fhiso

Proposal for compound calendars to resolve a difficulty with default calendars

Submitted by: Smith, Richard

Created: 2013-04-14

URL: Most recent version: http://fhiso.org/files/cfp/cfps38.pdf
                This version: http://fhiso.org/files/cfp/cfps38_v1-0.pdf

Description: Proposal to allow the default calendar to be dependent on the
                date.

Keywords: dates, calendars, calendar change, Gregorian calendar, Julian
                calendar, English calendar

Abstract

This paper considers the proposal in CFPS 13 to support multiple calendars, and proposes that a document should be able to specify the default calendar that is assumed when no explicit calendar specified on the date. The relevance of calendar changes to the choice of a default calendar is discussed, and it is concluded that appropriate default calendar might depend on the date being interpretted. The notion of a *compound calendar* is introduced as a lightweight means of providing a suitable default calendar when calendar changes are relevant, and an algorithm is provided for parsing dates in an arbitrary compound calendar.

1   Default calendars

Many different calendars are in use throughout the world and a FHISO standard will certainly need to support more than just the Gregorian calendar. A proposal for a general scheme for supporting multiple calendars can be found in CFPS 13 [1]. That paper allows the calendar to be explicitly identified with a calendar ID: for example, 'GR' for the Gregorian calendar. If no calendar ID is given, the Gregorian calendar is assumed.

However for many applications, the Gregorian calendar is not an ideal default. A mediævalist might prefer Julian calendar to be the default; an Israeli, the Hebrew calendar. This paper proposes that a document can override the default calendar. A date is thus interpreted *(i)* using the calendar ID on the date itself, if one is present; otherwise *(ii)* using the default calendar ID specified in the document header, if one is present; and otherwise *(iii)* using the Gregorian calendar. This allows dates to be represented in a briefer, more natural format without repetition of the calendar on every date. This advantage is small and may be outweighed by the disadvantage that dates can no longer be copied directly from the document source without important loosing context, namely the document default calendar.

However, a bigger advantage is that an application can use the default calendar intelligently. An application might only display the calendar name if it is not the document default. Or when a user enters a date, the application can assume it is in the document default calendar unless the user has indicated otherwise. The latter suggestion is particularly valuable as it is exceedingly common to find GEDCOM files containing dates that almost certainly ought to use the Julian calendar, but are not tagged as such and are therefore interpretted as Gregorian.

2 Calendar changes

Which default calendar should an English genealogy use? Until 1752, England used the Julian calendar with years beginning on 25 March; in 1752 it switched to the Gregorian calendar [2]. In a genealogy that includes dates before and after 1752, neither Gregorian nor Julian seems wholly appropriate for the document default. This paper suggests that the default should actually be a third possibility: the English calendar, a calendar that is identical to the Julian calendar for dates up to and including 2 September 1752, and thereafter is identical to the Gregorian calendar. Assuming the YYYY-MM-DD formatting of dates proposed in cfps 13, if the English calendar specified as the document default, '2013-04-14' is unambiguously a Gregorian date and '1066-10-14' is unambiguously Julian. The string '1752-09-10' is an invalid English date in exactly the same way that '2013-02-29' is an invalid Gregorian date.

The situation is different in other countries, and often more complicated. In Scotland, the calendar was changed twice: first in 1600 when the year was deemed to begin on 1 January; and second in 1752 when the Gregorian calendar was adopted. The situation is similar in many other countries. In certain areas, the calendar has changed several times, depending on which country ruled the area at the time. Nova Scotia used the Gregorian calendar from the time of its settlement by the French until 1710 when it was conquered by the British; it then changed to the Julian calendar until 1752 when it reverted to the Gregorian calendar [3].

Nor is the problem confined to the transition from Julian to Gregorian calendars. Eritrea changed from the Ethiopian calendar to the Gregorian one in the (Gregorian) year 1991. Sweden used its own unique calendar between 1700 and 1712 before reverting to the Julian calendar.

Not every calendar change has resulted in unambiguous dates. In Nova Scotia eleven days were repeated (presumably 6–16 October 1710, though in practice it probably varied) following the Siege of Port Royal when the Julian calendar was introduced. In France, the short-lived revolutionary calendar used the year numbers 1–12; however, if revolutionary years are always formated with two digits and years *anno domini* with four, these can be disambiguated.

3 Compound calendars

The term *compound calendar* is used for a calendar, such as the English calendar discussed above, which is comprised of several periods when different calendars are used. Because the fhiso may want to standardise and vendors support many such compound calendars, this paper paper proposes a general mechanism for specifying compound calendars and parsing dates in them. The intention is that

an application can simply read the definitions of the compound calendars and then be able to understand dates writen using them, without having to have dedicated code to handle each one.

Formally a compound calendar is defined a sequence of *periods of coverage*, each of which refers to the calendar ID in use at that time, and optionally the start and end dates of the coverage, expressed in that calendar. An omitted start date means the calendar applies arbitrarily far in the past, and an omitted end date means arbitrarily far in the future.

When parsing a string as a date expressed using a compound calendar, an application must try parsing the date using the calendar of each period of coverage. If the string is syntactically and semanitcally a valid date in that calendar, and is not before the start date (if one exists) nor after the end date (if it exists), then the date is said to lie within that period of coverage. If the string is not a date lying in any period of coverage then the date is invalid according to the compound calendar.

It is recommended that authors of compound calendar specifications take care to avoid allowing strings that can be interpretted as valid dates lying in more than one period of coverage. If an application encounters such a string, it is permitted to assume that each refers to the same underlying date, and if it is able to tell that to be untrue, should treat the date as invalid.

As an example, supposing JU25M is the calendar ID for the Julian calendar with years beginning on 25 March, and that months are still numbered with January as 01, then the following table defines the English calendar, EN. (This paper does not formally propose this calendar because there is currently no proposal for a suitable Julian calendar.)

Calendar ID	Start date	End date
JU25M		1752-09-02
GR	1752-09-14	

References

[1] Tony Proctor, 2013, *Proposal to Accommodate Alternative World Calendar Systems* (CFPS 13), http://fhiso.org/files/cfp/cfps13.pdf

[2] United Kingdom Parliament, 1750, *Calendar (New Style) Act 1750*, http://www.legislation.gov.uk/apgb/Geo2/24/23

[3] Mike Spathaky, 2006, *Old Style and New Style Dates and the change to the Gregorian Calendar: A summary for genealogists*, http://www.cree.name/genuki/dates.htm