Technical Report LPT-2008-27

Material

J. Morbach, A. Yang, W. Marquardt

July 2008

Table of Contents

List of Figures

1. Partial Model 'Material'

The partial model **material** provides an *abstract description of matter*. In this context, "matter" refers to "anything that has mass and occupies space" (Gold et al., 1982). As for the 'abstract description', the partial model does not consider all the characteristics of matter, but accounts only for those that are independent of the *shape*, *size*, or *amount* of a particular occurrence of matter.

The partial model **material** originates from the CLiP (cf. Bayer, 2003) partial model '**chemical_process_material**', which is described in an article by Yang et al. (2003); several passages of this article have been included in the present documentation (mostly in paraphrased form). In comparison to CLiP, the partial model **material** has a somewhat different structure[1], and it is modeled partly differently in order to correct certain flaws and logical contradictions of the CLiP model. In addition, **material** incorporates concepts of the ChEBI ontology (EBI, 2007; OLS, 2006) to describe matter at the molecular level, which was not enabled by the CLiP model.
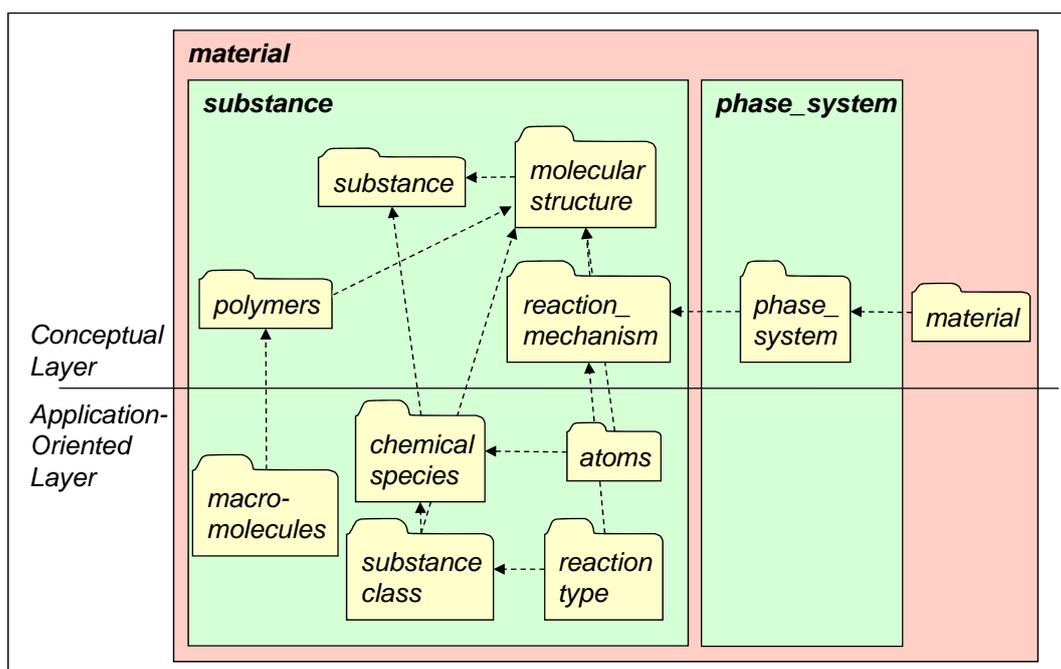
Fig. 1: Structure of the partial model **material**

Besides the ontology module *material*, **material** contains two further partial models, called **substance** and **phase_system** (cf. Sect. 3). While **substance** comprises several modules, **phase_system** consists of a single module.

---

[1] In particular, the CLiP partial model '**mathematical model of phase system**', which forms part of the CLiP material model, has been relocated to the OntoCAPE partial model **mathematical_model** (cf. Morbach et al. 2008b).
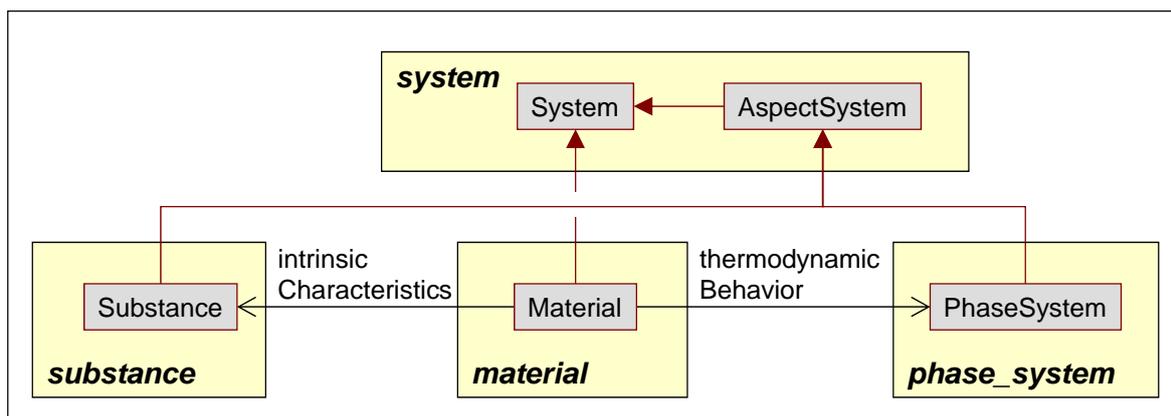
Fig. 2: Major classes of the partial model **material**

Material is the main class of the ontology module *material*. As shown in Fig. 2, *material* is a special type of *system*. A *material* can be viewed from two different perspectives, which are represented as *aspect systems* of *material*:

- *Substance* represents the intrinsic characteristics (i.e., the physicochemical nature) of a *material*.

- *Phase system* describes the macroscopic thermodynamic behavior of a *material*.

The critical concept for the *substance* and *phase system* is the *physical context* (which is introduced in ontology module *phase_system*). This class is defined as a set of independent *properties*[2] with known *values*, which is sufficient to determine other *properties* of interest; temperature, pressure, and molar concentrations are examples of *properties* that constitute a *physical context*. *Material* without consideration of the *physical context* is modeled as *substance*. The *properties* of *substance* are constants, such as molecular weight or critical properties; their *values* cannot be altered by mechanical, thermo-physical or chemical processes. In contrast, *material* within a certain *physical context* is modeled as *phase system*. The *values* of the *phase system properties* are subject to the respective *physical context*. *Phase system properties* include the thermodynamic state variables and other quantities derived thereof.

The two *aspect systems* are represented in separate partial models (i.e., **substance** and **phase_system**), which subdivide the partial model **material**. This partition facilitates the usage of **material** in different types of applications: The partial model **substance** can be used in applications where merely the intrinsic characteristics of *materials* are of interest. Correspondingly, applications that are interested in the thermodynamic behavior of *materials* will use the concepts of the partial model **phase_system**.

Similarly to materials, the concept of a chemical reaction can be divided into a context-dependent and context-independent part. Reaction mechanism and reaction stoichiometry are context-independent and thus form part of the partial model **substance**. The context-dependent reaction property, namely reaction equilibrium constant, is introduced in the partial model **phase_system**.

---

[2] often called 'state variables' in the thermodynamics literature

1.1. Relation Between 'Phase System' and 'Material Amount'

A *phase system* is an *abstraction of a concrete occurrence of matter*. By 'concrete occurrence', we mean the actual spatiotemporal setting, for example, the manufacturing of some material in a chemical plant or its usage as a construction material. The concept of a *material amount*, introduced in the partial model **CPS_behavior** (cf. Wiesner et al., 2008), is complementary to the *phase system* concept: A *material amount* represents the concrete occurrence of matter in, for example, a chemical plant. To better understand the distinction between *phase system* and *material amount*, the idea of an 'abstraction of a concrete occurrence of matter' is further explained in the following.

The "physical prototype" of a *phase system* is an *arbitrary amount* of (static or flowing) material in an equilibrium state and with a rather *simple geometry*. A *phase system* may take the form of either a *single phase* or a *multiphase system* where the constituting *single phases* are connected through *phase interfaces* with the most plain geometry (i.e., planar surfaces). Thus, the abstraction excludes all material characteristics that depend on the shape, size, or amount of a particular occurrence of a material. Consequently, *extensive properties* as well as the distribution of *intensive properties* in space and the distribution of any *property* in time are not associated with a *phase system*. Rather, they are modeled as *properties* of *material amount*. Furthermore, a *phase system* is meant to represent only the physical properties which are applicable to material in an equilibrium state. Consequently, any properties that describe the rates of chemical reaction or transport phenomena or the spatial gradients of physical quantities involved in non-equilibrium events, are associated not with *phase system* but rather with *material amount*. For *single phases*, in particular, only the following *properties* are considered: (i) intensive thermodynamic states, and (ii) *properties* that can be determined solely from (i). For *multiphase systems*, only those *properties* are taken into account that are (i) solely dependent on the *properties* of the constituting *single phases* and (ii) not associated to a particular surface geometry other than the most plain one as mentioned above.

The distinction between a *phase system* and a *material amount* has been made not only because it is conceptually feasible, but also because it is practically useful. It was inspired by the physical properties packages that are commonly used in the field of process modeling. These packages handle the computation of such material properties without reference to amounts and geometry. Due to their independence of amount and geometry, the packages have proven to be highly reusable in modeling various kinds of chemical processes. Correspondingly, it is expected that the concepts of the partial model **material** prove reusable for supporting the modeling of different *material amounts*.

1.2. Material (Ontology Module)

The only purpose of the ontology module *material* is to establish the relations between the concepts in the partial model **substance** (representing the intrinsic aspects of matter) and those in partial model **phase_system** (describing the thermodynamic behavior of matter). Thus, the module contains only very few concepts, which are listed below.

3

Concept Descriptions

Individual concepts of the module *material* are defined below.

Class Descriptions

Classes

Material

Description

The class *material* represents all kinds of matter.

Relations

- *Material* is a subclass of *system*.

- The intrinsic characteristics of a *material* are represented through *substance*.

- The thermodynamic behavior of a *material* is represented through *phase system*.

Relations

intrinsicCharacteristics

Description

The relation designates the *aspect system* which represents the intrinsic characteristics of some *material*.

Characteristics

- Specialization of hasAspectSystem

- Domain: *Material*

- Range: *AspectSystem*

- Inverse: representsIntrinsicCharacteristicsOf

representsIntrinsicCharacteristicsOf

Description

The relation links a *substance* to a *material*.

Characteristics

- Specialization of representsAspectSystemOf

- Domain: *AspectSystem*

- Range: *Material*

- Inverse: intrinsicCharacteristics

representsThermodynamicBehaviorOf

Description

The relation links a *phase system* to a *material*.

Characteristics

- Specialization of representsAspectSystemOf

- Domain: *AspectSystem*

- Range: *Material*

- Inverse: thermodynamicBehavior

thermodynamicBehavior

Description

The relation designates the *aspect system* which represents the thermodynamic behavior of some *material*.

Characteristics

- Specialization of hasAspectSystem

- Domain: *Material*

- Range: *AspectSystem*

- Inverse: representsThermodynamicBehaviorOf

Individuals

intrinsic_characteristics

Description

Explicitly designates an *aspect system* to represent the intrinsic characteristics of some *material*.

Characteristics

- Instance of *aspect*

thermodynamic_behavior

Description

Explicitly designates an *aspect system* to represent the thermodynamic behavior of some *material*.

Characteristics

- Instance of *aspect*

2. Substance (Partial Model)

The partial model **substance** comprises 4 ontology modules on the Conceptual Layer, and (currently) four additional modules on the Application-Oriented Layer (cf. Fig. 3). The following modules are located on the Conceptual Layer:

- *Substance* is the main ontology module of the **substance** partial model. It provides essential concepts for the description of pure substances and mixtures, primarily at the macroscopic scale.

- The ontology module molecular_structure is concerned with the characterization of pure substances at the atomic scale.

- *Polymers* supplements *molecular_structure* by concepts for the description of macromolecular structures.

- Finally, *reaction_mechanism* allows to represent the mechanism and the stoichiometry of chemical reactions.
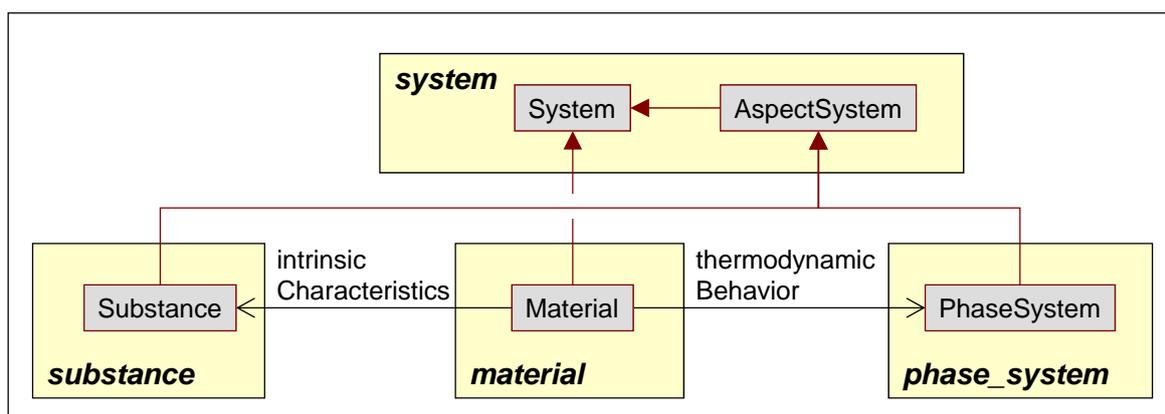
Fig. 3: Structure of the partial model **substance**

The Application-Oriented Layer comprises the following ontology modules:

- The module *chemical_species* instantiates concepts from the *substance* module to establish an information base about pure substances.

- *Atoms* refines *molecular_structure* by instance data about the chemical elements.

- Similarly, *macromolecules* refines *polymers* by instance data about the molecular structure of technical polymers.

- *Substance_class* categorizes chemical substances into classes with similar chemical properties, such as alcohols, esters, etc.

- *Reaction_type* describes important types of chemical reactions, like esterification or hydrohalogenation.

2.1. Substance (Ontology Module)

The ontology module *substance* originates from the CLiP model 'substance' (Yang et al. 2003). The major concepts of the *substance* module are shown in Fig. 4.
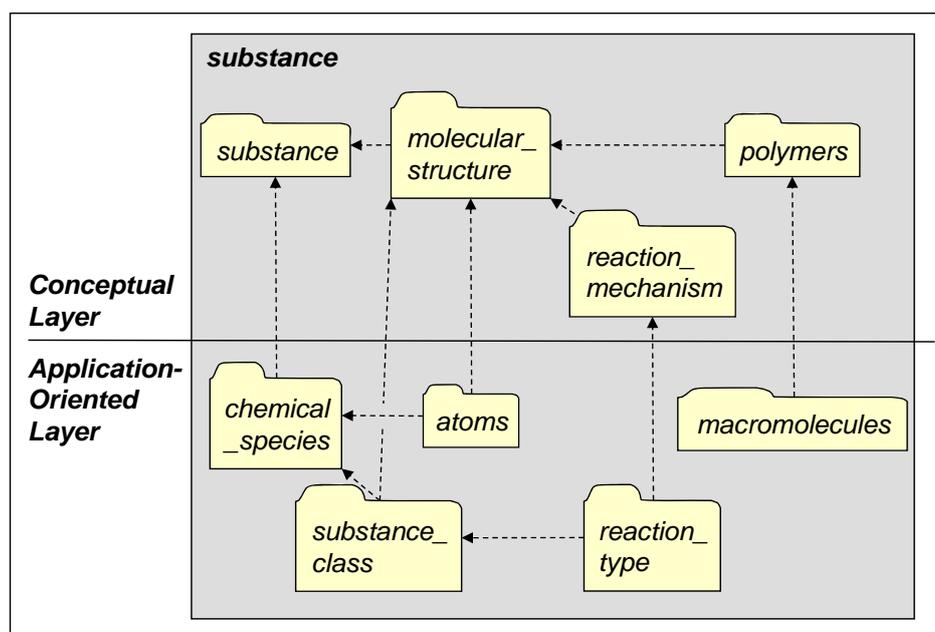
Fig. 4: Major concepts of *substance*

The key concept, *substance*, represents "a generalization of matter at or above the atomic level" (Yang et al., 2003). As mentioned earlier, *substance* is an *aspect system* of *material* and reflects its intrinsic, context-independent characteristics.

Three subclasses of *substance* are introduced: *Mixture* and *chemical component* both describe *substances* at the macroscopic scale, while *molecular entity* characterizes *substances* at the atomic scale.

The class *chemical component* subsumes the classes *chemical species* and *pseudo component*.

- A *chemical species* represents pure *substances* at the macroscopic scale. The IUPAC Compendium defines *chemical species* as an "ensemble of chemically identical *molecular entities* […]. The term is applied equally to a set of chemically identical atomic or molecular structural units in a solid array. […] The term is taken to refer to a set of *molecular entities* containing isotopes in their natural abundance. […] The wording of the definition […] is intended to embrace both cases such as graphite, sodium chloride or a surface oxide, where the basic structural units may not be capable of isolated existence, as well as those cases where they are." (McNaught and Wilkinson, 1997).

7

- The class *pseudo components* is introduced alongside *chemical species*. A *pseudo component* is an auxiliary concept, which represents the averaged properties of a number of *chemical species*. *Pseudo components* are often assumed to exist in the context of physical property calculations of complex multicomponent mixtures, such as petroleum (Hariu and Sage, 1969), fatty alcohols (Gutsche, 1986), or polymers (Kuma and Gupta, 1998).

Unlike a *pseudo component*, a *chemical species* has a specific chemical composition, which can be indicated by means of a chemicalFormula or by referring to the corresponding *molecular entity*.
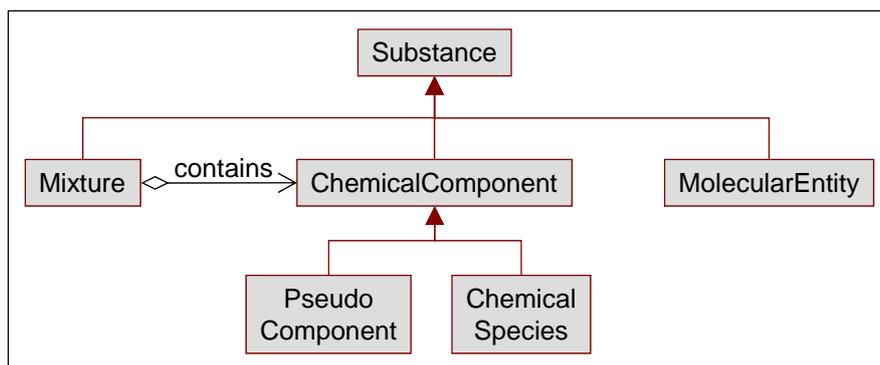
Fig. 5: Representation of pure substances at the macroscopic and the atomic scale

According to the IUPAC Compendium, a *molecular entity* denotes "any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer etc., identifiable as a separately distinguishable entity. Molecular entity is used [...] as a general term for singular entities, irrespective of their nature, while chemical species stands for sets or ensembles of molecular entities" (McNaught and Wilkinson, 1997). Thus, both *molecular entity* and *chemical species* describe pure substances – the former at the molecular scale, the latter at the macroscopic scale. The inverse relations hasMolecularStructure and hasMacroscopicAppearance interrelate a *chemical species* and its corresponding *molecular entity*; they may be used to navigate between the macroscopic and molecular perspectives (cf. Fig. 5).

The concept of *molecular entities* introduced in this module is further elaborated in the ontology module *molecular_structure*. This separation of macroscopic perspective (module *substance*) and molecular perspective (module *molecular_structure*) allows to characterize *substances* at the desired level of detail: For many applications in chemical engineering, the molecular structure of substances is not of interest and may therefore be omitted from an ontological description; if, on the other hand, the molecular properties are relevant to the application, they can be obtained easily by adding the ontology module *molecular_structure*.

