Summaries and Searches in OLAP Query Logs

Julien Aligon, PhD Student

supervised by Dr. Patrick Marcel and Pr. Arnaud Giacometti

UNIVERSITÉ FRANÇOIS - RABELAIS CAMPUS BLOIS

Laboratoire d'Informatique EA 2101

Motivations

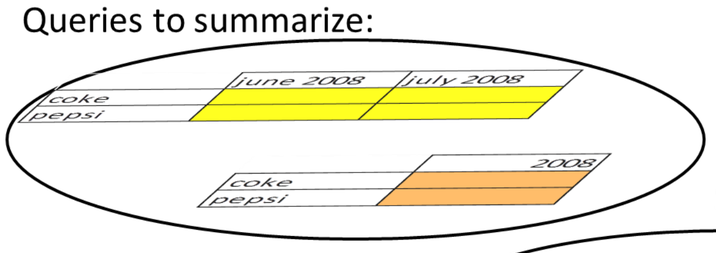
- How to allow a decision maker to have a rough idea of the queries launched by other decision makers?
- How to help the user to access a precise part of the log?
- How to help an administrator to manage and tune an OLAP server?

Quality Measure

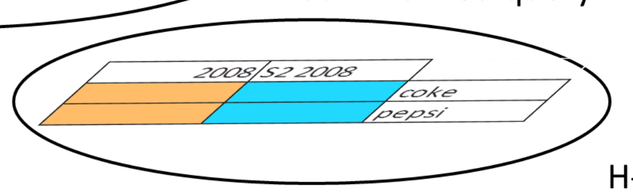
Problem: How to evaluate the quality of a summary?

Solution: A quality measure, based on precision and recall.

Queries to summarize:

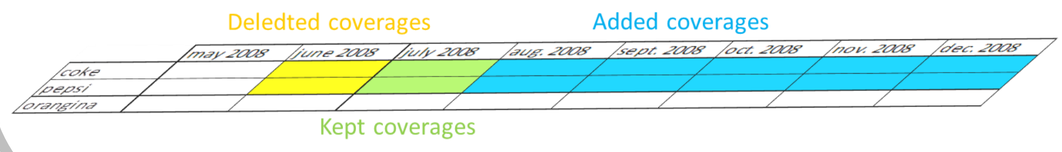


Summarized query:



$$\text{H-Recall} : \frac{2+2}{2+4}$$

The most detailed level:



$$\text{H-Precision} : \frac{2+2}{2+12}$$

SummarizeLog Algorithm

Problem: How to automatically compute the best summary?

Solution: With a greedy algorithm.

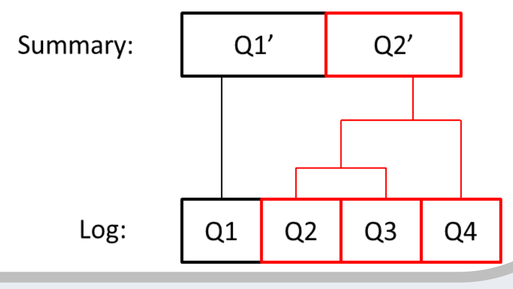
Parameters :

- The QSL operators to use
- Desired size N for the summary

While size of summary exceeds N

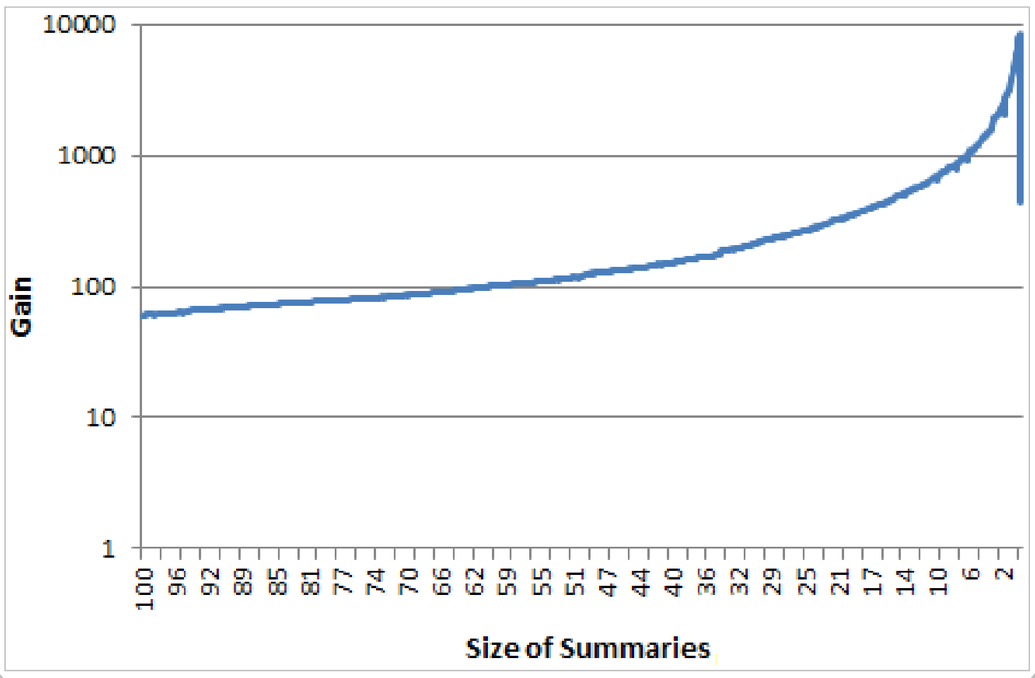
- Choose QSL espression maximizing the quality
- Replace the originals queries by the result of the QSL expression

➢ A summary partitions a log



Gain efficiency for searching

Gain of the computation time between a basic search (with disk access on each successive queries) and our proposal.



