

Ecolingua: a Formal Ontology for Data in Ecology

Virgínia Brillhante
Computing Science Department
Federal University of Amazonas
Av. Gen. Rodrigo O. J. Ramos, 3000
69060-020 - Manaus, AM - Brazil
virginia@dcc.ufam.edu.br

Abstract

Ecolingua is an ontology for ecological quantitative data, which has been designed through reuse of a conceptualisation of quantities and their physical dimensions provided by the EngMath family of ontologies. A hierarchy of ecological quantity classes is presented together with their definition axioms in first-order logic. An implementation-level application of the ontology is discussed, by way of which conceptual ecological models are synthesised from data annotations in Ecolingua through reuse of existing model structures.

Keywords: Ontology reuse, engineering and application; ecological data; model synthesis

1 Introduction

Ecolingua is an evolving, domain-specific ontology which brings a contribution towards a conceptualisation of the Ecology domain by formalising properties of quantitative data that typically feed ecological simulation models. Building on the EngMath family of ontologies [21], data classes are characterised in terms of physical dimensions, which are a fundamental property of physical quantities in general. Ecolingua is conceptualised within the frame-system view of the world, which is wide-spread in the ontologies field, and is formalised in first-order logic (FOL). It has been developed as part of a research project on model synthesis based on metadata and ontology-enabled reuse of model designs [4].

*The author wishes to thank FAPEAM (Fundação de Amparo a Pesquisa do Estado do Amazonas) for its partial sponsorship through the research project MOeMA-IS (Metadata, Ontologies and Sustainability Indicators Integrated to Environmental Modelling). Rodrigo Camelo is acknowledged for providing useful references. Thanks are also due to the anonymous reviewers for their valuable comments.

We start by referring to other developments on ontologies in the environmental sciences domain in Sect. 2, and to upper level ontologies related to Ecolingua in Sect. 3. This is followed by a discussion on the reuse of the EngMath ontologies in the development of Ecolingua in Sect. 4. In Sect. 5 we show the hierarchy of concepts in Ecolingua, how they are logically defined and how they are employed as an annotation language for data in Ecology. Section 6 starts presenting each Ecolingua concept beginning with contextual data classes. Section 7 is the core of the paper, presenting the concepts in Ecolingua's high-level layer, including quantitative data classes. Lower level Ecolingua appears in Sect. 8. Section 9 gives a summarised description of an Ecolingua application in the synthesis of conceptual ecological models which are built preserving consistency with the properties of their supporting data. Finally, conclusions and considerations on current and future work appear in Sect. 10.

2 Environmental Ontologies

So far, not much has been on the intersection between ontologies and environmental sciences. Ontology engineering in this domain is a challenging endeavour, given the difficulty to attain a conceptualisation which can encompass semantics of the many fields – biology, geology, law, computing science, etc. – that relate to environmental sciences. The cousin discipline of molecular biology, on the other hand, has been enjoying significant progress in knowledge representation through ontologies, with current technologies being explored and enhanced to suit needs that are characteristic of the domain [35].

Although taxonomies have been in use in the field of Ecology for a long time, the work by B. Niven is the earliest we are aware of where concepts are defined in the shape of what we call today a formal ontology. In

[33, 34] she proposes a formalisation of general concepts in animal and plant ecology, such as environment, niche and community. Niven’s motivation for doing this work was similar to that of ontology engineers nowadays: to provide a precise definition of a theory with all its assumptions explicitly laid down, so that it can be communicated to others exactly, and so that the knowledge domain is equipped with a framework within which specific models or systems can be constructed.

EDEN is an ontology of environmental concepts which has been built to give support to an agent-brokering system in the tasks of update and retrieval of environmental information from distributed resources [26]. The ontology is meant to express at a semantic level concepts that are believed to be embedded in environmental databases – examples of concepts in the paper are ‘site’ and ‘contaminant’ in relation to a land contamination database. For this reason, it was not constructed through traditional time-consuming knowledge acquisition methods involving domain experts. Rather, the approach was to build on knowledge “hard-coded” on readily available environmental information sources such as database schemas, user queries, data dictionaries and standardised vocabulary. The ontology is represented as a conceptual graph and does not contain axiomatic knowledge.

There have been efforts in developing an Environmental Pollutants ontology [18], whose aim was to include concepts on pollutants of various media (water, air, soil, etc.), methods for detecting pollutants, and pollutants concentrations regulated by environmental legislation. A Monoatomic-Ions sub-ontology – ions are indicators of pollution – was built [40] partly through reuse of a chemical elements ontology. The representation language used in this work is Ontolingua [20], which combines KIF (Knowledge Interchange Format), a prefix monotonic first-order logic language extended with set theory [17], as the underlying formalism and the frame system formalism in the front-end.

The SEEK (Science Environment for Ecological Knowledge) project [43] has been developing a number of OWL (Web Ontology Language) ontologies [49] as part of an initiative designed to foster global access to ecological data and information and use of integrated analytical tools. Still, most environmental ontologies developed in recent years have a relatively low degree of formality, i.e., they are not as rich in axiomatic definitions as they are in the structuring of their collection of concepts.

3 Related Upper Level Ontologies

Upper ontologies provide a structured set of defined concepts and their relations that are general and broad enough to be referred to in the construction of domain-specific ontologies. In practical ontology engineering, they are used for identification of concepts to which concepts in the domain ontology are related, with a view towards reusing relevant specifications.

The establishment of relations between classes in our domain ontology and classes in upper level ontologies, prior to the axiomatisation of the former in FOL, has been done using the Ontolingua Server [11, 37], motivated by the reusability potential of its extensive library of shareable ontologies¹. Suitable classes were found in the HPKB Upper Level, OKBC, Physical Quantities and Standard Dimensions ontologies, which are available in such library.

The HPKB Upper Level ontology [14] is a product of the DARPA HPKB (High Performance Knowledge Base) programme [24, 9], whose objective was to produce technology to enable rapid construction of large knowledge bases that are comprehensive in domain coverage, reusable by multiple applications and easily maintainable [5]. The ontology is now part of the open source, upper level version of the generic ontology Cyc [28, 42], called OpenCyc, containing over 40,000 concepts of commonsense knowledge. OpenCyc, in turn, was one of the starter documents for the IEEE Standard Upper Ontology (SUO) effort [25, 38]. The OKBC ontology is also a product of the HPKB programme, consisting of an ontology of the generic frame protocol [6, 8]. Parts of it have been incorporated into SUO’s Suggested Upper Merged Ontology (SUMO).

SUMO is an upper level ontology built to provide a comprehensive, formal and application-independent foundation for domain-specific ontologies [32]. It is now considered a stable ontology and is available from [46]. The language used to specify the ontology is a version of KIF called SUO-KIF. SUMO was created through a carefully crafted merging of a number of publicly available ontologies and theories² [31], including ontologies in the library of the Ontolingua Server,

¹The resulting ontology, translated to Prolog syntax using the Server’s translator, was not readily manageable, in summary, because the reused ontologies together did not make up a coherent, minimal theory. The yielded specification had to be reengineered and manually re-specified for the most part, as shown in [44]. The whole process of engineering *Ecolingua* from conceptualisation to implementation in Prolog is reported in [4].

²A history of SUMO revisions as it incorporated ontologies and theories can be found at <http://ontology.teknowledge.com/annotatedHistory.html>.

the environment that we initially used to create *Ecolingua*'s structure. A suite of resources associated with SUMO are available, including: a knowledge engineering environment for development and inference support, domain ontologies, mappings to the WordNet lexicon, templates for translating SUMO into English and other natural languages, etc.

Ecolingua draws most heavily on the *EngMath* family of ontologies, which includes ontologies of physical quantities and standard dimensions. The seminal work on *EngMath* is [21]. Subsequently, the ontologies were specified and made available in the Ontolingua Server and from there incorporated into SUMO. More details on our reuse of *EngMath* are given throughout the remainder of the paper.

DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) [16] is yet another upper level ontology, which, together with other modules, comprises a library of foundational ontologies [29] developed within the WonderWeb project [50]. As its name suggests, DOLCE has a cognitive bias. Its aim is to capture concepts that underly human natural language and commonsense. Differently from SUMO, however, it has not been proposed as a universal and comprehensive standard ontology, but as a minimal formal specification (given in first-order logic with modality) of general and reusable upper-level categories to be exploited as reference for building new ontologies, including those in the WonderWeb library of foundational ontologies.

DOLCE has been applied in Ecology [27]. In this work, an ontology development approach is proposed which harnesses existing modelling practices in the domain. System dynamics is, indeed, a established and widely used modelling approach in Ecology [13], supported by a well-known software tool called STELLA³. In light of that, Keet presents possible formal mappings between STELLA modelling elements and ontology elements. Entities and relationships were identified in example STELLA models as candidates for ontology elements and then defined in terms of DOLCE as foundation ontology. For instance, stocks, which represent state variables in the models, mapped into the *endurant* category in DOLCE; flows, which represent processes that occur in an ecosystem, mapped into the *perdurant* category; converters, which represent influence variables, mapped into the *state* or *quality* categories. The resulting ontologies (one per example model) were then built using Protégé [15, 41] and OWL-DL. By contrast, *Ecolingua* consists of concept definitions suitable for annotating *data* in *Ecol*

ogy. These annotations can then be used to substantiate automated synthesis of conceptual system dynamics models that are consistent with the annotated data. Such application of *Ecolingua* is briefly presented in Sect. 9.

