Revising and extending the Units of Measure "subontology"

Helena Sofia Pinto and **João P. Martins**
Grupo de Inteligência Artificial
Departamento de Eng. Informática
Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais
1049-001 Lisboa, Portugal

Abstract

In this paper we describe the revision and extension processes performed to the "subontology" for Units of Measure of the Merged ontology proposed by Ian Niles. We have used a principled approach to guide revision of knowledge. The methodology and the sets of criteria that were used in this revision are described.

1 Introduction and motivation

As proposed in the call for papers we were asked to revise SUO material and either add to the proposed standard or propose changes which were intended to "revise and improve" the ontology. We chose to do the latter. We picked up Ian Niles Merged ontology for revision (sent to all SUO participants on the 15th of January of 2001) and since we were already familiar with the domain of one of its "subontologies", the Units of Measure part of the ontology, we began the process of revising it. We believe that knowledge about units of measure needs to be represented in the SUO ontology in order to fulfill its scope and purpose.

We wanted to follow a principled approach in this revision in order to have some guarantees that the changes to be proposed would effectively improve it and would contribute to increase the confidence of future SUO users. For this reason, we picked up sets of criteria that had already been proposed in the literature to analyze an ontology for reuse purposes. These criteria were proposed specifically to analyze and study ontologies for integration[1] purposes [Pinto and Martins, 2000; Pinto, 1999]. Since these are criteria specifically for integration, which will be one of the main processes involved in reusing/using the SUO ontology, we believe that by following them while building it, will improve its reusability and usability in the future.

This paper is organized as follows. We begin by presenting the results of reviewing Units of Measure, refer how we are extending the ontology and make some general comments to the rest of Ian Nile's Merged ontology. Then we describe the steps followed to revise Units of Measure, presenting the most important conclusions found in knowledge acquisition and discussing the criteria used to review the ontology. Finally we present our conclusions and refer future work.

[1] Integration is one of the ontology reuse processes [Pinto *et al.*, 1999]. The other reuse process is merge. For a characterization of integration see [Pinto *et al.*, 1999; Pinto and Martins, 2000].

2 Revision of Units of Measure

We began by analyzing and studying the "subontology" proposed by Ian Niles for Units of Measure. We will use the terminology in the domain of standard units, which is described, in part, in Section 5.1. The conclusions that we reached from the analysis of this "subontology" from the domain point of view were:

missing knowledge There is some knowledge missing in the ontology:

- Some of the classifications of physical quantities are missing, for instance one can measure solid angles, areas, velocity, acceleration, magnetic flux density, luminous flux, capacitance, etc. Since, for instance, we have information that says that pressure measures are a subclass of units of measure, we should add information about missing physical quantities. For instance, it should be added the fact that velocity measures or magnetic flux density measures are subclasses of units of measure. The appropriate classes of physical quantities should be added to the ontology to improve its completeness, reusability and generality.
- In the case of SI units that measure the aforementioned physical quantities some are defined from SI base units (solid angles, velocity, acceleration) and others are derived SI units that have special names and symbols (magnetic flux density, luminous flux, capacitance). The instances of derived units with special names and symbols that are missing, such as tesla, the SI unit for magnetic flux density, should be added to the ontology to improve its completeness, reusability and generality.
- Some SI units are not asserted as such, for instance, `hertz`, `volt` or `watt`.
- The conversion function associated to the time unit `minute` is missing. It should be provided since this is an accepted unit but not the SI base unit for time (which is the `second`).