A *mixture* is generally a *substance* that contains two or more *chemical components*. The *mixture* concept can represent two different things: a loose collection of segregate *chemical components*, or a compound material formed by several blended *chemical components*. For the latter case, subclasses of *mixture* can be introduced to denote and classify typical mixtures of *chemical components*, such as alloys, polymer blends, types of crude oil, chocolate, sand, saltwater, or air. The composition of a *mixture* is not fixed,

as opposed to the composition of a *phase system* (cf. Sect. 3). Note that a *mixture* and its constituting *chemical species* are not connected via the hasSubsystem relation, but via the better scaling contains relation (cf. Morbach et al., 2008a). This measure is taken since *mixture* and *chemical species* typically have thousands of instances – a data set of this size would cause performance problems during reasoning if its members were connected via hasSubsystem.
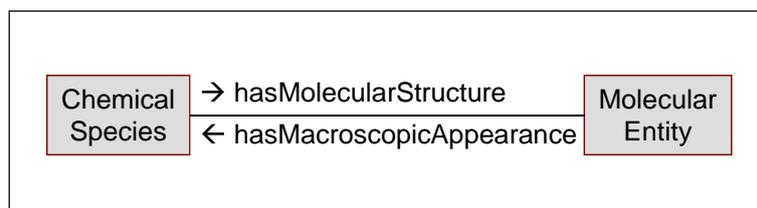
Fig. 6: Some *chemical component constants*

Since substance represents the intrinsic, context-independent characteristics of materials, the properties of substances are constants. Some exemplary chemical component constants are shown in Fig. 6.
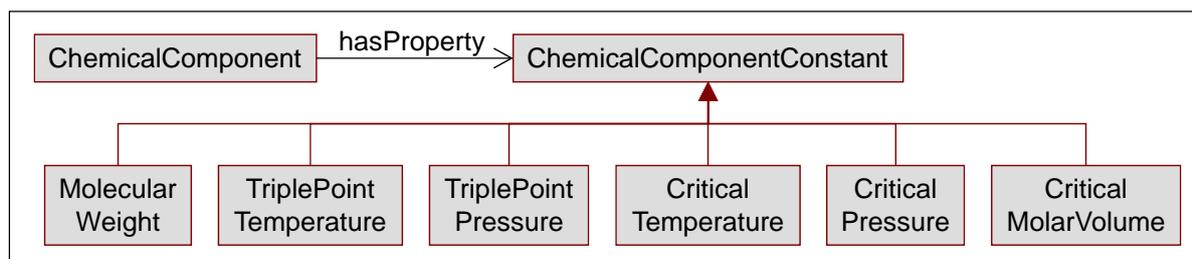
Fig. 7: Substance identifiers

Each *substance* has at least one substanceID, such as a (trivial or systematic) name (cf. Fig. 7). A substanceID[3] is an unambiguous, but not necessarily a unique identifier (i.e., each substanceID represents exactly one *substance*, but a *substance* can have more than one substanceID). Unique identifiers can be explicitly categorized as uniqueSubstanceIDs. The CAS_RegistryNumber is probably the most well-known uniqueSubstanceID.

---

[3] In OWL, the substanceIDs are modeled as inverse-functional attributes; uniqueSubstanceIDs are additionally characterized as functional.
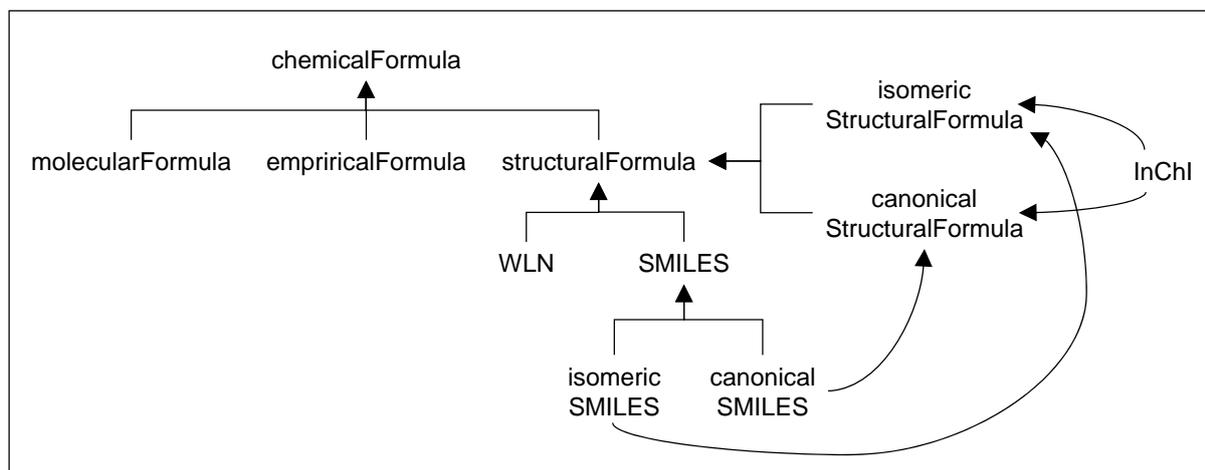
Fig. 8: Chemical formulas

A chemicalFormula provides information about the atoms that constitute a particular *molecular entity*. It can only be assigned to *molecular entities* and *chemical species*[4]. Different types of chemicalFormulas exist:

- An empiricalFormula indicates the relative number of each constituting chemical elements of a *molecular entity*.

- A molecularFormula specifies the (absolute) number of constituting atoms of a *molecular entity*, without indicating how they are linked.

- A structuralFormula supplies information about the types of bonds and the structural arrangement of the atoms of a *molecular entity* using a linear string notation. Different formats for the representation of structural formals exist – examples are WLN (Wisswesser Line Notation), SMILES (Simplified Molecular Input Line Entry Specification), and InChI (IUPAC International Chemical Identifier).

A structuralFormula is an unambiguous but not unique (Dietz, 1995) representations of a *molecular entity*. If the string representations of the structuralFormulas are generated by means of appropriate canonicalization algorithms, the representation will become unique; the canonicalSMILES notation is an example of such a canonicalStructuralFormula.

Not every structuralFormula is able to distinguish the different isomers of a *molecular entity*. Only a isomericStructuralFormula, like InChI or isomericSMILES, enables such a distinction.

Note that chemicalFormulas cannot be considered as substanceIDs since they are ambiguous (e.g., C_3H_6

---

[4] Conceptually, a chemicalFormula should be solely assigned to *molecular entity*. For practical usage, however, it is advantageous to assign a chemicalFormula directly to a *chemical species*, as many chemical engineering applications ignore the molecular perspective represented by *molecular entity*.

could refer to Propene as well as to Cyclopropane). Even though a structuralFormula can unambiguously identify a *molecular entity*, it is ambiguous for *chemical species* (see example in Fig. 8) and thus not a substanceID. Similarly, a canonicalStructuralFormula is a unique identifier for a *molecular entity*, but not for a *chemical species*, and can therefore not be considered as a uniqueSubstanceID.

Usage

Instances of *chemical species* and *molecular entity* should preferably have a uniqueSubstanceID to be easily identifiable in a software application. The indication of a chemicalFormula is optional for *chemical species* but mandatory for *molecular entities*.
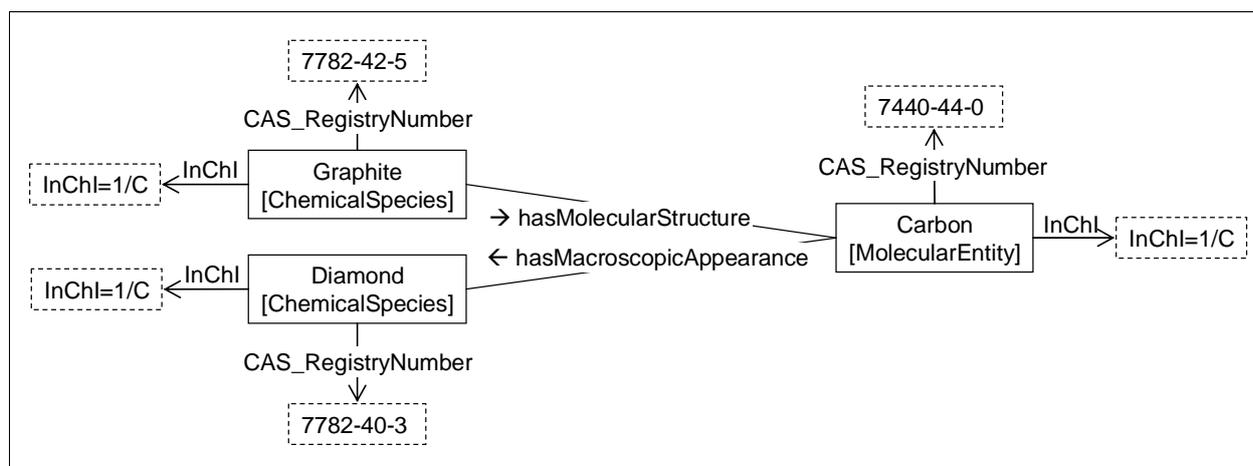
Fig. 9: Two different macroscopic appearances of a *molecular entity*

Note that a *molecular entity* can have different macroscopic appearances – **Carbon**, for instance, can macroscopically take the form of **Graphite** or that of a **Diamond** (cf. Fig. 9). In such a case, the individuals should have different substanceIDs (here: CAS_RegistryNumbers) to be distinguishable from each other. Generally speaking, the different polymorphic forms of a substance can be distinguished by different instances of substance.

More often, however, there is a one-to-one correspondence between a *chemical species* and the corresponding *molecular entity*. In this case, the same substanceID and the same instance identifier may be used for the *chemical species* and the corresponding *molecular entity* (however, the instance identifiers must have different namespace prefixes – cf. Fig. 18 in Sect. 2.5). This is acceptable since, in practice, the distinction between *molecular entities* and *chemical species* is often ignored. Thus, an identifier may refer to a *molecular entity* as well as to the corresponding *chemical species*.
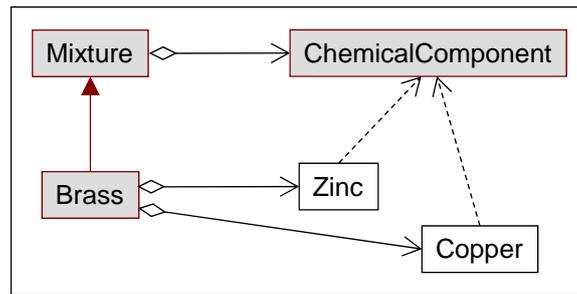
Fig. 10: Definition of *brass* as a subclass of *mixture*

Collections of segregate *chemical species* are represented as direct instances of the *mixture* class. By contrast, compound materials and typical mixtures are modeled as subclasses of *mixture*. The constituent *chemical species* can be indicated as part of the class definition, but this is not mandatory. Fig. 10 shows exemplarily the definition of *brass*; instances of *brass* would indicate the concrete occurrence of a brass material.

Concept Descriptions

Individual concepts of the module *substance* are defined below.

Classes

Chemical component

Description

The class *chemical component* subsumes *chemical species* and *pseudo components*.

Definition

A *chemical component* is either a *pseudo component* or a *chemical species*.

Relations

- *Chemical component* is a subclass of *substance*.

- All properties of a *chemical component* are *chemical component constants*.

Chemical component constant

Description

A *chemical component constant* is a constant *property* of a *chemical component*.

Chemical species

Description

A *chemical species* represents pure *substances* at the macroscopic scale. It consists of an "ensemble of chemically identical molecular entities […]. The term is applied equally to a set of chemically

12

identical atomic or molecular structural units in a solid array. […] The term is taken to refer to a set of molecular entities containing isotopes in their natural abundance. […] The wording of the definition […] is intended to embrace both cases such as graphite, sodium chloride or a surface oxide, where the basic structural units may not be capable of isolated existence, as well as those cases where they are." (McNaught and Wilkinson, 1997).

Relations

- *Chemical species* is a subclass of *chemical component*.

- The molecular structure of a *chemical species* can be represented by a *molecular entity*.

- A *chemical species* cannot have more than one empiricalFormula.

- A *chemical species* cannot have more than one molecularFormula.

- A *chemical species* cannot have more than one InChI.

- A *chemical species* cannot have more than one WLN.

- A *chemical species* cannot have more than one SMILES.

Critical molar volume

Description

The *critical molar volume* is the volume of one mole of a *chemical component* at the *critical temperature* and *critical pressure*.

Relations

- *Critical molar volume* is a subclass of *chemical component constant*.

Critical pressure

Description

The minimum pressure which would suffice to liquefy a substance at its *critical temperature*. Above the *critical pressure*, increasing the temperature will not cause a fluid to vaporize to give a two-phase system (McNaught and Wilkinson, 1997).

Relations

- *Critical pressure* is a subclass of *chemical component constant*.

Critical temperature

Description

The temperature, characteristic of each gas, above which it is not possible to liquefy a given gas (McNaught and Wilkinson, 1997).

Relations

- *Critical temperature* is a subclass of *chemical component constant*.

Mixture

<u>Description</u>

A *mixture* is a *substance* that contains two or more *chemical components*. The composition of a *mixture* is not fixed, as opposed to the composition of a *phase system*.

<u>Relations</u>

- *Mixture* is a subclass of *substance*.

- A *mixture* directlyContains only *chemical components*.

- A *mixture* cannot have a chemicalFormula.

<u>Usage:</u>

The *mixture* concept can represent two different things: (a) a loose collection of segregate *chemical components*, or (b) a compound material formed by several blended *chemical components*. As for (b), subclasses of *mixture* can be introduced to denote and classify typical mixtures of *chemical components*, such as alloys, polymer blends, or types of crude oil.

In case (a), the indication of the constituting *chemical components* of a *mixture* is mandatory

Note that the *molecular weight* of *mixtures* has to be described in the *phase_system* partial model, as it is context-dependent.

Molecular entity

The class *molecular entity* characterizes *substances* at the atomic scale. It represents "any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer etc., identifiable as a separately distinguishable entity. Molecular entity is used […] as a general term for singular entities, irrespective of their nature" (McNaught and Wilkinson, 1997).

<u>Relations</u>

- *Molecular entity* is a subclass of *substance*.

- The macroscopic appearance of a *molecular entity* is represented by a *chemical species*.

- A *molecular entity* has at least one chemicalFormula.

- A *molecular entity* cannot have more than one empiricalFormula.

- A *molecular entity* cannot have more than one molecularFormula.

- A *molecular entity* cannot have more than one InChI.

- A *molecular entity* cannot have more than one WLN.

- A *molecular entity* cannot have more than one SMILES.

Molecular weight

<u>Description</u>

The *molecular weight* of a *chemical component* is the ratio of the mass of one *molecule* of that substance, relative to the unified atomic mass unit (equal to 1/12 the mass of one atom of carbon-12). It is also known as (relative) molar mass or (relative) molecular mass.

If a *chemical component* consists of different *molecules* (as it is the case for a *polymer*), it represents the averaged *molecular weight* of this component.

<u>Relations</u>

- *Molecular weight* is a subclass of *chemical component constant*.

Pseudo component

<u>Description</u>

A *pseudo component* is an auxiliary concept, that defines a virtual chemical species with the averaged properties of a number of *chemical species*. *Pseudo components* are often assumed to exist in the context of physical property calculations of complex multicomponent mixtures, such as petroleum (Hariu and Sage, 1969), fatty alcohols (Gutsche, 1986), or polymers (Kuma and Gupta, 1998).

<u>Relations</u>

- *Pseudo component* is a subclass of *chemical component*.

- A *pseudo component* cannot have a chemicalFormula.

- A *pseudo component* cannot have a isomericSubstanceID.

Substance

<u>Description</u>

Substance represents matter at or above the atomic level. It reflects the intrinsic, context-independent characteristics of a *material.*

<u>Relations</u>

- *Substance* is a subclass of *system*.

- A *substance* must have at least one substanceID.

- A *substance* cannot have more than one CAS_RegistryNumber.

Triple point pressure

<u>Description</u>

The triple point of a *chemical component* is given by the temperature and pressure at which three phases (gas, liquid, and solid) of that *substance* coexist in thermodynamic equilibrium.

Relations

- *Triple point pressure* is a subclass of *chemical component constant*.

Triple point temperature

Description

The triple point of a *chemical component* is given by the temperature and pressure at which three phases (gas, liquid, and solid) of that *substance* coexist in thermodynamic equilibrium.

Relations

- *Triple point temperature* is a subclass of *chemical component constant*.

Relations

hasMacroscopicAppearance

Description

The relation hasMacroscopicAppearance points from a *molecular entity* (describing a pure substance from a molecular perspective) to the corresponding *chemical species* (describing the same substance from a macroscopic perspective).

Characteristics

- Specialization of isDirectlyRelatedTo
- Domain: *Molecular entity*
- Range: *Chemical species*
- Inverse: hasMolecularStructure
- Inverse functional

hasMolecularStructure

Description

The relation hasMolecularStructure points from a *chemical species* (describing a substance from a macroscopic perspective) to the corresponding *molecular entity* (describing the same pure substance from a molecular perspective).

Characteristics

- Specialization of isDirectlyRelatedTo
- Domain: *Chemical species*
- Range: *Molecular entity*

- Inverse: hasMacroscopicAppearance

- Functional

Attributes

CAS_RegistryNumber

<u>Description</u>

A CAS_RegistryNumber is a uniqueSubstanceID issued by the Chemical Abstracts Service (CAS) a division of the American Chemical Society. A CAS_RegistryNumber can be assigned to *chemical species* and *molecular entities* as well as to some *mixtures*. A CAS_RegistryNumber includes up to 9 digits, which are separated into 3 groups by hyphens (xxxxxx-xx-x). The first part of the number, starting from the left, has up to 6 digits; the second part has 2 digits. The final part consists of a single check digit or checksum that makes it easy to determine whether a CAS number is valid or not. See CAS (2007) for details.

<u>Characteristics</u>

- Specialization of uniqueSubstanceID

- Domain: *Substance*

- Datatype: string (built-in XML Schema datatype)

- Functional

- Inverse functional

canonicalStructuralFormula

<u>Description</u>

A canonicalStructuralFomula is a structuralFormula that is generated by means of canonicalization algorithms to obtain a unique representation of a *molecular entity*.

<u>Characteristics</u>

- Specialization of structuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

canonicalSMILES

<u>Description</u>

canonicalSMILES is the version of the SMILES specification that applies canonicalization rules to ensure

that each *chemical species* and/or *molecular entity* has a single, unique SMILES representation.

Characteristics

- Specialization of SMILES

- Specialization of canonicalStructuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

chemicalFormula

Description

A chemicalFormula is a substanceID that can only be assigned to *chemical species* and/or *molecular entities*. It gives information about the atoms that constitute a particular *molecular entity*. The attribute chemicalFormula subsumes all types of formulas, such as empiricalFormula, molecularFormula, structuralFormula, etc.

Characteristics

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

empiricalFormula

Description

An empiricalFormula is a chemicalFormula that indicates the relative number of each constituting chemical element of a *molecular entity*.

In an empiricalFormula, the letters representing the chemical elements are listed according to the following convention: In organic compounds, C is always cited first, H second and then the rest, in alphabetical order. In non-carbon-containing compounds, strict alphabetical order is adhered to.

Characteristics

- Specialization of chemicalFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

InChI

<u>Description</u>

The IUPAC International Chemical Identifier (InChI) is a non-proprietary identifier for chemical substances that can be used in printed and electronic data sources thus enabling easier linking of diverse data and information compilations (Stein et al. 2003). InChI does not require the establishment of a registry system. Unlike the CAS Registry System, it does not depend on the existence of a database of unique substance records to establish the next number for any new *molecular entity* being assigned an InChI. It uses a set of IUPAC structure conventions, and rules for normalization and canonicalization of the structure representation to establish the unique label for a *molecular entity*.

