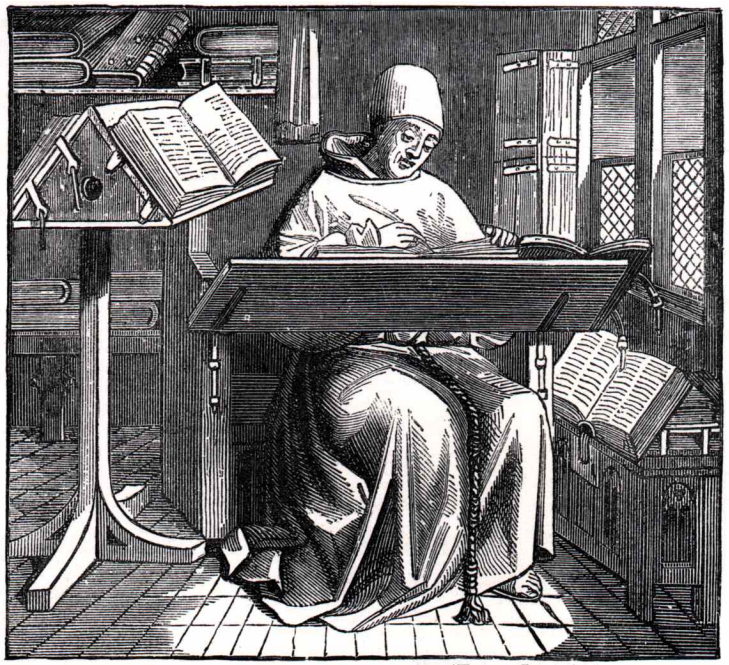
Writing Well [in NLP & ML]



SCRIPTORIUM MONK AT WORK. (From *Lacroix.*)

Course Webpage

https://sites.google.com/site/spflodd/home

Bad at writing?

- Writing is a skill:
  - You will get better by doing it
  - You will get better by getting feedback
  - You will get better by reading **good writing**!
- Not a native English speaker?
  - ***Not a problem!***
  - Good research writing is about **good ideas** and **clear thinking**, not a big mental lexicon

Your Job as a Writer

- **You are writing for your readers.**
  - To teach your reader something you figured out
  - To convince your reader of something
- You are **not** (primarily) writing for **you**
  - ***Not your job: to show how clever you are***
  - It is okay to be wrong—it is **not** okay to be unclear
- Okay… you are kind of writing for you
  - Writing helps you clarify your ideas
  - Writing lets you get feedback from others

Your Idea

- Figure out what **your idea** is
- Make sure the reader knows what your idea is. **Be 100% explicit.**
  - *"The main idea of this paper is…"*
  - *"The goals of this article are to characterize the core ideas of SMT and provide a taxonomy of various approaches."*
  - *"In this section we present the main contributions of this paper…"*
- This belongs at the very beginning of the paper
- Good ideas that are not distilled = bad paper!

Who is Your Reader?

- **Conference paper**
  - Reader: (*your home conference*) you, except you spent the last six months/year doing something else
  - Reader: (*a new conference*) pick an author who publishes there, and imagine him reading the paper
- **Journal article**
  - Reader: someone working in the journal subfield (NLP, ML, Theory), in particular: *those who work on different problems*
- **Dissertation / Book**
  - Reader: someone from a broad field (Computer Science, Physics)
  - Trick: *Imagine reading your dissertation in 10 years*
  - Anything you depend on that is "hot right now" needs to be contextualized and explained in terms of **stable common ground**

Your class research paper ≃ journal article

Structure [conference paper]

- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

Imagine Your Reader

- Your reader will determine
  - What notation are they be familiar with
  - What level of detail will be appropriate
  - What terminology will be appropriate
- **Respect** your reader
  - Don't bore your reader – ***get to the point***!
  - Organize your writing logically – *don't make the reader work more than necessary*!
  - People can be (irrationally) attached to their theories, methods, models – *don't be too harsh*!

Pitfalls when Imagining Your Reader

- Do ***not*** overestimate your readers
  - We are ***not*** as knowledgeable as you!
  - We are ***not*** as clever as you!
  - We will read your paper in **minutes**, **hours**, or **days**… *You have worked on it for **weeks**, **months**, or **years**!*
    - Lemma: Don't assume we have read all your citations. A citation is not shorthand for "go read X".