- The conversion function associated to the force unit of measure `pound-force` is missing. A conversion function into the SI corresponding unit (which is `newton`) should be provided.
- There should be information about the symbols associated with the several units represented in this ontology. They are an important attribute of units of measure, for instance the length unit meter has as its corresponding symbol m.
- Functions that express derived units with special names and symbols into SI base units should be included since this can allow appropriate symbolic manipulation and simplification of units of measure. For example, the frequency unit `hertz` can be expressed in terms of the time unit `second`. Therefore, if we take the inverse of a frequency we get a time or if we divide a force by a mass we get an acceleration (which is a quantity expressed in terms of base units). For instance, it would be interesting to know that `newton` (N) can be expressed in terms of SI base units as m kg s^{-2}. This kind of knowledge is missing in the ontology.
- Some of the multiples and submultiples of SI units are represented in the ontology while others are not, for instance `nano-second` is represented but `milli-second` is not. Maybe all appropriate multiples and submultiples should be added to the ontology. Either we add all accepted multiples and submultiples of all units or we decide in a case by case basis depending on the value that the corresponding physical quantities typically have.
- Maybe, we should include information about the physical quantities. For instance, we can say that speed quantities are expressed as the quotient of a length quantity with a time quantity. Therefore, if we divide a speed by a time interval we get a length. For the moment, we only have some of the categories of physical units represented in the ontology, for instance, the fact `MassMeasure` or `PressureMeasure` are subclasses of units of measure. We would like to add information, such as that pressure measures are a quotient of a force measure with an area measure or that force measures are a multiplication of a mass and an acceleration.

superfluous knowledge The mass unit `kilogram` should not be supplied with a conversion function into grams since it is the SI basic unit from which the other mass units should be measured. Therefore, all non-base SI mass units should be supplied with conversion functions into the SI base unit and not the other way around as in this case. Conversion functions for `coulomb` and `hertz` should not be provided because they are the derived units with special names and symbols in SI for electric charge and frequency, respectively.

"misplaced" knowledge We have not found misplaced knowledge.

knowledge sources changes We do not know which knowledge sources were used to build the Standard-Units ontology [Gruber and Olsen, 1994], the source of knowledge that was used to build this "subontology", therefore, we cannot evaluate if they should be changed. However, we advise that the knowledge sources to be used should be from the most reputable as possible sources, such as [Nor, 1991] or [Taylor, 1995]. This provides quality guarantees to SUO reusers/users. We should add that we have not analyzed the Standard-Units ontology, we only revised the part of Ian Nile's Merged ontology concerning Units of Measure.

documentation changes There are concepts which do not have documentation, for instance `Angstrom`. There are concepts with a very deficient documentation, for example, `meter` or `kilogram` (which is not even referred as the SI unit of mass). There is room for improvement in the documentation of this ontology.

To improve it we suggest that:

1. all SI base units should explicitly state their status in their documentation. The only unit that has this information is `ampere`.
2. all SI base units should include its internationally accepted natural language definition, for instance `meter` is currently defined as "The meter is the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second"[2];
3. all SI derived units should include in its documentation how they can be defined in terms of base units and in terms of other SI units including those derived with special names and symbols. In the documentation of `pascal` one can find the definition of the unit in terms of other SI units including those derived with special names and symbols. However, this is an exception. Moreover none of the derived units is provided with its definition in terms of SI base units.

terminology changes According to the rules and style conventions for spelling unit names the names of all units start with lower-case letter except at the beginning of a sentence or in capitalized material such as a title. Therefore, appropriate changes should be conducted in the ontology, at least, for SI units. Prefixes associated to units (kilo, milli, giga, etc.) should also be spelled starting with lower-case letters. Therefore, appropriate changes should be conducted in the ontology for accepted prefixes. For now we can leave non-SI units as they are currently spelled.

The physical quantities should have their standard names, such as electric current instead of `electrical current`.

The unit `Second-Duration` should be renamed second since this is the standard name for the unit. So should Minute-Duration into minute, Day-Duration into

[2] The previous definition of meter "The meter is the length equal to 1650763.73 wavelengths in vacuum of the radiation corresponding to the transition between the levels $2p_{10}$ and $5d_5$ of the krypton 86 atom, ^{86}Kr." was superseded in 1983.

day and Hour-Duration into hour since these units are accepted for use with the SI units and therefore have standardized names.

definition changes There are some problems in the definitions, namely:

- the definition of `coulomb` is incorrect since it is a unit of electric charge or quantity of electricity and not of electric current;
- moreover, this assertion (that coulomb is a unit of ... measure) says that coulomb is an `instance` of electric current measure when it should say that it is a `subclass` of electric charge measure or quantity of electricity measure;
- in the definition of `coulomb` and `hertz` an undefined function is used `PerFn`;
- there seems that there was not a coherent criteria as to which conversion factors should be used (number of digits?, ???), therefore there are a lot of these factors that should be changed for more accurate ones, for instance, the conversion factor that should be used to express calories in joule should be 4.1868 and not 4.186 as represented in the ontology.
- Some conversion functions are defined into units that are not the base unit, for instance megahertz is defined in terms of kilohertz when it should be defined in terms of hertz, hour is defined in terms of minute and not in terms of second. The conversion function should convert into base units or into derived units with special names and symbols but not for their multiples and submultiples.

practices changes We do not have information how the ontology was built therefore we cannot evaluate this criterion.

