Reflection on the new page limit for MSc dissertations

— Iain Murray, September 2019

In 2018/19 the Informatics MSc dissertations (course code INFR11077, not design Informatics or CDTs) had a 40 page limit on the main content, from Introduction to Conclusion. Appendices were allowed outside that limit. Students had to use a provided template, which carefully outlined the rules.

The number of content pages that we had to mark fell dramatically:

The mean fell from 49.4 to 35.7. Or 274 fewer content pages to read for those marking 20 projects.



Many students did use appendices. However, the overall size of the dissertations did fall too:



[The fixed x-axis of the plots cut off a single project from this year that had 226 pages including appendices, which contained full transcripts from a user study.]

Logistics

- I checked the page-counts and formats and it was very quick. (Preparing this document took longer.)
- I emailed six students $(\langle 2\% \rangle)$ with invalid submissions. Markers didn't need to do anything.
- I checked the projects before and after the deadline (because I was feeling kind) and then every day as late submissions dribbled in. It was quick, but *yet-another* task in my diary for a whole week.

Possible actions for next year

Please do not suggest changes to the policy until you have read the original policy and its rationale: http://web.inf.ed.ac.uk/sites/default/files/atoms/files/project_page_limits_v2.pdf

- I'm prepared to run dissertation checking next year, although anyone good with Unix could do it. At least if I polish the scripts slightly. It might make sense to make the task part of an existing admin role. Or maybe ITO could learn (they do run my scripts to drive GPG on the exam machines).
- If ITO could check the late submissions as they come in (only a few projects per day), the main task would be restricted to one quick session.
- If easy to do, the submission form should have a compulsory text box asking students to declare the number of content pages. Students then couldn't submit an over-long project without deliberate misrepresentation.¹ (Suggestion from Ian Stark.)
- To keep in mind: there are currently half as many UG projects as MSc projects each year. However, it looks like there will be >300 UG projects in 2022/23.
- Programme directors of new MSc programmes must follow these rules if using the INFR11077 course. They must also check that the IAT_{FX} thesis template on DICE and Overleaf covers their degree.

Some detail (you can stop reading here)

I wrote three scripts, which performed the following tasks:

- 1. Download all of the submitted .pdf MSc dissertations submitted for INFR11077. I ignored dissertations with other course codes (CDT students, Design Informatics), which had different rules.
- 2. Count the content pages in the dissertations.
- 3. Take a single page from within each dissertation, stamp the student number onto the page, and concatenate all into one document. It took <2 minutes to page down through this whole document (zoomed to show a page at a time). The two gross format violations popped out.

These scripts are slightly fragile, they're for a one-off task. Current limitations include:

- The user has to copy-paste the INFR11077 class list from Euclid to filter the submissions, and pipe it through a command to get student numbers.
- The user has to copy-as-curl a link out of the Firefox developer tools before being able to download from DICE authenticated pages. This fiddly step could be removed by using the (unsupported) CosigNego tools written by computing support.
- The download script will download submissions that were missing on the previous run, but the other tools don't have a slick way to give incremental updates. I found it quicker to do some manual things to check the trickle of late projects than write more code.

If someone else wanted to take over I could help polish a bit more. Some things aren't worth doing though:

- The manual inspection in Step 3. above can be automated further, but having used more complicated scripts (with some nasty dependencies) for conferences, I don't think it's worth it to save 2 minutes.
- Having run a paper checker on a submission server for conferences, I don't think we should do that. I wouldn't want to inflict that headache on support.
- There were ~4 dissertations that confused the scripts, which I checked by hand. It would be really hard to automatically understand all the weird things 1% of people might do though.

^{1.} There would still be a problem with those who accidentally break the style by copy-pasting long lists of packages into their .tex, containing packages like fullpage, which alters the margins. (Two of the six invalid projects did this, $\sim 0.5\%$ of students.)