An example of food ontology for diabetes control

Jaime Cantais[1], David Dominguez[2], Valeria Gigante[3],
Loredana Laera[1], and Valentina Tamma

[1] Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, UK
[2] ITACA, Universidad Politecnica de Valencia, 46520 Valencia, Spain
[3] Istituto Scientifico Universitario San Raffaele, Via Olgettina, 60, 20132 Milan, Italy

Abstract. This paper describes our experience in the rapid prototyping of a food ontology oriented to the nutritional and health care domain that is used to share knowledge between the different stakeholders involved in the PIPS project.

1 Introduction

PIPS (Personalised Information Platform for Health and Life Services) is an E-health Integrated project funded by the European Commission under the Framework 6 call, that aims to create novel healthcare delivery models by building an environment for Health, and Knowledge Services Support. This environment integrates different technologies in order to enable healthcare professionals to get access to relevant, updated medical knowledge, and European citizens to choose healthier lifestyles. The project aims to bring together healthcare suppliers, citizens, public organizations, food/drug industry and services, researchers, and health related policy makers in order to create a dynamic knowledge environment. This dynamic environment builds on traditional and new approaches for handling knowledge from current medical practice, evidence based medicine, and disparate knowledge sources from health/nutrition domains.

The philosophy underlying PIPS is to provide an integrated environment that enables the interaction of different types of users with conventional computers as well as small, and ubiquitous devices, such as mobile phones, and medical devices, at the aim to provide them with personalised advice. The PIPS platform combines a number of technologies in order to generate personalised advice, such as software agents, intelligent decision making, natural language generation, and knowledge management. This paper focus in particular on this last aspect: managing heterogeneous knowledge from different sources is one of the primary tasks in PIPS and this knowledge and the inferences performed on it are the building blocks used to generate personalised advice. Resources include structured, semi-structured and unstructured data; for this reason we take an ontological approach to sharing and reconciliation, whereby shared ontologies [1] are used to achieve a common understanding of the domains in which the system operates.

The general domain we consider is that of well-being and lifestyle, and the various project partners are each contributing different types of expertise to a specific aspect of the knowledge base we are modelling. Several different types of knowledge contribute to the domain: medical knowledge, knowledge about food and nutrition, about patients, and their clinical records, products and treatments. In order to share and combine all these aspects, we need to model ontologies for each of the topics involved, but also, and more interestingly, we need to model the different interactions between these areas; for example we aim to model how the type of nutrition affects the health of a person, and how the type of nutrition should change depending on the conditions affecting a patient. In this paper we present one of the ontologies we designed for the PIPS project, the Food Ontology.

2 The problem: Guiding a diabetic patient in the choice of food

Two different scenarios were designed in order to develop a prototype that would probe all the future capabilities of the system, although only one of them is related to the work described in this paper. This section gives a complete description of one of the two scenarios storyboards. In our scenario the main actor is Mary Johnson, 35 years old, a patient that presents an history of diabetes of type one, and is insulin-dependent. Mary will be guided through all the situations in these scenarios by the PIPS assistant, that is the interface between the patient and the PIPS system. The description of the system is outside the scope of this paper, for which it is sufficient to say that its components include a knowledge management module that stores a knowledge base of patient data (composed by ontologies and instance sets) and it is able to reason with ontological definitions as well as with values (T-box reasoning and A-box reasoning on OWL ontologies).

In our scenario, Mary is in the supermarket doing her shopping. She has a mobile phone with Internet capabilities, that includes a normal web browser. She connects to the PIPS portal using her mobile phone, and she checks her diet for the day, in order to buy the correct ingredients for the recommended meals. Mary is not in too keen on cooking, and she would rather have a ready meal of spaghetti bolognese. She then chooses a packet from a shelf. The specific product seems, at a first glance, to be compatible with Mary's diet. However, Mary wants to be sure that the chosen product will not cause any side effect to her condition. She types the bar code of the product in the mobile phone and the system retrieves all the nutritional information related to that product from the knowledge base. In addition to retrieving the nutritional information of commercial products, the system is able to compare them with the specific nutritional plan answering her needs, and to suggest the size of the portion she is allowed to eat safely. In order to perform this comparison, the system needs to perform some reasoning with the concepts and relationships defined in three ontologies, namely the *Diets ontology*, the *Product ontology* and the *Food ontology*, that one we review in this paper.