QSL: Query Summarizing Language

Problem: How to express various summaries of a log?

Solution: With a declarative language.

Q1:

		Paris	Marseille
Pepsi	July08		
Coke	July08		

Q2:

		North	South
Coke	July08		

$Q1 \triangleleft Q2 = Q1$

$Q1 \ U_B \ Q2$:

		Paris	Marseille	North	South
Pepsi	July08				
	July08				
Coke	July08				
	July08				
Coke	July08				
	July08				

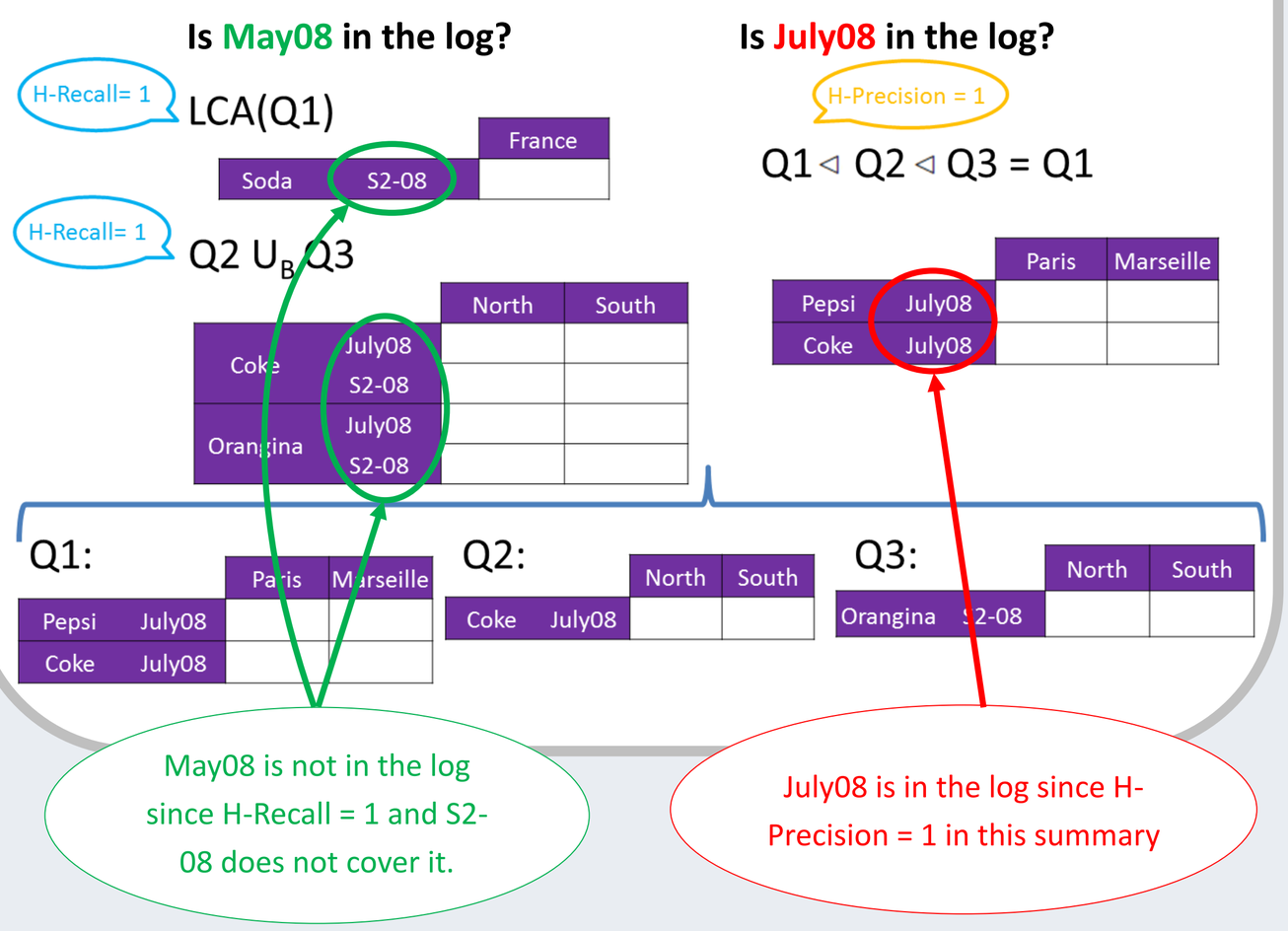
LCA(Q1) :

		France
Soda	S2-08	

Searches in a log

Problem: How to efficiently search in a log?

Solution: An Algorithm using the properties of the measure and the language.



Perspectives

- Take sessions into account.
- Extension to a collaborative context.
- Generalization to other types of logs (SQL, web logs).