Units in Astronomy & Astrophysics

Arnold Rots[1], Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA
Norman Gray[2], School of Physics and Astronomy, University of Glasgow, Glasgow, UK
Sébastien Derrière[3], Centre de Données astronomique de Strasbourg (CDS), Observatoire de Strasbourg, France

This paper has been prepared by the authors as the International Astronomical Union's (IAU) contribution to the Task Group on Digital Representation of Units of Measure (DRUM) of the Committee on Data (CODATA) of the International Science Council (ISC).

1 Introduction

The community of astronomers and astrophysicists has, like many scientific communities, always had a strong international outlook, which received an extra nudge by the establishment of large observatories in prime observing locations. Simply put: astronomers need data that cannot just be obtained from experiments conducted in an randomly located laboratory. This necessitated clear data exchange standards, and thus agreed upon unit definitions. Presently, there are three authoritative documents that define units to be used in the field. Those are generally based on SI units, but there are quirks (such as the continued use of certain *cgs* units), special-purpose units, and serious caveats.

In the following sections we shall present the standards documents, the special astronomical units, the quirks, the special quantities, and the caveats.

2 Standards Documents

There are essentially three documents that provide recommendations to astronomers regarding the use of units. These documents are in close agreement, though there may be differences in emphasis.

2.1 The IAU Style Manual

Published by the International Astronomical Union, the *IAU Style Manual* by G. A. Wilkins, Commission 5, in IAU Transaction XXB (1989) contains a section titled *SI Units* which has been slightly reformatted in Appendix A.

2.2 The FITS Standard

The Flexible Image Transport System (FITS), initially intended to provide a standard enabling the exchange of images in astronomy on 9-track tapes and subsequently expanded to provide a general standard for the exchange of digital data, includes prescriptions for what units to use and how to

[1] ORCID: 0000-0003-2377-2356; arots@cfa.harvard.edu
[2] ORCID: 0000-0002-1941-9202; norman.gray@glasgow.ac.uk
[3] ORCID: 0000-0002-3545-7065; sebastien.derriere@astro.unistra.fr

express them. The FITS standard documents may be found at:
https://fits.gsfc.nasa.gov/fits_documentation.html
Section 4.3, which deals with units is extracted in Appendix B. This section is important, because the FITS standard covers probably the majority of the digital astronomical data products that are publicly available.

2.3 The IVOA Recommendation on Units

The International Virtual Observatory Alliance (IVOA), an organization working on standards for data exchange and interoperability among astronomical data repositories, has issued a recommendation on the use of units in astronomy and astrophysics data:
https://www.ivoa.net/documents/VOUnits/20140523/index.html.
This is the most recent document and should be used as a primary reference. It also cross-references the other two standards mentioned above. There is one exception: due to a misunderstanding regarding IAU recommended units, it states incorrectly that “yr” is the preferred symbol for year, rather than “a”; to be clear: “a” is the preferred symbol for year.

Section 2 of this document and Tables 1 through 6 deal with the Units *per se*.

3 Special Units in Astronomy

The most obvious examples of units one will only find in an astronomical context are units of length, units of time, and units of flux density. In addition, some specific definitions related to celestial coordinates should be mentioned.

3.1 Length

Within the solar system and often in stellar astrophysics one will encounter the Astronomical Unit (AU). This represents the average earth-sun distance, and is currently defined to have an exact value, close to 0.149598 Tm. In Galactic and extragalactic astronomy, the parsec (pc) is the unit of choice. It is the distance at which a star exhibits a parallax of one second of arc as the earth moves in its orbit around the sun, i.e., 3800 × (180 / π) × 2 AU = 30.857 Pm. The lightyear (lyr) is mainly used in popular presentations: 9.461 Pm.

3.2 Time

In astronomy it is not uncommon to use additional units of time, larger than the second: minute, hour, day, year, and century. One needs to keep in mind that, depending on context, time may be used as a measure of angle as well as a measure of (chronographic) duration. Specifically, the units day and year used to be based on measures of angle: earth’s rotation around its axis and its orbital motion around the sun. To avoid the intricacies of leap seconds and Julian, Besselian, and tropical years, the unit “day” is now generally taken as 86400 s, the unit “year” as 365.25 days, and the unit “century” as 36525 days (essentially, the Julian year and century). In addition, there is the concept of sidereal time which is based on the definition of a day as one exact rotation of the earth (with respect to the fixed stars, rather than with respect to the mean sun, meaning that there are 366.25 days in a sidereal year), leading to a day that is roughly four minutes shorter. However, this only really applies to (earth-bound) telescope operations and will hardly be visible outside the field. The above are rather generic definitions: time in astronomy is a particularly subtle question, and we discuss it in more detail below, under "Caveats".

3.3 Flux Density

The most SI-like unit is the Jansky (Jy) at 10^{-26} W m^{-2} Hz^{-1}; the quirk is that -26 is not a multiple of 3. The traditional unit used in the optical part of the spectrum is the magnitude (mag). Its use goes back a few thousand years and their physically quantitative interpretation tends to be rather involved. Magnitudes are logarithms of flux ratios and their conversion to a flux density in SI units is specific to the actual band filter used. The Rayleigh (R) is used mainly for auroral and airglow measurements. In High Energy Astrophysics (X-ray, Gamma-ray, and beyond) where actual individual photons are detected, *counts* (a dimensionless unit) or count rate (*counts*/s) are commonly used. Since the photon energy is usually provided, the conversion is not too onerous. Whereas optical astronomers generally use wavelength for the spectral coordinate, and radio astronomers frequency, HEA commonly uses energy (eV).

3.4 Celestial Coordinates

The four main spherical coordinate systems on the sky are equatorial, ecliptic, Galactic, and Super-Galactic. Since these are being viewed from the inside, the longitude component increases to the left. Any of the reference positions listed in the next section may be used as origin, although some may not make sense. In addition, Cartesian coordinate systems can be used that are aligned with these coordinate systems, for instance for solar system or satellite orbit ephemerides. Position angles (for instance of elliptical shapes or linear polarization) are defined as measured from North through East (i.e., clockwise). Obviously, in most cases these will be modulo 180°, but there are cases (orbital parameters, galaxies with known rotation) where they are modulo 360°.

Decimal degrees are used for all. However, equatorial coordinates have additional properties. Its coordinates (Right Ascension and Declination) are often displayed in sexagesimal form, but in that case Right Ascension (the longitude component) is provided in hours, minutes, and seconds: hh:mm:ss.s… This may seem peculiar, but it reflects the comment above that time may represent an angle: Right Ascension represents the Local Sidereal Time at which this equatorial longitude passes through the local meridian. There is one other peculiarity. In principle, the equatorial coordinate frame is determined by the direction of the earth’s axis and the ascending node of the ecliptic. Because of a number of perturbations, dominated by the precession of the earth’s axis, both directions change in time. Up until the end of the 20th century this has necessitated the addition of a standard “Equinox” label to the coordinates, such as B1900, B1950, J2000 (*B* for Besselian years, *J* for Julian years). Currently, the common use is to notate coordinates in ICRS (International Celestial Reference System) which is close to J2000. If an “Epoch” is attached to a coordinate pair, it indicates a position at that particular time (e.g., for moving objects). In summary: *Equinox* refers to the coordinate frame, *Epoch* to a particular position in that frame.

4 Caveats

4.1 Reference positions

Astronomical observations are generally transformed, for convenience, to one or other of a small set of nominal reference positions, and the position in question is as important in the measurement as the coordinate system or units being used. We refer to this location as the reference position. Such positions are specified in 6-dimensional phase space.

The most commonly used reference positions are:

- Topocenter: the location of the measuring device; typically the location of an observatory or telescope, either on (or in) the earth or anywhere else in space.
- Geocenter: the center of the earth; to cancel out the effects of the earth's rotation
- Earth-moon barycenter: the center of mass in the earth-moon system
- Heliocenter: center of the sun; mainly for solar research
- (Solar system) barycenter: barymetric center of the solar system (still contained within the solar body); common for observations, to remove the effects caused by the earth's annual motion
- Local Standard of Rest: basically the Barycenter, but with velocities corrected for the local motion of the sun
- Galactocentric: spatially located at the center of the Galaxy and corrected for the sun's orbital motion in the Galaxy
- In addition, the centers of any of the planets in the solar system may be used

4.2 Velocities and Redshifts / Doppler shifts

Pinpointing objects in 6-dimensional phase space is obviously problematic in astronomy. Two spatial coordinates, celestial longitude and latitude, can clearly be measured, but the third – distance – needs to be inferred from different properties; for instance, if one knows the intrinsic luminosity of a star, one can derive its distance from the apparent luminosity. Similarly, the space velocity in a plane perpendicular to the line of sight can be measured, but its accuracy is obviously severely limited by the distance to the object in question and the accuracy of that distance. The way out for the line-of-sight velocity is the use of spectral lines to determine redshifts and derive Doppler velocities. Subsequently, one can determine the distance of an external galaxy from its Doppler velocity, based on a model of the universe's expansion.

The issue here is that, although distances are expressed in parsecs and line-of-sight velocities are expressed in units of km/s, these are really values derived from the measurement of non-spatial quantities. Therefore, one should exercise caution in treating them as spatial quantities. In addition, there are a few quirks involved in the concept of Doppler velocity, as explained below in the *Quirks* section. And, as far as linear sizes of objects in space are concerned, the conversion from angular size to linear size using a known distance is not straightforward for distant objects, as relativistic effects kick in.