3 The Food Ontology: existing resources and related efforts

The aim of Food Ontology is to represent an abstract model of the different types of foods available to the PIPS users, together with their nutritional information, including the type and amount of nutrients, and the recommended daily intake.

There are a number of existing coding systems that have been devised to classify foods and their nutritional properties, and several databases developed with the same purpose. However, very few ontological resources exist that describe food. The most renown food ontology is the *Wine and Food Ontology* [4]. This ontology was designed to match recipes with the most suitable wine, and it does not model any information regarding nutritional facts. In order to build an ontology of food for PIPS, we based our model on the Eurocode2 food categories [5]. Eurocode2 is a food coding system originally developed within the European FLAIR Eurofoods-Enfant Project with the aim to serve as a standard instrument for nutritional surveys in Europe, and provide food property information to be used when comparing different food intakes. The terms modelled in Eurocode 2 have been integrated with the food database developed by one of the PIPS partners, ITACA, a spin off of the Polytechnic University of Valencia specialised in data management for health care applications. This database holds nutritional information about different kinds of food as well as a mapping between these foods and their classification in Eurocode2.

4 Proposed formalisation

Before proceeding with the description of the ontology, we briefly described here the development process we used in order to design the Food ontology. We outline the development process because it helps to understand the scope of the ontology. We then proceed in describing the main concepts described in the ontology and illustrate them through Protégé screenshots. We do not include in this paper the OWL ontology because of lack of space, however the most recent version of the OWL ontology can be downloaded from `http://www.csc.liv.ac.uk/~jcantais/PIPSFood.owl`.

4.1 The development process

As mentioned in the previous section, the food ontology aims to represent the nutritional aspects of the different types of food available to the PIPS stakeholders. The ontology was developed through a collaborative process that included domain experts, database experts and ontology engineers. Ontologies in PIPS undergo fast prototyping and deployment phases in the ontology lifecycle, as all the other modules in the PIPS architecture.

The choice of the development process has been dictated by the need to allow novices in ontology modelling to contribute to the ontology design. We decided to use the “Ontology 101 development process” by Noy and McGuiness [2]. This is a quick but complete development process for building ontologies that was deemed suitable for the project needs, and was easy to understand by non expert. The ontology was modelled by using Protégé[6], a popular ontology editor that provides a graphical representation of the ontology. We chose this editor for the modelling in order to allow non experts to visually evaluate the ontology. However, Protégé was only used as a graphical modelling interface, while the translation of the ontologies in OWL (and specifically in OWL-DL [6]) was achieved through the following stepwise translation process:

[4] —`http://www.w3.org/TR/2002/WD-owl-guide-20021104/food.owl`—
[5] —`http://www.ianunwin.demon.co.uk/eurocode/index.htm|`—
[6] http://protege.stanford.edu

1. firstly we used the translation functionality provided by Protégé, in order to obtain a first draft of the ontology. However, this was only a starting point, because of the known limitations that Protege-OWL has with respect to complete OWL-DL expressivity (namely rdfs:domain statements that are neither named classes nor unions of named classes and individuals with anonymous rdf:types)[7].
2. then we validated the OWL file obtained through the use of ontology reasoners such as Racer [3] and Pellet [4],
3. finally, any amendment to the ontology (due to expressivity or reasoning problems) is made by hand.

The development process consists of seven steps that we describe in the remainder, where we also discuss their application to the development of the Food ontology:

1. **Determine the domain and scope of the ontology**: we identified a number of competency questions (following [5]) that we use to limit the scope of the ontology. Examples of the competency questions are:
   – What is the maximum amount of muesli per day that is recommend by specialists?
   – Is the vitamin C content of a lemon in average higher than the content of an orange?
   – What kind of oil is better in order to get the recommended daily intake of oleic acid?
2. **Consider reusing existing ontologies**: as we discussed, we reuse terms from the ITACA food database, and the Eurocode 2 coding system;
3. **Enumerate important terms in the ontology**: key terms used in this ontology are the nouns describing generic types of food, with no relation to specific brands, such as: food names such as milk, bread, meat, vegetable, etc. Also nutritional terms like fat, vitamin, protein, sodium, sugar, and so on. The brand names used to commercialise these products are used to design the *Product Ontology*, whose description is out of the scope of this paper;
4. **Define classes and the class hierarchy**: the Eurocode2 coding system provides the backbone of the class hierarchy;
5. **Define the properties of classes and slots**: the coding system is a simple taxonomy that does not contain any property. Since we are interested in the nutritional features of food, we associated with the top class Food the nutritional properties described in the ITACA database schema, so that every kind of food inherits these properties.
6. **Define the facets of the slots**: here we define the cardinality constraints, and value restrictions. Properties modelling food nutritients have minimum cardinality 0, in order to allow us to represent the fact that foods rarely have all nutrients. Value restrictions are asserted, as described in the next section;
7. **Create instances**: decisions concerning the modelling of instances (individuals in OWL) are dictated by the notion that, from the perspective of representing nutritional information about different kinds of food, there is no difference between, for example, two strawberries. The ontology needs to represent the nutritional properties of strawberries as well as their placement in the hierarchy. Therefore we

