Semantic Rules for Context-Aware Geographical Information Retrieval

Carsten Keßler[1], Martin Raubal[2], and Christoph Wosniok[1]

[1] Institute for Geoinformatics, University of Münster, Germany
{carsten.kessler|c.wosniok}@uni-muenster.de
[2] Department of Geography, University of California, Santa Barbara, USA
raubal@geog.ucsb.edu

Abstract. Geographical information retrieval (GIR) can benefit from context information to adapt the results to a user's current situation and personal preferences. In this respect, *semantics-based* GIR is especially challenging because context information – such as collected from sensors – is often provided through numeric values, which need to be mapped to ontological representations based on nominal symbols. The Web Ontology Language (OWL) lacks mathematical processing capabilities that require free variables, so that even basic comparisons and distance calculations are not possible. Therefore, the context information cannot be interpreted with respect to the task and the current user's preferences. In this paper, we introduce an approach based on *semantic rules* that adds these processing capabilities to OWL ontologies. The task of recommending personalized surf spots based on user location and preferences serves as a case study to evaluate the capabilities of semantic rules for context-aware geographical information retrieval. We demonstrate how the Semantic Web Rule Language (SWRL) can be utilized to model user preferences and how execution of the rules successfully retrieves surf spots that match these preferences. While SWRL itself enables free variables, mathematical functions are added via built-ins – external libraries that are dynamically loaded during rule execution. Utilizing the same mechanism, we demonstrate how SWRL built-ins can query the Semantic Sensor Web to enable the consideration of real-time measurements and thus make geographical information retrieval truly context-aware.

1 Introduction

Information retrieval methods play a crucial role when it comes to organizing and finding relevant information in large collections, potentially spread over thousands of Web servers. This applies also to *geographical* information retrieval (GIR), where innovative approaches are required to turn the vast amount of geotagged content and volunteered geographic information [1] into useful information sources. Recent research in information retrieval extends classical methods by employing context information to disambiguate queries [2] and to tailor the results to a user's preferences and her current situation [3].

GIR generally deals with indexing unstructured data sources such as Web pages for detection and disambiguation of place names, document ranking based on relevance with respect to a given query, as well as developmentof user interfaces [4]. Retrieval from the Semantic Web, however, makes use of structured information, i.e., ontologies or resources annotated by means of these ontologies [5]. This structured information is mostly based on nominal symbols encoded in the Web Ontology Language (OWL) [6]. OWL and its theoretic underpinnings in description logics (DL) provide the reasoning capabilities that enable the Semantic Web in the first place. However, they also pose a special challenge for encoding spatio-temporal information, which relies largely on numeric data collected from different kinds of sensors (e.g. for temperature or air pollution) and localized via different positioning techniques such as GPS.

In terms of measurement scales [7], ontologies are built from *nominal* entities and the relationships between them. Context information collected from sensors, in contrast, consists of *ordinal* (e.g. severe weather categories), *interval* (e.g. temperature in centigrades) and *ratio* (e.g. absolute humidity) scaled data. We propose to utilize *semantic rules* for the mapping between these two worlds as well as for the interpretation of the raw values with respect to the current user's preferences. This approach follows the distinction between task model, user model and context model[1] [8]. The Semantic Web Rule Language (SWRL) employed in this paper extends OWL with the first-order logic style rules required for user modelling. Moreover, it provides mathematic functions via *built-ins* – external libraries that are dynamically loaded during rule execution. We demonstrate how this mechanism can be utilized to query sensors on the Web directly from SWRL rules. This allows for a strict separation between static information stored in the ontology and dynamic, frequently changing context information. Moreover, rules can directly access the measured information, so that it is no longer required to manually update the information on individuals in the ontology.

In this paper, we use the task of finding suitable surf spots as a scenario, which is a prime example of a GIR task that cannot be satisfactorily solved without taking context information into account. Whether a spot is appropriate for a specific user depends not only on the user's current location (i.e., her distance to the spot), but also on the current wave and weather conditions, accessibility, crowdedness, and so forth. All of these factors must be regarded with respect to the skill level and personal preferences of the user posing the query: a suitable wave height for one person may be boring for another one, and conditions suitable for experienced surfers may even be dangerous for beginners. While the development of a user interface is out of scope for this research, we have selected four contextual aspects (location, wave height, bottom and crowdedness) that influence a user's decision during the selection of a surf spot. We have developed a prototypical surf spot ontology populated with samples at the central coast in California for demonstration purposes. The insights from this

[1] We use the term *model* in the sense of modelling information, not in the strict model-theoretic sense.

research are not specific to the task of selecting suitable surf spots, but can be easily transferred to other highly context-dependent tasks ranking from leisure activities [9] to critical professional applications [10].

The next section introduces related work on ontologies and context. Section 3 outlines the case study, the specific requirements it imposes, as well as the surf spot ontology. Section 4 presents und discusses our rule-based approach for context modelling. Section 5 demonstrates how a novel SWRL built-in can be utilized to process live sensor measurements during rule execution, followed by conclusions and directions for future work in Section 6.

2 Background: Ontologies and Context

This section points to related work in the areas of ontologies for the Semantic Web and context from the perspective of geographical information retrieval.

