Unified Code for Units of Measure (UCUM)

Gunther Schadow, MD, PhD,
Investigator, Regenstrief Institute
Associate Professor, Indiana University School of Informatics
Co-chair, Orders and Observations, Health Level-7
President, Pragmatic Data LLC

Abstract: *The Unified Code for Units of Measure (UCUM) provides human-friendly codes for all units of measures with precise semantics to facilitate unambiguous and computable communication between computer systems used in science, engineering and business world wide. UCUM is inspired by earlier standards (ISO 2955-1983, ANSI X3.50-1986) which it expands and corrects (resolving various ambiguities). The semantics of units is based on the intuition applied by most physicists when computing with quantities and units (but chemists and medical people may have to enhance their intuition first). It represents the meaning of equivalence, commensurability, conversion, base and derived units, including special units which require arbitrary conversion functions (e.g., logarithm). UCUM's formal semantics is defined algebraically, which leads to a very compact representation and efficient (constant time) reasoning. However, it is quite different from symbolic knowledge representation methods that many ontologists are familiar with. While symbolic models in UML are useful for discussing the design of the UCUM implementation, they do not replace the elegance and efficiency of the algebraic definition. This supports the conclusion that units of measure are essentially quantitative phenomena that require a focus on quantitative methods for their definition. UCUM does not, however, attempt to define base units in any formal way but uses those as primitives and refers to appropriate standards bodies (ISO, BIPM) for their definitions. UCUM has been adopted by many standards organization world wide in and outside of the medical domain. While the actual maintenance of the core code system is minimal, defining organization and governance are becoming more important. The key challenge in content is to deal with procedure defined (arbitrary) units that are common in biomedical sciences.*

UCUM - Unified Code for Units of Measure

- Scope
  - All units of measures being contemporarily used in international science, engineering, and business.
- Purpose
  - To facilitate unambiguous semantically defined and computable communication of physical quantities with units between computer systems.
- Inspired by ISO 2955-1983, ANSI X3.50-1986, and HL7 ISO+.
  - Predecessors (now obsolete) had many defects
    - "a" for "year" and "are"
    - "cd" for candela and centi-day
    - ounce: Avoirdupois? Troy? Apothecaries'?
    - many missing units, e.g., mm[Hg]
- Specifies all units with computable semantics
  - Automatic conversion
  - No need for large tables

UCUM is Easy and Precise

- Human-friendly Codes with Precise Semantics
  - Write units as you would when typing a scientific journal article draft in a plain-text email
    - ug/dL, not mcg/deciliters
    - g, kg, not gm, gms, kgs, etc.
    - cm3 not cc
  - For standard units, you will be correct most of the time.
  - Only beware of customary units, jargon, non-units, and arbitrary units.

http://unitsofmeasure.org

UCUM

Search

logged in as gschadow | Logout | Help/Guide | About Trac | Preferences

Wiki | Timeline | Roadmap | View Tickets | New Ticket | Search | Admin

Start Page | Index | History | Last Change

Shortcut to the Specification

The Unified Code for Units of Measure

Gunther Schadow, Clement J. McDonald

What is it?

The Unified Code for Units of Measure is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. The purpose is to facilitate unambiguous electronic communication of quantities together with their units. The focus is on electronic communication, as opposed to communication between humans. A typical application of The Unified Code for Units of Measure are electronic data interchange (EDI) protocols, but there is nothing that prevents it from being used in other types of machine communication. How does it relate?

The Unified Code for Units of Measure is inspired by and heavily based on ISO 2955-1983, ANSI X3.50-1986, and HL7's extensions called ISO+. *The respective ISO and ANSI standards are both entitled Representation of [...] units in systems with limited character sets where ISO 2955 refers to SI and other units provided by ISO 1000-1981, while ANSI X3.50 extends ISO 2955 to include U.S. customary units. Because these standards carry the restriction of limited character sets* in their names they seem to be of less value today where graphical user interface and laser printers are in wide-spread use, which is why the european standard ENV 12435 in its clause 7.3 declares ISO 2955 obsolete.

