Information Retrieval

WS 2013 / 2014

Lecture 14, Tuesday February 11th, 2014

(Course Evaluation, Exam, Work at our Chair)

Prof. Dr. Hannah Bast
Chair of Algorithms and Data Structures
Department of Computer Science
University of Freiburg

UNI FREIBURG

Overview of this lecture

UNI FREIBURG

- Organizational
  - Your experiences with ES#13 (Stat. Sign.) **+ BUG FIX**
  - Results of the official **evaluation** of this course
- Exam
  - Types of tasks + grading scheme
  - Let's solve some tasks together live
- Work at our chair
  - How we work
  - Current projects + what's behind them
  - Upcoming courses

Experiences with ES#13 (Stat. Significance)

UNI FREIBURG

- Summary / excerpts last checked February 11, 15:00
  - Many of you skipped this, because you already had enough points + the usual end semester stress

    Beware though that the (written) exam is already next week
  - No complaints about the mistakes on the slides, instead:

    "There is nothing more instructive than incompleteness"
  - The 20 points for the online evaluation raised the maximum number of points from the announced 260 to 280

    Don't worry, the 20 points will substitute your weakest sheet

    Everybody who registered for the exam is admitted by us !
  - "The EvaSys mails landed in my spam folder"

Results from ES#13 (Statistical Significance)

UNI FREIBURG

- Summary
  - Even if some precisions better for BM25, high p-values:

    TF-IDF: 0.2, 0.4, 0.4, 0.8, 0.4 BM25: 0.4, 0.4, 0.4, 0.8, 0.6

    R-Test: 30% T-Test: 10% Z-Test: 8%
  - For incomparable measurements, p-values as high as 50%
  - Bottom line: **five queries** are simply not enough to provide evidence for a quality difference between ranking methods
  - We also note that the R-Test is the most conservative, followed by the T-Test, followed by the Z-Test

    Understand that this reflects well how realistic these test are

Yet another mistake in Lecture 13

- The variance of M = the mean difference
  - Consider the Z-Test (for the T-Test it is analogous)
  - The slides said, the variance of M is σ^2 / n
  - But the variance of M is actually $4 \cdot \sigma^2 / n$
  - The test statistics is hence $\sqrt{n} \cdot \Delta\mu / (\mathbf{2} \cdot \sigma)$

    That is, **a factor of 2 less** than the slides used to say
  - Correspondingly, the p-values become (even) larger

I have corrected this on Slides 20 – 22 now

$\mathrm{var}(X-Y) \neq \mathrm{var}(X) - \mathrm{var}(Y)$

$\mathrm{var}(X-Y) = \mathrm{var}(X) + \underbrace{\mathrm{var}(-Y)}_{=\mathrm{var}(Y)} = \mathrm{var}(X) + \mathrm{var}(Y)$

n measurements $X_1, \dots, X_n$, $M_1 = \sum_{X_i \in A} X_i / (n/2)$, $\mathrm{var}(X_i) = \sigma^2$

$A = \{X_1, \dots, X_{n/2}\}$, $B = \{X_{n/2+1}, \dots, X_n\}$

$\Rightarrow \mathrm{var}\left(\sum_{X_i \in A} X_i\right) = n/2 \cdot \sigma^2$

$\Rightarrow \mathrm{var}\left(\underbrace{\sum_{X_i \in A} X_i / (n/2)}_{M_1}\right) = \frac{n/2}{(n/2)^2} \cdot \sigma^2 = 2\sigma^2/n$

$\mathrm{var}(M) = \mathrm{var}(M_1) + \mathrm{var}(M_2) = 2\sigma^2/n + 2\sigma^2/n = 4\sigma^2/n$ ∎

Results Course Evaluation 1/8

UNI FREIBURG

- Participants
  - Still participating in the course: **47**
  - Registered for exam: **44**

    18 x MSc Inf, 14 x BSc Inf, 7 x B.Sc. ESE, 5 x Erasmus
  - Participated in the evaluation: **43** … great, thanks !
  - Nominations for teaching award: **31** … thanks again !
  - In the following, a summary of your feedback
  - You find **all** the details linked on the course Wiki

