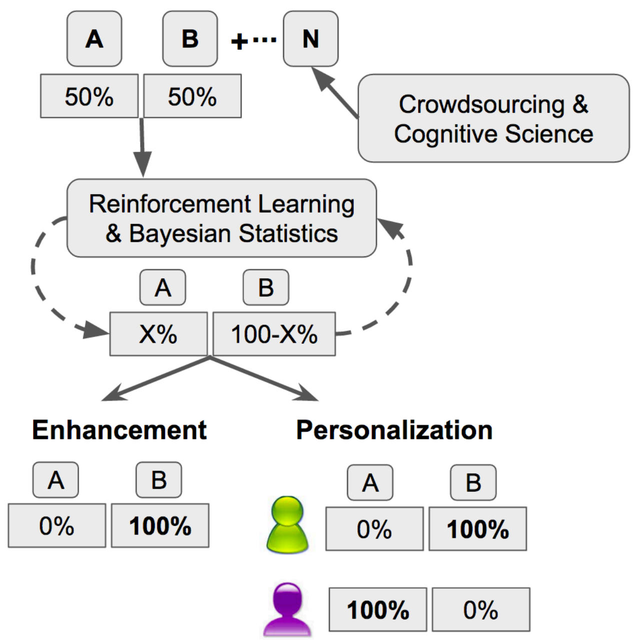
Combining Active Learning & Human Computation for A/B Experimentation: Perpetually Enhancing and Personalizing User Interfaces

Joseph Jay Williams
www.josephjaywilliams.com, williams@cs.toronto.edu
Intelligent Adaptive Interventions group
University of Toronto, Computer Science



[I'm originally from the Caribbean, Trinidad and Tobago]

Open up Gdoc at URL tiny.cc/uofa to questions/comments

Example Problem: Dynamic A/B test to enhance/personalize email

[Returning to course?]

Dear Sam,

We hope you have enjoyed the opportunity to explore Statistics and R for the Life Sciences. It has been a
we can improve the course for future students?

Each of the links below connects to a short survey. Please click on the link that best describes you.

- I plan on continuing with the course
- I am not continuing the course because it was not what I expected when I signed up.
- I am not continuing the course because the course takes too much time.
- I am not continuing the course because I am not happy with the quality of the course.
- I am not continuing the course because I have learned all that I wanted to learn.
- I am not continuing the course now, but I may at a future time.

Your feedback is very important to us. Thank you for registering for Statistics and R for the Life Sciences.

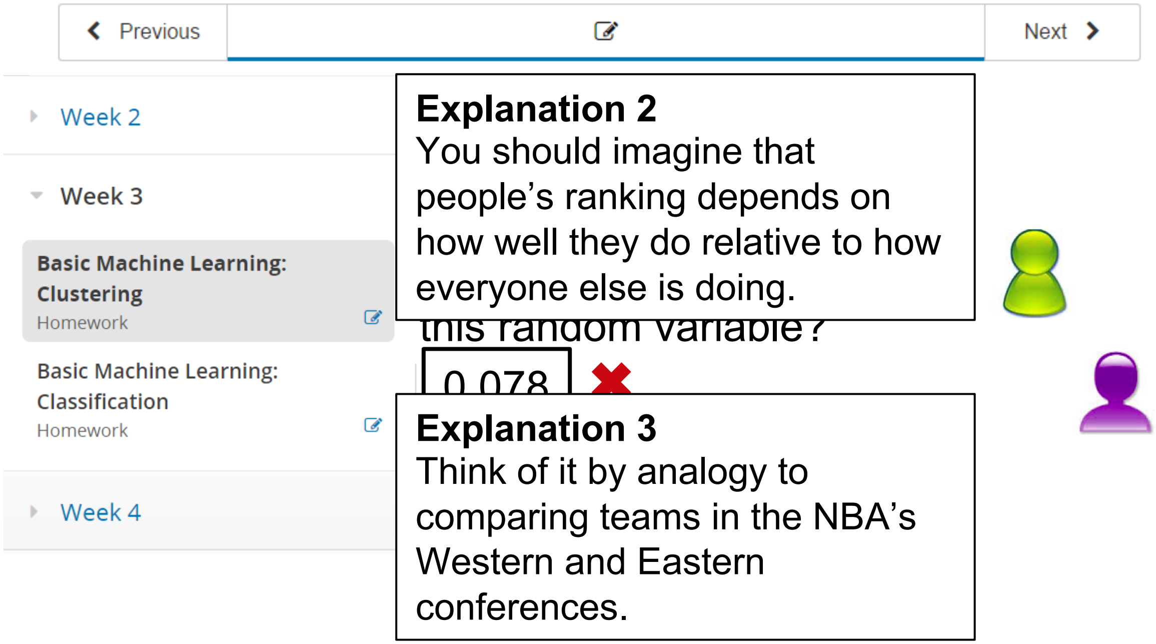
Follow the link to opt out of future emails:
Click here to unsubscribe

Action Space:
[Subject Line = 1, 2, 3]
[Intro Message = 1, 2, 3…]
[**Response Format** = 1, 2, 3…]

Context:
[**Age**]
[**Number Days Active**]
[Country]
[Activities in week 1]
…
…
..