<u>Characteristics</u>

- Specialization of isomericStructuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

isomericSMILES

<u>Description</u>

isomericSMILES is the version of the SMILES specification that includes extensions to support the specification of isotopes, chirality, and configuration about double bonds.

<u>Characteristics</u>

- Specialization of SMILES

- Specialization of isomericStructuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

isomericStructuralFormula

<u>Description</u>

An isomericStructuralFormula is a structuralFormula that allows to distinguish the different isomers of a *molecular entity*.

<u>Characteristics</u>

- Specialization of structuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: any (built-in XML Schema Datatype)

- Functional

molecularFormula

<u>Description</u>

A molecularFormula is a chemicalFormula that specifies the (absolute) number of constituting atoms of a *molecular entity*, without indicating how they are linked.

In a molecularFormula, the letters representing the chemical elements are listed according to the following convention: In organic compounds, C is always cited first, H second and then the rest, in alphabetical order. In non-carbon-containing compounds, strict alphabetical order is adhered to.

For polymers and other macromelocules, parentheses are placed around the repeating unit. For example, a hydrocarbon molecule that is described as CH3(CH2)50CH3, is a molecule with 50 repeating CH2 units. If the number of repeating units is unknown or variable, the letter n may be used to indicate this (e.g. CH3(CH2)nCH3).

For ions, the charge on a particular atom may be denoted with a right-hand "+" or "-", e.g., "Na+" or "Cu,2+". The total charge on a charged molecule or a polyatomic ion may also be shown in this way, e.g., "H3O+" or "SO4,2-".

<u>Characteristics</u>

- Specialization of chemicalFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

Name

<u>Description</u>

The name attribute holds the various names of a *substance*. Both trivial and systematic names can be indicated here.

<u>Characteristics</u>

- Specialization of substanceID

- Domain: *Substance*

- Datatype: string (built-in XML Schema Datatype)

- Inverse functional

SMILES

Description

SMILES (Simplified Molecular Input Line Entry System) is a line notation for unambiguously describing the structure of chemical molecules using ASCII strings (Weininger, 1988).

Characteristics

- Specialization of structuralFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

structuralFormula

Description

A structuralFormula is a chemicalFormula that supplies information about the types of bonds and the spatial arrangement of the atoms of a *molecular entity* using a linear string notation.

Characteristics

- Specialization of chemicalFormula

- Domain: *Chemical species* or *molecular entity*

- Datatype: string (built-in XML Schema Datatype)

- Functional

substanceID

Description

A substanceID is an identifier for a *substance*. The substanceID must be unambiguous but not necessarily unique.

Characteristics

- Domain: *Substance*

- Datatype: any (built-in XML Schema Datatype)

- Inverse functional

uniqueSubstanceID

Description

A uniqueSubstanceID is a unique identifier for a *substance*. The different isomers of a *substance* are not necessarily distinguished by a uniqueSubstanceID.

<u>Characteristics</u>

- Specialization of substanceID

- Domain: *Substance*

- Datatype: any (built-in XML Schema Datatype)

- Functional

- Inverse functional

<u>Usage:</u>

It is up to the user to decide, which level of differentiation is to be expressed by a uniqueSubstanceID. That is, depending on the respective application, a uniqueSubstanceID may or may not distinguish the different isomers (or even polymorphic forms) of a *substance*.

WLN

<u>Description</u>

The Wiswesser line notation (WLN), also known as Wiswesser line formula, is a precise and concise means of expressing structural formulas as character strings. The basic idea is to use letter symbols to denote functional groups and numbers to express the lengths of chains and the sizes of rings (Smith 1968).

<u>Characteristics</u>

- Domain: *Chemical species* or *molecular entity*

- Range: string (built-in XML Schema datatype)

- Functional

2.2. Molecular Structure

The ontology module *molecular_structure* characterizes the molecular structure of pure *substances*. It refines the definition of *molecular entity*. To this end, the ontology module adopts some ontological concepts from the ChEBI[5] Ontology (EBI, 2007; OLS, 2006), particularly of its sub-ontology *Molecular Structure*. The ChEBI Ontology is a non-proprietary ontology concerned with the classification of "small" chemical compounds. It consists of four sub-ontologies; the sub-ontology *Molecular Structure* provides (structural) descriptions of molecular entities and of parts thereof, based on the composition and/or the connectivity between the constituting atoms. The terminology and nomenclature of *Molecular Structure* is based on the 'IUPAC Compendium of Chemical Terminology' (McNaught and Wilkinson, 1997), informally known as the 'Gold Book'. The

---

[5] ChEBI: Chemical Entities of Biological Interest

compendium extracts the definitions of terms described in the various IUPAC glossaries and other IUPAC nomenclature documents, including 'Glossary of Terms Used in Physical Organic Chemistry' (Müller, 1994), 'Glossary of Class Names of Organic Compounds and Reactive Intermediates Based on Structure' (Moss et al., 1995), and 'Basic Terminology of Stereochemistry' (Moss, 1996).
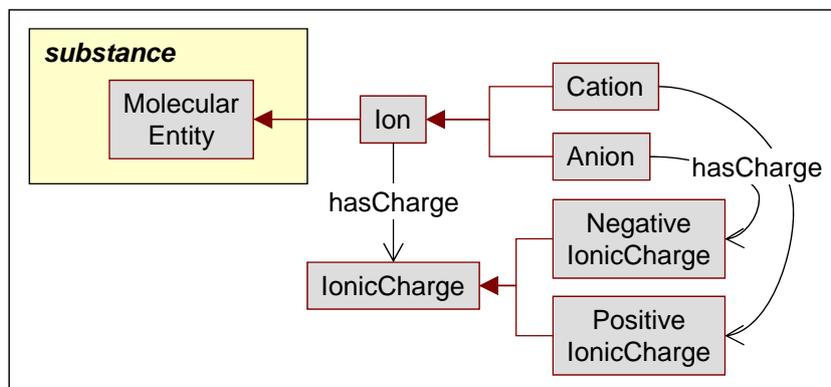
Fig. 11: Representation of ions

The class *ion* is introduced as a first refinement of *molecular entity*. An *ion* is defined as a *molecular entity* that has an *ionic charge* (cf. Fig. 11). *Ions* that have a *positive ionic charge* are classified as *cations*, those with a *negative ionic charge* are classified as *anions*.
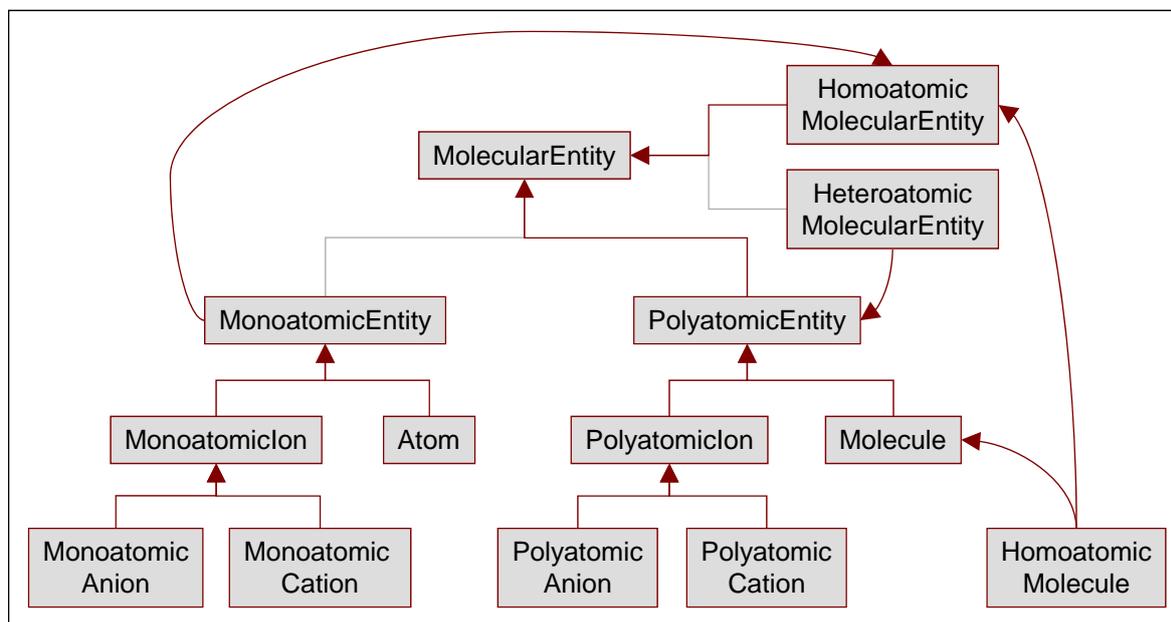
Fig. 12: Representation of ions

Fig. 12 presents a further refinement of the class *molecular entity*; the refinement is based on two different criteria:

- The differentiation *monoatomic* vs. *polyatomic entity* considers the number of atoms a *molecular entity* consists of. For example, the Na^+ ion is a *monoatomic entity*, while the N_2 molecule is a *polyatomic entity*.

23

- The differentiation *homoatomic* vs. *heteroatomic molecular entity* considers the number of elements a molecular entity consists of. The N_2 molecule, for instance, is a *homoatomic molecular entity*, as it consists of a single element.

Consequently, each *monoatomic entity* is a *homoatomic molecular entity*, and each *heteroatomic molecular entity* is a *polyatomic entity*.

Monoatomic ion and *polyatomic ion* (and their respective subclasses) are defined analogously to *ion*.

The charge of an *ion* is represented by one of the following instances of *ionic charge*: e, 2e, 3e, -e, -2e, or -3e (cf. Fig. 13). The individual e represents the value of the elementary_charge, a *physical constant* that denotes the electric charge of a single *monoatomic ion* (1.60217653E-19 C). The other *ionic charges* are (positive or negative) integer multiples of e.
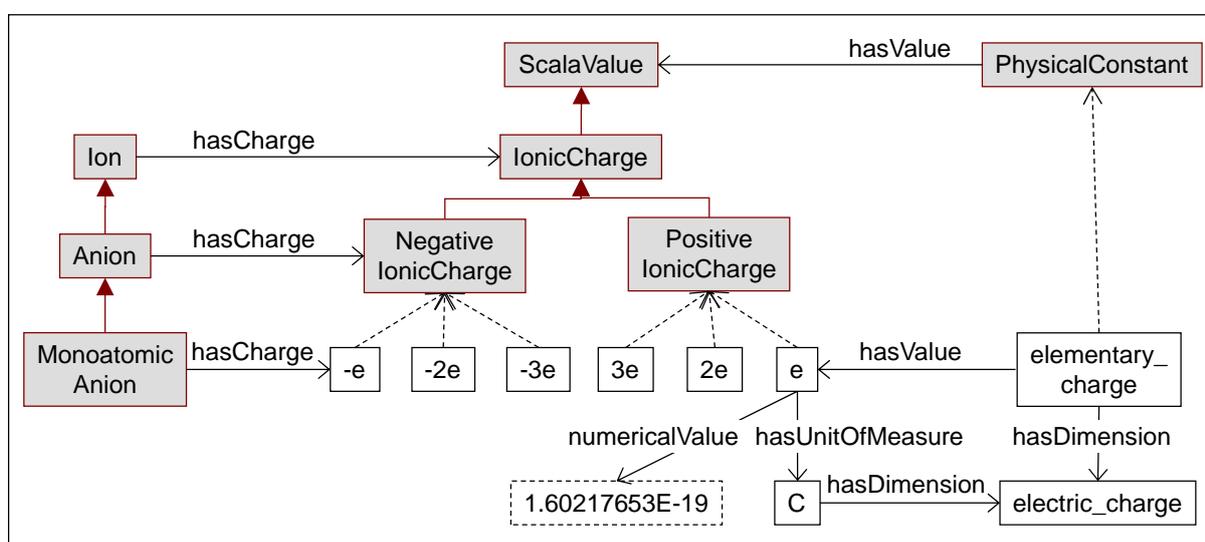
Fig. 13: Representation of ionic charge

Fig. 14 presents concepts that can be used to describe the constitution of *molecular entities*: A *molecular entity* may contain a *molecular group*, which denotes either a linked collection of *atoms* or a single *atom* within a *molecular entity*. As each *molecular group* is a *molecular entity*, it may again contain some *molecular* sub-*group*. A restriction is placed on the *atom* class such that it cannot contain a *molecular group*.
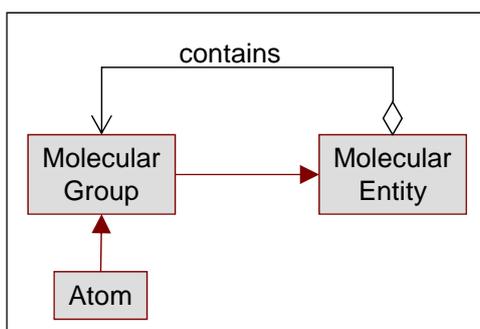
Fig. 14: Composition of Molecular Entities

Usage

The module *molecular_structure* is extended by the ontology modules *polymers*, *atoms*, and *substance_class*. The usage of the concepts of *molecular_structure* is explained in the specifications of these modules (cf. Secs. 2.3, 2.5, and 2.7).

Concept Descriptions

Individual concepts of the module *molecular_structure* are defined below.

Classes

Anion

Description

An *anion* is a monoatomic or polyatomic species having one or more elementary charges of the electron (McNaught and Wilkinson, 1997).

Definition

An *aninon* is an *ion* that has a *negative ionic charge*.

Relations

- *Anion* is a subclass of *ion*.

Atom

Description

An *atom* is the smallest particle still characterizing a chemical element (McNaught and Wilkinson, 1997).

Relations

- *Atom* is a subclass of *monoatomic entity*.
- *Atom* is a subclass of *molecular group*.
- An atom cannot have an *ionic charge*.

Cation

Description

A *cation* is a monoatomic or polyatomic species having one or more elementary charges of the proton (McNaught and Wilkinson, 1997).

Definition

A *cation* is an *ion* that has a *positive ionic charge*.

Relations

- *Cation* is a subclass of *ion*.

Heteroatomic molecular entity

Description

A *heteroatomic molecular entity* is a *molecular entity* consisting of two or more [distinct] chemical elements (OLS, 2006).

Relations

- *Heteroatomic molecular entity* is a subclass of *polyatomic entity*.

Homoatomic molecular entity

Description

A *homoatomic molecular entity* is a *molecular entity* consisting of one or more *atoms* of the same element (OLS, 2006).

Relations

- *Homoatomic molecular entity* is a subclass of *molecular entity*.

Homoatomic molecule

Description

A *homoatomic molecule* is a *molecule* consisting of *atoms* of the same element (OLS, 2006).

Relations

- *Homoatomic molecule* is a subclass of *homoatomic molecular entity*.

- *Homoatomic molecule* is a subclass of *molecule*.

Ion

Description

An *ion* is an atomic or molecular particle having a net electric charge (McNaught and Wilkinson, 1997).

Definition

An *ion* is a *molecular entity* that has some *ionic charge*.

Relations

- *Ion* is a subclass of *molecular entity*.

- An *ion* has exactly one *ionic charge*.

Ionic charge

<u>Description</u>

Ionic charge is a *scalar value* that represents the electric charge of an *ion*.

<u>Definition</u>

Ionic charge is either a *positive ionic charge* or a *negative ionic charge*.

<u>Relations</u>

-   *Ionic charge* is a subclass of *scalar value*.

Molecular entity (continued)

<u>Additional relations</u>

-   A *molecular entity* may have one *ionic charge*.

-   A *molecular entity* may contain some *molecular groups*.

<u>Usage</u>

The constitution of a *molecular entity* can be described by means of a chemicalFormula. Also, the constituting *atoms* or *molecular groups* can be explicitly indicated via the contains relation.

Molecular group

<u>Description</u>

A linked collection of atoms or a single atom within a *molecular entity*.

<u>Relations</u>

-   *Molecular group* is a subclass of *molecular entity*.

<u>Usage</u>

Molecular group can represent any characteristic part of a *molecular entity*, such as an active center, a functional group, a monomer unit within a polymer, etc.

Molecule

<u>Description</u>

A *molecule* is an electrically neutral entity consisting of more than one *atom* (McNaught and Wilkinson, 1997).

Monoatomic anion

<u>Description</u>

A *monoatomic anion* is an *anion* consisting of a single *atom*.

27

Definition

A *monoatomic anion* is a *monoatomic ion* that has a *negative ionic charge*.

<u>Relations</u>

- *Monoatomic anion* is a subclass of *monoatomic ion*.

- A *monoatomic anion* has a *negative ionic charge*.

Monoatomic cation

<u>Description</u>

A *monoatomic cation* is an *cation* consisting of a single *atom*.

<u>Definition</u>

A *monoatomic cation* is a *monoatomic entity* that has a *positive ionic charge*.

<u>Relations</u>

- *Monoatomic cation* is a subclass of *monoatomic ion*.

- A *monoatomic cation* has a *positive ionic charge*.

Monoatomic entity

<u>Description</u>

A *monoatomic entity* is a *molecular entity* consisting of a single *atom*.

<u>Definition</u>

Monoatomic entity is either an *atom* or a *monoatomic ion*.

<u>Relations</u>

- *Monoatomic entity* is a subclass of *homoatomic molecular entity*.

- A *monoatomic entity* cannot contain a *molecular group*.

Monoatomic ion

<u>Description</u>

A *monoatomic ion* is an *ion* consisting of a single *atom* (OLS, 2006)

<u>Definition</u>

Monoatomic ion is a *monoatomic entity* that has an *ionic charge*.

<u>Relations</u>

- *Monoatomic ion* is a subclass of *monoatomic entity*.

28

- A *monoatomic ion* has one *ionic charge*.

- A *monoatomic ion* is either an *monoatomic anion* or a *monoatomic cation*.

Negative ionic charge

<u>Description</u>

A *negative ionic charge* is a *scalar value* that represents the negative electric charge of an *ion*.

<u>Relations</u>

- *Negative ionic charge* is a subclass of *ionic charge*.

<u>Usage</u>

Instances of *negative ionic charge* are _-e or integer multiples of _-e.

Polyatomic anion

<u>Description</u>

A *polyatomic anion* is an *anion* consisting of more than one *atom* (OLS, 2006).

<u>Definition</u>

A *Polyatomic anion* is a *polyatomic entity* that has a *negative ionic charge*.

<u>Relations</u>

- *Polyatomic anion* is a *polyatomic ion*.

- A *polyatomic anion* has one *negative ionic charge*.

Polyatomic cation

<u>Description</u>

A *polyatomic cation* is a *cation* consisting of more than one *atom* (OLS, 2006).

<u>Definition</u>

A *polyatomic cation* is a *polyatomic entity* that has a *positive ionic charge*.

<u>Relations</u>

- *Polyatomic cation* is a *polyatomic ion*.

- A *polyatomic cation* has one *positive ionic charge*.

Polyatomic entity

<u>Description</u>

Any *molecular entity* consisting of more than one *atom* is a *polyatomic entity* (OLS, 2006).

Relations

- *Polyatomic entity* is a *molecular entity*.

Polyatomic ion

Description

A *polyatomic ion* is an *ion* consisting of more than one *atom* (OLS, 2006).

Definition

A *polyatomic ion* is an *polyatomic entity* that has an *ionic charge*.

Relations

- *Polyatomic ion* is a *polyatomic entity*.

- A *polyatomic ion* has one *ionic charge*.

Positive ionic charge

Description

A *positive ionic charge* is a *scalar value* that represents the positive electric charge of an *ion*.