The conclusions that we reached from the analysis of this subontology from the point of view of the ontologists that may come to reuse/use it (quality of the product for a possible user) were:

general structure The general structure of the ontology was adequate, namely:

- it is one well-balanced hierarchy;
- for the moment there is no need to further divide it into other subontologies, but if other systems of units (for instance, CGS) should be added, each system of units should be described in its own subontology with conversion functions into SI corresponding units;
- although some classes of units and their corresponding units are missing the knowledge already introduced contains enough specialization of concepts;
- knowledge is correctly "placed" in the structure so that inheritance mechanisms can infer appropriate knowledge from the ontology (apart from the problem in the coulomb definition that is preventing inheritance of properties from its corresponding class of measure);
- some concepts are missing but they don't harden the task of introducing new concepts in the ontology;
- the semantic distance between sibling concepts [Arpirez-Vega *et al.*, 1998] was minimized.

basic distinctions Maybe it would be a good idea to superimpose another set of basic distinctions. These would make explicit which units are base units, which units are derived ones and which units are derived units with special names and symbols. This information depends on the system of units. For example, the pairs (base quantity/base unit) in SI are (length/meter), (mass/kilogram), (time/second), (electric current/ampere), (thermodynamic temperature/kelvin), (amount of substance/mole) and (luminous intensity/candela).

structuring relation The structuring relation (class-superclass and instance-class) is adequate.

naming convention rules The naming convention rules are adequate.

definitions The assertion that coulomb is an "electrical current" unit (this is not correct, it is an "electrical charge" unit) is out of place, since all other assertions of the same kind are in the beginning of the subontology. Therefore, it should be placed next to the other definitions of this kind.

There is a problem in all definitions of time units since the physical quantity of `TimeMeasure-Duration` is defined and then all time unit instances are defined as instances of `TimeMeasure`.

There are some minor spelling errors, for instance in the definition of the conversion factor for pico ampere the expression "Pico-AmpereFn" is used.

Regarding the other definitions, *in general*, they follow unified patterns, are concise, consistent, correct (lexically and syntactically), precise and efficient. We believe there is room for improvement in the completeness aspect, for instance, an attribute that refers the symbol of units of measure and functions that express derived units into base units. This can, in part, be achieved by the introduction of the knowledge that was identified as missing. There is also room for improvement in the clarity issue. This can, in part, be achieved by the improvement of the documentation associated to each knowledge piece.

documentation As we said there is room for improvement in the documentation. From the ontologist point of view, the user, if he/she did not have any knowledge of the domain it would be rather difficult to acquire it from the ontology. The alphabetical order is not the most appropriate one to introduce knowledge of a domain to someone not familiar with it. Therefore, we should introduce all knowledge following some logical order instead.

Moreover, much more internal documentation should be added to make an introduction of the domain to those not familiar with it. Users should be able to learn about the domain from the ontology. As a matter of fact, one of the purposes of ontologies may be teaching [Arpirez-Vega *et al.*, 1998; 2000]. It is one explicit purpose of the SUO ontology.

Moreover, representation alternatives and the choices that were made are not discussed.

Finally, part of the documentation is not coherent, for instance, we only know that `ampere` is one of the base units of the SI system.

knowledge pieces represented In general, the knowledge pieces that are represented are all important and appropriate for this domain, but there is a lack of a lot of knowledge about other physical quantities and measure units.

There are other sets of criteria to analyze an ontology but, in general, they are subsumed by the criteria that were used in our revision. The design criteria proposed by Gruber [Gruber, 1995] (clarity, coherence, extendibility, minimal encoding bias and minimal ontological commitment) are, in general, satisfied (apart from clarity, specially of the documentation). The ontology also complies to most of the criteria proposed for evaluation[3] and assessment[4] of ontologies proposed in [Gómez-Pérez *et al.*, 1995; Gómez-Pérez, 1999]. All criteria proposed in [Gómez-Pérez and Rojas-Amaya, 1999] are subsumed by the ones that we used in our revision. Moreover, the criteria that were used are much broader.