Perpetually Improving User Interfaces

Vision: Systems that perpetually improve – like real teachers



Apply Bandit Algorithms (RL more generally?) to A/B testing

Approach: Making Experiments Collaborative, Dynamic, Personalized

A B + … N
Continually add conditions

50% 50%

Outcome Metric

Dynamic Analysis

A B
X% 100-X%

Enhancement

A B
0% **100%**

Personalization

A B

0% **100%**

100% 0%

MOOClet

github.com/kunanit/mooclet-engine

HCI, Ed, Health

Cognitive Science

Cognitive Science 2010
J. of Exp. Psych., 2013

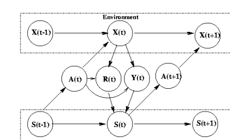
Crowdsourcing & Human Computation

CHI 2016,
ACM LAS 2016

Instructional Design

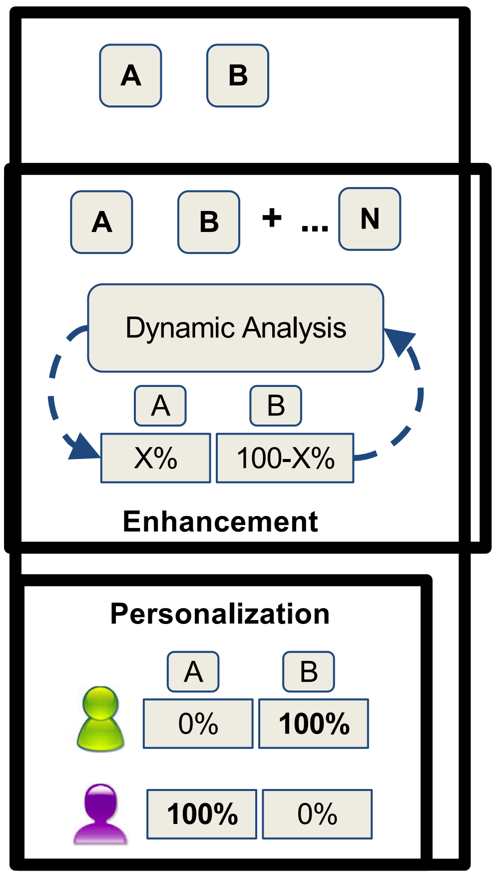
EDM 2015
IJAIED 2016

Bayesian Statistics & Machine Learning



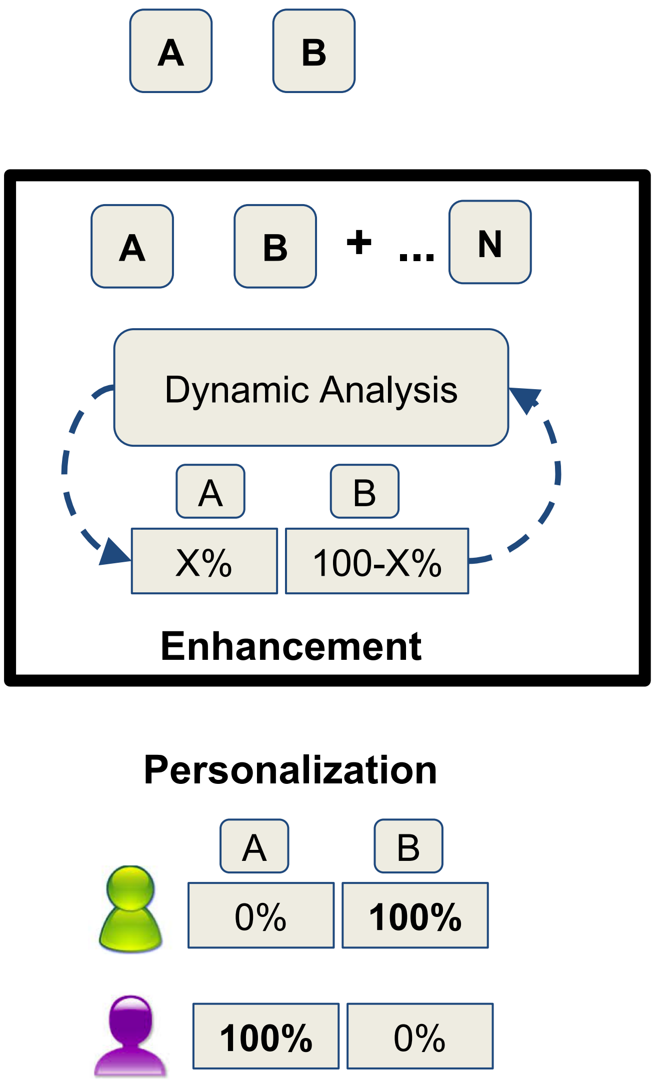
NIPS 2008, UAI 2013
ACIC 2016

Overview



- Vision: Perpetually Improving Systems
- Approach: Collaborative, Dynamic, Personalized Experimentation
- 1. Crowdsourcing & Dynamically Testing Student Explanations
- 2. Instructor-Centered Experimentation
- 3. Discovering how to personalize
- Future
  - Applications of bandit/RL algorithms to education, health, other areas
  - Interpretable & Interactive bandits/RL

Overview



- **1. Crowdsourcing & Dynamically Testing Student Explanations**
- 2. Instructor-Centered Experimentation
- 3. Discovering how to personalize
- Future

AXIS: Adaptive eXplanation Improvement System

```
x = matrix(rnorm(m*n),m,n)
What is the standard error?

Answer:
```

Explanation
A z-score is defined as the number of standard deviations a specific point is away from the mean.

CHI 2016

Explain why this answer is correct.

A B +⋯ N **Continually add conditions**

Analyze & Dynamically Adapt

A	B
X%	100-X%

AXIS: Generating Explanations at Scale with Learnersourcing and Machine Learning

Joseph Jay Williams[1] **Juho Kim**[2] **Anna Rafferty**[3] **Samuel Maldonado**[4]
Krzysztof Z. Gajos[1] **Walter S. Lasecki**[5] **Neil Heffernan**[4]

[1]Harvard University
Cambridge, MA
joseph_jay_williams@harvard.edu, kgajos@eecs.harvard.edu

[2]Stanford University & KAIST
Stanford, CA
juhokim@cs.kaist.ac.kr

[3]Carleton College
Northfield, MN
arafferty@carleton.edu

[4]WPI
Worcester, PA
{sjmaldonado,nth}@wpi.edu

[5]Computer Science & Engineering
University of Michigan, Ann Arbor
wlasecki@umich.edu

ABSTRACT
While explanations may help people learn by providing information about why an answer is correct, many problems on online platforms lack high-quality explanations. This paper

lems [18, 21]. For example, students learning math frequently memorize how to apply rote procedures to solve problems [8]. With only superficial changes to how problems are described (e.g., which side of the equation the variable x appears on)

Renkl, 1997

Learners Rate & Generate Explanations

Previous | | Next

Linda is training for a marathon, which is a race that is 26 miles long.

Her average training time for the 26 miles is 208 minutes, but the day of the marathon she was x minutes faster than her average time.

What was Linda's running speed for the marathon in miles per minute?

26/(208 - x) ✓

Explanation
Linda's speed is the distance she ran divided by the time it took. The distance Linda ran was 26 miles. The time it took her was 208 – x. Linda's speed was 26/(208 - x)

How helpful was the above information for your learning?

Completely Unhelpful										Perfectly Helpful
0	1	2	3	4	5	6	7	8	9	10

To help you learn, explain in your own words why the answer is correct.

Explanation
C

Explanation
A

Explanation
B

Explanation
C

Dynamic Experimentation: Exploration vs Exploitation

- Multi-Armed Bandit (Reinforcement Learning)
- Randomized Probability Matching (Thompson Sampling)

Action a $\in A$

Reward R

Policy π

Parameters θ

Explanation
The probability is 3/7 * 5/8, because the number of cookies is changing.

Rating
How helpful was the above information for your learning?
0 1 2 3 4 5 6 7 8 9 10

Exp1	Exp 2	Exp 3
15%	65%	20%

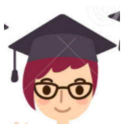
$\epsilon_i \sim Beta(\alpha_i, \beta_i)$ (Probability of Explanation *i* being Rated Helpful)

$R \sim Bin(10, \epsilon_i)$ (0 to 10 Rating by Student)

$$P(\theta|\mathrm{D}) \propto \prod P(r|a,\theta)P(\theta)$$

AXIS Deployment

- AXIS deployed with n=150



1
100

1	2
50	**50**

1	2
20	**80**

1	2	3
10	**60**	**30**

AXIS Policy: Probability distribution over explanations

1	2	3	4	5	6	7	8	9	10	11
18	**13**	**4**	**5**	**8**	**18**	**22**	**6**	**3**	**1**	**2**

Evaluation of AXIS explanations

- Do AXIS explanations help learning?