Relations

- *Positive ionic charge* is a subclass of *ionic charge*.

Usage

Instances of *positive ionic charge* are e or integer multiples of e.

Relations

hasCharge

Description

The relation hasCharge indicates the *ionic charge* of an *ion*.

Characteristics

- Specialization of hasCharacteristic

- Domain: *Molecular entity*

- Range: *Ionic charge*

- Functional

Attributes

atomicNumber

<u>Description</u>

The atomicNumber (also known as the proton number) is the number of protons found in the nucleus of an *atom*.

<u>Characteristics</u>

- Specialization of uniqueSubstanceID

- Domain: *Atom*

- Datatype: positiveInteger (built-in XML Schema datatype)

- Functional

- Inverse functional

Individuals

ElementaryCharge

<u>Description</u>

The ElementaryCharge is a fundamental *physical constant* that denotes the electric charge carried by a single proton, or equivalently, the negative of the electric charge carried by a single electron.

<u>Characteristics</u>

- ElementaryCharge is an instance of *physical constant*

- ElementaryCharge has the dimension electric_charge.

- ElementaryCharge has the value e.

e

<u>Description</u>

The *scalar value* e represents the value of the ElementaryCharge.

<u>Characteristics</u>

- e is an instance of *positive ionic charge*

- e has a numericalValue of 1.6021765E-19.

- e has the *unit of measure* Coulomb

_2e

<u>Description</u>

_2e is a *scalar value*; it is defined by the equation _2e = 2 * e.

<u>Characteristics</u>

- _2e is an instance of *positive ionic charge*

- _2e has a numericalValue of 3.204353E-19.

- _2e has the *unit* Coulomb

_3e

<u>Description</u>

_3e is a *scalar value*; it is defined by the equation _3e = 3 * e.

<u>Characteristics</u>

- _3e is an instance of *positive ionic charge*

- _3e has a numericalValue of 4.8065294E-19

_-e

<u>Description</u>

_-e is a *scalar value*; it is defined by the equation _-e = (-1) * e.

<u>Characteristics</u>

- _-e is an instance of *negative ionic charge*

- _-e has a numericalValue of -1.6021765E-19.

_-2e

<u>Description</u>

_-2e is a *scalar value*; it is defined by the equation _-2e = (-2) * e.

<u>Characteristics</u>

- _-2e is an instance of *negative ionic charge*

- _-2e has a numericalValue of –3.204353E-19.

_-3e

<u>Description</u>

_-3e is a *scalar value*; it is defined by the equation _-3e = (-3) * e.

- **_-3e** is an instance of *negative ionic charge*

- **_-3e** has a numericalValue of -4.8065294E-19

2.3. Polymers

The ontology module *polymers* is concerned with the structural description of *macromolecules*. A *macromolecule* is a *molecule* of high *molecular weight*, the structure of which essentially comprises the multiple repetitions of units derived, actually or conceptually, from *molecules* of low *molecular weight* (McNaught and Wilkinson, 1997).

Industrially, *macromolecules* are synthesized from *monomer molecules* which undergo polymerization (cf. Fig. 15). Besides the *macromolecules*, *oligomer molecules* are often formed as intermediates or by-products of the polymerization reaction. *Oligomers* are *molecules* of intermediate *molecular weight*, the structure of which essentially comprises a small plurality of units derived, actually or conceptually, from *molecules* of lower *molecular weight* (McNaught and Wilkinson, 1997).

The structure of a *macromolecule* can be described by indicating its building blocks, the so-called *constitutional units*. Three types of *constitutional units* are distinguished:

- *Monomer units* are *constitutional units* resulting from a *monomer molecule* that has been polymerized.

- *Repeating units* are the shortest *constitutional units* that can be found repeatedly in a *macromolecule*.

- Finally, *end-groups* are *constitutional units* that form the extremities of a *macromolecule* or *oligomer molecule*.
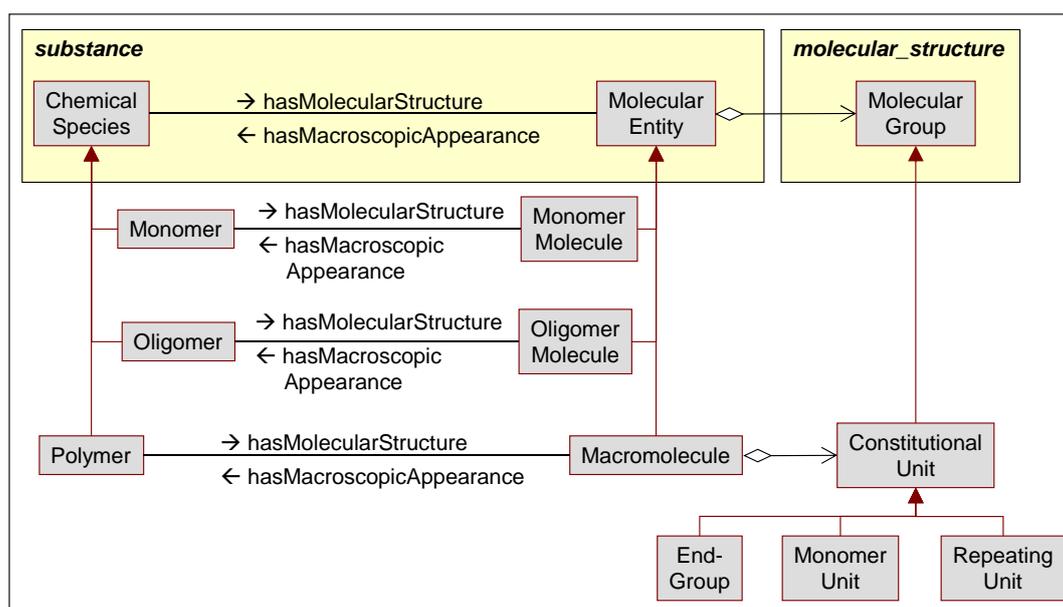

Fig. 15: Ontology module *polymers*

Monomers are *chemical species* composed of *monomer molecules*, and *oligomers* are *chemical species* composed of *oligomer molecules*. Similarly, *polymers*, such as polyethylene or polyamide, are *chemical species* that are composed of *macromolecules*. Note, however, that a *polymer* consists of a statistical distribution of different *macromolecules*, which vary with respect to chain length, side branches, cross-linkage, etc. Thus, an instance of *polymer* represents one specific configuration of *macromolecules*.

Usage

Depending on the application, an instance of *macromolecule* can represent either one specific molecule or a group of molecules of the same type, yet with different molecular structures and chain lengths. In the latter case, the chemicalFormula of the *macromolecule* is indicated in a generic form, with a variable number of repeating units (e.g., CH_3-$(C2H4)_n$-CH_3 for the polyethylene molecule shown in Fig. 16).
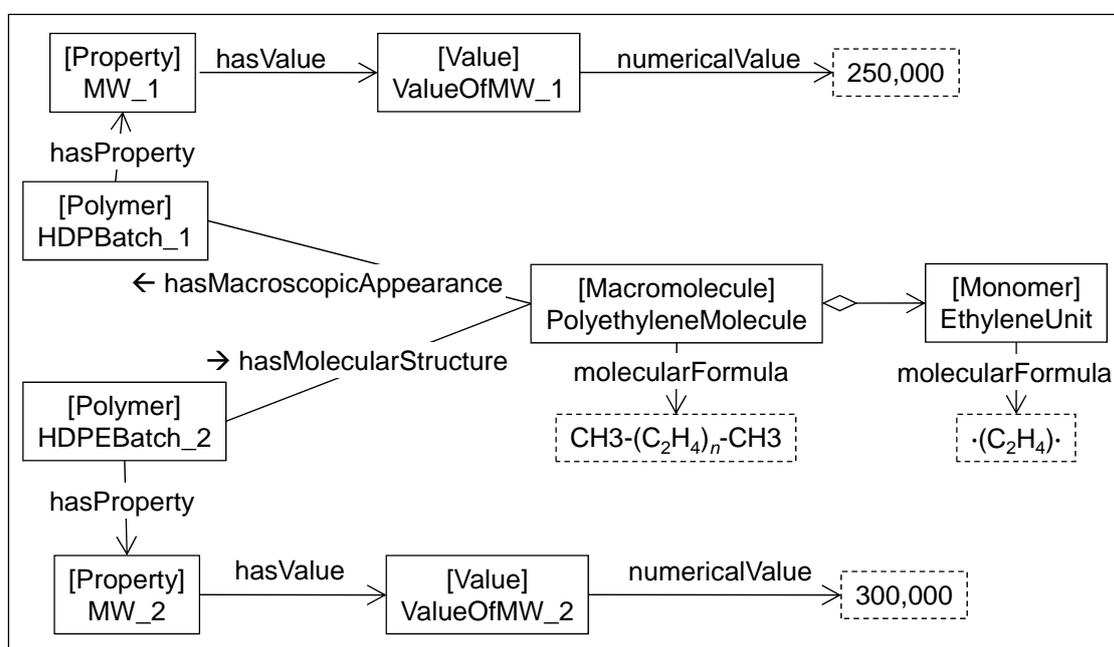

Fig. 16: : Two batches of HDPE with different *molecular weights*

Like any *chemical species*, a *polymer* has *chemical component constants*, such as the *molecular weight*. Since each instance of *polymer* represents one individual configuration of *macromolecules*, the *values* of its *chemical component constants* are valid for this specific configuration only. Fig. 16 shows exemplarily two different *polymers* (HDPE_batch_1 and HDPE_batch_2) that have different *molecular weights* (MW_1 and MW_2). Both *polymers* have the same (generic) molecular structure, which is represented by an instance of *macromolecule* (PolyethyleneMolecule) and further characterized through indication of its constituting *monomer unit* (EthyleneUnit).

Please note that polymer materials may also be modeled as *mixtures* consisting of a number of *pseudo components* in the substance module (cf. Sect. 2.1). However, this kind of representation is not covered by this ontology module.

Concept Descriptions

Individual concepts of the module *polymers* are defined below.

Classes

Constitutional unit

Description

A *constitutional unit* is a *molecular group* that constitutes a characteristic part of a *macromolecule*.

Relations

- *Constitutional unit* is a subclass of *molecular group*.

End-group

Description

An *end-group* is a *constitutional unit* that is an extremity of a *macromolecule* or *oligomer molecule* (McNaught and Wilkinson, 1997).

Relations

- *End-group* is a subclass of *constitutional unit*.

Macromolecule

Description

A *molecule* of high *molecular weight*, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from *molecules* of low *molecular weight* (McNaught and Wilkinson, 1997).

Relations

- *Macromolecule* is a subclass of *molecule*.
- A *macromolecule* contains some *constitutional unit*.
- A *macromolecule* can only contain *constitutional units*.
- The macroscopic appearance of a *macromolecule* can only be a *polymer*.

Monomer

Description

A *substance* composed of *monomer molecules* (McNaught and Wilkinson, 1997).

Relations

- *Monomer* is a subclass of *Chemical species*.

-   The molecular structure of a *monomer* can only be a *monomer molecule*.

Monomer molecule

Description

A *molecule* which can undergo polymerization thereby contributing *constitutional units* to the essential structure of a *macromolecule* (McNaught and Wilkinson, 1997).

Relations

-   *Monomer molecule* is a subclass of *molecule*.

-   The macroscopic appearance of a *monomer molecule* can only be a *monomer*.

Monomer unit

Description

The largest *constitutional unit* contributed by a single *monomer molecule* to the structure of a *macromolecule* or *oligomer molecule*.

Note: The largest constitutional unit contributed by a single monomer molecule to the structure of a *macromolecule* or *oligomer molecule* may be described as either monomeric, or by monomer used adjectivally.

Relations

-   *Monomer unit* is a subclass of *constitutional unit*.

Oligomer

Description

A *substance* composed of *oligomer molecules* (McNaught and Wilkinson, 1997).

Relations

-   *Oligomer* is a subclass of *chemical species*.

-   The molecular structure of an *oligomer* can only be an *oligomer molecule*.

Oligomer molecule

Description

A *molecule* of intermediate *molecular weight*, the structure of which essentially comprises a small plurality of units derived, actually or conceptually, from *molecules* of lower *molecular weight*. A *molecule* is regarded as having an intermediate *molecular weight* if it has properties which do vary significantly with the removal of one or a few of the units. (McNaught and Wilkinson, 1997).

Relations

- *Oligomer molecule* is a subclass of *molecule*.

- The macroscopic appearance of an *oligomer molecule* can only be an *oligomer*.

Polymer

Description

A *substance* composed of *macromolecules* (McNaught and Wilkinson, 1997).

Relations

- *Polymer* is a subclass of *chemical species*.

- The macroscopic appearance of a *polymer* can only be a *polymer molecule*.

Repeating unit

Description

A *repeating unit* is the shortest *constitutional unit* that can be found repeatedly in a *macromolecule*.

Relations

- *Repeating unit* is a subclass of *constitutional unit*.

2.4. Chemical Species

The ontology module *chemical_species* is an application-oriented extension of the ontology module *substance*. It contains several thousand instances of *chemical species*, their substanceIdentifiers, chemicalFormulas and *molecular weights*. The data stems from the NIST Chemistry WebBook (Linstrom and Mallard, 2005), a non-proprietary online database providing physical property data of pure substances.

The following entries of the NIST Chemistry WebBook are available in the *chemical_species* module (cf. Fig. 17):

- Species name: used as identifiers[6] for the instances of *chemical species*.

- Other names (synonyms) for the species: represented as values of the name attribute.

- Chemical formula: represented as values of the attribute molecularFormula.

- CAS Registry Number (if available): represented as values of CAS_RegistryNumber.

---

[6] In order to comply with the OWL syntax, certain characters in the species names had to be replaced by character symbols. Fig. 17. lists the transformation rules that were applied for this task.

- IUPAC International Chemical Identifier (if available): represented as values of the InChI attribute.

- Molecular weight: represented as a *scalar value* of *molecular weight*.

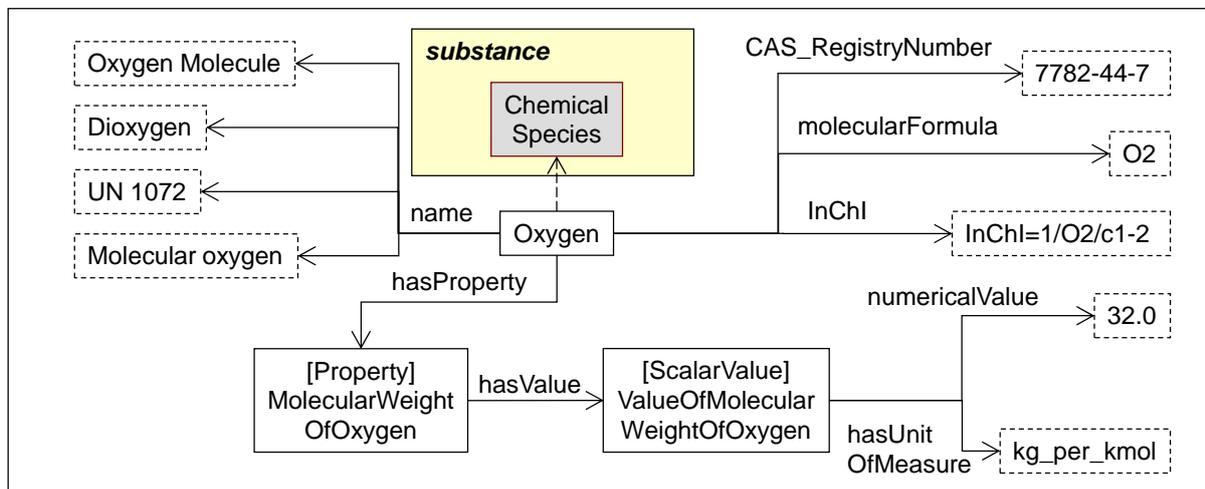Fig. 17 shows exemplarily the representation of Oxygen.

Fig. 17: Ontological representation of Oxygen

Usage

The ontology module *chemical_species* contains instance data of about approximately 25,000 *chemical species*. Most of the current ontology management systems and reasoners cannot cope with such a large amount of information. For that reason, the OWL file *chemical_species.owl* provides a reduced version of *chemical_*species, comprising only ca. 2,000 selected *chemical species*. The full-fledged version of the ontology module is available as the OWL file *chemical_species_all.owl*.

2.5. Atoms

The ontology module *atoms* is an application-oriented extension of the ontology module *molecular_structure*. It represents the chemical elements as instances of the *atom* class. Each *atom* refers to the corresponding instance of *chemical species* (represented in the ontology module *chemical_species*) via the relation hasMacroscopicAppearance. An example is given in Fig. 18.
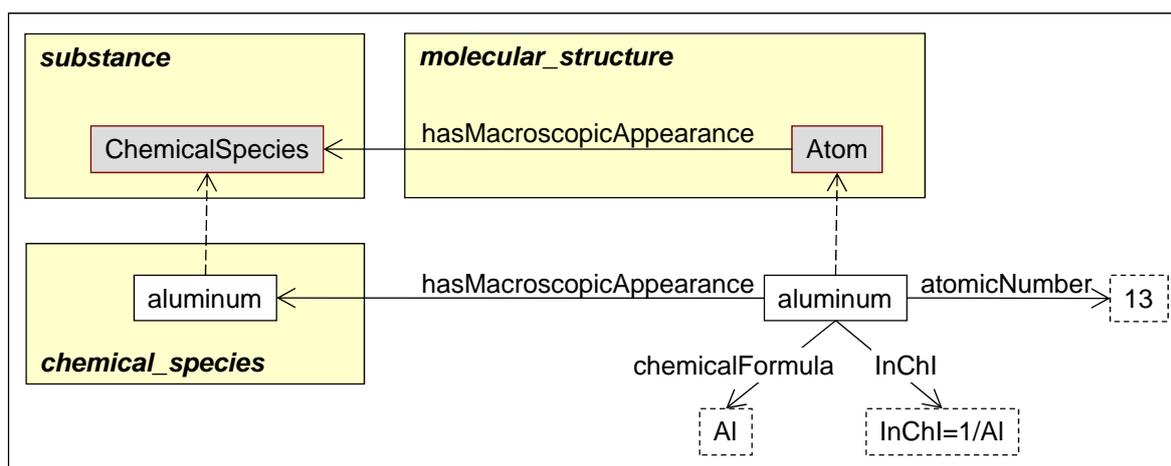
Fig. 18: Ontological representation of the **Aluminum** *atom*

The same instance identifiers are used for *chemical species* and corresponding *atoms*. This is done since, in practice, the distinction between *molecular entities* and *chemical species* is often neglected. Thus, the name of a compound may refer to the respective *molecular entity* or to the corresponding *chemical species* (e.g., **Aluminum** may mean a single atom of Al or a macroscopic amount of Al). By means of the namespace prefix, the instances can be kept apart (e.g., **chemical_species:Aluminum** vs. **atoms:Aluminum**).

Each *atom* is characterized by its atomicNumber. Additionally, the InChI and molecularFormula of the *atom* are indicated (this would not be necessary since the InChI and molecularFormula have already been specified in the *chemical_species* module, but for reasons of user convenience the information is duplicated here).

2.6. Macromolecules

The ontology module *macromolecules* provides a few examples (i.e., individuals) of *macromolecules* in order to demonstrate the usage of the module *polymers* (cf. 2.3). For the description of the respective individuals, please refer to the formal specification.

2.7. Substance Class

The ontology module *substance_class* is an application-oriented extension of the **substance** partial model; it is concerned with the classification of pure *substances* and *molecular entities*. Of course, there are different ways to classify substances, and each has its individual assets and drawbacks. The *substance_class* module merely suggests one possible approach to substance classification, which may or may not be applicable to a particular problem at hand.