After reviewing the ontology we introduced the changes that were suggested to solve the identified problems.

For instance, `meter` is now defined as:

```
; Length Base Unit
(instance-of meter LengthMeasure)
(instance-of meter SystemeInternation-
alUnit)
(documentation meter "SI length unit.
symbol:  m.  It is one of the base units
in SI. Its definition has evolved over
time.  It is currently defined as:  The
meter is the length of the path trav-
eled by light in vacuum during a time
interval of 1/299792458 of a second.")
```

For instance, `coulomb` is now defined as:

```
; Electric Charge Units
(instance-of coulomb ElectricChargeMea-
sure)
(instance-of coulomb SystemeInternation-
alUnit)
(documentation coulomb "SI charge unit.
symbol:  C. It is the quantity of elec-
tric charge transported through a cross
section of a conductor in an electric
circuit during each second by a current
of 1 ampere.  coulomb = s*A")
; NOTE: Coulomb does not have a conver-
sion function.
```

[3]Consistency, conciseness, expandability and sensitiveness are verified but not the completeness criterion.

[4]Quality, portability and usability are verified, but there is room for improvement of understandability and generality.

3 Extension of Units of Measure

Currently, we are introducing knowledge that was identified as missing or important to be represented in the previous section. There are some extensions that should be discussed within the SUO group, namely whether one should include knowledge about physical quantities (pressure is a quotient of a force measure with an area measure).

A possible extension to be performed later can be the introduction of knowledge about other systems of units. This should only be performed if there is any interest on this extension by the rest of the members of the SUO group.

4 General comments to other parts of Ian Nile's Merged ontology

One of the comments we would like to add is the fact that the documentation provided to "General Classes" should be improved in the style of the mini-essays proposed by Pat Hayes in an e-mail (8/Feb/2001) "More documentation, please" for Occurrent and Continuant. Moreover, the documentation of all these concepts should be associated directly to its definition, because for those not familiar with such philosophical and general concepts it is hard to understand what they may mean, when should someone classify a concept according to a particular distinction, etc. The documentation of every knowledge piece should directly be associated to the knowledge piece to improve the clarity of the ontology. If SUO ontology wants to be adopted as a standard it must be understood even with people without solid backgrounds in philosophy. The changes proposed in Ian Nile's e-mail (29/Jan/2001), "Proposed change to Merged ontology" are a step in that direction.

As to the part of the ontology where processes are defined, why is there no mention to the concept of an activity? For instance, software processes [IEEE-Std-1074-1995, 1996] and their corresponding life cycles are defined in terms of activities to be performed.

We believe that the axiom that states that only entities are instances of classes and only classes have instances has a syntactical error. The antecedent should be (instance-of? ?instance ?class) and not (instance-of? (?instance ?class)).

All axioms involving concepts that have in the meanwhile been removed in Ian Nile's e-mail (29/Jan/2001), "Proposed change to Merged ontology" should be removed. For instance, the axiom stating that if a mediating entity ?m has ?x and ?y then either ?x has ?y or vice-versa should be eliminated since the mediating concept does not belong to the ontology.

5 Methodology followed to revise Units of Measure

We began by reviewing the Units of Measure subontology using the criteria proposed in [Pinto and Martins, 2000; Pinto, 1999]. We had already some knowledge about the domain[5]

[5]We have Mechanical Engineering backgrounds and some knowledge in the Chemistry, Thermodynamics and Electrical domains, so we are familiar with most units appearing in the SI system of units and have extensively manipulated SI units.

so we could begin the task before acquiring some knowledge about the domain.

Since we detected some problems we did some knowledge acquisition. We mainly looked into on-line sources of knowledge to speed up the knowledge acquisition process. The best source of knowledge found was the NIST Reference on Constants, Units, and Uncertainty `http://physlab.nist.gov/cuu/Units/index.html`.

Then, we completed the analysis of the ontology following the aforementioned criteria, namely we verified every conversion function.

Finally we implemented all proposed changes in the ontology. At the same time as we changed the ontology to eliminate identified problems minor mistakes were found.

