

From

*“Combining Unsupervised and Supervised Classification to Build User Models for Exploratory Learning Environments”**

To FUMA:

Framework for User Modeling and Adaptation

Cristina Conati

Department of Computer Science
University of British Columbia



*** With Saleema Amershi, now Microsoft Research**

Great Progress on AI-driven Support to Problem Solving

[DuBulay, Mitrovic, Yacef; Handbook of AI in Education 2023]

ANDES Physics Workbench - [P11-2-Solution.Rd]

A 2000 kg car is neutral at the top of a 20-degree inclined driveway 20 m long along its parking brake and rolls down. Assume that the driveway is frictionless.

At what speed will it hit the garage door?

Answer:

Name	Definition	Units
T1	car hits garage door	
m	mass of car	kg
v1	magnitude of the instantaneous velocity of car at time T1	m/s
Fnet,x	magnitude of the net force on the car at time T1	N
a	magnitude of the average acceleration of car during T1 to T2	m/s ²
W	magnitude of the weight force on car at time T1 to T2	N

20 degrees

Physics

Geometry

In the contiguous representation, all interaction takes place in the diagram. Clicking on a question mark opens a nearby work field in which feedback is given. Correct answers are displayed in the diagram.

The table is visible, but non-interactive. It displays a passive record of answers.

How is the packet switching model of message transmission like the postal system?

Packets are sent to intermediate destinations before being routed to their final destination.

Computer Science

scenarios

Deep-sea divers Jacques Cousteau and his assistant are exploring shark-infested waters. Jacques Cousteau is 10 feet below the surface (10 feet) and his assistant is 110 feet below the surface (110 feet). Jacques Cousteau is swimming down at the rate of eight feet per minute, and his assistant is swimming towards the surface at six feet per minute.

1 Where does Jacques Cousteau begin his descent? Where does his assistant begin his ascent?
2 What is the depth, in feet, of each diver after four minutes?
3 In how many minutes will the assistant be 74 feet below the surface?
4 According to the algebraic model, at what time will the divers be at the same depth?

To write an expression, define a variable for the time from now and use this to the depth of Jacques Cousteau and the depth of his assistant.

Quantity Name	TIME	DEPTH OF JACQUES	DEPTH OF ASSISTANT
UNIT	MINUTES	FEET	FEET
Expression 1	0	-10	-110
Question 1		-10	-110
Question 2			
Question 3			
Question 4			

Graph

DEPTH

DEPTH OF JACQUES

Algebra

SQL-Tutor

Problem 3: Retrieve the name and address of all employees who work for the Research department.

Select: []
From: []
Where: []
Group by: []
Having: []
Order by: []

Feedback Level: Error Flag Submit Answer Reset

Schema for the COMPANY Database

The general description of the database is available here. Clicking on the name of a table brings up the table details. Primary keys in the attribute list are underlined, foreign keys are in italics.

Table name Attribute list

DEPARTMENT DNAME EMPLOYEE MGR MGRSTARTDATE
EMPLOYEE EID EMPNAME EMPJOB EMPJOB ADDRESS SEX SALARY SUPERVISOR DNO
DEPT_LOCATIONS DNAME DEPTJOB LOCATION
PROJECT EMPNO EMPJOB LOCATION DNO
WORKS_ON EID EMPJOB
DEPENDENT EID DEPENDENT_NAME SEX BDATE RELATIONSHIP

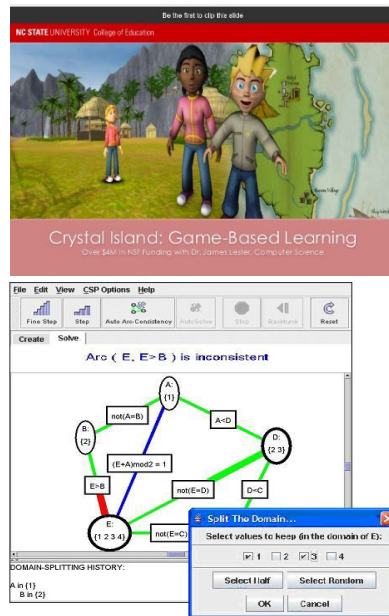
SQL

- Well defined problem solutions => guidance on problem solving steps
- Clear definition of correctness => basis for feedback

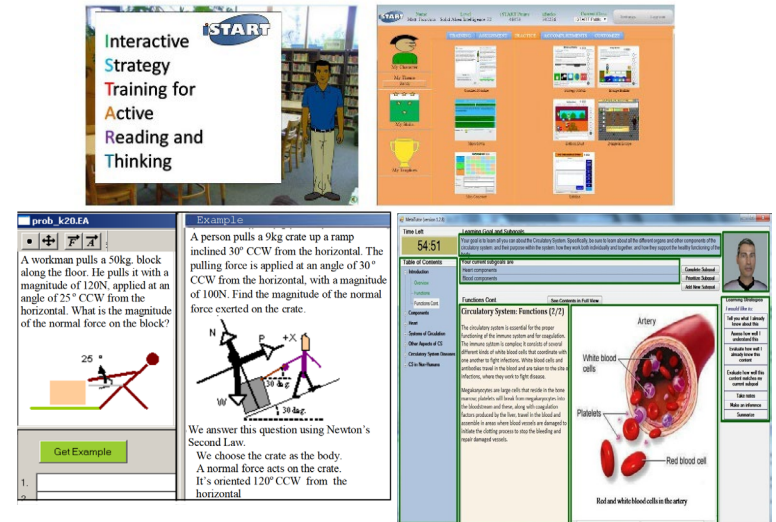
Beyond Problem Solving

- Exploratory Learning Environments (ELEs) that support active learning via student-driven exploration

Educational Games and Simulations



Meta-Cognitive Tutors



An Example

Tools for Learning Artificial Intelligence

A screenshot of the AIspace website interface. The header includes the "AIspace" logo and "Main Tools". The main content area is titled "Main Tools" and lists several AI-related tools with icons and brief descriptions. The tools listed are: Graph Searching, Consistency Based CSP Solver, Stochastic Local Search Based CSP Solver, Definite Clause Deduction, Belief and Decision Networks, Decision Trees, Neural Networks, and STRIPS To CSP conversion. Each tool entry includes a link to "Help" and "Bugs & Enhancements".

AIspace Main Tools

Graph Searching
Search is an important part of AI; many problems can be
[Help] [Bugs & Enhancements]

Consistency Based CSP Solver
Constraint satisfaction problems (CSPs) are pervasive
[Help] [Bugs & Enhancements]

Stochastic Local Search Based CSP Solver
This tool is designed to help you learn another strategy
[Help] [Bugs & Enhancements]

Definite Clause Deduction
Every representation and reasoning system needs a pro
[Help] [Bugs & Enhancements]

Belief and Decision Networks
Belief networks (also called Bayesian networks or caus
[Help] [Bugs & Enhancements]

Decision Trees
Learning is the ability to improve one's behaviour based
a test data set
[Help] [Bugs & Enhancements]

Neural Networks
Inspired by neurons and their connections in the brain, n
[Help] [Bugs & Enhancements]

STRIPS To CSP conversion
A STRIPS problem is a planning problem that views the
of said problem into a CSP for solving.
[Help] [Bugs & Enhancements]

- AISpace (Amershi et al., 2007)
 - Suite of interactive simulations of common Artificial Intelligence algorithms
 - Used regularly in our AI courses
 - Google "AISpace" if you want to try it out
- CSP (Constraint Satisfaction Problems) Applet
 - visualizes the working of the AC3 algorithm

The ACSP applet

CSP Applet Version 4.6.1.I --- s1.xml

File Edit View CSP Options Help

Fine Step Auto Arc-Consistency Stop Backtrack Reset

Create Solve

Click on a variable to split its domain.
Click on a constraint to reorder its variables.
Click on an arc to make it arc-consistent.

A: {1 2 3 4}

B: {1 2 3 4}

C: {1 2 3 4}

D: {1 2 3 4}

E: {1 2 3 4}

not(A=B)

not(B=D)

not(B=C)

A=D

E<A

E<B

C<D

E<D

E<C

DOMAIN-SPLITTING HISTORY:

AI-driven Support in ELE

- ❑ Not all students learn well from exploratory activities

[e.g., Van Joolingen et al., 2007]

- Important to provide **support** for those students who **need help**.
- While maintaining student **initiative** and **engagement**

- ❑ Challenge: No clear definition of **correct/effective** behaviors



what behaviors should drive **personalized** support?

how to provide such support **effectively** and **unobtrusively**?

FUMA for Data-Driven Personalization

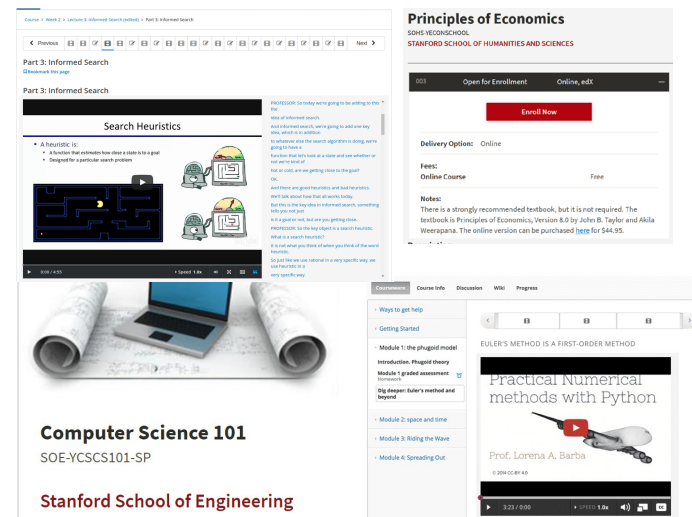
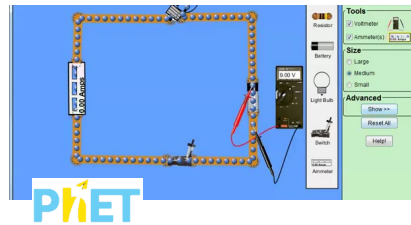
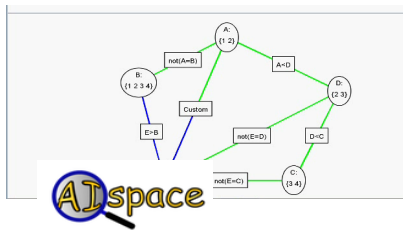
❑ FUMA (Framework for User Modeling and Adaptation)

- First version proposed by **Amershi and Conati 2009 (ToT Award 2022)**
- **Learn** from data what user behaviors should trigger **personalized help**
- **Recognize** and **react** to these behaviors in **real-time** during interaction

❑ Evaluated in several **ELEs**

Four MOOCs [Lallé et al., AIED 2020]

Two Interactive Simulations [Kardan Conati UMAP 2015 Fratamico et al JAIED 2017]



Environment for Game Design

[Lallé et al., LAK 2021, AIED 2023, Yalcin et al TiiS 2022]



Overview

- ❑ Overview of FUMA and initial results with the CSP applet
- ❑ Extension to other data and environments
 - Challenges and lessons learned
- ❑ What's next?

