulness, since meaning requires both the definition of a terminology in terms of a logical structure (using symbols) and grounding of symbols to a perceptual level [2][13].

In that, in order to facilitate the vision of the *Semantic Sensor Web (SSW)* [18] the challenge is to bridge between formal symbolic knowledge representations and the measured data collected by sensors by mapping a given set of arbitrary sensor data to a particular set of symbolic representations. This task is particularly challenging due to the potential infinite variety of possible data sets.

Conceptual Spaces (CS) [8] follow a theory of describing knowledge in geometrical vector spaces which are described by so-called quality dimensions to bridge between the perceived and the symbolic world. Representing instances as vectors, i.e. *members*, in a CS provides a means to compute similarities by means of spatial distance metrics. However, several issues still have to be considered when applying CS. For instance, CS as well as sensor data provide no means to represent arbitrary relations between data sets, such as part-of relations.

In order to overcome the issues introduced above, we propose a two-fold knowledge representation approach which extends symbolic knowledge representations through a refinement based on CS. This is achieved based on an ontology which allows to refine symbolic concepts as CS and to ground instances to so-called *prototypical members*, i.e. prototypical vectors, in the CS. The resulting set of CS is formally represented as part of the ontology itself. By computing similarities in terms of spatial distances between a given set of sensor measurements and the finite set of prototypical members, the most similar instance can be identified. In that, our approach provides a means to bridge between the real-world - as measured by sensor data - and symbolic representations.

The remainder of the paper is organized as follows: Section 2 introduces the symbol grounding problem in the context of sensor data, while our representational approach based on CS is proposed in Section 3. In Section 4, we introduce an implementation of our approach based on an existing SWS reference model and we introduce first proof-of-concept prototype in Section 5. Finally, we discuss and conclude our work in Section 6.

2 Sensor Data, Symbol Grounding and Spatial Representations

This section motivates our approach by introducing the so-called symbol grounding problem in the context of the SSW and introduces some background knowledge on metric-based spatial knowledge representation.

2.1. Sensor Data and the Symbol Grounding Problem

Sensor data usually consists of measurements which describe observations of phenomena in real-world environments. In order to ensure a certain degree of interoperability between heterogeneous sensor data, recent efforts, such as the OpenGIS Observations and Measurements Encoding Standard (O&M)[3], propose a standardized approach to represent observed measurements based on a common XML schema. However, in order to provide comprehensive applications capable of reasoning in real-time on observed real-world phenomena, i.e. the contextual knowledge produced by sensor-driven devices, one needs to bridge between the measurements provided by sensors and the formally specified knowledge as, for instance, exploited by the Semantic Web [18]. Figure 1 illustrates the desired progression from observed real-world phenomena, e.g. a certain color, to measurements provided by sensors, e.g. measurements of the hue, saturation and lightness (HSL) dimensions, to symbolic knowledge entities such as a particular OWL individual representing a specific color.

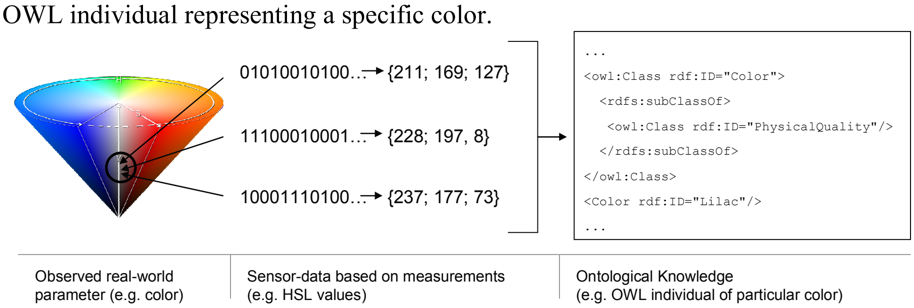

Fig. 1. Envisaged progression from real-world observations to ontological representations through sensor data.

However, whereas sensor data usually relies on measurements of perceptual characteristics to describe real-world phenomena, ontological knowledge presentations represent real-world entities through symbols what leads to a representational gap. Hence, several issues have to be taken into account. The symbolic approach – i.e. describing symbols by using other symbols, without a grounding in the real world or perceptual dimensions what is known as the *symbol grounding* problem [2] – of established SW representation standards, leads to ambiguity issues and does not entail meaningfulness, since meaning requires both the definition of a terminology in terms of a logical structure (using symbols) and grounding of symbols to a perceptual level [2][13]. Moreover, describing the complex notion of any specific real-world entity in all its facets through symbolic representation languages is a costly task and may never reach semantic meaningfulness.

