Datetime and Related Data Types

# Objectives:

1. Maximum duration – preferably from the Big Bang to billions of years in the future
2. Maximum precision – currently, smallest measurable time is an attosecond (10-18 second). Smallest possible time is 10-43 second.
3. Very fast computations of differences, sums and day to date conversions
4. Full support for leap seconds
5. Full support for time zones – both current and historical
6. Fully compatible relative time data type

# Architecture and Design:

1. The lowest level data type is TAI, which models International Atomic Time. This facilitates fast computations, due to its continuous nature across the entire time range. The SI second is defined as the time for 9,192,631,770 radiation cycles for a Cesium atom. With this definition of a second, the rotation of the earth, which is irregular, will not take precisely 86400 seconds. Currently, the average rotation takes 2 ms longer. As a result, a time standard based on the rotation of the earth to measure days will drift from an ‘atomic day’ by 2 ms / day.
2. The most common time standard based on the rotation of the earth is a variation of Universal Time, called UT1. UT1 is derived from GPS satellites, corrected for seasonal variations in the earth’s rotational speed. Given the unpredictable nature of UT1, it is not practical for designing precision time systems or devices. A more practical time standard, called Coordinated Universal Time, UTC, was developed for this purpose. UTC differs from TAI by an integral number of seconds. When needed, leap seconds, which may be positive or negative, are introduced to keep the difference between UTC and UT1 less than 0.9s. Leap seconds can be introduced at the end of any month, but, to date, have only been introduced at the end of June or December. UTC is somewhat unpredictable for future times and introduces some additional computational complexity due to the discontinuities introduced by leap seconds, but does strike a good balance between the desire for computational simplicity and efficiency, and close adherence to the earth’s rotation. All current time zones are defined in terms of UTC.
3. UTC is the next higher level data type. A corresponding leap second table is required to enable conversions between TAI and UTC. Each row of the table will represent the insertion of a leap second. The table will have 3 columns:
   1. TAI time at which the leap second is inserted
   2. The cumulative number of leap seconds *following* the insertion of the leap second
   3. The sign and number of leap seconds inserted as a signed integer
4. Both TAI and UTC will be represented as the number of ticks from the Big Bang. Each tick represents an attosecond. A representation based on 64-bit integers has two integers, one representing the number of whole seconds and the other representing the remaining fractional second as a number of attoseconds. A representation based on 32-bit integers has four integers, representing the number of gigaseconds, seconds, nanoseconds and attoseconds, respectively.
5. The precise time of the Big Bang is not known. Currently, the age of the Universe is thought to be 4.354 +/- 0.012 x 1017 seconds. This is an uncertainty of 37 million years. Given this uncertainty, the maximum age of the universe is 4.366 x 1017 seconds. We need to select a convenient reference time and define TAI to be 4.366 x 1035 attoseconds at that time.
6. TAI and UTC represent times, independent of a notion of *date*. A *date* has meaning only within the context of a *calendar*.

A *calendar* is a system for structuring time into a series of days, then further structuring days into a series of months, and, finally, structuring months into a series of years. We know that when used to define *time intervals*, a minute denotes a 60 second time interval, while an hour denotes a 60 minute time interval. [*Time Interval* is a distinct data type from *time*, as we will discuss below.] While it is tempting to then define a day as a 24 hour time interval, it is better to resist this urge when endeavoring to understand and define a calendar.

To best understand a calendar, we must first define the notion of *time of day*. *Time of day* is a construct consisting of an hour, minute, second and fraction of a second that counts time within a specified time interval referred to as a day, according to an algorithm that considers daylight savings time transitions and leap seconds. The start of the interval is denoted 00:00:00.0, and referred to as midnight. Ignoring daylight savings time transitions and leap seconds, the count proceeds as expected – after a second, the time of day is 00:00:01.0; 60 seconds later, the time of day is 00:01:01.0; 3600 seconds after that, the time of day is 01:01:01.0. [The elements of time of day are referred to as hours, minutes, seconds and fractions of seconds. While the hour and minute elements are similar to the hour and minute notions used to define *time intervals*, they are subtly different when used to designate a time of day, as we will see when considering leap seconds and transitions to and from daylight savings time.] The count proceeds accordingly until the time of day reaches 23:59:59.0. When another second elapses, the time of day is reset to 00:00:00.0, and the next day begins.