The classification criterion that is applied here is *chemical similarity* (Schummer, 1998): *Molecular entities* are categorized into *substance classes* like *alcohols*, *carboxylic acids*, *esters*, *aldehydes*, *amines*, etc. Note that a *molecular entity* may be assigned to multiple substance classes – for example, it may belong to the class of *acids* as well as to the class of *aromatics*.

39

Molecular entities are firstly classified into *organic compounds* and *inorganic compounds* (Fig. 19). Below, the various substance classes are defined, which may or may not have further subclasses of their own.
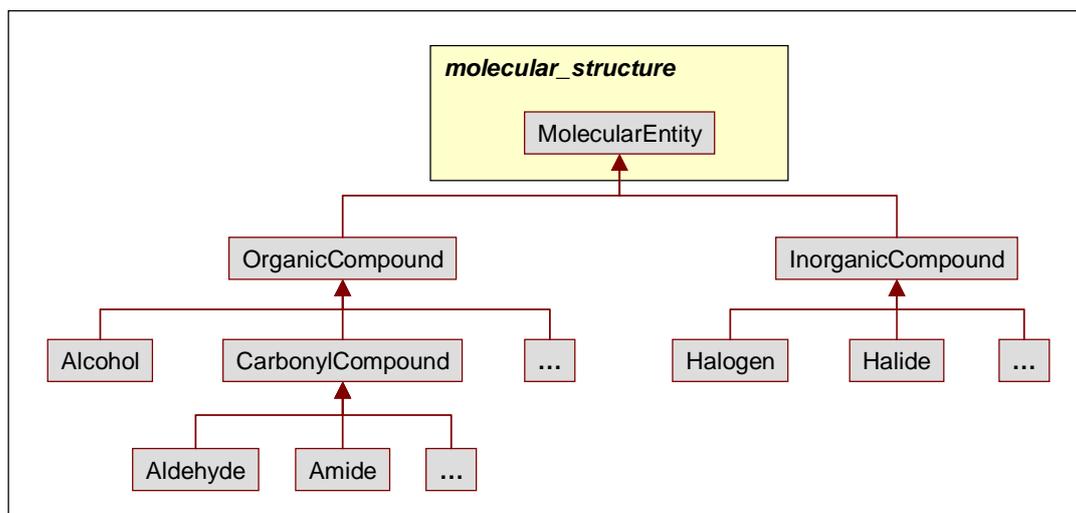

Fig. 19: Some exemplary substance classes

Each substance class has one characteristic *functional group*, as indicated in Fig. 20. A *functional group* is a *molecular group* that has "similar chemical properties whenever it occurs in different compounds. It defines the characteristic physical and chemical properties of families of organic compounds" (McNaught and Wilkinson, 1997). *Molecular entities* that belong to the same substance class contain the same *functional group*. For example, all *alcohols* contain a HydroxylGroup.
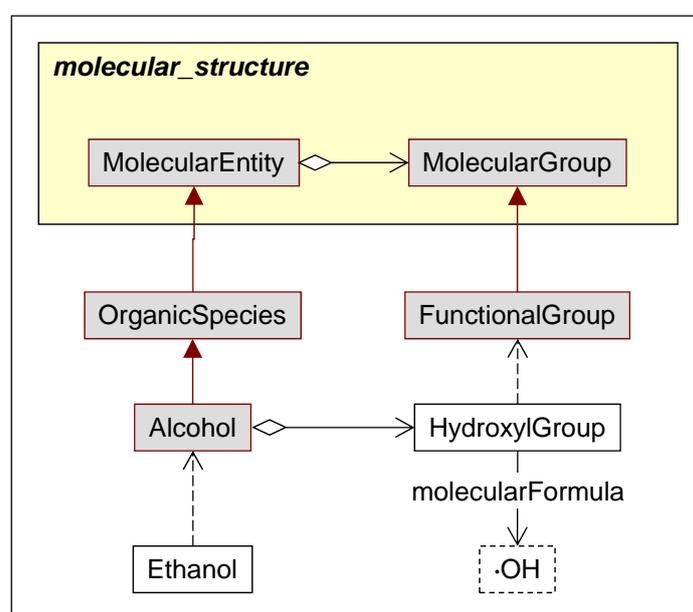

Fig. 20: *Functional groups*

Only a limited number of substance classes have been defined, so far. In a later version of OntoCAPE, it is planned to integrate further parts of the ChEBI ontology (EBI, 2007), which gives an extensive classification of *molecular entities*.

Concept Descriptions

Individual concept of the module *substance_class* is defined below.

Functional group

<u>Description</u>

"Organic compounds are thought of as consisting of a relatively unreactive backbone, for example a chain of hybridized carbon atoms, and one or several functional groups. The functional group is an atom, or a group of atoms that has similar chemical properties whenever it occurs in different compounds. It defines the characteristic physical and chemical properties of families of organic compounds" (McNaught and Wilkinson, 1997).

<u>Relations</u>

- *Functional group* is a subclass of *molecular group*.

For a description of the individual substance classes, refer to the OWL implementation of the ontology module *substance_class*.

2.8. Reaction Mechanism

The ontology module *reaction_mechanism* allows the representation of the mechanism, i.e., the stoichiometry of chemical reactions. *Chemical reaction* is the key concept of this module. In OntoCAPE, a *chemical reaction* is modeled as a subclass of *system* (cf. Morbach et al., 2008a) and is regarded generally as a mechanism that produces chemical products by consuming some reactants.

. A *chemical reaction* may be a *reaction network* or a *single reaction*. A *reaction network* represents a loose collection of a number of *single reactions*, usually for indicating their co-existence in a specific circumstance. In particular, we consider a collection of *single reactions* in a *reaction network* if we want to explicitly refer to each of the *single reactions* of a more complex reaction mechanism involving many parallel and consecutive reactions. A *single reaction* in turn may be a *composite reaction* or an *elementary reaction*. An *elementary reaction* refers to a molecular transformation for which no molecular intermediates have been detected or need to be postulated for a proper description. In contrast, a *composite reaction* refers to molecular transformations involving a number of reaction steps and intermediates with a known mechanism the details of which are however not of interest on the granularity of the reactor hosting the reaction. In other words, *composite reactions* refer to a net reaction of a more complex *reaction network* which is not described in detail in a certain context. If a detailed description is desired in addition, a *reaction network* could be formulated and associated with the *composite reaction*. *Single reactions* can also be classified into *irreversible reactions* and *reversible reactions*.
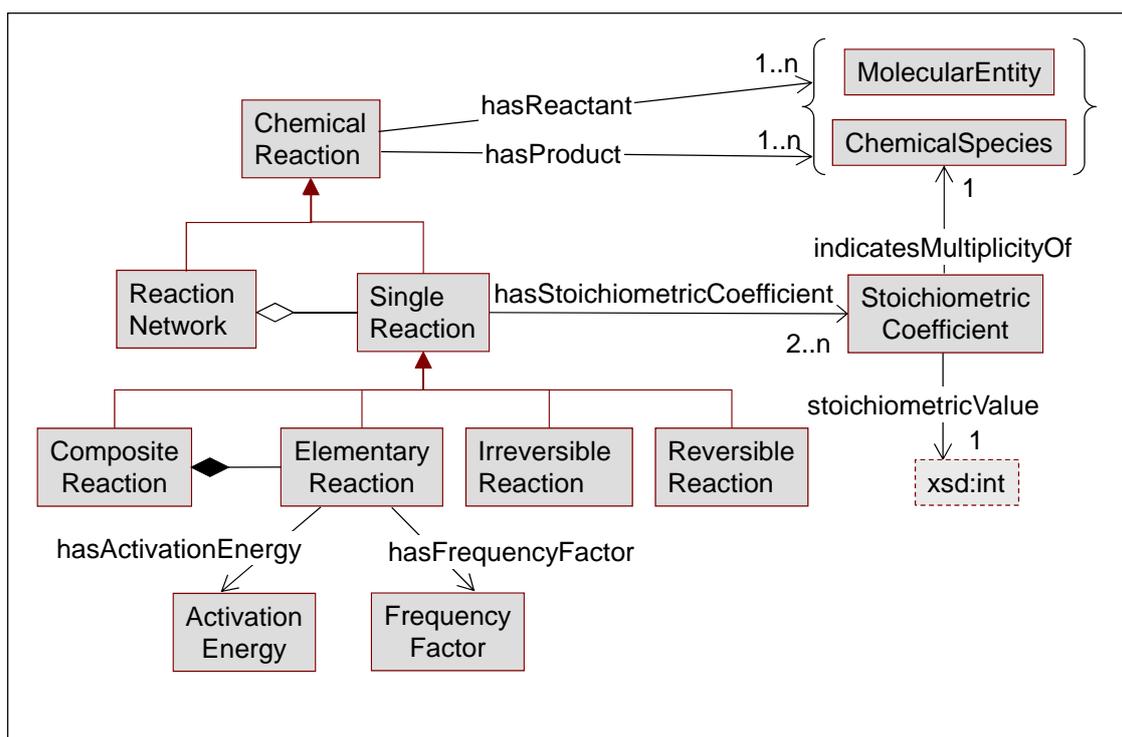
Fig. 21: Ontology module *reaction_mechanism*

The reactants and products of the *chemical reaction* are identified via the relations hasReactant and hasProduct, respectively. The reactants and products can be either *molecular entities* or *chemical species*, depending on whether one wants to describe the reaction on the molecular level (conversion of some *molecular entities* into some other *molecular entities*) or on the macroscopic level (conversion of *chemical species*). The stoichiometry of a *single reaction* is modeled by means of *stoichiometric coefficients*, each of which indicates the multiplicity of the respective reactant or product. The numerical value of a *stoichiometric coefficient* is specified by the attribute stoichiometricValue.

A *single reaction* is a *chemical reaction* that has a fixed stoichiometry which does not change with reaction conditions or extent of reaction (Levenspiel, 1999). The stoichiometry of a *reaction network* is given and fixed by a stoichiometric matrix. Such a matrix can be assembled implicitly from the *stoichiometric coefficients* of the individual (elementary or composite) reactions. Hence, this matrix may be either the true one (as it would be the case if there were only elementary reactions) or the one which is the result of an aggregation process of some *single reactions* into a *composite reaction* which is part of a *reaction network*. In case a *reaction network* is represented by a *single reaction* only, the apparent stoichiometry corresponding to this *single reaction* may change with the environment in which these reactions occur.

The reactants and products of the *chemical reaction* are identified via the relations hasReactant and hasProduct, respectively. The reactants and products can be either *molecular entities* or *chemical species*, depending on whether one wants to describe the reaction on the molecular level (conversion of some *molecular entities* into some other *molecular entities*) or on the macroscopic level (conversion of *chemical species*). The stoichiometry of a *single reaction* is modeled by means of *stoichiometric coefficients*, each of

which indicates the multiplicity of the respective reactant or product. The numerical value of a *stoichiometric coefficient* is specified by the attribute stoichiometricValue.

The *reaction_mechanism* module further introduces two constant *properties* for a *elementary reaction* in connection with the application of the Arrhenius Equation to depict the temperature-dependency of the reaction rate. These two properties are *activation energy* and *frequency factor*. Note that these two concepts are introduced only to address the situations where the activity energy is regarded as independent of temperature, i.e. where the linear Arrhenius behaviour is exhibited. Chemical kinetics of more complex situations will be directly and modelled through proper kinetics *laws* as part of the *mathematical models* to be covered by Morbach et al. (2008b).

Note that the *reaction_mechanism* module only captures chemical reaction related properties which are independent of the physical context. Properties that change with the physical context are introduced later either in the **phase_system** (cf. Section 3.4) or the **CPS_behavior** (cf. Wiesner et al., 2008) partial model, depending on whether a property is dedicated to equilibrium reactions or not.

Usage

The usage of the *reaction_mechanism* module is demonstrated by the following example shown in Fig. 22: Consider the overall (composite) reaction

$$2\ H_2 + 2\ NO \rightarrow N_2 + 2\ H_2O$$
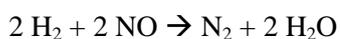
The reaction mechanism of the overall reaction is given by the following *elementary reactions*:

$$2\ NO \rightarrow N_2O_2$$

$$H_2 + N_2O_2 \rightarrow N_2O + H_2O$$

$$H_2 + N_2O \rightarrow N_2 + H_2O$$
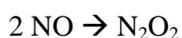
Fig. 22 illustrates how to represent the stoichiometry of the Overall Reaction.
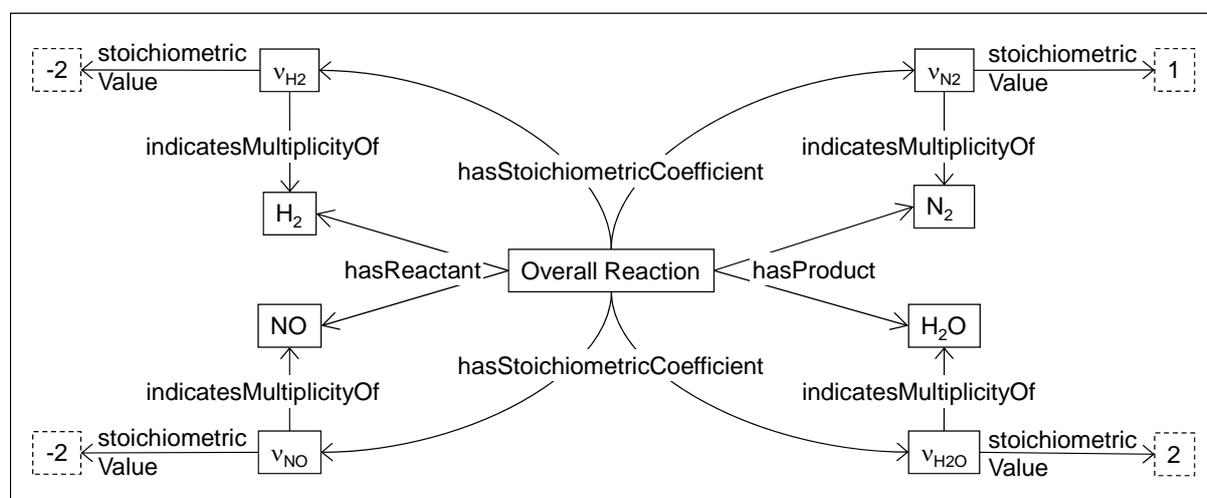
Fig. 22: Stoichiometry of the reaction **2 NO → N₂O₂**

In Fig. 23, the Overall Reaction is decomposed into the three *elementary reactions*. The reactants and products of the reactions are shown, as well. The *stoichiometric coefficients* are not represented for reasons of clarity.
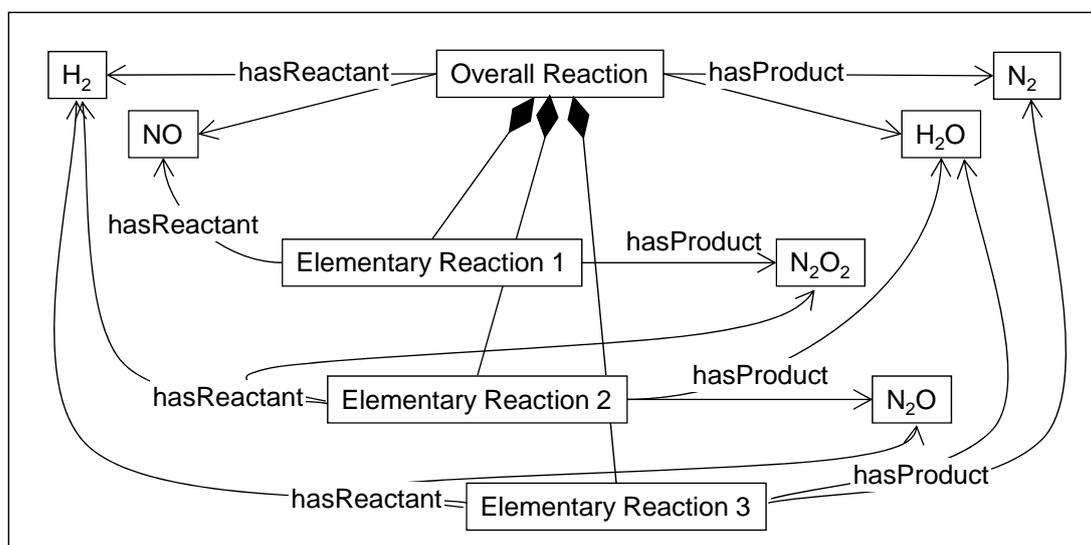

Fig. 23: The Overall Reaction is composed of three *elementary reactions*

Concept Descriptions

Individual concepts of the module *reaction_mechanism* are defined below.

Classes

Activation energy

Description

Activation energy is a constant *property* as involved in the Arrhenius Equation for depicting the temperature-dependency of reaction rates.

Relations

- *Activation energy* is a subclass of *property*.

Chemical reaction

Description

A *chemical reaction* converts some *chemical species* (or *molecular entities*) into some other *chemical species* (or *molecular entities*).

Relations

- *Chemical reaction* is a subclass of *system*.

- A *chemical reaction* has at least one reactant of type *chemical species* or *molecular entity*.

44

- A *chemical reaction* has at least one product of type *chemical species* or *molecular entity*.

- A *chemical reaction* has at least two *stoichiometric coefficients*.

Composite reaction

Description

A *composite reaction* is a *chemical reaction* that could in principle be decomposed into several *elementary reactions*. Examples are parallel reactions (simultaneously occurring *elementary reactions* that form different products from a single set of reactants) and stepwise reactions (a set of consecutive *elementary reactions* with at least one reaction intermediate).

Definition

A *composite reaction* is a *chemical reaction* that is composed of *elementary reactions*.

Relations

- *Composite reaction* is a subclass of *chemical reaction*.

- A *composite reaction* is composed of at least two *elementary reactions*.

- A *composite reaction* can only be composed of *elementary reactions*.

Elementary reaction

Description

An *elementary reaction* is a *chemical reaction* for which no reaction intermediates have been detected or need to be postulated in order to describe the *chemical reaction* on a molecular scale. An *elementary reaction* is assumed to occur in a single step and to pass through a single transition state (McNaught and Wilkinson, 1997).

Definition

An *elementary reaction* is an exclusive subsystem of a *composite reaction*.

Relations

- *Elementary reaction* is a subclass of *chemical reaction*.

- An *elementary reaction* is an exclusive subsystem of exactly one *composite reaction*.

- An *elementary reaction* can only be an exclusive subsystem of a *composite reaction*.

- An *elementary reaction* may have an *activation energy*.

- An *elementary reaction* may have a *frequency factor*.

Frequency factor

<u>Description</u>

Frequency factor is a constant *property* as involved in the Arrhenius Equation for depicting the temperature-dependency of reaction rates.

<u>Relations</u>

- *Frequency factor* is a subclass of *property*.

Irreversible reaction

<u>Description</u>

An *irreversible reaction* is *chemical reaction* that converts reactants to products which cannot be readily reversed to restore the reactants to its original state.

<u>Relations</u>

- *Irreversible reaction* is a subclass of *chemical reaction*.

Reaction network

<u>Description</u>

A *reaction network* refers to a set of *chemical reactions* that represent a loose collection of a number of *single reactions*, usually for indicating their co-existence in a specific circumstance. The stoichiometry of a *reaction network* is given and fixed by a stoichiometric matrix which can be assembled from the *stoichiometric coefficients* of the individual *elementary reactions* or *composite reactions*.

<u>Relations</u>

- *Reaction network* is a subclass of *chemical reaction*.

- A *reaction network* is aggregated of at least two *single reactions*.

- A *reaction network* can only be aggregated of *single reactions*.

Reversible reaction

<u>Description</u>

A *reversible reaction* is a *chemical reaction* that proceeds in both directions at the same time, as the product decomposes back into reactants as it is being produced.

