

CodeQuest with DataLog

Elnar Hajiyeve, Mathieu Verbaere, Oege de Moor

Programming Tools Group,
University of Oxford, United Kingdom
<http://progtools.comlab.ox.ac.uk/projects/codequest/>
{elinar.hajiyeve,mathieu.verbaere,oege}@comlab.ox.ac.uk

Querying code for:

- Style
- Fault detection
- Refactoring
- Aspect weaving

Projects

Projects	KSLOC
Jakarta RegExp	3
JQuery 3.1.3 + TyRuBa	38
JFreeChart 1.0.0	93
Eclipse 3.1	1608

Benchmarked systems:

- JQuery:** Eclipse querying plugin
- XSB:** Prolog compiler with tabling
- CodeQuest:** our own system

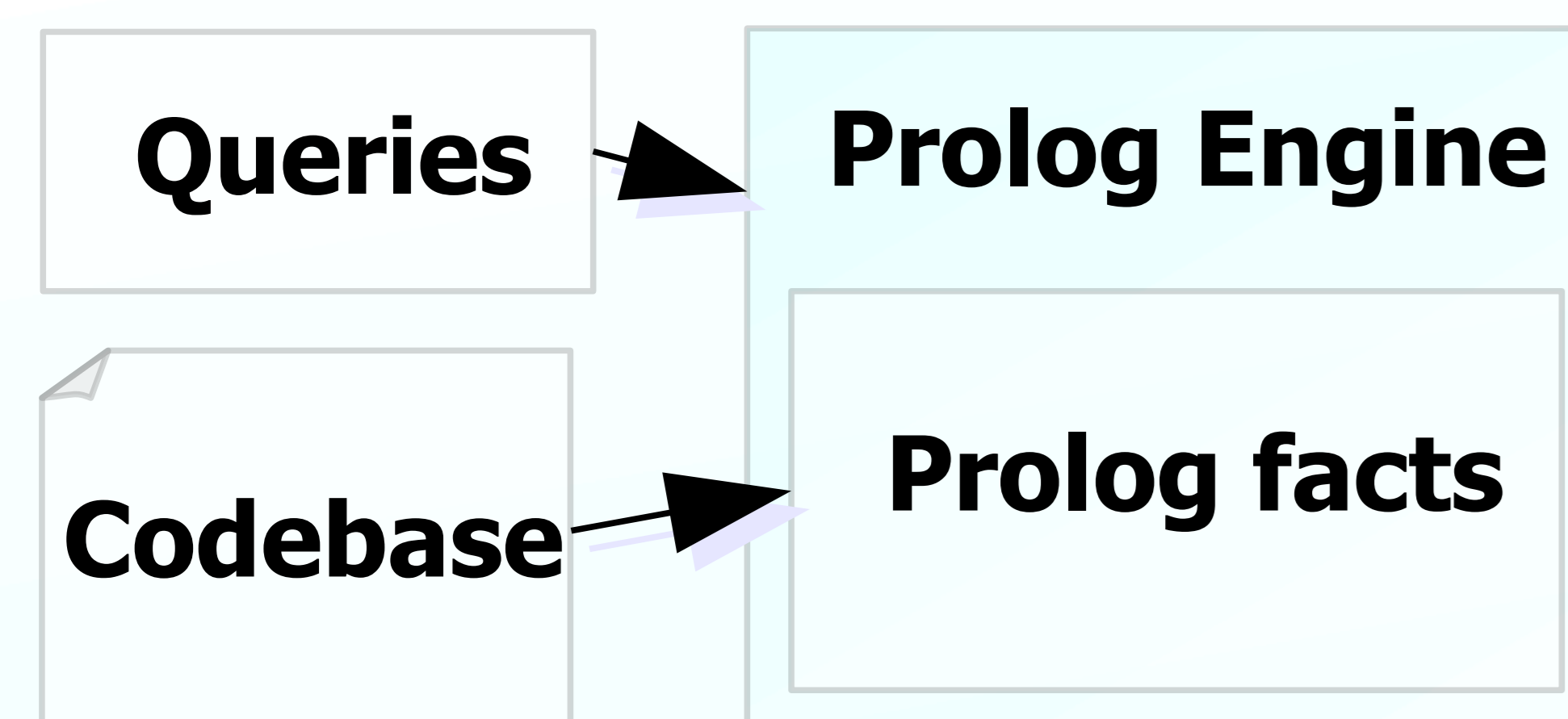
Our solution:

- Use DataLog = Prolog - data structures
- Expressive enough, efficient and scalable

Wish list:

- Expressive to cope with recursion on hierarchical structures
- Concise like logic programs
- Efficient query evaluation
- Scalable to store and query very large software projects
- Integrated with IDEs

Common answer: Prolog



Previous work:

Some systems that use this architecture:
JQuery, LogicAJ, ASTLog, etc.

Disadvantages of Prolog:

- Facts in memory do not scale
- Need mode annotations, cuts

Queries examples:

```

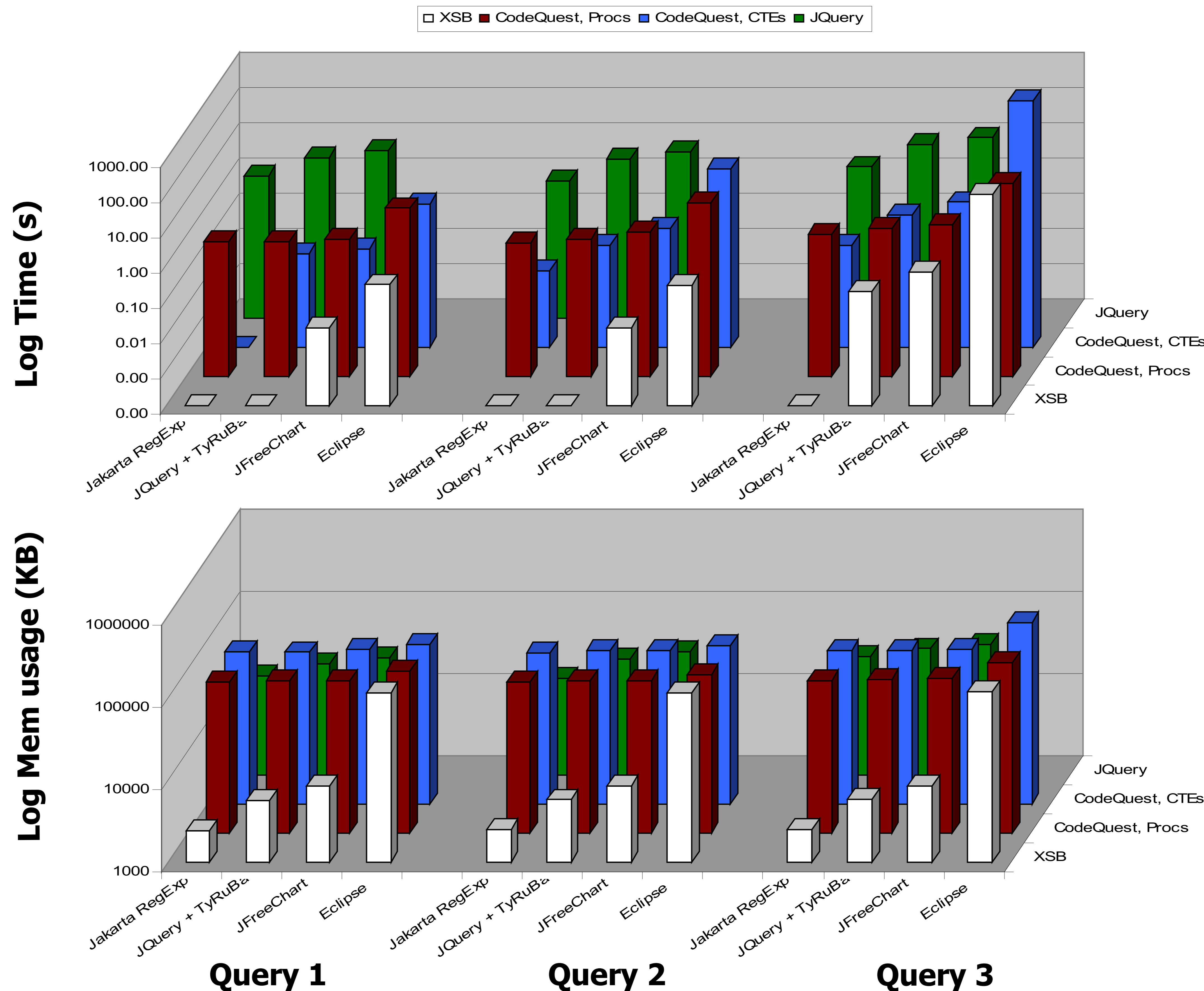
query1(T,F) :- type(T),
               hasChild(T,F),field(F),
               hasStrModifier(F,'public'),
               NOT(hasStrModifier(F,'final')).
  