2.1 Ontologies for the Semantic Web

Research on semantics investigates the meaning of symbols in communication. While views on semantics differ depending on the field of study, information systems research is mostly concerned with the question of how the meaning of pieces of information and ways to interact with them can be described unambiguously. Ontologies as explicit specifications of conceptualizations [11] provide such descriptions by formally constraining the potential interpretations of the symbols used for communication about concepts.

Three kinds of ontologies can be distinguished. *Foundational ontologies* such as DOLCE [12] are "axiomatic theories of domain-independent top-level notions such as object, attribute, event, parthood, dependence, and spatio-temporal connection" [13, p. 91]. While foundational ontologies provide generic specifications on an abstract level, *domain ontologies* specify the shared conceptualization of a specific community, such as the AGROVOC Ontology currently under development by the Food and Agriculture Organization of the United Nations[2] or various ontologies within geographic information science [5, 14, 15] which is the focus of this research. *Application ontologies*, such as the surf spot ontology introduced in Section 3.3, are the most specific ones, as they specify the conceptualizations that underlie specific applications. In an ideal case, these three groups of ontologies are used in a layered fashion and *semantic interoperability* between all parts of this ontological hierarchy is enabled.

As engineering artifacts, ontologies only become useful through standardized formats. The Web Ontology Language (OWL) is a W3C specification for publishing and sharing XML-encoded ontologies. The OWL DL sublanguage[3] maps to the $\mathcal{SHOIN}$ description logic (DL). A detailed discussion of DL is beyond the scope of this paper – see [16, 17] for detailed information on its semantics. The only aspect that needs to be kept in mind here is that the formal semantics underlying $\mathcal{SHOIN}$ maps to set theory.

[2] `http://www.fao.org/aims/agrovoccs.jsp`.

[3] OWL comprises the three sublanguages OWL Lite, OWL DL and OWL Full.

2.2 Context in Information Retrieval

From the numerous definitions of context [18], we subscribe to the one by Dey [19, p. 5], who defines context as "any information that can be used to characterise the situation of an entity". This view on context matches both the motivation for this research as well as our case study from geographical information retrieval. Research activities on context-aware information retrieval have increased remarkably in recent years. The ubiquitous and pervasive computing communities have developed numerous approaches to automatically provide users with information and services based on their current situation – see [20] for an overview. Existing context-aware applications range from smart spaces [21] over mobile tour guides [22] to generic prototyping platforms [23]. Gu et al. [24] propose to use an OWL ontology to model the different aspects forming a context. Lawrence [25] provides an overview of the different strategies to make Web search context-aware.

The Semantic Web community has investigated different approaches towards semantic models of context and contextualizing ontologies. CTXML [26] is an XML dialect that enables local, contextualized models within ontologies. It specifies an XML-encoding for contexts as concept hierarchies that can be assigned information on the owner, the groups the context is specified for, and on security and history. Mappings between different contexts based on distributed description logics [27] allow for the identification of different specifications that refer to the same concepts. CTXML eventually merged into C-OWL [28], which is an extension to OWL that enables contextual ontologies via bridge rules that map between the local ontologies. C-OWL thus adds the capability to handle potentially inconsistent knowledge. Korpipää and Mäntyjärvi [29] propose an ontology for contexts based on RDF. It is based on context types (e.g. Environment:Light:Intensity) that can be filled with values (e.g. *Dark*). However, they do not describe how the mapping from the raw sensor data onto the context values (*Dark* in the above example) works.

The use of SWRL for reasoning about context has already been proposed within the A-MUSE project [30]. However, the utilization of SWRL within the project was limited to reasoning about context information in the ontology. For example, SWRL was used to find out whether a person is currently attending a meeting in a certain room based on the property values of the individuals in the knowledge base. We extend this approach by using SWRL to *represent* context (instead of only reasoning about it) and by querying context information directly from a novel SWRL built-in.

3 Case Study: The Surf Spot Finder

This section introduces the case study used in the remainder of this paper. The specific requirements of an application for surf spot recommendation are analyzed and the geographic feature type ontology is sketched.

3.1 Scenario

Choosing a surf spot depends on weather conditions, tide, waves, personal abilities, location and social aspects. All of these characteristics, including their parts, must be rated differently for every surfer. Answering the question "How are the waves?" involves wave height, speed and frequency, but is also strongly tied to the location, type of sea bed and beach accessibility. Although some Web sites provide relevant information for surfers on wave[4] or weather[5] conditions, such information must be rated differently depending on the respective user's personal surfing skills. For example, for an experienced surfer the conditions may be rather rough and in contrast to conditions suitable for a beginner. This leads to the conclusion that there is no generic answer to the question "Where is my optimal surf spot at a given time?". In addition, social aspects play an important role. While a surfer may want to meet a friend to practice, most surfers prefer spots that are not too crowded.

In the remainder of this paper, we will use the scenario of an online surf spot finder service that provides users with surf spot recommendations based on the current conditions and their personal profiles. While the technical aspects of this tool are out of scope for this paper, we investigate the requirements for a representation of contextual factors relevant to this task. As such, the surf spot finder is a prime example of context-aware geographical information retrieval. The insights gained in this study can be transferred to a number of outdoor leisure activities such as hiking, diving and climbing [9], but also to critical applications that depend on the current conditions and skills of the people involved, such as emergency response services [10].

