Envisioning the Future of Cyberlearning

Ryan S.J.d. Baker
Teachers College, Columbia University
It’s an honor

• To be speaking in front of such an accomplished group of scientists
It’s an honor

• To be speaking in front of such an accomplished group of scientists

• Researching and building excellent cyberlearning systems
It’s an honor

• To be speaking in front of such an accomplished group of scientists

• Researching and building excellent cyberlearning systems

• With the potential to have a major impact on the future of American education
Individualization

• Folks have been talking about individualizing education for a long time (Rousseau, 1762; Parkhurst, 1922)
We’re starting to get there...
Dion wants to earn a minimum quiz average of 92% in his biology course. His grades so far are 88%, 95%, and 88%. Which inequality below represents the possible scores for his next quiz which will allow Dion to achieve his goal?

\[
\text{Sum of the values} \geq 92
\]

Number of values

\[
\frac{89 + 95 + 85 + x}{4} \geq 92
\]

Solve for \(x\).

\[
\frac{296 + x}{4} \geq 368
\]

- \(x > 99\)
- \(x < 99.5\)
- \(x \geq 99\)
- \(x \leq 99.5\)

The sun exerts a gravitational force on the earth as the earth moves in its orbit around the sun. Does the earth pull equally on the sun? Explain why.

Presentation of the question/problem

Student input (answers, comments, questions)

Dialog history with tutor turns student turns
Individualization requires

1. Determining something about the student
2. Knowing what matters
3. Doing the right thing about it
1. Determining something about the student
2. Knowing what matters
3. Doing the right thing about it
Determining Something About the Student

• We’ve made a ton of progress, accelerating in recent years
Stuff We Can Infer: Complex Cognitive Skill

- Programming (Corbett & Anderson, 1995)
- Physics (Martin & VanLehn, 1995)
- Mathematics (Feng et al., 1999)
- Databases (Mitrovic et al., 2001)
- Science Inquiry Skill (Sao Pedro et al., 2013)
Stuff We Can Infer: Deep Learning

• Retention (Jastrzembski et al., 2006; Pavlik et al., 2008; Wang & Beck, 2012)
• Transfer/Shallow Learning (Baker et al., 2011, 2012)
• Preparation for Future Learning (Baker et al., 2011; Hershkovitz et al., in press)
Stuff We Can Infer: Meta-Cognition

• Self-Efficacy/Uncertainty/Confidence (Litman et al., 2006; McQuiggan, Mott, & Lester, 2008; Arroyo et al., 2009)
• Unscaffolded Self-Explanation (Shih et al., 2008; Baker, Gowda, & Corbett, 2011)
• Help Avoidance (Aleven et al., 2004, 2006)
• Conscientiousness and Persistence (Ventura et al., 2012)
Stuff We Can Infer: Disengaged Behaviors

• Gaming the System (Baker et al., 2004, 2008, 2010; Walonoski & Heffernan, 2006; Beal, Qu, & Lee, 2007)

• Off-Task Behavior (Baker, 2007; Cetintas et al., 2010)

• Inexplicable “WTF” Behavior (Rowe et al., 2009; Wixon et al., 2012)

• Carelessness (San Pedro et al., 2011; Hershkovitz et al., 2011)
Stuff We Can Infer: Affect

- Boredom (D’Mello et al., 2008; Sabourin et al., 2011; Baker et al., 2012)
- Frustration (McQuiggan et al., 2007; D’Mello et al., 2008; Sabourin et al., 2011; Baker et al., 2012)
- Confusion (D’Mello et al., 2008; Lee et al., 2011; Sabourin et al., 2011; Baker et al., 2012)
- Engaged Concentration/Flow (D’Mello et al., 2008; Sabourin et al., 2011; Baker et al., 2012)
- Curiosity (Sabourin et al., 2011)
- Excitement (Arroyo et al., 2009)
- Situational Interest (Arroyo et al., 2009)
- Joy (Conati & Maclaren, 2009a, 2009b)
Sensor-free detection possible

• Recent systems have been able to infer these constructs solely from student interaction with the learning system
Example

• Automated detectors of student engagement and affect in ASSISTments (Pardos et al., 2013; Ocumpaugh et al., under review)
Process
Field Observations of Student Engagement and Affect

• Using BROMP observation protocol (Ocumpaugh et al., 2012)
  – over 40 coders trained, used in dozens of papers
• Synchronized to log files with Android app HART
Use data mining to find behaviors that co-occur with human observations

• 160 features of interaction distilled

• Small set of data mining algorithms compared:
  – Decision Trees
  – Decision Rules
  – Step Regression
  – Naïve Bayes
  – $K^*$
Model generalizability tested on new students from diverse populations

• Students in rural, urban, and suburban schools in Northeastern USA
  – Diverse in terms of SES, race, ethnicity
Result

• Models that can distinguish a bored, frustrated, confused, engaged, off-task, or gaming student
  – 63-82% of the time
  – Achieving agreement with human coders 1/3 to 2/3 as well as coders agree with each other
1. Determining something about the student

2. *Knowing what matters*

3. Doing the right thing about it
Off-Task Behavior

• We can detect off-task behavior (cf. Baker, 2007; Cetintas et al., 2009; Pardos et al., 2013)
Off-Task Behavior

• We can detect off-task behavior (cf. Baker, 2007; Cetintas et al., 2009; Pardos et al., 2013)

• Off-task behavior is continually a major focus of classroom management practice
But...