Results Course Evaluation 2/8

UNI FREIBURG

- Style of the course
  - Learned a lot: 70% fully agree, 23% agree, 7% ok
  - Level of contents: 67% appropriate, 28% high, 5% low
  - Well explained: 74% fully agree, 21% agree, 5% ok
  - Answers questions: 79% fully agree, 14% agree, 7% ok
  - Good atmosphere, nerd humor, cares about students
  - Also difficult stuff explained by simple examples first
  - Great emphasis on practical (and real) applications
  - One topic per lecture is nice … but see also complaints
  - "Frau Bast is unbelievably motivated … at least after 2 pm"

Results Course Evaluation 3/8

UNI FREIBURG

- Exercise Sheets
  - Effort ok for ECTS ... 1 = fully agree, 6 = fully disagree

    19% x 1 28% x 2 25% x 3 16% x 4 7% x 5 5% x 6 this course

    20% x 1 34% x 2 24% x 3 12% x 4 6% x 5 4% x 6 department average
  - Practical exercises that deepen / verify understanding
  - TIP files were good, or rather: invaluable
  - A lot of work, sometimes / for some too much
  - Great feedback from tutor, thanks again !
  - Some tutors too slow with their corrections

Results Course Evaluation 4/8

UNI FREIBURG

- Materials / Online Support
  - Helpful: 72% fully agree, 21% agree, 7% ok / partly
  - Consumed lecture by presence or by video recordings:

    21% by presence, 56% by recording, 23% both ways — this course

    47% by presence, 16% by recording, 30% both ways — department average
  - Great lecture recordings
  - Great support on the forum ... also on weekends and Xmas
  - Daphne, SVN, Jenkins works well + useful
  - Handwriting sometimes too small / hard to read on slides
  - "Prof. Bast is making funny noises during the recording"

Results Course Evaluation 5/8

UNI FREIBURG

- Complaints **from last year (WS 12/13)**
  - Quite a lot of mistakes on the slides (for some lectures)
  - Second half of course: big picture less clear / harder to follow
  - Naïve Bayes, SVM, SPARQL belong into other lectures
  - Too much mathematics, esp. when only the result is needed
  - "Requiring login for evaluation defeats the purpose"
  - Tutor feedback not only on the code, please
  - Unfair distribution / awarding of points
  - Colors are good, but switching colors is bad
  - Improve space management when writing on slides

Results Course Evaluation 6/8

UNI FREIBURG

- Complaints **from this year (WS 13/14)**
  - Do drawings / math in PowerPoint / LaTeX, not by hand
  - Present mathematical contents on the blackboard, this would make the lecture more agile
  - More theory / mathematics please, also in exercises
  - More background information, e.g. Eigenvector Decompos.

    Note: there is a trade-off between **breadth** and **depth**, based on previous student's feedback I go for breadth
  - Still too much effort for the exercise sheets
  - Some exercise sheets have nothing to do with Information Retrieval, e.g. using Octave, SQLite etc.

Results Course Evaluation 7/8

UNI FREIBURG

- Planned improvements **from last course (WS 12/13)**
  - Improve slides + explanations + time management

    I.p. for: Web Stuff, LSI, K-Means, SVM, SPARQL, T-test ✓
  - Better split of web app stuff over two lectures ✓
  - In general, there will be much fewer mistakes on slides ✓
  - Improve specification for exercise sheets ✓
  - Reduce time effort needed for sheets ✓
  - Solve the pen color problem ✓
  - Maybe more about semantic search next time ✗
  - And, of course, I will consider all the other feedback too ✓

Results Course Evaluation 8/8

UNI FREIBURG

- Planned improvements **for next course (WS 14/15)**
  - Fix remaining mistakes on slides … without introducing new ones
  - **Further** reduce time effort needed for sheets
  - Forever work on time management … depends on daily condition
  - **Different room** … HS 036 is **terrible** in the winter + with 10 people
  - In general: I have made numerous notes for each lecture of small things which I want to improve next time …

Exam 1/5

ORAL EXAMS
(B.Sc. Computer Science ONLY)
Date 1: FEB 21 (FRI)
Date 2: MAR 27 (THU)

UNI FREIBURG

- Where, when, how
  - Wednesday, **February 19, 2:00 – 4:00pm** in **HS 026**
  - There will be six tasks, out of which the five best will count

