Week 4 Video 6

Knowledge Inference: KT-IDEM and DKT

Fast-moving area

- This lecture goes into some of the latest developments in our understanding of knowledge inference during online learning

Multiple skills

- BKT can be extended to allow items to have multiple skills
- Addresses some of the same goals as PFA
- (Pardos et al., 2008; Koedinger et al., 2011)
- However, appears to be less effective than other ways of representing multiple skills (Xu & Mostow, 2012; Gonzalez-Brenes et al., 2014)

Item difficulty

- Another type of extension to BKT is modifications to include item difficulty
- Addresses some of the same goals as IRT

- (Pardos & Heffernan, 2011; Khajah, Wing, Lindsey, & Mozer, 2013; Gonzalez-Brenes et al., 2014)

KT-IDEM (Pardos & Heffernan, 2011)

- Rather than having a single P(G) and P(S) for each skill
- Each item has its own P(G) and P(S)

LFKT (Khajah et al., 2014)

- Rather than having a single P(G) and P(S) for each skill
- Guess and slip are contextually adjusted based on
  - Skill
  - Item
  - Student (past performance on other skills)

FAST+item (Gonzalez-Brenes et al., 2014)

- Substitutes logistic regression equations for the 4 parameters of Bayesian Knowledge Tracing
  - With coefficients for each skill
  - And coefficients for each item

- Better prediction of student correctness than traditional BKT or PFA

Deep Knowledge Tracing (DKT) (Piech et al., 2015)

- Based on “deep learning”, aka recurrent neural networks/long short term memory networks
- Fits on sequence of student performance across skills
  - Predicts performance on future items within system
- Can fit ***very*** complex functions
  - Very complex relationships between items over time

DKT

- Initial paper reported massively better performance than original BKT or PFA (Piech et al., 2015)

DKT

- (Xiong et al., 2016) reported that (Piech et al., 2015) had used the same data points for both training and test, due to miscommunication about the data set
- DKT doesn't do quite as well when this error is fixed – still moderately better than original PFA or BKT

DKT

- (Khajah et al., 2016) compared DKT to modern extensions to BKT on same data set
- (Wilson et al., 2016) compared DKT to temporal IRT on same data set
- Bottom line: All three approaches appear to perform comparably well

Additional Issue

- (Yeung & Yeung, 2018) report degenerate behavior for DKT
  - Getting answers right leads to lower knowledge
  - Wild swings in probability estimates in short periods of time
- They propose a regularization method to moderate these swings

Extension for Latent Knowledge Estimation

- (Zhang et al., 2017) propose an extension to DKT that uses an item-skill mapping as well as DKT
  - Latent skill still difficult to interpret
- (Lee & Yeung, 2019) propose an alternative to DKT that attempts to output more interpretable latent skill estimates
  - Still some degenerate behavior reported

Watch this space

- Ongoing rapidly moving discussion about algorithms
- Time will tell which approaches are best

Interpretability of modern approaches

- Is prediction of immediate future correctness the right indicator?
- Are skill estimates more useful when prediction of immediate future correctness is better?

- What is the ultimate goal – predicting immediate performance or understanding what knowledge students carry forward?

Next Up

- Memory Algorithms