Metrological support for quantities and units in digital systems

B. D. Hall[1] and M. Kuster[2]
[1]Measurement Standards Laboratory of New Zealand,
PO Box 31-310, Lower Hutt, New Zealand.
[2]Dumas, Texas, USA.

Abstract

Unit systems and dimensions emerged in the late 19th century. However, efforts to adapt these ideas for use in modern digital systems are proving a challenge. We provide an interpretation of units and dimensions that clarifies the main reasons for difficulties. We then suggest how a digital system would provide adequate support for quantities and units. A layer of metrological information is envisaged that would track details and allow familiar unit formats to be rendered. Three independent aspects of the data should be captured: 1) the quantity; 2) the measurement scale, scale type and conversion functions; and 3) the semantics of numerical data.

1 Introduction

The International System of Units (SI) has been adopted for some time as the basis for measurement throughout the world [1]. This system is essential to society because it facilitates reliable and effective communication about physical quantities in science, engineering, healthcare, commerce, trade, etc.

In order that the SI continue to serve society in an increasingly digital world, the responsible body, the International Committee for Weights and Measures (CIPM), has initiated a process of digital transformation of the SI. Digitalisation aims to provide unambiguous communication about quantities between autonomous machines. The transformation process will address questions of representation; but it will not affect the physical realisation and dissemination of measurement units.

The document that describes the SI is called the 'SI Brochure' [1]. Now in its 9th edition, it contains rules for writing units (another document offering extensive guidance is [2]). These rules are intended for people, not machines, and there are some difficulties associated with their digitalisation.

Digital systems use logical representations of real-world concepts to operate effectively. Difficulties in representing physical quantities have emerged which relate to how we think about quantities, dimensions, and units (here, scales of measurement are considered synonymous with units). In this short summary, we outline an interpretation of these concepts, and the relations between them, in order to address the representational difficulties. We propose to support quantities and units with additional information without substantially changing the user's perception of units.

In the next section we give an interpretation of units and dimensions based on the interoperability of quantity equations, such as physical laws. This helps to clarify the roles of units and dimensions and allows us to consider a more general representation. In section 3 we identify one of the problems not adequately addressed by the SI: unique identification of quantities. Then, in section 4, we outline properties of a general metrological basis that would support familiar unit systems. In section 5, we discuss some of the implications of our proposal before drawing conclusions in section 6.

2 Interoperability, units and dimensions

The way we currently express quantities was introduced in the 19$^{\text{th}}$ century by Maxwell to communicate clearly when different measurement systems were in use; it is an excellent case-study of interoperability. Maxwell wrote [3], *"If l is the numerical value of a length, it is understood to be expressed in terms of the concrete unit [L], so that the actual length would be fully expressed by l [L]."*

We may think about this statement in the following way. Suppose a numerical value $l = 18$ is obtained with inches as the unit of measure. By writing 18 in, someone interested in a different frame of reference can convert the value. The symbol "in" here stands for the context-dependent numerical value of the reference. For example, in=2.54 when the frame of reference uses centimetres; that is, when lengths are reported in centimetres, the value of one inch is 2.54. This flexibility is very useful, even today, because there is considerable diversity in unit preferences.

Since the early of the 20$^{\text{th}}$ century, scientists have expressed relationships between quantities without reference to units of measure. We may write, for example, that speed

$$\boldsymbol{v} = \boldsymbol{l}/\boldsymbol{t} \,, \tag{1}$$

where $\boldsymbol{t}$ is the time taken to cover a distance $\boldsymbol{l}$. An equation like (1) is called a quantity equation and the terms (shown in bold) are considered quantities not numbers. Such equations are considered true regardless of units. If numeric measures in a coherent system of units replace the terms, the equality will hold without conversion factors. A reformulation of (1) that represents actual computation uses the full expression of each quantity is

$$v[V] = l[L] \,/\, (t[T]) \tag{2}$$

where $[V]$, $[L]$, and $[T]$ are conversion factor variables (in a coherent unit system [V], [L], and [T] all equal one). Since all terms are now numeric, (2) may be separated into an equation for the value, (2a), and an equation for the conversion factor, (2b):

$$v = l/t \tag{2a}$$

$$[V] = [L]/[T] = [LT^{-1}] \tag{2b}$$

Readers will recognise the right-hand side of (2b) as the dimensions of speed. This is no coincidence: dimensional analysis may be developed in this way [4].

It is helpful to think of dimensions as numeric variables, because then the arithmetic operations in dimensional expressions apply to numbers not quantities. This is often a source of confusion.