Hence, in order to facilitate the vision of the SSW, the challenge is, to map a given set of sensor observation data to semantic (symbolic) instances which most appropriately represent the observed real-world entity within an ontology. In this

[3] http://www.opengeospatial.org/standards/om

respect, it is particularly obstructive that a potentially infinite amount of real-world phenomena, i.e. measurement data, needs to be mapped to a finite set of knowledge representations, e.g. ontological concepts or instances.

2.2. Exploiting Measurements through spatial Knowledge Representations

Sensor data usually consists of sets of measurements being observed from the surrounding environment. In that, spatially oriented approaches to knowledge representation which exploit metrics to describe knowledge entities naturally appear to be an obvious choice when attempting to formally represent sensor data. *Conceptual Spaces (CS)* [8] follow a theory of describing entities in terms of their quality characteristics similar to natural human cognition in order to bridge between the perceived and the symbolic world. CS foresee the representation of concepts as multidimensional geometrical *Vector Spaces* which are defined through sets of quality dimensions. Instances are supposed to be represented as vectors, i.e. particular points in a CS. For instance, a particular color may be defined as point described by vectors measuring HSL or RGB dimensions. Describing instances as points within vector spaces where each vector follows a specific metric enables the automatic calculation of their semantic similarity by means of distance metrics such as the Euclidean, Taxicab or Manhattan distance [11] or the Minkowsky Metric [19]. Hence, semantic similarity is implicit information carried within a CS representation what is perceived as one of the major contribution of the CS theory. *Soft Ontologies (SO)* [10] follow a similar approach by representing a knowledge domain *D* through a multi-dimensional *ontospace A*, which is described by its so-called *ontodimensions.* An item *I*, i.e. an instance, is represented by scaling each dimension to express its impact, presence or probability in the case of *I*. In that, a SO can be perceived as a CS where dimensions are measured exclusively on a ratio-scale.

However, several issues have to be taken into account. For instance, CS as well as SO do not provide any notion to represent any arbitrary relations [17], such as *part-of* relations which usually are represented within symbolic knowledge models. Moreover, it can be argued, that representing an entire knowledge model through a coherent CS might not be feasible, particularly when attempting to maintain the meaningfulness of the spatial distance as a similarity measure. In this regard, it is even more obstructive that the scope of a dimension is not definable, i.e. a dimension always applies to the entire CS/SO [17].

3 Grounding Ontological Concepts in Conceptual Spaces

We propose the grounding of ontologies in multiple CS in order to bridge between the measurements provided by sensor-driven devices and symbolic representations of the SW.

3.1. Approach: Spatial Groundings for Symbolic Ontologies

We claim that CS represent a particularly promising model when being applied to individual concepts instead of representing an entire ontology in a single CS. By representing instances as so-called *prototypical members* in CS, arbitrary sensor-data can be associated with specific ontology instances in terms of the closest – i.e. the most similar – prototypical member representation.

We propose a two-fold representational approach – combining SW vocabularies with corresponding representations based on CS – to enable similarity-based matchmaking between a given set of sensor data and ontological representations. In that, we consider the representation of a set of n concepts C of an ontology O through a set of n Conceptual Spaces CS. Instances of concepts are represented as prototypical members in the respective CS. The following Figure 2 depicts this vision:

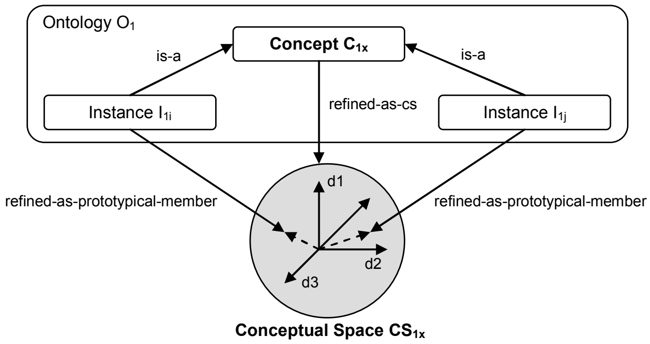

Fig. 2. Representing ontology instances through prototypical members in CS.

While benefiting from implicit similarity information within a CS, our hybrid approach allows overcoming CS-related issues by maintaining the advantages of ontology-based knowledge representations and provides a means to ground knowledge entities to cognitive dimensions based on measurements. To give a rather obvious example, a concept describing the notion of a geospatial location could be grounded to a CS described through quality dimensions such as its longitude and latitude. In previous work [3][4], we provided more comprehensive examples, even for rather qualitative notions, such as particular subjects or learning styles.

