Papers to review

In this doc, put links, etc. to papers we should review.

**Introductory paper to get us started and the one we plan on reviewing on 16/3/22**

<https://link.springer.com/chapter/10.1007/978-94-009-2195-5_12?fbclid=IwAR1lpyXZAfGDjGvAkNm5GdAJu5CpJrD8IqkyZ24NDrg6kxgfhewA7tMjKTE>

<https://link.springer.com/content/pdf/10.1007%2F0-306-47531-6_4.pdf?fbclid=IwAR37lwqfTVkaKCyrgWJuOZBFvgGKnqY24uKlH1ukDyucbMHwABOVHpZCBvw>

(Start looking at papers published more recently)

**Good paper to review on 23/3/22**

<https://www.sciencedirect.com/science/article/pii/S0360131511000418?via%3Dihub&fbclid=IwAR2GyqiC24yKcswEo-NVG0z14Oin0yAnf7W0hJEXDOIHoAub06N0NaaQxPU#bbib4>

The paper outlines three problems with CAT using current IRT models and proposes an elo rating system to try and solve these problems. Elo model seems to be a more simple model than IRT. Users and test items are given a baseline rating and their ratings (both the students’ and the items’) self-adjust based on whether the student answers the questions correctly.

IRT aims to give questions that the user has a 50% chance of answering correctly for maximum evaluation accuracy. If the probability is not 50%, more questions are required to get an accurate figure. However, the Elo system in the paper aims to give users questions with 75% expected correct answer rate. To mitigate this lack of accuracy, they add a time component where fast correct answers indicate higher ability than slower correct answers and would hence give more elo to the user.

**Similar to the above paper**

<https://www.fi.muni.cz/~xpelanek/publications/CAE-elo.pdf>

Important point made: the nature of the Elo system makes it more suitable for adaptive learning and small tests over high stakes examinations.

* Candidate ability estimation and item difficulty estimation are done on the fly, and after a period of calibration, it will eventually converge to a value with low error margin.
* Comparisons of estimation can be made relative to other estimations subject to the same Elo testing.

For estimating item difficulty at random item selection, the proportion correct method is as effective as IRT and Elo methods. However, when items are selected using adaptive selection, the proportion correct method gives poor estimates, whereas Elo would perform similarly to IRT despite its computational simplicity.

Once the Elo rating system has good estimates of item difficulties, it can estimate the skill of new students rather quickly - only 10 answers are required to estimate their ability.

“The elo rating system is closely related to the Rasch model (one parameter model) used in IRT.”

“Basic IRT models assume that a student’s skill is constant (such models are typically used in testing and it is assumed that there is no change of the skill during a test), whereas the Elo rating system is designed to track changing skill levels.” and does not make fixed assumptions about the nature of learning. Gives reasonably accurate estimates in any setting.

Points to make this Wednesday

* Elo is simpler, easier to understand (paper linked is pretty recent)
* IRT is complex, and there are many out there
* Elo is more for formative assessments (learning rather than testing), IRT is for [summative testing](https://onlinelibrary.wiley.com/doi/10.1111/jcal.12577) (eg GMAT)
* AI? Here's a paper that discusses the use of machine learning in CAT. Could not access (2020 paper from IEEE CSCI) <https://www.computer.org/csdl/proceedings-article/csci/2020/762400a649/1uGYygGpBS0>

<https://www.researchgate.net/publication/225603921_Towards_Computerized_Adaptive_Assessment_Based_on_Structured_Tasks>

<https://www.researchgate.net/publication/239773489_Using_Item_Response_Theory_and_Adaptive_Testing_in_Online_Career_Assessment>

<https://www.researchgate.net/publication/259622247_OPTIMAL_ITEM_POOL_DESIGN_FOR_A_HIGHLY_CONSTRAINED_COMPUTERIZED_ADAPTIVE_TEST>

<https://www-jstor-org.ezproxy.auckland.ac.nz/stable/3701289?seq=1>

<https://www.researchgate.net/publication/326803834_How_to_build_a_Computerized_Adaptive_Test_with_free_software_and_pedagogical_relevance>

<https://eric.ed.gov/?id=EJ1272239>

<http://jattjournal.net/index.php/atp/article/view/155939/109385>

<https://eric.ed.gov/?q=Computerized+Adaptive+Testing&id=EJ1296015>