3.2 Requirements

We have selected four contextual aspects to illustrate the modeling challenges with respect to context-aware GIR. The selection also reflects the temporal, spatial and thematic components that characterize geographic information [31]:

1. The *Location* of the surf spot influences the choice depending on the maximum distance the surfer is willing to travel.
2. The current *Wave height* plays an important role with respect to the surfer's skills. This dynamic aspect must be measured by sensors at regular intervals (e.g. those of the Coastal Data Information Program[6]).
3. The *Bottom* points to potential dangers: sandy bottoms are safer for beginners, whereas rocky bottoms can be dangerous, but may be acceptable for experienced surfers.
4. The current *Crowdedness* is a social aspect influencing the decision process, as surfers mostly prefer less crowded spots. Although this aspect is inherently dynamic, real-time measurement seems impractical; we thus model crowdedness distinguishing between weekdays and weekends.

[4] `http://magicseaweed.com` or `http://www.wannasurf.com`.
[5] `http://www.windguru.com`.
[6] `http://gcmd.nasa.gov/records/GCMD_SIO_CCS_CDIP_CALSWELLMODEL.html`.

While these aspects will be covered in the surf spot ontology introduced in the next section, the user models will be specified through SWRL rules as discussed in Section 4. For demonstration purposes, we define two different users with different preferences concerning the four properties of the surf spots described above. In the following, u_{beg} will refer to a user model corresponding to a typical beginner, whereas u_{pro} will refer to the model of an experienced or professional surfer. Table 1 provides an overview of the two user models. We assume that both users are located in downtown Santa Barbara (lat:34.41213°, lon:−119.68913°).

	Location	**WaveHeight**	**Bottom**	**Crowdedness**
User u_{beg}	$\leq 0.5°$ away	< 1.5m	Sandy	No preference
User u_{pro}	$\leq 1.5°$ away	> 2.5m	No preference	Low

Table 1. Overview of the user preferences to be formalized in Section 4. For reasons of simplicity, the distance to the user's location is expressed in decimal degrees; a representation in km (road network distance) would be required for end users.

3.3 Surf Spot Ontology

We have developed a surf spot ontology[7] for this research using the Protégé[8] ontology editor. The ontology consists of a number of concepts such as Surfing (subconcept of Activity), WaveFrequency and RideLength in addition to the four properties of specific surf spots listed in Section 3.2. Figure 1 gives an overview of the concepts specified in the ontology.

4 Representing Context

This section introduces the rule-based mapping approach and presents an implementation of the user models for the surf spot finder case study in SWRL.

4.1 Rule-based Mapping Approach

Context information only becomes meaningful when it is interpreted with respect to the user. It depends on the user's personal preferences whether a shop has interesting offers, a blog article is worth reading or a surf spot is adequate with respect to the user's expectations and abilities. However, ontologies on the Semantic Web lack support for free variables that are required to express such dependency on the user profile. It is not possible to specify that a user finds surf spots with rocky bottoms too dangerous, or that another one finds waves

[7] Available from `http://ifgi.uni-muenster.de/~lette/resources/surfing.zip`.
[8] Protégé is free and open source software; see `http://protege.stanford.edu`.

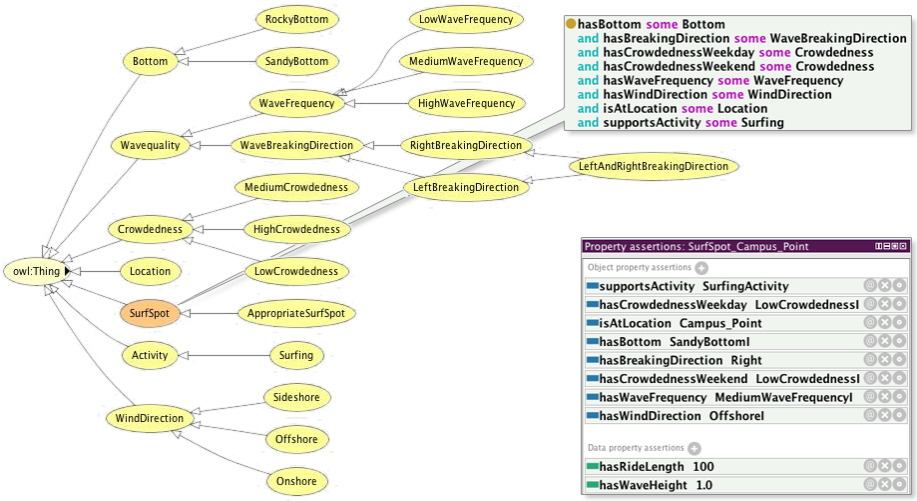

Fig. 1. The graph shows the concept hierarchy of the surf spot ontology. The box on top shows the properties of the central SurfSpot concept. The box at the bottom shows the properties of the SurfSpot individual *SurfSpot_Campus_Point*.

below 2m unattractive. We propose to utilize the Semantic Web Rule Language (SWRL) for personalized mappings between the numeric sensor world and information stored in ontologies, as shown in Figure 2.

