1 Units (V2.0 2004 Mar 9)

The VO may or may not decide to standardize internally on specified units (cf. FITS requiring SI units and degrees). In any case, astronomers, and therefore the VO software interfaces to astronomers, will want to support unit conversions.

In this document I present a simple data model for units.

1.1 Base, compound and derived units

A **system of units** is based on a set of atomic **base units**. In the SI system, these are kg, m, s, A, cd, sr. In the Planck system, there is only one base unit, typically GeV. Other systems include the astronomical system of pc, Msun, yr, etc. and the legendary furlong-elephant-fortnight system (based on the international standard elephant kept in a very large glass jar in Paris). Our data model should not prescribe a set of base units; rather we need to provide a mechanism for specifying the base units desired (both defaults and on a per-user basis).

We make compound units by combining existing units in several ways:

- Exponentiation: e.g m^{-3}; possibly by a fractional power e.g $\mathrm{Hz}^{-1/2}$.
- Prefixing with SI multiples: e.g $\mathrm{cm} = 10^{-2}\mathrm{m}$.
- multiplication by arbitrary constant: e.g. $\mathrm{unit} = 4.2 \times 10^{-12}\mathrm{erg\ cm}^{-2}\mathrm{s}^{-1}$.
- product of units: e.g. $\mathrm{kg\ m}^{-2}$.

Derived units associate a new unit name with a compound unit, e.g.

$$M_{\odot} = 1.989 \times 10^{30}\mathrm{kg}$$

or

$$\mathrm{Jy} = 10^{-26}\mathrm{W\ m}^{-2}\mathrm{Hz}^{-1}$$

A **simple unit** is either a base unit or a derived unit, possibly with an SI prefix, raised to some exponent.

The general compound unit is then the concatenation (product) of some number of **simple units** and a scaling factor:

$$k \prod_{i=1}^{n} p_i u_i^{e_i}$$

where k is an overall numerical factor, u_i is a named unit (base or derived), e_i is the exponent of unit u_i, and p_i is the SI prefix for unit i.

For example, the unit

$$10^{-14}\mathrm{erg\ cm}^{-2}s^{-1}\mathring{A}^{-1}$$

consists of the scale factor 10^{-14}, and the product of four simple units. The first simple unit, erg, is a derived unit in the CGS and SI base systems. The second simple unit, cm^{-2}, is a power of the simple unit cm, which in the SI system is an SI multiple of the base unit m, but in the CGS system is itself a base unit; in an astronomical system in might be a derived unit defined as a scaling on the parsec.

1.2 Methods

The different operations we can make on compound units are:

- multiply units $U1 * U2$, recognizing SI prefixes
- divide units $U1/U2$
- exponentiate unit to power f, $U1^f$
- resolve to base units
- return number of components n
- return prefix, unit and exponent for component i, $1 <= i <= n$

Example of multiplication:

$$(4.2 \times 10^{-12}\mathrm{erg\ cm^{-2}s^{-1}}) * (\mathrm{mK\ km/s}) = 4.2 \times 10^{-7}\mathrm{erg\ mK\ cm^{-1}s^{-2}}$$

We have to decide how to combine different prefixes. In this case I have assumed that when the same unit appears twice with different prefixes, the first prefix (cm in this case) is retained, and that units with positive exponents appear before units with negative exponents. Alternatively, we could make the rule that all prefixes are removed into the leading unit, giving

$$4.2 \times 10^{-8}\mathrm{erg\ K\ m^{-1}s^{-2}}$$

or

$$4.2 \times 10^{-2}\mu\mathrm{erg\ K\ m^{-1}s^{-2}}$$

This has the disadvantage that if your base unit is prefixed, e.g. cm or kg, it can be stripped of that prefix.

In any case this multiplication operation pays no attention to declarations of base units (so erg is not resolved to its component parts). You don't always want to reduce things to base units (we often want to say erg/s rather than gcm^2s.). We therefore separate out a separate functionality, 'resolve', which depends on the basis units selected.

Example with basis units kg, m, s:

$$\mathrm{resolve}(4.2 \times 10^{-12}\mathrm{erg\ cm^{-2}s^{-1}}) = 4.2 \times 10^{-15}\mathrm{kg\ s^{-3}}$$

Example with basis units Mpc, Msun, yr:

$$\mathrm{resolve}(10^{-6}\mathrm{atoms\ cm^{-3}s^{-1}}) = 0.78\ \mathrm{M_\odot\ Mpc^{-3}yr^{-1}}$$

2 Model

Although one could break down the unit model down to the level for the individual prefixes etc., I believe it is more practical to stop with a very simple model in which the unit is a string of the above form, specifically in text

```
1.4E26 erg cm^-2 s^-1
```

or

```
1.4x10^26 erg cm^-2 s^-1
```

and we implement the four methods described above, with the parser opaque to the model.