4 Quantities in Ecology and EngMath Reuse

The bulk of ecological data consists of numeric values representing measurements of attributes of entities and processes in ecological systems. The most intrinsic property of such a measurement value lies on the physical nature, or dimension, of what the value quantifies [10]. Differences between the measured objects (such as entities or processes) or the units of measure used do not necessarily determine that in their physical nature the measurements are diverse between them. Grams of algae and gigatons of metal, say, are not intrinsically different and their common physical dimension is implicit in 'grams' and 'gigatons', both units of the mass dimension. On the other hand, a measure of mass is intrinsically different from a measure of distance because they belong to different physical dimensions.

The understanding of this fundamental relation between ecological quantitative data and physical dimensions drew our attention towards the *EngMath* family of ontologies which provides a well-founded and well-defined conceptualisation of quantities and physical dimensions. *EngMath* has been designed with a view towards enabling sharing and reuse of engineering models. With this purpose in mind, knowledge sharing resources of the time were employed to specify *EngMath*, namely, the KIF language and the Web-based Ontolingua environment. More recently, *EngMath* ontologies have also been incorporated into the IEEE Suggested Upper Merged Ontology (SUMO). *EngMath*'s high degree of formality has rendered it one of the few ontologies that, according to the literature, have indeed been reused either in the engineering of new ontologies or in the design of applications that involve mathematical modelling [3, 48, 32, 1].

EngMath divides physical quantities into constant quantities and function quantities. Constant quantities are mappings of physical things to quantities, whereas function quantities are mappings, with any finite number of arguments, of constant quantities to other constant quantities (e.g. the altitude of a particle at particular times is a unary function quantity). The ontology also draws a distinction between scalars and higher-order tensors. Scalar quantities have size

³ISEE Systems, <http://www.iseesystems.com>

(or magnitude) but not direction, their magnitudes are real numbers and as such have a linear order. Higher-order tensors, such as vectors, cannot be fully characterised by ordered magnitude alone, they require an additional statement such as direction or orientation. Within the Ecology domain, all defined properties in EngMath’s conceptualisation of constant, scalar, physical quantities are applicable to ecological measurements, namely:

1. Every ecological measurement has an intrinsic physical dimension – e.g., vegetation biomass is of the mass dimension, the height of a tree is of the length dimension;
2. The physical dimension of an ecological measurement can be a composition of other dimensions through multiplication and exponentiation to a real power – e.g., the amount of a fertiliser applied to soil every month has the composite dimension mass/time;
3. Ecological measurements can be dimensionless – e.g., number of individuals in a population; and can be non-physical – e.g., profit from a fishing harvest;
4. Comparisons and algebraic operations (including unit conversion) can be meaningfully applied to ecological measurements, provided that their dimensions are homogeneous – e.g., you could add or compare an amount of some chemical to an amount of biomass as they are both of the mass dimension, but it would make no sense to add or compare, an amount of biomass to, say, an amount of time.

Also relevant to Ecolingua is the EngMath conceptualisation of units of measure. They are defined as physical quantities like all others, with the same properties. The only thing special about a unit of measure is that it is established by convention as an absolute amount of something to be used as a standard reference for quantities of the same dimension. Therefore, *one can identify the physical dimension of a quantity from the unit of measure in which it is expressed* [30]. Being Ecolingua an ontology for annotation of ecological data, instantiation of its terms, as we shall see in Sect. 7, requires the specification of a quantity’s unit of measure. In this way, annotating a data set in Ecolingua is straightforward in the sense that it is of course commonplace to have the units of measure of the data on hand, whereas the data’s physical dimensions are not part of the everyday vocabulary of ecologists and modellers.

Lower level Ecolingua (Sect. 8) includes an axiomatisation of units and scales of measurement, including their dimensions, base and derived units, and the SI system [47]. Such axiomatisation allows for automatic elicitation of a quantity’s physical dimension from the unit or scale in which it is specified. The magnitude (value) of physical quantities is defined in EngMath as “a binary function that maps a quantity and unit of measure to a numeric value (a dimensionless quantity) . . . [for example,] the magnitude of 50 kg in kilogrammes is 50” Ecolingua is not concerned with magnitudes as they play no role on the data physical dimensions.

5 Class Hierarchy and Axiomatisation

For identification of concepts and relations in the ecological data domain, it was taken as initial case study a data set generated from a tropical forest logging experiment in the Amazon, Brazil. This data set was chosen due to its diversity and large scale which should provide a good coverage of concepts, and also because it has been applied in modelling, underpinning the construction of a large simulation model used to support decision-making on sustainable logging strategies for forests in that part of the world [2].

The core of Ecolingua consists of a classification of quantitative ecological data, the main objects in the domain, according to their physical dimension – a fundamental property of all physical quantities. The terms of the defined classes are then employed (or are instantiated) to annotate specific data sets. The annotation is not done automatically. Human judgement determines the ontological classes and relations appropriate for the objects in each data set.

Ecolingua class hierarchy is shown in Fig. 1 with its two main sections of Contextual Data classes and Quantity classes.

The hierarchy comprises Ecolingua classes and external classes defined in other upper level ontologies, namely, Physical-Quantities and Standard-Dimensions of the EngMath family of ontologies, and Okbc-Ontology and Hpkb-Upper-Level (see Sect. 3). External classes are denoted `Class@Ontology`. The arcs in the hierarchy represent subclass relations, bottom up, e.g., the ‘Weight of’ class is a subclass of the ‘Quantity’ class. Ecolingua leaf classes are mutually disjoint.

We distinguish two different types of subclass relations, indicated by the bold and dashed arcs in Fig. 1. Bold arcs correspond to full, formal subclass relations. Dashed arcs correspond to relations we call *referential*

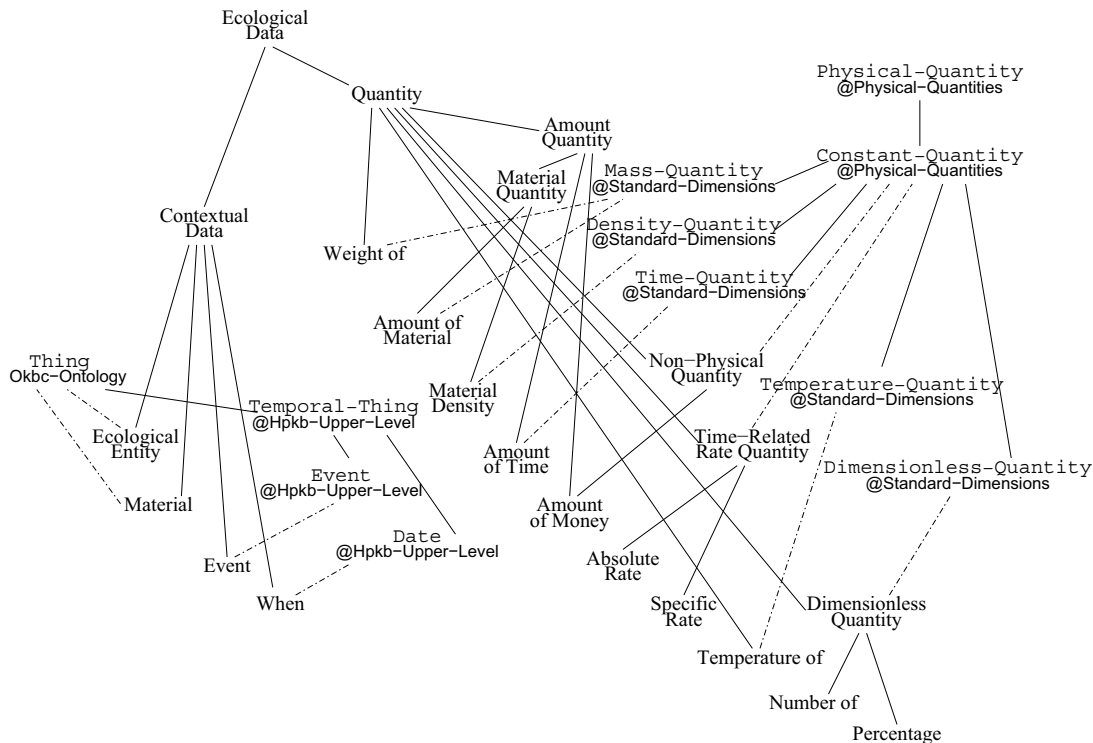


Figure 1: Ecolingua class hierarchy

between Ecolingua classes and external classes (mostly of the EngMath family of ontologies) in that they do not hold beyond the conceptual level, i.e., definitions that the (external) class involves are not directly incorporated by the (Ecolingua) subclass.

In the forthcoming definitions of Ecolingua classes, textual and axiomatic KIF definitions of their referential classes are given. Ecolingua axioms defining quantity classes incorporate the physical dimension, when specified, of its referential class in EngMath through the unit of measure concept, as explained in Sect. 4, and contextualise the quantity class in the Ecology domain through concepts such as ecological entity, compatibility between materials and entities, etc.