It is also notable that terms in a quantity equation should be fully expressible as the pair of factors, $q\,[Q]$. This restricts the measurement scales compatible with an equation. For example, the acceptable scales for $\boldsymbol{l}$ in (1) belong to a class of length scales that are inter-related by simple multiplicative factors (e.g., we may express the numeric value l in metres, feet, furlongs, etc). We may refer to this class as the 'dimension of length'. In other words, the terms in a quantity equation each have an associated dimension (see [5] for more detail).

The close relation between quantity equations and dimensions has led to some confusion. However, the notion of a 'quantity' is much broader than a 'dimension'. We assume that a quantity is some property of certain objects that may be measured to establish a definite order. However, different types of scale are conceivable (for temperature: two

objects may be compared to determine which has the higher temperature, or temperatures may be reported in degrees Fahrenheit, or in degrees kelvin—each a different type of scale for temperature but only kelvin belongs in the SI dimension for temperature).

3 Difficulties for a digital SI

One of the real difficulties facing digitalisation of the SI is unique identification of quantities. This is a requirement. To quote Maxwell again, *"There must be as many different units as there are different kinds of quantities to be measured, ..."* [3]. Unfortunately, many more quantities exist than SI units; so, unit names or symbols do not reliably identify quantities. It is also worth noting that, if everyone adopted SI units (i.e., if no other measurement units were used), units and dimensions would become redundant; the quantity (e.g., SI length) would suffice.

In his *Treatise on Electricity and Magnetism*, Maxwell related the quantities used to three base quantities: length, time, and mass. The SI also follows this approach, using seven base quantities. Derived quantities are introduced by establishing a relation to other base and derived quantities; from these defining equations, dimensional expressions follow (as shown above), which we usually simplify and express as products of base quantity dimensions.

Dimensional expressions play a prominent role in the SI. The names of derived units are generated by substituting base unit names for dimensional variables. For example, the name of the SI unit of speed is metres-per-second ($\mathrm{m} \cdot \mathrm{s}^{-1}$), and the unit of acceleration is metres-per-second-per-second ($\mathrm{m} \cdot \mathrm{s}^{-2}$). This convention conveniently reminds us of the dimensional expression but does not reliably identify quantities.

In scientific writing, if the unit does not remove ambiguity readers will infer a quantity from contextual information. For instance, the turns-ratio of a transformer and the voltage gain of an electronic amplifier are of quite different physical characters, yet both are dimensionless quantities and have the SI unit "one". When distinguishing them matters, some *ad hoc* notation will be used. Another example is ITS-90, whose temperatures may be reported in kelvin, or degrees Celsius. The ITS-90 is designed to be an approximate scale for thermodynamic temperature [6]. However, measurements reported with respect to ITS-90 are strictly of a different quantity, so information must be provided in the reporting context to disambiguate the quantity being represented by the unit.

4 A metrological information layer

To address the shortcomings of unit representations we propose that digital systems use additional information. We wish to find some means of tracking quantities and scales so that digital machines will resolve difficulties such as those mentioned here. The extra information would provide a more complete metrological picture and allow familiar unit representations, like the SI, to be rendered correctly. This supports the goal of digital transformation: to fully exploit new technologies to provide enhanced benefits to users, rather than simply emulating existing systems and processes.

The full expression of a quantity should consist of a triplet of information: $l\,[L]\ <l>$, where $<l>$ identifies the quantity and $l\,[L]$ retain their meanings—value and unit. We note that there are already areas of science where SI unit symbols are annotated to identify a quantity, such as "kg N/ha", for the mass of nitrogen per hectare, or "%rh", for relative humidity expressed in percent. These notations are non-standard but suit a user community's purposes. They have arisen because scientists feel that

SI notation does not express enough detail: the very same reason that digital systems require additional information.

An independent means of identifying quantities addresses the issue of the many dimensionless quantities that have unique natures but are indistinguishable in terms of SI units. For instance, angle is defined as a ratio of lengths in the SI but so is an aspect ratio; and we would not want to confuse emissivity (a ratio of thermal radiations) with refractive index (a ratio of speeds of light); etc. Dimensionless quantities share the common feature that numeric values do not change when base units are changed. Thus, dimensionless quantities might be more aptly called 'unitless quantities' [7].

We propose that a digital Metrological (M) information layer be used to hold information about quantities, scales, and numerical formats. This M-layer would identify quantities, thereby addressing the problem posed by the many dimensionless quantities. The layer would register relationships between quantities and scales (units). It would support different types of scale, by pairing scales with appropriate conversion functions. Dimensional expressions, like $L\,T^{-1}$, ingeniously encode conversion factor functions for ratio scales, but other scales will use different functional forms. Without ambiguity, the M-layer would allow associating more than one scale with a given quantity (e.g., the kelvin and degree Rankine are both units of thermodynamic temperature) and help to manage the association of certain scales with more than one quantity (e.g., heat capacity and entropy are both indicated in $\mathrm{J \cdot K^{-1}}$; torque and work are both indicated in $\mathrm{N \cdot m}$).

