A Guide to Computer Dating

Ian Stewart

30/09/2008

8 Whitefield Close
Westwood Heath
Coventry CV4 8GY
UK

In 46 BC the Roman calendar was getting out of synch with the seasons. On the advice of the Greek astronomer Sosigenes, Julius Caesar introduced an extra day into every fourth 'leap' year to make the average length of the year 365 $^{1}/_{4}$ days. Misunderstanding the rule, his priests counted the fourth year of one cycle as the first in the next, so every *third* year became a leap year. The mistake wasn't fully sorted out for fifty years. For all our supposed sophistication, we have not learned from Caesar's priests. The programming language COBOL usually allocates only two digits for year dates — 96 for 1996, for example — so the year 2000 will be treated as 1900 unless the programs are changed. Even though the problem is a known one, it is likely that a lot of business software will go berserk on the eve of the new millennium.

We need not make the same mistake again. About ten years ago Nachum Dershowitz and Edward M. Reingold of the Department of Computer Science, University of Illinois at Urbana-Champaign, decided to develop calendar and diary features for the Unix-based editor GNU-Emacs. Out of this project grew a unique resource: computer code for converting dates from one calendric system to another. The fourteen calendars catered for are the Gregorian, ISO, Julian, Coptic, Ethiopic, Islamic, Persian, Bahá'í, Hebrew, Mayan, French Revolutionary, Chinese, old Hindu, and modern Hindu. Their book *Calendric Calculations*, published by Cambridge University Press, is an absolute goldmine for chronologists.

Calendars vary from culture to culture because they are all attempts to perform the impossible: to rationalise the irrational. Our units for time are based on three distinct astronomical cycles: the day, month, and year. A normal 24-hour *mean solar day* is the period between successive occasions when the Sun is overhead. (One rotation on its axis, relative to the 'fixed stars', takes 23 hours 56 minutes and 4 seconds — but the Earth is also revolving around the Sun, and it takes a further four minutes for the extra rotation to compensate for the Sun's apparent slippage across the sky.) The period between successive new moons is the *mean synodic month*, which lasts 29.530588853 days. The period required for the Sun to retuirn to the same position in its apparent path is the *mean tropical year* of 365.242189 days.

If the month were 29.5 days and the year 365.25, then the Moon would repeat its motion exactly every 59 days (2 ∞ 29.5) and the Sun every 1461 days (4 ∞ 365.25). So every 86199 days — 59∞1461 — the system of Earth, Moon, and Sun would return to precisely the same relative position. A calendar with an 86199-day cycle would remain in step forever — ignoring slow changes to the lengths of the day, month, and year caused by forces such as tidal friction. Unfortunately for calendar designers, the ratios between days, months, and years behave like irrational numbers: they are not expressible as exact fractions. (At least , not using smallish integers: 29530588853/1000000000 would lead to an impractically long cycle.) So in practice the lunar and solar cycles never return to *exactly* the same state at exactly the same time of day.

The day is central to timekeeping because of the day-night cycle, the lunar month is important for many cultures for religious reasons, and the year determines the cycle of the seasons, so a comprehensive calendar has to include all three. In practice most cultures either plump for a solar calendar, and fudge the months, or a lunar calendar, and ignore problems with the seasons. Whatever the choice, the calendar-designer must find practical ways to deal with small cumulative errors, hence the complicated paraphernalia of leap days, months of variable length ('Thirty days hath September...'), and so on. To find out just how complex it can get you should either consult a copy of *Calendrical Calculations* or visit its home page on the World Wide Web:

`http://emr.cs.uiuc.edu/home/reingold/calendar-book/index.html`.

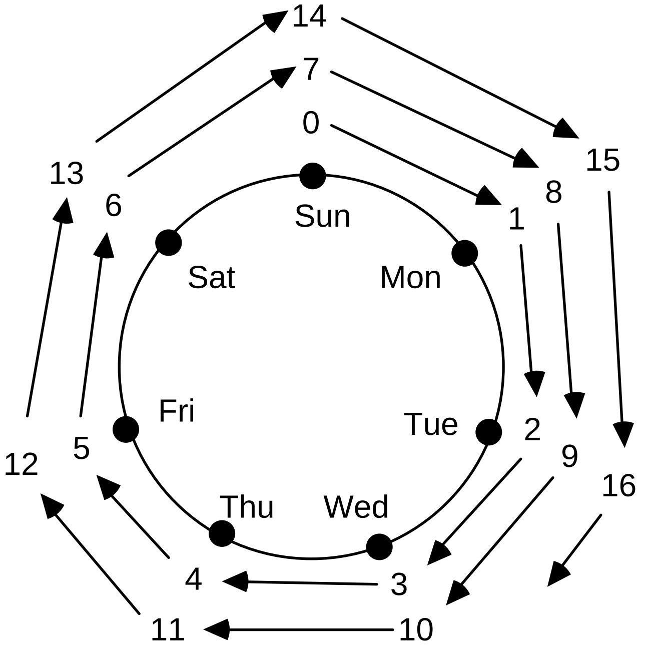
Here I shall try to convey some of its unique and fascinating flavour, but skipping many fine points.

