Writing up and presenting your work

Bill MacCartney and Christopher Potts
Stanford Linguistics

CS 244U: Natural language understanding
May 13

Two workshops

- Apr 15: Workshop 1: Project planning and system evaluation

  May 6: Lit review due [link]

  May 25: Project milestone due by 11:59 pm [link]

- Today: Workshop 2: Writing up and presenting your work

  Jun 1, 3: Three-minute in-class presentations [link]

  Jun 10: Final project due by 3:15 pm [link]

$\Big($Policy on submitting related final projects to multiple classes [link]$\Big)$

Inspiration

It's nice if you do a great job and earn an A on your final project, but let's think bigger:

- Many important and influential ideas, insights, and algorithms began as class projects.

- Getting the best research-oriented jobs will likely involve giving a job talk. Your project can be the basis for one.

- You can help out the scientific community by supplying data, code, and results (including things that didn't work!).

EMNLP short papers

- The EMNLP short paper deadline is June 15, 2015 [link]

- These are 4 pages plus up to 2 pages for references.

- The short paper track is highly selective, but it is more open to exploratory work and initial, suggestive results.

- Let us know if you think you might revise your paper for submission here, and we'll try to give you tips on how to get from 8 pages down to 4.

Inspiring past projects

https://www.stanford.edu/class/cs224u/restricted/
past-final-projects/

- Semantic role labeling
- Unsupervised relation extraction
- Solving standardized test problems
- Humor detection
- Biomedial NER
- Sentiment analysis in political contexts
- Learning narrative schemas
- Supervised and unsupervised compositional semantics
- . . .

Lectures not given

- Dependency parsing for NLU [link]
- Natural logic and textual inference [link]
- Sentiment analysis [link]
- Dialogue agents [link]
- Using WordNet (in Python NLTK) [link]
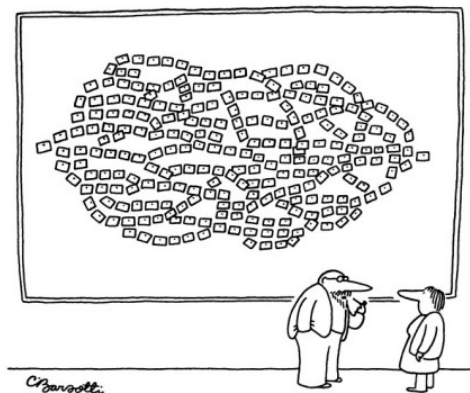- Using the Penn Discourse Treebank [link]

Plan for today

Overview

On writing papers

On conference submissions

On giving talks

Your presentations

On writing papers

"It's plotted out. I just have to write it."

The outline of a typical NLP paper

Eight two-column pages plus 1-2 pages for references. Here are the typical components (section lengths will vary):

Title info 1. Intro	2. Prior lit.	3. Data	4. Your model

4. Your model	5. Results	6. Analysis	7. Conclusion

A commonly-used structure for NLP papers

1. Opening: general problem area, goals, and context.

2. Related work (if it helps with set-up; else move to slot 6)

3. Model/proposal
   a. Data (separate section if detailed/new/...)
   b. Experimental set-up

4. Results

5. Discussion

6. Related work (if here largely for due diligence, or if understandable only after the results have been presented)

7. Conclusion: future work — not what you will do per se, but rather what would be enlightening and important to do next.

Similar to the format for experimental papers in psychology and linguistics, except that they tend to have much longer openings and section 3 often has more sub-parts on the methods used.

Stuart Shieber on the 'rational reconstruction' format

http://www.stanford.edu/class/cs224u/slides/schieber-writing.pdf

Stuart Shieber on the 'rational reconstruction' format

http://www.stanford.edu/class/cs224u/slides/schieber-writing.pdf

- Continential style: "in which one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was" [. . .] "Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius."

Stuart Shieber on the 'rational reconstruction' format

http://www.stanford.edu/class/cs224u/slides/schieber-writing.pdf

- Continential style: "in which one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was" [. . .] "Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius."

- Historical style: "a whole history of false starts, wrong attempts, near misses, redefinitions of the problem." [. . .] "This is much better, because a careful reader can probably follow the line of reasoning that the author went through, and use this as motivation. But the reader will probably think you are a bit addle-headed."