5.1 Rationale, Logical Form and Usage of Ecolingua Concepts

We consider Ecolingua a semi-formal ontology. Most of its concepts have textual and axiomatic definitions, the latter intended to more precisely constrain the interpretation, and thus appropriate uses, of the concepts. The ontology does not qualify as fully formal, in that logical properties such as soundness and completeness are not claimed or proved. Some Ecolingua

concepts have a hard-to-define meaning, such as *ecological entity* and *material*. Such difficulty explains the scarcity of axiomatic definitions in general ontologies of commonsense knowledge.

Ecolingua axioms are represented as first-order logic, well-formed formulae of the form $Cpt \rightarrow Ctt$. That is, if Cpt holds then Ctt must hold. Variables not explicitly quantified in the axioms are assumed universally quantified. The antecedents, Cpt , are atomic sentences representing an Ecolingua concept consisting of a predicate and the terms over which it applies. The Cpt sentences make up Ecolingua vocabulary. One annotates an ecological data set by instantiating these sentences. The consequents, Ctt , are logical sentences that constrain the interpretation of each Cpt , consisting of predicates over Cpt 's terms that must hold when Cpt does. Most relations in Ctt are “internal”, or non-descriptive, in that they are not used as vocabulary to annotate data. We call the set of these relations lower level Ecolingua, which is shown in Sect. 8.

To make the above clearer, let us exemplify how Ecolingua annotation terms can be used in practice. Consider Table 1. The table is typical of ecological data in format and content: it is filled with values of a variable of interest measured under certain condi-

<i>Net Production Rate of Plants (g Produced/kg Plant Biomass-Day)</i>					
Water Temperature (°C)	Net Production				
	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5
10	0.225	0.422	0.072	0.358	0.002
15	0.639	0.601	0.578	0.143	0.231
20	0.595	0.353	1.180	0.377	1.555

Table 1: Data on processes occurring within a pond system including the net production rate of plants as a function of water temperature (extract from [19])

tions. Also, it consists of processed data, resulting of the application of some statistical method to extensive raw data collected in sampling campaigns. We assume processed data is given for *Ecolingua*-annotation.

At the level of quantity conceptualisation, two quantities play a role in Table 1, namely, net production rate of plants and water temperature, with a functional relationship between them. The two quantities map into the *Ecolingua* quantity classes ‘Specific Rate’ (Sect. 7.2) and ‘Temperature of’ (Sect. 7.3), respectively. The water temperature quantity, for instance, could be annotated in *Ecolingua* as `temperature_of(water_temp, water, °C)`. This would be a data annotation compliant with *Ecolingua* under the interpretation specified in the concept’s axiom, i.e., that water is an ecological entity and that °C is a scale of temperature. The axioms that define these concepts must also be solved for the water temperature data annotation to be established. In turn, the functional relationship between water temperature and net production rate of plants expressed in the table can be annotated as `influences(water_temp, plant_production_rate, ?)`.

6 Contextual Data

An ecological quantity alone does not tell much if it is not contextualised, if, for example, no reference is given on what it quantifies. We define a few general concepts of contextual data around the ecological quantities which form the greater part of the ontology. The concepts are the following:

- The **Ecological Entity** class – `eco_entity(E)` – e.g. of instantiated sentence: `eco_entity(vegetation)`.

Referential subclass of the *Thing@Okbc-Ontology* class (Fig. 1). We quote here an excerpt of the natural-language definition given in *Okbc-Ontology*: “**Thing** is the class of everything in the

universe of discourse that can be in a class. This includes ... all other objects defined in user ontologies. **Thing** is ... the practical root of all ontologies.” (Also denoted Entity and \top [45].)

An ecological entity *E* is any distinguishable thing, natural or artificial, with attributes of interest in an ecological system (e.g., vegetation, water, an animal, a group of individuals or population, machinery), the system itself (e.g., a forest, a lake), or its boundaries (e.g., atmosphere). Ecological quantities usually consist of measurements of attributes of entities – e.g., carbon content of vegetation, temperature of an animal’s body, birth rate of a population, volume of water in a lake.

- The **Material** class – `material(Mt)` – e.g. of instantiated sentence: `material(biomass)`.

Also a referential subclass of the *Thing@Hpkb-Upper-Level* class (Fig. 1).

A material *Mt* is anything that has mass and can be contained in an ecological entity – e.g., biomass, chemicals, timber.

- The **Compatibility** relation between materials and entities – `compatible(Mt, E)` – e.g. of instantiated sentence: `compatible(biomass, vegetation)`.

A material and entity are compatible if it occurs in nature that the entity contains the material. For example, biomass is only thought of in relation to living entities (plants and animals), not in relation to inorganic things.

If Mt and E are compatible then Mt is a material and E is an entity:

$$\text{compatible}(Mt, E) \rightarrow \text{material}(Mt) \wedge \text{eco_entity}(E)$$

- The **Event** class – `event(Ev)` – e.g. of instantiated sentence: `event(rainy_season)`.

Referential subclass of the `Event@Hpkb-upper-level` class (Fig. 1) defined only in natural language as “[`Event`] is one important subset of `Temporal-Thing` (the collection of all things which have a particular temporal extent, things about which one might sensibly ask ‘When?’). The elements of `Event` are events or actions, things that we say are ‘happening’, changes in the state of the world ...”

An event Ev is any happening of ecological interest with a time duration – e.g., seasons, sampling campaigns, logging or harvest events.

- The **When** class – `when(T)` – e.g. of instantiated sentence: `when(13_Oct_2005)`.

Referential subclass of the `Date@Hpkb-upper-level` class (Fig. 1) defined only in natural language as “a `Date` is any `Time-Interval` which can be defined purely by its location on the calendar. Thus a `Date` could be a particular calendar day, ..., a particular calendar month, a particular decade, etc.”

A `when T` is any description that identifies a moment or period of time – e.g., 7.55 pm on the 5th Aug. 2002, 1990–2000 decade.

- The **Time of Event** relation – `time_of_event(Ev, T)` – e.g. of instantiated sentence: `time_of_event(rainy_season, dec_to_may)`.

Every event occurs at some moment or period of time – e.g., months of the rainy season, harvest date.

If a time of event Ev is T then Ev is an event and T is a when description:

$$\text{time_of_event}(Ev, T) \rightarrow \\ \text{event}(Ev) \wedge \text{when}(T)$$

7 Quantitative Data

Herein we show the defined concepts in `Ecolingua` that represent classes of quantitative data, as well as some relationships between them.

7.1 Amount Quantity

Many quantities in Ecology represent an amount of something contained in a thing or place, for example, carbon content in leaves, water in a lake, energy stored

in an animal’s body. Clearly, not exclusive to Ecology, amount of something appears to be a commonsense notion having had mention in works on qualitative modelling [12] and commonsense reasoning [7, 16].

7.1.1 Material Quantity

Quantities that represent an amount of material things are of the mass dimension (intuitively a ‘quantity of matter’ [30]). For such quantities, the amount of material class is defined as a referential subclass of the `Mass-Quantity@Standard-Dimensions` class defined in KIF as:

$$(\Leftrightarrow (\text{Quantity-Dimension } ?X \text{ Mass-Dimension}) \\ (\text{Mass-Quantity } ?X))$$

- **Amount of Material** class – *If A identifies a measure of amount of material Mt in E specified in U then Mt is a material, E is an entity which is compatible with Mt , and U is a unit of mass (e.g. of instantiated sentence: `amt_of_mat(tree_timber, timber, tree, kg)`):*

$$\text{amt_of_mat}(A, Mt, E, U) \rightarrow \\ \text{material}(Mt) \wedge \text{eco_entity}(E) \wedge \\ \text{compatible}(Mt, E) \wedge \text{mass_unit}(U)$$

Other quantities represent measurements of amount of material in relation to space, e.g., amount of biomass in a crop acre, or of timber harvested from a hectare of forest. The dimension of such quantities is mass over some power of length, most commonly area (length^2) and volume (length^3). For these quantities, we define the material density class as a referential subclass of the `Density-Quantity@Standard-Dimensions` class defined in KIF as:

$$(\Leftrightarrow (\text{Density-Quantity } ?X0) \\ (\text{Quantity-Dimension } ?X0 \text{ Density-Dimension}))$$

- **Material Density** class – *If A identifies a measure of density of Mt in E specified in U then Mt is a material, E is an entity which is compatible with Mt , and U is equivalent to an expression Um/Ul , where Um is a unit of mass and Ul is a unit of some power of length (e.g. of instantiated sentence: `mat_dens(fish_biomass, biomass, fish, kg/ha)`):*

$$\text{mat_dens}(A, Mt, E, U) \rightarrow \\ \text{material}(Mt) \wedge \text{eco_entity}(E) \wedge \\ \text{compatible}(Mt, E) \wedge \\ \exists Um, Ul . \text{eqv_expr}(U, Um/Ul) \wedge \\ \text{mass_unit}(Um) \wedge \\ \text{length}^n_unit(Ul)$$

The `eqv_expr` relation used above and in upcoming axioms holds between expressions that are syntactically identical or between expressions that might be written differently but are algebraically equivalent, for example, the expressions $Um \times (1/Ul)$ and $Ul^{-1} \times Um$ are equivalent to Um/Ul .