Provided our refinement of ontology concepts as CS and of instances as prototypical members, a given set of sensor data which measures the quality dimensions of a particular CS_i represents a vector v in CS_i which can be mapped to an appropriate ontology instance I in terms of the spatial distance of the prototypical member of I and v. Figure 3 illustrates the approach based on the color example introduced in Section 2.1. While measurements obtained from sensors are well-suited to be represented as vectors, i.e. members, in a CS, we facilitate similarity-based computation between a given set of sensor data and sets of prototypical members which represent ontological instances. For instance, the example in Figure 3 depicts the utilisation of a CS based on the HSL dimensions to map between color measurements obtained through sensors and prototypical members representing certain color instances. Based on the spatial distance between one measured color vector and different prototypical members, the closest vector, i.e. the most similar one, can be identified. In that, CS provide a means to bridge between observed sensor data and symbolic ontological representations.

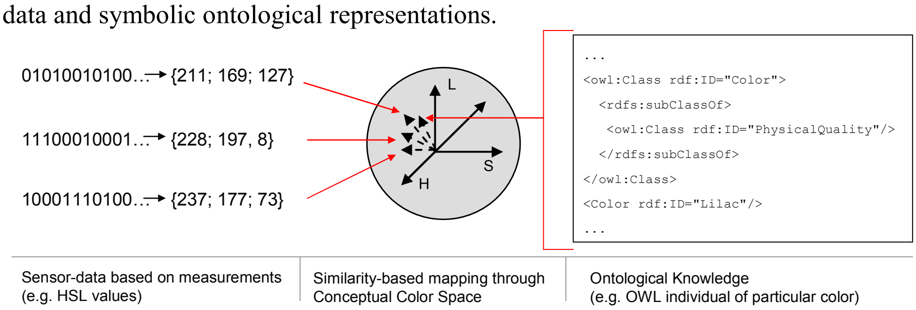

Fig. 3. Similarity-based mapping between distinct sets of sensor-based color measurements and ontological color instances based on a common CS using the HSL dimensions.

3.2. A formal Ontology to represent Conceptual Spaces

In order to be able to refine and represent ontological concepts through CS, we formalised the CS model into an ontology, currently being represented through OCML [12]. Hence, a CS can simply be instantiated in order to represent a particular concept.

Referring to [16][8], we formalise a CS as a vector space defined through quality dimensions d_i of CS. Each dimension is associated with a certain metric scale, e.g. ratio, interval or ordinal scale. To reflect the impact of a specific quality dimension on the entire CS, we consider a prominence value p for each dimension. Therefore, a CS is defined by

$$CS^n = \{(p_1 d_1, p_2 d_2, \ldots, p_n d_n) | d_i \in CS, p_i \in \Re\}$$

where P is the set of real numbers. However, the usage context, purpose and domain of a particular CS strongly influence the ranking of its quality dimensions. This clearly supports our position of describing distinct CS explicitly for individual concepts. Please note that we do not distinguish between dimensions and domains [8] but enable dimensions to be detailed further in terms of subspaces. Hence, a dimension within one space may be defined through another CS by using further dimensions [16]. In this way, a CS may be composed of several subspaces and consequently, the description granularity can be refined gradually. Dimensions may be correlated. For instance, when describing an apple the quality dimension describing its sugar content may be correlated with the taste dimension. Information about correlation is expressed through axioms related to a specific quality dimension instance.

A particular (prototypical) member M – representing a particular instance – in the CS is described through valued dimension vectors v_i:

$$M^n = \{(v_1, v_2, \ldots, v_n) | v_i \in M\}$$

With respect to [16], we define the semantic similarity between two members of a space as a function of the Euclidean distance between the points representing each of the members. Hence, with respect to [16], given a CS definition *CS* and two members *V* and *U*, defined by vectors $v_0, v_1, \ldots, v_n$ and $u_1, u_2, \ldots, u_n$ within *CS*, the distance between *V* and *U* can be calculated as:

$$dist(u,v) = \sqrt{\sum_{i=1}^{n} p_i \left(\left(\frac{u_i - \bar{u}}{s_u} \right) - \left(\frac{v_i - \bar{v}}{s_v} \right) \right)^2}$$

where $\bar{u}$ is the mean of a dataset *U* and s_u is the standard deviation from *U*. The formula above already considers the so-called Z-transformation or standardization [13] which facilitates the standardization of distinct measurement scales utilised by different quality dimensions in order to enable the calculation of distances in a multi-dimensional and multi-metric space. Please note, as mentioned in Section 2.2, different distance metrics could be applied depending on the nature and purpose of the CS.