Problem

```
x = matrix(rnorm(m*n),m,n)
What is the standard error?
```

Answer: [ ]

Problem

```
x = matrix(rnorm(m*n),m,n)
What is the standard error?
```

Answer: [ ]

AXIS Explanation

Problem

```
x = matrix(rnorm(m*n),m,n)
What is the standard error?
```

Answer: [ ]

Filtered Explanation

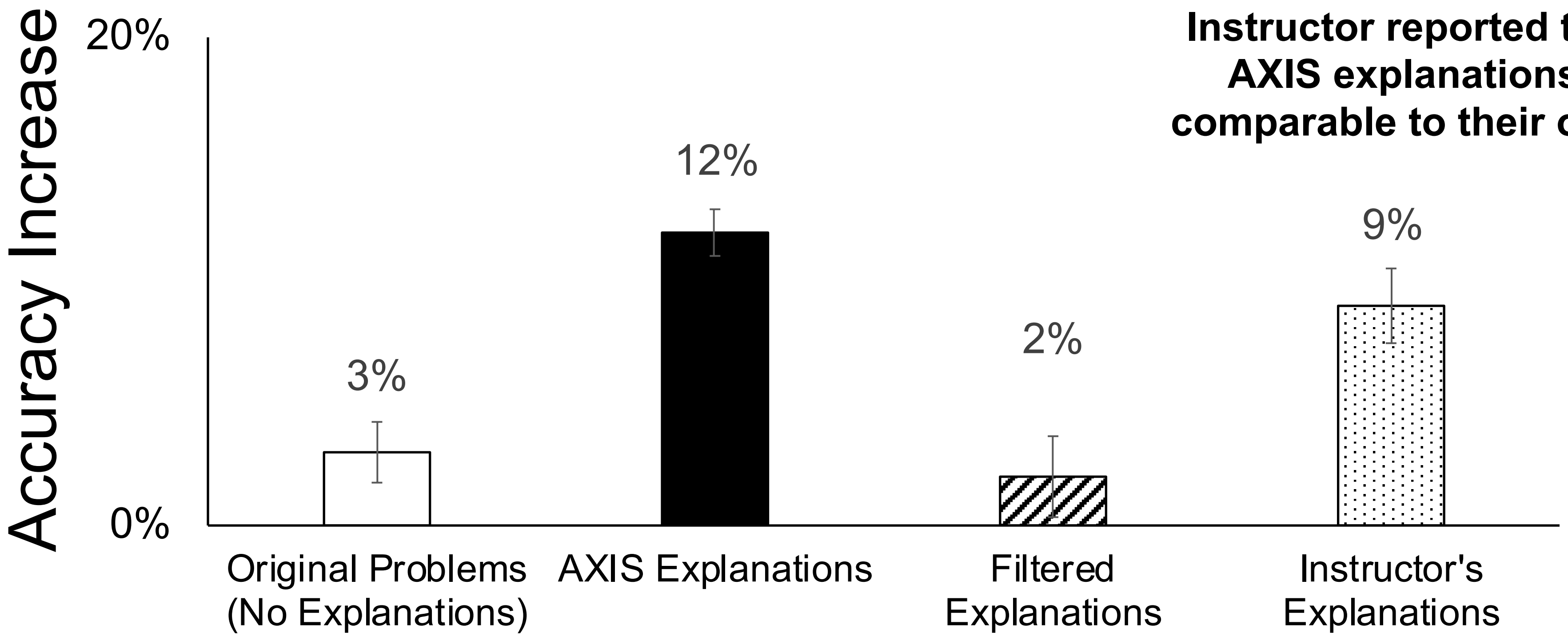
Problem

```
x = matrix(rnorm(m*n),m,n)
What is the standard error?
```

Answer: [ ]

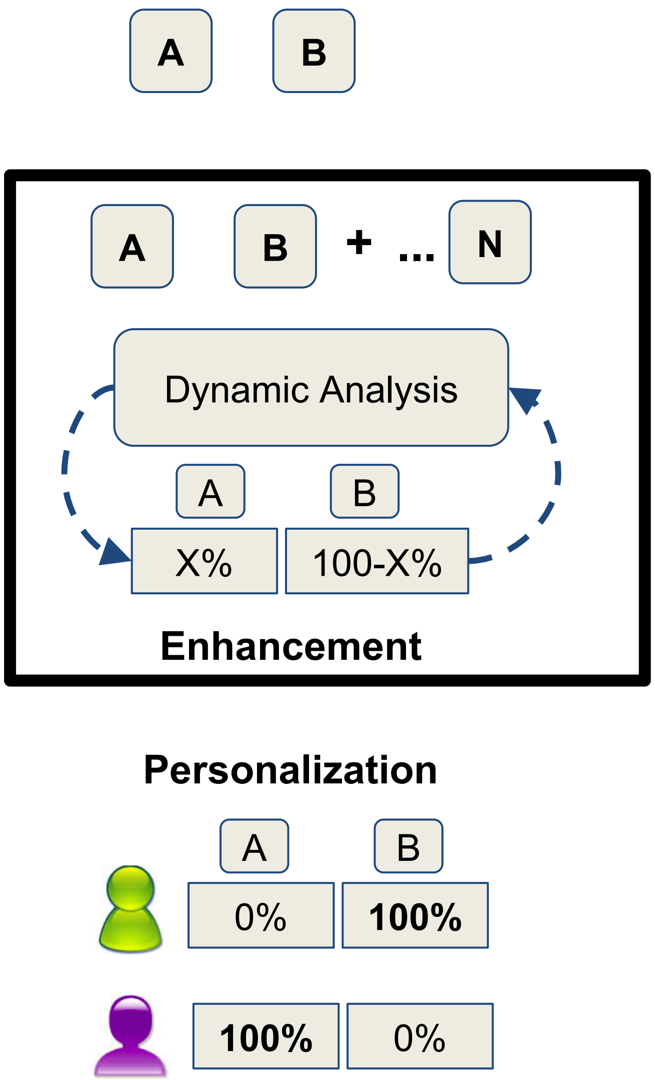
Instructor Explanation

Impact of AXIS Explanations on Learning



Instructor reported the AXIS explanations comparable to their own

Overview



- 1. Crowdsourcing & Dynamically Testing Student Explanations
- **2. Instructor-Centered Experimentation**
- 3. Discovering how to personalize
- Future

2. Instructor-Centered Experimentation

Version	Probability of Explanation	Mean Student Rating	Number of Students
1. Quantitative Explanation	0.23	7.26	46.00
2. Analogical Explanation	0.77	7.48	56.00

CHI 2018

Enhancing Online Problems Through Instructor-Centered Tools for Randomized Experiments