7.1.2 Amount of Time

Quantities can also represent amounts of immaterial things, time being a common example. The duration of a sampling campaign and the gestation period of females of a species are examples of ecological quantities of the amount of time class. The class is a referential subclass of the `Time-Quantity@Standard-Dimensions` class defined in KIF as:

$$(\Leftrightarrow (\text{Quantity.Dimension } ?X \text{ Time-Dimension}) \\ (\text{Time-Quantity } ?X))$$

- **Amount of Time** class – *If A identifies a measure of an amount of time of Ev specified in U then Ev is an event and U is a unit of time* (e.g. of instantiated sentence: `amt_of_time(rainy_season_duration, rainy_season, month)`):

$$\text{amt_of_time}(A, Ev, U) \rightarrow \text{event}(Ev) \wedge \text{time_unit}(U)$$

7.1.3 Non-Physical Quantity

Despite the name, the ‘physical quantity’ concept in `EngMath` allows for so-called non-physical quantities. These are quantities of new or non-standard dimensions which can be defined preserving all the properties of physical quantities (Sect. 4). That is, the quantities have an intrinsic dimension which makes them comparable and which is amenable to algebraic operations and combination with other dimensions. In [21] a monetary dimension is given as an example of a valid non-physical dimension: “one can accumulate sums of money, do currency conversion, compare relative wealth” and one can combine amount of money with time, for example, to give a rate of inflation.

The class of non-physical quantities is a referential subclass of `Constant-Quantity@Physical-Quantities` class defined as: “A `Constant-Quantity` is a constant value of some `Physical-Quantity`, like 3 meters or 55 miles per hour. ...” Ecological data often involve measurements of money concerning some economical aspect of the system-of-interest, e.g., profit given by a managed natural system or balance of a cooperative’s bank account.

- **Amount of Money** class – *If A identifies a measure of amount of money in E specified in U then E*

is an entity and U is a unit of money (e.g. of instantiated sentence: `amt_of_money(profit, fish_pond, $)`):

$$\text{amt_of_money}(A, E, U) \rightarrow \\ \text{eco_entity}(E) \wedge \text{money_unit}(U)$$

7.2 Time-Related Rate Quantity

In general, rates express a quantity in relation to another. In Ecology, rates commonly refer to instantaneous measures of processes of movement or transformation of something occurring over time, for example, decay of vegetation biomass every year, consumption of food by an animal each day. `Ecolingua` defines a class of rate quantities of composite physical dimension, including time, as a referential subclass of the `Constant-Quantity@Physical-Quantities` class, which is a generalisation of dimension-specific quantity classes (Fig. 1).

The absolute rate class is for measures of processes where an amount of some material is processed over time. These quantities have a composite dimension of mass, $\text{mass}/\text{length}^n$ or money (the dimensions of amount quantities with the exception of time) over the time dimension. To the mass or $\text{mass}/\text{length}^n$ dimensions will correspond adequate units of measure (e.g., kg, ton/ha) which we call units of material.

- **Absolute Rate** class – *If R identifies a measure of the rate of processing Mt from E_{from} to E_{to} specified in U then Mt is a material, E_{from} and E_{to} are entities which are different from each other and compatible with Mt, and U is equivalent to an expression Ua/Ut , where Ua is a unit of material and Ut is a unit of time* (e.g. of instantiated sentence: `abs_rate(consumption, biomass, algae, fish, kg)`):

$$\text{abs_rate}(R, Mt, E_{\text{from}}, E_{\text{to}}, U) \rightarrow \\ \text{material}(Mt) \wedge \text{eco_entity}(E_{\text{from}}) \wedge \\ \text{eco_entity}(E_{\text{to}}) \wedge E_{\text{from}} \neq E_{\text{to}} \wedge \\ \text{compatible}(Mt, E_{\text{from}}) \wedge \text{compatible}(Mt, E_{\text{to}}) \wedge \\ \exists Ua, Ut . \text{eqv_expr}(U, Ua/Ut) \wedge \\ (\text{mat_unit}(Ua) \vee \text{money_unit}(Ua)) \wedge \\ \text{time_unit}(Ut)$$

Sometimes, processes are measured in relation to an entity involved in the process. We call these measures specific rates. Let us take the example of food consumption by animals. Suppose measures of how much food the animals consume are given in kg/day. That would be the absolute rate of consumption. Now, the rate of food consumption in relation to the animals’ weight would be a specific rate. For example, a measure given in, say, kg/kg/day is a specific rate meaning

how much food in kilograms per kilogram of the animals' weight is consumed per day.

Ecology people refer to process rates meaning, interchangeably, their absolute or specific forms [19]. Specific rates are usually expressed as a dimensionless quantity (a unitless number) over time. Such representation leaves implicit that the dimensionless quantity is in fact a ratio of two quantities of the same dimension whose units have been cancelled out. This is illustrative of representational issues that can benefit from ontologies. Ecolingua makes this hidden assumption explicit, and in this way ensures that a specific rate's dimension is correctly composed of a ratio of equal dimensions over the time dimension.

- **Specific Rate** class – *If R identifies a measure of a specific rate, related to R_{abs} , of processing Mt specified in U then: R_{abs} measures the absolute rate of processing Mt from E_{from} to E_{to} specified in U_{abs} , which is an expression equivalent to Ua/Ut where Ua is a unit of measure of material; and U is equivalent to an expression $Ub/Uc/Ut$ where both Ub and Uc are units of measure of material and are of the same dimension D (e.g. of instantiated sentence: `spf_rate(consumption_rate, consumption, biomass, kg/kg/day)`):*

$$\begin{aligned} \text{spf_rate}(R, R_{\text{abs}}, Mt, U) \rightarrow & \exists E_{\text{from}}, E_{\text{to}}, U_{\text{abs}}, Ua, Ut . \\ & \text{abs_rate}(R_{\text{abs}}, Mt, E_{\text{from}}, E_{\text{to}}, U_{\text{abs}}) \wedge \\ & \text{eqv_expr}(U_{\text{abs}}, Ua/Ut) \wedge \text{mat_unit}(Ua) \wedge \\ & \exists Ub, Uc, D . \text{eqv_expr}(U, Ub/Uc/Ut) \wedge \\ & \text{mat_unit}(Ub) \wedge \text{mat_unit}(Uc) \wedge \\ & \text{unit_dimension}(Ub, D) \wedge \\ & \text{unit_dimension}(Uc, D) \end{aligned}$$

Note that the axiom does not constrain the units of the absolute and specific rates, U and U_{abs} , to be of the same dimension. One can be a unit of mass/time while the other is a unit of mass/lengthⁿ/time. However, the dimensions of the two units, Ub and Uc , part of the specific rate's composite unit of measure, must be the same. For example, the absolute rate could be specified in mg/ha/day and the specific rate in g/kg/day.

7.3 Temperature Quantity

Another fundamental physical dimension is temperature, which has measurement scales rather than units [30] (see Sect. 8.3). Temperature in a green house or of water in a pond, are two examples of temperature quantities in ecological data sets. The referential superclass of the class below is

Temperature-Quantity@Standard-Dimensions defined in KIF as:

$$\begin{aligned} (\Leftrightarrow & (\text{Temperature-Quantity } ?X0) \\ & (\text{Quantity.Dimension } ?X0 \\ & \text{Thermodynamic-Temperature-Dimension})) \end{aligned}$$

- **Temperature** of class – *If T identifies a measure of the temperature of E specified in S then E is an entity and S is a scale of temperature (e.g. of instantiated sentence: `temperature_of(pond_water_temp, pond_water, °C)`):*

$$\begin{aligned} \text{temperature_of}(T, E, S) \rightarrow \\ \text{eco_entity}(E) \wedge \text{temperature_scale}(S) \end{aligned}$$

7.4 Weight (of something) Quantity

Strictly speaking, weight is a gravitational force with which a body tends towards the centre of the earth, i.e., a composite physical dimension of the form mass \times length \times time⁻². But in Ecology, as in many other contexts, when colloquially referring to how heavy a thing is, by 'weight' people mean a quantity of mass. For example, the weight of an animal, the weight of a fishing harvest. It is in this everyday sense of weight that we define a class of weight (of something) quantities, which has Mass-Quantity@Standard-Dimensions as referential superclass.

- **Weight** of class – *If W identifies a measure of the weight of E specified in U then E is an entity and U is a unit of mass (e.g. of instantiated sentence: `weight_of(indiv_fish_weight, individual_fish, kg)`):*

$$\text{weight_of}(W, E, U) \rightarrow \text{eco_entity}(E) \wedge \text{mass_unit}(U)$$

Note that for quantities of both this class and the Amount of Material class, the specified unit must be a unit of mass. But the intuition of a measure of weight does not bear a containment relationship between a material and an entity like the intuition of an amount of material does.

7.5 Dimensionless Quantity

Another paradoxically named concept in the Eng-Math ontology is that of dimensionless quantities, in that they do have a dimension, the so-called identity dimension. Real numbers are an example. The class of dimensionless quantities has a referential superclass of the same name, Dimensionless-Quantity@Physical-Quantities, defined in KIF as:

$$\begin{aligned}
& (\Leftrightarrow (\text{Dimensionless-Quantity } ?X) \\
& \quad (\text{And } (\text{Constant-Quantity } ?X) \\
& \quad \quad (= (\text{Quantity.Dimension } ?X) \\
& \quad \quad \quad \text{Identity-Dimension})))
\end{aligned}$$

This concept applies to quantities in Ecology that represent counts of things, such as number of individuals in a population or age group.