FUMA

[Amershi and Conati 2009, Kardan and Conati 2011, 2015]

Behavior Discovery

Actions Logs
Other Data

Vector of
Interaction
Features

Clustering

Groups together students that have **similar interaction behaviors**

Association
Rules
Mining

Extract rules describing **distinguishing interaction patterns** in each cluster

Interpret in terms of learning

- Performance Measures
- Experts

User Classification

New User's
Actions

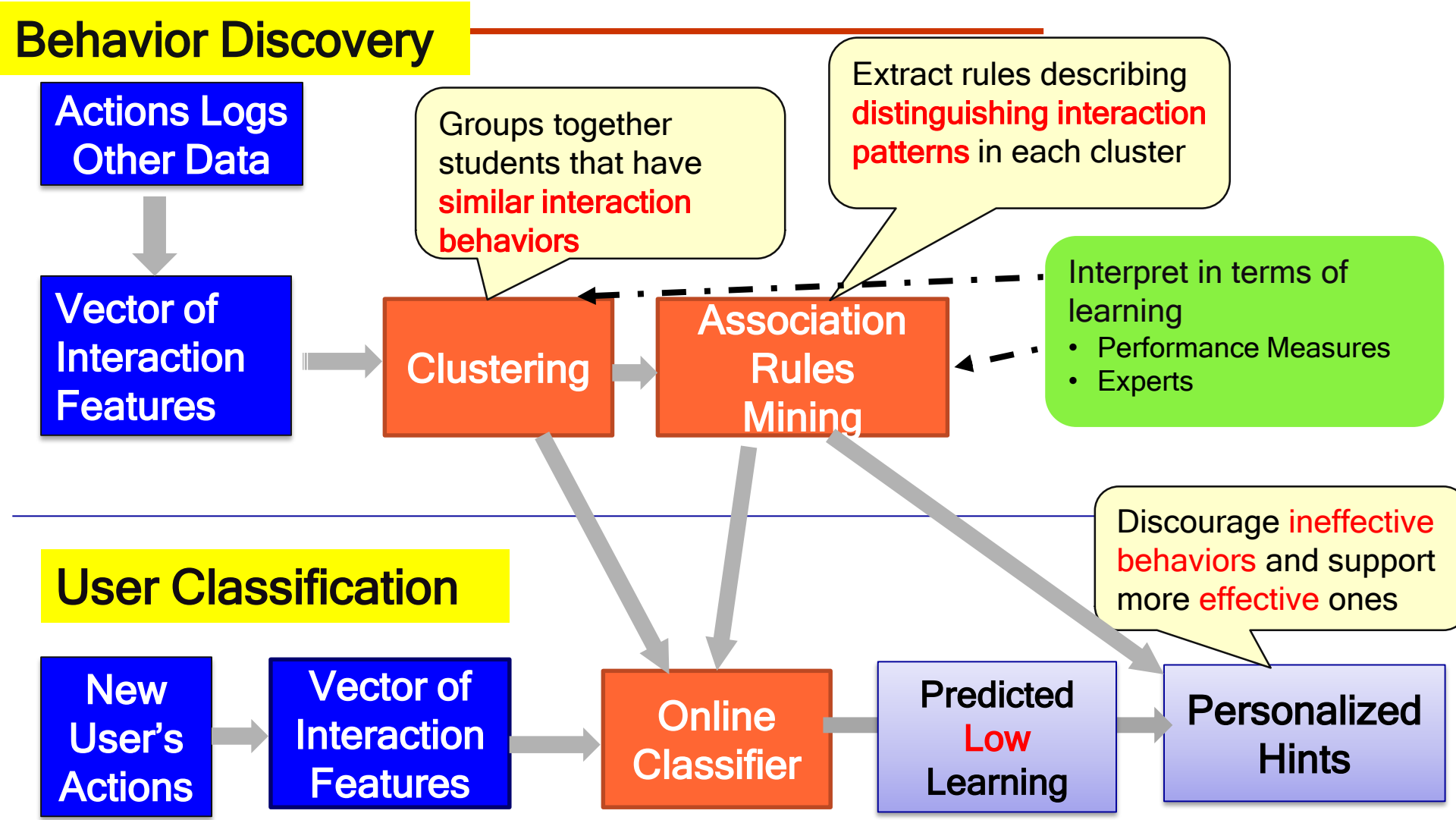
Vector of
Interaction
Features

Online
Classifier

Predicted
Low
Learning

Personalized
Hints

Discourage **ineffective behaviors** and support more **effective** ones



Test Bed - CSP Applet

[Amershi and Conati 2009, Kardan and Conati 2011, 2015]

CSP Applet Version 4.6.0 --- scheduling1.xml

File Edit View CSP Options Help

Fine Step Step Auto Arc-Consistency AutoSolve Stop Step Back Backtrack Reset

Create Solve

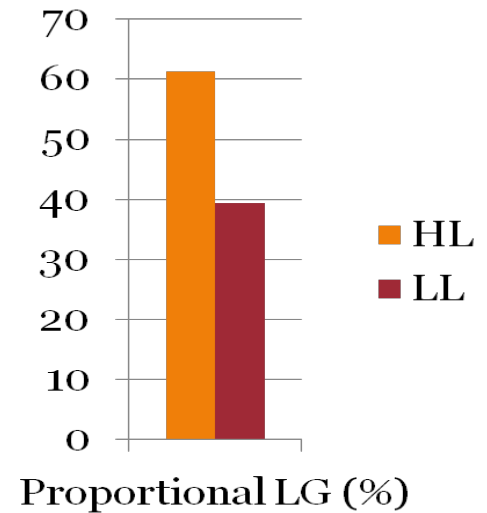
Arc (B, E<B) is inconsistent

DOMAIN-SPLITTING HISTORY:

Behavior Discovery



- Dataset:
 - 64 subjects, 13,000+ actions, 17+ hour
- 7 types of actions → 21 features
 - Action **frequency**
 - **Time** between actions (Mean and SD)
- Found **two clusters** with different learning
 - lower learning (LL) and higher learning (HL)
- Sample Rules



HL members:

- Use **Direct Arc Click** action **frequently** (R1).

LL members:

- Use **Direct Arc Click** **sparsely** (R3)
- Leave **little time** between a **Direct Arc Click** and the next action (R2)

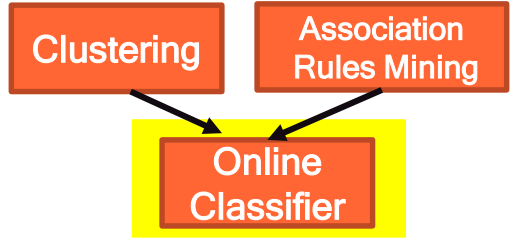
From Behavior Patterns to Hints

Intervention Code	Intervention Description
DAC_fr	Using Direct Arc Click more often
DAC_PA	Spending more time after performing Direct Arc Clicks
Reset_fr	Using Reset less frequently
AAC_fr	Using Auto Arc-consistency less frequently
DS_fr	Using Domain Splitting less frequently (only when appropriate)
FS_PA	Spending more time after performing Fine Steps
BT_fr	Using Back Track less frequently (only when appropriate)
FS_fr	Using Fine Step less frequently
Reset_PA	Spending more time after performing after resetting for planning