4.3 Time

Time is a multi-faceted concept in astronomy for a number of reasons. We refer to Rots et al. (2015)[4] for more details.

There are two aspects to timekeeping: specifying a particular point in time (a timestamp) and measuring the length of time intervals. The connection between the two is that the timestamp is really an interval that has elapsed since a commonly accepted fiducial zero point. Most commonly used for timestamps are ISO-8601 strings (`CCCC-MM-DDThh:mm:ss.s`… or [+-]`CCCCC-MM-DDThh:mm:ss.s`…) or Julian Days (JD; days since −04713-11-24T12:00:00). One should note that −00001-12-31T23:59:59 is the second that is followed by +00001-01-01T00:00:00. Hours, minutes, and seconds are counted zero-relative in ISO-8601, while days, months, and years are counted one-relative.

There are also two modes of timekeeping: counting seconds in a chronographs and measuring an angle. The first problem that arises is that, although we tend to count time intervals chronographically in elapsed SI seconds, originally time intervals were expressed as rotation angle measures: as fractions of the earth's rotation (from one mean sun meridian passage to the next) or of the earth's orbital rotation around the sun. Neither rotation is constant. That is how we have ended up with Tropical, Besselian, and Julian years and time scales like UT1 and TAI. UTC is the time scale that bridges the two domains: it is a count of SI seconds (i.e., atomic seconds, just like TAI), but kept in sync with the earth's rotation (as represented by UT1) by inserting leap seconds as needed.

The second problem is a typical astronomical one. In most terrestrial applications recording the timestamp of an event is a simple and unambiguous matter, as time signals are distributed and synchronized across the surface of the globe. And relating the elapsed time between two events is similarly a simple matter. In astronomy, however, one does need to know the spatial position at which related events with critical timing information were observed; the most extreme example would be a set of two observations made six months apart, where the optical pathlengths to a celestial source could be different by up to the diameter of the earth's orbit. Commonly, astronomers refer timestamps of time-sensitive data and ephemerides to the barycenter of the solar system. As an aside, pathlength corrections may not only involve geometric pathlengths, but also relativistic effects caused by photons' traveling through varying gravitational potentials.

The third issue is that of time scales. Neither ISO-8601, nor JD implies the use of a particular timescale. For instance, UTC, TAI (International Atomic Time, maintained by the international time services), TT (Terrestrial Time, used in astronomy), and GPS run synchronously, but have various offsets from each other in the tens of seconds, while UTC's offset increases with time due to the insertion of leap seconds. UTC is most commonly used, even though it is eminently unsuitable because not all days have the same length. Time intervals derived from timestamps in TAI, TT, and GPS will be identical and can be directly converted to SI seconds; that is not necessarily the case for UTC.

[4] https://ui.adsabs.harvard.edu/link_gateway/2015A%26A...574A..36R/PUB_PDF

Finally, General Relativity introduces two additional problems: the rate of time's progression is dependent on one's motion and the gravitational potential at one's location. It may generally suffice to be aware that there are a variety of time scales in use that differ subtly, with as bottom line: even though all SI seconds are the same, they may *appear* different since they are realized in different time scales at different locations, with different motions and different gravitational potentials. This will not be an issue for most applications restricted to activities on the earth's surface, but in case anyone wonders about these effects we will briefly elaborate on this.

As we travel through space on the surface of the globe, we share a common gravitational potential and space motion, and it is perfectly reasonable to keep our clocks synchronized. The same applies with adequate precision to most satellites and spacecraft traveling with the earth. However, elsewhere this breaks down because of the eccentricity of the earth's orbit. Seen from the solar system's barycenter the earth's clocks appear to slow down and speed up over the course of a year, as the earth's orbital speed, as well as its location in the solar system's gravitation potential vary. This is resolved by the definition of TDB (Barycentric Dynamical Time) which on average (i.e., averaged over a year) runs synchronously with TAI; but it means that time intervals between two timestamps will generally not be the same in TAI and in TDB.

As well as TDB and TAI, there are multiple other timescales which are appropriate for different astronomical applications, including for example *coordinate* time scales TCG (referenced to the geocenter) and TCB (solar system barycentric). These are defined, by the IAU, via a network of mutual conversions between them. The underlying issue is that time *appears* to run at different rates in these time scales due to the difference in gravitational potential at their reference locations. This leads to what is ultimately relevant to a discussion of units: although *all* of these timescales fundamentally have the same *unit* – the SI second – there are subtle differences between them which are important for certain applications, even though the timescale being used may not be obvious from the unit string; hopefully, it will be clear from the context.

4.4 Polarization

When it comes to measuring polarized radiation in astronomy, the issue is not so much one of the units used to express the values, as it is the definition of the observed parameters themselves and how they relate to an inertial frame. Directly measured values will be related to the frame of the measuring device (telescope), which means that the relation between its position angle and the celestial position angle (for definition, see previous section) may change in time. In published results, polarization may be expressed in Stokes parameters (I, Q, U, V), linear polarization percentage and position angle, or circular polarization.

5 Quirks

5.1 *cgs* Units

Considering the often large exponents in astronomical quantities, it is peculiar that many parameters are still expressed in *cgs* units, particularly cm and erg. It has proven difficult to change habits in this respect. On the other hand, the exponents would still be large if SI units were used throughout.

5.2 Doppler shifts

We noted the fact that line-of-sight velocities derived from Doppler shifts should be treated with caution, as they are not on par with directly measured linear velocities. However, one should also be aware that there are three flavors of Doppler velocities in use: *OPTICAL, RADIO*, and *RELATIVISTIC*.

OPTICAL: $$V_{opt} = c \cdot \frac{\Delta\lambda}{\lambda_0} = -c \cdot \frac{\Delta\nu}{\nu}$$

RADIO: $$V_{rad} = -c \cdot \frac{\Delta\nu}{\nu_0} = c \cdot \frac{\Delta\lambda}{\lambda}$$

RELATIVISTIC: $$V_{rel} = c \cdot \frac{\lambda^2 - \lambda_0^2}{\lambda^2 + \lambda_0^2} = -c \cdot \frac{\nu^2 - \nu_0^2}{\nu^2 + \nu_0^2}$$

λ and ν are the measured wavelength and frequency, respectively, of an observed spectral line; λ_0 and ν_0 are the rest wavelength and rest frequency of that spectral line.

The *OPTICAL* definition is generally used for observational data in all extra-galactic contexts and for optical Galactic data. The *RADIO* definition is used for Galactic data based on spectral lines in the radio part of the spectrum. If the *RELATIVISTIC* definition is used, one should keep in mind that this still does not represent a true linear velocity, particularly at large distances.

6 Appendix A

Reprinted here is the section “SI Units” from the "IAU Style Manual" by G.A. Wilkins, Comm. 5, in IAU Transactions XXB (1989), which may be consulted for further details (download PDF file, 4.7MB)

SI Units

The international system (SI) of units, prefixes, and symbols should be used for all physical quantities except that certain special units, which are specified later, may be used in astronomy, without risk of confusion or ambiguity, in order to provide a better representation of the phenomena concerned. SI units are now used to a varying extent in all countries and disciplines, and this system is taught in almost all schools, colleges and universities. The units of the centimetre-gram-second (CGS) system and other non-SI units, which will be unfamiliar to most young scientists, should not be used even though they may be considered to have some advantages over SI units by some astronomers.

General information about SI units can be found in the publications of national standards organisations and in many textbooks and handbooks.

There are three classes of SI units: (a) the seven base units that are regarded as dimensionally independent; (b) two supplementary, dimensionless units for plane and solid angles; and (c) derived units that are formed by combining base and supplementary units in algebraic expressions; such derived units often have special names and symbols and can be used in forming other derived units. The units of classes (a) and (b) are listed in Table 1. The units of class (c) of greatest interest to astronomers are given in Table 2 for those with simple names and symbols, and in Table 3 for those with compound names and symbols. In forming compound names division is indicated by per, while in the corresponding symbols it is permissible to use either a negative index or a solidus (oblique stroke or slash); thus the SI: unit of velocity is a metre per second and the corresponding symbol is m s^{-1} or m/s.

The space between the base units is important in such a case since m/s could be interpreted as a frequency of 1000 Hz; a space is not necessary if the preceding unit ends in a superscript; a full stop (period) may be inserted between units to remove any ambiguity; the solidus should only be used in simple expressions and must never be used twice in the same compound unit.

Table 1. The names and symbols for the SI base and supplementary units.

Quantity	*SI Unit: Name*	*Symbol*
length	metre	m
mass	kilogram	kg
time [(1)]	second	s
electric current	ampere	A

thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd
plane angle	radian	rad
solid angle	steradian	sr

[1] The abbreviation sec should not be used to denote a second of time.

Table 2. Special names and symbols for SI derived units.

Quantity	*SI Unit: Name*	*Symbol*	*Expression*
frequency	hertz	Hz	s^{-1}
force	newton	N	kg m s^{-2}
pressure, stress	pascal	Pa	N m^{-2}
energy	joule	J	N m
power	watt	W	J s^{-1}
electric charge	coulomb	C	A s
electric potential	volt	V	J C^{-1}
electric resistance	ohm	Omega	V A^{-1}
electric conductance	siemens		A V^{-1}
electric capacitance	farad	F	C V^{-1}
magnetic flux	weber	Wb	V s
magnetic flux density	tesla	T	Wb m^{-2}
inductance	henry	H	Wb A^{-1}
luminous flux	lumen	lm	cd sr
illuminance	lux	lx	lm m^{-2}

Table 3. Examples of SI derived units with compound names.