Stuart Shieber on the 'rational reconstruction' format

http://www.stanford.edu/class/cs224u/slides/schieber-writing.pdf

- Continental style: "in which one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was" [. . .] "Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius."

- Historical style: "a whole history of false starts, wrong attempts, near misses, redefinitions of the problem." [. . .] "This is much better, because a careful reader can probably follow the line of reasoning that the author went through, and use this as motivation. But the reader will probably think you are a bit addle-headed."

- Rational reconstruction: "You don't present the actual history that you went through, but rather an idealized history that perfectly motivates each step in the solution." [. . .] "The goal in pursuing the rational reconstruction style is not to convince the reader that you are brilliant (or addle-headed for that matter) but that **your solution is trivial**. It takes a certain strength of character to take that as one's goal."

David Goss's hints on mathematical style

"Two basic rules are: 1. Have mercy on the reader, and, 2. Have mercy on the editor/publisher. We will illustrate these as we move along."

http://www.math.osu.edu/~goss.3/hint.pdf

On conference submissions

http://xkcd.com/541/

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.

2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.
2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.
3. The program chairs assign reviewers their papers, presumably based in large part on their bids.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.
2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.
3. The program chairs assign reviewers their papers, presumably based in large part on their bids.
4. Reviewers read the papers, write comments, supply ratings.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.

2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.

3. The program chairs assign reviewers their papers, presumably based in large part on their bids.

4. Reviewers read the papers, write comments, supply ratings.

5. Authors are allowed to respond briefly to the reviews.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.

2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.

3. The program chairs assign reviewers their papers, presumably based in large part on their bids.

4. Reviewers read the papers, write comments, supply ratings.

5. Authors are allowed to respond briefly to the reviews.

6. The program chair might stimulate discussion among the reviewers about conflicts, the author response, etc. At this stage, all the reviewers see each other's names, which helps contextualize responses and creates some accountability.

Typical NLP conference set-up

1. You submit a completed 8-page paper, along with area keywords that help determine which committee gets your paper.

2. Reviewers scan a **long** list of titles and abstracts and then bid on which ones they want to do. The title is probably the primary factor in bidding decisions.

3. The program chairs assign reviewers their papers, presumably based in large part on their bids.

4. Reviewers read the papers, write comments, supply ratings.

5. Authors are allowed to respond briefly to the reviews.

6. The program chair might stimulate discussion among the reviewers about conflicts, the author response, etc. At this stage, all the reviewers see each other's names, which helps contextualize responses and creates some accountability.

7. The program committee does some magic to arrive at the final program based on all of this input.

Typical ACL set-up

These rating categories have prose descriptions attached to them to help clarify the program committee's intentions:

Appropriateness:	1-5
Clarity:	1-5
Replicability:	1-5
Originality / Innovativeness:	1-5
Soundness / Correctness:	1-5
Meaningful Comparison:	1-5
Thoroughness:	1-5
Impact of Ideas or Results:	1-5
Impact of Resources:	1-5
Overall Recommendation:	1-5
Presentation Format:	Poster/Talk/Both possible
Best paper possibility?	Yes/Maybe/No
Resubmission as short paper:	recommended/not recommended

Presentation types and venues

Presentation types

- Oral presentations vs. poster presentations
- Workshops vs. main conferences

Some important NLP conferences (broadly construed)

- ACL
- NAACL
- EMNLP
- EACL
- COLING
- CoNLL

- WWW
- WSDM
- KDD
- ICWSM
- AAAI
- CogSci

- ICML
- NIPS

Typical linguistics/cog-sci set-up

1 You submit an abstract or short form paper.

2 The reviewers write comments and give rankings.

3 The program committee does some magic to arrive at the final program based on this input.

On abstracts

- Important for creating a first impression, reviewer bidding, and reviewer assigning.
- A general structure:
    1. The opening is a broad overview — a glimpse at the central problem.
    2. The middle take concepts mentioned in the opening and elaborates upon them, probably by connecting with specific experiments and results from the paper.
    3. The close establishes links between your proposal and broader theoretical concerns, so that the reviewer has fresh in her mind an answer to the question "Does the abstract offer a substantive and original proposal".

Overview
000000

On writing papers
0000

On conference submissions
00000

On giving talks
0000000

Your presentations
00

On giving talks

http://www.condenaststore.com/-sp/Sign-GONE-LECTURIN-New-Yorker-Cartoon-Prints_i8476260_.htm

Basic structure

- Beginning
  - What problem are you solving?
  - Why is it important?
  - What approaches have been tried, and why have they not fully solved the problem?