TABLE OF CONTENTS

APPENDICES

Syntax

- Unit atom: g, m, min, [in_i], m[Hg]
- Simple unit: prefix * atom
  - kg, mm, mm[Hg], but ***<u>not</u>*** mmin, m[in_i]
  - Prefixes only for "metric" units
- Unit terms
  - Exponents: m2, cm3
  - Multiplication: kg.m
  - Division: m/s
  - Combination: kg.m/s2
  - Numeric factors, Parentheses: mg/(12.h)
- Annotations:
  - kg{potatoes}, mg{creat}
  - do not change the meaning

Semantics – The Intuition

- Unit is a product of integer-powers of base units:
  - 1 N = 1 kg.m/s2 = 10^{3} m^{1} s^{-2} g^{1}
  - 1 dyn/s/cm5 = 10^{8} m^{-4} s^{-1} g^{1}
  - 1 mmol/L = 10^{-3} m^{-3} mol^{1}
  - 1 U/L = 1/60 10^{-3} m^{-3} s^{-1} mol^{1}
- Represented as:
  - (Factor) x [vector of exponents]
    - E.g., 1/60 10^{-3} x [-3, -1, 0, 1, 0, 0, 0]

Semantics - Examples

- 1 N = 1 kg.m/s2 = 10^3 m**^{1}** s**$^{-2}$**g**^{1}**
  - 10^3 x **[1**, **-2**, **1**, 0, 0, 0, 0**]**
- 1 dyn/s/cm5 = 10^8 m**$^{-4}$**s**$^{-1}$**g**^{1}**
  - 10^8 x **[-4**, **-1**, **1**, 0, 0, 0, 0**]**
- 1 mmol/L = 10^{-3}m**$^{-3}$** mol**^{1}**
  - 10^{-3} x **[-3**, 0, 0, **1**, 0, 0, 0**]**
- 1 U/L = 1/60 10^{-3}m**$^{-3}$**s**$^{-1}$** mol**^{1}**
  - 1/60 10^{-3} x **[-3**, **-1**, 0, **1**, 0, 0, 0**]**

Semantics – The Formalities

§16 preliminaries

- **1** The semantics of *UCUM* is defined by the algebraic operations of multiplication, division and exponentiation between units, by the equivalence relations of equality and commensurability of units, and by the multiplication of a unit with a scalar.
- **2** Every expression in *UCUM* is mapped to one and only one semantic element. But every semantic element may have more than one valid representant in *UCUM*.
- **3** The set of expressions in *UCUM* is infinite.

§17 equivalence relations

- 1 The set of expressions in UCUM has 2 binary, symmetric, reflexive, and transitive relations:

  (1) “equals” = and

  (2) “is commensurable with” ~.

  All expressions that are equal are also commensurable but not all commensurable expressions are equal.

§18 algebra of units

- **1** The equivalence classes generated by the equality relation = are called *units*.
- **2** The set of units U has a binary multiplication operator $\cdot$ that is associative and commutative and has the neutral element **1** (called "*the unity*"). For each unit $\mathbf{u} \in U$ there is an inverse unit $\mathbf{u}^{-1}$ such that $\mathbf{u} \cdot \mathbf{u}^{-1} = \mathbf{1}$.
- **3** The division operation $\mathbf{u} / \mathbf{v}$ is defined as $\mathbf{u} \cdot \mathbf{v}^{-1}$.
- **4** The exponentiation operation with integer exponents n is defined as $\mathbf{u}^n = \mathbf{\Pi}_1{}^n\, \mathbf{u}$.
- **5** The product $\mathbf{u}' = r\mathbf{u}$ of a real number scalar r with the unit $\mathbf{u}$ is also a unit, where $\mathbf{u}' \sim \mathbf{u}$.

§19 dimension and magnitude

- **1** The equivalence classes generated by the commensurability relation ~ are called *dimensions*.

  The set *D* of dimensions is infinite in principle, but only a finite subset of dimensions are used in practice. Thus, implementations of *UCUM* need not be able to represent the infinite set of dimensions.
- **2** Two commensurable units that are not equal differ only by their magnitude.
- **3** The quotient **u** / **v** of any two commensurable units **u** ~ **v** is of the same dimension as the unity (**u** / **v** ~ **1**). This quotient is also equal to *the unity* **1** multiplied with a scalar *r*: **u** / **v** = *r***1**, where *r* is called the *relative magnitude* of **u** regarding **v**.

§20 base units

- **1** Any system of units is constructed from a finite set **B** of mutually independent base units

  $$\mathbf{B} = \{ \mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n \},$$

  on which any other unit $\mathbf{u} \in U$ is defined as

  $$\mathbf{u} = r_1\mathbf{b}_1^{u1} \cdot r_2\mathbf{b}_2^{u2} \cdot ... \cdot r_n\mathbf{b}_n^{un},$$

  where $r = r_1 \cdot r_2 \cdots r_n$ is called the *magnitude* of the unit **u** regarding **B**.
- **2** With respect to a basis **B** every unit can thus be represented as a pair $(r, \hat{u})$ of magnitude r and dimension $\hat{u} = (u_1, u_2, ..., u_n)$.
- **3** Two sets of base units are equivalent if there is an isomorphism between the sets of units that they generate.