SWRL [32] adds rule capabilities to OWL ontologies. SWRL rules are Horn-like clauses [33], consisting of an antecedent (body) and a consequent (head), each containing a set of atoms. For example, the rule

`hasParent(?x,?y)` $\wedge$ `hasBrother(?y,?z)` $\rightarrow$ `hasUncle(?x,?z)`

states that if `x` has the parent `y` and `y` has the brother `z`, then `x` has the uncle `z`. The SWRL submission to W3C introduces a concrete XML syntax for use with OWL, which is achieved by combination with RuleML[9] statements. SWRL can be extended by additional reasoning mechanisms via built-ins. Among the built-ins that have been implemented so far are libraries for mathematics, string handling and date support, enabling rules such as

`Person(?p)` $\wedge$ `hasAge(?p,?a)` $\wedge$ `swrlb:lessThan(?a,18)` $\rightarrow$ `Minor(?p)`.

Built-ins thus play a central role for the mapping from numeric sensor inputs to ontologies via SWRL. The following section demonstrates how SWRL rules using built-ins can be utilized to represent a user model as shown in Figure 3.

4.2 User Model Formalization in SWRL

The antecedent of a SWRL rule can be used to represent a conjunction of the different aspects of a user's preferences. Consider the following SWRL rule in prefix syntax, which is a complete user model for u_{beg} as described in Table 1:

[9] `http://www.ruleml.org/`.

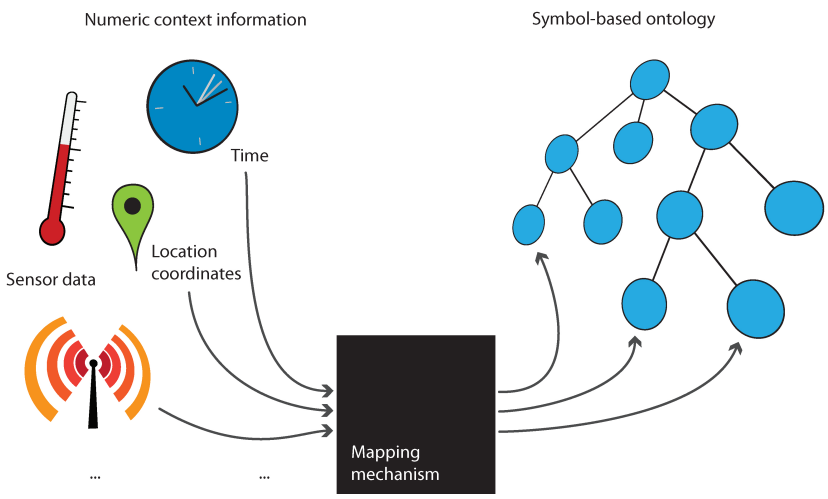

Fig. 2. Context-aware geographical information retrieval on the Semantic Web requires a mapping from the context model (left) to the static ontological information (right).

```
01   SurfSpot(?spot) ∧
02   hasWaveHeight(?spot, ?height) ∧
03   swrlb:lessThan(?height, 1.5) ∧
04   isAtLocation(?spot, ?location) ∧
05   hasLatitude(?location, ?lat) ∧
06   hasLongitude(?location, ?lon) ∧
07   swrlb:subtract(?distLat, ?lat, 34.412132) ∧
08   swrlb:subtract(?distLon, ?lon, -119.68913) ∧
09   swrlb:abs(?distLatAbs, ?distLat) ∧
10   swrlb:abs(?distLonAbs, ?distLon) ∧
11   swrlb:lessThanOrEqual(?distLatAbs, 0.5) ∧
12   swrlb:lessThanOrEqual(?distLonAbs, 0.5) ∧
13   hasBottom(?spot, ?bottom) ∧
14   SandyBottom(?bottom)
15   → AppropriateSurfSpot(?spot)
```

The first line constrains the individuals to whom this rule is applied to the ones of the class `SurfSpot` and assigns the variable name `?spot` to them. Lines 2–3 constrain the matching individuals to those that have a wave height smaller than 1.5m. Lines 5–12 take care of the distance calculation of the spots to the user. After assigning variable names to the latitude and longitude of a spot (5–6), the absolute distance to the user's current location is calculated (7–10) to evaluate whether the spot is not further away than the user is willing to travel[10] (11–12). The user preference for surf spots with sandy bottoms is expressed

[10] For a lack of space, distance is calculated separately for latitude and longitude. However, the calculation in Euclidian (or any other) metric is also possible in SWRL.

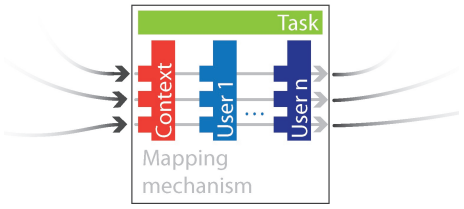

Fig. 3. The contents of the mapping mechanism black box shown in Figure 2: The mapping consists of task model (green, not further specified here), context model (red) and user models (blue). The context model provides access to dynamically changing information from sensors as described in Section 5, which is then fed into the user models. The user models, each consisting of a SWRL rule, filter the incoming context information according to the users' preferences.