Conventional temperature scales offer an example of what we mean by different types of scale for the same quantity. The SI unit of thermodynamic temperature is the kelvin, but temperature is often reported in degrees Celsius or degrees Fahrenheit. Maxwell's full expression of a quantity as a pair of terms would not apply to these scales, because the conversion mechanism is different. To convert a value in degrees Celsius to one in degrees Fahrenheit requires both an offset and a scale factor. So, it is not meaningful to talk about these scales belonging to a dimension; they are classified as 'interval' scales, whereas the kelvin scale is a 'ratio' scale citeStevens.

Scale types also relate to the semantics of quantity values. Although real numbers may represent magnitudes, such as 17.5 °C, there are restrictions on how such numbers may be used meaningfully. For instance, a ratio of interval scale indications does not represent the corresponding ratio of quantities. So, if the average temperature in summer is reported as 20.4 °C and in winter it is 10.2 °C we would be incorrect to say that the average summer temperature is twice that of winter. Another striking example is the representation of angles on a cyclic scale. Given a range of (−180°,180°], we should take care with angles near the extremes: values like −179° and 179° lie at opposite ends of the scale, but the corresponding angles are adjacent (e.g., the average of these two angles is 0° not 180°, which might have been expected). Note also that the unit, "degrees", does not specify the range (which might have been [0°,360°) degrees or $[-\infty^\circ, +\infty^\circ]$). So, we should handle numerical representation as an independent aspect of the quantity value in an M-layer.

5 Discussion

In this short summary paper, we propose that an M-layer be considered an essential component for digitalisation of quantities and units. We have shown that this layer is needed to manage the diversity of measurement scales. It has the potential to organise this diversity in a way that is consistent with metrological principles. An M-layer will not add to complexity of implementations, rather it should lead to conceptual clarity

in a problem area that has so far resisted attempts at automation.

We began by explaining the limited role of dimensions, which we showed are essentially numeric conversion factors. This clarified two basic facts that are not usually recognised: 1) the notion of dimensions applies only to ratio scales, and 2) rather than products or quotients of actual physical quantities, arithmetic operations applied to dimensional variables relates to scaling under changes of unit. We emphasize this numerical interpretation of dimensions because the role they play in unit systems is not widely appreciated. Dimensions are useful when changing scales (e.g., between Imperial and SI units), but they are not a proxy for quantities.

By clarifying the limited role of dimensions, we showed how there is an opportunity to provide better support for a broader range of measures. Some independent means of identifying quantities, such as a quantity register with unique identifiers, as well as a more general approach to unit conversion, will enhance interoperability. To render quantities in a familiar format, software would apply the unit definition(s) of interest for a given user community's context. Examples include the SI (any edition), Imperial units, and units used in specific localities or fields (e.g., medicine). This leads us to suggest an M-layer containing fundamental metrological information.

5.1 Existing software

There have been many attempts to develop software support for units and quantities and several ontologies and other systems exist, intended to support measurement units. Space does not permit us to adequately review these projects here. However, it is notable that, although none equate to our proposal, most of the elements that we think critical already exist, scattered across these diverse efforts. Ideally, we would attempt some integration of features, rather than start again from scratch, but it is too soon to speculate on that.

Briefly, we note that the 'Unified Code for Units of Measure' (UCUM) allows annotations of unit symbols, by which a human reader may identify quantities [9]. However, these annotations are not exploited by the coding system and so have no effect on the semantics. While the system does support some non-ratio scales, it still (incorrectly) attributes dimensions to them; conversion functions, to and from a ratio scale for the same quantity, are required. The 'Quantity, Unit, Dimension and Type' collection of ontologies (QUDT) contains a great deal of information about quantities and units [10]. This tool might be used to provide unique identifiers for quantities and units in an M-layer. However, QUDT uses a dimensional signature for a quantity internally and, to avoid some of the problems associated with dimensionless quantities, this signature is implemented as a ratio of dimension vectors. The 'Ontology of units of Measure' (OM) is concerned with units for quantities on ratio scales but recognises Celsius and Fahrenheit as 'scales' (a term it regards as being distinct from 'units') and associates their semantics with temperature differences, not absolute temperature [11].

In summary, none of these projects use a conceptual model exactly like the M-layer we envisage. It is difficult to foresee whether this conceptual discord would cause a problem. The feasibility of the proposed M-layer will depend on widespread acceptance of infrastructure that is supported by appropriate international organisations.