Joseph Jay Williams
National University of Singapore
williams@comp.nus.edu.sg

Anna N. Rafferty
Carleton College
Northfield, MN, USA
arafferty@carleton.edu

Dustin Tingley
Harvard University
Cambridge, MA, USA
dtingley@gov.harvard.edu

Andrew Ang
Harvard University
Cambridge, MA, USA
andrew_ang@harvard.edu

Walter S. Lasecki
University of Michigan
Ann Arbor, MI, USA
wlasecki@umich.edu

Juho Kim
KAIST
Daejeon, South Korea
juhokim@cs.kaist.ac.kr

ABSTRACT

Digital educational resources could enable the use of randomized experiments to answer pedagogical questions that instructors care about, taking academic research out of the laboratory and into the classroom. We take an instructor-centered approach to designing tools for experimentation that lower the barriers for instructors to conduct experiments. We explore this approach through DynamicProblem, a proof-of-concept system for experimentation on components of digital problems, which provides interfaces for authoring of experiments on explanations, hints, feedback messages, and learning tips. To rapidly turn data from experiments into practical improvements, the system uses an interpretable machine learning algorithm to analyze students' ratings of which conditions are helpful, and present conditions to future students in proportion to the evidence they are higher rated. We evaluated the system by collaboratively deploying experiments in the courses of three mathematics instructors. They reported benefits in reflecting on their pedagogy, and having a new method for improving online problems for future students.

explanations, hints, and pedagogical tips into online problems [10], just as occurs in face-to-face tutoring. However, when instructors give face-to-face explanations, they naturally vary the explanations that they give, enabling them to compare which explanations students find helpful and which leave them confused based on students' body language and verbal responses. This adaptation is rare with online problems: instructors typically create one feedback message for a problem, and that message is provided to every student who interacts with the problem. This eliminates the opportunity to iteratively improve the message over time, and the opportunity for the instructor to reflect on what types of messages are most helpful for students.

One way to address this challenge is through experimentation in online problems: randomized experiments can be used to compare alternative versions of explanations, and data can be collected about how helpful students find each explanation. This affordance of online educational resources is increasingly being used by researchers in psychology and education to

Researchers Analyzing Data from Dynamic Experiments

- What are the consequences of dynamic experiments for drawing statistical inferences?
  (see URL tiny.cc/jbanditspower, AIED 2018, Journal of Ed. Data Mining, under review)

Statistical consequences of using multi-armed bandits to conduct adaptive educational experiments

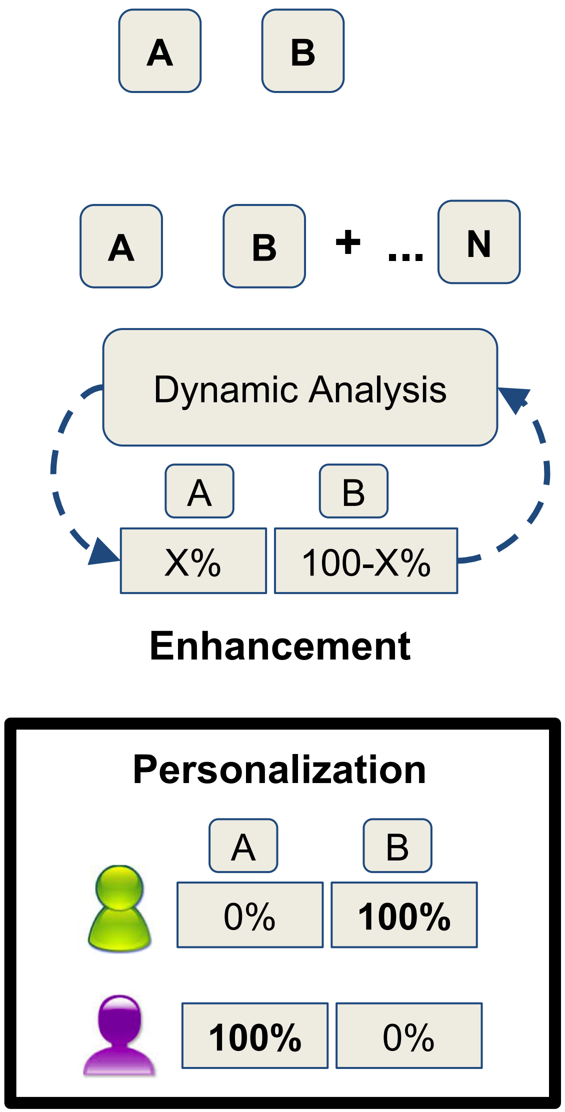
Anna N. Rafferty
Department of Computer Science
Carleton College
arafferty@carleton.edu

Huiji Ying
Department of Computer Science
Carleton College
yingh@carleton.edu

Joseph Jay Williams
Department of Computer Science
University of Toronto
williams@cs.toronto.edu

Randomized experiments can provide key insights for improving educational technologies, but many students may experience conditions associated with inferior learning outcomes in these experiments. Multi-armed bandit (MAB) algorithms can address this issue by accumulating evidence from the experiment as it is running and modifying the experimental design to assign more helpful conditions to a greater proportion of future students. Using simulations, we explore the statistical impact of using MAB for

Overview



- 1. Crowdsourcing & Dynamically Testing Student Explanations
- 2. Instructor-Centered Experimentation
- **3. Discovering how to personalize**
- Future

Discover how to personalize emails

Question about course participation

HX Research Team <noreply@qemailserver.com>

to me

Dear Sam,

Brief

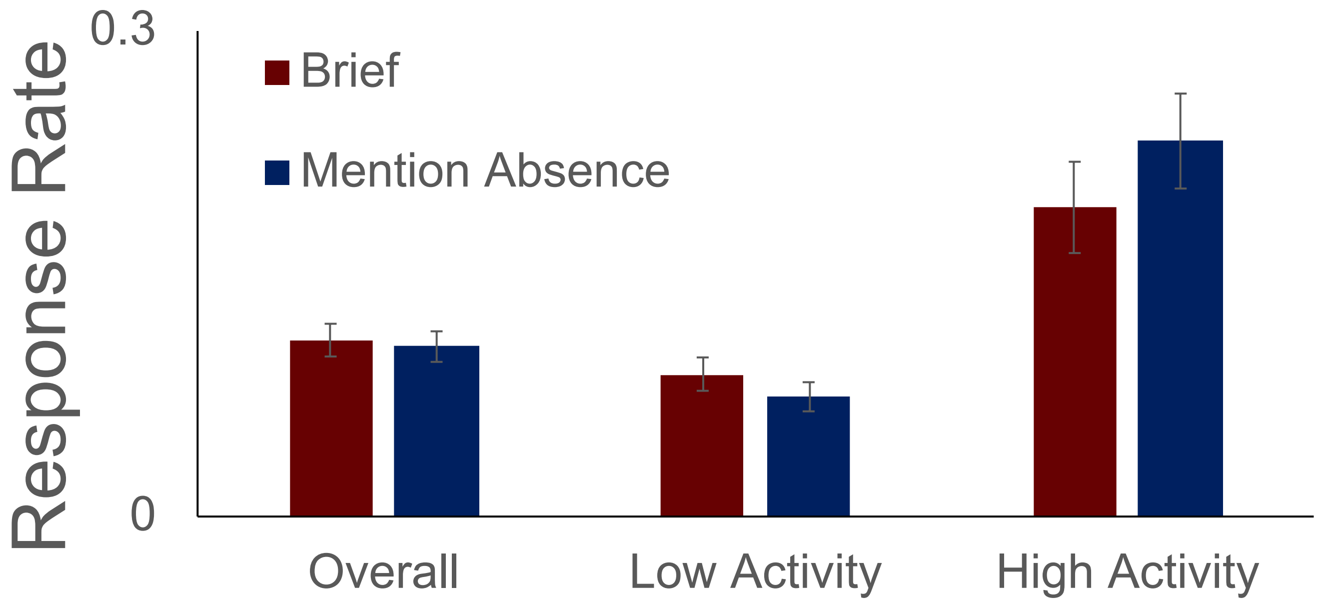
Would you please take this short survey, so we can improve the course for future students?