```

```

query2(M,T) :- method(M),
               writes(M,F),hasType(F,FT),
               hasSubtypeStar(T,FT).
  
```

```

query3(M1,M2) :- overrides(M2,M1),
                 hasStrModifier(M1,'abstract'),
                 NOT(hasStrModifier(M2,'abstract')).
  
```



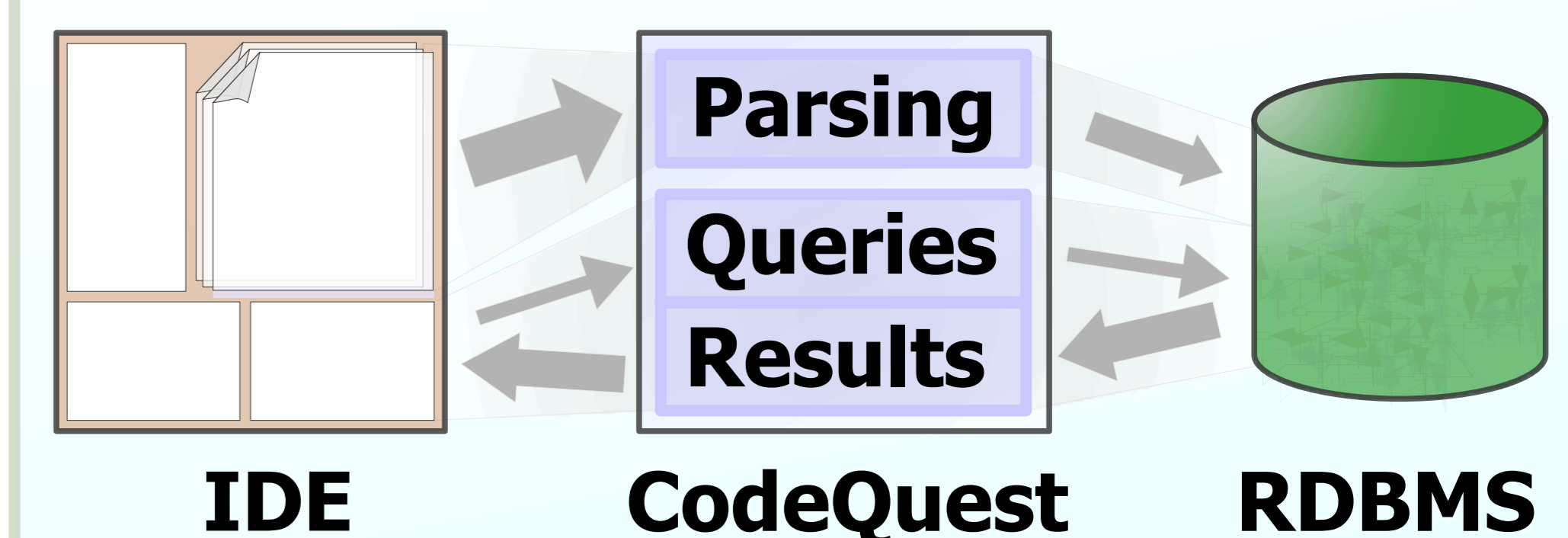
Storing facts in DBMS

is common in the software maintenance community (e.g. Paul and Prakash '96)

Our contributions:

- Express code queries with DataLog
- Optimising compiler from DataLog to SQL
- Incremental database update

CodeQuest architecture:



Future work:

- Aggressive optimisations for recursion
- Integration with Visual Studio
- Support of various databases
- Integration with *abc* compiler