Base Units

Kind of Quantity	Variable	Unit	Vector $\vec{b}_i$
Length	s	1 m	(1, 0, 0, 0, 0, 0, 0)
Time	t	1 s	(0, 1, 0, 0, 0, 0, 0)
Mass	m	1 g	(0, 0, 1, 0, 0, 0, 0)
Charge	Q	1°C	(0, 0, 0, 1, 0, 0, 0)
Temperature	T	1 K	(0, 0, 0, 0, 1, 0, 0)
Luminous intensity	I_v	1 cd	(0, 0, 0, 0, 0, 1, 0)
Angle	ς	1 rad	(0, 0, 0, 0, 0, 0, 1)

- base units are arbitrary and conventional
- linear transformation can convert between isomorphic systems of units

Derived Units

Kind of Quantity	Definition*	Unit	u	Factor ν	Vector $\vec{u}$
1	1	*The unity*	1	1	(0, 0, 0, 0, 0, 0, 0)
Area	$A = s_1 \cdot s_2$		1 m^2	1	(2, 0, 0, 0, 0, 0, 0)
Volume	$V = A \cdot s$	liter	1 L	10^{-3}	(3, 0, 0, 0, 0, 0, 0)
Velocity	$v = s/t$		1 m/s	1	(1, −1, 0, 0, 0, 0, 0)
Angular velocity	$\omega = s/t$		1 rad/s	1	(0, −1, 0, 0, 0, 0, 1)
Volume current	$\dot{V} = V/t$		1 L/min	6×10^{-2}	(3, −1, 0, 0, 0, 0, 0)
Acceleration	$a = v/t$		1 m/s^2	1	(1, −2, 0, 0, 0, 0, 0)
Force	$F = m \cdot a$	newton	1 N	10^3	(1, −2, 1, 0, 0, 0, 0)
Work	$W = F \cdot s$	joule	1 J	10^3	(2, −2, 1, 0, 0, 0, 0)
Moment of force	$M = F \cdot s$		1 Nm	10^3	(2, −2, 1, 0, 0, 0, 0)
Power	$P = W/t$	watt	1 W	10^3	(2, −3, 1, 0, 0, 0, 0)
Electric current	$I = Q/t$	ampere	1 A	1	(0, −1, 0, 1, 0, 0, 0)
Electric potential	$U = WQ$	volt	1 V	1	(2, −2, 1, −1, 0, 0, 0)

- Any unit can be derived from the base by means of the algebra defined above.

§21 special units

- **1** Units that imply a measurement on a scale other than a ratio scale (e.g., interval scale, logarithmic scale) do not represent proper units as elements of the group $(U, \cdot)$ but are called *special units*. The set of special units is denoted S, where $S \cap U = \{\}$.
- **2** A special unit $\mathbf{s} \in S$ is defined as the triple $(\mathbf{u}, f_{\mathbf{s}}, f_{\mathbf{s}}^{-1})$ where $\mathbf{u} \in U$ is the "corresponding" proper unit of **s** and where $f_{\mathbf{s}}$ and $f_{\mathbf{s}}^{-1}$ are mutually inverse real functions, applied as follows: let $r_{\mathbf{s}}$ be the numeric measurement value expressed in the special unit **s** and let m be the corresponding dimensioned quantity (the measurement with proper unit **u**.) Now, $r\mathbf{s} = f_{\mathbf{s}}(m/\mathbf{u})$ converts the measurement to the special unit and $m = f_{\mathbf{s}}^{-1}(r\mathbf{s}) \times \mathbf{u}$ does the inverse.
- **3** Although not elements of U, special units are said to be "of the same dimension" or "commensurable with" their corresponding proper unit **u** and the class of units commensurable with **u**. This can be expressed by means of a binary, symmetric, transitive and reflexive relation ≈ on $U \in S$.