Mention Absence

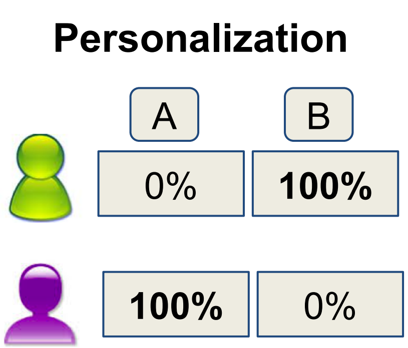
It has been a while since you logged into the course, so we are eager to learn about your experience. Would you please take this short survey, so we can improve the course for future students?

Click here to take the survey.

Optimization through Personalization



- 14.5% more responses



Contextual Bandits (Li et al, 2010); SMART; JITAI

Example Problem: A/B experiment to enhance/personalize email

[Returning to course?]

Dear Sam,

We hope you have enjoyed the opportunity to explore Statistics and R for the Life Sciences. It has been a
we can improve the course for future students?

Each of the links below connects to a short survey. Please click on the link that best describes you.

- I plan on continuing with the course
- I am not continuing the course because it was not what I expected when I signed up.
- I am not continuing the course because the course takes too much time.
- I am not continuing the course because I am not happy with the quality of the course.
- I am not continuing the course because I have learned all that I wanted to learn.
- I am not continuing the course now, but I may at a future time.

Your feedback is very important to us. Thank you for registering for Statistics and R for the Life Sciences.

Follow the link to opt out of future emails:
Click here to unsubscribe

Action Space:
[Subject Line = 1, 2, 3]
[Intro Message = 1, 2, 3…]
[**Response Format** = 1, 2, 3…]

Context:
[**Age**]
[**Number Days Active**]
[Country]
[Activities in week 1]
…
…
..

Contextual bandits & personalizing to subgroups

Balancing Student Success and Inferring Personalized Effects in Dynamic Experiments

Hammad Shaikh
University of Toronto
hammy.shaikh@mail.utoronto.ca

Arghavan Modiri
University of Toronto
modiri.arghavan@gmail.com

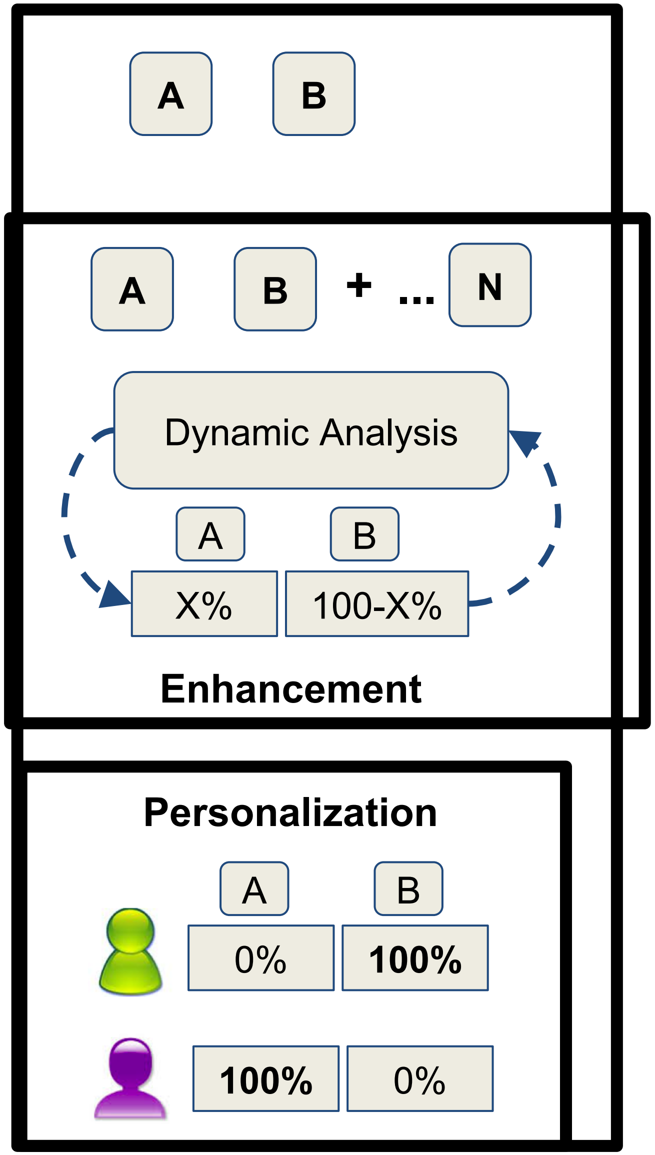
Joseph Jay Williams
University of Toronto
williams@cs.toronto.edu

Anna N. Rafferty
Carleton College
arafferty@carleton.edu

ABSTRACT

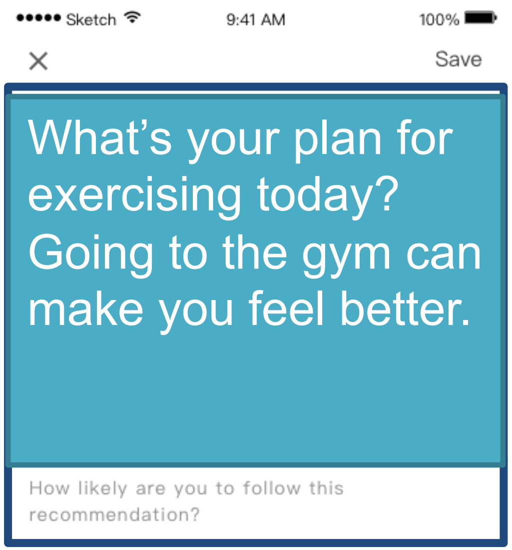
Randomized controlled trials (RCTs) can be embedded in educational technologies to evaluate how interventions affect student outcomes and how effectiveness varies with characteristics like prior knowledge. But RCTs often assign many students to ineffective conditions. Adaptive algorithms like contextual multi-armed bandits (MABs) could change how students are assigned to conditions over time, offering the potential to both evaluate effectiveness for subgroups of students and direct more students to interventions that are effective for them. We use simulations to compare contextual conditions for subgroups of learners. For instance, students with lower prior knowledge benefit more from concrete versus abstract examples, while the opposite holds for students with higher prior knowledge [2]. *Contextual* MABs have the potential to learn how effective conditions are for individuals [7]. These algorithms estimate condition effectiveness as a function of features, such as prior knowledge. We propose using contextual MABs to conduct personalized experiments that assign students to conditions based on their characteristics and simultaneously estimate how these characteristics impact the (differential) effectiveness of the conditions.