    Grading scheme: see exam page linked on the Wiki
  - The exam is **open book** = you can bring books, paper, **...**

    But please be ecological when printing out slides
  - Electronic devices of any kind are not allowed
  - There will be a sub-forum for questions about the exam

  Please bring: **student id**, **colored pens**, **brain**

Exam 2/5

UNI FREIBURG

- Types of questions
  - **Type 1:** Do the steps of an algorithm, or a variant thereof, like we did in the lecture … see colored pens
  - **Type 2:** Write a small program, or understand what a given small program does
  - **Type 3:** Small calculations or proofs … see brain
  - **In general:** the emphasis is on (basic) understanding, not on learning things by heart
  - **Preparation:** understand how it was done in the lecture, then put the solution away, then try **yourself**

    If you did / do all the exercises, you are well prepared

Exam 3/5

- Example of a Type 1 Question

Doc 1 bla bla
Doc 2 bla bli bla
Doc 3 blu blu
Doc 4 bla blu bla
Doc 5 bli blu blu

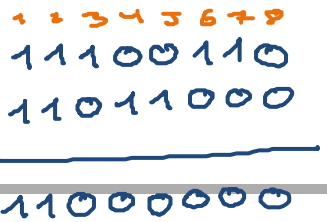
non-gap 2g

bla : (Doc1, 2) (Doc2, 2) (Doc4, 2)
bli : (Doc2, 1) (Doc5, 1)
blu : (Doc3, 2) (Doc4, 1) (Doc5, 2)

with gaps gap 2g

bla : (+1, 2) (+1, 2) (+2, 2)
bli : (+2, 1) (+3, 1)
blu : (+3, 2) (+1, 1) (+1, 2)

Exam 4/5



- Example of a Type 2 Question

boolean type

logical AND

```
ArrayList<Bool> intersect(ArrayList<Bool> list1,
                          ArrayList<Bool> list2) {
    result = new ArrayList<Bool>();
    for(int i=0; i< list1.size(); i++) {
        Bool bit = list1.get(i) && list2.get(i);
        result.add(bit);
    }
    return result;
}
```

I assumed that the bit arrays have the same size.

Exam 5/5

- Example of a Type 3 Question

$\Pr(X = i) = 2^{-i} =: p_i$, $H(X) = -\sum_{i=1}^{\infty} p_i \cdot \log_2 p_i$

$= -\sum_{i=1}^{\infty} 2^{-i} \cdot \underbrace{\log_2 2^{-i}}_{-i}$

$= \sum_{i=1}^{\infty} i / 2^i = 2$

↳ used without proof as suggested.

list with n gaps $\Rightarrow 2n$ bits (expected)

bit array repr. $\Rightarrow n$ bits , $n = \#docs$.

better if $n < 2n$

Work at our Chair 1/3

UNI FREIBURG

- How we work
  - 1/3 Theory (new algorithms, performance analysis, etc.)

    E.g. an efficient index for semantic search, or for computing shortest paths in very large transportation networks
  - 1/3 Algorithm Engineering (good implementation)

    An idea that looks great in theory might not work that well, or even not at all, in practice

    On the other hand, hacking around without theoretical understanding often leads to nowhere good either
  - 1/3 Software Engineering (good software)

    Writing a program for yourself which runs once now is one thing

    Writing software together with others that can still be used in five years from now is a totally different story

Work at our Chair 2/3

UNI FREIBURG

- Current projects
  - Multi-modal route planning

    Arbitrary combination of car, transit, bike, flights, …

    Good models, efficient algorithms, a working system
  - Semantic search

    Search with "understanding" of the query and documents

    Show example of **Broccoli** and **Freebase Easy**
  - Research paper management

    Automatic metadata + reference extraction

    Show demo video of IceCite prototype

Work at our Chair 3/3

UNI FREIBURG

- Upcoming courses
  - **Programming in C++** … in the SS 2014

    2nd semester BSc Info + 4th semester BSc ESE
  - **Randomized Algorithms** … in the SS 2014

    Spezialvorlesung given by Sabine Storandt (postdoc at my chair)
  - **Algorithms and Data Structures** … next time in SS 2015

    Basic course for 2nd semester BSc Informatik students
  - **B.Sc. / M.Sc. projects or theses**

    Offered all the time, just ask … and **ask again** if no response !