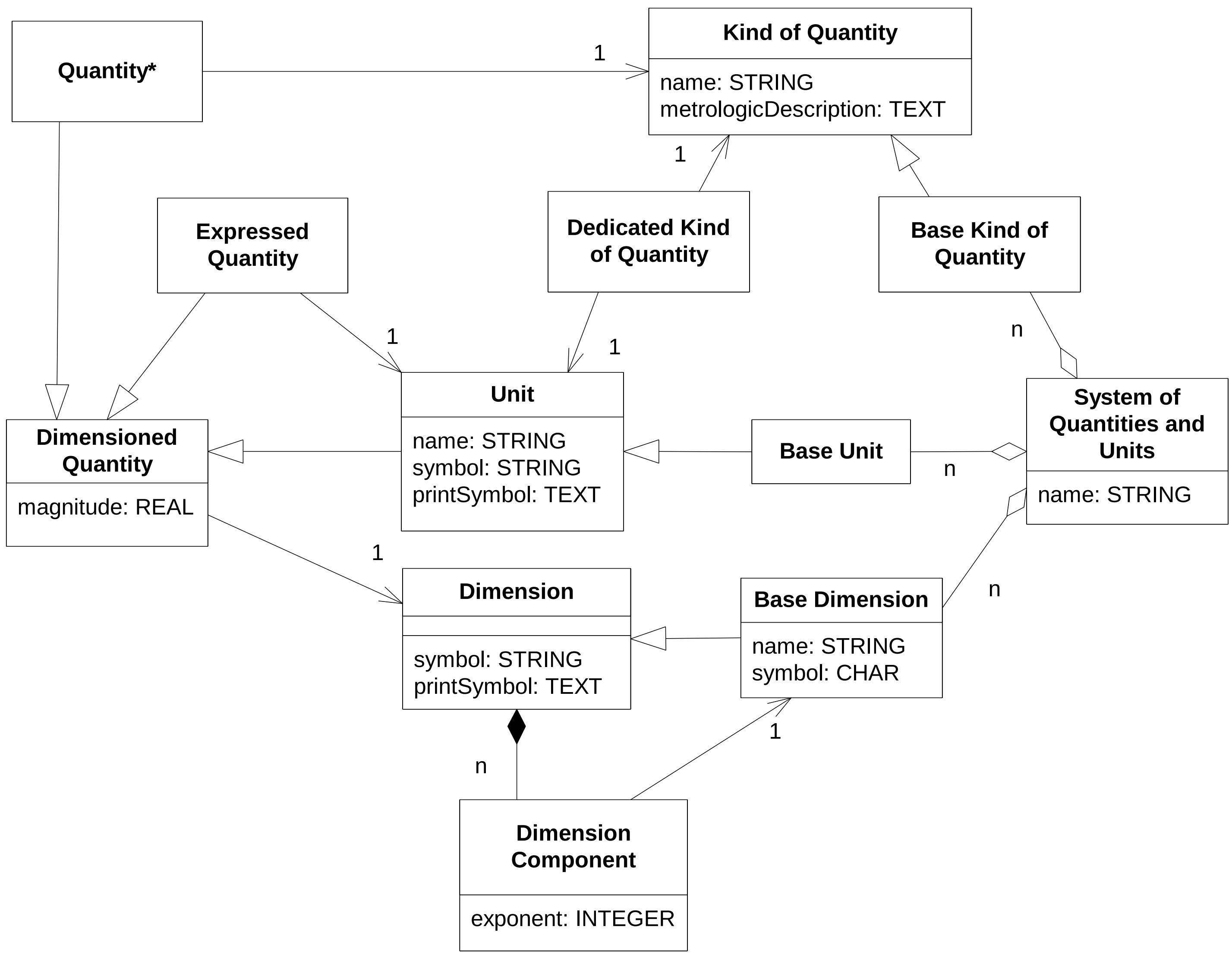
Examples for special units

$\mathbf{u'}$	$\mathbf{u}$	ν	$f(x)$	$f^{-1}(x)$
1°C	1 K	1	$x - 273.15$	$x + 273.15$
1°F	1 K	5/9	$x - 459.67$	$x + 459.67$
pH	1 mol/L	1	$-\log_{10}x$	10^{-x}
1 Np	1	1	$\ln x$	e^x
1 bel	1	1	$\log_{10}x$	10^x
1 db(SPL)	1 Pa	2×10^{-5}	$20 \cdot \log_{10}x$	$10^{x/20}$
1 db(mV)	1 mV	1	$20 \cdot \log_{10}x$	$10^{x/20}$
1 db(W)	1 W	1	$10 \cdot \log_{10}x$	$10^{x/10}$