Overview



- Vision: Perpetually Improving Systems
- Approach: Collaborative, Dynamic, Personalized Experimentation
- 1. Crowdsourcing & Dynamically Testing Student Explanations
- 2. Instructor-Centered Experimentation
- 3. Discovering how to personalize
- **Future**

Goodlife collab. example: Text msgs to go to gym



Examples of Potential Research Directions

- **Human Computation/Interactive ML/ Human in the loop**: When humans should: Add new actions, provide prior knowledge, add/remove data points, modify reward function, modify exploration-exploitation tradeoff.
- **Interpretable ML**: Understanding moment-by-moment policy, which contextual variables are important, when policy is more/less reliable.
- Exploration-Exploitation tradeoffs & Personalization
- Get Scientists to use dynamic A/B Experiments: Balance regret minimization against hypothesis testing
- Multiple rewards, uncertainty about their relative importance and reliability

Ongoing & Future Activities

- Interested in deploying bandits/RL to real-world? (e.g. good applications to add to a paper)
- Students can listen in on Weekly Research Group: URL tiny.cc/joiniaiextended
- Grad course on Dynamic A/B experiments (www.josephjaywilliams.com/gradcourse)
- Research Programmer Sam Maldonado (AdapComp/MOOClet web service for Active Learning → A/B testing)
- ONR Grant: Personalizing Explanations in Online Problems Using Multi-Armed Contextual Bandits

Thank You!

- Juho Kim, Krzysztof Gajos, Anna Rafferty
- Harvard VPAL (Vice Provost for Advances in Learning) Research
- Tania Lombrozo & Tom Griffiths
- Candace Thille & John Mitchell
- Jascha Sohl-Dickstein, PERTS, Khan Academy
- Sam Maldonado
- Lytics Lab

Ongoing & Future Activities

- Interested in deploying bandits/RL to real-world? (e.g. good applications to add to a paper)
- Students can listen in on Weekly Research Group: URL tiny.cc/joiniaiextended
- Grad course on Dynamic A/B experiments (www.josephjaywilliams.com/gradcourse)
- Research Programmer Sam Maldonado (AdapComp/MOOClet web service for Active Learning → A/B testing)
- ONR Grant: Personalizing Explanations in Online Problems Using Multi-Armed Contextual Bandits

Examples of Problems

[Returning to course?]

Dear Sam,

We hope you have enjoyed the opportunity to explore Statistics and R for the Life Sciences. It has been a
we can improve the course for future students?

Each of the links below connects to a short survey. Please click on the link that best describes you.

- I plan on continuing with the course
- I am not continuing the course because it was not what I expected when I signed up.
- I am not continuing the course because the course takes too much time.
- I am not continuing the course because I am not happy with the quality of the course.
- I am not continuing the course because I have learned all that I wanted to learn.
- I am not continuing the course now, but I may at a future time.

Your feedback is very important to us. Thank you for registering for Statistics and R for the Life Sciences.

Follow the link to opt out of future emails:
Click here to unsubscribe

Action Space:
[Subject Line = 1, 2, 3]
[Intro Message = 1, 2, 3…]
[**Response Format** = 1, 2, 3…]

Context:
[**Age**]
[**Number Days Active**]
[Country]
[Activities in week 1]
…
…
..

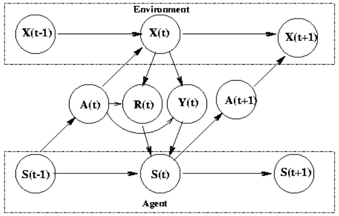
Bridging Designers, Scientists, Machine Learning

Designers

Social-Behavioral Scientists

A B +··· N

Statistics & Machine Learning



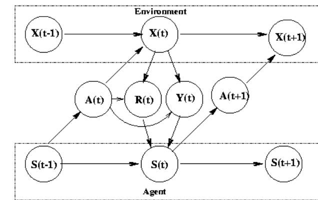
Enhancing and Personalizing Online Resources through Tools for Experimentation (under review)

Interpretable & Interactive Interfaces to ML

Designers

Social-Behavioral Scientists

Statistics & Machine Learning



Interpretable ML

50%	50%
X%	100-X%

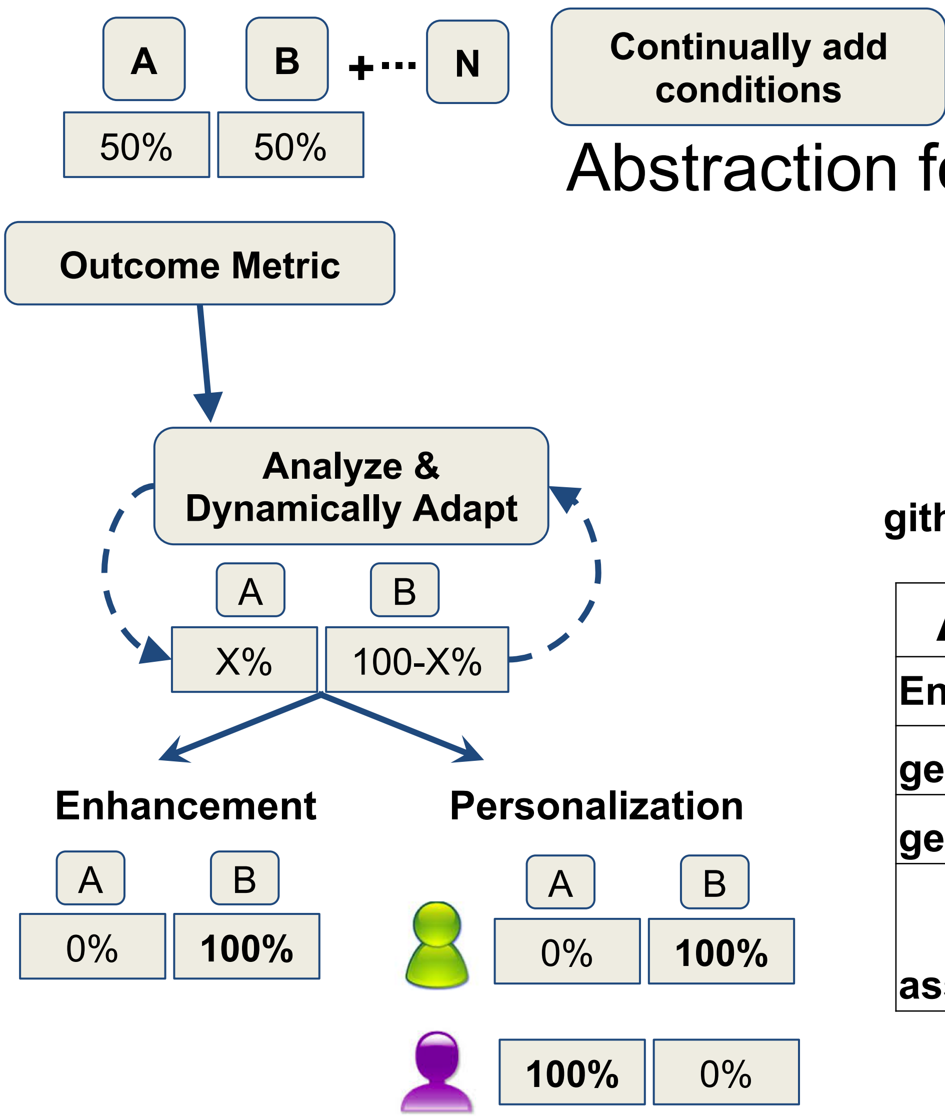
Probability of Version	Mean Student Rating	Number of Students
0.49	4.3	3
0.51	5.2	5