Given the above, a day can be defined as the time interval beginning with time of day, 00:00:00.0 and ending when the time of day resets to 00:00:00.0. While this may seem contrived, we will see that, in fact, a day can only be properly defined in this mechanistic fashion.

Next consider the case where a leap second is inserted within a particular day. The leap second will be inserted at the end of a minute. [In most cases, where the time zone is offset from GMT by an integral number of hours, the leap second will be inserted at the end of an hour.] If we assume the insertion takes place after 11:59:59.0, the time of day representing the next second will be denoted 11:59:60.0, instead of 12:00:00.0. The second after that will be denoted 12:00:00.0. In this case, given our mechanistic definition of a day, the day will have a duration of 86401 seconds! Additionally, the time interval beginning 11:59:00.0 and ending 12:00:00.0, will be 61 seconds long. While we could say this is a minute time interval with a length of 61 seconds, it is clearer and more consistent to state that the time interval from the 11:59:00.0 to 12:00:00.0 time of day is 61 seconds or one minute and one second. In this way, the minute and hour *time intervals* have consistent lengths. Days, months and years are not used as time interval types.

Similarly, there can be a situation where a leap second is removed within a particular day. The leap second is removed at the end of a minute. If we assume the removal takes place in the 11:59 minute, the time of day representing the next second after the 11:59:58.0 second will be denoted 12:00:00.0 – there is no second with the time of day 11:59:59.0. The day will have a duration of 86399 seconds.

Most time zones that implement daylight savings time shift time forward by 1 hour on a day in the spring and back by 1 hour on a day in the fall. If we consider the situation where the shift takes place at 2 a.m., on the spring day, the time of day for the second after the 01:59:59.0 second is 03:00:00. This results in a day with a length of 82800 seconds (or 23 standard 3600 second hours). Similarly, on the fall day, the time of day for the second after 01:59:59.0 second is 01:00:00. The count proceeds as normal until the 01:59:59.0 second is reached a second time, after which the time of day is 02:00:00. The resulting day has 90000 seconds (or 25 standard 3600 second hours).

From the above, we see that while most days have 86400 seconds, days may also have 86399, 86401, 82800 or 90000 seconds. In theory, other day lengths are possible, should there be the requirement for the insertion or removal of multiple leap seconds or were leap seconds to be inserted or removed on a day where a transition to or from daylight savings time takes place, though none of these situations have arisen as of this writing.

We can now define a *date* as the coordinates of a particular day, consisting of the index of the year and the number of the month within that year in which the day is contained, as well as the index of that day within the month. A *datetime* denotes a particular point of time on a specific date, by adding a time of day to the date.

The definition of a specific calendar requires several elements:

1. The number of days in each month
2. The number of months in each year
3. Special features, e.g., leap days and leap months
4. An epoch or reference point

While many calendars have been used throughout history, most of the world now uses the Gregorian calendar, which was introduced in 1582. The calendar was adopted in Venice in 1582, then gradually adopted throughout the world until its 1752 introduction in the United Kingdom. The Gregorian calendar is a refinement to the Julian calendar, which had been the most widely used calendar. The refinement eliminated leap years in years that were multiples of 100, except for years that were multiples of 400, shortening the average year from 365.25 days to the more accurate 365.2425 days. The motivation for the change was to ensure the Vernal equinox always occurred on March 21st. [By 1582, the equinox was falling on March 11th.]

For our purposes, we will use both the Julian and Gregorian calendars. Each has the same number of days per month and months per year. Neither has a leap month and the leap days differ as described. The epoch for both is 12h Jan 1, 4713 BCE, with no time zone adjustment. When the calendar is unspecified, a date will be assumed to be Julian on or before October 14, 1582 and Gregorian on or after October 15, 1582. [October 5 – 14 are only defined on the Julian calendar.]

1. Prior to the introduction of TAI, days were not counted in SI seconds – they were governed by the rotation of the Earth. The most common form of day counting, which has been employed for both the Julian and Gregorian calendars, is the Julian day.

The epoch for the Julian day count is 12h Jan 1, 4713 BCE. Days are counted continuously – a leap day is just another day. To determine the Julian day for a date on the Julian calendar, we continue the pattern of leap days inserted as February 29th, starting in 4713 BCE, continuing every 4 years. To determine the Julian day for a date on the Gregorian calendar, we apply the Julian convention until October 4, 1582. To account for the accumulated difference in calendars due to the different treatment of leap years, the day after October 4, 1582 is assigned the date, October 15, 1582. Also, the Gregorian leap year convention is adopted, with the next leap year occurring in 1584.