Quantity	*SI unit: Name*	*symbol*
density (mass)	kilogram per cubic metre	kg m^{-3}
current density	ampere per square metre	A m^{-2}
magnetic field strength	ampere per metre	A m^{-1}
electric field strength	volt per metre	V m^{-1}
dynamic viscosity	pascal second	Pa s

heat flux density	watt per square metre	W m^{-2}
heat capacity, entropy	joule per kelvin	J K^{-1}
energy density	joule per cubic metre	J m^{-3}
permittivity	farad per metre	F m^{-1}
permeability	henry per metre	H m^{-1}
radiant intensity	watt per steradian	W sr^{-1}
radiance	watt per square metre per steradian	W m^{-2} Sr^{-1}
luminance	candela per square metre	cd m^{-2}

Table 4. SI prefixes and symbols for multiples and submultiples.

Submultiple	*Prefix*	*Symbol*	*Multiple*	*Prefix*	*Symbol*
10^{-1}	deci	d	10	deca	da
10^{-2}	centi	c	10^{2}	hecto	h
10^{-3}	milli	m	10^{3}	kilo	k
10^{-6}	micro	mu	10^{6}	mega	M
10^{-9}	nano	n	10^{9}	giga	G
10^{-12}	pico	p	10^{12}	tera	T
10^{-15}	femto	f	10^{15}	peta	P
10^{-18}	atto	a	10^{18}	exa	E

Note: Decimal multiples and submultiples of the kilogram should be formed by attaching the appropriate SI prefix and symbol to gram and g, not to kilogram and kg.

4.12 SI prefixes: Decimal multiples and submultiples of the SI: units, except the kilogram, are formed by attaching the names or symbols of the appropriate prefixes to the names or symbols of the units. The combination of the symbols for a prefix and unit is regarded as a single symbol which may be raised to a power without the use of parentheses. The recognised list of prefixes and symbols is given in Table 4. These prefixes may be attached to one or more of the unit symbols in an expression for a compound unit and to the symbol for a non-SI unit. Compound prefixes should not be used.

4.13 Non-SI units: It is recognised that some units that are not part of the international system will continue to be used in appropriate contexts. Such units are listed in Table 5; they are either defined exactly in terms of SI units or are defined in other ways and are determined by measurement. Other non-SI units, such as Imperial units and others listed in Table 6, should not normally be used.

Table 5. Non-SI units that are recognised for use in astronomny.

Quantity	*Unit: Name*	*Symbol*	*Value*
time [(1)]	minute	min or ‘	60 s
time	hour	h	3600 s = 60 min
time	day	d	86 400 s = 24 h
time	year (Julian)	a	31.5576 Ms = 365.25 d
angle [(2)]	second of arc	"	(pi/648 000) rad
angle	minute of arc	'	(pi/10 800) rad
angle	degree	º	(pi/180) rad
angle [(3)]	revolution(cycle)	c	2pi rad
length	astronomical unit	au	0.149 598 Tm
length	parsec	pc	30.857 Pm
mass	solar mass	$M_\odot$	1.9891 x 10^{30} kg
mass	atomic mass unit	u	1.660 540 x 10^{-27} kg
energy	electron volt	eV	0.160 2177 aJ
flux density	jansky [(4)]	Jy	10^{-26} W m^{-2} Hz^{-1}

[1] The alternative symbol is not formally recognised in the SI system.

[2] The symbol mas is often used for a milliarcsecond (0".001).

[3] The unit and symbols are not formally recognised in the SI system.

[4] The jansky is mainly used in radio astronomy.

[5] The degree Celsius (oC) is used in specifying temperature for meteorological purposes, but otherwise the kelvin (K) should be used.

5.14 Time and angle : The units for sexagesimal measures of time and angle are included in Table 5. The names of the units of angle may be prefixed by 'arc' whenever there could be confusion with the units of time. The symbols for these measures are to be typed or printed (where possible as superscripts) immediately following the numerical values; if the last sexagesimal value is divided decimally, the decimal point should be placed under, or after, the symbol for the unit; leading zeros should be inserted in sexagesimal numbers as indicated in the following examples.

2d 13h 07m 15.259s 06h 19m 05.18s 120o 58' 08".26

These non-SI units should not normally be used for expressing intervals of time or angle that are to be used in combination with other units.

In expressing the precision or resolution of angular measurement, it is becoming common in astronomy to use the milliarcsecond as the unit, and to represent this by the symbol mas; this is preferable to other abbreviations, but its meaning should be made clear at its first occurrence. The more appropriate SI Unit would be the nanoradian (1 nrad = 0.2 mas). In general, the degree with decimal subdivision is recommended for use when the radian is not suitable and when there is no requirement to use the sexagesimal subdivision. If it is more appropriate to describe an angle in terms of complete revolutions (or rotations or turns or cycles), then the most appropriate symbol appears to be a letter c; this may be used in a superior position as in 1c = 360º =2pi rad = 1 rev, but it may be used as in 1 c/s = 1Hz.

The use of units of time for the representation of angular quantities, such as hour angle, right ascension and sidereal time, is common in astronomy, but it is a source of confusion and error in some contexts, especially in formulae for numerical calculation. The symbol for a variable followed by the superscript for a unit may be used to indicate the numerical value of that variable when measured in that unit.

5.15 Astronomical units: The IAU System of Astronomical Constants recognises a set of astronomical units of length, mass and time for use in connection with motions in the Solar System; they are related to each other through the adopted value of the constant of gravitation when expressed in these units (IAU 1976). The symbol for the astronomical unit of length is au; the astronomical unit of time is 1 day (d) of 86 400 SI seconds (s); the astronomical unit of mass is equal to the mass of the Sun and is often denoted by $M_\odot$, but the special subscript makes this symbol inconvenient for general use.

An appropriate unit of length for studies of structure of the Galaxy is the parsec (pc), which is defined in terms of the astronomical unit of length (au). The unit known as the light-year is appropriate to popular expositions on astronomy and is sometimes used in scientific papers as an indicator of distance.

The IAU has used the julian century of 36 525 days in the fundamental formulae for precession, but the more appropriate basic unit for such purposes and for expressing very long periods is the year. The recognised symbol for a year is the letter a, rather than yr, which is often used in papers in English; the corresponding symbols for a century (ha and cy) should not be used. Although there are several different kinds of year (as there are several kinds of day), it is best to regard a year as a julian year of 365.25 days (31.5576 Ms) unless otherwise specified.

It should be noted that sidereal, solar and universal time are best regarded as measures of hour angle expressed in time measure; they can be used to identify instants of time, but they are not suitable for use as precise measures of intervals of time since the rate of rotation of Earth, on which they depend, is variable with respect to the SI second.

5.16 Obsolete units: It is strongly recommended that the non-SI units listed in Table 6 are no longer used. Some of the units listed are rarely used in current literature, but they have been included for use in the study of past literature. Imperial and other non-metric units should not be used in connection with processes or phenomena, but there are a few situations where their use may be justified (as in "the Hale 200-inch telescope on Mount Palomar"). The equivalent value in SI units should be given in parentheses if this is likely to be helpful.

Table 6. Non-SI units and symbols whose continued use is deprecated.

Quantity	*Unit: Name*	*Symbol*	*Value*
length	angstrom	Å	10^{-1O} m = 0.1 nm
length	micron	mu	10^{-6} m
length	fermi		1 fm
area	barn	b	10^{-28} m^2
volume	cubic centimetre	cc	10^{-6} m^3
force	dyne	dyn	10^{-5} N
energy	erg	erg	10^{-7} J
energy [(2)]	calorie	cal	4.1868 J
pressure	bar	bar	10^{5} Pa
pressure	stand. atmosphere	atm	101 325 Pa
acceleration (grav.)	gal	Gal	10^{-2} m s^{-2}
gravity gradient	eotvos	E	10^{-9} s^{-2}
magnetic flux density	gauss	G	corresponds to 10^{-4} T
magnetic flux density	gamma		corresponds to 10^{-9} T
magn. field strength	oersted	Oe	corr. to (1000/4pi) A m^{-1}

[1] Non-metric units, such as miles, feet, inches, tons, pounds, ounces, gallons, pints, etc., should not be used except in special circumstances.

[2] There are other obsolete definitions and values for the calorie.

The definitions of the SI units and an extensive list of conversion factors for obsolete units are given by Anderson (Physics Vade Mecum, American Institute of Physics 1981). In particular, wavelengths should be expressed in metres with the appropriate SI prefix; e.g., for wavelengths in the visual range the nanometre (nm) should be used instead of the angstrom (Å), which is a source of confusion in comparisons with longer and shorter wavelengths expressed in recognised SI units. The notation of the form of a Greek Lambda foIlowed by a numerical value (which represents the wavelength in angstroms) should also be abandoned.

The name micrometre should be used instead of micron. In all cases, the spelling metre should be used for the unit, while the spelling meter should be used for a measuring instrument (as in micrometer). The word kilometre should be pronounced ki-lo-me-te, not kil-lom-e-ter.