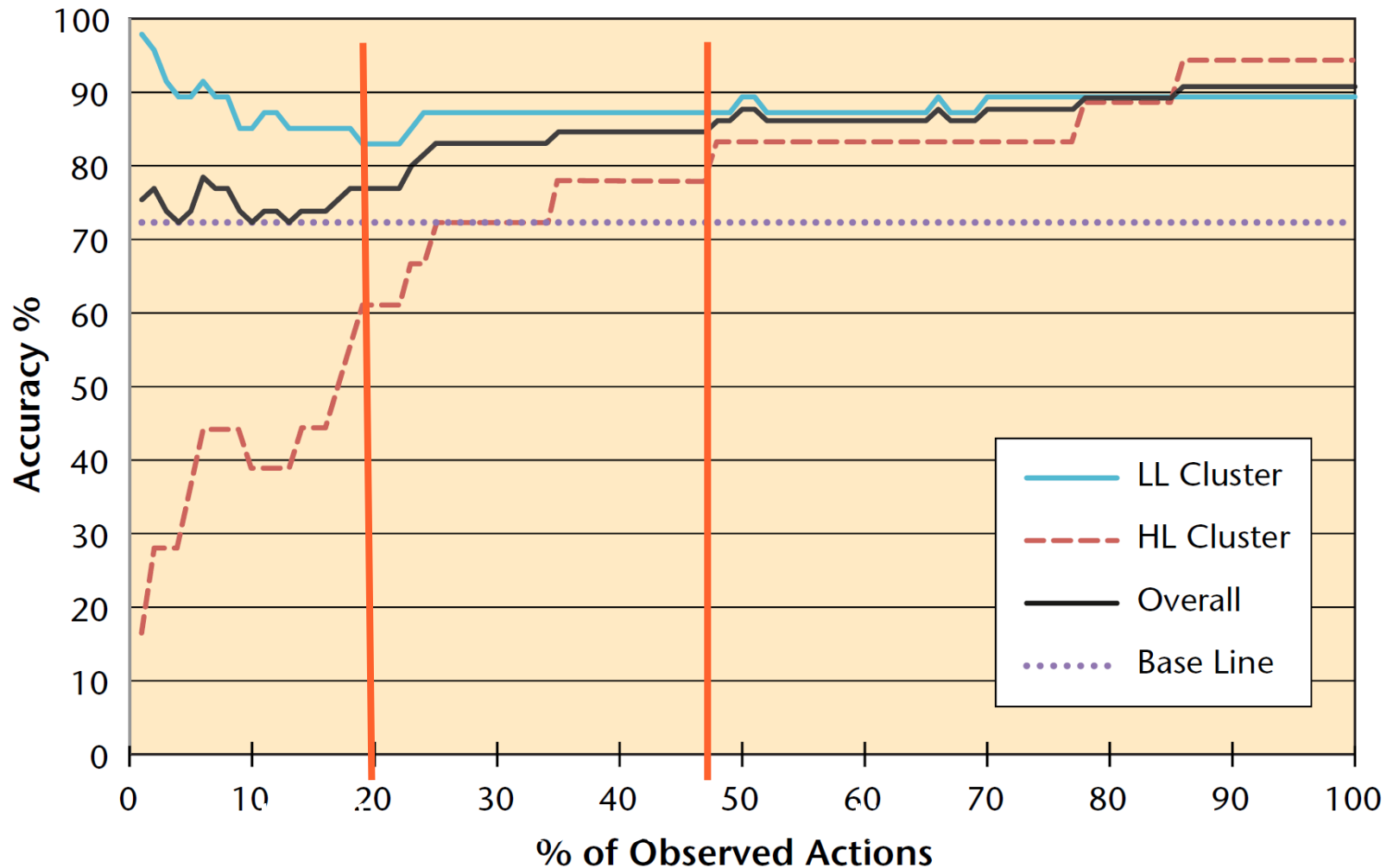
Table 2. Description of hints

Classifier Evaluation on CSP Applet

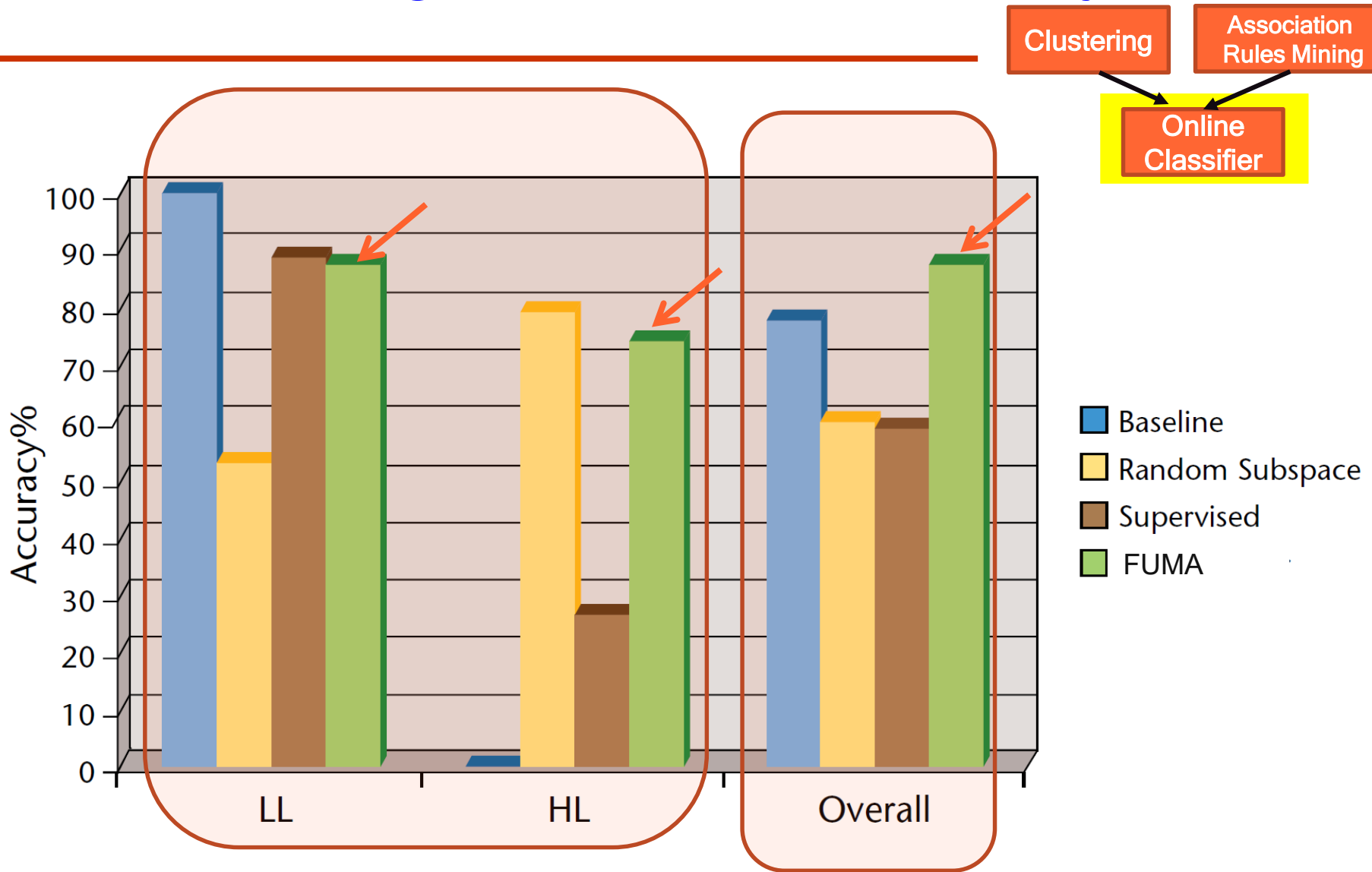
[Kardan and Conati 2012]



Accuracy as a function of observed actions

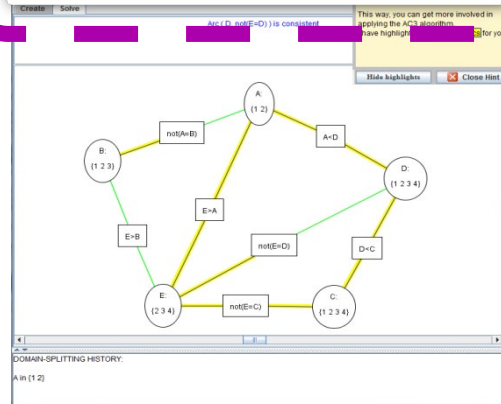
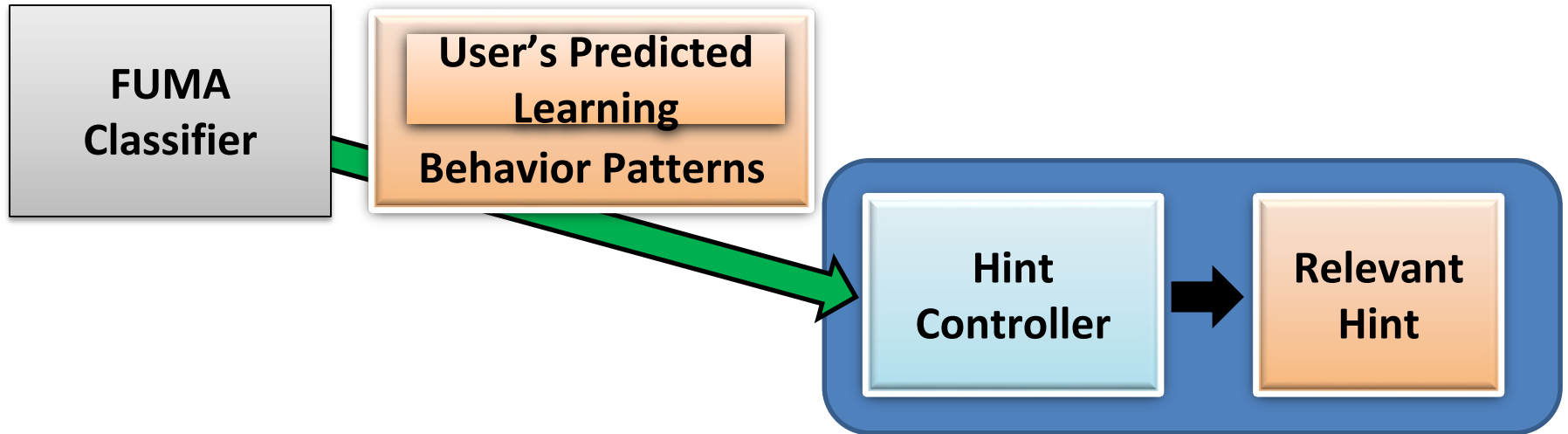


Average Overtime Accuracy



Providing Personalized Support

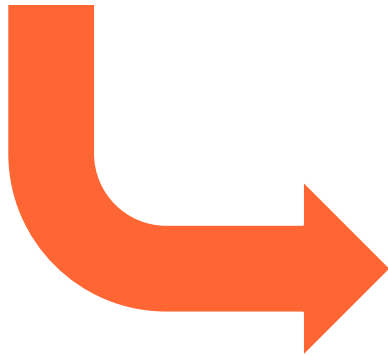
(Kardan and Conati CHI 2015)



Incremental Hints: Level 1

Classifier User Model detects a Low Learner that

- Uses *Direct Arc Click* sparsely (R3)



A screenshot of a software interface. At the top, there is a light gray header bar. Below it, a yellow dialog box with a gray border is open. The dialog box contains the text: "Did you know you can tell AC-3 which arc to make consistent by clicking on that arc?". At the bottom of the dialog box, there are two buttons: "Hide Highlights" and "Close Hint" (with a red 'X' icon). In the background, a portion of the AC-3 interface is visible, showing a blue line and a green line meeting at a node labeled "A<D". The text "because of arc (D, A<D)" is partially visible in blue.

Incremental Hints: level 2

◀ cktrack ↻ Reset

Arc (D, not(E=D)) is consistent

As I suggested earlier, you can choose which arc to make consistent next by clicking on it.

This way, you can get more involved in applying the AC3 algorithm. I have highlighted the **relevant arcs** for you.

Hide highlights ✖ Close Hint

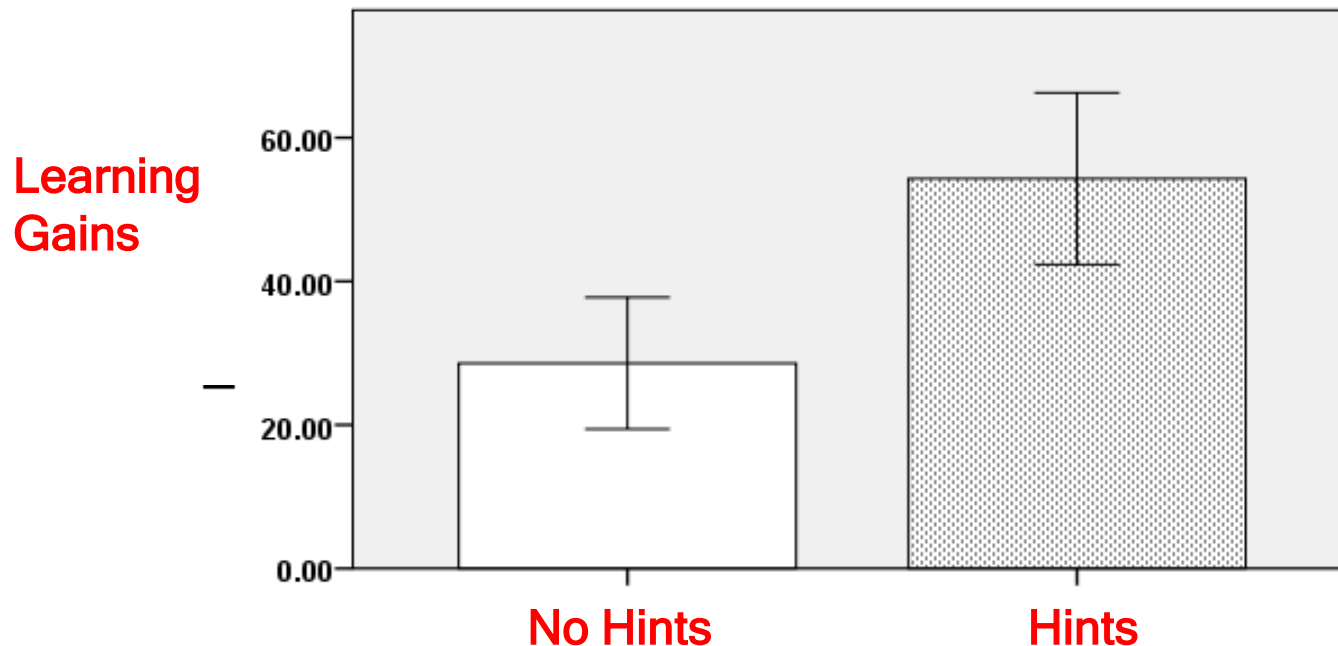
The diagram shows a constraint network with two nodes: A and D. Node A is labeled 'A: {1 2}' and node D is labeled 'D: {1 2 3 4}'. There are four arcs: 'E>A' (green), 'A<D' (yellow), 'not(E=D)' (green), and 'D<C' (yellow). The arcs 'A<D' and 'D<C' are highlighted in yellow, indicating they are relevant arcs. The arc 'not(E=D)' is highlighted in green, indicating it is consistent. The arc 'E>A' is also highlighted in green, indicating it is consistent. The arc 'A<D' is highlighted in yellow, indicating it is not consistent.