- **Number of** class – *If N measures the number of E specified in U then E is an entity and N is a dimensionless quantity specified in U (e.g. of instantiated sentence: `number_of(fish_stocked, fish, 1)`):*

$$\begin{aligned}
& \text{number_of}(N, E, U) \rightarrow \\
& \quad \text{eco_entity}(E) \wedge \text{dimensionless_qty}(N, U)
\end{aligned}$$

Percentages can also be defined as dimensionless quantities. Food assimilation efficiency of a population, mortality and birth rates are examples of ecological quantities expressed as percentages.

- **Percentage** class – *If P is a percentage that quantifies an attribute of E specified in U then E is an entity and P is a dimensionless quantity specified in U (e.g. of instantiated sentence: `percentage(mortality_rate, fish, 1)`):*

$$\begin{aligned}
& \text{percentage}(P, E, U) \rightarrow \\
& \quad \text{eco_entity}(E) \wedge \text{dimensionless_qty}(P, U)
\end{aligned}$$

7.6 Influence relation and constancy of quantities

Besides quantities themselves, influences that hold (or are suspected to hold) between quantities are also part of ecological data specifications. They are denoted in expressions such as ‘A is dependent on B’, ‘A is a function of B’, ‘A increases with B’, etc., and in data tables, like Table 1, where values of a quantity A are shown as a function of some other quantity B. The axiom below defines such relations between quantities.

- **Influence** relation between quantities – *If Q influences Q' where the influence is qualified with $Sign$ then Q and Q' are different quantities of defined classes, the sign $Sign$ of the influence is positive, negative or undetermined, and Q' is not a constant quantity (e.g. instantiated sentence: `influences(harvest, tree_timber, -)`):*

$$\begin{aligned}
& \text{influences}(Q, Q', Sign) \rightarrow \\
& \quad \text{eco_qty}(Q) \wedge \text{eco_qty}(Q') \wedge Q \neq Q' \wedge \\
& \quad \text{Sign} \in \{+, -, ?\} \wedge \neg \text{constant}(Q')
\end{aligned}$$

A positive influence sign means that the two quantities change in the same direction, e.g., an influence between an absolute rate ‘plant production’, the influencer quantity, and an amount of material ‘plant biomass’, the influenced quantity: the more plant production the more biomass, the less plant production the less biomass. The influence is negative when the two quantities change in opposite directions, e.g., influencer ‘plant mortality’ (also an absolute rate) and influenced ‘plants biomass’. A ‘?’ sign means the direction of change is unknown or unclear, like when there may be threshold values of the quantities where the direction may reverse, e.g., influencer ‘water temperature’ (a temperature quantity) and influenced ‘plant production’.

The property of constancy of quantities is defined as a contrapositive of the influence relation. If a quantity is constant then it does not suffer influence of any other.

- **Constancy** property of quantities – *If Q is constant then Q is a quantity of a defined class and there is no other quantity Q' that influences Q (e.g. of instantiated sentence: `constant(fish_stocked)`):*

$$\begin{aligned}
& \text{constant}(Q) \rightarrow \\
& \quad \text{eco_qty}(Q) \wedge \\
& \quad \neg \exists Q', Sign . Q' \neq Q \wedge \text{influences}(Q', Q, Sign)
\end{aligned}$$

8 Lower Level Ecolingua

Lower level Ecolingua comprises the subset of the ontology’s concepts which are not descriptive, i.e., not directly employed to annotate ecological data instances. The definitions in this level compose the interpretation constraints of the annotation concepts in Sect. 7. The axiomatisation of lower level Ecolingua consists of a rendition of parts of the EngMath ontology, augmented in specification detail⁴.

The Suggested Upper Merged Ontology (SUMO) has a Units of Measure subontology [39] which deals with physical measurements and has also been derived mostly from EngMath. As we do in Ecolingua, it takes the SI system as the base system of units. The coverage of units of measure used in different domains provided by this ontology is much more extensive than what we present here.

Our definitions in lower level Ecolingua concerning physical quantities and their units and scales of measure are written on the basis of the intrinsic relation

⁴The EngMath quotations and references in the sequel are from [21].

that exists between a unit of measure and its physical dimension. Given a unit in which some ecological quantity has been measured, the axiomatisation that follows allows the physical dimension of such quantity to be elicited, such as mass, length, time, or a composite dimension, such as mass/area (a density dimension). Once its physical dimension is known, the quantity can be associated with ecological model elements: quantities of certain dimensions are suitable to fill certain types of model elements. This, in short, is the role of **Ecolingua** in an ecological model synthesis application (Sect. 9), where the axioms given here in Horn clause notation are employed assuming a resolution-based system.

8.1 Ecolingua Quantities and their Units and Scales

The next clause specifies all defined **Ecolingua** quantities and links up the annotation, higher level ontology terms with their lower level definitions. As seen earlier, the fundamental characteristic of a quantity is its physical dimension, when it has one. Basic physical dimensions are orthogonal, which leads to the definition of types of units of measure according to such dimensions. An `identity_dimension` constant is defined to specify the dimension of “dimensionless” quantities.

Ecolingua quantities

Q is an **Ecolingua** quantity if Q identifies an instance of any of the **Ecolingua** quantity classes.

$$\begin{aligned} \text{eco_qty}(Q) \leftarrow & \\ & \text{amt_of_mat}(Q, Mt, E, U) \vee \text{mat_dens}(Q, Mt, E, U) \vee \\ & \text{amt_of_time}(Q, U) \vee \text{amt_of_money}(Q, E, U) \vee \\ & \text{abs_rate}(Q, Mt, E_{\text{from}}, E_{\text{to}}, U) \vee \\ & \text{spf_rate}(Q, R_{\text{abs}}, Mt, U) \vee \text{temperature_of}(Q, E, S) \vee \\ & \text{weight_of}(Q, E, U) \vee \text{number_of}(Q, E, U) \vee \\ & \text{percentage}(Q, E, U) \end{aligned}$$

Dimensionless Quantity

Q is a dimensionless quantity specified in U if the dimension of unit U is the identity dimension.

$$\begin{aligned} \text{dimensionless_qty}(Q, U) \leftarrow & \\ & \text{unit_dimension}(U, \text{identity_dimension}) \end{aligned}$$

Unit of mass

U is a unit of mass if U is a unit of measure of the mass dimension.

$$\begin{aligned} \text{mass_unit}(U) \leftarrow & \\ & \text{unit_of_measure}(U) \wedge \text{unit_dimension}(U, \text{mass}) \end{aligned}$$

Unit of power of length

U is a unit of a power of length if U is a unit of measure of the length dimension (power is 1) or, more generally, of the length to a real power dimension.

$$\begin{aligned} \text{length}^n\text{_unit}(U) \leftarrow & \\ & \text{unit_of_measure}(U) \wedge \\ & \left(\text{unit_dimension}(U, \text{length}) \vee \right. \\ & \left. (\text{unit_dimension}(U, \text{length}^N) \wedge \text{realnum_expr}(N)) \right) \end{aligned}$$

Unit of material

U is a unit of material if U is a unit of mass or if U is equivalent to an expression Um/Ul , where Um is a unit of mass and Ul is a unit of a power of length.

$$\begin{aligned} \text{mat_unit}(U) \leftarrow & \text{mass_unit}(U) \\ \text{mat_unit}(U) \leftarrow & \\ & \text{eqv_expr}(U, Um/Ul) \wedge \text{mass_unit}(Um) \wedge \\ & \text{length}^n\text{_unit}(Ul) \end{aligned}$$

Unit of money

U is a unit of money if U is a unit of measure of the money dimension.

$$\begin{aligned} \text{money_unit}(U) \leftarrow & \\ & \text{unit_of_measure}(U) \wedge \text{unit_dimension}(U, \text{money}) \end{aligned}$$

Unit of time

U is a unit of time if U is a unit of measure of the time dimension.

$$\begin{aligned} \text{time_unit}(U) \leftarrow & \\ & \text{unit_of_measure}(U) \wedge \text{unit_dimension}(U, \text{time}) \end{aligned}$$

Identity unit

EngMath: `Identity-Unit` is an individual in the `Physical-Quantities` ontology defined in KIF as (= `Identity-Unit 1`).

`identity_unit(1)`

Scale of temperature

S is a temperature scale if S is a scale of the temperature dimension.

`temperature_scale(S)` \leftarrow
`scale_dimension(S , temperature)`

8.2 Units of Measure

As seen in Sect. 7, the definitions of quantity classes involve units and scales of measure. In this section, we start presenting an axiomatisation of these special quantities, which is important and practical for applications of `Ecolingua` because it allows for the elicitation of the physical dimension of a given quantity from the unit or scale in which it is expressed.

U is a unit of measure if U is the base unit for a fundamental dimension in the defined system of units (Sect. 8.5).

`unit_of_measure(U)` \leftarrow `bu_fdim(U , D)`

Units named by convention are units of measure too [30]. A conventional unit of mass:

`unit_of_measure(g)`

Some conventional units of length:

`unit_of_measure(cm)`
`unit_of_measure(mile)`
`unit_of_measure(ha)`

And a couple of conventional units of time:

`unit_of_measure(hour)`
`unit_of_measure(year)`

EngMath: Every product of units is also a unit of measure, and the product is commutative.