3.3. Representing Ontologies through Conceptual Spaces

The derivation of an appropriate space CS_i to represent a particular concept C_i of a given ontology *O* is understood a non-trivial task which aims at the creation of a CS instance which most appropriately represents the real-world entity represented by C_i. We particularly foresee a transformation procedure consisting of the following steps:

S1. Representing concept properties pc_{ij} of C_i as dimensions d_{ij} of CS_i.
S2. Assignment of metrics to each quality dimension d_{ij}.
S3. Assignment of prominence values p_{ij} to each quality dimension d_{ij}.
S4. Representing instances I_{ik} of C_i as members in CS_i.

Given the formal ontological representation of the CS model (Section 3.2), we are able to simply instantiate a specific CS by applying a transformation function

$$trans : C_i \Rightarrow CS_i$$

which is aimed at instantiating all elements of a CS, such as dimensions and prominence values *(S1 – S3*). *S1* aims at representing each concept property pc_{ij} of C_i as a particular dimension instance d_{ij} together with a corresponding prominence p_{ij} of a resulting space CS_i:

$$trans : \{(pc_{i1}, pc_{i2}, \ldots, pc_{in}) | pc_{ij} \in PC_i\} \Rightarrow \{(p_{i1}d_{i1}, p_{i2}d_{i2}, \ldots, p_{in}d_{in}) | d_{ij} \in CS_i, p_{ij} \in \Re\}$$

Please note that we particularly distinguish between data type properties and relations. While the latter represent relations between concepts, these are not represented as dimensions since such dimensions would refer to a range of concepts (instances) instead of quantified metrics, as required by *S2*. Therefore, in the case of relations, we propose to maintain the relationships represented within the original ontology *O* without representing these within the resulting CS_i. In that, the complexity of CS_i is reduced to enable the maintainability of the spatial distance as appropriate similarity measure. The assignment of metric scales to dimensions (*S2*) which naturally are described using quantitative measurements, such as size or weight, is rather straightforward. In such cases, interval scale or ratio scale, could be used, whereas otherwise, a nominal scale might be required. *S3* is aimed at assigning a prominence value p_{ij} – chosen from a predefined value range – to each dimension d_{ij}. Since the assignment of prominences to quality dimensions is of major importance for the expressiveness of the similarity measure within a CS, most probably this step requires incremental ex-post re-adjustments until a sufficient definition of a CS is achieved.

With respect to *S4*, one has to represent all instances I_{ki} of a concept C_i as member instances in the created space CS_i:

$$trans : I_{ik} \Rightarrow M_{ik}$$

This is achieved by transforming all instantiated properties pi_{ikl} of I_{ik} as valued vectors in CS_i.

$$trans : \{(pi_{ik1}, pi_{ik2}, ..., pi_{ikn}) | pi_{ikl} \in PI_{ik}\} \Rightarrow \{(v_{ik1}, v_{ik2}, ..., v_{ikn}) | v_{ikl} \in M_{ik}\}$$

Hence, given a particular CS, representing instances as members becomes just a matter of assigning specific measurements to the dimensions of the CS. In order to represent all concepts C_i of a given ontology *O*, the transformation function consisting of the steps *S1-S4* has to be repeated iteratively for all C_i which are element of *O.* The accomplishment of the proposed procedure results in a set of CS instances which each refine a particular concept together with a set of member instances which each refine a particular instance. Please note that applying the procedure proposed here requires additional effort which needs to be further investigated within future work.

4 Implementation - Exploiting Sensor Data for Semantic Web Service Discovery

In previous work [3][4], we applied our two-fold approach to Semantic Web Services (SWS) technology [6] which aims at the automated discovery, orchestration and invocation of Web services based on comprehensive semantic annotations of services. Current results of SWS research are available in terms of reference models such as OWL-S [14], SAWSDL[4] or WSMO [24]. In [3][4], our CS representation was deployed to refine instances which are part of SWS annotations in order to enable interoperability between heterogeneous SWS and SWS requests. In contrast, here we propose the utilization of our CS-based representational approach to facilitate interoperability between observations and measurements provided by sensors and symbolic SWS representations based on extensions which are described in this section.

The representational model described above had been implemented by and aligned to established SWS technologies based on WSMO [24] and the Internet Reasoning Service IRS-III [1]. Further details on the IRS-III Service Ontology and its extension through our CS formalisation can be found in [5]. However, please note that in principle the representational approach described above could be applied to any SWS reference model and is particularly well-suited to support rather light-weight approaches such as SAWSDL or WSMO Lite [21].