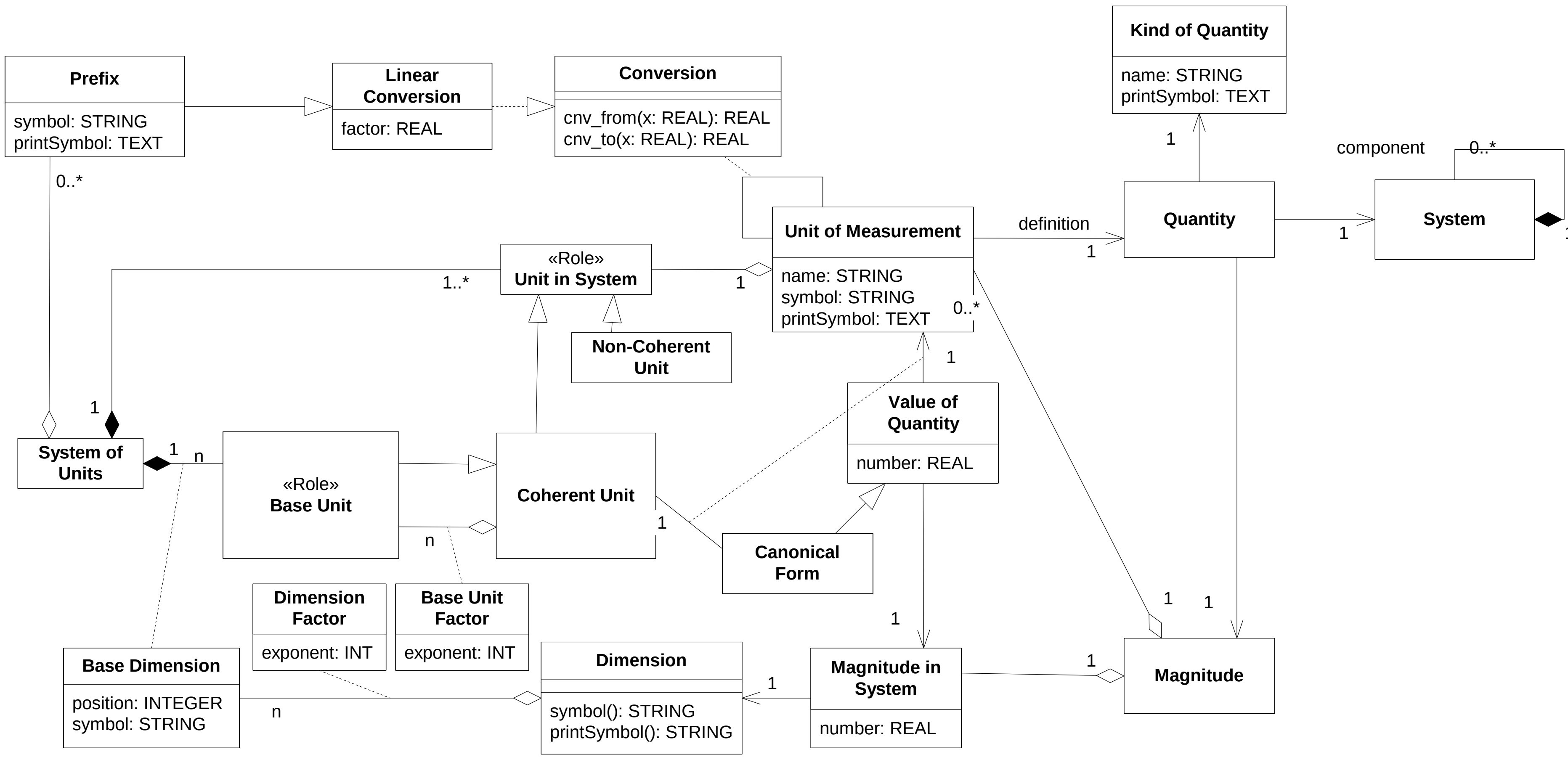
Of Note

- UCUM semantics is an algebra of units
- Quite different from symbolic knowledge representation methods that many ontologists are familiar with.
- Straight-forward implementation.
- Constant time reasoning.