[7] —`http://protege.stanford.edu/mail_archive/msg18490.html`—

decided to model basic food types as instances. For these foods weinstantiate the properties modelling the nutritional information.

4.2 The proposed formalisation

The Food ontology resulting from the development process described above has a total of 177 classes, 53 properties and 632 instances. The ontology is translated in OWL-DL, and we defined disjointness and cardinality constraints, as well as functional properties. In the remainder of this section we describe the hierarchical structure of the ontologies, and the most significant class properties together with their constraints.

As it can be seen in Figure 1, the PIPS Food ontology imports the PIPS Top Level ontology, that creates a link between the main domains described by the PIPS ontologies (such as diabetes, diet, products, etc). The definition of a top level specific to the PIPS application was deemed necessary in order to permit the reasoning component of the knowledge management module to traverse the PIPS ontologies as if they were modelled in one ontology only. The top level ontology, as well as the other PIPS ontologies are outside the scope of this paper (more information can be found at `http://www.csc.liv.ac.uk/semanticweb/PIPS-Ontologies.html`).

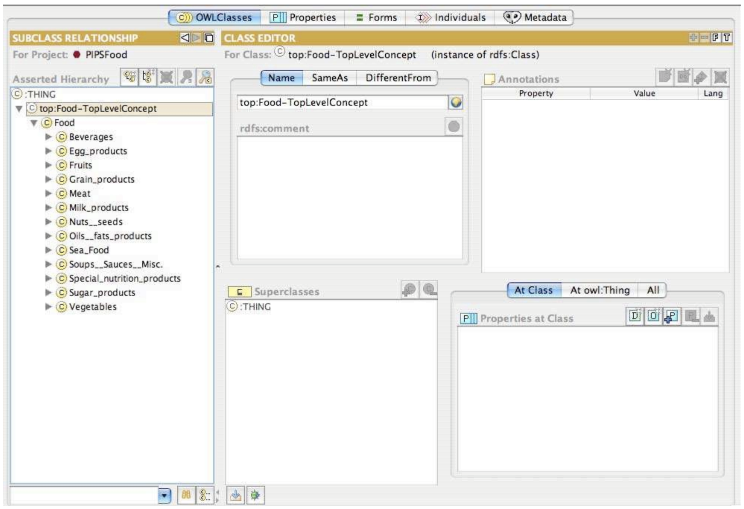

Fig. 1. The main food categories in the PIPS Food Ontology

`Food` is the root concept for this ontology, all the other concepts inherit the properties associated with it. These properties allow us to describe an aliment in terms of its nutrients, and we have 50 properties to describe them. We associate a data property with each nutrient, these have numerical range and max cardinality is in most cases 1, meaning that a nutrient can be present in the food description, and if it is present only one value can be associated with it. In addition to the nutrients, we have three special properties, `hasMaxAmount`, `hasMedAmount` and `hasMinAmount`, that represent the maximum, medium and minimum daily intake as recommended by nutritionists. These amounts are general enough to be prescribed per type of food.

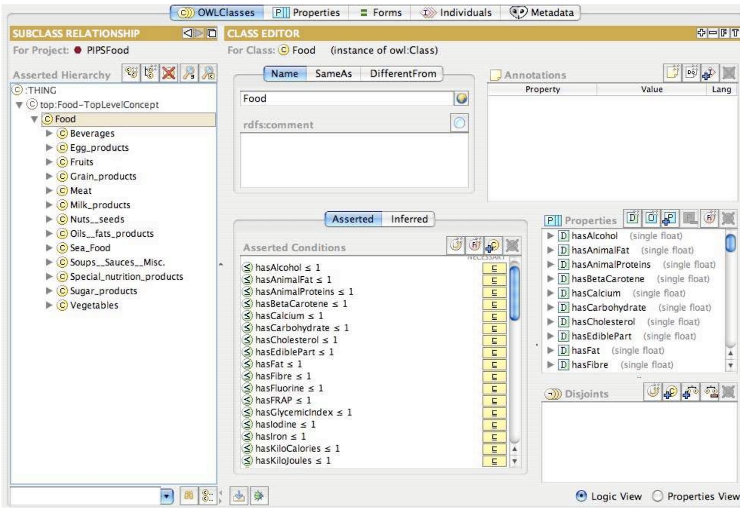

Fig. 2. The concept Food, with its properties and the constraints defined over them