If wavenumbers are used they should be based on the metre, not the centimetre; in any case the unit (m-l or cm-l) should be stated since they are not dimensionless quantities. The uses of frequency (in Hz) at radio wavelengths and energy (in eV) at X-ray wavelengths are appropriate for some purposes, but they serve to obscure the essential unity of the electromagnetic spectrum, and so it may be helpful to give the wavelength as well at the first occurrence; the correspondences between these units and wavelength are as follows:

wavelength in metres =2.997 924 58 x 10^{8} / frequency in Hertz

or = 1.239 842 4 x $l0^{6}$ / energy in electron-volts

5.17 Magnitude: The concept of apparent and absolute magnitude in connection with the brightness or luminosity of a star or other astronomical object will continue to be used in astronomy even though it is difficult to relate the scales of magnitude to photometric measures in the SI system. Magnitude, being the logarithm of a ratio, is to be regarded as a dimensionless quantity; the name may be abbreviated to mag without a full stop, and it should be written after the number. The use of a superscript m is not recommended. The method of determination of a magnitude or its wavelength range may be indicated by appropriate letters in italic type as in U, B, V. The photometric system used should be clearly specified when precise magnitudes are given.

7 Appendix B

FITS documentation may be found at https://fits.gsfc.nasa.gov/fits_documentation.html.
Following is section 4.3 of the FITS Standard paper, version 4, which deals with units.

4.3. Units

When a numerical keyword value represents a physical quantity, it is *recommended* that units be provided. Units *shall* be represented with a string of characters composed of the restricted ASCII-text character set. Unit strings can be used as values of keywords (e.g., for the reserved keywords BUNIT, and TUNIT*n*), as an entry in a character-string column of an ASCII-table or binary-table extension, or as part of a keyword comment string (see Sect. 4.3.2, below).
The units of all *FITS* header keyword values, with the exception of measurements of angles, *should* conform with the recommendations in the IAU Style Manual (McNally 1988). For angular measurements given as floating-point values and specified with reserved keywords, the units *should* be degrees (i.e., deg). If a requirement exists within this Standard for the units of a keyword, then those units *must* be used.
The units for fundamental physical quantities recommended by the IAU are given in Table 3. Table 4 lists additional units that are commonly used in astronomy. Further specifications for time units are given in Sect. 9.3. The recommended plain-text form for the IAU-recognized *base units* are given in Column 2 of both tables.6 All base units strings *may* be preceded, with no intervening spaces, by a single character (two for deca) taken from Table 5 and representing scale factors mostly in steps of 103. Compound prefixes (e.g., ZYeV for 1045 eV) *must not* be used.

4.3.1. Construction of units strings

Compound units strings *may* be formed by combining strings of base units (including prefixes, if any) with the recommended syntax described in Table 6. Two or more base units strings (called str1 and str2 in Table 6) *may* be combined using the restricted set of (explicit or implicit) operators that provide for multiplication, division, exponentiation, raising arguments to powers, or taking the logarithm or square-root of an argument. Note that functions such as log actually require dimensionless arguments, so that log(Hz), for example, actually means log(x/1 Hz). The final units string is the compound string, or a compound of compounds, preceded by an *optional* numeric multiplier of the form 10**k, 10ˆ*k*, or 10±*k* where *k* is an integer, *optionally* surrounded by parentheseswith the sign character *required* in the third form in the absence of parentheses. Creators of *FITS* files are encouraged to use the numeric multiplier only when the available standard scale factors of Table 5 will not suffice. Parentheses are used for symbol grouping and are strongly *recommended* whenever the order of operations might be subject to misinterpretation. A space character implies multiplication, which can also be conveyed explicitly with an asterisk or a period. Therefore, although spaces are allowed as symbol separators, their use is discouraged. Note that, per IAU convention, case is significant throughout. The IAU style manual forbids the use of more than one slash ('/') character in a units string. However, since normal mathematical precedence rules apply in this context, more than one slash *may* be used but is discouraged. A unit raised to a power is indicated by the unit string followed, with no intervening spaces, by the *optional* symbols ** or ˆ followed by the power given as a numeric expression, called expr in Table 6. The power *may* be a simple integer, with or without sign, *optionally* surrounded by parentheses. It *may* also be a decimal number (e.g., 1.5, 0.5) or a ratio of two integers (e.g., 7/9), with or without sign, which *must* be surrounded by parentheses. Thus meters squared *may* be indicated by m**(2), m**+2, m+2, m2, m^2, m^(+2), etc. and per meter cubed *may* be indicated by m**-3, m-3, m^(-3), /m3, and so forth. Meters to the three-halves power *may* be indicated by m(1.5), m^(1.5), m**(1.5), m(3/2), m**(3/2), and m^(3/2), but *not* by ms/2 or m1.5.

Table 3: IAU-recommended basic units.

Quantity	Unit	Meaning	Notes
SI base & *supplementary units*			
length	m	meter	
mass	kg	kilogram	g gram allowed
time	s	second	
plane angle	rad	radian	
solid angle	sr	steradian	
temperature	K	kelvin	
electric current	A	ampere	
amount of substance	mol	mole	
luminous intensity	cd	candela	
IAU-recognized derived units			
frequency	Hz	hertz	s^{-1}
energy	J	joule	N m
power	W	watt	J s^{-1}
electric potential	V	volt	J C^{-1}
force	N	newton	kg m s^{-2}
pressure, stress	Pa	pascal	N m^{-2}
electric charge	C	coulomb	A s
electric resistance	Ohm	ohm	V A^{-1}
electric conductance	S	siemens	A V^{-1}
electric capacitance	F	farad	C V^{-1}
magnetic flux	Wb	weber	V s
magnetic flux density	T	tesla	Wb m^{-2}
inductance	H	henry	Wb A^{-1}
luminous flux	lm	lumen	cd sr
illuminance	lx	lux	lm m^{-2}

Table 4: Additional allowed units.

Quantity		Unit	Meaning	Notes

plane angle		deg	degree of arc	π/180 rad
		arcmin	minute of arc	1/60 deg
		arcsec	second of arc	1/3600 deg
		mas	milli-second of arc	1/3 600 000 deg
time		min	minute	60 s
		h	hour	60 min = 3600 s
		d	day	86 400 s
	†	a	year (Julian)	31 557 600 s (365.25 d), peta a(Pa) forbidden
	†	yr	year (Julian)	a is IAU-style
energy∗	†	eV	electron volt	$1.6021765 \times 10^{-19}$ J
	‡	erg	erg	10^{-7} J
		Ry	Rydberg	13.605692 eV
mass∗		solMass	solar mass	1.9891×10^{30} kg
		u	unified atomic mass unit $1.6605387 \times 10^{-27}$ kg	
luminosity		solLum	Solar luminosity	3.8268×10^{26} W
length	‡	Angstrom	angstrom	10^{-10} m
		solRad	Solar radius	6.9599×10^{8} m
		AU	astronomical unit	1.49598×10^{11} m
		lyr	light year	9.460730×10^{15} m
	†	pc	parsec	3.0857×10^{16} m
events count		count		
		ct	count	
		photon	photon	
		ph	photon	
flux density	†	Jy	jansky	10^{-26} W m^{-2} Hz^{-1}
	†	mag	(stellar) magnitude	
	†	R	rayleigh	$10^{10}/(4\pi)$ photons m^{-2} s^{-1} sr^{-1}
magnetic field	†‡	G	gauss	10^{-4} T
area pixel (image/detector)		pixel		
		pix	(image/detector) pixel	
	†‡	barn	barn	10^{-28} m
Miscellaneous units				

		D	debye	⅓ × 10^{-29} C.m
		Sun	relative to Sun	e.g., abundances
		chan	(detector) channel	
		bin	numerous applications (including the one-dimensional analog of pixel)	
		voxel	three-dimensional analog of pixel	
	†	bit	binary information unit	
	†	byte	(computer) byte	eight bits
		adu	Analog-to-digital converter	
		beam	beam	area of observation as in Jy/beam

Notes.

(†)Addition of prefixes for decimal multiples and submultiples are allowed.

(‡)Deprecated in IAU Style Manual (McNally 1988) but still in use.

(∗)Conversion factors from CODATA Internationally recommended values of the fundamental physical constants 2002 (http://physics.nist.gov/cuu/Constants/).

Table 5: Prefixes for multiples and submultiples.

Submult	Prefix	Char		Mult	Prefix	Char
10^{-1}	deci	d		10	deca	da
10^{-2}	centi	c		10^{2}	hecto	h
10^{-3}	milli	m		10^{3}	kilo	k
10^{-6}	micro	u		10^{6}	mega	M
10^{-9}	nano	n		10^{9}	giga	G
10^{-12}	pico	p		10^{12}	tera	T
10^{-15}	femto	f		10^{15}	peta	P
10^{-18}	atto	a		10^{18}	exa	E
10^{-21}	zepto	z		10^{21}	zetta	Z
10^{-24}	yocto	y		10^{24}	yotta	Y

8 Appendix C

The IVOA standard on units may be found at
https://www.ivoa.net/documents/VOUnits/20140523/index.html
Following are Sections 1 and 2 of the IVOA standard.

Units in the VO
Version REC-1.0

IVOA Recommendation 1.0

This version:
http://www.ivoa.net/Documents/VOUnits/20140523/

Latest version:
http://www.ivoa.net/Documents/VOUnits/

Previous versions:
http://www.ivoa.net/Documents/VOUnits/20140513/

Editor(s):
Sébastien Derrière and Norman Gray

Authors:
Markus Demleitner
Sébastien Derrière
Norman Gray
Mireille Louys
François Ochsenbein

code.google.com/p/volute, rev2670, 2014-07-02 19:42:18 +0100 (Wed, 02 Jul 2014)

Contents

List of Tables

Abstract

This document describes a recommended syntax for writing the string representation of unit labels (‘VOUnits’). In addition, it describes a set of recognised and deprecated units, which is as far as possible consistent with other relevant standards (BIPM, ISO/IEC and the IAU).