in lines 13–14. Finally, those surf spots satisfying lines 2–14 are reclassified as `AppropriateSurfSpot` (see Figure 1). Instead of this reclassification, one could also imagine using query built-ins that count the number of spots matching the conditions. The user model for u_{pro} has been defined accordingly:

```
01   SurfSpot(?spot) ∧
02   hasWaveHeight(?spot, ?height) ∧
03   swrlb:greaterThan(?height, 2.5) ∧
04   isAtLocation(?spot, ?location) ∧
05   hasLatitude(?location, ?lat) ∧
06   hasLongitude(?location, ?lon) ∧
07   swrlb:subtract(?distLat, ?lat, 34.412132) ∧
08   swrlb:subtract(?distLon, ?lon, -119.68913) ∧
09   swrlb:abs(?distLatAbs, ?distLat) ∧
10   swrlb:abs(?distLonAbs, ?distLon) ∧
11   swrlb:lessThanOrEqual(?distLatAbs, 1.5) ∧
12   swrlb:lessThanOrEqual(?distLonAbs, 1.5) ∧
13   hasCrowdednessWeekend(?spot, ?crowdedness) ∧
14   LowCrowdedness(?crowdedness)
15   → AppropriateSurfSpot(?spot)
```

This rule implicitly contains temporal information on the weekday in line 13, which could also be integrated automatically via SWRL's date and time built-ins. Either of the two user models can be activated by selecting the corresponding rule. Activation of both rules would find all spots that suit both users. This feature could be used to accommodate the needs of user groups. The SWRL-tab available in the Protégé ontology editor provides a user-friendly way of developing, editing and testing[11] SWRL rules along with the corresponding OWL ontology [34].

[11] Using the jess rule engine, see `http://www.jessrules.com/`.

4.3 Discussion

After applying the two rules for the respective user context models, the Jess rule engine correctly reclassified the matching surf spots. For each user, only one of the 15 individuals of the surf spot class satisfied all given requirements. In the following, we discuss the advantages of a rule-based approach and specifically SWRL with its built-in mechanism over other Semantic Web technologies.

As mentioned before, widespread Semantic Web technologies lack support for free variables as required for the representation of a threefold model consisting of task, user and context as shown in Figure 3. This applies both to RDF/OWL (including C-OWL) as well as to alternative approaches such as the ISO Topic Map standard [35]. Although the languages have not been developed to support such functionality initially, these restrictions place constraints on their applicability to real-world problems. Their underlying realist view on semantics [36] ignores the fact that a large number of applications in information retrieval, linguistics or location-based services, for example, can only produce useful results if they take the user's current context into account. Even the context extension to OWL, C-OWL, is limited to fixed contexts (i.e., local micro theories) and mappings between them. It thus cannot live up to the requirements of the notion of *dynamic context awareness* [37] from AI, which heavily relies on sensor input. The same applies to Topic Maps, which allow users to define *scopes* specifying in which context assertions are valid [38]. Scopes form a mechanism that is very useful when stating assertions such as the topic name *München* is only valid in the scope of the German language or the topic name *Jaguar* appears in different scopes (animal, car, operating system). However, same as for C-OWL, scopes neither support rules nor free variables. Overall, these limitations result in the rare use of semantic technologies in context-aware software in AI, which is still largely based on statistics-driven approaches.

As demonstrated above, SWRL helps to overcome these limitations by specifying user models as rules that contextualize OWL ontologies in terms of these models. Moreover, modelling in OWL and SWRL requires a thoughtful separation between the notions modeled in OWL and those modeled in SWRL, forcing clean ontological specifications. For example, consider the data property assertions shown in the bottom right box of Figure 1: any of these could also be modeled as object property assertions. This would require, for example, to define a concept *RideLength* through subconcepts such as *RideLengthMax50m*, *RideLengthMax100m*, and so forth. While this approach appears feasible at first sight, it would render mathematical processing of the actual values impossible, such as checking whether the waves are too high (or too low) for the current user. Hence, our approach forces a clean ontology design based on the distinction between the static information stored in the ontology, and the dynamic aspects which go into the rules. This separation may even become more important when decidability plays a role: since the integration of SWRL renders OWL undecidable [39], it may be desirable to retain the static part of the ontology for reasoning purposes. Nonetheless, context information about individual surf spots

is still hard-coded into the ontology. We address this problem in the following section by establishing a link to real sensors.

5 Linking Rules to Real-world Sensors

We have demonstrated how SWRL rules can be employed as user models. Dynamic values, however, still have to be updated manually for an ontology's individuals. In this section, we sketch an approach to this problem that enables querying of sensors from within SWRL rules.

5.1 The SWRL Built-in Bridge

The SWRLTab in Protégé provides a built-in bridge[12], which enables the development of new built-ins, such as for reasoning (temporal, spatial, etc.), mathematical calculations and querying. Technically, new built-ins must be individuals of the class *BuiltIn*, which is defined in the SWRL ontology. This ontology has to be imported to enable the definition of rules. The processing functionality that evaluates the single built-ins (such as *swrlb:lessThan*) must be implemented in Java and is dynamically loaded by Protégé.

As such, new SWRL built-ins bear the potential to link OWL ontologies to external information such as databases – or (geo)sensors [40]. Querying built-ins already exist for (properties of) individuals in the ontology. In the following, we outline the functioning of a new built-in to collect observations from the Sensor Web, which is currently under development. The presented approach can be regarded as a more light-weight variant of the Semantic Sensor Web [41], which builds on fully annotated sensor web services.