5.1 Knowledge acquisition results

Units of measure can be divided into two categories: those that are base units and those that are derived from them. The Systeme International (SI) system of units considers the following *base* units:

length meter (m)

mass kilogram (kg)

time second (s)

electric current ampere (A)

temperature kelvin (K)

luminous intensity candela (mol)

amount of substance mole (cd)

From these units all other units can be defined. Units defined from base units are called *derived units*. In SI some derived units are, for example:

area square meter (m^2)

volume cubic meter (m^3)

speed, velocity meter per second ($m\ s^{-1}$)

acceleration meter per second square ($m\ s^{-2}$)

mass density kilogram per cubic meter ($kg\ m^{-3}$)

specific volume cubic meter per kilogram ($m^3\ kg^{-1}$)

Some *derived units have special names and symbols*. There are 22 SI derived units that were given special names:

plane angle radian[6] (rad) which can be expressed in terms of SI base units as $m\ m^{-1} = 1$

solid angle steradian[7] (sr) which can be expressed in terms of SI base units as $m^2\ m^{-2} = 1$

frequency hertz (Hz) which can be expressed in terms of SI base units as s^{-1}

force newton (N) which can be expressed in terms of SI base units as $m\ kg\ s^{-2}$

pressure or stress pascal (Pa) which can be expressed in terms of SI base units as $m^{-1}\ kg\ s^{-2}$ and in terms of other SI units as $N\ m^{-2}$

[6] In former editions this was considered a supplementary unit

[7] In former editions this was considered a supplementary unit

energy, work, quantity of heat joule (J) which can be expressed in terms of SI base units as $m^2\ kg\ s^{-2}$ and in terms of other SI units as N m

power, radiant flux watt (W) which can be expressed in terms of SI base units as $m^2\ kg\ s^{-3}$ and in terms of other SI units as $J\ s^{-1}$

electric charge, quantity of electricity coulomb (C) which can be expressed in terms of SI base units as s A (one also sees A s)

electric potential difference, electromotive force volt (V) which can be expressed in terms of SI base units as $m^2\ kg\ s^{-3}\ A^{-1}$ and in terms of other SI units as $W\ A^{-1}$

capacitance farad (F) which can be expressed in terms of SI base units as $m^{-2}\ kg^{-1}\ s^4\ A^2$ and in terms of other SI units as $C\ V^{-1}$

electric resistance ohm (Ω) which can be expressed in terms of SI base units as $m^2\ kg\ s^{-3}\ A^{-2}$ and in terms of other SI units as $V\ A^{-1}$

electric conductance siemens (S) which can be expressed in terms of SI base units as $m^{-2}\ kg^{-1}\ s^3\ A^2$ and in terms of other SI units as $A\ V^{-1}$

magnetic flux weber (Wb) which can be expressed in terms of SI base units as $m^2\ kg\ s^{-2}\ A^{-1}$ and in terms of other SI units as V s

magnetic flux density tesla (T) which can be expressed in terms of SI base units as $kg\ s^{-2}\ A^{-1}$ and in terms of other SI units as $Wb\ m^{-2}$

inductance henry (H) which can be expressed in terms of SI base units as $m^2 kg\ s^{-2}\ A^{-2}$ and in terms of other SI units as $Wb\ A^{-1}$

celsius temperature degree Celsius (°C) which can be expressed in terms of SI base units as K

luminous flux lumen (lm) which can be expressed in terms of SI base units as $m^2\ m^{-2}\ cd = cd$ and in terms of other SI units as cd sr

iluminance lux (lx) which can be expressed in terms of SI base units as $m^2\ m^{-4}\ cd = m^{-2}\ cd$ and in terms of other SI units as $lm\ m^{-2}$

activity (of a radionuclide) becquerel (Bq) which can be expressed in terms of SI base units as s^{-1}

absorbed dose, specific energy (imparted), kerma gray (Gy) which can be expressed in terms of SI base units as $m^2\ s^{-2}$ and in terms of other SI units as $J\ kg^{-1}$

dose equivalent sievert (Sv) which can be expressed in terms of SI base units as $m^2\ s^{-2}$ and in terms of other SI units as $J\ kg^{-1}$

catalytic activity katal (kat) which can be expressed in terms of SI base units as s^{-1} mol