`unit_of_measure(U)` \leftarrow
`eqv_expr(U , $V \times W$)` \wedge
`unit_of_measure(V)` \wedge `unit_of_measure(W)`

EngMath: Every real-valued exponentiation of a unit is also a unit of measure.

`unit_of_measure(U)` \leftarrow
`eqv_expr(U , V^X)` \wedge $U \neq V$ \wedge
`realnum_expr(X)` \wedge `unit_of_measure(V)`

EngMath: Units of measure are positive.

`unit_of_measure(U)` \leftarrow
`eqv_expr(U , $X \times V$)` \wedge `arithm_expr(X)` \wedge
`unit_of_measure(V)` \wedge `eval(X , 1)`

Unit identities

Identities of conventional units specify how they are expressed in terms of other units. The \equiv symbol means that, for example, g is not merely equal to or equivalent to $0.001 \times \text{kg}$ but that g *is* $0.001 \times \text{kg}$ [30]. Examples of identities are shown below for the mass, length and time dimensions.

$g \equiv Expr$ \leftarrow `eqv_expr($Expr$, $0.001 \times \text{kg}$)`
 $ha \equiv Expr$ \leftarrow `eqv_expr($Expr$, $10000 \times \text{m} \times \text{m}$)`
 $\text{year} \equiv Expr$ \leftarrow `eqv_expr($Expr$, $365 \times \text{day}$)` \vee
`eqv_expr($Expr$, $366 \times \text{day}$)`

8.3 Scales of Measure

Scales in general are not the same as units of measure. They are not a standard quantity against which other quantities of the same physical dimension can be compared. They have reference origins and can have negative values. Units are like distances between points on a scale [21].

Scales subsume units of measure in that the latter can be seen as scales of type ratio which are characterised by having an absolute zero. For example, K (kelvin) used for measuring temperature is an absolute (or ratio) scale and thus can also be called a unit of measure; $^{\circ}\text{C}$ (degree Celsius) and $^{\circ}\text{F}$ (degree Fahrenheit), on the other hand, are scales that, rigorously speaking, should not be called units of measure because they do not have an absolute zero [30, 10].

This distinction between scales and units is made in the `EngMath` ontology. In `Ecolingua`, we define K as a unit of measure (Sect. 8.5) and define $^{\circ}\text{C}$ and $^{\circ}\text{F}$ through scale correspondences in relation to K :

`scale_corresp('K', $^{\circ}\text{C}$, X $^{\circ}\text{C} = (273.15 + X)$ 'K')`
`scale_corresp($^{\circ}\text{C}$, $^{\circ}\text{F}$, X $^{\circ}\text{C} = (9/5 \times X + 32)$ $^{\circ}\text{F}$)`

$$\begin{aligned} \text{scale_corresp}(S, S', \text{Formula}) &\leftrightarrow \\ \text{scale_corresp}(S', S, \text{Formula}) & \end{aligned}$$

We cannot have identities between the temperature scales (like we have for units) because the zeros of the scales do not coincide. The left and right-hand sides of the equality in the formulae above represent points in the different scales at the same temperature.

8.4 Relating Units and Scales of Measure to their Physical Dimensions

The relation between units or scales of measure and their physical dimensions does not only take place between base units and the corresponding fundamental dimensions in the defined system of units (e.g., length and m (meter) in SI). Other units can be derived linked to the base units and physical dimensions can be composed as well (e.g., km/h for speed, which is a composition of the length and time dimensions).

Dimensions of Units

Unit U is of dimension D if, within a defined system of units (Sect. 8.5), U is the base unit of a fundamental dimension D ; or if U is a derived unit of a fundamental or composite dimension D ; or if U is a non-physical unit of a fundamental or composite dimension D .

$$\begin{aligned} \text{unit_dimension}(U, D) &\leftarrow \text{bu_fdim}(U, D) \\ \text{unit_dimension}(U, D) &\leftarrow \\ &\text{du_fdim}(U, D) \vee \text{du_cdim}(U, D) \\ \text{unit_dimension}(U, D) &\leftarrow \\ &\text{npu_fdim}(U, D) \vee \text{npu_cdim}(U, D) \end{aligned}$$

And, in the case of a unit of the identity dimension, we render the KIF axiom

$$\begin{aligned} (= (\text{Quantity.Dimension Identity-Unit}) \\ \text{Identity-Dimension}) \end{aligned}$$

of the Physical-Quantities ontology as:

$$\begin{aligned} \text{unit_dimension}(U, \text{identity_dimension}) &\leftarrow \\ \text{identity_unit}(U) & \end{aligned}$$

Dimension of the Temperature Scales

Scale S is of the temperature dimension if S is a unit of measure which is the base unit for the fundamental

dimension temperature in the defined system of units or if it is a derived unit for temperature.

$$\begin{aligned} \text{scale_dimension}(S, \text{temperature}) &\leftarrow \\ \text{unit_of_measure}(S) &\wedge \\ \left(\text{bu_fdim}(S, \text{temperature}) \vee \right. & \\ \left. \text{du_fdim}(S, \text{temperature}) \right) & \end{aligned}$$

Scale S is of the temperature dimension if there is a formula of correspondence between a measure in S and a measure in S' and S' is of the temperature dimension.

$$\begin{aligned} \text{scale_dimension}(S, \text{temperature}) &\leftarrow \\ \text{scale_corresp}(S, S', \text{Formula}) &\wedge \\ \text{scale_dimension}(S', \text{temperature}) & \end{aligned}$$

8.5 System of Units

EngMath: “A system of units is a class of units defined by composition from a base set of units, such that every instance of the class is the ‘standard’ unit for a physical dimension and every physical dimension has an associated unit To define a system of units, the model builder chooses a set of fundamental dimensions that are orthogonal (i.e., not composable from each other) . . . For each of the fundamental dimensions, the model builder chooses a standard unit of that dimension; these are called the base-units of the system . . . For example, SI is a system of units that defines a set of seven fundamental dimensions with the base-units meter [m, for length], kilogram [kg, for mass], second [s, for time], ampere [A, for electric current], Kelvin [K, for temperature], mole [mol, for amount of substance], and candela [cd, for luminous intensity].”

Mass, length, time, temperature and money are common dimensions in ecological data and are chosen to compose the current set of fundamental dimensions in Ecolingua. Base units are assigned to this set based on SI and other derived units can be composed in relation to them.

Base units of fundamental dimensions

For the base units we borrow kg for mass, m for length, s for time and K for temperature from SI and define the base unit \$ for the money dimension.

`bu_fdim(kg, mass)`
`bu_fdim(m, length)`
`bu_fdim(s, time)`
`bu_fdim('K', temperature)`
`bu_fdim($, money)`

EngMath: “The set of fundamental dimensions determines the space of possible quantities that can be described in this system — those whose physical dimensions are some algebraic combination of the fundamental dimensions.”

Derived units are part of this space of possible quantities. Direct comparison with a standard (through one of the base units) is feasible for only a limited number of physical quantities. So the units for most quantities must be derived from a few base units [30].

EngMath: Every derived unit in the system is a composition, using multiplication and exponentiation to a real power, of units from the base set.

In the sequel, we give axioms that generically define units of fundamental or composite dimensions that are derived from the base units. This is a notable difference between *Ecolingua* and SUMO’s Units of Measure subontology which lacks definitions such as these [39, 46]. The axioms allow establishing the relation between any valid derived unit and its fundamental or composite dimension in the base system (provided that derived units with special names are included in the ontology).

Derived units of fundamental dimensions

U is a derived unit of a fundamental dimension D if an identity exists between U and an expression $Expr$ which is a derived unit of D .

$$\begin{aligned} \text{du_fdim}(U, D) \leftarrow \\ U \equiv Expr \wedge \text{du_fdim}(Expr, D) \end{aligned}$$

U is a derived unit of a fundamental dimension D if U is equivalent to an expression $X \times V$ where X is a positive real number and V is the base unit or a derived unit of D .

$$\begin{aligned} \text{du_fdim}(U, D) \leftarrow \\ \text{eqv_expr}(U, X \times V) \wedge U \neq V \wedge \\ \text{posrealnum_expr}(X) \wedge \\ (\text{bu_fdim}(V, D) \vee \text{du_fdim}(V, D)) \end{aligned}$$

Derived units of composite dimensions

U is a derived unit of a composite dimension D if an identity exists between U and an expression $Expr$ which is a derived unit of D .

$$\text{du_cdim}(U, D) \leftarrow U \equiv Expr \wedge \text{du_cdim}(Expr, D)$$

The composition of a derived unit accord with the composition of the dimension. Physical dimensions are composed from other dimensions using multiplication and exponentiation to a real power.

$$\begin{aligned} \text{du_cdim}(U, D) \leftarrow \\ \text{eqv_expr}(U, V \times W) \wedge \text{eqv_expr}(D, A \times B) \wedge \\ \text{unit_dimension}(V, A) \wedge \text{unit_dimension}(W, B) \\ \text{du_cdim}(U, D) \leftarrow \\ \text{eqv_expr}(U, V^X) \wedge U \neq V \wedge \text{realnum_expr}(X) \wedge \\ \text{eqv_expr}(D, A^X) \wedge D \neq A \wedge \text{unit_dimension}(V, A) \end{aligned}$$

Non-physical units and dimensions

We define \$ as the unit of the non-physical fundamental dimension money.

$$\text{npu_fdim}(\$, \text{money})$$

U is the unit of a non-physical composite dimension D if U is equivalent to an expression $V \times \$$ and D is equivalent to an expression $A \times \text{money}$ where A is the dimension of the unit V .

$$\begin{aligned} \text{npu_cdim}(U, D) \leftarrow \\ \text{eqv_expr}(U, V \times \$) \wedge \\ \text{eqv_expr}(D, A \times \text{money}) \wedge \\ \text{unit_dimension}(V, A) \end{aligned}$$

Dimension identities

Similarly to units, there exist identities of dimensions as well. A few of them are specified below.

$$\begin{aligned} \text{area} &\equiv Expr \leftarrow \text{eqv_expr}(Expr, \text{length}^2) \\ \text{volume} &\equiv Expr \leftarrow \text{eqv_expr}(Expr, \text{length}^3) \\ \text{frequency} &\equiv Expr \leftarrow \text{eqv_expr}(Expr, 1/\text{time}) \end{aligned}$$

8.6 Reals and Real-Valued Expressions

Reals and real-valued expressions are part of lower level *Ecolingua* specification in the composition of dimensions and units of measure.