How to Deliviet the Hints Effectively?

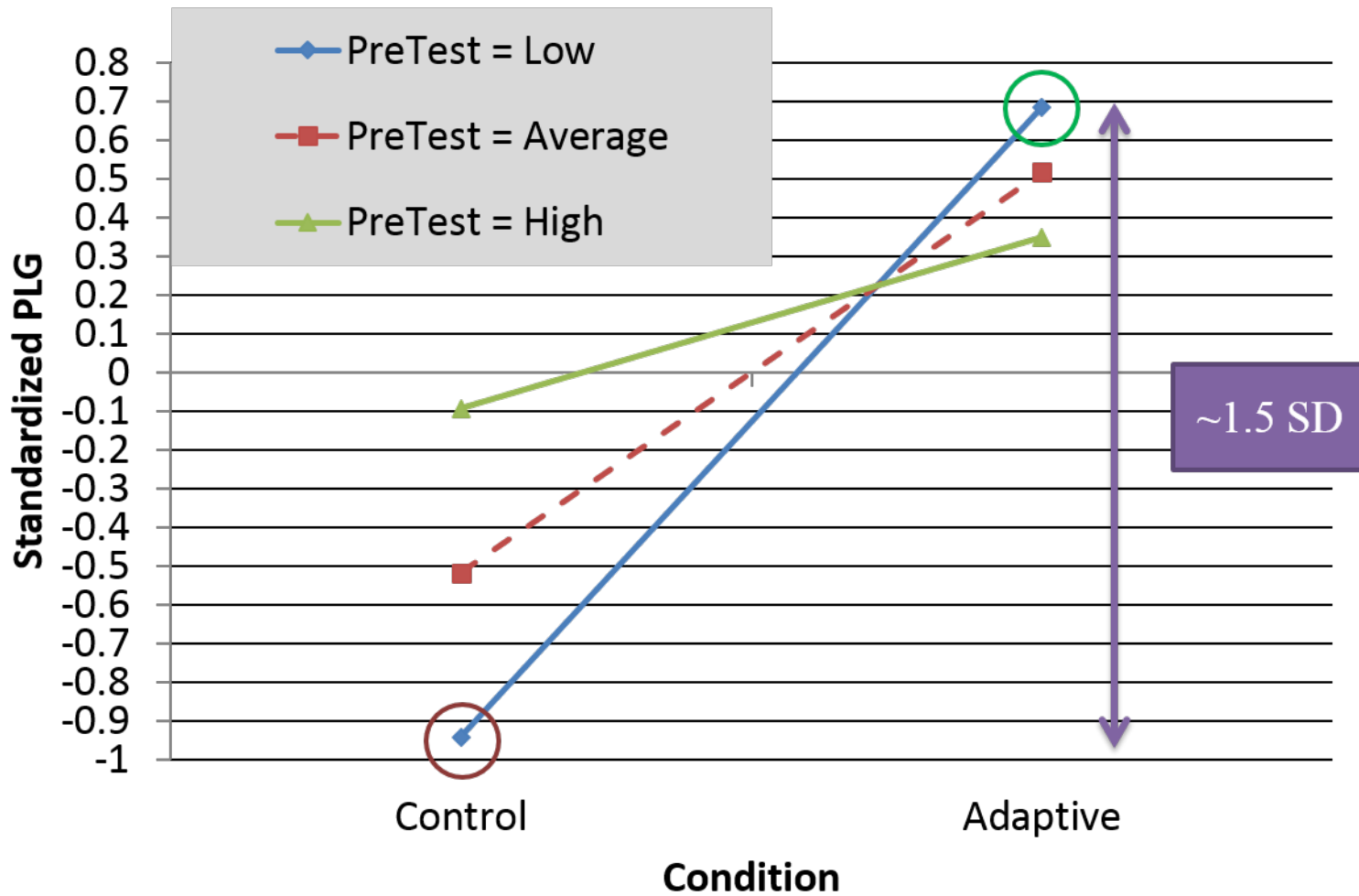
Evaluation

(Kardan and Conati CHI 2015)

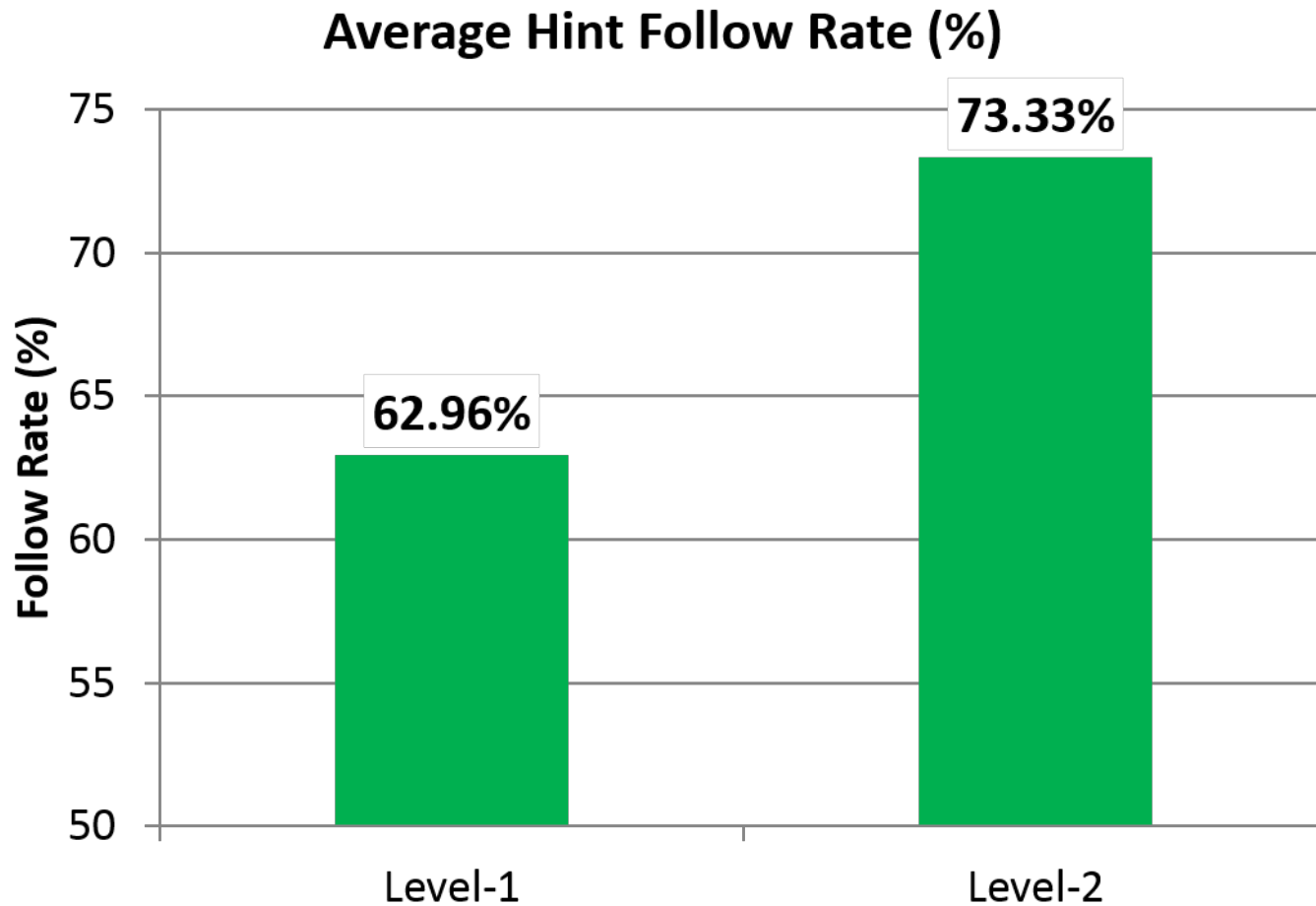
- ❑ User study :
 - Two groups of 18 students worked with the CSP applet
 - One group with **personalized hints**, and one **without**
- ❑ Students in the ACSP group learned more



Learning Gain: PreTest×Condition



Results: Acceptance of Interventions



Overview

- Overview of FUMA and initial results with ACSP applet [Amershi and Conati 2009, Kardan and Conati 2012, 2015]
- Extensions to
 - multimodal data
 - More complex OELEs
- What's next?

Experimenting With Multimodal Data

Actions Logs
Gaze Data

We tried with the CSP applet
[Kardan and Conati UMAP 2013]

Vector of
Interaction
Features

Clustering

Associatio
n Rules
Mining

Behavior Discovery

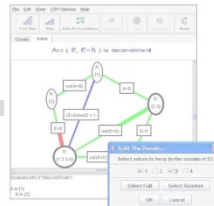
User Classification

New
user's
Actions

Feature
Vector
Calculation

Online
Classifier

Adaptive
Interventions



User Study to Collect Gaze Data

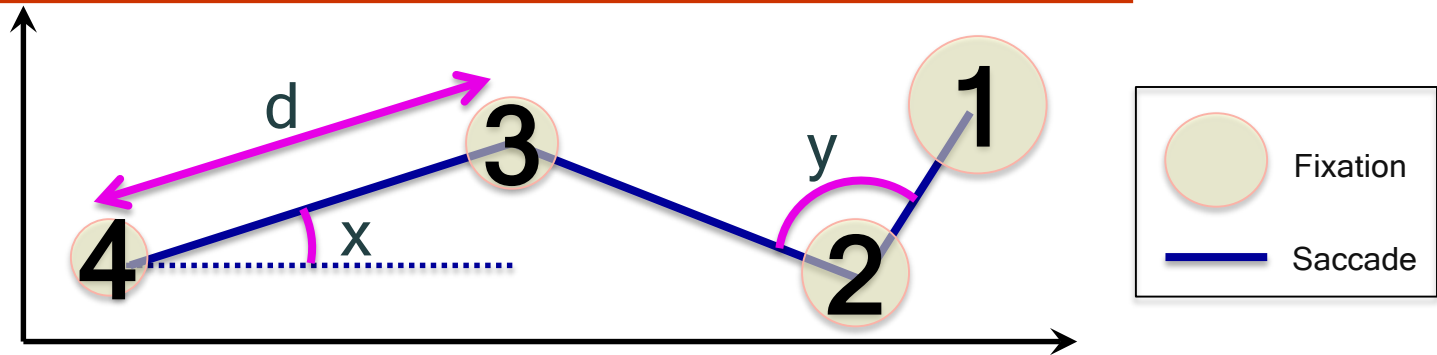
- 45 participants
- Tobii T120 eye tracker to capture user gaze

The screenshot displays the CSP Applet interface. The main window shows a constraint network diagram with three nodes: A, B, and C, each with a domain of {1, 2, 3, 4}. The constraints are A=B, B=C, and not(A=C). A mouse cursor is pointing at the not(A=C) constraint. The interface includes a menu bar (File, Edit, View, CSP Options, Help) and a toolbar with buttons for Fine Step, Auto Arc-Consistency, Stop, Backtrack, and Reset. The bottom status bar shows the text "DOMAIN-SPLITTING HISTORY:".