The simplest calendric system would ignore years and months and number consecutive days, choosing some convenient 'epoch' (start day). Astronomers use one such system, the Julian day, but Dershowitz and Reingold prefer an invention of their own: the 'fixed date' or 'rata die', abbreviated to RD. Day 1 of the RD system is January 1 in year 1 of the Gregorian calendar, the calendar we now use. There was no actual year 1 in the Gregorian calendar, since that was introduced in 1582 by Pope Gregory XIII, so we extrapolate backwards. That particular day was a Monday, which is convenient since we can take day 0 to be the previous Sunday and number the days of the week from 0-6 starting on Sunday. *Calendrical Calculations* uses the RD value as a common reference system: for example, to convert a date in the Hebrew calendar to one in the Chinese, you convert from Hebrew to RD and then from RD to Chinese. This way only 28 conversion functions (one each way for each of the 14 calendars) are needed.

Here are two simple warm-up problems which exemplify the type of mathematics required:

1. What day of the week will 1,000,000 RD be?
2. How many complete mean tropical years will elapse between 0 and 1,000,000 RD?

To answer question 1, observe that the days of the week form a repeating cycle of length seven, 'wrapping round' as shown in **Fig.1**.



Finding the day by counting modulo 7.

Therefore any RD that is a multiple of 7 must be a Sunday, any that leaves remainder 1 on division by 7 is a Monday, and so on. We say that the day number is the RD number *modulo* 7. 'Modulo' is Latin for 'to the modulus': x mod 7 means 'find the remainder upon dividing x by 7'. Since 1,000,000 = 7 ∞ 142857 + 1, this remainder is 1 when x = 1,000,000, so 1,000,000 RD is a Monday.

To answer the second question, divide 1,000,000 by 365.242189 to get 2737.9094. This tells us that 1,000,000 RD occurs 2737 complete (mean tropical) years after 0 RD, a number that we find by omitting everything after the decimal point. Mathematically this is performed by the 'floor function' |_ x _|, which is the greatest integer less than or equal to x.

[Note to typesetter: |_ and _| are L-shaped symbols like square brackets [ ] with their tops removed.]

Now consider converting a Gregorian date, such as December 25 1996, to its RD value. Recall Pope Gregory's leap year rule, which makes the average length of a year more accurate: multiples of 4 have an extra day on 29 February, unless they are multiples of 100, however multiple of 400 are also leap years. Dershowitz and Reingold show that this leads to the calculation rule in **BOX 1**. For example, let M = 12, D = 25, Y = 1996. Then (a) = 728175, (b) = 498 - 19 + 4 = 483, (c) = 336, (d) = -1, and (e) = 25. So the RD value of December 25 1996 is 728175 + 483 + 336 -1 + 25 = 729018. As a simple application, the day of the week is therefore 729018 mod 7 = 3, so Christmas 1996 happens on a Wednesday.

==

BOX 1 To find the RD value of month M, day D and year Y Gregorian compute:

(a) 365(Y-1)

(b) |_(Y-1)/4_| - |_(Y-1)/100_| + |_(Y-1)/400_|

(c) |_(367M-362)/12_|

(d) 0 if M _ 2, -1 if M>2 and Y is a leap year, and -2 otherwise

(e) D

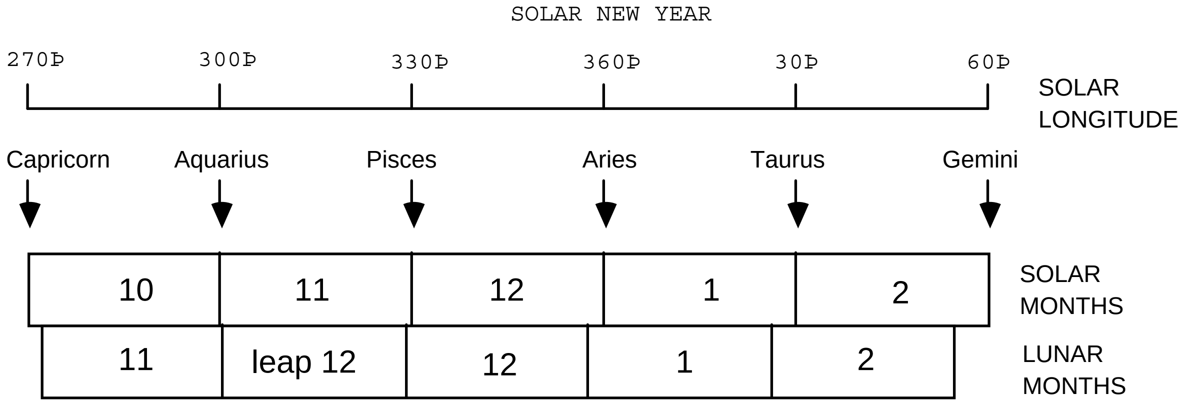
and add them together.