5.2 Querying Sensors from SWRL

The Open Geospatial Consortium has introduced a number of service specifications to enable a fully interoperable Sensor Web. One of these standards is the Sensor Observation Service (SOS) [42], which allows clients to request observations for a specific sensor via the `GetObservation` request. Following the idea of "SQWRL for sensors", we are currently developing the `sos:obs` built-in. While the process of service discovery and description (via the SOS' `GetCapabilities` request) is beyond the scope of this paper, Figure 4 shows an overview of the workflow of an SWRL-enabled OWL-ontology using the `sos:obs` built-in. Note that in combination with a user model as implemented in Section 4, the static information about individuals in the ontology is completely separated from any dynamically changing information collected from sensors, as well as the user model.

Using `sos:obs` requires the ontology to contain an individual representing the SOS to be queried, which holds information on the service's capabilities. This

[12] `http://protege.cim3.net/cgi-bin/wiki.pl?SWRLBuiltInBridge`.

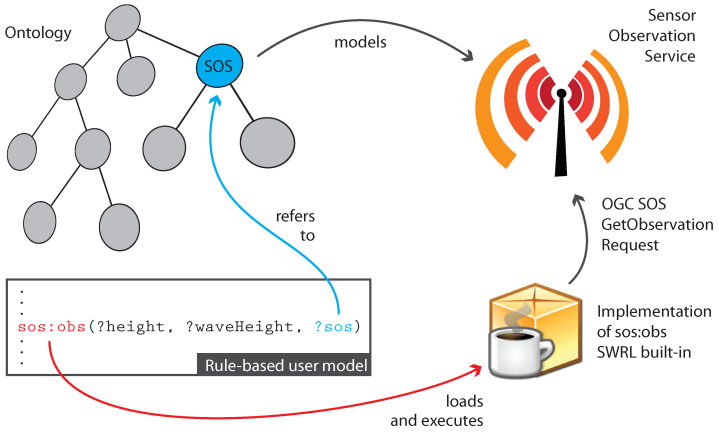

Fig. 4. Workflow of the `sos:obs` built-in: An SOS has to be modeled in the ontology, so that it can be referred to from SWRL rules. During rule execution, the implementation of the built-in is loaded and executed. It queries the SOS and returns the value to the rule engine for further processing.

instance is an individual of the class generated from the SOS XML schema[13]. The usage pattern for the built-in binds the value observed for the given property (`?observedProperty`) of an SOS individual (`?sos`) to `?observedValue`:

```
sos:obs(?observedValue, ?observedProperty, ?sos)
```

With this built-in, it is possible to remove any dynamically changing information represented from the surf spot individuals. The ontology then only comprises static information about the individual surf spots as well as the sensors measuring the dynamic aspects of the surf spots. The dynamic context information gets requested on demand, i.e. when the rules representing the user model are executed. Accordingly, the user model for u_{beg} is expressed through the following rule (the part of the code not shown here remains unchanged):

```
01   SurfSpot(?spot) ∧
02   hasSOS(?spot, ?sos) ∧
03   hasObservedProperty(?sos, ?waveHeight) ∧
04   sos:obs(?height, ?waveHeight, ?sos) ∧
05   swrlb:lessThan(?height, 1.5) ∧
⋮
17   → AppropriateSurfSpot(?spot)
```

[13] `http://schemas.opengis.net/sos/1.0.0/sosGetCapabilities.xsd`.

6 Conclusions and Future Work

The Semantic Web is currently based on nominal data and it is therefore inherently difficult to integrate information from the Sensor Web. In this paper, we have demonstrated how semantic rules can be employed to bridge these two worlds to enable context-aware geographical information retrieval from the Semantic Web. Using the Semantic Web Rule Language (SWRL), we showed how the user aspect in a context model can be fully represented in a SWRL rule. SWRL's support for free variables allows for a reclassification of the individuals in an ontology – in our example, surf spots matching the conditions of a user model were correctly reclassified as appropriate with respect to the given user. Moreover, we have outlined an approach to dynamically read observation values from the Sensor Web during execution of a rule based on SWRL built-ins. This approach allows for a strict separation of static knowledge about the individuals in an ontology and any dynamic information through an explicit link to sensors.

Since the built-in is still under development, its completion is the next step in this research. Execution time is thereby a crucial aspect: any rule making use of the built-in will have to wait for the sensor to deliver the data in order to execute the rest of the rule. Caching mechanisms may be required to speed up processing, especially for observed values that do not change very fast. For phenomena such as wave height, hourly observation is certainly sufficient (at least for the surfing scenario; a tsunami warning system, in contrast, would require more frequent updates). Such large-scale systems also raise the question of scalability. Once implemented, a system with a broad user basis could also be employed to apply standard information retrieval measures such as recall and precision.

While our approach to modeling context is independent of the employed information retrieval method, the background for this research is in context modeling for semantic similarity measurement [43, 44]. Previous approaches to enable context-aware similarity measurement [45, 46] regarded context as a subset of the static information at hand to enhance the cognitive plausibility of the results [47]. Combinations of semantic rules and similarity measurement show promise for adaptation of information retrieval results to the current context that goes beyond previous approaches. One could think of a reclassification of the ontology via rules before measuring similarity, which would put the ontology in the context of the current similarity query. Vice versa, subsets of the ontology which are most similar with respect to a given query could serve as input to rules, which could then adapt these preliminary results to the current context. As such, the approach discussed in this paper is not specific to geographical information retrieval; nonetheless, this field with its broad range of numeric base data provides numerous use cases for its application.