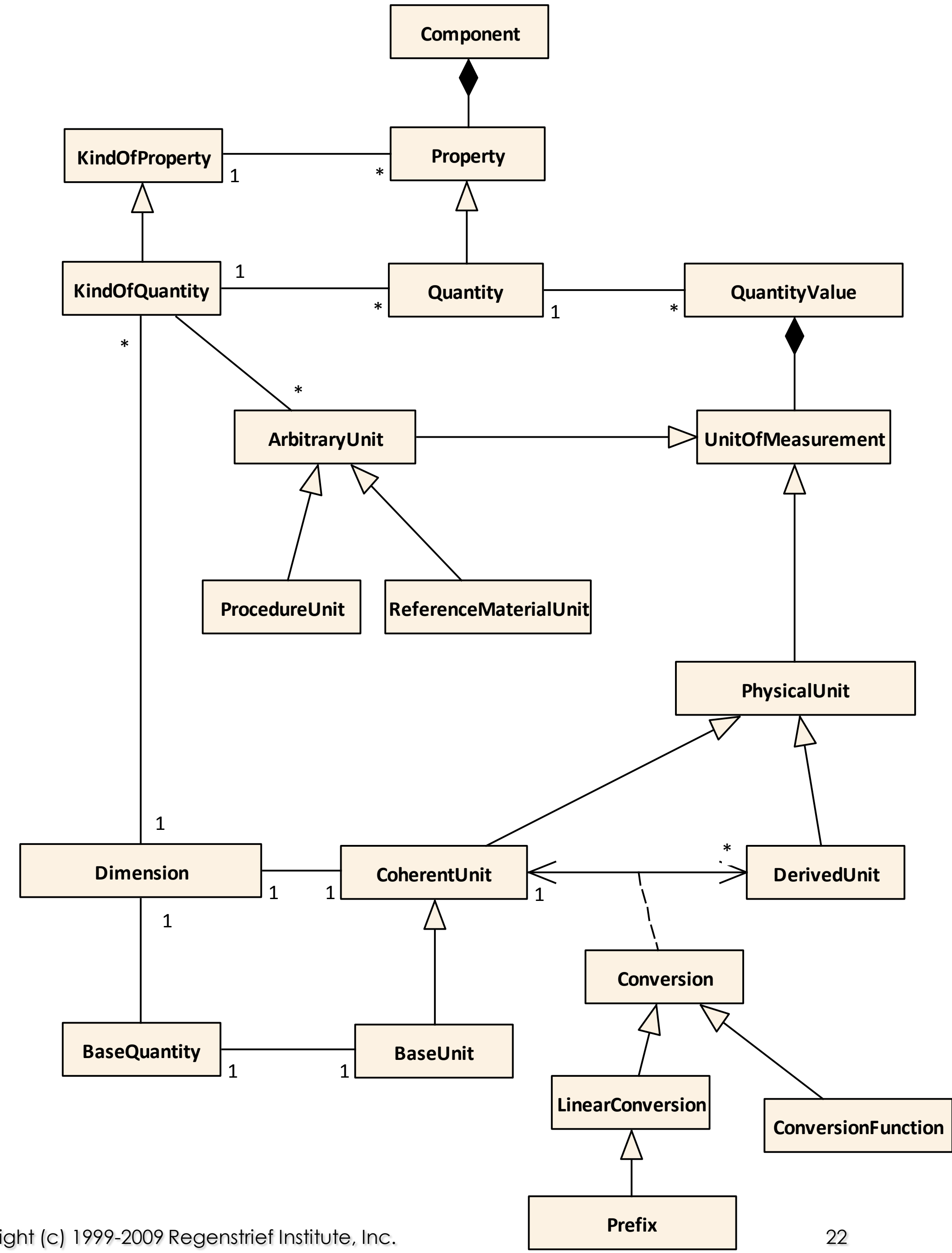
UCUM Implementation Model



Another UML model attempt for VIM, Units and UCUM



UCUM model in ISO11240 draft (TC215)



Conclusions regarding Semantics of Units of Measures

- Units of measure are essentially quantitative phenomena that require a focus on quantitative methods for their definition.
  - While symbolic models in UML are useful for discussing the design of the UCUM implementation and illustrating certain concepts
  - Symbolic schematics do not seem to replace the elegance and efficiency of the algebraic definition.
- Limitation
  - UCUM does not attempt to define base units in any formal way but uses those as primitives and refers to appropriate standards bodies (ISO, BIPM) for their definitions.
  - But we do not know of another formal system that would do that.

Ongoing UCUM Projects

1. Organization and Procedures Project
2. Managing Requests for new entries
3. Figuring out how Procedure Defined (Arbitrary) Units fit into UCUM's semantic framework
4. Testing and Quality Assurance of UCUM implementations
5. Translations, internationalization

Thank you

- http://unitsofmeasure.org