University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Lukas Yu

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, 10 April 2017

Bachelor's Thesis: Integration of Generally Valid Predicates into PostgreSQL

Ongoing time points, such as now, are common in relations that include a valid-time attribute. The end point of the valid time is equal to now if a tuple is valid until the present time. Since the present time continuously changes, the valid time of a tuple does so as well.

Thus, evaluating a predicate on ongoing time points does not result in a single boolean value, i.e., *true* or *false*, but in an ongoing boolean whose boolean value depends on the reference time. The reference time allows instantiating, i.e., replacing, ongoing data types to fixed data types at any possible time. For instance, consider the two ongoing time points now and $17/04/02$ and the less-than predicate $17/04/01 < now$. At reference time $17/04/01$, now, which always instantiates to a time point equal to the reference time, instantiates to the fixed time point $17/04/01$. Thus, the less-than predicate evaluates to *false*. At reference time $17/04/02$, now instantiates to time point $17/04/02$ and thus, the predicate evaluates to *true*. The result of the predicate is the ongoing boolean $\mathrm{b}[[17/04/02, \infty), (-\infty, 17/04/02)]$ that instantiates to the boolean value *false* up to reference time $17/04/01$ and to *true* from reference time $17/04/02$ onwards.

In this project, the student implements generally valid logical connectives and predicates for ongoing time points and time intervals into the kernel of the widely-used open-source database system PostgreSQL. Then, the results of the logical connectives and predicates remain valid as time passes by and the connectives and predicates can be used to formulate complex conditions.

Tasks

1. Implement the ongoing time points $a+b$ [3] in PostgreSQL by extending the existing *date* data type in order to support ongoing time points with the granularity of days.
2. Extend the ordering predicates for the *date* data type to allow ordering ongoing time

points. A time point $a{+}b$ is ordered before another time point $c{+}d$ if $a < c$ or if $a = c$ and $c < d$ ($a < b$, $c < d$).

3. Implement a new data type *ongoingBoolean* in PostgreSQL. Use a space-efficient representation of an ongoing boolean. The new data type must be usable as a data type of a table's attribute.
4. Develop efficient algorithms for the generally valid logical connectives conjunction, disjunction, and negation for ongoing booleans based on your representation of an ongoing boolean. Implement these algorithms as functions in PostgreSQL.
5. Integrate the generally valid comparison predicates for ongoing time points, $<$, $\leq$, $=$, and $\neq$, into PostgreSQL. Use the definitions provided in [3]. Make the predicates callable in SQL statements.
6. Integrate the generally valid comparison predicates for ongoing time intervals given in [3] into PostgreSQL. Make the predicates callable in SQL statements.
7. Empirically evaluate the performance of your approach. Compare your implementation with the state-of-the-art approaches, which instantiate ongoing time points before evaluating the predicates [2, 1]. Determine the cases in which your approach performs well and the cases in which it does not and explain your findings.
8. Write the thesis (approximately 50 pages).
9. Present your thesis in a DBTG meeting (25 minutes presentation).

References

[1] L. Anselma, B. Stantic, P. Terenziani, and A. Sattar. Querying Now-Relative Data. *Journal of Intelligent Information Systems*, 41(2):285–311, 2013.

[2] Clifford, James and Dyreson, Curtis and Isakowitz, Tomás and Jensen, Christian S. and Snodgrass, Richard Thomas. On the Semantics of Now in Databases. *ACM Transactions on Database Systems*, 1997.

[3] Mülle, Yvonne and Böhlen, Michael H. Generally Valid Queries in Databases with Ongoing Time Points. to be published.

Supervisor: Yvonne Mülle (muelle@ifi.uzh.ch)

Start date: 10 April 2017

End date: 30 September 2017

Presentation date: 5 September 2017, 14:00

University of Zurich
Department of Informatics

Prof. Dr. Michael Böhlen
Professor