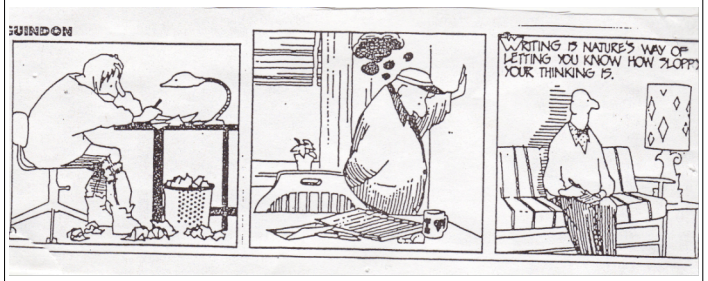
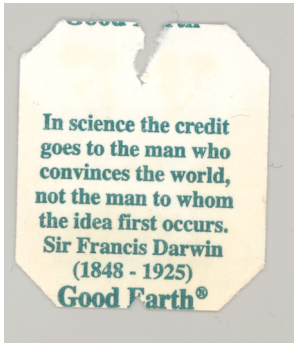
Mini Lecture on Writing*

*At least when Kobus is the reader

Reasons for Writing





Wisdom from tea dipper handle

General Comments

- Critical skill --- develop it!
  - Engage the challenge of improving over time
  - Don't treat writing as something to be avoided
- Lots of information available
  - web, books, etc.
  - UA has quite a few resources (classes, clinics)
- Get feedback
  - Must learn the meta lessons
    - Most colleagues are happy to "trash" your writing, but it gets to be a bore if the same problems keep appearing

What to say?

- Writing has purpose
  - What is it for the case at hand?
- Answer the implied questions
- Ideally, what you want to say
  - Add value

Some important points

- Focus on the purpose, story, and organization
  - Cut and paste is evil
- The reader does not know what you are talking about (until you tell them)
  - Clarity above everything else
- The reader is not obligated to read
  - Make it interesting and enjoyable
  - Make it easy for them to skim (organization)

Getting it right

- Study the order
  - Often the difference between OK and good is the order
- Try reading at different scales
  - Headings
  - Headings and sub-heads
  - Heading, sub-heads, and first sentence of each paragraph
    - Paragraphs should lead with the main points. These must get through even if the reader is going fast.

Getting it right

- Read it as if you are a different person
- Get feedback from different persons
  - If they are having trouble understanding it, it will often be your problem to figure out what to do about it

Potentially useful tricks

- Think of writing as problem solving
  - How best to get the message across to the target audiences for the desired purpose?
- Annotate sentences informally
  - Explain the purpose and meaning of each sentence as if you were explaining to someone how you are solving your writing problem.
  - Sometimes this will tell you what to write instead!
- Read what you wrote out loud (suggested by Tasneem)

Class paper example outline

- Abstract
- Introduction
  - Basic high level description of what you are doing
    - what is the basic problem
      - why is it interesting
      - why is it important
  - What is new about the work
  - What is the literary context of the work

Class paper example outline

- Problem statement
  - What is the data
  - What is the problem, relatively precisely, perhaps formally
  - Key assumptions

The problem statement might be part of the introduction, or it might be obvious by the time the introduction is done. BUT, before the reader embarks into the meat of the matter they MUST know what problem you are solving. They should be able to ask themselves, “how would I do this?”

Class paper example outline

- Model
  - Assumptions should be clear
  - A model links ideas/theory to observations
    - How this happens should be clear
- Inference
  - How you find the parameters of the model
  - Can be a serious section, or so trivial that it can be combined with the previous
- Implementation details
  - This may be part of experiments instead
  - Preprocessing and feature selection might go here

Class paper example outline

- Experiments
  - Measures
  - Data
  - Experimental details (e.g. preprocessing)
  - Results
    - Don't forget to say what they mean!
    - A table of numbers is very boring
- Discussion
  - Often not relevant, but may be critical if you are trying to argue for a different way of thinking

Class paper example outline

- Conclusion
  - Over-rated
    - It does not deserve much space or writing time
  - If your paper warrants "discussion" then you might do "discussion and conclusions" or just "discussion"
  - "Future work" is boring. Be brief or omit it entirely*.
    - If needed, you can claim obvious extensions within the text with wording text like "we have yet to experiment with different numbers of clusters".
    - If you have future work, make sure it is not something you could have arguably done for this paper---it should add value.

*Looking back at my conference papers I have a few sentences on it about 1/2 the time (sometimes with regret). I don't seem to use it in journal papers.