TAI was initiated at 0h Jan 1, 1958 UT2, enabling the usage of Gregorian dates as a coordinate system, and enabling conversions between Julian days and TAI. Thus, 1 January 1958 00:00:00 TAI is Julian day 2436204.5, exactly.

UTC was initiated at 0h Jan 1, 1961. The TAI instant 1 January 1961 00:00:01.422818 exactly was identified as UTC instant 1 January 1961 00:00:00.000000 exactly. The 1.422818 second drift represented the observed difference between the invariant TAI and the variable UT2 that occurred since TAI was initiated. At that point, a decision was made to begin to align TAI and UTC by having UTC tick exactly one second for every 1.000000015 second of TAI. To keep UTC within 0.1 second of UT2, time steps occurred every few months and frequency changed at the end of each year.

There was dissatisfaction with the difference in frequency between TAI and UTC and with the frequent, irregular jumps in UTC. As a result, the decision was made in 1971 to use the SI second for UTC, matching it to TAI, and to eliminate the irregular jumps. Jumps would be integral seconds as described in the current leap second system. Leap seconds would be selected to approximate UT1 as opposed to UT2. The UTC instant 1 Jan 1972 00:00:00 was identified as 1 Jan 1972 00:00:10 TAI exactly. As a result, the difference between UTC and TAI will always be an integer number of seconds.

1. In order to precisely establish the epoch for our TAI data type, we need to select an age of the universe within the known range that is convenient for data analysis. We elected to choose a time interval prior to the Julian day epoch where the trailing digits of the number of days and the number of seconds computed for a date coordinate are the same as if the Julian day epoch were used.

The age of the universe in days is 5.039 x 1012 within an uncertainly of .014 x 1012. This corresponds with an age in seconds of 4.354 x 1017 with an uncertainty of .012 x 1017. We observe that multiplying the number of seconds by a multiple of 5 produces a result with the minimum number of digits, 3. If we choose 5.05 x 1012 days as the time interval, we are within the range of uncertainty. The number of seconds is 4.3632 x 1017.

The current Julian day count is on the order of 2,000,000, so the resulting day count with the selected epoch is of the form 5,050,002,xxx,xxx. Similarly, the number of TAI seconds corresponding to the current Julian day count is on the order of 200,000,000,000 so the resulting number of seconds with the selected epoch is 436,320,2xx,xxx,xxx,xxx. In this way, the Julian day count is the trailing 7 digits, while the second count is the trailing 12 digits.

Given the selected epoch, the Big Bang occurred at 12h Dec 26 13,826,151,189 BCE. The UTC instant, 1 Jan 1972 00:00:00 occurs at a day count of 5,050,002,444,317.5 and a TAI second count of 436,320,211,189,032,010.0.

1. Given the evolution of the time standards, accurate conversion from UTx date to TAI seconds is possible beginning 1961, but only convenient beginning 1972. Given that we know 10 seconds of drift occurred from 1961 to 1972, we could arbitrarily add 1 leap second at the end of each of the years 1962 through 1971.
2. The next higher level datatype is the DateTime. DateTime has the following elements:
   1. Calendar type – Julian or Gregorian
   2. Year + for CE, - for BCE – both are 1-based
   3. Month
   4. Day of Month
   5. Hour – 0 - 23
   6. Minute – 0 - 59
   7. Second - 0 – 60 – to facilitate leap seconds
   8. Milliseconds – 0 – 999
   9. Microseconds – 0 – 999
   10. Nanoseconds – 0 – 999
   11. Attoseconds – 0 – 999,999,999
   12. Time zone
   13. Standard or Daylight Savings Time – Only required for the hour of the transition back to standard time

The datetime can be converted to TAI seconds for any date through the next potential leap second date. After that date, the conversion can be conditional, with the accumulated number of leap seconds stored with the TAI seconds to be refined later.