[4] http://www.w3.org/2002/ws/sawsdl/spec/

In order to facilitate the representational approach described in Section 3, we aligned the CS Ontology (Section 3.2) with the IRS-III Service Ontology to allow for the refinement of individual concepts – used as part of formal SWS descriptions – as formally expressed CS. In that, instances being used to represent SWS characteristics such as interfaces or capabilities can be refined as vectors.

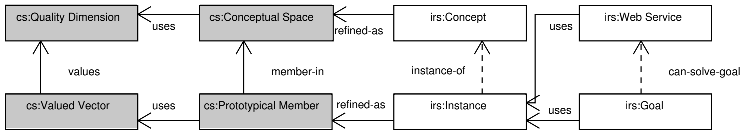

Fig. 4. Core concepts of the CS Ontology aligned to the IRS-III Service Ontology.

Figure 4 depicts the core concepts of CSO and their alignment with the IRS-III Service Ontology. Concepts (instances) as being used by IRS service or goal descriptions are refined as CS (members) within the CSO. In that, following the procedure proposed in Section 3.3, service capabilities are refined in multiple CS. To take into account the representational gap between measurement data as provided by sensors and symbolic SWS goal representations, we introduced a novel way of requesting goal achievements through IRS-III. Instead of simply invoking a goal by providing the goal achievement request SWS_i, including the actual input data, we also foresee the on-the-fly provisioning of underlying assumptions in terms of sets of measurements, i.e. vectors $\{V_1, V_2, \ldots, V_n\}$, which in fact describe the actual contextual environment of the request.

In order to facilitate automated similarity computation between SWS and SWS requests, we extended the matchmaking capabilities of IRS-III through a set of additional functionalities:

F1. Instantiation of member M_i in CSO for each V_i provided as part of SWS_i
F2. Similarity computation between goal request SWS_i and potentially relevant SWS

Given the ontological refinement of SWS descriptions into CS as introduced in Section 3.2 this new functionality enables to automatically achieve IRS-III goals without being restricted to complete matches between a particular goal achievement request and the available SWS. When attempting to achieve a goal, our new function is provided with the actual goal request SWS_i, named *base,* and the SWS descriptions of all x available services that are potentially relevant for the base – i.e. linked through a dedicated mediator:

$$SWS_i \cup \{SWS_1, SWS_2, \ldots, SWS_x\}$$

Each SWS contains a set of concepts $C=\{c_1..c_m\}$ and instances $I=\{i_1..i_n\}$. We first identify all members $M(SWS_i)$ – in the form of valued vectors $\{v_1..v_n\}$ refining the instance i_l of the base as proposed in Section 3.2. In addition, for each concept c within the base the corresponding conceptual space representations $MS=\{MS_1..MS_m\}$ are retrieved. Similarly, for each SWS_j related to the base, prototypical members $M(SWS_j)$ – which refine capabilities of SWS_j and are represented in one of the conceptual spaces $CS_1..CS_m$, – are retrieved:

$$CS \cup M(SWS_i) \cup \{M(SWS_1), M(SWS_2), \ldots, M(SWS_x)\}$$

Based on the above ontological descriptions, for each member v_l within $M(SWS_i)$, the Euclidean distances to any prototypical member of all $M(SWS_j)$ which is represented in the same space MS_j as v_l are computed. In case one set of prototypical members $M(SWS_j)$ contains several members in the same MS – e.g. SWS_j targets several instances of the same kind – the algorithm just considers the closest distance since the closest match determines the appropriateness for a given goal. For example, if one SWS supports several different locations, just the one which is closest to the one required by SWS_i determines the appropriateness.

Consequently, a set of x sets of distances is computed as follows $Dist(SWS_i)=\{Dist(SWS_i,SWS_1), Dist(SWS_i,SWS_2) .. Dist(SWS_i,SWS_x)\}$ where each $Dist(SWS_i,SWS_j)$ contains a set of distances $\{dist_1..dist_n\}$ where any $dist_i$ represents the distance between one particular member v_i of SWS_i and one member refining one instance of the capabilities of SWS_j. Hence, the overall similarity between the base SWS_i and any SWS_j could be defined as being reciprocal to the mean value of the individual distances between all instances of their respective capability descriptions and hence, is calculated as follows:

$$Sim(SWS_i, SWS_j) = \left(\overline{Dist(SWS_i, SWS_j)}\right)^{-1} = \left(\frac{\sum_{k=1}^{n}(dist_k)}{n}\right)^{-1}$$

Finally, a set of x similarity values – computed as described above – which each indicates the similarity between the base SWS_i and one of the x target SWS is computed:

$$\{Sim(SWS_{i,}SWS_1), Sim(SWS_{i,}SWS_2),.., Sim(SWS_i, SWS_x)\}$$

As a result, the most similar SWS_j, i.e. the closest associated SWS, can be selected and invoked. In order to ensure a certain degree of overlap between the actual request and the invoked functionality, we also defined a *threshold similarity value T* which determines the similarity threshold for any potential invocation.