<u>Relations</u>

- *Reversible reaction* is a subclass of *chemical reaction*.

Single reaction

<u>Description</u>

A *single reaction* is a *chemical reaction* that has a fixed stoichiometry which does not change with reaction conditions or extent of reaction (Levenspiel, 1999).

<u>Relations</u>

- *Single reaction* is a subclass of *chemical reaction*.

- *A single reaction has some stoichiometric coefficients.*

Stoichiometric coefficient

<u>Description</u>

A *stoichiometric coefficient* indicates the multiplicity of a *chemical species* or *molecular entity* that participates in a *chemical reaction*.

<u>Relations</u>

- *Stoichiometric coefficient* is a specialization of the meta class *n-ary relation*.

- A *stoichiometric coefficient* indicates the multiplicity of exactly one *chemical species* or *molecular entity*.

- The numerical value of the *stoichiometric coefficient* is specified by the attribute stoichiometricValue.

Relations

hasActivationEnergy

<u>Description</u>

The relation hasActivationEnergy identifies the *activation energy* of an *elementary reaction*.

<u>Characteristics</u>

- Specialization of the meta relation hasProperty

- Domain: *elementary reaction*

- Range: *activation energy*

hasFrequencyFactor

<u>Description</u>

The relation hasFrequencyFactor identifies the *frequency factor* of a *elementary reaction*.

Characteristics

- Specialization of the meta relation hasProperty

- Domain: *elementary reaction*

- Range: *frequency factor*

hasStoichiometricCoefficient

Description

The relation hasStoichiometricCoefficient identifies the *stoichiometric coefficients* of a *chemical reaction*.

Characteristics

- Specialization of the meta relation isOriginOf

- Domain: *chemical reaction*

- Range: *stoichiometric coefficient*

hasProduct

Description

The relation hasProduct denotes the products of a *chemical reaction*. A product is a *molecular entity* or a *chemical species* that is formed during a *chemical reaction*.

Characteristics

- Specialization of isDirectlyRelatedTo

- Domain: *chemical reaction*

- Range: *molecular entity* or *chemical species*

hasReactant

Description

The relation hasReactant denotes the reactants of a *chemical reaction*. A reactant is a *molecular entity* or a *chemical species* that is consumed in the course of a *chemical reaction*.

Characteristics

- Specialization of isDirectlyRelatedTo

- Domain: *chemical reaction*

- Range: *molecular entity* or *chemical species*

indicatesMultiplicityOf

<u>Description</u>

The relation indicatesMultiplicityOf indicates the multiplicity of the reactants and products participating in a *chemical reaction*.

<u>Characteristics</u>

- Specialization of the meta relation indicatesMultiplicityOf

- Domain: *stoichiometric coefficient*

- Range: *molecular entity* or *chemical species*

- Functional

Attributes

stoichiometricValue

<u>Description</u>

The attribute stoichiometricValue specifies the numerical value of a *stoichiometric coefficient*. It is positive for products and negative for reactants..

<u>Characteristics</u>

- Specialization of the meta relation relationAttribute

- Domain: *stoichiometric coefficient*

- Datatype: int (built-in XML Schema datatype)

- Functional

2.9. Reaction Type

The *reaction_type* module is an application-oriented extension of the ontology module *reaction_mechanism*. It describes some important types of *chemical reactions*, such as *esterification* or *hydrohalogenation*. Fig. 24 shows exemplarily the definition of the class *transesterification*: *transesterification* is a *chemical reaction* that has an *ester* and an *alcohol* as reactants and a different *ester* and a different *alcohol* as products (the classes *ester* and *alcohol* are defined in the ontology module *substance_class*). It is a subclass of the more generic *esterification* reaction, which is defined as a reaction that has an *ester* as a product.
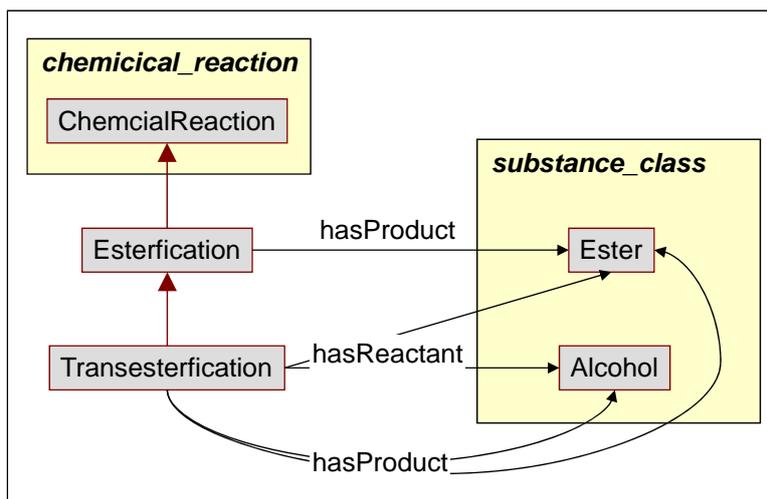
Fig. 24: Definition of the *transesterification* reaction

Concept Descriptions

For the individual concepts of the module we refer to the OWL implementation due to the high number of concepts.

3. Phase System

3.1. High-level concepts

The partial model **phase_system** describes the macroscopic thermodynamic behavior of *material*, subject to a certain *physical context*. Following the conceptualization presented in the upper_level of OntoCAPE, the behavioral aspect of a *system* can be characterized quantitatively by *properties* as well as qualitatively by *phenomena*. In the case of *phase system*, which represents the thermodynamic behavior of *material*, one phenomenon, namely thermodynamic equilibrium, constantly exists according to the definition of *phase system*. There is no other phenomenon applicable to *phase system*. Therefore, there is no practical need to characterize a *phase system* from a phenomenon point of view. Consequently, a *phase system* is only characterized by means of *properties*.

High-Level Concepts

Phase system is the key concept of this partial model. It subsumes the classes *single phase* and *multiphase system* (cf. Fig. 25).

- A *single phase* represents finite, homogeneous region of matter within which the values of its *physical quantities* are uniformly constant, i.e., they do not experience any change in passing from one point in the volume to another.

- A *multiphase system* is composed of several *single phases*; a *single phase* that is a subsystem of a *multiphase system* is characterized as a *single phase in multiphase system*.

Optionally, the interface between two adjacent *single phases* can be indicated through the class *phase interface*. A *phase interface* refers to the *single phases* via the relation isDirectlyConnectedTo. If a *phase interface* connects two instances of *single phase in multiphase system*, it is a subsystem of the overall *multiphase system*. Moreover, the connectivity rules stated in the ontology module *network_system* (cf. Morbach et al., 2008a) must be obeyed.
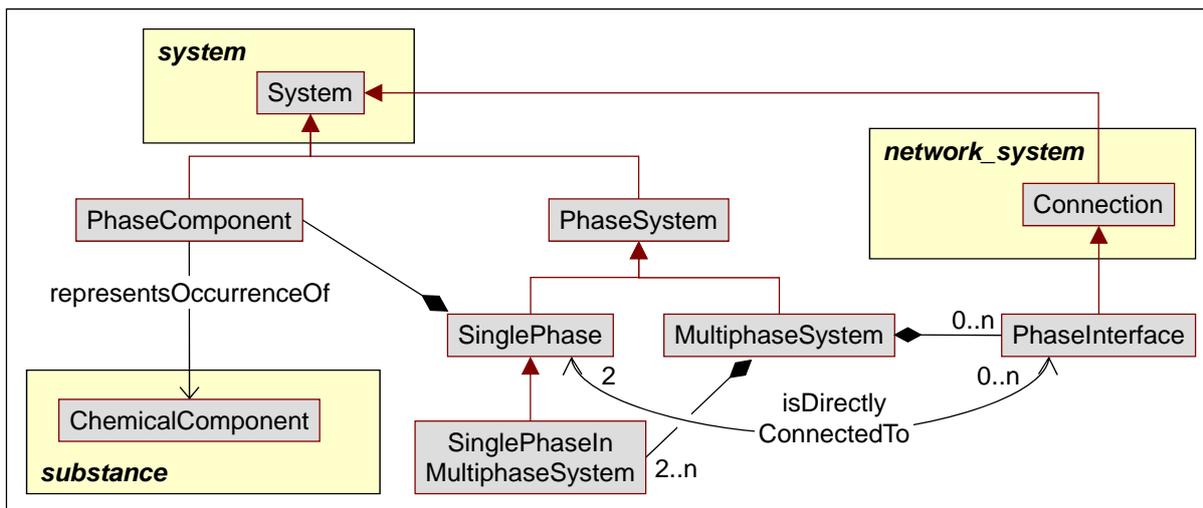
Fig. 25: Major concepts of partial model **phase_system**

Finally, the class *phase component* represents the occurrence of a *chemical component* in a *phase system*. The relation representsOccurenceOf establishes the relationship between a *phase component* and the corresponding *chemical component*.

The *state of aggregation* (a.k.a. state of matter) is a characteristic of *a single phase* that describes its physical state (cf. Fig. 26). Well-known examples of *states of aggregation* are solids, liquids, and gases. There are other, less familiar states, such as plasma or Bose-Einstein condensates.
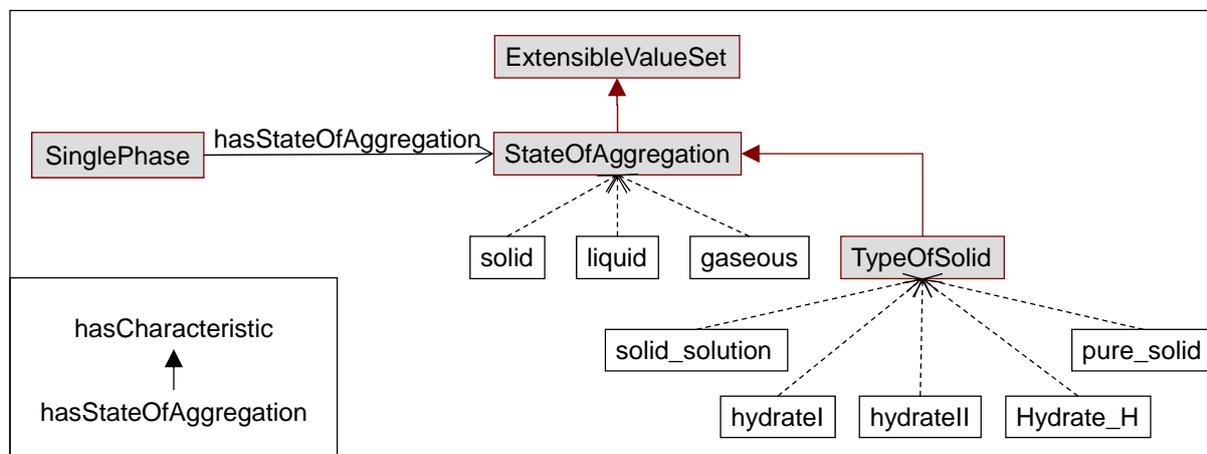

Fig. 26: State of aggregation

State of aggregation is linked to a *single phase* via the relation hasStateOfAggregation, a specialization of hasCharacteristic. *State of aggregation* is a subclass of *extensible value set*; **solid**, **liquid**, and **gaseous** are predefined instances of *state of aggregation*. The solid state can be further characterized by choosing one of the following instances of *type of solid* instead of **solid**: SolidSolution, HydrateI, HydrateII, HydrateH, or PureSolid (Drewitz et al., 2006).

3.2. Properties of Phase Systems

As explained in Sect. 1.1, the scope of the properties defined within the partial model **phase_system** is confined to those which equilibrium thermodynamics is concerned with (e.g., Modell and Reid, 1983); properties within this scope are traditionally called "physical properties", including those of stable equilibrium thermodynamic states as well as the transport coefficients that correspond to a certain thermodynamic state (such as dynamic viscosity[7], thermal conductivity, or diffusion coefficients). A common feature of all these properties is that they have no coordinate in time or space and are independent of the geometry and amount of material. Properties that don't belong to this category will be associated with *material amount* instead as defined in the *behavior* module in Wiesner et al. (2008).

As shown in Fig. 27, four distinct types of *intensive properties* are defined in partial model **phase_system**: *phase system properties*, *phase interface properties*, *phase component properties*, and *phase reaction properties*.
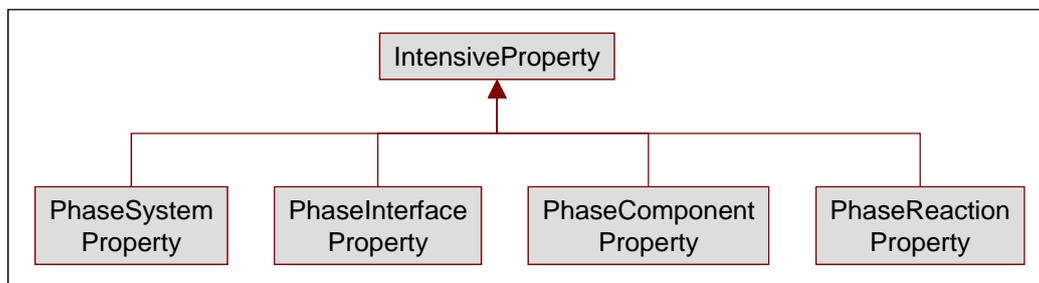

Fig. 27: High-level classification of the *intensive properties* defined in partial model **phase_system**

Fig. 28 gives an overview on the *phase system properties* currently defined in OntoCAPE. In accordance with the above discussion, *phase system properties* include (1) *thermodynamic state properties* such as *temperature*, *pressure*, *specific volume*, *density*, etc.; (2) *transport phenomena properties* such as *thermal conductivity*, *dynamic viscosity*, *and diffusion coefficient*; and (3) specific *properties* of *single phases in multiphase systems*, such as *phase ratios*. For detailed descriptions of these individual *phase system properties*, we refer the reader to the concept definition section below.

---

[7] This is restricted to the dynamic viscosity of Newtonian fluids only. For a non-Newtonian fluid, this property is dependent of the gradient of velocity, thus going beyond the scope of equilibrium thermodynamics (i.e. that of phase system). Therefore, it should be regarded as a property of a material amount in that case.
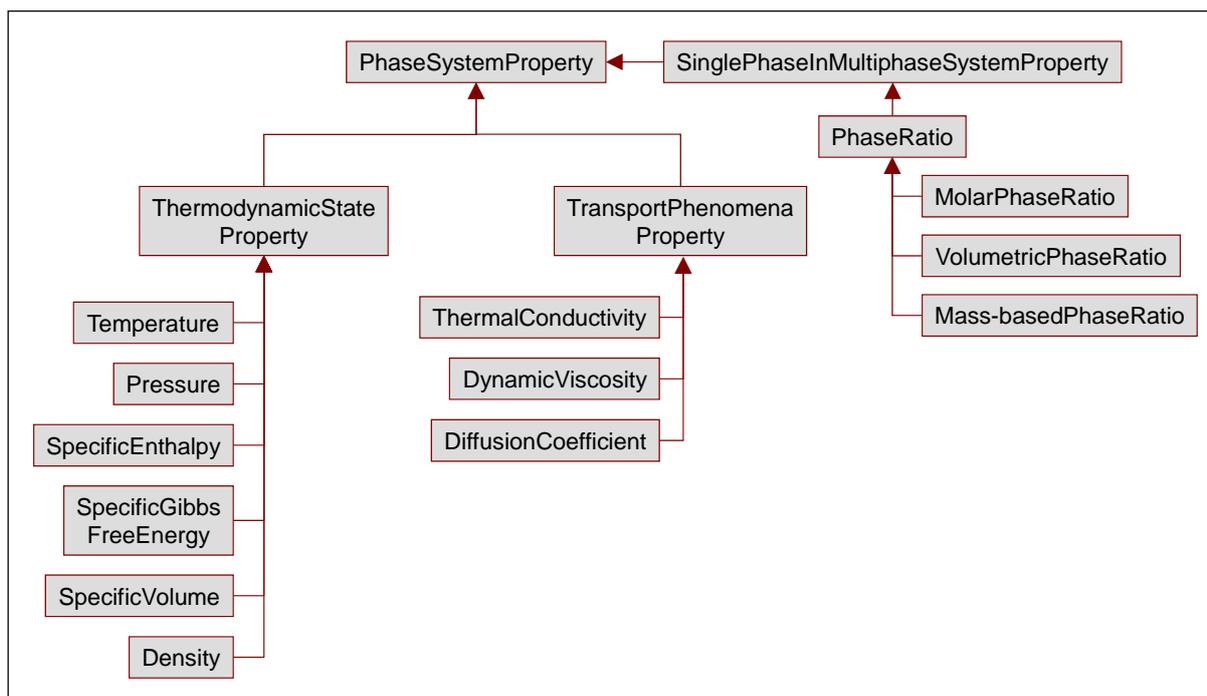
Fig. 28: Phase system properties

Properties associated with a *phase interface* are summarized as *phase interface properties*. Currently, OntoCAPE includes one single *phase interface property*, namely *surface tension* (cf. Fig. 29).
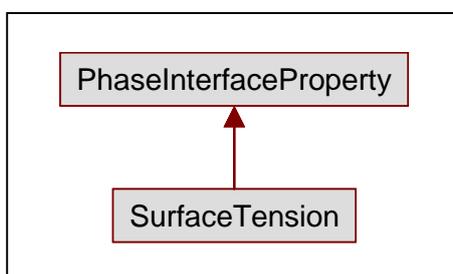

Fig. 29: Phase interface properties

A *phase component* reflects the behavior of a *chemical component* within a *phase system* by means of *phase component properties*, such as *activity coefficient*, *fugacity*, or *partial molar volume*. Thus, all *properties* of individual *chemical components* that appear within certain *physical contexts* are attributed to the *phase component* class. Fig. 30 shows the *phase component properties* that have been defined in OntoCAPE so far.
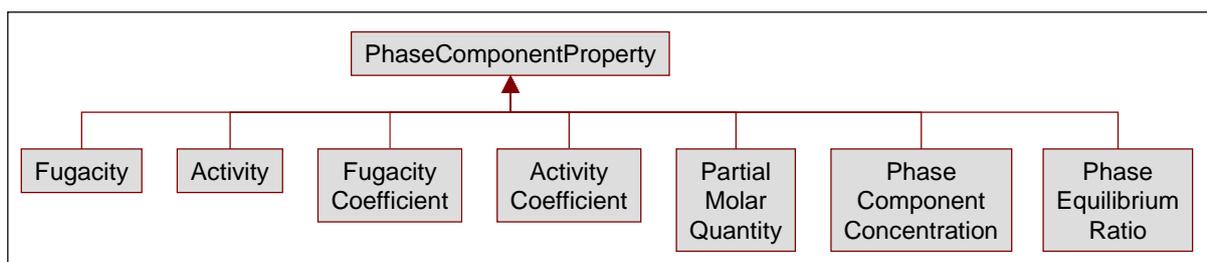

Fig. 30: Phase component properties

Note that *phase equilibrium ratio* (a.k.a. K value) is associated with a *phase component* and two *single phases* involved in a *multiphase system*. However, it is primarily regarded as a property of a *phase component* in this ontology, because it indicates a trait of a *phase component* in the context of phase equilibrium (e.g. the "lightness" in vapor-liquid equilibrium; cf. Smith and Van Ness, 1975). Furthermore, this property is applicable only to a *phase component* which is associated with a *multiphase system*. Additionally this property refers to a first and a second *single phase* to articulate the nature of the equilibrium ratio. For details of other *phase component properties*, we refer to the individual concept definitions below.