Start CSP Applet Version 4...



Eye-tracking measures



General Measures

- Number of Fixations
- Fixation rate
- Fixation Duration
- Saccade Length (d)
- Relative Saccade Angles (y)
- Absolute Saccade Angles (x)
-

Measures specific to Areas of Interest (AOI)

Areas Of Interest

CSP Applet Version 4.6.1 --- scheduling2.xml

File Edit View CSP Options Help

Fine Step Step Auto Arc-Consistency AutoSolve Stop Step Back Backtrack Reset

Create Solve

Arc (C, D<C) is inconsistent

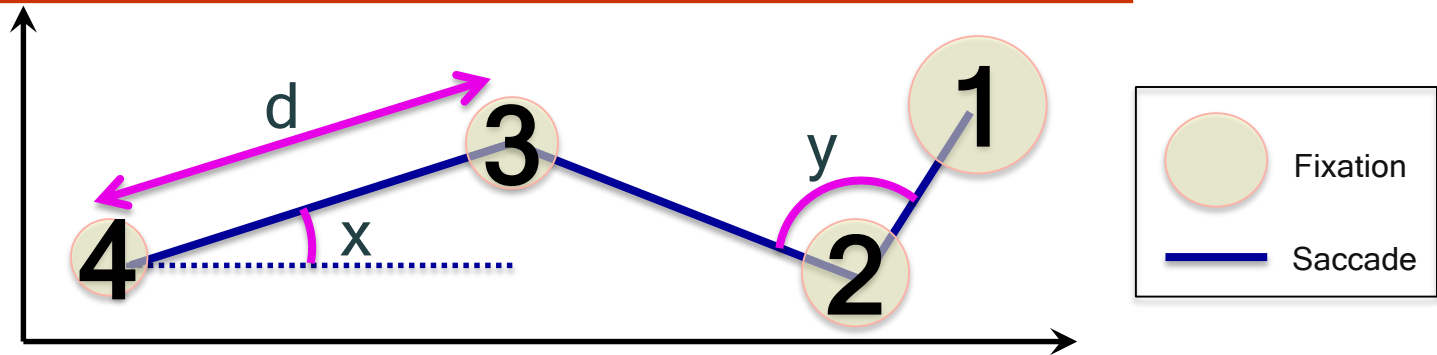
Diagram illustrating a Constraint Satisfaction Problem (CSP) network. The nodes are A, B, C, D, and E, each with a domain of {1 2 3 4}. The constraints are:

- not(A=B)
- A < D
- Custom
- not(E=D)
- D < C
- E > B
- not(E=C)

The constraint D < C is highlighted in red, indicating it is the current focus of the solver.

DOMAIN-SPLITTING HISTORY:

Eye-tracking measures



General Measures

- Number of Fixations
- Fixation rate
- Fixation Duration
- Saccade Length (d)
- Relative Saccade Angles (y)
- Absolute Saccade Angles (x)
-

Measures specific to Areas of Interest (AOI)

- Proportional number of fixations
- Proportional time spent
- Time to first fixation
- Transitions between two AOIs
-

51 features based on summary statistics (e.g. mean, st.dev.) of these measures

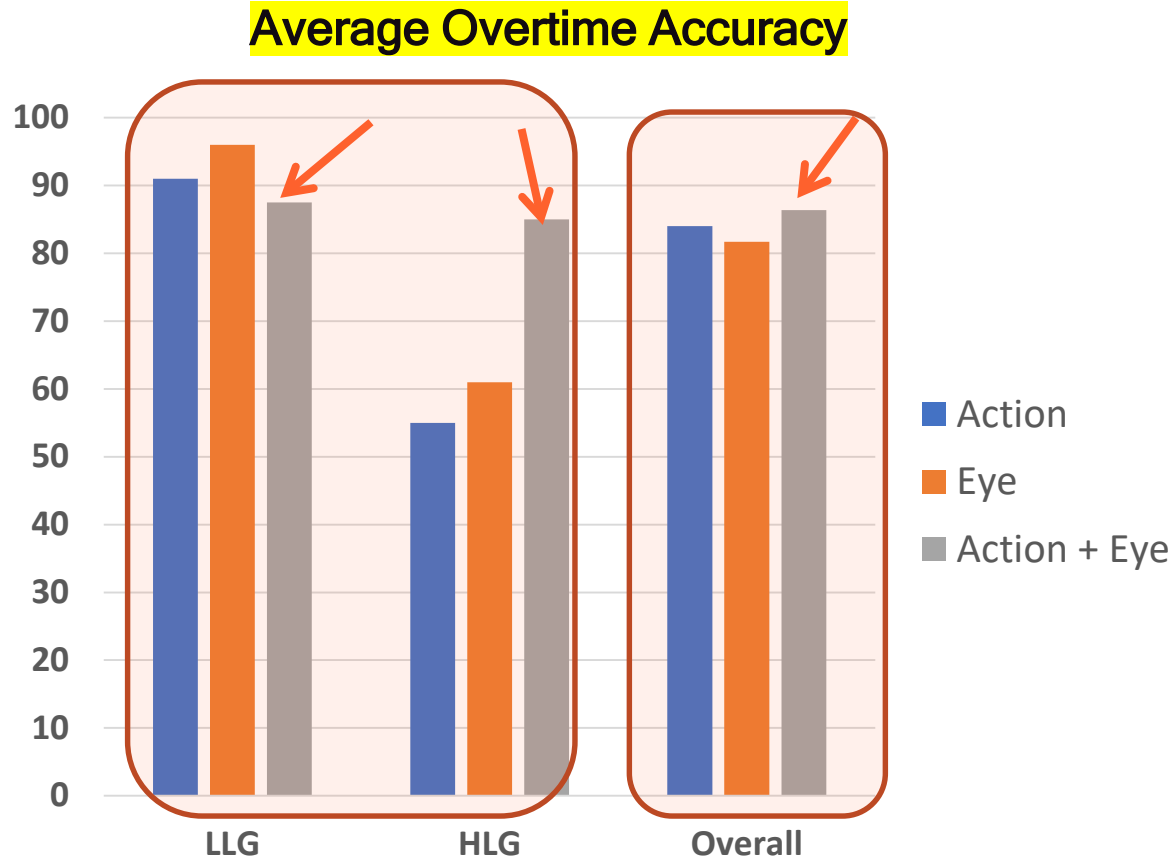
Apply FUMA to Action and Gaze Data

- Still found 2 clusters

- Higher and Lower Learners

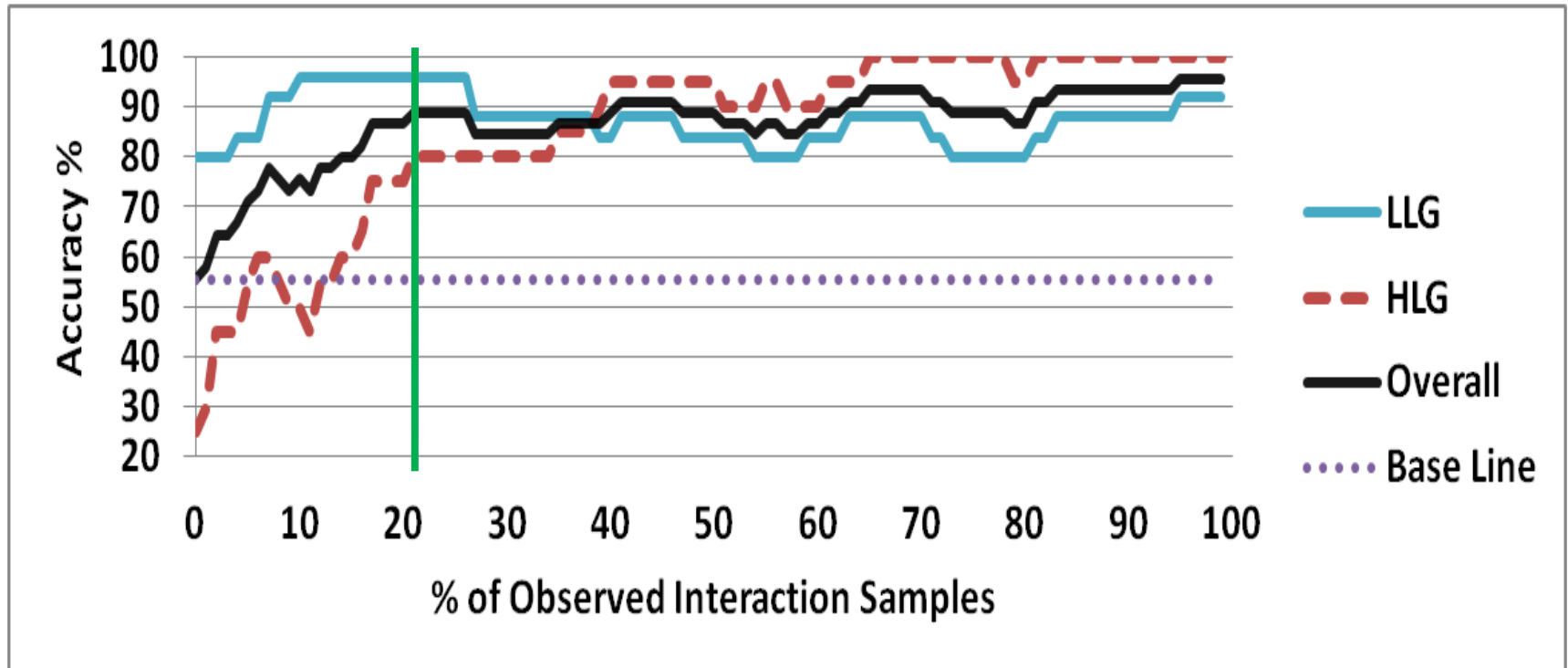
- Compared classifiers based on

- Action only
- Gaze only
- Gaze + action



Merging Action and Gaze Data

[Kardan and Conati 2013]



- ❑ Action + Gaze classifier achieves 80% classification accuracy over both classes after seeing 22% of the data



Multimodal Data: Lessons Learned

- ❑ Combining **action** and **gaze** data increases classification accuracy
- ❑ But the **associations rules** from these **multimodal** clusters are **harder** to turn into actionable hints
 - ❑ They may include features such as **average saccade angles** or **fixation rate**
- ❑ Solutions to investigate
 - ❑ Use multimodal data for classification/user modeling, but only action features to build hints
 - ❑ Use only higher-level gaze features (E.g. transitions between AOI)
 - ❑ Other?
- ❑ More future work:
 - ❑ investigate the **tradeoff** between classification **accuracy** and rule **interpretability** with **other multimodal data**

Overview

- Overview of FUMA and initial results with ACSP applet [Amershi and Conati 2009, Kardan and Conati 2012, 2015]
- Extensions to
 - Multimodal data
 - More complex OELEs
- What's next?