# Notes from Wikipedia:

UTC was officially initiated at the start of 1961 (but the name Coordinated Universal Time was not adopted by the [International Astronomical Union](http://en.wikipedia.org/wiki/International_Astronomical_Union) until 1967).[[34]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTENelsonMcCarthy200515-34)[[35]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTENelsonMcCarthy2001515-35) The TAI instant 1 January 1961 00:00:01.422818 exactly was identified as UTC instant 1 January 1961 00:00:00.000000 exactly, and UTC ticked exactly one second for every 1.000000015 s of TAI.[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)] Time steps occurred every few months thereafter, and frequency changes at the end of each year. The jumps increased in size to 100 ms, with only one 50 ms jump having ever occurred. This UTC was intended to permit a very close approximation of UT2, within around 0.1 s.[*[citation needed](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed" \o "Wikipedia:Citation needed)*]

In 1967, the [SI](http://en.wikipedia.org/wiki/SI) second was redefined in terms of the frequency supplied by a caesium atomic clock. The length of second so defined was practically equal to the second of [ephemeris time](http://en.wikipedia.org/wiki/Ephemeris_time).[[36]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTEMarkowitz1988-36) This was the frequency that had been provisionally used in TAI since 1958. It was soon recognised that having two types of second with different lengths, namely the UTC second and the SI second used in TAI, was a bad idea. It was thought that it would be better for time signals to maintain a consistent frequency, and that that frequency should match the SI second. Thus it would be necessary to rely on time steps alone to maintain the approximation of UT. This was tried experimentally in a service known as "Stepped Atomic Time" (SAT), which ticked at the same rate as TAI and used jumps of 200 ms to stay synchronised with UT2.[[2]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTEMcCarthySeidelmann2009227-2)

There was also dissatisfaction with the frequent jumps in UTC (and SAT). In 1968, [Louis Essen](http://en.wikipedia.org/wiki/Louis_Essen), the inventor of the caesium atomic clock, and G. M. R. Winkler both independently proposed that steps should be of 1 s only.[[37]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTEEssen1968161.E2.80.935-37) This system was eventually approved, along with the idea of maintaining the UTC second equal to the TAI second. At the end of 1971, there was a final irregular jump of exactly 0.107758 TAI seconds, so that 1 January 1972 00:00:00 UTC was 1 January 1972 00:00:10 TAI exactly, making the difference between UTC and TAI an [integer number](http://en.wikipedia.org/wiki/Integer) of seconds. At the same time, the tick rate of UTC was changed to exactly match TAI. UTC also started to track UT1 rather than UT2. Some time signals started to broadcast the DUT1 correction (UT1 − UTC) for applications requiring a closer approximation of UT1 than UTC now provided.[[38]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTESeidelmann199285.E2.80.9387-38)[[39]](http://en.wikipedia.org/wiki/Coordinated_Universal_Time#cite_note-FOOTNOTENelsonLombardiOkayama200546-39)

TT differs from [Geocentric Coordinate Time](http://en.wikipedia.org/wiki/Geocentric_Coordinate_Time) (TCG) by a constant rate. Formally it is defined by the equation

TT = (1 − LG) TCG + E

where TT and TCG are linear counts of [SI](http://en.wikipedia.org/wiki/SI) [seconds](http://en.wikipedia.org/wiki/Second) in Terrestrial Time and Geocentric Coordinate Time respectively, LG is the constant difference in the rates of the two time scales, and E is a constant to resolve the [epochs](http://en.wikipedia.org/wiki/Epoch_(reference_date)) (see below). LG is defined as exactly 6.969290134 × 10−10. (In 1991 when TT was first defined, LG was to be determined by experiment, and the best available estimate was 6.969291 × 10−10.)

The equation linking TT and TCG is more commonly seen in the form

TT = TCG − LG × (JDTCG − 2443144.5003725) × 86400

where JDTCG is the TCG time expressed as a [Julian Date](http://en.wikipedia.org/wiki/Julian_Date). This is just a transformation of the raw count of seconds represented by the variable TCG, so this form of the equation is needlessly complex. The use of a Julian Date does specify the epoch fully, however (see next paragraph). The above equation is often given with the Julian Date 2443144.5 for the epoch, but that is inexact (though inappreciably so, because of the small size of the multiplier LG). The value 2443144.5003725 is exactly in accord with the definition.

Time coordinates on the TT and TCG scales are conventionally specified using traditional means of specifying days, carried over from non-uniform time standards based on the rotation of Earth. Specifically, both Julian Dates and the [Gregorian calendar](http://en.wikipedia.org/wiki/Gregorian_calendar) are used. For continuity with their predecessor [Ephemeris Time](http://en.wikipedia.org/wiki/Ephemeris_Time) (ET), TT and TCG were set to match ET at around Julian Date 2443144.5 (1977-01-01T00Z). More precisely, it was defined that TT instant 1977-01-01T00:00:32.184 exactly and TCG instant 1977-01-01T00:00:32.184 exactly correspond to the [International Atomic Time](http://en.wikipedia.org/wiki/International_Atomic_Time) (TAI) instant 1977-01-01T00:00:00.000 exactly. This is also the instant at which TAI introduced corrections for [gravitational time dilation](http://en.wikipedia.org/wiki/Gravitational_time_dilation).