The concentration of a *phase component* within a *phase system* can be represented by different concepts (Fig. 31): A *phase component fraction* is a relative concentration, defined as the mass, molar amount, or volume of one *phase component* divided by the mass, molar amount, or volume of all *phase components* in that *phase system*. *Volume-based concentrations* are defined as the mass, molar amount, or volume of a *phase component* of a *phase system* divided by the total volume of that *phase system*. The class *phase component concentration* subsumes the different concentration measures for a *phase component*. The composition of a *phase system* is described through the *composition* concept. A *composition* is a *property set* that comprises the *phase component concentrations* of all *phase components* that constitute the *phase system* (cf. Fig. 31).
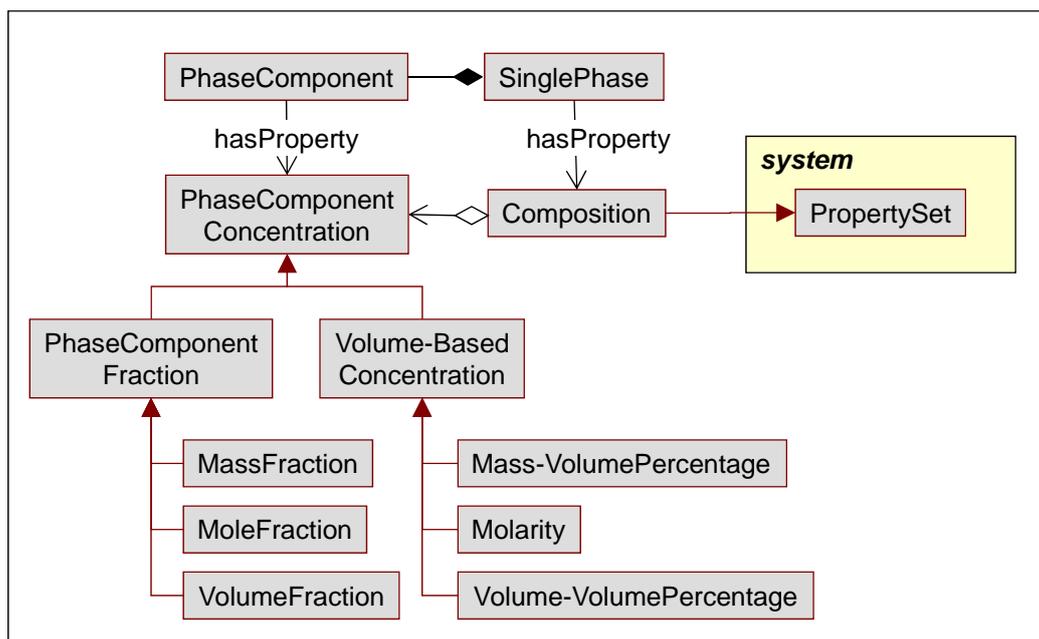

Fig. 31: Composition of a phase system

It is worth noting that, although the *composition* concept introduced above is generally applicable to both *single phases* and *multiphase systems*, in practice the *composition* of a *multiphase system* is usually depicted only by means of *mass fractions* or *mole fractions* of the constituting *phase components*.

3.3. Physical Context

A *phase system* is subject to a certain *physical context*. According to Yang et al. (2003), a *physical context*

"stands for a set of *phase system properties* with known values in order to characterize a certain *phase system*. Different sets of *properties* might be selected here, but in any case the selected *properties* should be sufficient for determining other *properties* of interest of the *phase system*." In OntoCAPE, a *physical context* is modeled as a *property set* (cf. Fig. 32).

A *physical context* comprises a minimum set of *intensive thermodynamic state variables* required to fully characterize a *phase system* in thermodynamic equilibrium. Conventionally, a *physical context* is described by *temperature*, *pressure*, and *phase component fractions*, although other *intensive thermodynamic state variables* may be used alternatively.
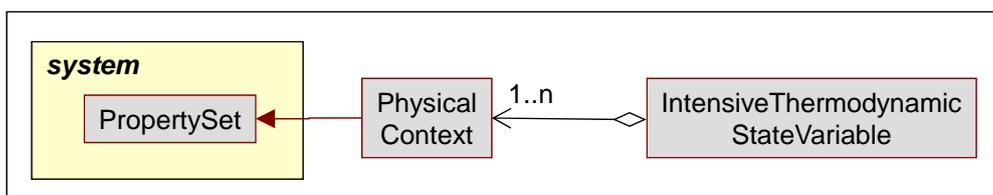
Fig. 32: Physical context

An *intensive thermodynamic state variable* (often simply called "state variable" in the literature) denotes any *intensive property* that can be used to characterize the macroscopic thermodynamic state of a *phase system* (cf. Fig. 33). It subsumes the classes *thermodynamic state property* and *phase component property*. *Temperature*, *pressure*, and *molar fractions* are the most commonly used *intensive thermodynamic state variables*.

The minimum number of *intensive thermodynamic state variables* that are required to determine the thermodynamic equilibrium state of a *phase system* is given by the Gibbs Phase Rule (e.g., Sandler, 1999) and depends on the characteristic of the *phase system*. Moreover, not all *intensive thermodynamic state variables* are independent, as stated by the Gibbs-Duhem Equation (e.g., Sandler, 1999). Thus, the combination of *intensive thermodynamic state variables* appearing in a *physical context* has to be chosen appropriately. In the case of a *multiphase system*, the *physical context* should be composed of the *intensive thermodynamic state variables* of the constituting *single phases*; by contrast, using the *intensive thermodynamic state variables* of the composite *multiphase system* is not recommended for practical use (Sandler, 1999).
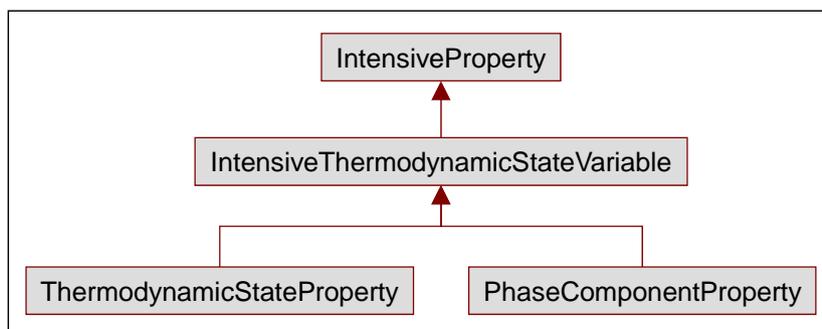
Fig. 33: Intensive thermodynamic state properties

56

3.4. Reactions in Phase Systems

The occurance of a *chemical reaction* (defined in the module *reaction_mechanisms*; cf. Sec. 2.8) in a *phase system* or at a *phase interface* is represented by the hasReaction relation that links the latter two to the former, as shown in Fig. 34. Furthermore, the **phase_system** partial model introduces a specific *phase reaction property*, namely *reaction equilibrium constant*. A *phase system* or *phase interface* may have one or more *reaction equilibrium constants*. Each instance of this property refers to a particular *reversible reaction* which again has been defined in the *reaction_mechanism* module.
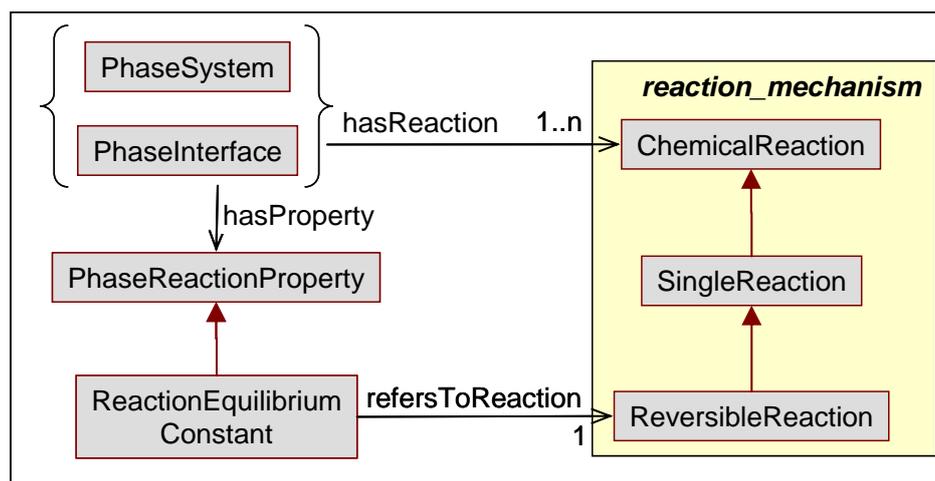

Fig. 34: Phase reactions

Concept definitions

Individual concepts of the module *phase_system* are defined below.

Classes

Activity

Description

The *activity* of some *phase component* is the ratio of the *phase component*'s *fugacity* to the *fugacity* in its standard state.

Relations

- *Activity* is a subclass of *phase component property*.

Activity coefficient

Description

The *activity coefficient* of a *phase component* is the ratio of its *fugacity* in the actual *phase system*, to its *fugacity* in an ideal mixture.

Relations

- *Activity Coefficient* is a subclass of *phase component property*.

Composition

Description

Composition represents the composition of a *phase system* by assembling the concentrations of the different *phase components* that constitute the *phase system*.

Relation

- *Composition* is a subclass of *property set*.
- A *composition* comprises a number of *phase component concentrations*.
- A *composition* can only comprise *phase component concentrations*.

Density

Description

The mass of a *phase system* divided by its volume. It is the reciprocal of *specific volume*.

Relations

- *Density* is a subclass of *thermodynamic state property*.

Diffusion coefficient

Description

Proportionality constant, relating the flux of amount of some *phase component* to its concentration gradient.

Relations

- *Diffusion coefficient* is a subclass of *intra-phase transport property*.

Dynamic viscosity

Description

For a laminar flow of a fluid, the ratio of the shear stress to the *velocity gradient* perpendicular to the plane of shear (McNaught and Wilkinson 1997).

Relations

- *Dynamic viscosity* is a subclass of *intra-phase transport coefficient*.

Fugacity

Description

The fugacity of a *phase component*.

- *Fugacity* is a subclass of *phase component property*.

Fugacity coefficient

Description

Ratio of *fugacity* to the partial pressure of a *phase component*.

Relations

- *Fugacity coefficient* is a subclass of *phase component property*.

Intensive property

Description

An *intensive property* is a *physical quantity*, the *value* of which does not depend on the system size or the amount of material in the system.

Relations

- *Intensive property* is a subclass of *physical quantity*.

Intensive thermodynamic state variable

Description

An *intensive thermodynamic state variable* is an *intensive property* that characterizes the (macroscopic) thermodynamic state of a *phase system*.

Definition

An *intensive thermodynamic state variable* is either a *thermodynamic state property* or a *phase component property*.

Relations

- *Intensive thermodynamic state variable* is a subclass of *intensive property*.

Mass-based phase ratio

Description

The ratio of the mass of a *single phase in multiphase system* divided by the mass of the *multiphase system*.

Relations

- *Mass-based phase ratio* is a subclass of *phase ratio*.

Mass fraction

Description

The mass of a *phase component* divided by the total mass of all *phase components* of a *phase system*.

Relations

- *Mass fraction* is a subclass of *phase component fraction*.

- A *mass fraction* has the dimension **mass_fraction**.

Molarity

Description

The *molarity* (a.k.a. amount concentration) is the molar amount of a *phase component* divided by the volume of the *phase system.*.

Relations

- *Molarity* is a subclass of phase *Volume-Based Concentration*.

Molar phase ratio

Description

The ratio of the (molar) amount of substance of a *single phase in multiphase system* to the (molar) amount of substance of the *multiphase system*.

Relations

- *Molar phase ratio* is a subclass of *phase ratio*.

Mole fraction

Description

The number of moles of a *phase component* divided by the total number of moles of all *phase components* in the *phase system*.

Relations

- *Mole fraction* is a subclass of *phase component fraction*.

- A *mole fraction* has the dimension **molar_fraction**.

Multiphase system

Description

A *multiphase system* is a *phase system* that consists of two or more *single phases*.

Definition

A *multiphase system* is a *phase system* which is composed of some *single phases*.

Relations

- *Multiphase system* is a subclass of *phase system*.

- A *multiphase system* is composed of at least two *single phases*.

- A *multiphase system* can only be composed of *single phases* or *phase interfaces*.

Partial density

<u>Description</u>

Mass of a *phase component* divided by the volume of the *single phase*.

<u>Relations</u>

- *Partial density* is a subclass of phase *Volume-Based Concentration*.

Partial molar enthalpy

<u>Description</u>

The partial derivative of the *specific enthalpy* with respect to the number of moles of one *phase component*.

<u>Relations</u>

- *Partial molar enthalpy* is a subclass of *partial molar quantity*.

Partial molar quantity

<u>Description</u>

A *partial molar quantity* is the partial derivative of the considered molar quantity with respect to the number of moles of one *phase component*. The temperature, pressure, and the number of moles of all other *phase components* are held constant when forming the derivative. The class subsumes all kinds of partial molar quantities, such as *partial molar enthalpy* and *partial molar volume*.

<u>Relations</u>

- *Partial molar quantity* is a subclass of *phase component property*.

Partial molar volume

<u>Description</u>

The partial derivative of the *specific volume* with respect to the number of moles of one *phase component*.

<u>Relations</u>

- *Partial molar volume* is a subclass of *partial molar quantity*.

Phase component

Description

A *phase component* represents the occurrence of a *chemical component* in a *phase system*.

Relations

- *Phase component* is a subclass of *system*.

- A *phase component* is an exclusive subsystem of a *single phase*.

- A *phase component* can only be an exclusive subsystem of a *single phase*.

- A *phase component* represents the occurrence of exactly one *chemical component*.

- A *phase component* can only have *phase component properties*.

- A *phase component* has some *phase component concentration*.

Phase component concentration

Description

Phase component concentration subsumes the different concentration measures for a certain *phase component* within a *single phase*.

Relations

- *Phase component concentration* is a subclass of *phase component property*.

Phase component fraction

Description

In general, the ratio of two *physical quantities* of the same kind, the numerator quantity applying to one particular *phase component* of a *phase system*, and the denominator to the sum of quantities of all *phase components* of that *phase system*. Concretely, a *phase component fraction* subsumes the following concentration measures: *mass fraction*, *volume fraction*, and *mole fraction*.

Relations

- *Phase component fraction* is a subclass of *phase component property*.

Phase component property

Description

A *phase component property* is an *intensive property* that characterizes a *phase component*.

Definition

A *phase component property* is an *intensive property* of some *phase component*.

- *Phase component property* is a subclass of *intensive property*.

- A *phase component property* can only be the property of a *phase component*.

Phase equilibrium ratio

<u>Description</u>

A *phase equilibrium ratio* of a *phase component* in a *multiphase system* is a *phase component property* that denotes the ratio of the *mole fraction* of this *phase component* in one specific *single phase* to that in another specific *single phase*; both *single phases* are part of the *multiphase system*.

<u>Relations</u>

- *Phase equilibrium ratio* is a subclass of *phase component property*.

Phase interface

<u>Description</u>

A *phase interface* represents the interface between two *single phases*.

<u>Relations</u>

- *Phase interface* is a subclass of *connection*.

- A *phase interface* is directly connected to exactly two *single phases*.

- A *phase interface* can only be directly connected to *single phases*.

- A *phase interface* can only be an exclusive subsystem of a *multiphase system*.

- A *phase interface* may have some *phase reactions*.

Phase interface property

<u>Description</u>

A *phase interface property* is an *intensive property* that characterizes the interface between two *single phases*.

<u>Definition</u>

A *phase interface property* is an *intensive property* of some *phase interface*.

<u>Relations</u>

- Phase ratio*interface property* is a subclass of *intensive property*.

Phase Ratio

<u>Description</u>

A *phase ratio* characterizes the proportion of a *single phase* in a *multiphase system* on a mass, molar, or volume basis.

<u>Definition</u>

A *phase ratio* is either a *mass-based phase ratio*, or a *volumetric phase ratio*, or a *molar phase ratio*.

<u>Relations</u>

- *Phase ratio* is a subclass of *single phase in multiphase system property*.

Phase reaction property

<u>Description</u>

A *phase reaction property* is an *intensive property* that is a property of the occurance of a *chemical reaction* in a *phase system* or at a *phase interface*.

<u>Relations</u>

- *Phase reaction property* is a subclass of *intensive property*.

Phase system

<u>Description</u>

A *phase system* represents the macroscopic thermodynamic behavior of *material* in some *physical context*.

<u>Definition</u>

A *phase system* is either a *single phase* or a *multiphase system*.

<u>Relations</u>

- *Phase system* is a subclass of *system*.

- A *phase system* has exactly one *physical context*.

- A *phase system* cannot have more than one *reaction network*.

- A *phase system* may have a number of *phase system properties*.

Phase system property

<u>Description</u>

A *phase system property* is an *intensive property* that characterizes a *phase system*. It can be described or calculated without reference to the shape, size, or amount of a particular occurrence of a material. In the case of calculation, this is consistent with the usage of general-purpose physical property packages,

where such information is not required as the input for the calculation.

Definition

A *phase system property* is an *intensive property* of a *phase system*.

Relations

- *Phase system property* is a subclass of *intensive property*.

- A *phase system property* can only be a property of a *phase system*.

Physical context

Description

A *physical context* of a *phase system* is a set of *intensive thermodynamic state properties* with known *values* which are sufficient for determining other *properties* of interest of the *phase system*. Conventionally, a *physical context* is described by *temperature*, *pressure*, and *molar fractions*, although other *intensive thermodynamic state variables*, such as *specific volume*, *specific enthalpy*, *specific Gibbs free energy*, or *volume-based concentrations* may be used alternatively.

Relations

- *Physical context* is a subclass of *property set*.

- A *physical context* comprises an *intensive thermodynamic state property*.

Pressure

Description

The (total absolute) pressure of a *phase system*.

Relations

- *Pressure* is a subclass of *thermodynamic state property*.

Reaction equilibrium constant

Description

According to the IUPAC Compendium (McNaught and Wilkinson, 1997), the *reaction equilibrium constant* is a *physical quantity* characterizing the chemical equilibrium of a chemical reaction. It is defined by an expression of type $K = \prod_i x_i^{\nu_i}$,

where ν_i is the *stoichiometric coefficient* of a reactant (negative) or product (positive) of the reaction, and x_i stands for a *physical quantity* which can be the equilibrium value either of *pressure*, *fugacity*, *molarity*, *molar fraction*, or *activity*. Depending on the chosen quantity, one obtains one of the following types of *reaction equilibrium constant*: pressure based, fugacity based, concentration based, amount fraction based, molarity based, relative activity based, or standard equilibrium constant, respectively. These different

types can be modeled as subclasses of the *reaction equilibrium constant*.

Relations

- *Reaction equilibrium constant* is a *phase reaction property*.
- *Reaction equilibrium constant* refers to a *reversible reaction*.

Single phase

Description

A *single phase* is a finite volume of material within which the physical properties are uniformly constant, i.e., do not experience any abrupt change in passing from one point in the volume to another.

Relations

- *Single phase* is a subclass of *phase system*.
- A *single phase* is composed of some *phase components*.
- A *single phase* can only be composed of *phase components*.
- A *phase system* has exactly one *composition*.
- A *single phase* has exactly one *state of aggregation*.
- A *single phase* may have some *phase reactions*.

Single phase in multiphase system

Description

A *single phase* that is part of a *mulitphase system*.

Definition

A *single phase* which is an exclusive subsystem of a *multiphase system*.

Relations

- *Single phase in multiphase system* is a subclass of *single phase*.
- A *single phase multiphase system* is an exclusive subsystem of a *multiphase system*.
- A *single phase multiphase system* can only be an exclusive subsystem of a *multiphase system*.
- A *single phase multiphase system* may have some *phase ratio* to characterize its fraction of the overall *multiphase system*.

Single phase in multiphase system property

Description

Class used for grouping the *properties* of a *single phase in multiphase system*.

<u>Definition</u>

The *property* of a *single phase in multiphase system*.