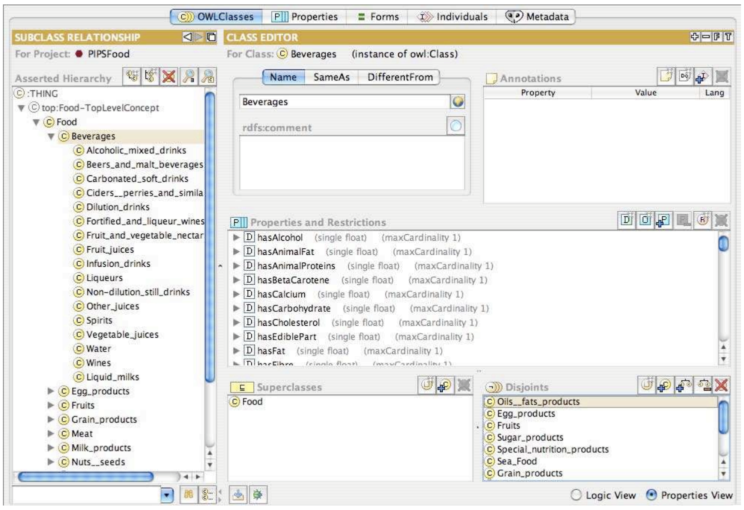

Fig. 3. The disjointness constraints defined for the class Beverages

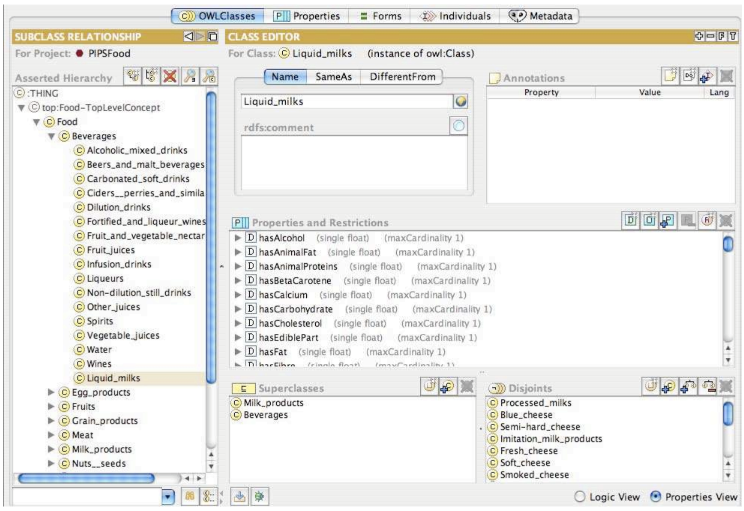

Fig. 4. Example of multiple inheritance

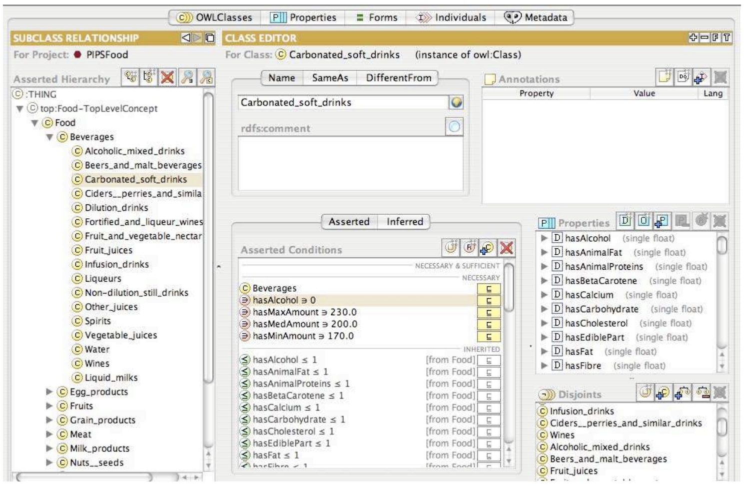

Fig. 5. Value restrictions defined for the class Carbonated soft drinks

For instance, there are recommended amounts of fruit per day, applicable to any type of fruit. Figure 2 illustrates some of the properties for the concept `Food`.

