Informatics 2A: Lecture 1

Introduction and Course Administration

John Longley

Mirella Lapata

Subject of the course

The official full title of the course is:

Informatics 2A: Processing Formal and Natural Languages

The course is about ways of describing, specifying and processing both computer languages and human languages.

Remarkably, many important ideas and methods are common to both of these — though there are also major differences.

Lecture 2 will give a overview and roadmap of the intellectual content of the course. Today's lecture is mainly admin.

Course staff

Lecturers:

- John Longley (`jrl@staffmail.ed.ac.uk`)
  Office hour: Thursday 1-2pm, IF 5.12

- Mirella Lapata (`mlap@inf.ed.ac.uk`)
  Office hour: Monday 1-2pm, IF 4.16

Inf2 Year Organizer:

- Guido Sanguinetti (`gsanguin@inf.ed.ac.uk`)

Course Secretary:

- Kendal Reid, ITO (`kr@inf.ed.ac.uk`, AT 4.02)

Teaching Assistant:

- Christos Christodoulopoulos (`s0787820@sms.ed.ac.uk`)
  Office hour: Wednesday 1-2pm, IF 3.39

Communication mechanisms

Course website:

`http://www.inf.ed.ac.uk/teaching/courses/inf2a/`
This is the main anchor point for all course information and material. Bookmark it now.

Email list: `inf2a-students`. Important administrative announcements (e.g. changes to deadlines) will be posted here.

Discussion forum: for technical aspects of course material (see later).

Course reps: for feedback from you to course staff (see later).

It is your responsibility to check (especially) the website and your email to stay in touch with what's going on.

Prerequisites and pass criteria

Prerequisite for Inf2A: pass in all Inf1 courses (except for visiting students).

Assessment on the course will be 75% exam, 25% coursework.

For a pass in Inf2A, you need all of the following:

- At least 40% combined total mark.

- At least 35% in the exam.

- At least 25% on the assessed coursework.

For (safe) progression to Honours degree programmes, you need to pass *all* Inf2 courses with at least 50% *on the first attempt*.

Inf2A exam

The main exam takes place in December 2011.
The resit is in August 2012.

Exam dates are set by Registry, not us. We'll let you know once they are announced.

The exam is pen-and-paper, and lasts 2 hours. It will consist of 20 compulsory multiple choice questions (50% total), and a choice of 2 out of 3 longer questions (25% each).

Assessed coursework

There will be two assessed coursework assignments, carrying equal weight.

Assignment 1: issued Mon 24 Oct, due in Mon 7 Nov, 4pm
Assignment 2: issued Mon 14 Nov, due in Mon 28 Nov, 4pm

Both assignments will be electronic in nature, and are to be submitted online from DICE machines.

Assignments will be marked and returned within 2 weeks of submission.

Plagiarism

The University's definition:

> Plagiarism is the act of including in ones work the work of another person without providing adequate acknowledgement of having done so, either deliberately or unintentionally.

Plagiarism is a very serious offence. Discussion of general ideas with other students is encouraged, but e.g. sharing of even small snippets of code for assessed assignments is absolutely out.

For further clarification, consult

```
http://www.ed.ac.uk/schools-departments/academic-services/
          staff/discipline/plagiarism
http://www.inf.ed.ac.uk/admin/ITO/
          DivisionalGuidelinesPlagiarism.html
```

Course activities: Lectures

Congratulations on finding your way here. Lectures will be in AT LT5 at 16:10 on Tuesdays, Thursdays and Fridays.

The website contains links to the slides for each lecture. These links will become live immediately after (or just before) the lecture takes place.

Printed copies of lecture slides: please bear in mind the cost to the planet. If you print them yourself, use the 4up option.

Question: If I didn't hand out copies, would you typically print one off for yourself?

Tutorials

Tutorials for Inf2a don't start until Week 3 (beginning Monday 3 Oct). So Tutorial n happens in Week $n + 2$.

Each tutorial will cover material from the previous week's lectures. A *tutorial sheet*, consisting of problems to be discussed in Tutorial n, will be handed out at the Friday lecture in Week $n + 1$ (and made available on the website).

You will soon receive an email from Kendal Reid via inf2a-students, advertising the preliminary allocation of students to tutor groups. If you can't make the time of your allocated group, please email Kendal suggesting some groups you *could* manage. Or if you change tutor groups for any other reason, please let Kendal know.

N.B. If you miss two tutorials in a row, your DoS will be notified and you may be chased up!

Python and Lab Sessions

In parallel with the lecture material, you are also expected to pick up a new programming language (Python) and to learn to use the associated Natural Language Toolkit (NLTK). These skills will be needed for the two assessed course assignments.

This can be done with the help of four worksheets (available via the website), one for each of Weeks 3–6. You can work through these on your own, or you may attend the Lab Session to which you have been assigned (or a different one). A Lab Demonstrator will be on hand at these sessions to offer help.

Class reps

We require around six volunteers to serve as class reps for the course. Their job is to serve as a channel for feedback on the course content and delivery to the course staff. (You can be a rep for more than one Inf2 course.)

Besides collecting and collating such feedback, you will be expected to attend one Staff-Student Liaison Meeting for Inf2A, and perhaps invited to other Informatics-wide liaison meetings. You will receive guidance from EUSA on how to be a class rep.

Any volunteers? If so, come and give me your name and email address at the end of this lecture. These details will then be posted on the website.

Discussion forum

The Inf2A homepage contains a link to a web-based discussion forum. This is a place where you can post technical queries relating to the course material — they may be answered by lecturers, the TA, tutors or other students.

Of course, you mustn't (parts of) solutions to assignments or tutorial questions!

To join, follow the link from the Inf2A homepage and click on 'Subscribe'. Then read the introductory message from me, which explains how to set things up so that you can post to the forum.

Recommended reading

The following textbook is highly recommended for this course and many other Natural Language courses in later years:

- D. Jurafsky and J. Martin, Speech and Language Processing (2nd edition), Prentice-Hall, 2009.

For the formal language side, the recommended text is:

- D. Kozen, Automata and Computability, Springer, 2000.

Lectures will stick closely to the terminology and notation of these texts. Another useful resource is:

- S. Bird, E. Klein and E. Loper, Natural Language Processing with Python, O'Reilly, 2009. Available online at `http://www.nltk.org/book`

Technical wizardry

- Inf2A lectures are being videoed this year! A link to the videos will appear on the website soon.

- Many of the lectures will make use of Personal Response Systems, a.k.a. clickers. You can collect your clicker (for free) from the Main Library in George Square. The same clicker will work in all courses.

Clicker availability

Where are Clickers available for loan?

1. The Main Library

2. Somewhere else

Needing help?

- If you are suffering from personal circumstances that may be adversely affecting your work, contact your DoS.

- If you wish to apply for a coursework deadline extension (for a good reason!), contact the Year Organizer Guido Sanguinetti (*not* the lecturers). Normally, extensions will only be granted if applied for *prior to* the stated deadline.

- If you are having difficulties understanding the course material, seek help from (roughly in order) your friends, your tutor, the teaching assistant, the lecturers. Or try posting to the discussion forum.

- If you wish to voice a complaint about the course material or delivery, contact one of the class reps.

Enjoy the course!

Next lecture: Overview and roadmap of the intellectual content of the course (JL, ML).

Any questions?