The calculation has the following interpretation: (a) is the number of non-leap days in prior years. (b) is the number of leap days in prior years (one every fourth, except that every 100th is omitted, but you put back every 400th). (c) is a cunning formula for the number of days in prior months of year Y, based on the assumption that February has 30 days, which it doesn't — hence the correction term (d). In step (e) the number D is of course the number of days in the current month — the only days not yet counted.

==

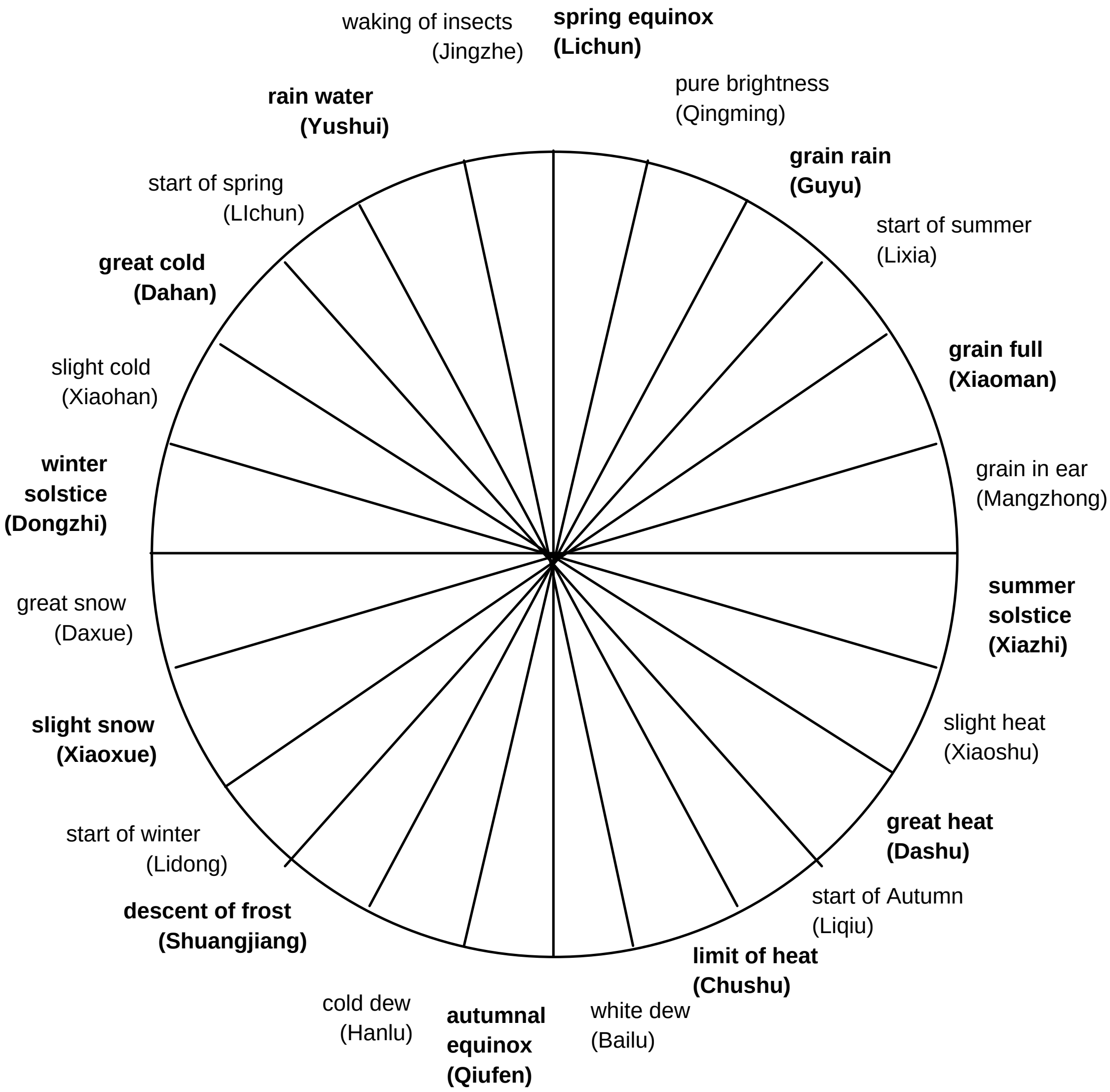
To see the complexity that the software in *Calendrical Calculations* handles with ease, consider the modern Persian calendar. It was adopted in 1925, but its epoch is March 19 622 AD — the vernal equinox prior to the epoch of the Islamic calendar. It is closely based on the more ancient Jalalai calendar invented by a committeee of astronomers that included Omar Khayyam. There are twelve months: the first six (Fravardin, Ordibehest, Xordad, Tir, Mordad, Sahrivar) have 31 days, the next five (Mehr, Aban, Azar, Dey, Bahman) have 30, and the last (Esfand) has 29 in an ordinary year and 30 in a leap year. The leap year pattern, taken unchanged from the Jalali calendar, is highly intricate. It follows a cycle of 2820 years, containing 683 leap years. The 2820 years are divided into twenty-one 128-year subcycles, followed by one of 132. Each 128-year subcycle is divided into sub-subcycles of lengths 29 + 33 + 33 + 33; whereas the 132-year subcycle is divided as 29+33+33+37. Finally, in each sub-subcycle years 5, 9, 13, and so on, going up in fours, are leap years. The Persian calendar is in error by 1.7 minutes at the end of one 2820-year cycle, so it would take 2.39 million years to slip a day relative to the true astronomical cycles!