There are other SI derived units whose names and symbols include derived units with special names and symbols, such as:

dynamic viscosity pascal second (Pa s)

moment of force newton meter (N m)

surface tension meter per meter (N m^{-1})

angular velocity radian per second (rad s^{-1})

angular acceleration radian per second squared (rad s^{-2})

heat capacity, entropy joule per kelvin (J K^{-1})

The accepted SI prefixes are:

Prefix	Factor	Symbol
yotta	10^{24}	Y
zetta	10^{21}	Z
exa	10^{18}	E
peta	10^{15}	P
tera	10^{12}	T
giga	10^{9}	G
mega	10^{6}	M
kilo	10^{3}	k
hecto	10^{2}	h
deka	10^{1}	da

Prefix	Factor	Symbol
deci	10^{-1}	d
centi	10^{-2}	c
milli	10^{-3}	m
micro	10^{-6}	μ
nano	10^{-9}	n
pico	10^{-12}	p
femto	10^{-15}	f
atto	10^{-18}	a
zepto	10^{-21}	z
yocto	10^{-24}	y

5.2 Sets of criteria used to analyze the Units of Measure "subontology"

We revised the Units-of-Measure subontology following the integration-oriented criteria of candidate ontologies proposed in [Pinto and Martins, 2000; Pinto, 1999]. There are two sets of criteria: those that are to be performed by domain experts which were designed for technical evaluation and those that are to be performed by ontologists that aim at reusing the ontology (by means of integration) which were designed for user assessment.

These criteria have been used in two integration experiences: building the Reference ontology [Pinto, 1999; Arpirez-Vega *et al.*, 2000; 1998] reusing the KA^2 ontology [Benjamins and Fensel, 1998; Benjamins *et al.*, 1999] and building some of the subontologies needed to build an Environmental Pollutants ontolology, namely a Monoatomic Ions [Pinto, 1999; Amaya, 1998; Gómez-Pérez and Rojas-Amaya, 1999] ontology reusing Chemicals [Fernández, 1996].

In the following paragraphs we describe the criteria that were used.

The integration-oriented technical evaluation criteria that were used to analyze the Units of Measure subontology were:

missing knowledge What knowledge is missing (by knowledge we mean any knowledge piece, such as classes, instances, relations, etc.)?

Sometimes some knowledge pieces that are relevant, important and usually used to characterize the domain are not represented in the ontology. This includes not only classes and instances, but also important basic distinctions made of the domain concepts (classification criteria of the concepts described in the ontology that are widely accepted to characterize the domain and usually represented in the upper-levels of the ontology), relations that are relevant to represent knowledge about the domain (which relations should be specified and for which concepts should they be specified), etc. The domain experts should also analyze what important knowledge about the domain is missing in the ontology in view to the particular use that the ontology is going to have.

superfluous knowledge What knowledge should be removed?

Sometimes some knowledge pieces represented in the ontology are superfluous, either because they are not important, or not relevant, or not usually used to describe the domain in question or because they are not needed for the particular use that the ontology is going to have.

"misplaced" knowledge What knowledge should be relocated?

Sometimes knowledge pieces should be placed elsewhere in the ontology so that the domain is best characterized.

knowledge sources changes Which knowledge sources changes should be made?

Sometimes some of the knowledge sources used to acquire knowledge are not the most reputable or up-to-date. Knowledge from those sources that is represented in the ontology should be replaced by more reputable, standard and more up-to-date knowledge.

documentation changes Which documentation changes should be made?

Sometimes the documentation of the domain is not correct (syntactically and semantically), precise, complete (comprehensive) or reflects the last discoveries in the field and should be changed. The documentation should explain the domain and the knowledge pieces represented in the ontology so that a non-expert could learn enough about the domain to be able to understand the knowledge pieces that are represented in the ontology.

terminology changes Which terminology changes should be made?

Sometimes the terminology used is not the most usually accepted in a (sub)field or in a related field, or it is not the standard terminology and should be changed.

definition changes Which definition changes should be made?

Sometimes the definitions used are not the most usually accepted, standard or composed of the definitional characteristics of the knowledge pieces.

practices changes Which practices changes should be made?