TT and TCG expressed as Julian Dates can be related precisely and most simply by the equation

JDTT = EJD + (JDTCG − EJD) (1 − LG)

where EJD is 2443144.5003725 exactly.

The difference between ET and UT is called [ΔT](http://en.wikipedia.org/wiki/%CE%94T); it changes irregularly, but the long-term trend is [parabolic](http://en.wikipedia.org/wiki/Parabola), decreasing from ancient times until the nineteenth century,[[22]](http://en.wikipedia.org/wiki/Ephemeris_time#cite_note-morr3-22) and increasing since then at a rate corresponding to an increase in the solar day length of 1.7 ms per century (see [leap seconds](http://en.wikipedia.org/wiki/Leap_second)).

[International Atomic Time](http://en.wikipedia.org/wiki/International_Atomic_Time) (TAI) was set equal to [UT2](http://en.wikipedia.org/wiki/Universal_Time) at 1 January 1958 0:00:00 . At that time, ΔT was already about 32.18 seconds. The difference between Terrestrial Time (TT) (the successor to ephemeris time) and atomic time was later defined as follows:

1977 January 1.000 3725 TT = 1977 January 1.000 0000 TAI, *i.e.*

TT − TAI = 32.184 seconds

This difference may be assumed constant—the rates of TT and TAI are designed to be identical.

**Dual dating[[edit](http://en.wikipedia.org/w/index.php?title=Gregorian_calendar&action=edit&section=8" \o "Edit section: Dual dating)]**

*Main article:* [*Dual dating*](http://en.wikipedia.org/wiki/Dual_dating)

During the period between 1582, when the first countries adopted the Gregorian calendar, and 1923, when the last European country adopted it, it was often necessary to indicate the date of some event in both the Julian calendar and in the Gregorian calendar, for example, "10/21 February 1750/51", where the dual year accounts for some countries already beginning their numbered year on 1 January while others were still using some other date. Even before 1582, the year sometimes had to be double dated because of the different beginnings of the year in various countries. Woolley, writing in his biography of [John Dee](http://en.wikipedia.org/wiki/John_Dee_(mathematician)) (1527–1608/9), notes that immediately after 1582 English letter writers "customarily" used "two dates" on their letters, one OS and one NS.[[37]](http://en.wikipedia.org/wiki/Gregorian_calendar#cite_note-37)

**Old Style and New Style dates[[edit](http://en.wikipedia.org/w/index.php?title=Gregorian_calendar&action=edit&section=9" \o "Edit section: Old Style and New Style dates)]**

*Main article:* [*Old Style and New Style dates*](http://en.wikipedia.org/wiki/Old_Style_and_New_Style_dates)

"Old Style" (OS) and "New Style" (NS) are sometimes added to dates to identify which system is used in the [British Empire](http://en.wikipedia.org/wiki/British_Empire) and other countries that did not immediately change. Because the [Calendar Act of 1750](http://en.wikipedia.org/wiki/Calendar_Act_of_1750) altered the start of the year,[[38]](http://en.wikipedia.org/wiki/Gregorian_calendar" \l "cite_note-38) and also aligned the British calendar with the Gregorian calendar, there is some confusion as to what these terms mean. They can indicate that the start of the [Julian year](http://en.wikipedia.org/wiki/Julian_year_(calendar)) has been adjusted to start on 1 January (NS) even though contemporary documents use a different start of year (OS); or to indicate that a date conforms to the Julian calendar (OS), formerly in use in many countries, rather than the Gregorian calendar (NS).[[29]](http://en.wikipedia.org/wiki/Gregorian_calendar#cite_note-EC-NA-29)[[39]](http://en.wikipedia.org/wiki/Gregorian_calendar#cite_note-39)[[40]](http://en.wikipedia.org/wiki/Gregorian_calendar#cite_note-40)[[41]](http://en.wikipedia.org/wiki/Gregorian_calendar#cite_note-41)

1. The next higher level data type is calendar time, CalT.

* Consider Terestial Time – a historically revised form of TAI.