From 100 Students to 100,000

Jennifer Widom
Stanford University
“Introduction to Databases”

**CS145 @ Stanford**

140  (enrolled students)

**DB-Class @ world**

60,000  (enrolled students)

26,000  (of them submitted 1+ assignments)

6,500  (completed the entire course)

**DB-Class cumulative** [July ‘12]

115,000  (accounts)

480,000  (assignment submissions)

3,350,000  (video views)
History

1) “Flipped classroom”
2) Complete course materials online
3) Public course offering
“Flipped Classroom”

Lectures ➔ self-paced videos with embedded quizzes

Class time ➔ interactive activities
Isolation Level **Read Committed**

- A transaction may **not** perform dirty reads
- Still does not guarantee global serializability

```
Update Student Set GPA = (1.1) * GPA Where sizeHS > 2500
```

Concurrent with...

```
Set Transaction Isolation Level Read Committed;
Select Avg(GPA) From Student;
Select Max(GPA) From Student;
```

\[ T_1; T_2 \]
\[ T_2; T_1 \]
Sample Video (demo)

```sql
select Student.sID, count(distinct cName)
from Student, Apply
where Student.sID = Apply.sID
group by Student.sID
union
select sID, 0
from Student
where sID not in (select sID from Apply);
```

(Number of colleges applied to by each student)
Consider a relation $R(A,B,C)$ with multivalued dependency $A \rightarrow B$. Suppose there at least 3 different values for $A$, and each value of $A$ is associated with at least 4 different $B$ values and at least 5 different $C$ values.

What is the minimum number of tuples in $R$?

- 60
- 15
- 12
- 27
Consider a relation $R(A,B,C)$ with multivalued dependency $A \rightarrow B$. Suppose there are at least 3 different values for $A$, and each value of $A$ is associated with at least 4 different $B$ values and at least 5 different $C$ values.

**What is the correct option?**

- 6
- 12
- 18
- 27

**Explanation**

Multivalued dependency $A \rightarrow B$ says that for each value of $A$, we must have every combination of $B$ and $C$ values. So for each of the 3 values of $A$, we must have at least $4 \times 5 = 20$ different tuples.

**Correct**
Classroom Time

• Professor-led interactive problem-solving

• Lectures outside of core material
  – Guest lectures from industry
  – Guest lectures from Stanford
  – Research presentations
  – Advanced/exotic topics

• Help/review sessions
Step #2: Course Materials Publicly Available

Videos with embedded quizzes

+ Comprehensive exercises

+ Slides, scripts, readings, software guides

= Complete course

- Completed summer 2011
- 1000’s of views
The Inflection Point

Online materials ➔ Structured course offering
+ Schedule with deadlines
+ Automated exercises
+ Online exams
+ “Statement of accomplishment”
★ Community

Transformed 1000’s to 10,000’s
(Still free of charge)
Automated Exercises

- **Quizzes using technology from Gradiance**
  - “Root” question + bank of correct & incorrect answers
  - Explanations
  - Different variant each time
- **Automatically checked programming**
  - Particularly well-suited for database class
  - Relational algebra, SQL, triggers, views, DTDs, XPath & XQuery, XSLT

Students encouraged to repeat until perfect score
Sample Quiz

Authorization Quiz

This quiz uses an approach pioneered and patented by the Gradiance Corporation. Each multiple-choice quiz problem is based on a “root question,” from which the system generates different correct and incorrect choices each time you take the quiz. Thus, you can test yourself on the same material multiple times. We strongly urge you to continue testing on each topic until you complete the quiz with a perfect score at least once.

After submitting your selections the system will score your quiz, and for incorrect answers will provide an “explanation” (sometimes for correct ones too). These explanations should help you get the right answer the next time around. To prevent rapid-fire guessing, the system enforces a minimum of 10 minutes between each submission of solutions.

-----------------------------------------------

Question 1  
Unanswered

Consider a set of users U, V, W, X, and Y. Suppose user U creates a table T and thus is the owner of T. Now suppose the following set of statements is executed in order:

1. User U: grant select on T to V, W with grant option
2. User V: grant select on T to W
3. User W: grant select on T to X, Y
4. User U: grant select on T to X
5. User U: revoke select on T from V restrict
6. User U: revoke select on T from W cascade

Which of the following statements is true?