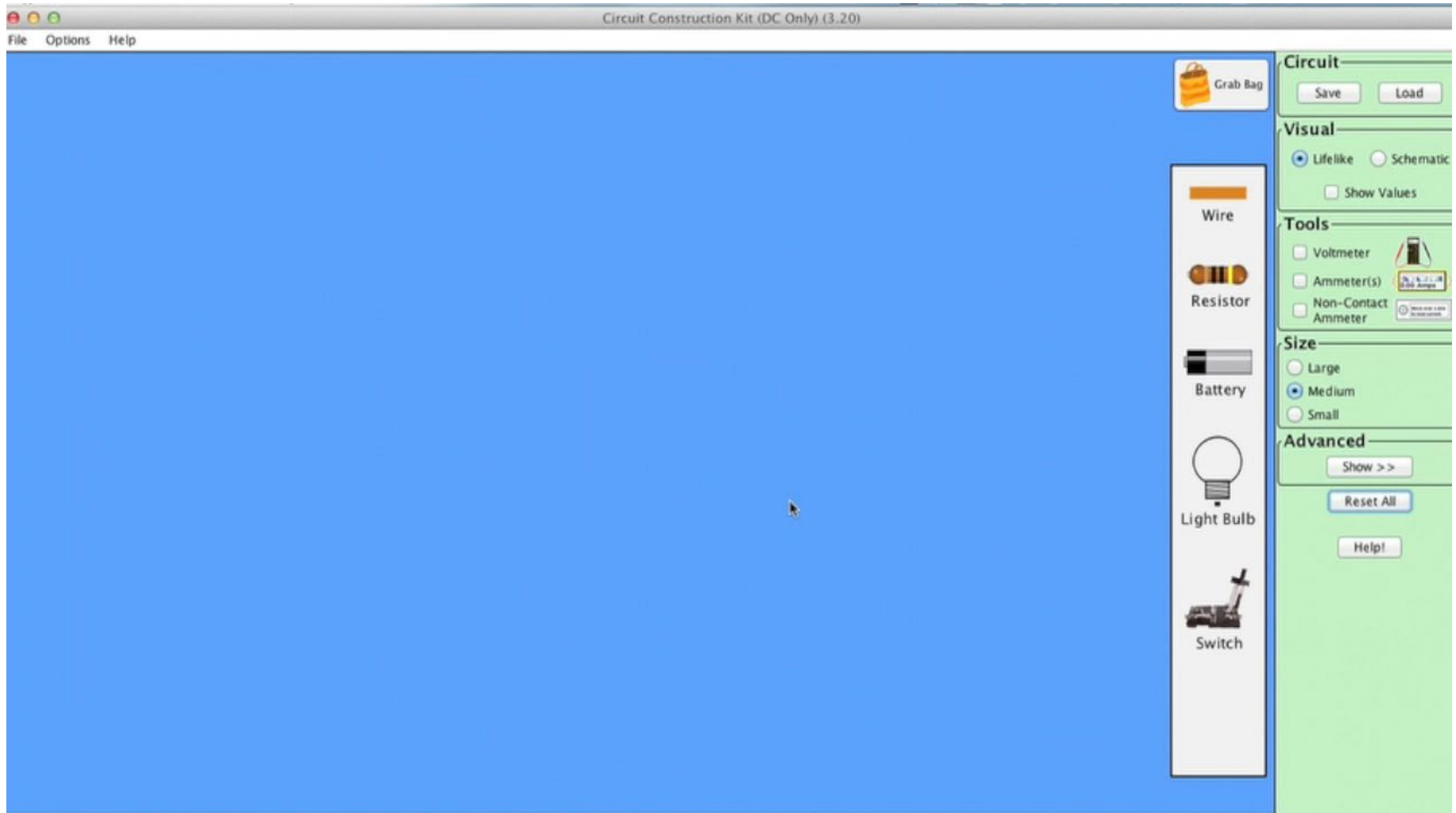
PhET DC Circuit Construction Kit (CCK)

- ❑ Part of large suite of simulations developed of U. of Colorado
- ❑ Allows students to explore building electrical circuits



The screenshot shows the PhET DC Circuit Construction Kit (CCK) simulation interface. The main window is titled "Circuit Construction Kit (DC Only) (3.20)" and has a menu bar with "File", "Options", and "Help". The simulation area is a blue workspace with a grid of blue dots. A circuit is built using orange wire, a battery, a light bulb, a resistor, and an ammeter. A voltmeter is connected across the light bulb, displaying "9.00 V". A callout box shows "0.90 Amps". The right sidebar contains a "Grab Bag" icon, a "Circuit" section with "Save" and "Load" buttons, a "Visual" section with "Lifelike" (selected) and "Schematic" radio buttons, and a "Show Values" checkbox. The "Tools" section includes checkboxes for "Voltmeter", "Ammeter(s)", and "Non-Contact Ammeter". The "Size" section has radio buttons for "Large", "Medium" (selected), and "Small". The "Advanced" section has a "Show >>" button, a "Reset All" button, and a "Help!" button. At the bottom, there are pause and play buttons.

Interaction Demo



Complex Interaction

25 **actions**, eg:

On circuit elements

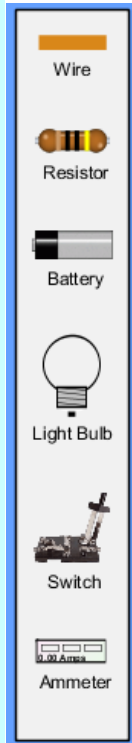
- Add
- Move
- Remove
- Join

Measurement

- Voltage
- Current

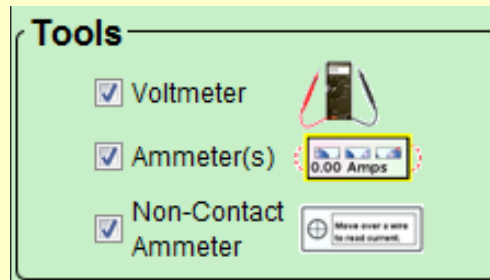
On Interface

- Simulation settings
- Window



22 **components**, eg:

- Basic circuit elements
- Measurement tools



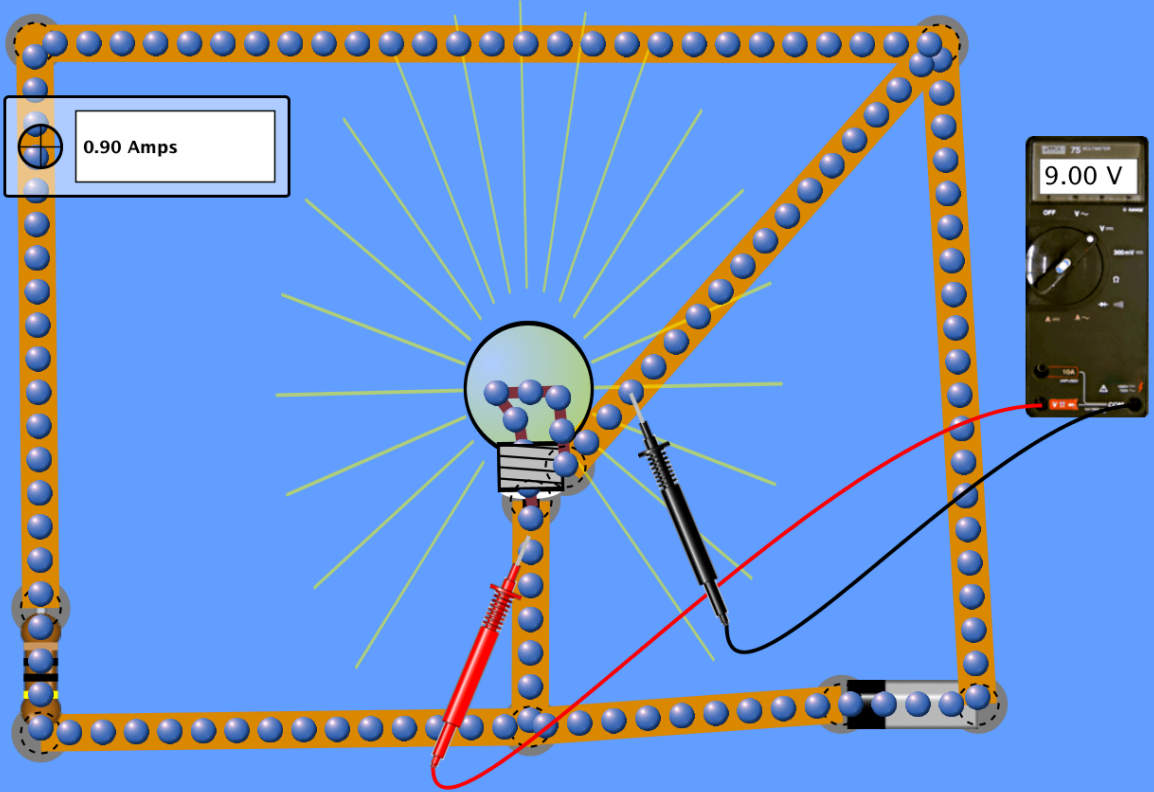
Context-dependent **outcomes**

- Light intensity change
- Current change
- Fire
- Measurement Reading change
- None

- Many ways to interact

- Context plays an important role

» Different outcomes depending on the state of the circuit



Grab Bag



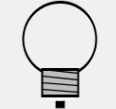
Wire



Resistor



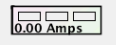
Battery



Light Bulb



Switch



Ammeter

Circuit

Save Load

Visual

Lifelike Schematic

Show Values

Tools

Voltmeter

Ammeter(s)

Non-Contact Ammeter

Size

Large

Medium

Small

Advanced

Show >>

Reset All

Help!