Iterative Drafting

- Various useful approaches
  - Outlines, storybording, just spewing
- Don't worry about the prose on early drafts
  - First figure out what to say
  - Focus on the strategy
    - message for the expected audience(s)
    - order and organization
  - At the deadline approaches, cleanup is easier than creating content
- On early drafts don't worry much about length
- Most of what you write will be replaced or changed

Making Sentences*

- Be direct, specific, informative, and concise
- Use first person active voice
- Don't waste space on the obvious or mundane
  - This is different than what seems obvious because you are familiar with what **you** did or what **you** want to say.
- Provide details and depth
  - The reader might not care about details, but make sure you are not skipping content
  - How you actually solved the problem does matter
  - Make it easy for the fast reader to skip details

*Not so critical in early drafts

Figures and Tables

- Figures / tables should have a clear purpose
- Write good (often substantive*) captions
  - Tell the reader that what they should conclude from the figure/table
  - Can be a good place to ensure that the reader has easy access to details about the particular experiment
- Alternative channel of explanation
  - Use figures and tables and their captions as a second channel to explain what you are doing in a different way
  - One should be able to get quite a bit of information by just looking at the figures and their captions.

*Some writers will try to save space by explaining figures in the text---I disagree.

Figures and Tables

- Use adequate resolution
- Make sure it is clear if bigger is better or worse
- Consider providing error estimates
  - Bars in graphs
  - Errors in parenthesis (with explanation)
- Be careful with significant figures
  - Don't have digits that are obviously meaningless

Example table and caption.

Algorithm	Freq	Range
Empirical distribution	2.05 (1.8)	0.74 (0.3)
Dependent MMMM	4.89 (1.2)	2.50 (0.5)
Correspondence MMMM	3.80 (1.4)	2.25 (0.6)
Restricted dependent MMMM	8.00 (1.3)	5.65 (1.1)
Restricted correspondence MMMM	7.44 (1.5)	5.37 (1.1)
Result using the initial estimate	5.88 (0.7)	5.26 (1.3)
The main method proposed here	**8.67** (0.9)	**7.40** (1.0)
Theoretical maximum	22.3 (3.3)	20.3 (3.2)

Table 1. Performance on training data. This table shows quantitative region labeling results for the method proposed in the text compared with two versions of the multimodal mixture model (MMMM). These are averages over 16 splits of 27,128 images including 1014 semantically region labeled images among them. The numbers in parentheses is an error estimate based on the variance of results over the splits. As discussed in the text this is the performance on training data. The maximum achievable result (last row) with the ground truth scoring system is a function of the segmentation quality. These results show a substantive improvement over the stiff base line (the restricted dependent version of the multi-modal mixture model) in the case of the semantic range measure, and a modest improvement in the case of frequency correct. Increasing the frequency correct number in an absolute sense by improving the labeling of less common words is difficult on this data set because it is heavily weighted towards common words.

Example figure and caption.

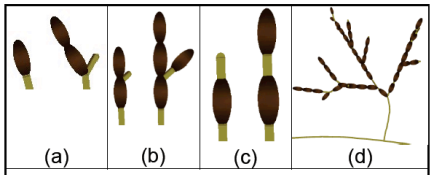


Figure 2. Building *Alternaria* sporulation structures from simple components (spores, hyphae) and simple rules that are applied stochastically. The pairs (a,b,c) illustrate possible growth rules where the structure on the left is replaced by the structure on the right. The application of rule (a) creates a branch; rules (b,c) add spores; (d) the result of applying the rules stochastically 60 times.

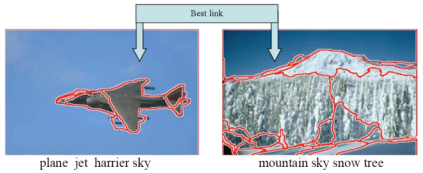


Figure 1. A friendly example illustrating some of the ideas in this paper. The associated words are shown below each image. Regions that are close together in feature space will propagate shared words, of which there is only one ("sky") in this example. Assuming such evidence for sky in the left image, exclusion reasoning means that we have some chance to get the other labels correct, as the remaining words will be assigned to the remaining regions that are most unlike sky. In this work we assume (incorrectly) that the synonyms "plane" and "jet" and the specific term "harrier" all refer to different regions. This problem can be mitigated with language processing (e.g. [19]), which would allow us to establish the labels for all regions in the left hand image.

A second example figure and caption. This was on the first page of the paper. It can help draw the reader in if there is a "marketing oriented" figure on the first page.

Citations

- Be generous
  - You did not invent computer science
  - The reader (reviewer) might be someone you should have cited (or one of their friends).
- Try to add content to literature reviews.
  - It takes effort to write about other people's work (for me, anyway), but it the reader will appreciate added insight and thoughtful connections.
- Try to avoid using citations as noun
  - Instead of
    - "[2] shows that X is good."
  - Use
    - X is good [2]

Math

- Math tells a stories too!
  - Clarity above all
  - Try to be modular
  - Order counts (top down makes sense to me)
  - The math and the English should work well together
    - A non-mathematical reader should be able to understand the flow of the logic
- Work hard to get good notation
- Define your symbols and explain your notation
- Small stuff
  - Number all equations (there is debate)
  - Punctuate equations (even more debate)