Sometimes the procedures used to gather knowledge (knowledge acquisition) and to build the ontology (ontological engineering) are not the most correct ones or follow the accepted best practices in the domain area.

The integration-oriented user assessment criteria that were used to analyze the Units of Measure subontology were:

general structure Is the general structure of the ontology adequate?

It is important to assess whether the general structure of the ontology is appropriate, that is, it complies with the user requirements needed of it. To analyze the structure of the candidate ontology six criteria should be taken into consideration:

- Is the structure adequate (one hierarchy, several hierarchies, a graph, etc.) and preferably well-balanced?
- Is the ontology divided into adequate and enough modules, that is, is the ontology divided into natural and appropriate (quality and quantity) subontologies?
- Is there adequate and enough specialization of concepts, that is, are the needed concepts and their specializations represented?
- Is knowledge correctly "placed" in the structure so that inheritance mechanisms can infer appropriate knowledge from the ontology?
- Is there enough diversity represented in the ontology so that new concepts are more easily introduced?
- Are similar concepts represented close to one another whereas less similar concepts are represented further apart (minimization of the semantic distance between sibling concepts [Arpirez-Vega *et al.*, 1998])?

If the ontology has not the adequate structure, then the changes to be made can be so extensive that it may be more cost effective to build an ontology from scratch. That is why its analysis is so important.

basic distinctions Are the relevant and required (quantity and quality) basic distinctions represented?

Changing the basic distinctions (usually represented at the top-levels of the ontology) upon which the ontology is based can also imply a vast revision of the ontology.

structuring relation Is the privileged relation upon which the ontology is structured the required one?[8]

Changing the privileged relation according to which the ontology is organized can also have deep consequences. The whole knowledge would, most probably, have to be revised, since the new relation organizes knowledge in a completely different way. Knowledge about a given domain that should be represented using one relation has nothing to do with what should be represented using another relation. Probably it is preferable to build a new ontology from scratch (if none of the available ones meets our needs).

naming convention rules Do the names of the knowledge pieces follow standardization rules?

Whenever possible, naming convention rules should be enforced all over the resulting ontology so that terminology becomes coherent. This increases reusability and usability of the resulting ontology (it is easier to find relevant knowledge and it is easier to introduce new knowledge).

definitions Do the definitions of the knowledge pieces follow unified patterns, are clear, concise, consistent, complete, correct (lexically and syntactically), precise and accurate? Are they efficient?

All these questions deal with the way knowledge is represented in the ontology.

documentation Is the documentation clear, helpful and adequate? Does it discuss alternative representations and the choices that were made to represent knowledge? Is it coherent in relation to the definition of the knowledge piece?

Although documentation is one of the constituents of an ontology knowledge piece it usually is its most neglected component. This should not be so because documentation is crucial to improve clarity of an ontology.

knowledge pieces represented Are all and only the appropriate knowledge pieces represented (or included)?

This issue should be analyzed taking into account the knowledge pieces that domain experts have found important to be represent.[9] If they have found some knowledge pieces lacking/superfluous they should be added/deleted to/from the ontology. The ontologists should analyze whether the proposed changes affect coherence, for instance. The use made of the knowledge in the ontology also influences the way those knowledge pieces are represented and which knowledge pieces need to be represented. The ontologist analysis has to focus both aspects: the relevant and needed knowledge pieces are represented and they are usefully and adequately represented.

6 Conclusions and future work

In this paper we describe the process of revision that was performed to the Units of Measure "subontology". We are currently introducing knowledge that was identified as missing or important to be represented. We believe that the comprehensive revision done to this ontology has contributed to improve its quality, generality and therefore its reusability/usability.

The use of sets of criteria to guide ontology review allowed us to be more systematic and focused on the kind of problems to look for. We have used the broadest criteria found in the literature for reuse processes. Reuse is one of the purposes of the SUO ontology.

[8] An ontology can be thought of as structured or organized according to one privileged relation, for example, ISA or part-of.

[9] The result of knowledge acquisition.

As soon as all changes are introduced we plan to submit the improved ontology to the SUO group for comments, criticism, revision.