From a theoretical point of view, the problem of monotonicity needs to be addressed: depending on the sensor measurements imported via `sos:obs`, the same reasoning steps may lead to different results. Concerning SWRL, the number of tools for rule engineering and execution is still very limited. In practice, the rules quickly become verbose, even with the prefix syntax used in this paper. The XML-syntax used for storage is even more extensive, so that neither of

the two formats is adequate for interaction of non-experts. These users, such as surfers in our scenario, cannot be expected to edit rules by hand. Instead, user interfaces that hide the complexity of the rules from the user are required.

Acknowledgements

This research is partly funded by the SimCat project (DFG Ra1062/2-1 and DFG Ja1709/2-2), see `http://sim-dl.sourceforge.net`.

References

1. Goodchild, M.F.: Citizens as voluntary sensors: Spatial data infrastructure in the world of web 2.0. International Journal of Spatial Data Infrastructures Research **2** (2007) 24–32
2. Kraft, R., Chang, C.C., Maghoul, F., Kumar, R.: Searching with context. In: WWW '06: Proceedings of the 15th international conference on World Wide Web, New York, NY, USA, ACM Press (2006) 477–486
3. Brown, P.J., Jones, G.J.F.: Context-aware retrieval: Exploring a new environment for information retrieval and information filtering. Personal and Ubiquitous Computing **5** (2001) 253–263
4. Jones, C.B., Purves, R.S.: Geographical information retrieval. International Journal of Geographical Information Science **22**(3) (2008) 219–228
5. Abdelmoty, A.I., Smart, P.D., Jones, C.B., Fu, G., Finch, D.: A critical evaluation of ontology languages for geographical information retrieval. Journal of Visual Languages and Computing **16**(4) (2005) 331–358
6. Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D., Patel-Schneider, P.F., Stein, L.: OWL web ontology language reference. W3C recommendation, W3C (February 2004)
7. Stevens, S.S.: On the theory of scales of measurement. Science **103**(2684) (1946) 677–680
8. Raubal, M., Panov, I.: A formal model for mobile map adaptation. In Gartner, G., Rehrl, K., eds.: Location Based Services and TeleCartography II – From Sensor Fusion to Context Models. Selected Papers from the 5th International Symposium on LBS & TeleCartography 2008, Salzburg, Austria. Lecture Notes in Geoinformation and Cartography, Berlin, Springer (2009) 11–34
9. Wilkes, M., Janowicz, K.: A graph-based alignment approach to similarity between climbing routes. In: First International Workshop on Information Semantics and its Implications for Geographic Analysis (ISGA '08) at GIScience 2008. (2008)
10. Rinner, C.: Multi-criteria evaluation in support of emergency response decision-making. In: Joint CIG/ISPRS Conference on Geomatics for Disaster and Risk Management. (2007)
11. Gruber, T.: A translation approach to portable ontology specifications. Knowledge Acquisition **5**(2) (1993) 199–220
12. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., Schneider, L.: WonderWeb deliverable D17. the WonderWeb library of foundational ontologies and the DOLCE ontology. Technical report, ISTC-CNR (2002)
13. Schneider, L.: Designing foundational ontologies. In Song, I.Y., Liddle, S.W., Ling, T.W., Scheuermann, P., eds.: Conceptual Modeling – ER 2003. LNCS 2813, Springer, Berlin (2003) 91–104

14. Mark, D., Smith, B., Tversky, B.: Ontology and geographic objects: An empirical study of cognitive categorization. In Freksa, C., Mark, D.M., eds.: Spatial Information Theory: A Theoretical Basis for GIS. LNCS 1661, Springer (1999) 283–298
15. Frank, A.: Ontology for spatio-temporal databases. In Koubarakis, M., et al., eds.: Spatio-Temporal Databases – The CHOROCHRONOS Approach. LNCS 2520. Springer (2003) 9–77
16. Horrocks, I., Patel-Schneider, P.F., van Harmelen, F.: From $\mathcal{SHIQ}$ and RDF to OWL: The making of a web ontology language. Journal of Web Semantics **1**(1) (2003) 7–26
17. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F., eds.: The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2003)
18. Bazire, M., Brézillon, P.: Understanding context before using it. In Dey, A.K., Kokinov, B., Leake, D., Turner, R., eds.: Modeling and Using Context - 5thInternational and Interdisciplinary Conference CONTEXT 2005. LNCS 3554. Springer, Paris, France (2005) 29–40
19. Dey, A.K.: Understanding and using context. Personal Ubiquitous Computing **5**(1) (February 2001) 4–7
20. Loke, S.: Context-Aware Pervasive Systems: Architectures for a New Breed of Applications. Auerbach Publications (2006)
21. Wang, X., Dong, J.S., Chin, C., Hettiarachchi, S., Zhang, D.: Semantic space: An infrastructure for smart spaces. IEEE Pervasive Computing **3**(3) (July 2004) 32–39
22. Chou, S., Hsieh, W., Gandon, F., Sadeh, N.: Semantic Web technologies for context-aware museum tour guide applications. In: Advanced Information Networking and Applications, 2005. AINA 2005. 19th International Conference on. Volume 2. (2005)
23. Raento, M., Oulasvirta, A., Petit, R., Toivonen, H.: ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications. IEEE Pervasive Computing **4**(2) (2005) 51–59
24. Gu, T., Wang, X., Pung, H., Zhang, D.: An Ontology-based Context Model in Intelligent Environments. In: Proceedings of Communication Networks and Distributed Systems Modeling and Simulation Conference. (2004)
25. Lawrence, S.: Context in web search. IEEE Data Engineering Bulletin **23**(3) (2000) 25–32
26. Bouquet, P., Donà, A., Serafini, L., Zanobini, S.: ConTeXtualized local ontologies specification via CTXML. In: AAAI-02 workshop on Meaning Negotiation. Volume 28. (July 2002)
27. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. Journal on Data Semantics **1** (2003) 153–184
28. Bouquet, P., Giunchiglia, F., Harmelen, F., Serafini, L., Stuckenschmidt, H.: C-OWL: Contextualizing ontologies. In: The Semantic Web – ISWC 2003. LNCS 2870, Springer (2003) 164–179
29. Korpipää, P., Mäntyjärvi, J.: An ontology for mobile device sensor-based context awareness. In Blackburn, P., Ghidini, C., Turner, R.M., Giunchiglia, F., eds.: Modeling and Using Context. LNAI 2680, Springer (2003) 451–458
30. Zwaal, H., Hutschemaekers, M., Verheijen, M.: Manipulating context information with SWRL (A-MUSE Deliverable D3.12). Online: `https://doc.telin.nl/dscgi/ds.py/Get/File-62333/Manipulating%20context%20information%20with%20SWRL` (last accessed February 22, 2009) (2006)