5 Application: Measurement-based SWS discovery of Weather Forecast Web Services

Our measurement-based SWS discovery approach (Section 4) was actualised within an initial proof-of-concept prototype application which mediates between different weather forecast Web services. This example use case illustrates how measurements can be dynamically mapped to symbolic representations, SWS in this case, by means of similarity-computation within CS.

Here, SWS_1, SWS_2 and SWS_3 provide weather forecast information for different locations. Each service has distinct constraints, and thus distinct SWS descriptions. In detail, SWS_1 is able to provide forecasts for France and Spain while SWS_2 and SWS_3 are providing forecasts for the United Kingdom. All services show different Quality of Service (QoS) parameters. Three distinct service ontologies O_1, O_2, and O_3 had been created, each defining the capability of the respective service by using distinct vocabularies. For example, SWS_2 considers concepts representing the notions of location and QoS together with corresponding instances (see also Table 1):

$$\{(country, QoS), (UK, QoS2)\} \subset O_2 \subset SWS_2$$

By applying the representational approach proposed in Section 3, each concept of the involved heterogeneous SWS representations had been refined as a shared CS, while instances - defining the capabilities of available SWS - were defined as prototypical members. For example, a simplified CS (*CS_1: Location Space* in Figure 5) was utilized to refine geographical notions (e.g. country) by using two dimensions indicating the geospatial position of the location:

$$\{(p_1 l_1, p_2 l_2)\} = \{(latitude, longitude)\} = CS_1$$

The two dimensions latitude and longitude are equally ranked, and hence, a prominence value of 1 has been applied to each dimension. Note that each of the depicted concepts and instances, such as *O_2:UK* and *O_3:UK*, are distinct and independent from each other, and thus might show heterogeneities, such as distinct labels, for instance *United Kingdom* and *Great Britain*. In the case of *O_2:UK* and *O_3:UK,* these two instances are refined by two distinct prototypical members: $L_1(SWS_2) = \{(v_1 = 55.378051, v_2 = \text{-}3.435973) | v_i \in CS_1\}$ and $L_1(SWS_3) = \{(v_1 = 55.378048, v_2 = \text{-}3.435963) | v_i \in CS_1\}$. Each member has been defined by different individuals applying similar, but non-equivalent geodata.

In addition, a second space (*CS_2: QoS Space* in Figure 5) has been defined by three dimensions – latency (in ms), throughput (number of Web services), availability (in %): $\{(p_1 r_1, p_2 r_2, p_3 r_3)\} = \{(latency, throughput, availability)\} = CS_2$

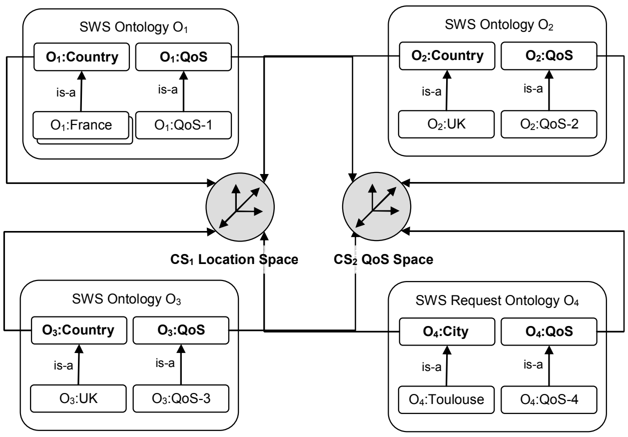

Fig. 5. Grounding assumptions of distinct weather forecast SWS to common CS.

Potential service consumers define a goal (e.g. *SWS_4* in Figure 5) together with the set of input parameters and the underlying assumptions in terms of measurements. After accomplishment of *F.1*, i.e. the dynamic instantiation of members in their corresponding CS to represent the sensor data provided with the actual goal request *SWS_4*, all involved goals and SWS were grounded in the same set of CS as depicted in Figure 5.