Layered Representation to Capture Complex Interaction with FUMA

[Fratamico et al. AIED 2015, JAIED 2017]

4 layers:

- ❑ Actions (A), Components (C), Outcomes (O)
 - from **logs**
- ❑ Families (F): **engineered**
 - Abstract **actions** into **8** more **general activities** that students can perform in CCK, e.g.
 - » Build (add, changeResistance, join)
 - » Revise (changeResistance, join, split, remove)
 - » Test (startMeasure, endMeasure, traceMeasure)

Representing the User Interaction

- ❑ Different **combinations** of the 4 layers represent **interaction-events** at different **granularities**, e.g.:
 - All 4 layers (OFAC)
 - » **current_change.revise.join.wire**
 - » Student generated a **current change** while **revising** the circuit by **joining** two **wires**
 - Outcome, Action, Component: (OAC)
 - » **current_change.join.wire**
 - » Does not include high level information on family
- ❑ Tested FUMA on 11 of these combinations, based on
 - **Quality** of the derived **clusters**
 - Classification **accuracy**
 - **Usefulness** of the generated association rules for **adaptive interventions**

FUMA

Behavior Discovery

Actions Logs
Other Data

Vector of
Interaction
Features

Clustering

Association
Rules
Mining

Interpret in terms of
learning

User Classification

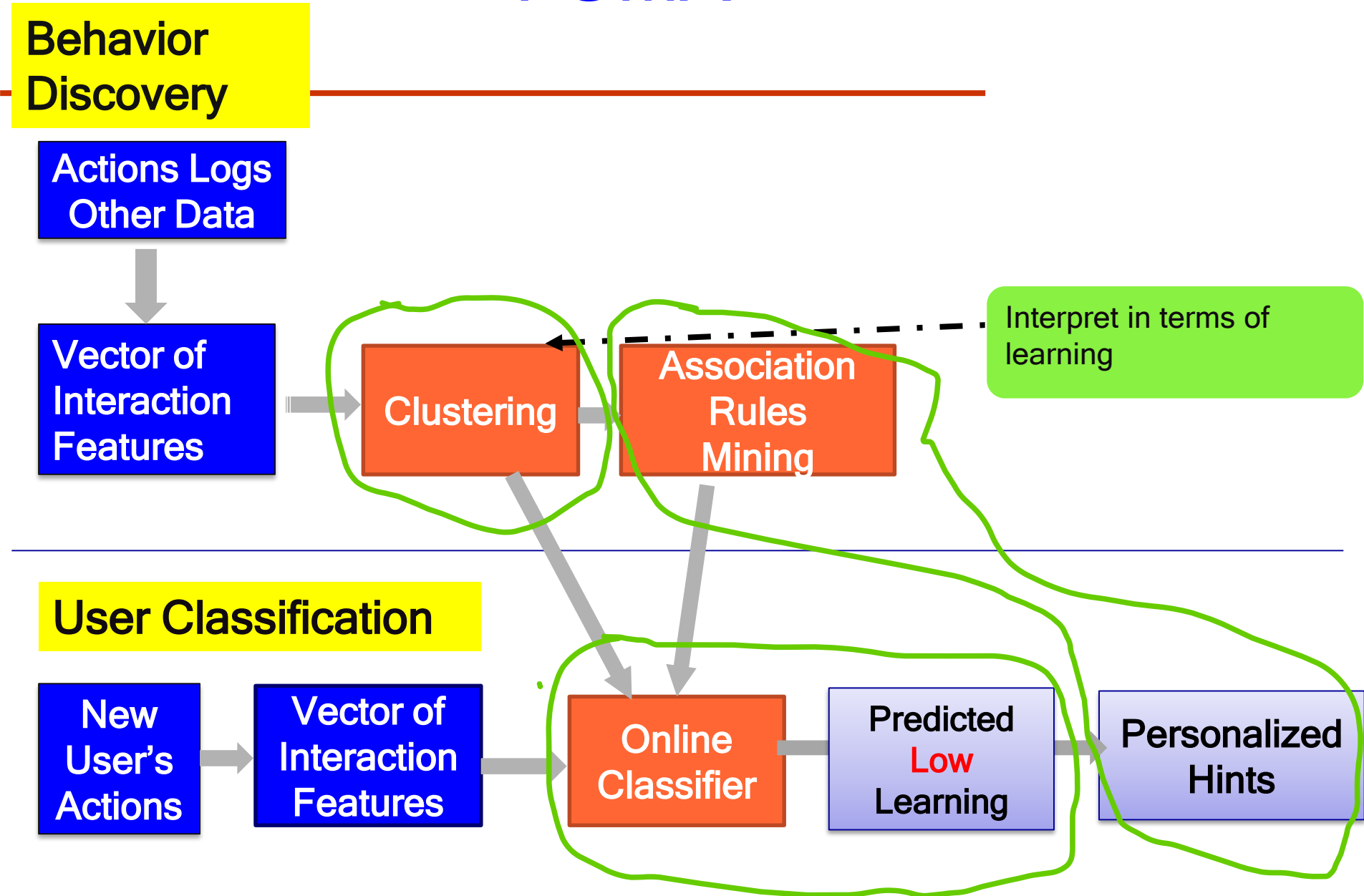
New
User's
Actions

Vector of
Interaction
Features

Online
Classifier

Predicted
Low
Learning

Personalized
Hints



User Study

Data collected from a lab study with CCK

- ❑ 96 UBC students taking a first year physics course
- ❑ Were given a general learning goal:
 - *Explore how resistors affect the behavior of circuits by exploring different combinations of resistors and resistances*
- ❑ Collected pre and post test data

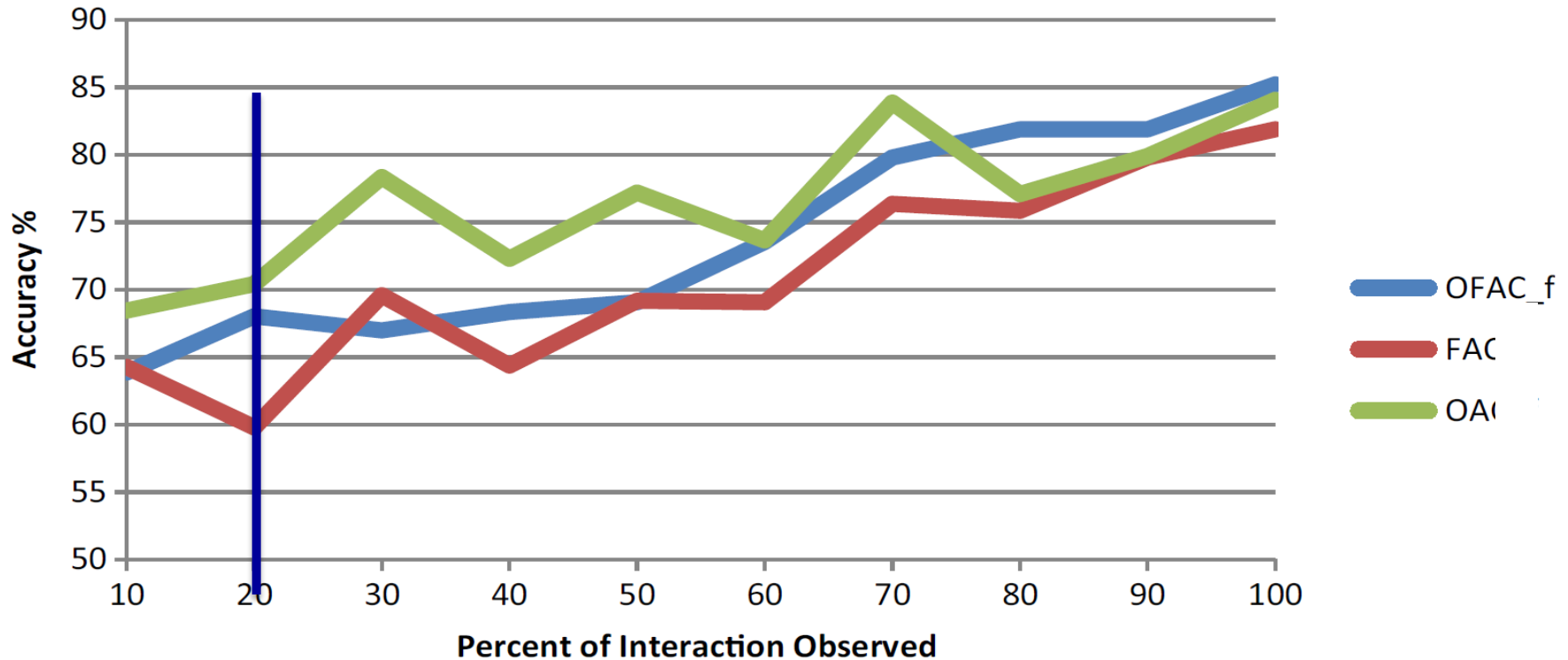
Quality of Clusters

- 3 of the 11 feature sets generated clusters (2) with significant difference in learning gains

Feature Set	Effect Size (partial η^2)
Family.Action.Component (FAC)	.041
Outcome.Action.Component (OAC)	.076
Outcome.Family.Action.Component (OFAC)	.065

OAC achieves the best cluster quality in term of highest difference in learning gains

Classification Accuracy



❑ OAC is the best classifier

- Achieves 70% accuracy after seeing 20% of interaction data (~ 5 min)
- OFAC gets there after seeing 50% of the data

Generated Association Rules

- ❑ All features sets identified **4 general behavior patterns** that instructors confirmed to **impact learning** with CCK
 - test frequently
 - frequently change resistance of resistors
 - pause to reflect in between actions
 - limit the usage of light bulbs and changes to their light intensity
- ❑ **OFAC** generated **more specific** rules (22)
 - Against the **15** generated by **OAC**
- ❑ Better suited to provide **incremental** feedback, e.g.
 - Start at the “Family” level; (e.g. “Test more”)
 - **Incrementally** go into **more detail** on how to do it

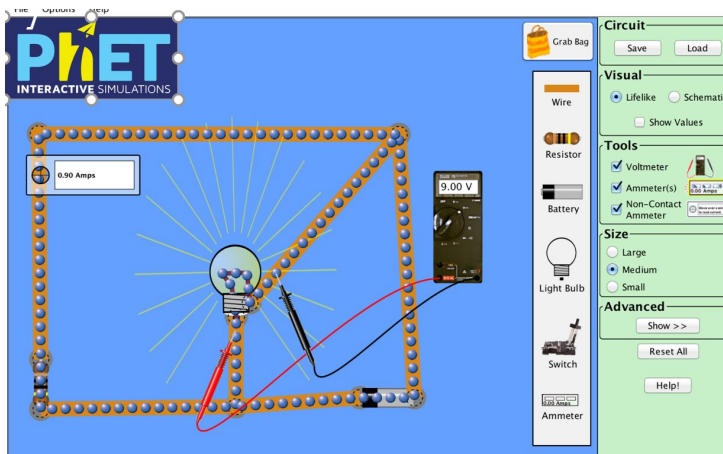
Summary of Results

- ❑ **OAC** best for **classification accuracy**, specifically for providing **timely hints**
- ❑ **OFAC** best for **usefulness** of the generated association rules
 - can provide **richer** hints
- ❑ Need to empirically explore **tradeoff** between there two factors
- ❑ Investigate if this tradeoff exist with **other** complex ELEs