The intention is that units written to conform to this specification will likely also be parsable by other well-known parsers. To this end, we include machine-readable grammars for other units syntaxes.

Status of this document

This document has been produced by the IVOA Semantics Working Group. It has been reviewed by IVOA Members and other interested parties, and has been endorsed by the IVOA Executive Committee as an IVOA Recommendation. It is a stable document and may be used as reference material or cited as a normative reference from another document. IVOA's role in making the Recommendation is to draw attention to the specification and to promote its widespread deployment. This enhances the functionality and interoperability inside the Astronomical Community.

The place for discussions related to this document is the Semantics IVOA mailing list `semantics@ivoa.net`.

A list of current IVOA recommendations and other technical documents can be found at `http://www.ivoa.net/Documents/`.

Note on conformance

Text within the following document is classified as either ‘normative’ or ‘informative’.

Normative text means information that is required to implement the Recommendation; an implementation of this Recommendation is conformant if it abides by all the prescriptions contained in normative text. **Informative** text is information provided to clarify or illustrate a requirement but which is not required for conformance.

The sections and subsections of this Recommendation are labeled, after the section heading, to specify whether they are normative or informative. If a subsection is not labeled, it has the same normativity as its parent section. References are normative if they are referred to within normative text.

When found within normative sections, the key words **must**, **must not**, **required**, **shall**, **shall not**, **should**, **should not**, **recommended**, **may**, **optional**, thus formatted, are to be interpreted as described in RFC 2119 (Bradner, 1997).

Acknowledgements

We thank all those participants in IVOA and EuroVO workshops who have contributed by exposing use cases and providing comments during this document's long development. These include Bob Hanisch, Rick Hessman, Paddy Leahy, Jeff Lusted, Jonathan McDowell, Marco Molinaro, Pedro Osuna, Anita Richards, Bruno Rino, Arnold Rots, Jesus Salgado, Mark Taylor, Brian Thomas, and recent contributors on the DM and Semantics forums.

1 Introduction (informative)

This document describes a standardised use of units in the VO (hereafter simply 'VOUnits'). It aims to describe a syntax for unit strings which is as far as possible in the intersection of existing syntaxes, and to list a set of 'known units' which is the union of the 'known units' of those standards. We *recommend*, therefore, that applications which write out units should do so using *only* the VOUnits syntax, and that applications reading units should be able to read *at least* the VOUnits syntax, plus all of the units of Sect. 2.4. It is not, however, quite possible for VOUnits to be in the intersection of existing syntaxes; there is futher discussion of this point in Sect. 2.12.1.

We also provide, for information, a set of self- and mutually-consistent machine-readable grammars for all of the syntaxes discussed.

The introduction gives the motivation for this proposal in the context of the VO architecture, from the legacy metadata available in the resource layer, to the requirements of the various VO protocols and standards and applications.

This document is organised as follows. Sect. 2 details the proposal for VOUnits. Sect. 3 lists some use cases and reference implementations. In Appx. A, there is a brief review of current practices in the description and usage of units; in Appx. B there is a detailed discussion of the differences between the various syntaxes; and in Appx. C there are formal (yacc-style) grammars for the four syntaxes discussed.

The normative content of this document is Sect. 2 and Appx. C.4.

1.1 Units in the VO Architecture

Generally, every quantity provided in astronomy has a unit attached to its value or is unitless (e.g., a ratio, or a numerical multiplier).

Units lie at the core of the VO architecture, as can be seen in Fig. 1. Most of the existing data and metadata collections accessible in the resource layer have some legacy units, which are mandatory for any scientific use of the corresponding data. Units can be embedded in data (e.g., FITS headers) or be implied by convention and/or (preferably) specified in metadata.

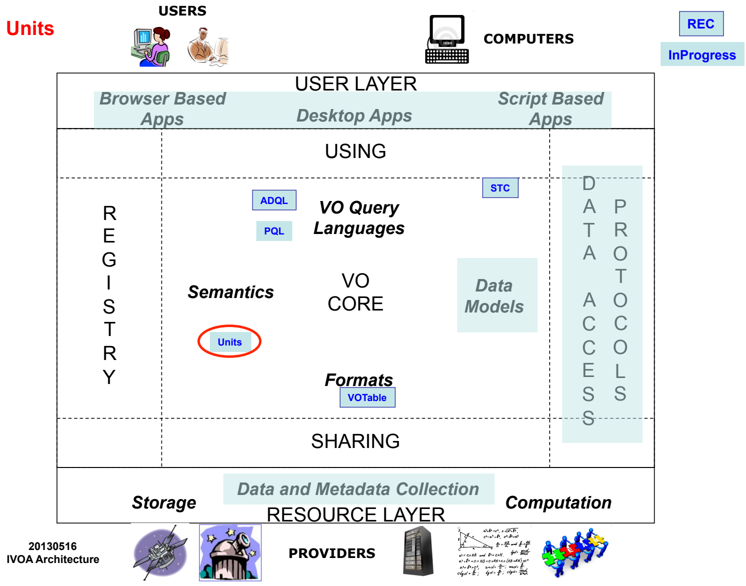

Figure 1: Units is a core building block in the VO. Most parts of the architecture rely on it: the User Layer with tools and clients, the Resource Layer with data. Protocols, registries entries, and data models also re-use these Units definitions.

Units also appear in the VOTable format (Ochsenbein et al., 2011), through the use of a `unit` attribute that can be used in the `FIELD`, `PARAM` and `INFO` elements. Because of the widespread dependency of many other VO standards on VOTable, these standards inherit a dependency on Units.

The Units also appear in many Data Models, through the use of dedicated elements in the models and schemas. At present, each VO standard either refers to some external reference document, or provides explicit examples of the Units to be used in its scope, on a case-by-case basis.

The registry records can also contain units, for the description of table metadata. The definition of VO Data Access protocols uses units by specifying in which units the input parameters have to be expressed, or by restricting the possible units in which some output must be returned.

And last but not least, tools can interpret units, for example to display heterogeneous data in a single diagram by applying conversions to a reference unit on each axis.

1.2 Adopted terms and notations

Discussions about units often suffer from misunderstandings arising from cultural differences or ambiguities in the adopted vocabulary. For the sake of clarity, in this document, the following concepts are used:

A **quantity** is the combination of a (numerical) *value*, measured for a *concept* and expressed in terms of a given *unit*; there may be other structure to a quantity, such as uncertainty or even provenance. In the VO context, the nature of the concept can be expressed with a UCD or a utype. This document does not address the full issue of representing quantities, but focusses on the *unit* part.

A **unit** can be expressed in various forms: in natural language (e.g., *metres per second squared*), with a combination of symbols with typographic conventions (e.g., $\mathrm{m\ s^{-2}}$), or by a simplified text label (e.g., `m.s-2`). VOUnit deals with the label form, which is easier to standardize, parse and exchange. A VOUnit corresponds in the most general case to a combination of several (possibly prefixed) symbols with mathematical operations expressed in a controlled syntax.

A **unit** consists of a sequence of **unit components**, each of which represents a **base unit**, possibly modified by a multiplicative **prefix** (of one or two characters), and raised to an integer or rational power. The whole unit may (in some syntaxes) be prefixed by a numerical **scale-factor**.

Each of the **base units** (for example, the metre) is represented by a **base symbol** (for example `m`). Each syntax has a number of **known units** (Sect. 2.4), for each one of which there is at least one symbol which identifies only that unit.

A **symbol** is either a base symbol or a base symbol with a scaling prefix.

For example, in the unit of `1.663e-1mm.s**-1`, the scale-factor is 1.663×10^{-1}, the two unit-components are `mm` and `s**-1`; the first symbol has base symbol `m` and prefix `m` (for 'milli'), and the second has base symbol `s`, no prefix, and the power -1.

1.3 Purpose of this document

The purpose of this document is to provide a reference specification of how to write VOUnits, in order to maximize interoperability within the VO; the intention is that VOUnit strings should be reliably parsable by humans *and* computers, with a single interpretation. This is broadly the case for the other existing unit-string syntaxes, although there are some slight ambiguities in the specifications of these syntaxes (cf Appx. C). We therefore include a set of self- and mutually-consistent machine-readable grammars for all of the syntaxes discussed.

The unit syntax(es) described here are intended to be human-readable, to the extent that, for example, a string such as `mm.s**-2` is human-readable (without this restriction, we could easily define a much more regular machine-to-machine grammar). Having an explicit unit-string grammar means that data providers can write human-readable strings in the confidence that the result will *additionally* be machine-readable in a reliable and checkable way. Or, where a string is not fully machine readable (because a data provider needs to use a custom unit such as 'jupiterMass'; see Sect. 2.11), that the string is at least partially machine readable, and that that partial readability is non-ambiguous.

We aim not to reinvent the wheel, and to be as compliant as possible with legacy metadata in major archives, and astronomers' habits.