In that, assumptions of available SWS had been described independently in terms of simple conjunctions of instances which were individually refined in shared CS as shown in Table 1. As shown in Table 1, the request *SWS_4* assumes a SWS which provides weather forecast for the location UK (*$L_1(SWS_4)$*) and ideal QoS (*$Q_1(SWS_4)$*) demanding zero latency but high throughput and availability.

Table 1. Assumptions of involved SWS and SWS requests described in terms of vectors in MS_1 and MS_2.

	Assumption $Ass_{SWSi} = (L_{1SWSi} \cup L_{2SWSi} \cup .. \cup L_{nSWSi}) \cup (Q_{1SWSi} \cup Q_{2SWSi} \cup .. \cup Q_{mSWSi})$	
	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}$={(46.227644, 2.213755)} $L_{2(SWS1)}$={(40.463667, -3.74922)}	$Q_{1(SWS1)}$={(155, 2, 91)}
SWS_2	$L_{1(SWS2)}$={(55.378051, -3.435973)}	$Q_{1(SWS2)}$={(15, 50, 98)}
SWS_3	$L_{1(SWS3)}$={(55.378048, -3.435963)}	$Q_{1(SWS3)}$={(78, 5, 95)}
SWS_4	$L_{1(SWS4}$={(55.378048, -3.435963)}	$Q_{1(SWS4)}$={(0,100,100)}

Though no exact SWS matches these criteria, at runtime similarities are calculated between SWS_4 and the related SWS (SWS_1, SWS_2, SWS_3) through the similarity-based discovery function described in Section 4. This led to the calculation of the following similarity values:

Table 2. Automatically computed similarities between SWS request SWS_4 and available SWS.

	Similarities
SWS_1	0.010290349
SWS_2	0.038284954
SWS_3	0.016257476

Given these similarities, our introduced goal achievement method automatically selects the most similar SWS (i.e. SWS_2 in the example above) and triggers its invocation.

6 Discussion and Conclusions

In order to contribute to the vision of the SSW, i.e. the convergence of sensor data and formal knowledge representations as part of the Semantic Web, we proposed a representational model which grounds ontological representations in CS to overcome the symbol grounding problem. The latter is perceived to be as one of the major obstacles towards the SSW. While ontological instances are represented as prototypical members within a CS, arbitrary sensor data which measures the dimensions of the CS can be associated with the most appropriate instance by identifying the most similar, i.e. the closest, prototypical member to the vector which represents the sensor data. Our approach is facilitated through a dedicated CS Ontology which allows to refine any arbitrary concept (instance) as CS (prototypical member). In that, our representational model allows to bridge between sensor measurements and symbolic knowledge representations by means of similarity computation between vectors within CS.

In addition, we implemented our approach by applying it to the field of SWS and utilising it for measurement-based SWS discovery while bridging between symbolic SWS representations and sensor-based measurement data. Therefore, we extended the matchmaking algorithm of an existing SWS Broker, IRS-III, with new capabilities allowing for measurement-based matchmaking based on our two-fold representational model. A first proof-of-concept prototype application utilises our approach to enable measurement-based discovery of weather forecast Web services based on measured parameters such as the geospatial location and the service QoS.

The proposed approach has the potential to further support interoperability between heterogeneous sensor data and symbolic knowledge representations. While our approach supports automatic mapping between ontology instances and sensor-based measurements it still requires a common agreement on shared CS. In addition, incomplete similarities are computable between partially overlapping CS.

However, the authors are aware that our approach requires considerable effort to establish CS-based representations. Future work has to investigate on this effort in order to further evaluate the potential contribution of the proposed approach. Moreover, while overcoming issues introduced in Section 2, further issues remain. For example, whereas defining instances, i.e. vectors, within a given CS appears to be a straightforward process of assigning specific quantitative values to quality dimensions, the definition of the CS itself is not trivial. Nevertheless, distance calculation relies on the fact that resources are described in equivalent geometrical spaces. However, particularly with respect to the latter, traditional ontology and schema matching methods could be applied to align heterogeneous spaces. In addition, we would like to point out that the increasing usage of upper level ontologies, such as DOLCE [9] or SUMO [15], and emergence of common schemas for sensor data such as the OpenGIS Observations and Measurements Encoding Standard, leads to an increased sharing of ontologies at the concept level. As a result, our proposed hybrid representational model becomes increasingly applicable by further contributing to the vision of the SSW.

7 References

[1] Cabral, L., Domingue, J., Galizia, S., Gugliotta, A., Norton, B., Tanasescu, V., Pedrinaci, C.: IRS-III: A Broker for Semantic Web Services based Applications. In proceedings of the 5th International Semantic Web Conference (ISWC 2006), Athens, USA (2006).

