Hi, and welcome to the first meeting of CS 6742, "Natural language processing and social interaction."

I'm the instructor; I'm a professor in computer science and in information science, and these days,

NLP and social interaction research in NLP and social interaction is what I do for a living,
(and for fun, come to think of it).

I'd give you more specifics now, but

As for this course and why it exists:

We all know that spoken/written language is one of the main channels by which we human beings communicate with each other.

(If you're wondering what some of the other channels are, there's facial expression, gesture, actions or and even inactions, if you think about it

Posture...)

Anyway, about all this language-based communication: I doubt it's escaped your notice that more and more of such communication is becoming computationally accessible, since it's now...
Aside: "distancing" of medium vs. probability of lying? No.

Jeff was just referred to by one of our colleagues as a professor of communications.

Now, actually, there are differences and similarities worth noting.

But just to point out some interesting work done by Jeff Hancock and colleagues, mention some work done by Jeff Hancock and

We can perform a rich categorization of the kinds of conversations, @ first blush:

In 2011, April 2011, Twitter was seeing 400 million tweets per day.

About 15% of Twitter is a spam or unverified communication

Let me briefly mention some work done by Jeff Hancock and colleagues. One interesting topic that we won't get into today is

This is actually lower than before 10 AM, but think things are changing.

We can perform a rich categorization of the kinds of conversations, @ first blush:

In 2011, April 2011, Twitter was seeing 400 million tweets per day.

About 15% of Twitter is spam or unverified communication.

(This is actually lower than before 10 AM, but think things are changing.)
Ok, so but anyway, there are yes, many of the conversations we have are one-on-one.

But your pretty we also have group conversations pretty frequently, where more people are involved, and that may also be semi-public depending on where they take place.

private/public audience

private-ish intended audience, public location = group discussions:

social-network conversation thread, blog comment threads

when "blog" here is broadly construed

some may be meant for more a subset of participants

may be meant for others more for an audience as well

but may also be an element of the public being an intended audience as well.

Ok, let's circle back:

And that thought leads us to:

public audience = analog to "soap boxes"

blog posts (even w/ what comments)

what else?

one you might not think of: reviews on

review aggregation sites (Yelp, Amazon, etc.)

now, here, you may feel like challenging me:

That is, you might ask, "does someone leaving a review on Amazon really qualify as a kind of social interaction?"

And I'd say:

Not surprisingly, I'd like to argue that we should consider that within scope.

The more obvious way to see this is to think of such a review as an act that may cause someone to change their mind, make a decision, such as to buy.

But you'll see another perspective in the reading I'll be assigning to you today. More on that later.
Now, given that we have all the interactions manifested in natural-lang. format, what can we do with them?

We'll be pursuing two opportunities in this course:

Two opportunities of interest:

1) Analysis: use NLP to understand interactions (and interactors) better. ex: are there language cues that correlate with successful interactions? (non-) ideas?

(phrasings of )

Can we learn which participants are engaged/crucial/agree?

(dis) (not) (dis)

Standard approach: mining from large-scale corpora ("bodies" of "text")

2) Facilitation: use NLP to improve interactions, or cause them

ex: make a hidden state evident (confusion, antagonism, etc...) Maybe even intervene?

I want to bring to your attention a particular W:

It's nice when these go hand-in-hand.

But speaking of...

But speaking of opportunities, I want to take this opportunity to talk about another facet of this course.

This word "opportunities" here—notice it doesn't say "solutions". This is meant to indicate that this class is a research-oriented course.
And the reason is that it's fun to say that.

This can't be a survey of known approaches to this problem because there aren't really a set of established approaches.

... just understanding processing language "in the wild" is hard.

**Example review excerpt:**
"Read the book" positive or negative opinion?

[Entry due to Bob Blah]
5 people didn't see all non-native speakers of English.

[or, "She ran the gamut of emotions from A to B"]

[or, "If you like this fragrance, please wear it @ home and close your window shut."]

We're close to you having to know that scents smells can be treated this way.

You have to know that the neural process is quit the size of a rose.

From a human or tell the bio of context degenerate.

OK, so that's one reason this is a research-focused course.

Is because the very topic matter is one in which, in a sense, very little is known.

There's also a second reason why this is a research-oriented course, and that's it's b/c I know not all of you are going to do research in NLP for a living.

Why reason #2:

... 6000-level courses are training for PhD-level research.

So much of the way this course is organized is driven by the desire to provide training in the research process, whatever your eventual research interests turn out to be.

So this end product will be a research project, and the goal is throughout the semester, is to work through the process of doing a project.
How does this agenda inform the course structure?

First the early part of the course is going to interleave presentation of fundamental material with you doing a pilot research study right off the bat.

- show webpg. show prior year to demonstrate lots of info
- show course structure

So as you can see, your first assignment is released as of today.

You’re tasked w/ coming up with a research idea based on those readings, and do a pilot empirical study.

Notice that simultaneous with that assignment, I’ll be doing some lectures on online reviews and conversation threads.

After that... some break lectures.
then some student-led versions of AI, where we read papers, try out some research ideas inspired by them.
and then there’s the final project portion of the course.

AI is representative of the spirit of the course, so intended or not

ended @ about 11:10.
About 30 showed up.