Writing for a Reader: Questions

- Are you introducing a new problem?
  - Is the problem obviously important?
  - Do you need to convince them it's important?
- Are you introducing a new technique?
  - Benefits relative to alternative techniques
  - Costs relative to alternative techniques [be honest]
- What is difficult to understand?
  - Algorithms [correctness, complexity]
  - Theorems [proofs, intuitions]
  - Models [assumptions]

For Example

- Pick examples that
  - Illustrate the easy case easily
  - Illustrate the simplest complicated case easily
  - Are **concrete**

John	proved	correctness	>	w1	w2	w3
PN	VBD	NN		t1	t2	t3

- Use a running example
  - Return to the same example throughout the paper
- Structure
  - Concrete → Abstract

An ounce of intuition is worth a pound of formalism!

Structuring a Paper

- Start with the known, move to the new
- Starting out
  - Identify a practical problem in need of solving
  - Identify an example illustrating some unexplained phenomenon
    - unexplained pattern of results
    - inconsistency between theory and reality
- Progress logically to new material
  - What is your proposed solution/explanation?
  - How do you express your solution formally?
  - Why did you choose this solution?
  - What did you do to realize this solution (experiment, proof, etc.)?
  - Results
  - Analysis

Structuring a Paper

- What is logical structure?
  - Getting you to the idea/insight/contribution in the most direct way
- What is ***not*** logical structure?
  - Recapitulating how you got to an idea
  - Don't make your reader suffer the way you did!
  - Example:
    IBM Model 3 was invented several years before **IBM Model 1** [numbering models is not great]

The Introduction

- Identify the problem you are solving
- **Clearly list your contributions**
  - Your contributions drive the structure of the whole paper
  - **For a survey paper**: Your contribution is a convenient way of understanding a bunch of related techniques / problems
- For an 8-page paper: intro = 1 page MAX
  - *No, your paper is not special*
- For a 16-page paper: intro = 1 + 1/2 pages MAX
- For a 32-page paper: intro = 1 + 1/2 + 1/3 pages MAX

Do not make the reader guess what your contributions are!

No "rest of this paper is..."

- Not:

> "The rest of this paper is structured as follows. Section 2 introduces the problem. Section 3 ... Finally, Section 8 concludes".

- Instead, **use forward references from the narrative in the introduction**.
  The introduction should survey the whole paper, and therefore forward reference every important part.

The most common of these approximations is the max-derivation approximation, which for many models can be computed in polynomial time via dynamic programming (DP). Though effective for some problems, it has many serious drawbacks for probabilistic inference:

1. It typically differs from the true model maximum.
2. It often requires additional approximations in search, leading to further error.
3. It introduces restrictions on models, such as use of only local features.
4. It provides no good solution to compute the normalization factor $Z(f)$ required by many probabilistic algorithms.

In this work, we solve these problems using a Monte Carlo technique with none of the above drawbacks. Our technique is based on a novel Gibbs sampler that draws samples from the posterior distribution of a phrase-based translation model (Koehn et al., 2003) but operates in linear time with respect to the number of input words (Section 2). We show that it is effective for both decoding (Section 3) and minimum risk training (Section 4).

Problems of standard approach that we are solving.

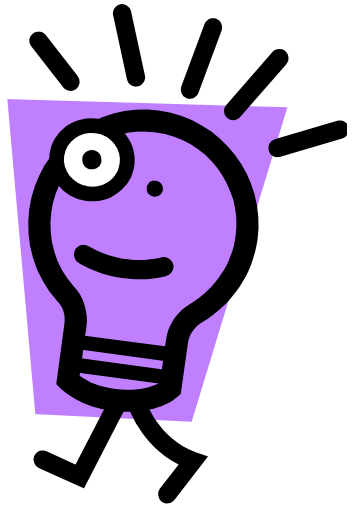
We didn't mention the conclusion!

Structure [Conference Paper]

- Abstract (4 sentences)
- Introduction (1 page)
- ~~Related work~~
- The problem (1 page)
- My idea (2 pages)
- The details (5 pages)
- Related work (1-2 pages)
- Conclusions and further work (0.5 pages)

No related work yet!

Related work



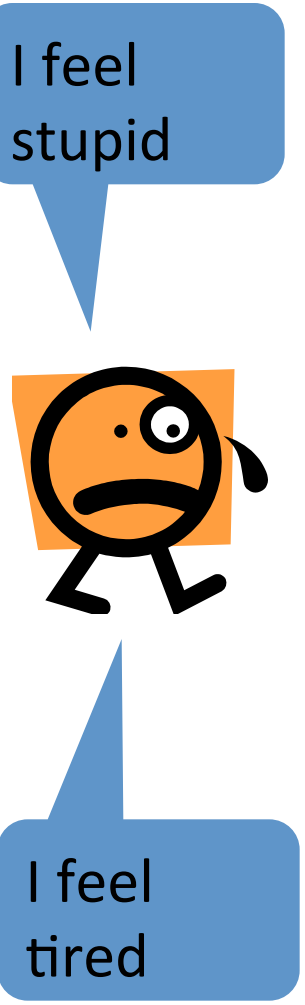
Your reader

Your idea

We adopt the notion of transaction from Brown [1], as modified for distributed systems by White [2], using the four-phase interpolation algorithm of Green [3]. Our work differs from White in our advanced revocation protocol, which deals with the case of priority inversion as described by Yellow [4].