X is a real-valued expression if X is a real number or if X is an arithmetic expression that evaluates to a real number.

$$\begin{aligned} \text{realnum_expr}(X) \leftarrow \\ X \in \mathbb{R} \vee \\ (\text{arithm_expr}(X) \wedge \text{eval}(X, Y) \wedge Y \in \mathbb{R}) \end{aligned}$$

Likewise for positive reals:

$$\begin{aligned} \text{posrealnum_expr}(X) \leftarrow \\ X \in \mathbb{R}^+ \vee \\ (\text{arithm_expr}(X) \wedge \text{eval}(X, Y) \wedge Y \in \mathbb{R}^+) \end{aligned}$$

9 A Practical Application of Ecolingua

In ecological modelling, as in other domains, using data derived from observation to inform model design adds credibility to model simulation results. Also, a common methodological approach that facilitates understanding of complex systems is to firstly design a conceptual (or qualitative) model which is later used as a framework for specification of a quantitative model. However, data sets given to support modelling of ecological systems contain mainly quantitative data which, in its low representational level, do not directly connect to high-level model conceptualisation. As a result, conceptual models are often led astray from their supporting data sets. In this context, an ontology of properties of domain data can play the role of a conceptual vocabulary for representation of data sets, by way of which the data’s level of abstraction is raised to facilitate connections with conceptual models. Ecolingua was initially built to support an application of synthesis of conceptual system dynamics models [13] stemming from data annotated through the ontology, where existing models are reused to guide the synthesis process. The application is depicted in Fig. 2 and is briefly discussed in the sequel.

Figure 2 shows the synthesis process starting with a given modelling data set to support the design of a conceptual ecological model. Ecolingua vocabulary is then manually employed to annotate the data set yielding metadata. Here is an example to illustrate how data, ontology and metadata inter-relate. Suppose a forest logging operation where the amount of timber of each logged tree is measured. The actual measures of tree timber, e.g., 500 kg, are *data*. $\text{amt_of_mat}(A, Mt, E, U)$ is an Ecolingua term, defined as a relation between an amount A of a material Mt

in an entity E expressed in the unit of measure U . Our amount of timber quantity can be represented as $\text{amt_of_mat}(t, \text{timber}, \text{tree}, \text{kg})$, which is an instance of *metadata*.

The synthesis mechanism tries and matches the structure (or topology) of the existing model with the metadata set, whose content marks up the structure to give a new model that is consistent with the given metadata. This is done by solving constraint rules that represent modelling knowledge in the mechanism. Matching the existing model with metadata means to reduce its structure to the concepts in Ecolingua, which define properties of ecological data in terms of the fundamental notion of physical dimensions. It all comes down to how similar the two data sets – the new model’s annotated with Ecolingua, and the data set that once backed up the existing model – are with respect to ontological properties. The more properties they share, the more of the existing model’s structure will be successfully matched with the new model’s metadata.

9.1 Automatically Checking for Ecolingua-Compliant Metadata

Besides providing a vocabulary for ecological data annotation, Ecolingua is employed by the synthesis system to check compliance of the manually specified metadata with the underlying ontology axioms, ensuring that only compliant metadata are carried forward into the model synthesis process. Metadata terms are retrieved from the metadata set when predicates of the form $\text{data}(C_{\text{term}})$ (e.g., $\text{data}(\text{amt_of_mat}(\text{tree_timber}, \text{timber}, \text{tree}, \text{kg}))$) are solved as the system runs. We define that when one of such predicates is solved, its metadata term also holds, in that it exists as part of the metadata set; thus, $\text{data}(C_{\text{term}}) \rightarrow C_{\text{term}}$. As we know from Sect. 5.1, Ecolingua axioms are written as $C_{pt} \rightarrow C_{tt}$. Since in the synthesis system such axioms are only reasoned upon when a metadata term C_{term} with logical value true unifies with C_{pt} , the use of the axiom can be reduced to solving C_{tt} , as its logical value alone will correspond to the logical value of the whole expression. Therefore, each Ecolingua axiom can be represented in the synthesis system as a clause of the form $c_ctt(C_{pt}, C_{tt})$ ⁵. The following Ecolingua compliance checking mechanism is thus defined and applied.

Let C_{term} be an instance of an Ecolingua concept C_{pt} . As defined by the Ecolingua axioms

⁵An additional gain is that the formula conforms to the Horn Clause representational formalism adopted throughout the synthesis system.

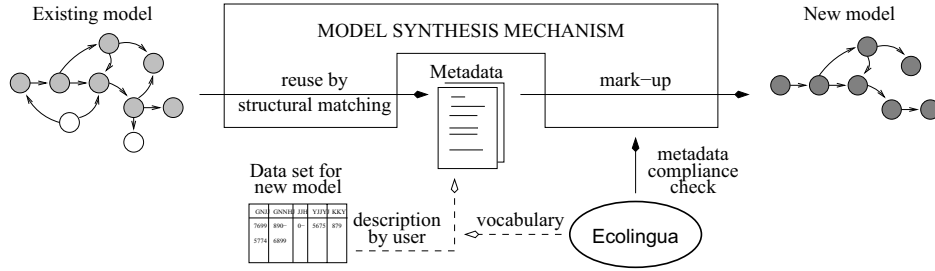


Figure 2: Application of Ecolingua in model synthesis through reuse

formula, C_{term} being true and unified with C_{pt} implies that the consequent constraint C_{tt} must be true. If however, the concept in question is one that lacks an axiomatic definition, it suffices to verify that C_{term} unifies with an Ecolingua concept:

```

onto_check( $C_{term}$ ) ←
  c_ctt( $C_{term}$ ,  $C_{tt}$ ) ∧  $C_{tt}$ 
onto_check( $C_{term}$ ) ←
  ¬ c_ctt( $C_{term}$ , ¬ $C_{tt}$ ) ∧ eco_concept( $C_{term}$ )

```

An implementation of this ontology-supported model synthesis mechanism has been empirically evaluated on the run time efficiency criterion. In the evaluation, around 600 runs of the mechanism were performed over manually Ecolingua-annotated and artificially generated metadata sets, having as benchmark twelve conceptual models found in three well-known textbooks in system dynamics ecological modelling [22, 19, 13]. This sample of models ranged from simple yet non-trivial models to reasonably complex models, amongst the largest in the textbooks. These models as well as other models were successfully synthesised. The mechanism is able to synthesise multiple model solutions for one given metadata set, which is justified in the ecological modelling domain, where, as it happens in modelling disciplines in general, one data set does not uniquely determine a model solution, as different interpretations of the data may yield different models. The evaluation and a complete description of the synthesis mechanism can be found in [4].

10 Concluding Remarks

We have defined classes of quantitative data in Ecology, using the well-known formalism of first-order logic. The definitions draw on the EngMath ontology to characterise quantity classes with respect to their physical dimension, which can be captured through

the unit of measure in which instances of the quantity classes are expressed in.

The ontology has been employed to enable a technique of synthesis of conceptual ecological models from metadata and reuse of existing models. The synthesis mechanism that implements the technique involves proofs over the ontology axioms written in Prolog in order to validate metadata that is given to substantiate the models. This is an application where an ontology is not used at a conceptual level only, as we commonly see, but at a practical, implementational level, adding value to a knowledge reuse technique.

As the ontology is founded on the universal concept of physical dimensions, its range of application can be widened. However, while the definitions presented here have been validated by an ecological modelling expert at the Institute of Ecology and Resource Management, University of Edinburgh, Ecolingua concepts and axioms are not yet fully developed. Quantities of space, energy and frequency dimensions, for example, as well as precise definitions, with axioms where possible, of contextual ecological concepts such as ecological entity, event, the compatibility relation between entities and materials, are not covered and will be created, or reused, as the ontology evolves.

Ecolingua is currently being ported into the state-of-the-art ontology languages OWL and SWRL (Semantic Web Rule Language) [23] using Protégé. Preliminary results appear in [36]. We are also considering relating such specification of Ecolingua to the OWL version of SUMO. In doing that, we aim at fostering further developments on Ecolingua as well as applications of it over the World Wide Web.