In particular:

- We describe (Appx. A) a number of existing unit syntaxes, and mention some ambiguities in their definition. Application authors should expect to encounter each of the syntaxes mentioned in this document (FITS, OGIP and CDS); all of these are broadly endorsed by this specification.
- In addition to the unit syntaxes described above, there are multiple specifications of base and known units (we refer, in particular, to specifications from BIPM, ISO/IEC and the IAU); these are broadly, but not completely, mutually consistent.
- Where there are some ambiguities in, or contradictions between, these various specifications, we recommend that application authors should resolve them as indicated in this specification.
- This document defines a syntax, called 'VOUnits', which is as far as is feasible in the intersection of the three existing syntaxes, and which we recommend that applications should use when writing unit strings. This aim is not quite possible in fact, and the extensions to it, and the mild deviations from it, are discussed below in Sect. 2 and Appx. C; there is a summary of the various units in Table 2 on page 14.

1.4 What this document will not do

This Recommendation does **not** prescribe what units data providers employ, except to the extent that we avoid giving a standard interpretation for a unit in some cases (for example we do not acknowledge the degree celsius or the century as units). Since we do not forbid 'unrecognised' units, this need not restrict data providers. Nor do we demand that a given quantity be expressed in a unique way (e.g., all distances in `m`). So long as data is labelled in a recognised system, a translation layer can be provided. Data providers can customise the translation tools if required. Depending on preference and the operations required, the user may have a choice of units for his or her query and for the result. In particular, the Recommendation does not require that only recognised units are used. While it is obviously desirable for data providers to use recognised and non-deprecated units where possible, there are occasions when this is unnecessary or undesirable.

This Recommendation does not discuss *quantities* at all. That is, we do not discuss the combination of number and unit which refers to a particular physical measurement, such as '$2\,\mathrm{m\,s^{-1}}$'. Though this might appear to be a trivial extension, it raises questions of the representation of decimal numbers, the representation of uncertainties, questions of unit conversion, and other data-modelling imponderables which have in the past, possibly surprisingly, generated a great deal of discussion within the IVOA without, so far, a generally acceptable resolution.

This Recommendation describes only isolated units, and not arrays, records or other combinations of units. Several VO protocols require embedding complex objects into result tables, and give string serializations for those: geometries in TAP results are the most common example. This specification does not cover this situation, although we hope that where individual unit strings are required in such instances, their syntax will conform to, or include, this specification by reference.

In general, this Recommendation is concerned almost exclusively with the syntactic question of what is and is not a valid unit string, leaving most questions of interpretation or enforcement to a higher layer in an application stack. Specifically:

- The specification does not forbid 'unknown' units. An implementation of this specification should be able to recognise, and communicate, that a unit is unknown, but it is not required to reject a unit string on the grounds that it is unrecognised.
- Similarly, although Table 2 on page 14 forbids some units from having SI prefixes, a VOUnit implementation should not itself reject a unit string which incorrectly includes a prefix, but should instead just make available the information that this has been detected, and that it is deprecated.
- The list of known units in Sect. 2.4 is not specific about the precise definitions of the units in question; for example, it refers to the 'second' without distinguishing between the various possible definitions that the second may have. See that section for further discussion.
- This Recommendation does not specify how an application should compare units for equivalence; for example, an application may or may not wish to deem `m/s` and `km/s` to be 'equivalent'. This Recommendation, similarly, does not specify how to compare units with scale-factors (cf Sect. 2.6).

2 The VOUnits syntax (normative)

The rules for VOUnits are defined in this section. Various aspects are addressed:

- how the labels are encoded;

- what base symbols are allowed and how they are spelled;
- what prefixes are allowed and how they are used;
- how symbols are combined.

A formal grammar summarizing these conventions is given in Appx. C.4.

The text below is expected to be compatible with the prescriptions of the SI standard (BIPM, 2006), except where noted.

2.1 String representation and encoding

VOUnits may occur in legacy contexts, in which the presence of non-ASCII characters may cause considerable technical inconvenience (for example FITS cards). There are only a few non-ASCII characters which we might wish to include in unit strings (for example Å or µ), and we can find substitutes for these sufficiently easily, that we feel there is little real benefit in permitting non-ASCII characters in VOUnit strings.

All the VOUnit characters in the specification below are printable ASCII characters (that is, in the range hexadecimal 20 to 7E); any extensions to this standard **should** be restricted to this same range.

All VOUnit strings **must** be regarded as case-sensitive (the strings in the other syntaxes are also case-sensitive).

2.2 Parsing unit strings – overview

The unit strings `unknown` and `UNKNOWN` (that is, in all-lowercase or all-uppercase) are reserved for cases when the unit is unknown; that is, it is known that there should be a unit, but the unit string has been lost or not been specified. These strings are not, however, part of the list of known units or the VOUnits grammar, and applications **must** check for their presence before unit parsing.

An empty unit string positively indicates that the corresponding quantity is dimensionless. Since an empty string does not conform to the grammars below, this also **must** be checked for before unit-parsing starts.

A **symbol** within a unit-component **should** be parsed as follows:

1. If it corresponds to a known **base symbol**, then it **must** be recognised as such (for example the `Pa` must be parsed as the known Pascal, and never as the peta-year).
2. If the symbol starts with a multiplicative prefix, then this is recognised independently of whether the resulting base symbol is a known or unknown unit – thus `Mm` and `Mfurlong` are parsed as millions of metres and furlongs, but note that this implies, for the sake of consistency, that `furlong` is parsed as the femto-'urlong'.
3. In the VOUnits syntax (a significant divergence from the other syntaxes), base symbols **may** be put between single quotes `'...'` (ASCII

character 27_{16}). Such symbols **must** be parsed as unrecognised unit symbols which are not further examined. See Sect. 2.11 for discussion.

A library which implements this specification **should** be able to distinguish known and unknown units, and identify deviations from the restrictions on their use, below. It **should** be able to communicate such information to a caller, but it **should not** unilaterally reject unit strings which use unknown units or use known units in disapproved ways (of course, a higher-level application is free to reject unit strings for any reason it pleases).

2.3 Base units

There is good agreement for the base symbols across the different schemes (see Table 10 on page 28).

The VOUnits base symbols are listed in Table 1

`m`	(metre)	`g`	(gram)	`J`	(joule)	`Wb`	(weber)
`s`	(second)	`rad`	(radian)	`W`	(watt)	`T`	(tesla)
`A`	(ampere)	`sr`	(steradian)	`C`	(coulomb)	`H`	(henry)
`K`	(kelvin)	`Hz`	(hertz)	`V`	(volt)	`lm`	(lumen)
`mol`	(mole)	`N`	(newton)	`S`	(siemens)	`lx`	(lux)
`cd`	(candela)	`Pa`	(pascal)	`F`	(farad)	`Ohm`	(ohm)

Table 1: VOUnits base units

For masses, the SI unit is `kg`. However, existing specifications recommend not using scale-factors with `kg`, but attaching them only to `g` instead.

Recognising a known unit takes priority over parsing for prefixes. Thus the string `Pa` represents the Pascal, and not the peta-year, and the string `mol` will always be the mole, and never a milli-'ol', for some unknown unit 'ol'.

2.4 Known units

In Table 2 on page 14, we indicate the 'known units' for each of the described syntaxes, which go beyond the physically motivated set of base units. There are a few units (namely `angstrom` or `Angstrom`, `pix` or `pixel`, `ph` or `photon` and `a` or `yr`) for which there are recognised alternatives in some syntaxes, and in these cases 'p' marks the preferred one.

This list of known units is not specific about the precise definitions of the units in question; for example, it refers to the 'second' without distinguishing between the various possible definitions that the second may have: they may be mean-solar or atomic seconds, and be defined at different points in spacetime. Generally, when data is exchanged in those areas where such distinctions matter – such as data connected with pulsar timings – the fine semantic details will be indicated by the data provider through other mechanisms. That said, a VOUnits processor must interpret the symbols of Table 2 on the following page compatibly with the indicated units: a `m` is always a metre of one type or another, and may not be interpreted as, for example, a minute.

Unrecognised units ***should*** *be accepted by parsers*, as long as they are parsed giving preference to the syntaxes and prefixes described here. Thus, for example, the string `furlong/week` **should** parse successfully (though perhaps with suitably prominent warnings) as the femto-'urlong' per week.

The Unity library (Sect. 3.2) recognises units with respect to a subset of the QUDT unit framework Hodgson et al. (2013), with some astronomy-specific additions. This is a particularly comprehensive collection of units, and we commend it to the IVOA community as a *lingua franca* for this type of work.

Sections 2.5 to 2.8 below, discussing the set of known units, are longer than one might expect would be necessary. Most of the discussion concerns rather arcane edge-cases, or attempts to reconcile the minor deviations between the relevant existing standards. In all cases, we have attempted to be as uninnovative and unsurprising as possible.

Future versions of this specification may add to the set of known units.

2.5 Binary units

The symbol 'b' is sometimes used for 'bits', but this is the SI symbol for 'barn', and this Recommendation aligns with the SI standard in this respect. Since the same symbol is sometimes used for 'bytes', it is probably best avoided in any case.