Interactive ML

Designer influences exploration vs exploitation tradeoff?

Encode designer/scientist's prior knowledge via Bayesian models?

Dynamic Experiments as Testbeds for Algorithms



AdapComp/MOOClet

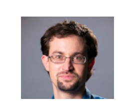
Abstraction for A/B Experimentation via Ac

Python/Django web app

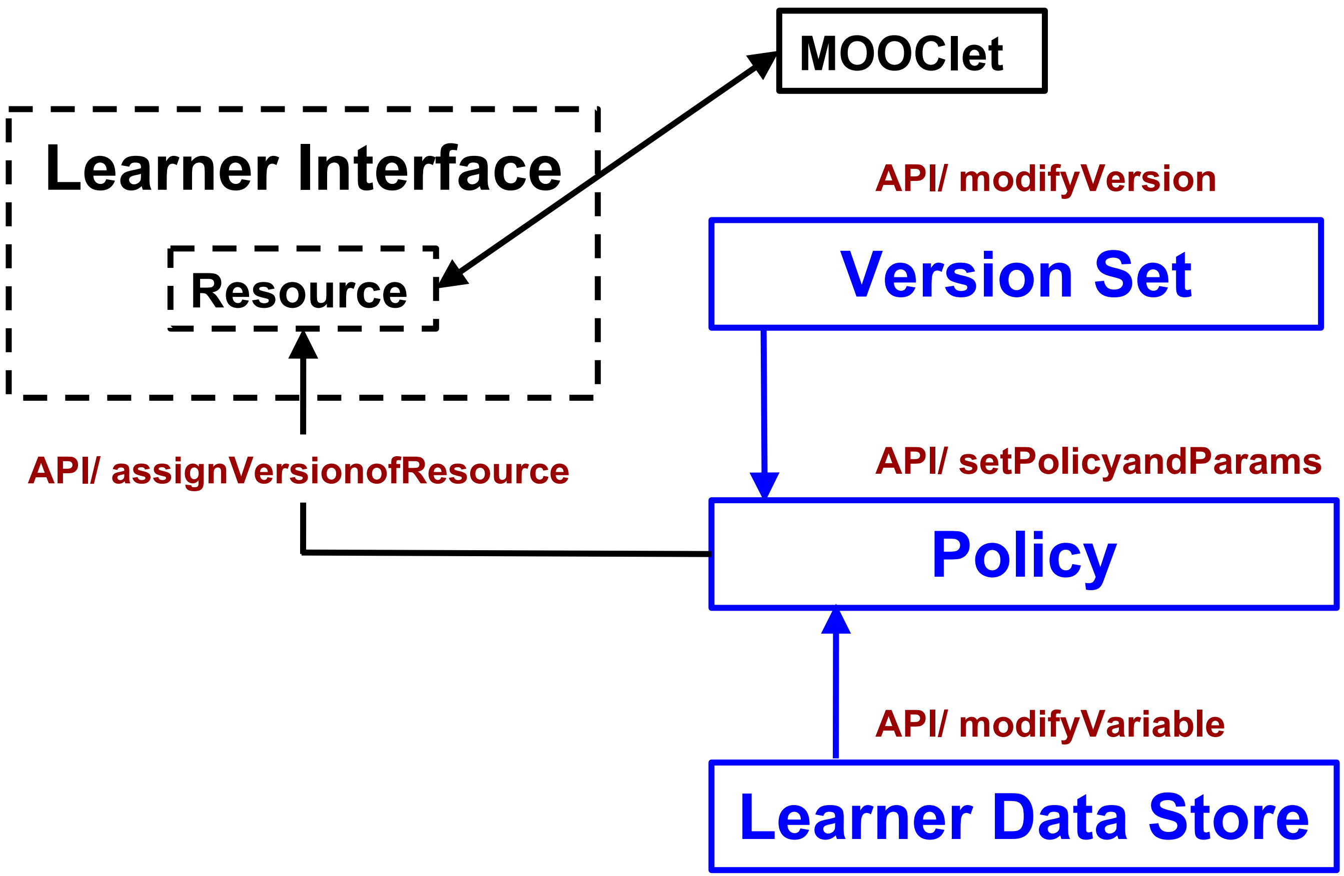
github.com/kunanit/mooclet-engine

AdapComp API Specification	
Endpoint	**Parameters**
getLearnerContext	learner_id
getPastRewards	adapcomp_id
assignLearnerCondition	learner_id, adapcomp_id, condition

ACM LAS 2017, RecSys



MOOClet Engine: Separates Versions, Policy, Data



github.com/kunanit/mooclet-engine

test.mooclet.vpal.io/moocletengine/api/

```
class Mooclet(models.Model):
    name = models.CharField(max_length=100,default='')
    policy = models.ForeignKey('Policy',blank=True,null=True)
    environment = models.ForeignKey(Environment)
    mooclet_id = models.PositiveIntegerField(blank=True)
```