References

- [1] M. Annamalai and L. Sterling. Dealing with Mathematical Relations in Web-Ontologies. In Proceedings of the *Workshop on Ontologies in*

- Agent Systems* (OAS'03) at the 2nd International Joint Conference on Autonomous Agents and Multi-Agent Systems, Melbourne, Australia, pages 1–8, 2003.
- [2] Y. Biot, N. Higuchi, V. Brillhante, J. de Freitas, J. Ferraz, N. Leal, S. Ferreira and T. Desjardins. INFORM – the INpa FORest Model. Project BIONTE Technical Report, National Institute for Research of the Amazon (INPA) – Brazil and Overseas Development Administration (ODA) – UK, May 1996.
- [3] P. Borst, H. Akkermans and J. Top. Engineering Ontologies. *International Journal of Human-Computer Studies*, 46:365–406, 1997.
- [4] V. Brillhante. Ontology and Reuse in Model Synthesis. PhD thesis, School of Informatics, University of Edinburgh, 2003.
- [5] M. Ceruti, C. Anken, A. Lin and S. Rubin. Applications of High-Performance Knowledge-Based Technology. In Proceedings of the *IEEE Intl. Conference on Systems, Man and Cybernetics*, Nashville, USA, 2000.
- [6] V. K. Chaudhri, A. Farquhar, R. Fikes, P. D. Karp and J. P. Rice. Open Knowledge Base Connectivity 2.0. KS-98-06, Knowledge Systems Laboratory, Stanford University, California, USA, Jan 1998.
- [7] E. Davis. Representations of Commonsense Knowledge. Morgan Kaufmann Publishers, Inc., San Mateo, California, USA, 1990.
- [8] S. Decker, D. Fensel, F. van Harmelen, I. Horrocks, S. Melnik, M. C. Klein and J. Broekstra. Knowledge Representation on the Web. In Proceedings of the *2000 International Workshop on Description Logics* (DL'00), Aachen, Germany, pages 89–97, 2000.
- [9] Defense Advanced Research Projects Agency. <http://www.darpa.mil/>, Oct 2005.
- [10] B. Ellis. Basic Concepts of Measurement. Cambridge University Press, London, 1966.
- [11] A. Farquhar, R. Fikes and J. Rice. The Ontolingua Server: a Tool for Collaborative Ontology Construction. KSL-96-26, Computer Science Department, Stanford University, 1996.
- [12] K. D. Forbus. Qualitative Process Theory. *Artificial Intelligence*, 24:85–168, 1984.
- [13] A. Ford. Modeling the Environment: an Introduction to System Dynamics Modeling of Environmental Systems. Island Press, Washington, D.C., USA, 1999.
- [14] G. Frank, A. Farquhar and R. Fikes. Building a Large Knowledge Base from a Structured Source. *IEEE Intelligent Systems*, 14(1):47–54, 1999.
- [15] N. Fridman, R. Fergerson and M. Musen. The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility. In R. Dieng and O. Corby, editors, Proceedings of the *12th International Conference on Knowledge Engineering and Knowledge Management* (EKAW'00), volume 1937 of Lecture Notes in Computer Science (LNCS), Springer. Juan-les-Pins, France, pages 17–32, 2000.
- [16] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari and L. Schneider. Sweetening Ontologies with DOLCE. In Proceedings of the *13th International Conference on Knowledge Engineering and Knowledge Management* (EKAW'02), volume 2473 of Lecture Notes in Computer Science (LNCS), Springer. Siguenza, Spain, pages 166–181, 2002.
- [17] M. Genesereth and R. Fikes. Knowledge Interchange Format, Version 3.0, Reference Manual. Logic-92-1, Logic Group, Computer Science Department, Stanford University, California, USA, 1992.
- [18] A. Gómez-Pérez and M. D. Rojas-Amaya. Ontological Reengineering for Reuse. In Proceedings of the *11th European Knowledge Acquisition Workshop* (EKAW'99), volume 1621 of Lecture Notes in Artificial Intelligence (LNAI), Springer. Dagstuhl Castle, Germany, pages 139–156, 1999.
- [19] W. E. Grant, E. K. Pedersen and S. L. Marín. Ecology and Natural Resource Management: Systems Analysis and Simulation. John Wiley & Sons, Inc., 1997.
- [20] T. Gruber. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.
- [21] T. Gruber and G. Olsen. An Ontology for Engineering Mathematics. In Proceedings of the *4th International Conference on Principles of Knowledge Representation and Reasoning*, Bonn, Germany, 1994.

- [22] J. Haefner. *Modeling Biological Systems: Principles and Applications*. Chapman & Hall, 1996.
- [23] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosz and M. Dean. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission, May 2004. <http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/>, Oct 2005.
- [24] High Performance Knowledge Bases. <http://reliant.tekknowledge.com/HPKB/>, Oct 2005.
- [25] IEEE Standard Upper Ontology Working Group. <http://suo.ieee.org/>, Oct 2005.
- [26] V. Kashyap. Design and Creation of Ontologies for Environmental Information Retrieval. In *Proceedings of the 12th International Conference on Knowledge Acquisition, Modeling and Management*, Banff, Canada, 1999.
- [27] C. M. Keet. Factors Affecting Ontology Development in Ecology. In *Proceedings of the 2nd International Workshop on Data Integration in the Life Sciences (DILS'05)*, volume 3615 of *Lecture Notes in Computer Science (LNCS)*, Springer-Verlag, San Diego, USA, pages 46–61, 2005.
- [28] D. B. Lenat and R. V. Guha. *Building Large Knowledge Based Systems: Representation and Inference in the Cyc Project*. Addison Wesley, Reading, Massachusetts, 1990.
- [29] C. Masolo, S. Borgo, A. Gangemi, N. Guarino and A. Oltramari. *Ontology Library (final)*. WonderWeb Deliverable D18, Laboratory for Applied Ontology – ISTC – CNR, Dec 2003.
- [30] B. Massey. *Measures in Science and Engineering: their Expression, Relation and Interpretation*. Ellis Horwood Limited, 1986.
- [31] I. Niles and A. Pease. Origins of the Standard Upper Merged Ontology: a Proposal for the IEEE Standard Upper Ontology. In *Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology* Seattle, Washington, USA, pages 37-42, 2001.
- [32] I. Niles and A. Pease. Towards a Standard Upper Ontology. In C. Welty and B. Smith, editors, *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS'01)*, Ogunquit, Maine, USA, 2001.
- [33] B. N. Niven. Formalization of the Basic Concepts of Animal Ecology. *Erkenntnis*, 17:307–320, 1982.
- [34] B. N. Niven. Formalization of Some Concepts of Plant Ecology. *Coenoses*, 7(2):103–113, 1992.
- [35] Open Biomedical Ontologies. <http://obo.sourceforge.net>, Oct 2005.
- [36] R. C. de Oliveira, J. S. Pereira and V. Brilhante. Reengineering a FOL Ontology into OWL and SWRL using Protégé. In *Proceedings of the Workshop on Protégé with Rules at the 8th Intl. Protégé Conference*, Madrid, Spain, 2005.
- [37] Ontolingua Server. <http://ontolingua.stanford.edu>, Knowledge Systems Laboratory, Department of Computer Science, Stanford University, Oct 2005.
- [38] OpenCyc.org. Formalized Common Knowledge. <http://www.opencyc.org>, Oct 2005.
- [39] S. Pinto. Revising and Extending the Units of Measure “Subontology”. In *Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology*, Seattle, Washington, USA, 2001.
- [40] S. Pinto. Towards Ontology Reuse. In *Proceedings of the AAAI-99 Workshop on Ontology Management*, WS-99-13, AAAI Press, Orlando, Florida, USA, pages 67–73, 1999.
- [41] The Protégé Ontology Editor and Knowledge Acquisition System. <http://protege.stanford.edu/>, Oct 2005.
- [42] S. Reed and D. Lenat. Mapping Ontologies into Cyc. In *Proceedings of the Workshop on Ontologies and the Semantic Web at the 18th National Conference on Artificial Intelligence (AAAI'02)*, Edmonton, Alberta, Canada, 2002.
- [43] Science Environment for Ecological Knowledge (SEEK), OWL Ontologies. <http://cvs.ecoinformatics.org/>, Oct 2005.
- [44] F. C. da Silva, W. Vasconcelos, D. S. Robertson, V. Brilhante, A. C. de Melo, M. Finger and J. Agustí. On the Insufficiency of Ontologies: Problems in Knowledge Sharing and Alternative Solutions. *Knowledge-Based Systems*, 15(3):147–167, 2002.
- [45] J. F. Sowa. *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Brooks/Cole, 2000.

- [46] Suggested Upper Merged Ontology.
<http://www.ontologyportal.org/>, Oct 2005.
- [47] Le Système International d'Unités (SI), The International System of Units (SI). Bur. Intl. Poids et Mesures, Sévres, France, 6th edition, 1991.
- [48] M. Uschold, M. Healy, K. Williamson, P. Clark and S. Woods. Ontology Reuse and Application. In Proceedings of the *1st International Conference on Formal Ontology in Information Systems* (FOIS'98), IOS Press. Trento, Italy, pages 179–192, 1998.
- [49] Web Ontology Language, W3C.
<http://www.w3.org/2004/OWL/>, Oct 2005.
- [50] WonderWeb: Ontology Infrastructure for the Semantic Web. <http://wonderweb.semanticweb.org>, Oct 2005.