References

[Amaya, 1998] M. Dolores Rojas Amaya. Ontologia de Iones Monoatómicos en Variables Fisicas del Medio Ambiente. Proyecto Fin de Carrera, Fac. de Informática, UPM, 1998.

[Arpirez-Vega *et al.*, 1998] J. Arpirez-Vega, A. Gomez-Perez, A. Lozano-Tello, and H. Sofia Pinto. $(ONTO)^2$Agent: An Ontology-Based WWW Broker to Select Ontologies. In *Proceedings of ECAI98's Workshop on Application of Ontologies and Problem Solving Methods*, pages 16–24, 1998.

[Arpirez-Vega *et al.*, 2000] J. Arpirez-Vega, A. Gomez-Perez, A. Lozano-Tello, and H. Sofia Pinto. Reference Ontology and $(ONTO)^2$Agent: the Ontology Yellow Pages. *Knowledge and Information Systems*, 2(4):387–412, 2000.

[Benjamins and Fensel, 1998] Richard Benjamins and Dieter Fensel. The Ontological Engineering Initiative $(KA)^2$. In Nicola Guarino, editor, *Formal Ontology in Information Systems*, pages 287–301. IOS Press, 1998.

[Benjamins *et al.*, 1999] Richard Benjamins, Dieter Fensel, Stefan Decker, and Asunción Gómez-Pérez. $(KA)^2$: Building Ontologies for the Internet, a Mid Term Report. *International Journal of Human Computer Studies*, 51:687–712, 1999.

[Fernández, 1996] Mariano Fernández. CHEMICALS: Ontologia de Elementos Químicos. Proyecto Fin de Carrera, Fac. de Informática, UPM, 1996.

[Gómez-Pérez and Rojas-Amaya, 1999] Asunción Gómez-Pérez and Dolores Rojas-Amaya. Ontological Reengineering for Reuse. In D. Fensel and R. Studer, editors, *Proceedings of the European Knowledge Acquisition Workshop, EKAW99*. Springer Verlag, 1999.

[Gómez-Pérez *et al.*, 1995] A. Gómez-Pérez, N. Juristo, and J. Pazos. Evaluation and Assessment of the Knowledge Sharing Technology. In N.J.I. Mars, editor, *Towards Very Large Knowledge Bases*, pages 289–296. IOS Press, 1995.

[Gómez-Pérez, 1999] Asunción Gómez-Pérez. Evaluation of Taxonomic Knowledge in Ontologies and Knowledge Bases. In *Proceedings of the Knowledge Acquisition Workshop, KAW99*, 1999.

[Gruber and Olsen, 1994] Thomas Gruber and G. R. Olsen. An Ontology for Engineering Mathematics. In J. Doyle, E. Sandewall, and P. Torasso, editors, *KR94 Proceedings*, pages 258–269. Morgan Kaufmann, 1994.

[Gruber, 1995] Thomas Gruber. Towards Principles for the Design of Ontologies for Knowledge Sharing. *International Journal of Human Computer Studies*, 43(5/6):907–928, 1995.

[IEEE-Std-1074-1995, 1996] IEEE-Std-1074-1995. IEEE Standard for Developing Software Life Cycle Processes. New York (USA), April 1996.

[Nor, 1991] Le Systéme International d'Unités (SI), The Intenational System of Units (SI). Bur. Intl Poids et Mesures, Sévres, France, 1991. 6th Edition.

[Pinto and Martins, 2000] H. Sofia Pinto and J.P. Martins. Reusing Ontologies. In *Proceedings of AAAI 2000 Spring Symposium Series, Workshop on Bringing Knowledge to Business Processes, SS-00-03*, pages 77–84. AAAI Press, 2000.

[Pinto *et al.*, 1999] H. Sofia Pinto, A. Gómez-Pérez, and J. P. Martins. Some Issues on Ontology Integration. In *Proceedings of IJCAI99's Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends*, pages 7.1–7.12, 1999.

[Pinto, 1999] H. Sofia Pinto. Towards Ontology Reuse. In *Proceedings of AAAI99's Workshop on Ontology Management, WS-99-13*, pages 67–73. AAAI Press, 1999.

[Taylor, 1995] Barry N. Taylor. Guide for the Use of the International System of Units (SI). United States Department of Commerce, National Institute of Standards and Technology, April 1995. NIST Special Publication 811.