- Y does not have privilege SELECT ON T after statement 6
- W does not have privilege SELECT ON T after statement 5
- X does not have privilege SELECT ON T after statement 5
- X does not have privilege SELECT ON T after statement 6

-----------------------------------------------

Question 2  
Unanswered

The following SQL statement over tables R(t,c), S(f,g), and T(a,b) requires certain privileges to execute:

```
UPDATE T
SET a=1, b=2
WHERE a = ALL (SELECT i FROM R)
OR EXISTS (SELECT f FROM S WHERE f > T.a)
```

Which of the following privileges is not useful for execution of this SQL statement?

- SELECT ON (T(b)
- SELECT ON (R(a))
XML Course-Catalog XPath and XQuery Exercises

In these exercises, you will be working with a small XML data set drawn from the Stanford course catalog. There are multiple departments, each with a department chair, some courses, and professors and/or lecturers who teach courses. The XML data is here.

Instructions: You are to write each of the following queries using either XPath or XQuery. Each question contains a link to a Query Workbench. The workbench will help you develop and debug your queries by running them using Saxon over the sample data set. Once you’re satisfied with your solution, click the “Keep This Solution” button in the Query Workbench — your query will be copied automatically to the box on this page for final submission.

You may perform these exercises as many times as you like, so we strongly encourage you to keep working with them until you complete the exercises with full credit.

NOTE: REMEMBER TO CLICK “Save and Submit” WHEN YOU ARE DONE!
Please be patient as it does take time to check all of the exercises.

Question 1
Unanswered
Return all Title elements (of both departments and courses).
Note: Your solution will need to reference doc("courses.xml") to access the data.

Query Workbench

Question 2
Unanswered
Return last names of all department chairs.
Note: Your solution will need to reference doc("courses.xml") to access the data.

Query Workbench
Personal Touch
Need for Perfection

Each order of magnitude more students

⇒ order of magnitude higher need to avoid ambiguities and errors
Meanwhile Back at Stanford

• Automated exercises and level of perfection well appreciated

Q: “Added value” for tuition?
  – In-class activities and lectures
  – Hand-graded written exercises
  – Programming project
  – Real exams

Statistically-significant increase in evaluations
The Public Students

Deeply appreciative

Q&A Forum - View

What is your background?

162

I'm 22 years old, high school dropout from the United States. I've been coding since I was 14, and have built a lot of applications, but I'm hoping to start over anything I can from these courses (I enrolled for several others as well).

I think it's awesome Stanford is offering these courses online for free, and has put so much effort into providing an excellent user experience. I'm really impressed so far.

How about the rest of you, where are you from, and what is your background?

Tags: greetings introductions welcomequestions

Posted by Calvin Freeberg (+179) [Edit] [Delete]
4 months 3 weeks ago [Favorite]
A Top Student

• Posted 900 answers on Q&A Forum
• No discrimination against “dumb” questions
• Each answer correct, of perfect length, with examples when appropriate, perfect English
A Top Student

Q&A Forum - View

Amy Cunningham

108
Now that the course is approaching its end, I would like to express my gratitude towards Amy Cunningham for her great contribution to this forum. Her clear and constructive answers helped us a lot in understanding the materials and exercises. Amy, thank you very much!

Tags: authorization

Posted by Anonymous User [Edit] [Delete]
4 months 10 hours ago [Feature?]
Downsides

1) Complainers
2) Cheaters
What’s Happened Since

• ~12 additional Stanford courses
  Mostly CS, but some others

• Very high interest across campus and beyond

• Many blossoming efforts
  – Universities and companies
  – Platforms and content
  – Non-profit and for-profit
Big Questions for University (selected subset!)

- Course-content ownership — university or professor?
- Teaching credit for putting courses online? Money?
- Who’s responsible for course production? Hosting?
- Is it a university’s mission to educate the world?
- If everything is online, what’s the $50K/year buying?
- What’s the future of higher education?
From 100 to 100,000

**CS145 @ Stanford**
- 140 (enrolled students)

**DB-Class @ world**
- 60,000 (enrolled students)
- 26,000 (of them submitted 1+ assignments)
- 6,500 (completed the entire course)

**DB-Class cumulative [April ‘12]**
- 115,000 (accounts)
- 480,000 (assignment submissions)
- 3,350,000 (video views)