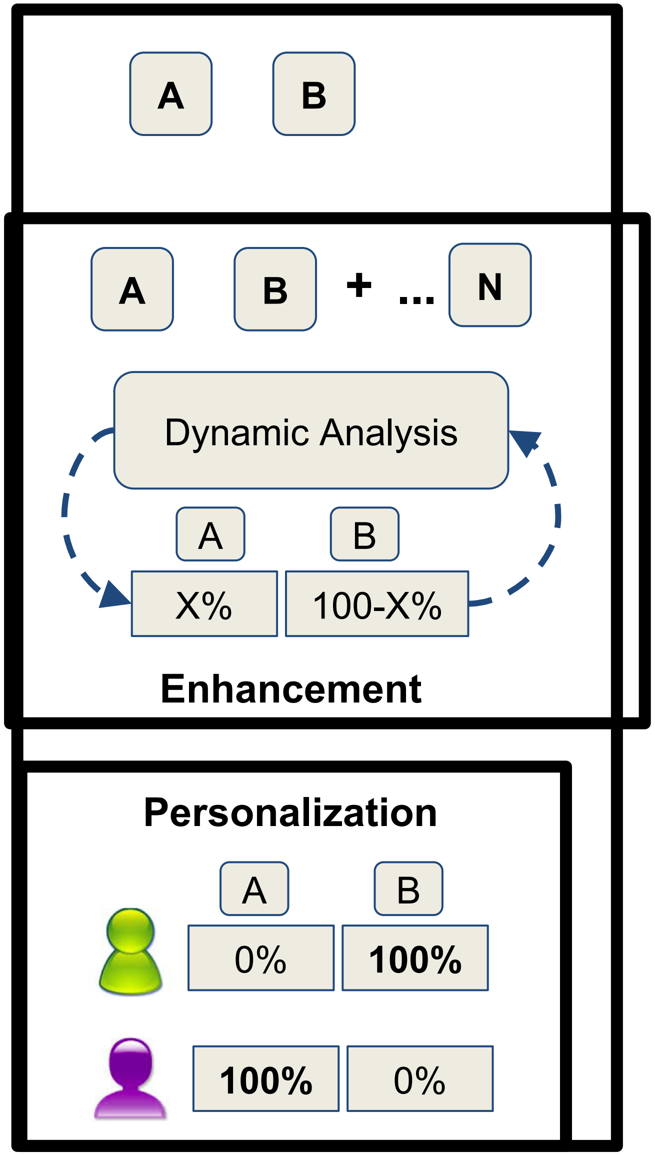
Examples of Potential Research Directions

- **Human Computation/Interactive ML/ Human in the loop**: When humans should: Add new actions, provide prior knowledge, add/remove data points, modify reward function, modify exploration-exploitation tradeoff.
- **Interpretable ML**: Understanding moment-by-moment policy, which contextual variables are important, when policy is more/less reliable.
- Exploration-Exploitation tradeoffs & Personalization
- Get Scientist to use dynamic A/B Experiments: Balance regret minimization against hypothesis testing
- Multiple rewards, uncertainty about their relative importance and reliability

Ongoing & Future Activities

- Intelligent Adaptive Interventions Research Group
- Students can listen in on Weekly Research Group: URL tiny.cc/joiniaiextended
- Grad course on Dynamic A/B experiments (www.josephjaywilliams.com/gradcourse)
- Research Programmer Sam Maldonado (AdapComp/MOOClet web service for Active Learning → A/B testing)
- Office of Naval Research Grant: Personalizing Explanations in Online Problems Using Multi-Armed Contextual Bandits

Review



- Vision: Perpetually Improving Systems
- Approach: Collaborative, Dynamic, Personalized Experimentation
- 1. Crowdsourcing & Dynamically Testing Student Explanations
- 2. Instructor-Centered Experimentation
- 3. Discovering how to personalize
- Future
  - Applications of bandit/RL algorithms to education, health, other areas
  - Interpretable & Interactive bandits/RL

Thank You!

- Juho Kim, Krzysztof Gajos, Anna Rafferty
- Harvard VPAL (Vice Provost for Advances in Learning) Research
- Tania Lombrozo & Tom Griffiths
- Candace Thille & John Mitchell
- Jascha Sohl-Dickstein, PERTS, Khan Academy
- Sam Maldonado
- Lytics Lab

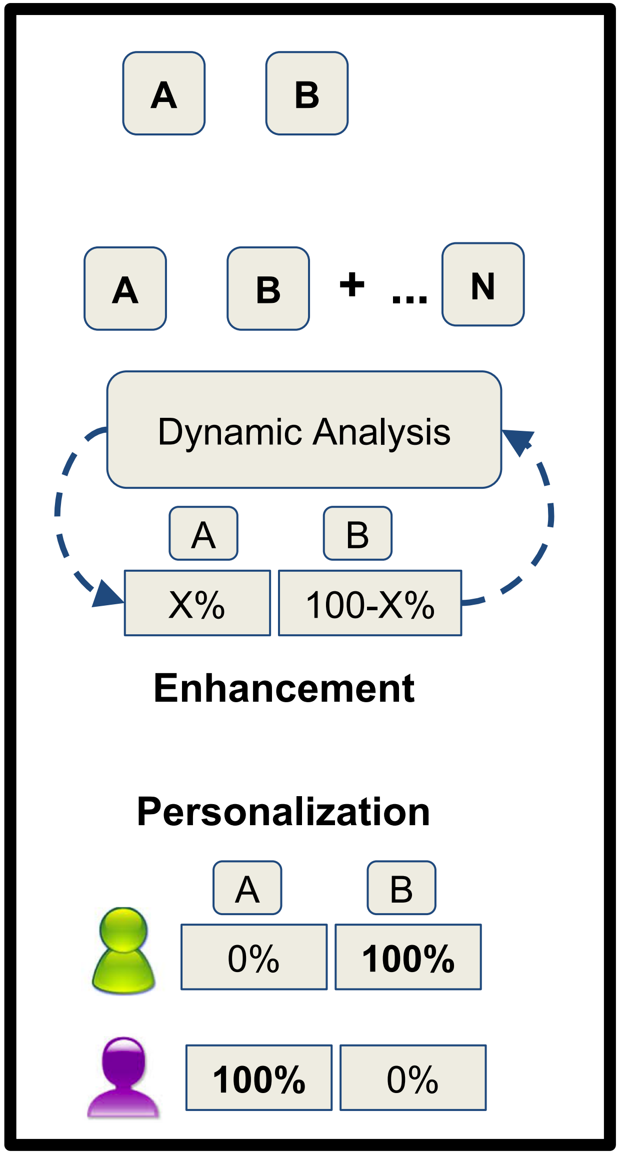
Examples of Potential Research Directions

- **Human Computation/Interactive ML/ Human in the loop**: When humans should: Add new actions, provide prior knowledge, add/remove data points, modify reward function, modify exploration-exploitation tradeoff.
- **Interpretable ML**: Understanding moment-by-moment policy, which contextual variables are important, when policy is more/less reliable.
- Exploration-Exploitation tradeoffs & Personalization
- Get Scientist to use dynamic A/B Experiments: Balance regret minimization against hypothesis testing
- Multiple rewards, uncertainty about their relative importance and reliability

Ongoing & Future Activities

- Intelligent Adaptive Interventions Research Group
- Students can listen in on Weekly Research Group: URL tiny.cc/joiniaiextended
- Grad course on Active Learning for A/B experiments (www.josephjaywilliams.com/gradcourse)
- Econ & HCI Postdoc Weiwen Leung
- Research Programmer Sam Maldonado (AdapComp/MOOClet web service for Active Learning → A/B testing)
- Office of Naval Research Grant: Personalizing Explanations in Online Problems Using Multi-Armed Contextual Bandits

Review



- Vision: Perpetually Improving Systems
- Approach: Collaborative, Dynamic, Personalized Experimentation
- 1. Crowdsourcing & Dynamically Testing Student Explanations
- 2. Instructor-Centered Experimentation
- 3. Discovering how to personalize
- Future