Off-Task Behavior is:

• More weakly correlated with learning and other outcomes than many other constructs

• Off-Task Behavior can foster positive collaborative relationships (cf. Goldman, 1996; cf. Barron, 2003; Kreijns, 2008)
  – E.g. a collaboration strategy

• Off-Task Behavior can disrupt boredom (Baker et al., 2011)
  – E.g. an emotional regulation strategy
So...

• Reduction of off-task behavior should probably not be a focus of Cyberlearning systems

• Though carefully leveraging and managing it may be beneficial and useful...
Focusing on what matters

Behavior in software

- Longer-term learning outcomes
  (Feng et al., 2009; Pardos et al., 2013)
- Choice of STEM or other societally useful careers
- College enrolment
- Dropout

Choice of STEM or other societally useful careers

(Feng et al., 2009; Pardos et al., 2013)
Focusing on what matters

Longer-term learning outcomes

Choice of STEM or other societally useful careers

Behavior in software

Dropout

College enrollment

(Feng et al., 2009; Pardos et al., 2013)

(Dekker et al., 2009; Bowers, 2010; Arnold, 2010; Ming & Ming, 2012)
Example

(San Pedro, Baker, Bowers, & Heffernan, in press)

• Automated detectors of engagement, affect, and learning in ASSISTments

• Can predict

• Whether a student will go to college or not, ~6 years later
  – 69% of the time for new students
Example
(San Pedro, Baker, Bowers, & Heffernan, in press)

• And the model can indicate what aspects of a student’s behavior are predictive of college attendance

• Alex is less likely to go to college
  – Top predictive factors: he is getting confused and gaming the system...

• Maria is less likely to go to college
  – Top predictive factors: she is getting bored and careless...
1. Determining something about the student
2. Knowing what matters
3. *Doing the right thing about it*
What do we do?

• When we know that a student is bored... or gaming the system... or has shallow learning... or etc. etc. etc.
Huge Space of Potential Interventions
Huge Space of Potential Interventions

- Theory can be our guide for which intervention to use...
Huge Space of Potential Interventions

• Theory can be our guide for which intervention to use...
  – But which theory?
Huge Space of Potential Interventions

• Theory can be our guide for which intervention to use...
  – But which theory?
  – And is it applicable in the current situation?
Huge Space of Potential Interventions

- Theory can be our guide for which intervention to use...
  - But which theory?
  - And is it applicable in the current situation?
  - And, knowing that detection will always be wrong sometimes...
Huge Space of Potential Interventions

• Theory can be our guide for which intervention to use...
  – But which theory?
  – And is it applicable in the current situation?
  – And, knowing that detection will always be wrong sometimes...
  – What are the relative costs of incorrectly applied interventions, compared to the benefits of correctly applied interventions?
What we need
What we need is **more data**!
What we need

• Lots and lots and lots and lots and lots and lots and lots of randomized studies comparing individualized interventions
Deployed automatically through online learning systems

- Try at small-scale and ramp up successes automatically
- Try, fail, and try again quickly!
Deployed automatically through online learning systems

• Try at small-scale and ramp up successes automatically
• Try, fail, and try again quickly!

• Some systems now used for hundreds of studies conceived by outside experimenters, a powerful tool for progress
  – PSLC LearnLabs
  – ASSISTments
  – MOOC platforms possibly moving in this direction?
In these studies...

• Automated detectors can be used not just to drive interventions
In these studies...

- Automated detectors can be used not just to drive interventions
- But also to understand the results of interventions
What is the effect of my new confusion intervention on...

- Future confusion
- Future boredom
- Future gaming the system
- Future learning of the same topic
- Future learning of new topics
- Etc. etc.
With these methods...

- We can not only create more individualized learning experiences
- But understand the full range of effects of our interventions
With these methods...

• We can not only create more individualized learning experiences

• But understand the full range of effects of our interventions

• Creating a feedback-loop that makes online learning more and more effective and engaging each year!
Thanks!

• Twitter @BakerEDMLab
• Facebook “Baker EDM Lab”
• Publications http://www.columbia.edu/~rsb2162/publications.html
• “Big Data and Education” on Coursera https://www.coursera.org/course/bigdata-edu
• Masters in Cognitive Studies with Focus on Learning Analytics, Starting Fall 2013