The old Hindu lunisolar calendar follows a very different pattern. The months follow the Moon's phases closely, and an additional leap month is 'intercalated' to keep the months in step with the solar year. Unlike most such systems, however, the cycle of intercalation does not follow a fixed, simple pattern. The overall structure involves a cycle lasting 1,577,917,500 days. The 'year' (strictly the *arya sideral year*) is one 4320000th of this, or 365.258 days. The *solar month* is one twelfth of a year, and there are twelve named months. The lunar month is one 53433336th of the 1577917500-day cycle, equal to 29.531 days. The basic idea is to run both month-lengths simultaneously. Usually a lunar month overlaps a boundary between solar months, but every so often a lunar month is completely contained in a solar month. In that case, it is considered to be a lunar leap month, and its name is also given to the following lunar month (**Fig.2**).



The Hindu lunisolar calendar. A leap month is added whenever a lunar month fits inside a solar one.

Finally we'll take a look at the Chinese calendar, which is based on astronomical events, not arithmetical rules. The Chinese calendar has been reformed at least fifty times: the version implemented in *Calendric Calculations* is the most recent, dating from 1645, the second year of the Qing dynasty. Months are lunar, beginning at the day of the new moon, and years contain either twelve or thirteen months. The arrangement of the months, however, depends upon the passage of the sun through the signs of the zodiac. The solar year is divided into twelve major solar terms called *zhongqi* and twelve minor solar terms called *jieqi*. Each term corresponds to a 15_ segment of solar longitude, the major ones starting at multiples of 30_ and the minor ones in the gaps between those. These terms occupy roughly the same period in any year (**Fig.3**).



Solar terms of the Chinese year (major terms in **boldface**).

The basic rule that determines the calendar is that the winter solstice *always* occurs during the eleventh month of the year. In a year that contains only twelve complete lunar months, therefore, the months are always numbered 12, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11. In a year that contains thirteen, however, one of the numbers is duplicated in a leap-month. Which? It is the first month that does not contain a major solar term. (Since there are thirteen lunar months and only twelve major solar terms, at least one month must fail to contain a major solar term — an application of the so-called 'pigeonhole principle': if there are more holes than pigeons, then at least one hole must be pigeonless.)

Since present-day calendars are so complex, what of the future? Now the mathematics required is dynamics, plus the sciences of astronomy, physics, climatology... All of the various astronomical cycles are slowly changing their lengths because of tidal gravitational forces. Moreover, there is the 'precession of the equinoxes', which is not steady, but has occasional glitches related to ice ages — so a future calendar must be linked to climate. In fact a future calendar must be interactive, adjusted according to what actually happens, not just based on preset rules, because astronomers Jack Wisdom (MIT) and Jacques Laskar (Bureau des Longitudes, Paris) have discovered that the motion of the solar system is chaotic, so if you set up a fixed calendar to keep pace with the seasons, the infamous 'butterfly effect' will cause it to drift away from reality. Independence Day in 10,000,000 AD may still be the fourth of July, but nobody can predict how many days from now that will be.