The Food Ontology organises foods in 13 main categories, each describing either a type of unprocessed aliment, such as meat, or fruit; major miscellaneous categories, such as beverages, or sauces; and food types determined by the main ingredient, such as milk products, egg products, etc. Figure 1 shows the various food types. Disjointness constraints are defined for some of these 13 classes, for instance `Beverages` is disjoint from `Fruit` or `Sea Food` (see Figure 3).

The hierarchical structure is mainly based on single inheritance, however, there are exceptions, such as the class `Liquid Milk`, that is defined as subclass of `Beverages` and `Milk products`, as illustrated in Figure 4.

Constraints on the values associated with class properties are added as we traverse the hierarchy, by posing restrictions on the values associated to some properties. For instance, the class `Carbonated soft drinks` has value restrictions on the property `hasAlcohol` (the amount of alcohol of a soft drink must be 0), and on the recommended intakes, as illustrated in Figure 5.

Finally, we defined individuals for each of the classes. For this ontology, as already mentioned, we consider individuals to be specific types of foods rather than a brand of a specific type of food. Therefore, individuals of the class `Carbonated soft drinks` include `coke`, and `tonic water`, as shown in Figure 6.

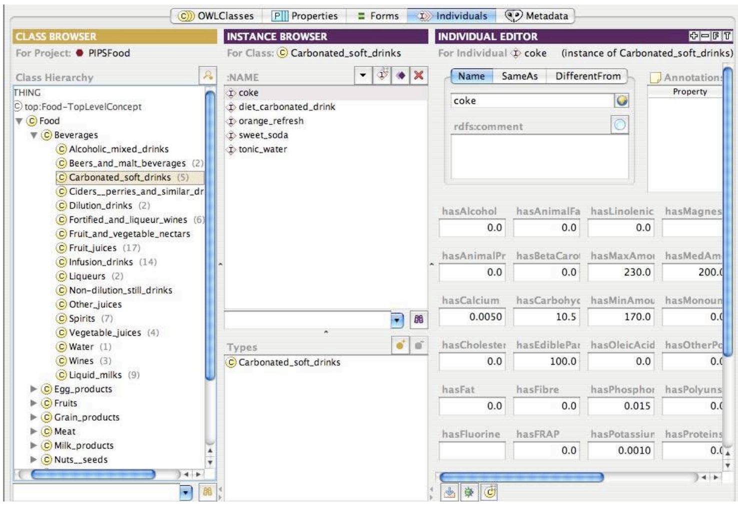

Fig. 6. Individuals defined for the class Carbonated soft drinks

5 Conclusions

This paper describes a Food ontology from the nutritional and health care viewpoint. This ontology is used to share knowledge between the different stakeholders involved in the PIPS project. We have presented the problem we addressed with the design of the Food ontology, namely the provision of nutritional advice to diabetic patients. We described briefly the development process we used to design the ontology, and we described the main features of the Food ontology.

Acknowledgements

The work presented in this paper is funded by PIPS: Personalised Information Platform for Health and Life Services - An FP6/IST eHealth Integrated Project (No. 507019)

References

1. Studer, R., Benjamins, V., Fensel, D.: Knowledge engineering, principles and methods. Data and Knowledge Engineering **25** (1998) 161–197
2. Noy, N., McGuinness, D.: Ontology development 101: A guide to creating your first ontology. Technical Report SMI-2001-0880, Stanford Medical Informatics (SMI), Department of Medicine, Stanford University School of Medicine (2001)
3. Haarslev, V., Moller, R.: Racer system description. In: Proceedings of the International Joint Conference on Automated Reasoning, IJCAR'2001. (2001)
4. Pellet OWL reasoner. (`http://www.mindswap.org/2003/pellet/index.shtml`)
5. Grüninger, M., Fox, M.: Methodology for the design and evaluation of ontologies. In IJCAI'95 Workshop on Basic Ontological Issues in Knowledge Sharing. (1995)
6. OWL. (`http://www.w3.org/TR/2003/ CR-owl-features-20030818/`)