5.2 Ordinal data

Another type of measurement scale not mentioned above is the ordinal scale. Examples include the Rockwell hardness scale and the Beaufort wind scale. An ordinal scale may be used to order objects that possess the quantity but meaning cannot be attributed to ratios or differences of values on an ordinal scale. Ordinal scales will likely become

more prevalent in the immediate future. They may often be appropriate to represent data from low-cost measuring systems, because of limited calibration or other system characterisation.

5.3 Data processing

Digital representation of quantities is a prerequisite for data processing, but processing poses additional challenges. Although the topic is beyond the scope of this summary, we briefly mention a few of the difficulties here.

The mathematical operations applicable to data are constrained by properties of the scale and numerical formats. A few examples were given above, but there is a lot more to consider. An M-layer could support dynamic manipulation of data by providing a means of capturing the semantics of intermediate data processing results. The subtle relationships between scale types and quantities may affect this. Temperature, again, provides an example: a difference in temperatures on either the kelvin or Celsius scales should, in fact, be represented on a ratio scale, where it may (according to the SI) be expressed in either degrees Celsius or kelvin. That is, subtraction may result in a different type of scale.

Perhaps the most difficult problem in data processing is reliably inferring the quantity resulting from a computation. This has been a long-standing issue for digital systems, often tackled by developing programming language support for units. To date, no satisfactory solution seems to have emerged. However, this problem may be addressed by including more contextual information with calculations, which data processing algorithms would use to identify output quantities. Hall recently showed that declaring quantities of interest, in terms of defining equations, resolves many problems [12]. This gives the appearance of derived-quantity definitions in a unit system. However, the approach only identifies a small set of related quantities in the problem context, which makes it more closely akin to the method of dimensional analysis used in mathematical physics [13].

6 Conclusions

Digital representations of physical data should incorporate additional metrological information to complement that already captured in standard unit notations. A layer of metrological information is envisaged that would enable familiar unit formats to be rendered to users yet track details so systems can unambiguously exchange and process data. The layer should handle three independent data aspects: 1) the quantity; 2) the measurement scale, scale type and conversion functions; and 3) the semantics of numerical values.

Importantly, we have shown that an independent identifier, other than unit system dimensions, should identify quantities. We have also clarified the role of dimensions as being related to unit changes, but only for ratio scales. More generally, the M-layer should support conversions between different types of scale, which would be handled by appropriate forms of conversion function.

The semantics of numeric values are also important for data processing. This is not addressed by conventional unit notations at all, but has, to some extent, been considered in Stevens's classification of scale types [8]. Digital support for the numerical 'value' representations of quantity values, and the operations that may be applied to them, deserves more attention.

Acknowledgment

The authors are grateful to Vladimir Bubanja and Peter Saunders for careful reviews of this manuscript. This work was partly funded by the New Zealand government.

References

[1] BIPM, *The International System of Units (SI)*, 9th ed. BIPM, Sèvres, France, 2019.

[2] A. Thompson, B.N. Taylor, *Guide for the Use of the International System of Units (SI)*, SP-811, Nat. Inst. Stds. Technol. 2008.

[3] J.C. Maxwell, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford, 1873.

[4] G.I. Barenblatt, *Dimensional Analysis*, Gordon and Breech, New York, 1987.

[5] B. Ellis, "On the nature of dimensions", *Philosophy of Science*, vol. 31, no. 4, pp. 357-380, 1964.

[6] CCT-WG4, "Estimates of the differences between thermodynamic temperature and the ITS-90", http://www.bipm.org/utils/en/pdf/Estimates_Differences_T-T90_2010.pdf (Accessed 31 March 2021)

[7] I. Johansson, "Metrological thinking needs notions of parametric quantities, units and dimensions", *Metrologia* vol. 47, no. 3, pp 219–230, 2010.

[8] S.S. Stevens, "On the theory of scales", *Science*, vol. 103, no. 2684, pp. 677-680, 1946.

[9] Unified Code for Units of Measure (UCUM), https://ucum.nlm.nih.gov/ (accessed 26 March 2021).

[10] Quantity, Unit, Dimension and Type (QUDT), https://www.qudt.org/, (accessed 26 March 2021).

[11] Ontology of units of Measure (OM), https://github.com/HajoRijgersberg/OM (accessed 26 March 2021).

[12] B.D. Hall, "Software for calculation with physical quantities" *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0IoT)*, Rome, Italy, 2020 [doi: 10.1109/MetroInd4.0IoT48571.2020.9138281].

[13] B.D. Hall, "Software representation of measured physical quantities", in *Advanced Mathematical and Computational Tools for Metrology XII*, Series on Advances in Mathematics for Applied Sciences (World Scientific Publishing), vol 90, pp. 273–284 2021.