31. Goodchild, M.F., Egenhofer, M.J., Kemp, K.K., Mark, D.M., Sheppard, E.: Introduction to the Varenius project. International Journal of Geographical Information Science **13**(8) (1999) 731–745
32. Horrocks, I., Patel-Schneider, P.F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A semantic web rule language combining OWL and RuleML. W3C Member Submission 21 May 2004, available online at `http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/` (last accessed Feb. 4, 2009) (2004)
33. Horn, A.: On sentences which are true of direct unions of algebras. Journal of Symbolic Logic **16** (1951) 14–21
34. O'Connor, M., Knublauch, H., Tu, S., Grosof, B., Dean, M., Grosso, W., Musen, M.: Supporting rule system interoperability on the semantic web with SWRL. In Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A., eds.: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland. LNCS 3729 974–986
35. Biezunski, M., Bryan, M., Newcomb, S.R.: ISO/IEC 13250:2000 topic maps: Information technology – document description and markup languages. Online: `http://www.y12.doe.gov/sgml/sc34/document/0129.pdf` (last accessed Feb. 10, 2009) (December 1999)
36. Gärdenfors, P.: Conceptual Spaces - The Geometry of Thought. Bradford Books, MIT Press, Cambridge, MA (2000)
37. Mascolo, C., Capra, L., Emmerich, W.: Mobile computing middleware. In Gregori, E., Anastasi, G., Basagni, S., eds.: Advanced lectures on networking – NETWORKING 2002 Tutorials, Springer-Verlag New York (2002) 20–58
38. Pepper, S., Grønmo, G.O.: Towards a general theory of scope. In: Extreme Markup Languages 2001, Montréal, Québec August 14-17. (2001)
39. Horrocks, I.: OWL rules, OK? In: W3C Workshop on Rule Languages for Interoperability. (2005)
40. Stefanidis, A., Nittel, S.: GeoSensor Networks. CRC Press, Boca Raton, FL (2004)
41. Sheth, A., Henson, C., Sahoo, S.S.: Semantic sensor web. Internet Computing, IEEE **12**(4) (2008) 78–83
42. Open Geospatial Consortium Inc.: Sensor observation service. OpenGIS implementation standard, OGC 06-009r6. Technical report (2007)
43. Keßler, C., Raubal, M., Janowicz, K.: The effect of context on semantic similarity measurement. In Meersman, R., Tari, Z., Herrero, P., Et, eds.: On The Move - OTM 2007 Workshops, Part II. Lecture Notes in Computer Science 4806, Springer-Verlag Berlin Heidelberg (2007) 1274–1284
44. Janowicz, K.: Kinds of contexts and their impact on semantic similarity measurement. In: 5th IEEE Workshop on Context Modeling and Reasoning (CoMoRea 2008); Hong Kong, March 17–21, 2008, IEEE Computer Society (2008) 441–446
45. Rodríguez, A., Egenhofer, M.: Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. International Journal of Geographical Information Science **18**(3) (2004) 229–256
46. Janowicz, K., Keßler, C., Schwarz, M., Wilkes, M., Panov, I., Espeter, M., Bäumer, B.: Algorithm, implementation and application of the SIM-DL similarity server. In Fonseca, F., Rodriguez, M.A., eds.: Second International Conference on GeoSpatial Semantics, GeoS 2007. LNCS 4853, Springer 128–145
47. Janowicz, K., Keßler, C., Panov, I., Wilkes, M., Espeter, M., Schwarz, M.: A study on the cognitive plausibility of SIM-DL similarity rankings for geographic feature types. In Bernard, L., Friis-Christensen, A., Pundt, H., eds.: The European Information Society. Proceedings of the 11th AGILE International Conference on Geographic Information Science. Lecture Notes on Geo-Information, Springer (2008) 115–134