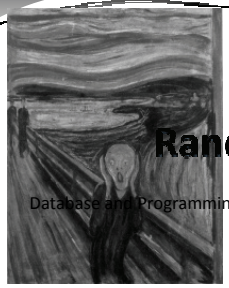

Review & Random Thoughts

CPS 296.1
Database and Programming Languages: Crossing the Chasm
Jun Yang
Duke University
April 22, 2010

Language is important

- Key to ease of use
  - Improving usability reduces human efforts in developing, testing, and maintaining code, arguably the most expensive
  - ☞ *Examples?*
- Key to efficiency (sometimes indirectly)
  - Language properties influence implementation/optimization
  - ☞ *Examples?*

Declarativeness

Killing two birds with one stone

- Ease of use + efficiency
- For DB folks: SQL doesn't have a patent on declarativeness
- Many functional/logical languages support declarative programming
- Though DB folks have really pushed the envelope on optimization

☞ *Any potential pitfalls?*

Image from http://n8tip.com/those-batchers

One size doesn't fit all

- But how do you code an app with many needs?
- Traditional solution: just write in different languages
  - ☞ *Problems?*
- Support an extensible set of restricted and declarative constructs really well
  - Ideally, with extensible, cost-based optimization
- Support other low-level constructs for completeness
  - Ideally, seamlessly with the high-level constructs

☞ *Promising examples?*

☞ *Anything new or different from UDFs in ORDBMS?*

Tough nuts to crack

- Seamlessness may be just an illusion
  - Processing boundary is still dictated by code (just less obviously)
- Extensibility is hard
  - Adding constructs to a language may be a solved problem now, but how to optimize them is still a wide open question
- Nesting is hard
  - Collections of collections?
  - SQL inside PL inside SQL inside PL?
- Too many optimization tricks but little strategy

Image from: http://www.1st-art-gallery.com/Joseph-Decker/Squirrel-With-Nuts.html

Lessons & ideas

- Love and hate with Java
- Right optimizations at right levels/times
  - Intermediate representations, JIT
- Lazy/deferred evaluation
- Higher-order functions (e.g., map/reduce) and lambda expressions
- To schematize or not
  - Omitting schema ≠ schemaless; learn from type inference
- Perfection is expensive; focus on the important stuff
  - Declarativeness really helps with complex, costly data flows

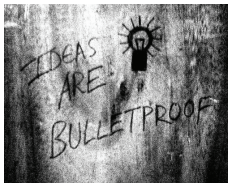

Image from: http://www.flickr.com/photos/ieshraq/2537467671/

Final reminders

Final "exam": Monday May 3, 9am-12pm, in this room

- Each of the six teams will get 25 minutes to present/demo
- Everybody votes for the best project
  - There will be a small prize for the best project team
  - And I will pay out to the Tuesday Jeopardy winners

- Final project reports will be due on Wednesday May 5, 12pm
  - 8-12 pages in two-column 10-pt standard proceedings format
  - Write it as if you are writing a conference submission
  - Don't forget the references
  - Email me the PDF and a link for downloading your project code

Image from: http://www.india-server.com/news-images/oscar-22242.jpg

7