ISO/IEC 80000-13, item 13-9.c notes that the term 'byte' 'has been used for numbers of bits other than eight' in the past, but that it should now always be used for eight-bit bytes; we recommend the same interpretation here. The same source notes the theoretical confusion between the symbol `B` for 'byte' and for 'Bel'. We believe it would be perverse in our present context to recommend against using 'B' for byte, and resolve this here in favour of 'byte' by mandating that `B` **must** be parsed as indicating the 'byte', that the `dB` is an unprefixable special-case unit (as discussed below), and by implication that the 'dB' **must not** be interpreted as a tenth of a byte.[1]

2.6 Scale factors

Units **may** be prefixed by any of the 20 SI scale-factors, and a subset **may** be prefixed by the eight binary scale-factors. The SI scale-factors – provided in Table 3a – are the same as those of BIPM (2006), of ISO/IEC 80000-1,

[1]We have no evidence that this has been a common source of confusion within the IVOA, or indeed anywhere else.

unit	*description*	*fits*	*ogip*	*cds*	*vou*
%	percent			·	
A	ampere	s	s	s	s
a	julian year	s		s	s
adu	ADU	·			s
Angstrom	angstrom	d		·	dp
angstrom	angstrom		·		d
arcmin	arc minute	·	·	·	s
arcsec	arc second	·	·	s	s
AU	astronomical unit	·	·	·	p
au	astronomical unit				·
Ba	besselian year	d			d
barn	barn	sd	·	s	sd
beam	beam	·			s
bin	bin	·	·		s
bit	bit	s		s	sb
byte	byte	s	·	s	sbp
B	byte				sb
C	coulomb	s	s	s	s
cd	candela	s	s	s	s
chan	channel	·	·		s
count	number	·	·		sp
Crab	crab		s		
ct	number	·		·	s
cy	julian century	·			
d	day	·	·	·	s
dB	decibel				·
D	debye	·		·	s
deg	degree (angle)	·	·	·	s
erg	erg	d	·		sd
eV	electron volt	s	s	s	s
F	farad	s	s	s	s
g	gramme	s	s	s	s
G	gauss	sd	·		sd
H	henry	s	s	s	s
h	hour	·	·	·	s
Hz	hertz	s	s	s	s
J	joule	s	s	s	s
Jy	jansky	s	s	s	s
K	kelvin	s	s	s	s
lm	lumen	s	s	s	s
lx	lux	s	s	s	s
lyr	light year	·	·		s
m	meter	s	s	s	s
mag	magnitudes	s	·	s	s
mas	milliarcsecond	·		·	·
min	minute (time)	·	·	·	s
mol	mole	s	s	s	s
N	newton	s	s	s	s
Ohm	ohm	s		s	s
ohm	ohm		s		
Pa	pascal	s	s	s	s
pc	parsec	s	s	s	s
ph	photon	·			s
photon	photon	p	·		sp
pix	pixel	·		·	s
pixel	pixel	p	·		sp
R	rayleigh	s			s
rad	radian	s	s	s	s
Ry	rydberg	·		s	s
s	second (time)	s	s	s	s
S	siemens	s	s	s	s
solLum	luminosity	·		·	s
solMass	solar mass	·		·	s
solRad	solar radius	·		·	s
sr	steradian	s	s	s	s
T	tesla	s	s	s	s
ta	year tropical	d			d
u	AMU	·			s
V	volt	s	s	s	s
voxel	voxel	·	·		s
W	watt	s	s	s	s
Wb	weber	s	s	s	s
yr	julian year	sp	·	sp	sp

Table 2: Known units in the various syntaxes. In the table, and for a given syntax, a '·' indicates that the unit is recognised, an 's' that it is additionally permitted to have SI prefixes, a 'b' that binary prefixes will be recognised, and a 'd' that it is recognised but deprecated. For those units which have alternative symbols for a given unit, a 'p' indicates the preferred one.

da	deca, 10^1	d	deci, 10^{-1}
h	hecto, 10^2	c	centi, 10^{-2}
k	kilo, 10^3	m	milli, 10^{-3}
M	mega, 10^6	u	micro, 10^{-6}
G	giga, 10^9	n	nano, 10^{-9}
T	tera, 10^{12}	p	pico, 10^{-12}
P	peta, 10^{15}	f	femto, 10^{-15}
E	exa, 10^{18}	a	atto, 10^{-18}
Z	zetta, 10^{21}	z	zepto, 10^{-21}
Y	yotta, 10^{24}	y	yocto, 10^{-24}

Ki	kibi, 2^{10}
Mi	mebi, 2^{20}
Gi	gibi, 2^{30}
Ti	tebi, 2^{40}
Pi	pebi, 2^{50}
Ei	exbi, 2^{60}
Zi	zebi, 2^{70}
Yi	yobi, 2^{80}

Table 3: VOUnits prefixes: (a, left) decimal prefixes; (b, right) binary prefixes

§6.5.4, and of Pence et al. (2010, Table 5) (see also Table 11 on page 29 for further comparisons).

Writers of unit strings **must not** use compound prefixes (that is, more than one SI prefix). Prefixes are concatenated to the base symbol without space, and **must not** be used without a base symbol.

The SI prefixes of Table 3a ***must** always refer to multiples of 1000*, even when applied to binary units such as bit or byte; this follows the stipulations (and clarifying note) of BIPM (2006, §3.1), and of ISO/IEC 80000-1, §6.5.4. If data providers wish to use multiples of 1024 (ie, 2^{10}) for units such as bytes or bits, they **must** use the the binary prefixes of ISO/IEC 80000-13, §4, reproduced in Table 3b (these originated in IEEE 1541).

Note: the ‘s’ and ‘b’ annotations in Table 2 on the preceding page are not symmetric: the ‘s’ annotation indicates that SI prefixes are permitted in the given syntax, which means that they are also recognised when preceding unknown units (which have no restrictions on them); in contrast, binary prefixes are recognised exclusively on units with a ‘b’ annotation, which means that they are *not* recognised with unknown units. That is, the `Mifurlong` is the mega-`ifurlong` and the `Kifurlong` is the unknown unit `Kifurlong`.

Note: The letter `u` is used instead of the μ symbol to represent a factor of 10^{-6}, following the character set defined in Sect. 2.1.

2.7 Astronomy symbols

Table 12 on page 30 lists symbols used in astronomy to describe times, angles, distances and a few additional quantities. The subset of these used by this specification are listed in Table 4.

Minutes, hours, and days of time **must** be represented in VOUnits by the

`min`	(minute of time)	`deg`	(degree of angle)	`Jy`	(jansky)
`h`	(hour of time)	`arcmin`	(arcminute)	`pc`	(parsec)
`d`	(day)	`arcsec`	(arcsecond)	`eV`	(electron volt)
`a`, `yr`	(year)	`mas`	(milliarcsecond)	`AU`	(astronomical unit)
`u`	(atomic mass)				

Table 4: Additional astronomy symbols

symbols `min`, `h` and `d`; however the `cd` is the candela, not the centi-day.[2] The year **may** be expressed by `yr` (common practice), or `a`, as recommended by ISO (ISO/IEC 80000-3, Annex C) and the IAU (IAU Commission 5, 1989, Table 6). However peta-year must only be written `Pyr`, to avoid the collision with the pascal, `Pa`.

There are no VOUnit symbols for degrees celsius or century. Temperatures are expressed in kelvin (`K`), and a century corresponds to `ha` or `hyr`. Note that *this is a mild deviation from the SI standard*, which states that the 'hectare', with unit symbol `ha`, is a 'non-SI unit accepted for use' as a measure of land area (BIPM, 2006, table 6), and which acknowledges neither 'a' nor 'yr' as a symbol for year.[3]

The astronomical unit **should** be expressed in upper-case, `AU`, in order to follow legacy practice. It may also be written `au`, in the VOUnits syntax, on the ground that it would be perverse to prefer the atto-atomic-mass to the astronomical unit, in an astronomical unit specification. *This is a deviation* from the SI recommendation of 'ua' (BIPM, 2006, Table 7), but conformant with the IAU's recommendation of 'au' (IAU Division I, 2012).[4]

Because of the near-degeneracy between the decimal prefixes `d` and `da`, there is an ambiguity when parsing the unit `dadu` – is this the deka-`du` or the deci-`adu`? The only cases where this ambiguity is possible are those involving known units starting with 'a' (`da` is unambiguously a deci-year for the same reason that `d` is unambiguously a day, because the presence of a bare unit prefix would be ungrammatical). We can think of no cases where the prefix is useful enough that resolving the ambiguity is worth the specification effort, so we deem the parse of `da.*` to be **unspecified**. In consequence, data providers **must not** use the `da` prefix, and **should not** use the `d` prefix (as noted in Sect. 2.8, the decibel, `dB` is listed as a 'known unit', as opposed to a deci-Bel).

[2] We therefore rule out interpreting dB/cd as 0.9 mbit/s.

[3] If large telescope arrays feel they must talk of attojoules per hectare per century, for some reason, they're going to have to be careful how they do so; it's probably best not to even think about atto-Henrys.

[4] If you feel a burning desire to write about micro-years or atto atomic-mass, this document is not the place you need to look for help.

2.8 Other symbols, and other remarks

Table 13 on page 31 corresponds to Table 7 in the IAU document, and the IAU strongly recommends no longer using these units. Data producers are strongly advised to prefer the equivalent notation using symbols and prefixes listed in Tables 10, 11 and 12.

However, in order to be compatible with legacy metadata, VOUnit parsers **should** be able to interpret symbols `angstrom` or `Angstrom` (for ångström), `barn`, `erg` and `G` (for gauss).

Table 14 on page 32 compares other miscellaneous symbols. The last set of VOUnits symbols, derived from this comparison, is in Table 5.