No related work yet!

- Problem 1: the reader knows nothing about the problem yet; so your (carefully trimmed) description of various technical tradeoffs is absolutely incomprehensible
- Problem 2: describing alternative approaches gets between the reader and your idea



Tips for Good Writing

Get Started

- Writing is the best way to develop your ideas
- You may not have a completely focused idea when you **start**, but you **must** have a completely focused idea when you **finish**.

Ask People for Help

Get your paper read by as many colleagues and friends as possible

- Explain what you want (“**I got lost here**” is much more important than “bayes should be capitalized”.)
- **Suggestion**: Ask your reader to explain your contribution back to you. Did they get it right?
- An expert can check details, but the logic of any paper should be comprehensible to a non-expert.
- **Remember**: Each reader can only read your paper for the first time once!

Advice: Verbs

- ***Feel / think***. Avoid "*We think…*", "*We believe…*". Obviously you do. **Don't hedge**
- ***Present / describe*** and friends. *We now present the wombat feature*... Did you invent it? Are you reviewing it? **Present** is ambiguous. **Use a non-ambiguous verb**
- **Use strong verbs**. "*We introduce the novel GAGA algorithm*" is stronger than "*We propose the GAGA algorithm.*" Good verbs: **introduce, validated, verified, demonstrate, show, proved**
- **The passive voice is okay, really**

Advice: Nouns

- Avoid **pronoun this**. "*This raises questions…*" Prefer instead **demonstrative this**: "*This pattern of results raises questions…*"
- (Smith et al., 2012) is **not** a noun. However, *Smith et al. (2012) are the authors of a very nice paper*.

Advice: Adjectives & Adverbs

- Avoid value-judgment adjectives. Bad:
  - We present an important algorithm.
- Good [**verifiably true**]:
  - We present a novel algorithm.
- **Better** [**true and precise**]:
  - We present a novel, polynomial time decoding algorithm that is a linear program relaxation of the ILP.
- Use adverbs ***sparingly***.

Advice: Discourse Connectives

- The end of every sentence is an opportunity for a reader to get bored and give up.
- But, discourse connectives signal the logical relationship that the next sentence will have to what came before. This keeps them going:

However,
As a result,
Therefore,
Similarly,
On the other hand,

Use simple, direct language

NO	YES
The object under study was displaced horizontally	The ball moved sideways
On an annual basis	Yearly
Endeavour to ascertain	Find out
It could be considered that the speed of storage reclamation left something to be desired	The garbage collector was really slow

Advice: Notation

- Explain any notation you use in a natural way, ***then*** give the formula.
- Also applies to writing: don't define a concept after you start referring to it.

Advice: Formatting

- This is a minor detail, but it makes a huge difference
- Bold-face = new term
- Italics = emphasis
- Italics or special font = text/code example
- Use \textsc{smallcaps} sparingly
- Use ``left and right'' quotes (LaTeX cares)

Advice: Abbreviations

- Okay for terms you will use many times; introduce the first time and then be consistent
  - We define a prior based on a finite mixture of Dirichlet Processes (DPs). (Use "DP" forever after this)
- Don't give abbreviations you never use or or use just once or twice
- Never introduce an abbreviation twice
- When in doubt, don't abbreviate.

Summary

- If you remember nothing else from today:
  - Identify your contributions
  - Use clear, concrete examples
  - Move from concrete to the abstract
  - Use precise, non-hedging language

Thanks To & Further Material

- Philip Resnik (UMD, my advisor)
- Simon Peyton Jones (MSR Cambridge)

  http://research.microsoft.com/en-us/um/people/simonpj/papers/giving-a-talk/giving-a-talk.htm

  [Bonus: how to give a research talk;
  how to write a research proposal;
  video of him talking about good writing]

  [*Several slides are taken from SPJ's posted talk*]

- Jason Eisner (JHU, Noah's advisor)

  http://www.cs.jhu.edu/~jason/advice/how-to-write-a-thesis.html

Further Material

- Geoffrey K. Pullum (Edinburgh)

  http://www.lel.ed.ac.uk/grammar/passives.html