<u>Relations</u>

- *Single phase in multiphase system property* is a subclass of *phase system property*.

- A *single phase in multiphase system property* can only be a property of a *single phase in multiphase system*.

Specific enthalpy

<u>Description</u>

Specific enthalpy denotes enthalpy per unit mass (i.e., the enthalpy of a *phase system* divided by its mass). It is defined by the equation $h = u + pv$, where u represents the specific internal energy, p. the *pressure*, and v the *specific volume* of the *phase system*.

<u>Relations</u>

- Specific enthalpy is a subclass of *thermodynamic state property*.

Specific Gibbs free energy

<u>Description</u>

Specific Gibbs free energy denotes the Gibbs free energy per unit mass (i.e., the Gibbs free energy of a *phase system* divided by its mass). It is defined by the equation $g = u + p*v - T*s$, where u represents the specific internal energy, p the *Pressure*, v the *specific volume*, T the *temperature*, and s the specific entropy of the *phase system*.

<u>Relations</u>

- Specific *Gibbs free energy* is a subclass of *thermodynamic state property*.

Specific volume

<u>Description</u>

The volume of a *phase system* divided by its mass. It is the reciprocal of *density*.

<u>Relations</u>

- *Specific volume* is a subclass of *thermodynamic state property*.

State of aggregation

<u>Description</u>

The *state of aggregation* (a.k.a. state of matter) describes the physical state of a *single phase*; solid, liquid, and gaseous are predefined instances of *state of aggregation*.

Relations

- *State of aggregation* is a subclass of *extensible value set*.

Surface tension

Description

Work required to increase a surface area divided by that area. When two phases are studied it is often called interfacial tension (McNaught and Wilkinson, 1997).

Relations

- *Surface tension* is a subclass of *phase interface property*.

Temperature

Description

The temperature of a *phase system*.

Relations

- *Temperature* is a subclass of *Thermodynamic state property*.

Thermal conductivity

Description

The *thermal conductivity* λ is the coefficient relating the heat flux q to the temperature gradient ∇T:

$q = -\lambda \nabla T$.

Relations

- *Thermal conductivity* is a subclass of *intra-phase transport coefficient*.

Thermodynamic state property

Description

A *thermodynamic state property* is a *phase system property* that can serve as an *intensive thermodynamic state variable* (i.e., characterize the thermodynamic state of a *phase system*).

Relations

- *Thermodynamic state property* is a subclass of *phase system property*.

Transport phenomena property

Description

A *transport phenomena property* is a *phase system property* that subsumes *dynamic viscosity*, *heat conductivity*, and *diffusion coefficient*.

Relations

- *Transport phenomena property* is a subclass of *phase system property*.

Type of solid

Description

Type of solid allows to further characterize the solid state of matter, according to the CAPE-OPEN Open Interface Specification "Thermodynamic and Physical properties" (Drewitz et al. 2006).

Definition

Type of solid is an enumerated class that can take one of the following instance values: HydrateI, HydrateII, HydrateH, PureSolid, SolidSolution.

Relations

- *Type of solid* is a subclass of *state of aggregation VR*.

Volume-based concentration

Description

Volume-based concentration denotes the concentration of a certain *phase component* in a *phase system* with respect to the volume of the *phase system*.

Definition

A *volume-based concentration* can be one of the following: *mass-volume percentage*, *molarity*, or *volume-volume percentage*.

Relations

- *Volume-based concentration* is a subclass of *phase component concentration*.

Volume fraction

Description

The volume of a *phase component* divided by the sum of volumes of all *phase components* of the *phase system* prior to mixing; for ideal mixtures, this equals to the *volume-volume percentage*.

Relations

- *Mole fraction* is a subclass of *phase component fraction*.
- A *mole fraction* has the dimension molar_fraction.

Volume-volume percentage

Description

Volume of a *phase component*, divided by the total volume of the *phase system*. For ideal mixtures, this

is the same as the *volume fraction*.

<u>Relations</u>

- *Volume-volume percentage* is a subclass of *volume- based concentration*.

Volumetric phase ratio

<u>Description</u>

The ratio of the volume of a *single phase in multiphase system* to the volume of the *multiphase system*.

<u>Relations</u>

- *Volumetric phase ratio* is a subclass of *phase ratio*.

Relations

has_composition

<u>Description</u>

Auxiliary relation, introduced as workaround for a qualified cardinality restriction (QCR).

<u>Characteristics</u>

- Specialization of hasProperty

- Domain: *Single phase system*

- Range: *Composition*

has_physical_context

<u>Description</u>

Auxiliary relation, introduced as workaround for a qualified cardinality restriction (QCR).

<u>Characteristics</u>

- Specialization of hasProperty

- Domain: *Phase system*

- Range: *Physical context*

hasReaction

<u>Description</u>

The relation hasReaction links a *chemical reaction* to the *phase system* or *phase interface* where it takes place.

<u>Characteristics</u>

-   Specialization of isDirectlyRelatedTo

-   Domain: *Single phase* or *phase interface*

-   Range: *Phase reaction*

hasStateOfAggregation

Description

The relation hasStateOfAggregation indicates the *state of aggregation* of a *single phase*.

Characteristics

-   Specialization of hasCharacteristics

-   Domain: *Single phase*

-   Range: *State of Aggregation*

-   Functional

representsOccurrenceOf

Description

The relation representsOccurenceOf establishes the relationship between a *phase component* and the corresponding *chemical component*, or between a *phase reaction* and the corresponding *chemical reaction*.

Characteristics

-   Specialization of isDirectlyRelatedTo

-   Domain: *Phase component* or *phase reaction*

-   Range: *Chemical component* or *chemical reaction*

-   Functional

refersToReaction

Description

The relation refersToReaction links a *reaction equilibrium constant* to a *reversible reaction*.

Characteristics

-   Specialization of isDirectlyRelatedTo

-   Domain: *Reaction equilibrium constant*

-   Range: *Reversible reaction*

refersToFirstSinglePhase

The relation refersToFirstSinglePhase indicates the first one of the two *single phases* to which a *phase*

equilibrium ratio of a *phase component* refers.

refersToSecondSinglePhase

The relation refersToSecondSinglePhase indicates the second one of the two *single phases* to which a *phase equilibrium ratio* of a *phase component* refers.

References

Bayer B (2003) *Conceptual Information Modeling for Computer Aided Support of Chemical Process Design*. Fortschritt-Berichte VDI: Reihe 3, Nr. 787. VDI-Verlag, Düsseldorf.

CAS (2007) CAS Registry Overview. Website, online available at http://www. cas.org/EO/regsys.html. Accessed February 2007.

Dietz A (1995) Yet another representation of molecular structure? *J. Chem. Inf. Comput. Sci.* **35**(5):787-802.

Drewitz W, Szczepanski R, Pinõl D, Banks P, Halloran M, van Baten J, Pons M, eds. (2006) Thermodynamic and physical properties, version 1.1. CAPE-OPEN Interface Standards Specification. Online available at http://www. colan.org/index.html. Accessed March 2007.

EBI – European Bioinformatics Institute (2008) Chemical Entities of Biological Interest (ChEBI). Online available at http://www.ebi.ac.uk/chebi/. Acces-sed June 2008.

Gold V, Loening KL, McNaught AD, Sehmi P (1987) *Compendium of Chemical Terminology*. Blackwell, Oxford.

Gutsche B (1986) Phase equilibria in oleochemical industry – application of continuous thermodynamics. *Fluid Phase Equilib.* **30**:65-70.

Hariu O, Sage R (1969) Crude split figured by computer. *Hydrocarbon Process.*, *Int. Ed.* **48** (4):143-148.

Kumar A, Gupta R (1998) *Fundamentals of Polymers*. McGraw-Hill, New York.

Levenspiel O (1999). *Chemical Reaction Engineering*. John Wiley & Sons, Inc., New York.

Linstrom PJ, Mallard WG, eds. (2005) NIST Chemistry WebBook. NIST Standard Reference Database Number 69, June 2005, National Institute of Standards and Technology, Gaithersburg, MD. Online available at http://web-book.nist.gov.

McNaught AD, Wilkinson A, eds. (1997) Compendium of Chemical Terminology, 2nd Edition. Blackwell Science, Oxford, UK. Electronic version online available at http://goldbook.iupac.org/.

Modell M, Reid RC (1983) Thermodynamics and its Applications. Second ed., Prentice-Hall, Englewood Cliffs.

Morbach J, Bayer B, Wiesner A, Yang A, Marquardt W (2008a) *OntoCAPE 2.0 – The Upper Level*. Technical Report (LPT-2008-25), Lehrstuhl für Prozesstechnik, RWTH Aachen University. Online available at http://www.avt.rwth-aachen.de/AVT/index.php?id=541&L=0&Nummer= LPT-2008-25.

Morbach J, Yang A, Marquardt W (2008b) *OntoCAPE 2.0 – Mathematical Models*. Technical Report (LPT-2008-28), Lehrstuhl für Prozesstechnik, RWTH Aachen University. Online available at

http://www.avt.rwth-aachen.de/AVT/index.php?id=541&L=0&Nummer=LPT-2008-28.

Moss GP, ed. (1996) Basic terminology of stereochemistry. IUPAC Recommendations 1996, *Pure Appl. Chem.* **68**:2193-2222.

Moss GP, Smith PAS, Tavernier D, ed. (1995) Glossary of class names of organic compounds and reactive intermediates based on structure. IUPAC Recommendations 1995, *Pure Appl. Chem.* **67**:1307-1375.

Müller P, ed. (1994) Glossary of terms used in physical organic chemistry. IUPAC Recommendations 1994, *Pure Appl. Chem.* **66**:1077-1184.

OLS – Ontology Lookup Service (2006) ChEBI Ontology Browser. Online available at http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName= CHE BI. Accessed February 2007.

Sandler SI (1999) *Chemical and Engineering Thermodynamics*. Third ed., John Wiley, New York.

Schummer J (1998) The chemical core of chemistry I: a conceptual approach. HYLE – *Int. J. Philosophy Chem.* **4**(2):129-162.

Smith EG (1968) The Wiswesser Line-Formula Chemical Notation. McGraw-Hill, New York.

Smith JM, Van Ness HC (1975) Introduction to Chemical Engineering Thermodynamics, 3rd Edition. McGraw-Hill, New York.

Stein SE, Heller SR, Tchekhovski D (2003) An open standard for chemical structure representation – the IUPAC Chemical Identifier. In: *Proceedings of the 2003 Nimes International Chemical Information Conference*:131-143.

Weininger D (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**:31-36.

Wiesner A, Morbach J, Bayer B, Yang A, Marquardt W (2008a) *OntoCAPE 2.0 – Chemical Process System*. Technical Report (LPT-2008-29), Lehrstuhl für Prozesstechnik, RWTH Aachen University. Online available at http://www.avt.rwth-aachen.de/AVT/index.php?id=541&L=0 &Nummer =LPT-2008-29.

Appendix A   Character replacements for generating OWL identifiers

The ontology module *chemical_species* (cf. Sec. 2.4) contains data about *chemical species*, which stems from the NIST Chemistry WebBook (Linstrom & Mallard, 2005). The 'species names' given in the NIST Chemistry WebBook are used as identifiers for the instances of *chemical species*. In order to comply with the OWL syntax, certain characters had to be replaced by character symbols. These replacements are given below.

' $\Rightarrow$ _APOS_

, $\Rightarrow$ _COMMA_

: $\Rightarrow$ _COLON_

* $\Rightarrow$ _STAR_

% $\Rightarrow$ _PER_

= $\Rightarrow$ _EQ_

+ $\Rightarrow$ _PLUS_

($\Rightarrow$ _LP_

) $\Rightarrow$ _RP_

[$\Rightarrow$ _LSB_

] $\Rightarrow$ _RSB_

{ $\Rightarrow$ _LB_

} $\Rightarrow$ _RB_

($\Rightarrow$ _LP_

SPACE $\Rightarrow$ __

As the NIST WebBook is represented in the HTML format, the following HTML predefined entities for special characters have also been replaced:

< $\Rightarrow$ _LT_

> $\Rightarrow$ _GT_

ð $\Rightarrow$ _ETH_

" $\Rightarrow$ _QUOT_

& $\Rightarrow$ _AND_

·	⇒	.
á	⇒	_a_ACUTE_
±	⇒	_PLUSMN_

Moreover, the following HTML predefined entities for Greek characters have been replaced:

«BETA»	⇒	_BETA_CAP_
«DELTA»	⇒	_DELTA_CAP_
«OMEGA»	⇒	_OMEGA_CAP_
«XI»	⇒	_XI_CAP_
«alpha»	⇒	_ALPHA_
«beta»	⇒	_BETA_
«gamma»	⇒	_GAMMA_
«delta»	⇒	_DELTA_
«epsilon»	⇒	_EPSILON_
«eta»	⇒	_ETA_
«lambda»	⇒	_LAMBDA_
«mu»	⇒	_MU_
«pi»	⇒	_PI_
«rho»	⇒	_RHO_
«tau»	⇒	_TAU_
«omega»	⇒	_OMEGA_
«omicron»	⇒	_OMICRON_
«xi»	⇒	_XI_
«psi»	⇒	_PSI_
«equiv»	⇒	_EQUIV_

The below decimal entities sometimes used in place of Greek HTML characters are replaced as follows:

Β ⇒ _BETA_CAP_

Δ ⇒ _DELTA_CAP_

α ⇒ _ALPHA_

β ⇒ _BETA_

γ ⇒ _GAMMA_

δ ⇒ _DELTA_

ε ⇒ _EPSILON_

η ⇒ _ETA_

λ ⇒ _LAMBDA_

μ ⇒ _MU_

π ⇒ _PI_

ρ ⇒ _RHO_

τ ⇒ _TAU_

ω ⇒ _OMEGA_

Finally, HTML tags (e.g., those for indicating subscript and superscript) have been replaced.

<sub>   ⇒   _sub_

</sub> ⇒   _subc_

<sup>   ⇒   _sup_

</sup> ⇒   _supc_

In addition to the above replacements, an underscore character must be inserted as the first character if the name of the species starts with a digit.

Appendix B   Documentation Format

Classes

Classes are characterized by the following categories:

<u>Description</u>: A lexical description of the class, for example "A chemical reactor is an apparatus for holding substances that are undergoing a chemical reaction." The description explains the meaning of the class to the user.

<u>Definition</u>: Unlike a description, a definition can be transcribed into a formal ontology language, where it establishes the set of necessary and sufficient conditions from which the membership of an ontological concept (class or individual) to the class can be inferred. Classes for which such a definition can not be indicated are called primitive classes.

<u>Relations</u>: The following characteristics are indicated, if existent:

- *Specialization*. A list of parent classes from which the current class is derived via specialization.
- *Disjointness*. A list of classes which are disjoint with the present class. Disjointness between classes means that an instance of the first class cannot simultaneously be an instance of the second class.
- *Restrictions*. Restrictions of binary relations (or attributes) specify the existence of a relation (or attribute) as well as its cardinality and value range with respect to the current class.

<u>Usage</u>: Some recommendations for the use of the class may be given if such advice is required.

Relations

Binary relations are characterized by the following categories:

<u>Description</u>: Similar to that of classes mentioned above.

<u>Characteristics</u>: The following characteristics are listed, if existent:

- *Specialization*. A listing of the relations from which the relation is derived via specialization.
- *Domain*. The domain of the relation.
- *Range*. The value range of the relation.
- *Inverse*. The inverse of a relation.
- Further characteristics, such as if the relation is *transitive*, *symmetric*, or (*inverse*) *functional*.

<u>Usage</u>: As above.

Attributes

Attributes are characterized by the following categories:

Description: As above.

Characteristics: The following characteristics are listed, if existent:

- *Specialization*. A listing of the attributes from which the attribute is derived via specialization.

- *Domain*. The domain of the attribute.

- *Range* or *datatype*. The value range of the attribute, which is usually indicated by referring to a built-in XML Schema Datatype (Biron et al., 2004).

- Further characteristics, such as if the attribute is *functional*.

Usage: As above.

Individuals

Predefined individuals are characterized by the following categories:

Description: As above.

Characteristics: The following characteristics are indicated, if existent:

- *Instance of*. The classes from which the individual is instantiated.

- *Different from*. A list of individual which are explicitly declared to be different from the present individual.

- *Relations*. Instances of binary relations the individual is involved in.

- *Attributes*. Attribute values of the individual.

Usage: As above.

Notation Conventions

Classes and relations of the Meta Model are named according to the CamelCase[8] naming convention: UpperCamelCase notation is used to denote identifiers of classes, while relation identifiers are represented in lowerCamelCase notation. No particular naming convention is followed for identifiers of individuals (i.e., instances of classes).

In this document, class identifiers are highlighted by *italicized sans-serif font*; for better readability, the UpperCamelCase notation is not applied in the text, but the individual words that constitute the class identifiers are written separately and in lowercase (e.g., *class identifier*). If relations are explicitly referred to in the text, they are written in lowerCamelCase notation and are additionally highlighted by

---

[8] CamelCase is the practice of writing compound words joined without spaces; each word is capitalized within the compound. While the UpperCamelCase notation also capitalizes the initial letter of the compound, the lowerCamelCase notation leaves the first letter in lowercase.

sans-serif font. Individuals are accentuated by bold sans-serif font. Partial models are denoted **bold serif font**, *italicized serif font* refers to ontology modules.

In figures, a graphical notation in the style of UML class diagrams is used; the basic elements are depicted in Fig. 35. Grey shaded boxes represent *classes*, white boxes represent *individuals*. *Attributes* are denoted by grey shaded boxes with dashed boundary lines, *attribute values* by white boxes with dashed boundary lines. *Specialization* is depicted through a solid line with a solid arrowhead that points from the subclass to the superclass. A dashed line with an open arrowhead denotes *instantiation*. *Binary relations* are depicted though solid lines. Three basic relation types are distinguished: a line with one open arrowhead represents a *unidirectional* relation; a line with two open arrowheads represents a *symmetric* relation; a line without any arrowheads represents a bidirectional relation[9]. Finally, graphic elements for two special types of relation are introduced: an *aggregation* relation is depicted through a line with a white diamond-shaped arrowhead pointing towards the aggregate class. Similarly, a black diamond-shaped arrowhead indicates a *composition* relation.
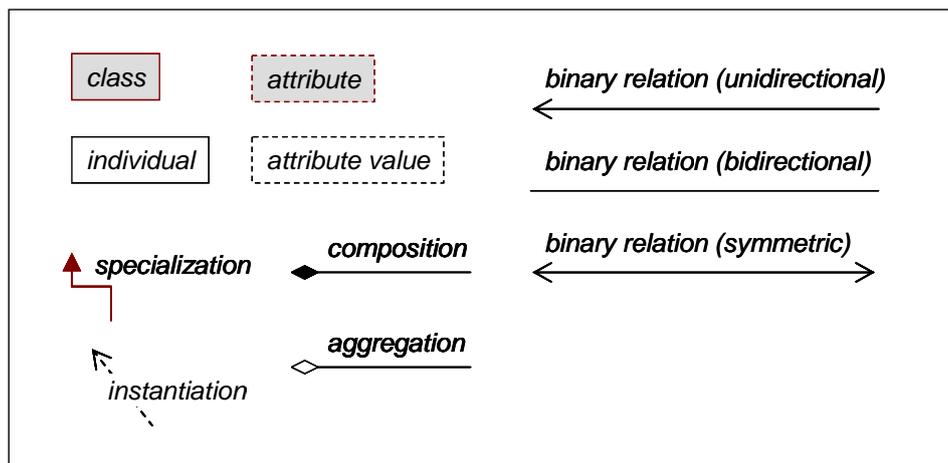
Fig. 35: Basic elements of graphical notation

---

[9] In OWL, a bidirectional relation is modeled through a unidirectional relation and its inverse.

Index of Concepts