`mag` (magnitude)	`pix` or `pixel`	`solMass` (solar mass)	`R` (rayleigh)
`Ry` (rydberg)	`voxel`	`solLum` (solar luminosity)	`chan` (channel)
`lyr` (light year)	`bit`	`solRad` (solar radius)	`bin`
`ct` or `count`	`byte` (8 bits)	`Sun` (relative to the Sun, e.g. abundances)	`beam`
`ph` or `photon`	`adu`	`D` (Debye)	`unknown` (Sect. 2.2)

Table 5: Miscellaneous VOUnits.

A few symbols which might theoretically be ambiguous are listed in Table 6, with their consensus VOUnit interpretation.

VOUnit	Correct interpretation	Incorrect
`Pa`	pascal	peta-year
`ha`	hecto-year	hectare
`cd`	candela	centi-day
`dB`	decibel	deci-byte
`B`	byte	bel
`au`	astronomical unit	atto-atomic-mass

Table 6: Possibly ambiguous units

It can be noted that some of the units listed in Table 14 on page 32 are questionable. They arise in fact from a need to describe quantities, when the only piece of metadata available is the unit label. Count, photon, pixel, bin, voxel, bit, byte are concepts, just as apple or banana. The associated quantities could be fully described with a UCD, a value and a void unit

`str1.str2`	Multiplication
`str1/str2`	Division
`str1**expr`	Raised to the power expr
`fn(str1)`	Function applied to a unit string

Table 7: Combination rules and mathematical expressions for VOUnits. See Appx. C.4 for the complete grammar.

label. It is possible to count a number of bananas, or to express a distance measured in bananas, but this does not make a banana a reference unit.

The FITS document provides the most general description of all the compared schemes, and VOUnits adopts similar definitions, for the sake of legacy metadata. The VOUnits symbol for magnitudes is `mag`. Note that all symbols like `count`, `photon`, `pixel` are always used in lower case and singular form.

The decibel, `dB` is listed in the SI specification (BIPM, 2006, Table 8) amongst a set of 'other non-SI units', and mentioned by ISO/IEC 80000-3, §0.5 in a 'Remark on logarithmic quantities'. The `dB` is included in the list of 'known units' of Table 2 on page 14 and so **must** be parsed as a unit by itself – as opposed to being parsed as the prefix 'd' qualifying the unit 'Bel' – and both the decibel and Bel **must not** be used with other scaling prefixes.

If there is no unit associated with a quantity (for example a quantity that is a character string, or unitless), data providers **should** indicate this with an empty string rather than blanks or dashes.

2.9 Mathematical expressions containing symbols

Table 15 on page 33 summarizes how, in the various existing syntaxes, mathematical operations may be applied on unit symbols for exponentiation, multiplication, division, and other computations.

The combination rules are where the largest discrepancies between the different schemes appear. The FITS document discusses the problem of trying to best accommodate the existing schemes (Pence et al., 2010, §4.3.1), without really resolving the problem. This and other ambiguities are discussed in the detailed syntaxes of Appx. C.

VOUnits follow a subset of the FITS rules, as summarized in Table 7.

As illustrated in Table 7, units may include a limited set of functional dependencies on other units. The set of functions recognised within VOUnits is the same as the set recommended by FITS, and listed in Table 8. As with unrecognised units, *parsers **should** accept unrecognised functions without error*, even if they deprecate them at some later processing stage. As described in Sect. 2.11, functions may be quoted to indicate that they **must not** be

`log(str1)`	Common Logarithm (to base 10)
`ln(str1)`	Natural Logarithm
`exp(str1)`	Exponential (e^{str1})
`sqrt(str1)`	Square root

Table 8: Functions of units.

interpreted as in this table. Note that since functions such as 'log' require dimensionless arguments, when a quantity x is (for example) represented by numbers labelled with units `log(Hz)`, that indicates that the numbers are related to x by the function $\log(x/(1\,\mathrm{Hz}))$.

2.10 The numerical scale-factor

A VOUnits unit string **may** start with a numerical scale-factor to indicate a derived unit. For example, the inch might appear as the unit of `25.4mm`. See Appx. C.4 for the syntax of the VOUnits numerical string.

A data provider may choose to use such a unit in order to represent a unit which is not listed as one of the VOUnit 'known units'. For example, given a VOTable column of masses relative to Jupiter's mass, one might label it as having units of `1.898E27kg` rather than `'jupiterMass'` (an 'unknown unit'). The *advantage* of doing so is that the data consumer can translate the column data into well-known physical units without further information, and the data source is thus self-contained. The *disadvantage* of doing so is (i) that the intention might be obscured (this is a type of provenance information); and (ii) that the measurements may be relative to (in this example) the actual mass of Jupiter rather than merely expressed in those terms, so that the measurements should change if the actual mass were to be refined as a result of a recalibration, or if (in the case of a pulsar period for example) the unit were time-varying. The data provider retains the choice of which strategy to take.

A data provider may need to provide further metadata information, to clarify the meaning of such a unit, or they may judge that the meaning is adequately clear to the intended audience, without further complication. Such further information is out of scope of the VOUnits Recommendation, in the same way that even 'known' units may be ambiguous in some contexts (cf, Sect. 2.4).

This Recommendation does not prescribe how many significant figures should be in a scale-factor, nor whether it should be interpreted as single- or double-precision, nor how units with scale-factors should be compared for equality. All of these are implementation choices for the software which is handling the units.

2.11 Quoting unknown units

The VOUnits syntax permits the use of 'unknown units' (that is, units not listed in Table 2 on page 14). There need be no syntactic indication that a unit is 'unknown'; this is convenient, but creates some minor ambiguities.

In the VOUnits syntax, base symbols may be put between single quotes `'...'` (a significant divergence from the other syntaxes). Such symbols **must** be parsed as unrecognised unit symbols which are not further examined.

This has two consequences. Firstly, it means that an unknown symbol which happens to start with an SI prefix is not broken into a base symbol and prefix: thus `'furlong'` is parsed as expected, whereas `furlong` would be the femto-'urlong'. Secondly, a quoted symbol is parsed as an unrecognised unit, even if it would otherwise indicate a known unit; thus the unit `'m'` is parsed as an unknown unit 'm', and does not indicate the metre.

This facility means that a data provider may label data with units of, for example, `'martianDay'` or the `'B'`, while still remaining conformant with the VOUnits Recommendation, and without risking the leading `m` being misparsed as an SI prefix, or the 'B' being misparsed as a 'byte'.

Quoted units can take prefixes (they are 'unknown units', so there are no restrictions on their usage), so that `m'furlong'` is a milli-furlong, and `m'm'` is a milli-'m'. The only permissible prefixes are those of Table 3.

2.12 General rationale (informative)

2.12.1 Deviations from other syntaxes

The aspiration of the VOUnits work was that the syntax should be as much as possible in the intersection of the various pre-existing syntaxes, so that a unit string which conformed to the VOUnits syntax would be parsable in each of those other syntaxes. This has not been possible in fact, for four reasons.

1. The CDS syntax permits only a dot to indicate a product, and the OGIP syntax only a star, while FITS permits both. The VOUnits syntax uses a dot, so that non-trivial OGIP unit strings are therefore necessarily invalid VOUnits strings in this one respect.
2. The VOUnits syntax permits (but does not require) a scale-factor at the beginning of the string, which is not a power of 10. Only the CDS syntax permits a similar factor. See Sect. 2.10 for discussion.
3. Only the VOUnits syntax permits quoted units.
4. Only the VOUnits syntax permits the use of the binary prefixes of Table 3.

The first is both unavoidable in specification, and largely unavoidable in practice; the others are VOUnit extensions which a data provider may of course decline to take advantage of.

The scale-factor and quoted-units extensions are intended to support the case where the data provider wishes to distribute data including a unit which is 'unknown', but which the provider nonetheless feels is necessary or useful; this should be done only after weighing the considerations of Sects. 2.10 and 2.11. For the sake of consistency, and in order to allow constructions such as `M'jupiterMass'`, the grammar permits quoted units to take scaling prefixes; this is not often likely to be a good idea.

A VOUnits string which avoids the three extensions above will be parsable, with the same meaning, in the CDS and FITS syntaxes, and will be parsable by an OGIP parser if dots are replaced by stars.

2.12.2 Restrictions to ASCII

As described above, VOUnit unit strings are restricted to printable ASCII characters. While the two most prominent uses of these strings will be within VOTable attributes (`unit="..."`) and in XML serialisations of a data model (for example `<unit>...</unit>`), we also intend them to be usable within FITS files and within databases. Neither of the latter two contexts is necessarily unicode-friendly, so permitting non-ASCII characters in a unit string (such as Å or µ) is more likely than not to cause trouble.

Similarly, forbidding spaces within VOUnit strings removes one (minor) complication when recognising them in use.

2.12.3 Other units, and unit-like expressions

As noted above, the VOUnits syntax does not include structures such as arrays or tuples of numbers. We include in this category sexagesimal coordinates, calendar dates (in ISO-8601 form or otherwise), RA-Dec pairs, and other structured quantities serialised as strings. Each of these is well-specified elsewhere, and would require a separate parser if encountered in data.

Existing VO standards already recommend that coordinates be expressed in decimal degrees.

Quantities like the Modified Julian Date (MJD) are also not recognized VOUnits. As described in Sect. 1.2, the quantity MJD can be seen as a concept (described by the appropriate UCD or utype), and the corresponding value will most likely be expressed in days, so the VOUnit will be `d`. There is no need to overload VOUnits to incorporate the description of concepts themselves.

The notion of unit conversion and quantity manipulation is discussed in Sect. 3.3.