Computer Standards & Interfaces 35 (2013) 529–535

Contents lists available at SciVerse ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

Quantities, units and computing

Marcus P. Foster

CSIRO Information Management and Technology, Private Bag 33 Clayton South, VIC 3169, Australia

ARTICLE INFO

Article history:
Received 9 August 2012
Received in revised form 26 November 2012
Accepted 8 February 2013
Available online 26 February 2013

Keywords:
Physical quantity
Units
Ontology
Notation
Precedence
Character set

ABSTRACT

Quantities and units are concepts central to our measurement and manipulation of the physical world, but their representation in information systems is barely codified and often ignored. The lack of formalization of metrological concepts, operations, symbols and characters has resulted in multiple reinvention (or more dangerously, omission) of these entities in informatics systems. At best, this creates ambiguity and inconvenience; at worst, the potential for an engineering disaster. The computer representation of quantities and SI units is reviewed at these four levels. Three implementations (languages, calculators and sensor data transfer) supporting units of measure are examined. Some suggestions for a hierarchy of metrological-informatics standards are given.

Contents

Abbreviations: BIPM, International Bureau for Weights and Measures; CIPM, International Committee on Weights and Measures; CGPM, General Conference on Weights and Measures; GUM, Guide to the expression of uncertainty in measurement; IEEE, Institute of Electrical and Electronics Engineers; ISO, International Organization for Standardization; ISQ, International System of Quantities; SI8, 8th edition of the 'SI Brochure'; VIM3, International vocabulary of measurement, 3rd edition.

E-mail address: marcus.foster@csiro.au.

0920-5489/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csi.2013.02.001

1. Introduction

Quantities and units are concepts central to our measurement and manipulation of the physical world, but their representation in computational and information systems is barely codified and often ignored [1]. The lack of formalization of metrological concepts, operations, symbols and characters has resulted in multiple reinvention (or more dangerously, omission) of these entities in informatics systems. At best, this creates ambiguity and inconvenience; at worst, the potential for an engineering disaster. The most-cited example is the loss of the spacecraft Mars Climate Orbiter, which most commentators [2,3] attribute to a confusion between metric and US customary units. Even the NASA review summarized the root problem as "Failure to use metric units in the coding of a ground software file, 'Small Forces', used in trajectory models". While this is technically correct, with respect the root problem could be more usefully described as *failure to include a unit in a quantity expression*, since the thrust calculated for a trajectory correction in 'English units' was transmitted as a pure number to the spacecraft thruster software, which interpreted it as a metric unit.

This suggests that a more useful standard for calculating, storing or transmitting the value of a quantity is needed: include the magnitude, unit and kind-of-quantity ('impulse' in this case), so that the receiving computer can reject or convert it. Although this seems self-evident, it is not trivial. Thirty years ago, Finkelstein [4] described 'measurement' as "a special case of representation by symbols, with strong relation to description by language, computer data representation and the like", and optimistically predicted "there will be significant advances in the areas of database and intelligent knowledge-based systems and one can confidently expect that measurement theory will progress in strong relation with them". There has been scant codification of metrological 'computer data representation' since.

An overview of measurement standards and of possible reasons for the lack of metrological-informatics standards is given, before a suggested framework for analyzing and developing these is presented in Section 1.3.

1.1. Measurement standards

Humans have used local units of measurement for several thousand years. Trade, and increasingly science and technology, required rigorous, reproducible unit standards, and led to the creation of the metric unit platinum *etalons* kilogram and meter in 1799, the establishment of the BIPM in 1875 and the international system of units, the SI, in 1960. The current 8th edition of the SI [5] is referred to here as 'SI8'.

VIM3 [6] describes metrological entities and relationships in text and concept diagrams. It defines (1.13) a 'system of units' as "a set of base units and derived units, together with their multiples and submultiples, defined in accordance with given rules, for a given system of quantities". The SI, however, evolved from a system of practical metric units and its associated quantity system [7] was not formalized for another 20 years. The current version of this document, ISO 80000 [8] is sometimes called the international system of quantities, ISQ. None of these basic standards SI8, VIM or ISQ are presented in machine-readable form.

The ISQ uses the term 'quantity calculus' to describe the (trivial) algebra of quantities and units. Its first axiom (Maxwell's) is that the value of a quantity Q is the product of a numerical value $\{Q\}$ and a 'unit' $[Q]$ (i.e. a unit quantity): $Q = \{Q\} \cdot [Q]$.

Quantity calculus includes commutative and associative laws and scalar multiplication [9]. Different quantities can have the same dimension (e.g. torque and moment; entropy and heat capacity), and are said to be different 'kinds of quantity' (VIM3, 1.2). Quantity calculus has some limitations for practical measurement. It describes operations allowable on ratio scales, but does not apply to ordinal scales like Rockwell Hardness and to interval scales such as Celsius, Fahrenheit, etc.

1.2. Why so few metrological-informatics standards?

One reason for the neglect of units in informatics is because unit and quantity systems carry cultural, political and historical compromises, and are not as systematic, unambiguous or rigorous as computer science would expect. Some longstanding metrological discussions [10], not necessarily supported by the author, are:

- the utility of the concept 'dimension' is limited, e.g. the Hz and rad/s have the same dimension T^{-1} but are different quantities;
- if derived units are algebraically reduced, they become units of different quantities;
- the unit and dimension 'one' have multiple meanings in the SI;
- the SI base unit 'kilogram' contains a prefix, and should be renamed;
- the SI base unit (thermodynamic) mole is different to historical chemical usage and is redundant;
- angle is as tangible a geometric quantity as length and should be considered a base quantity;
- the SI should have separate units for temperature and temperature difference;
- the candela is a physiological unit defined for one wavelength, and should not be a base unit in the SI.

There are also different conceptual and philosophical understandings about measurement and measurement units:

- practical measurement has a social function, must be fit for purpose, and must concord with the theory of scales, conventions and uncertainty [11];
- the different definitions of measurement by standard bodies need to be resolved before VIM4 is drafted [12];
- the VIM focuses on ratio scales of measurement, and should be expanded to include non-numerical and ordinal measurement as described by Stevens' [13] nominal, ordinal and interval scales [14];
- the VIM concepts of 'quantity' and 'kind of quantity' overlap and require redefinition [15,16];
- the SI mole is not a 'true' measuring unit [17];
- the base quantity 'amount of substance' should be recast as 'numerosity' [18];
- the SI needs to distinguish true and 'parametric' quantities, units and dimensions [19];
- dimensionless quantities are misrepresented and should be called 'unitless' [19].

1.3. Quantities and informatics — a framework

Against this background, the representation of quantities, dimensions and SI units is reviewed at the character/keyboard, symbolic (plain and formatted text), operational and semantic levels. Fig. 1 is an informal representation of these informatics levels, and their relationship to the key metrological entities and operations. Using this framework, three implementations supporting units of measure (languages, calculators and data transfer) are examined. As the focus here is less on metrology and more on the parsing of character strings, the VIM3 concept 'value of a quantity' (e.g. '1.23×10^4 N·s')