- Middle
  - What data?
  - What approach? (model type, feature representations)
  - hHow to evaluate success?

- End
  - Quantitative results, graphs that slope upward.
  - Qhich features/techniques/ resources contributed most?
  - What kinds of things do we still get wrong? Examples.

(Mirrors paper stucture, but talk structure has to be simpler.)

Pullum's Golden Rules

Geoff Pullum's Five Golden Rules (well, actually six) for giving academic presentations

1. Don't ever begin with an apology.
2. Don't ever underestimate the audience's intelligence.
3. Respect the time limits.
4. Don't survey the whole damn field.
5. Remember that you're an advocate, not the defendant.
6. Expect questions that will floor you.

http://www.lel.ed.ac.uk/~gpullum/goldenrules.html

Honesty

Patrick Blackburn's fundamental insight:

Where do good talks come from?

Honesty.

"A good talk should never stray far from simple, honest communication."

Slide contents: two schools of thought

Minimalism

1. Your slides should be as spare as possible without sacrificing clarity.
2. The audience should spend most of the time listening to and looking at you.
3. Individual slides do not stay up for long or get used in more than one way.

Comparative

1. Your slides should be as full as possible without sacrificing clarity.
2. Your talk should make it easy for people to spend time studying your slides.
3. Individual slides stay up for a long time and get used to make multiple comparisions and establish numerous connections.

Slide contents: two schools of thought

A personal matter

- The minimalist view seems right for telling a story — often the best mode when time is of the essence and the audience is mainly there to learn about what your paper contains.

- The comparative view seems right for teaching; it's the closest slides come to a full, well-organized chalkboard.

- Find the style that works for you. As long as you think long and hard about what it will be like to listen to your talk, and adjust accordingly, you'll shine.

PowerPoint used for evil (not inevitable!)

http://www.edwardtufte.com/
tufte/powerpoint
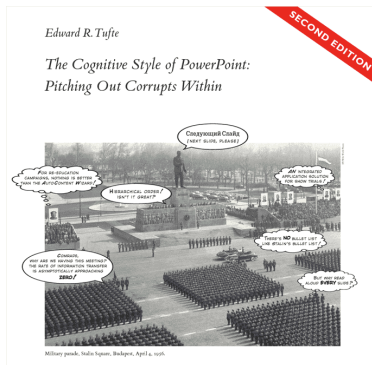
Peter Norvig: Gettysburg Address as PowerPoint

http://norvig.com/Gettysburg/

More mundane things

- Turn off any notifications that might appear on the screen.
- Make sure your computer is out of power-saver mode so that the screen doesn't shut off while you are talking.
- Shut down running applications that might get in your way.
- Make sure your desktop is clear of files and notes that you wouldn't want the world to see.
- If using PowerPoint: have a PDF back-up just in case.
- Projectors fail often; always be prepared to give the talk without slides.

The question period

- This is the most important part of the conference presentation.

- It should be a chance for the audience to gain a deeper understanding of your ideas. When the entire question period has this aim, it is a joy.

- But sometimes other things happen.

- Try to pause for one second before answering each question.

- Never say "I have no idea" and leave it at that.

- When floored, say: "I have no idea, but what . . . "

- Most questions won't make total sense to you. Your questioner doesnt know the work all that well.

- You'll be a hit if you can warp every question you get into one that makes sense and leaves everyone with the impression that the questioner raised an important issue.

Your presentations

The Night Before the Big Meeting Frank Receives a Visit from the PowerPoint Fairy.

Lightening talk limitations

- You have only 3 minutes!
- Prepare — you have only 3 minutes, and thus you can't waste time repeating yourself, figuring out how you want to state things and so forth.
- Practice — nothing in your slides should surprise you; for every slide, you should have a rhetorical plan of action.
- Coordinate — if you more than one person from your group is speaking, practice the transitions carefully so that they don't waste time.

Practical details

- Give your presentation file an informative name — "talk.pdf" is bad.
- Send us your slides in advance if possible.
- Otherwise, bring them on a thumb drive.
- Don't rely on the presentation machine having specific fonts. (PDF is safest)
- When you are on deck, wait against the wall near the projector.