[2] Cregan, A. (2007), Symbol Grounding for the Semantic Web. 4th European Semantic Web Conference 2007, Innsbruck, Austria.

[3] Dietze, S., Gugliotta, A., Domingue, J., (2008) Conceptual Situation Spaces for Situation-Driven Processes. 5th European Semantic Web Conference (ESWC), Tenerife, Spain.

[4] Dietze, S., Gugliotta, A., Domingue, J., (2008) Bridging the Gap between Mobile Application Contexts and Semantic Web Resources. Chapter in: Context-Aware Mobile and Ubiquitous Computing for Enhanced Usability: Adaptive Technologies and Applications, Editor: Dragan Stojanovic, Information Science Publishing (IGI Global), November 2008.

[5] Dietze, S., and Domingue, J. (2009) Enriching Service Semantics through Conceptual Vector Spaces, Workshop: Workshop on Ontology, Conceptualization and Epistemology for Information Systems, Software Engineering and Service Science (ONTOSE'09) at The 21st International Conference on Advanced Information Systems (CAiSE'09), Amsterdam, NL

[6] Fensel, D., Lausen, H., Polleres, A., de Bruijn, J., Stollberg, M., Roman, D., Domingue, J. (2006): Enabling Semantic Web Services – The Web service Modelling Ontology, Springer 2006.
[7] Fielding, R.T.: Architectural Styles and the Design of Network-based Software Architectures. PhD thesis, University of California, Irvine (2000)
[8] Gärdenfors, P. (2000), Conceptual Spaces - The Geometry of Thought. MIT Press, 2000.
[9] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.(2002), Sweetening Ontologies with DOLCE. In: A. Gómez-Pérez , V. Richard Benjamins (Eds.) Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web: 13th International Conference, EKAW 2002, Siguenza, Spain, October 1-4, 2002.
[10] Kaipainen, M., Normak, P., Niglas, K., Kippar, J., Laanpere, M., Soft Ontologies, spatial Representations and multi-perspective Explorability, Expert Systems, November 2008, Vol. 25, No. 5.
[11] Krause, E. F. (1987). Taxicab Geometry. Dover.
[12] Motta, E. (1998). An Overview of the OCML Modelling Language, the 8th Workshop on Methods and Languages, 1998.
[13] Nosofsky, R. (1992), Similarity, scaling and cognitive process models, Annual Review of Psychology 43, pp. 25- 53, (1992).
[14] OWL-S 1.0 Release. http://www.daml.org/services/owl-s/1.0/.
[15] Pease, A., Niles, I., Li, J.(2002), The suggested upper merged ontology: A large ontology for the semanticweb and its applications. In: AAAI-2002Workshop on Ontologies and the Semantic Web. Working Notes (2002).
[16] Raubal, M. (2004). Formalizing Conceptual Spaces. in: A. Varzi and L. Vieu (Eds.), Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004).Frontiers in Artificial Intelligence and Applications 114, pp. 153-164, IOS Press, Amsterdam, NL.
[17] Schwering, A. (2005). Hybrid Model for Semantic Similarity Measurement, in R. Meersman and Z. Tari (Eds.): CoopIS/DOA/ODBASE 2005, LNCS 3761, pp. 1449 – 1465, 2005..
[18] Sheth, A., Henson, C., Sahoo, S., Semantic Sensor Web, IEEE Internet Computing, July/August 2008, p. 78-83.
[19] Spencer, B., Liu, S. Inferring data transformation rules to integrate semantic web services. In S. A. McIlraith, D. Plexousakis, and F. van Harmelen, editors, Int'l Semantic Web Conference, volume 3298 of Lecture Notes in Computer Science, pages 456–470. Springer, 2004.
[20] Suppes, P., D. M. Krantz, et al. (1989). Foundations of Measurement - Geometrical, Threshold, and Probabilistic Representations. San Diego, California, USA, Academic Press, Inc.
[21] Vitvar, T., Kopecký, J., Viskova, J., Fensel, D. WSMO-Lite Annotations for Web Services. ESWC 2008: 674-689.
[22] World Wide Web Consortium, W3C (2004a): Resource Description Framework, W3C Recommendation 10 February 2004, http://www.w3.org/RDF/.
[23] World Wide Web Consortium, W3C (2004b): Web Ontology Language Reference, W3C Recommendation 10 February 2004, http://www.w3.org/TR/owl-ref/.
[24] WSMO Working Group, D2v1.0: Web service Modeling Ontology (WSMO). WSMO Working Draft, (2004). (http://www.wsmo.org/2004/d2/v1.0/).