FUMA: Evaluated in several ELEs

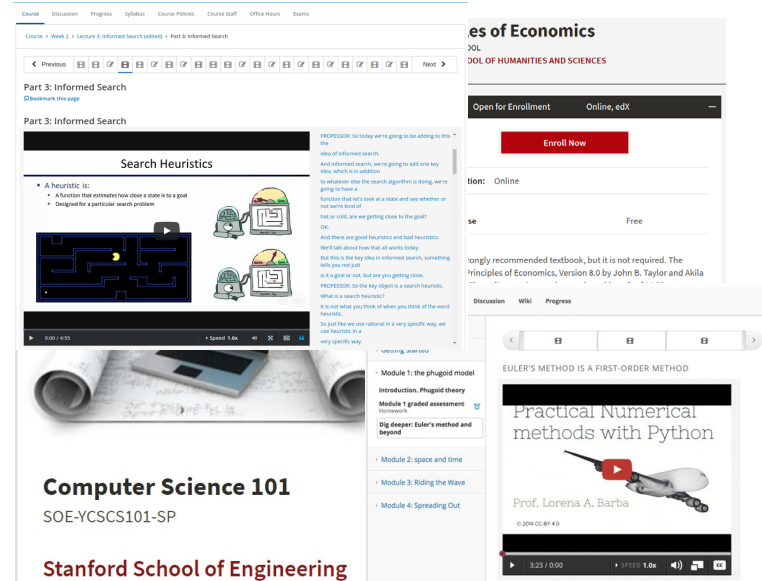
PHET Circuit Construction Kit (CCK)

Kardan et al., 2014; Fratamico et al., 2017;



Four MOOCs

Lallé et al., 2020,



ELE for Game Design

Lallé et al., 2021,



Unity-CT



Collaboration with **UME Academy**:

- Use the popular **Unity game engine** to teach Computational Thinking (CT) skills to K-12 kids

- **Free-form** interaction to create small games

<https://ume.academy>

Unity-CT



Collaboration with **UME Academy**:

- Use the popular **Unity game engine** to teach Computational Thinking (CT) skills to K-12 kids

- **Free-form** interaction to create small games

<https://ume.academy>

Unity-CT



Collaboration with **UME Academy**:

- Use the popular **Unity game engine** to teach Computational Thinking (CT) skills to K-12 kids



- Classes facilitated by a UME instructor

- Can we have AI agents that helps with this facilitation?

<https://ume.academy>

FUMA for Unity CT [Lalle et al LAK 2021]



Actions Logs
Other Data

Vector of
Interaction
Features

Clustering

Association
Rules
Mining

Behavior
Discovery

User Classification

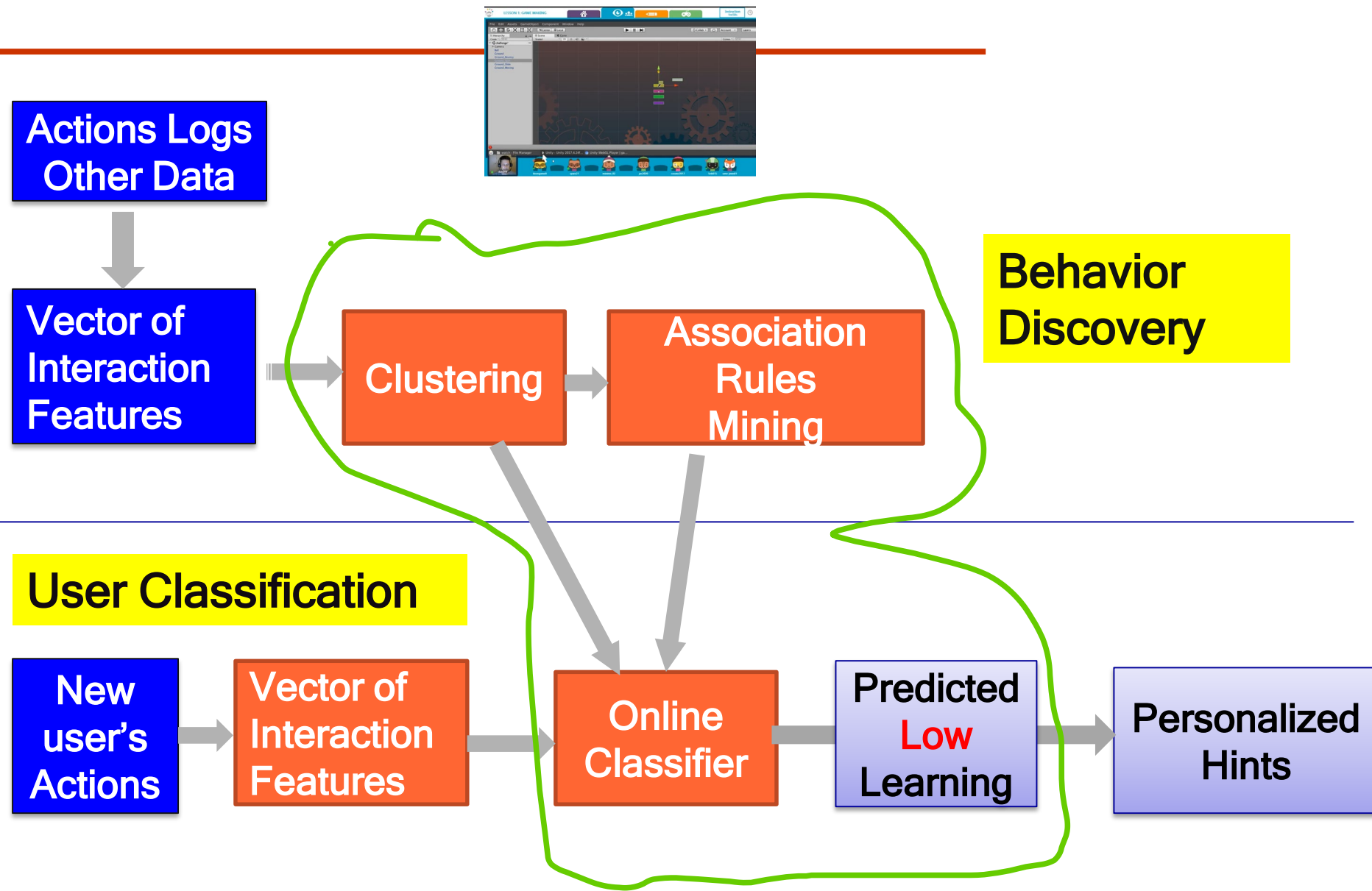
New
user's
Actions

Vector of
Interaction
Features

Online
Classifier

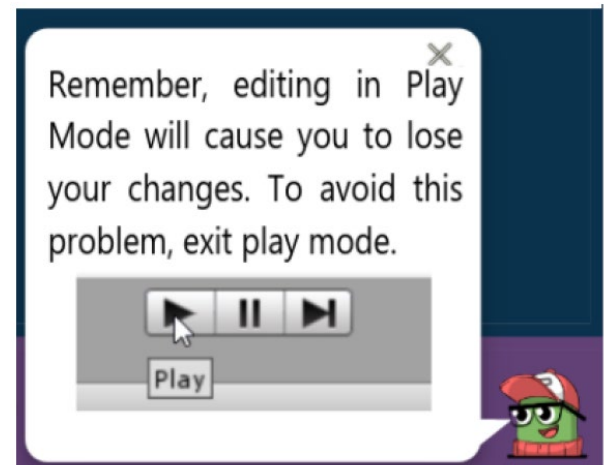
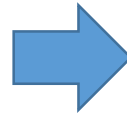
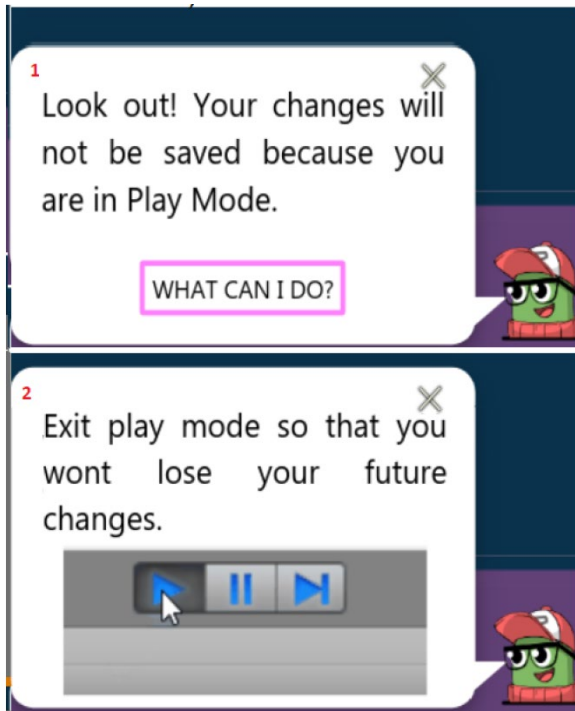
Predicted
Low
Learning

Personalized
Hints



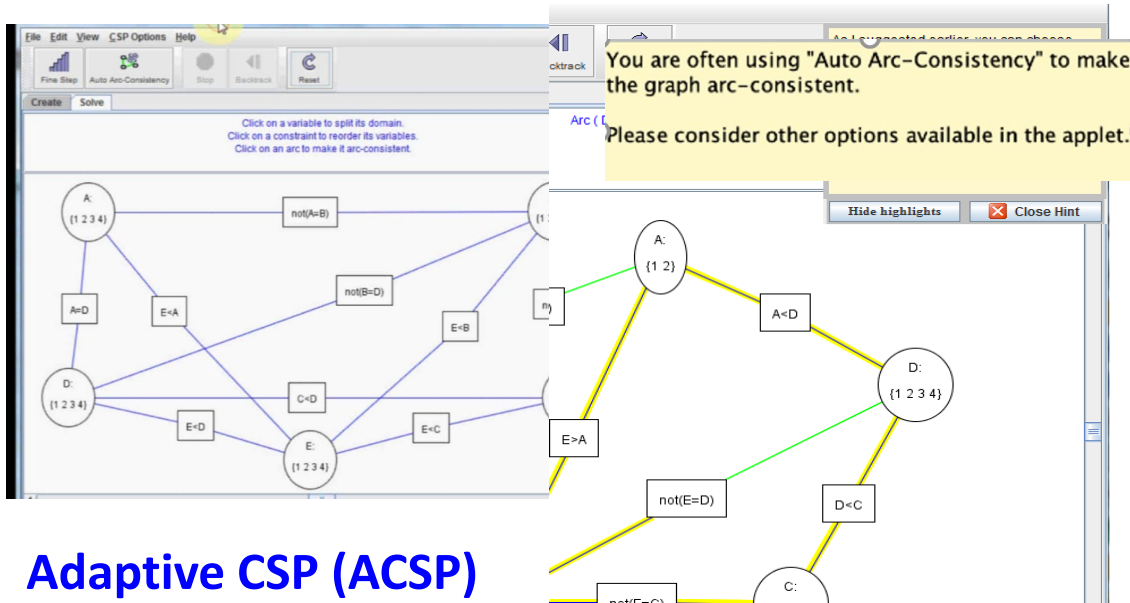
Current Work

- Designing the delivery of adaptive support, with UME **UX expert** and **instructors**
- For instance, what to do about **repeated hints** (Yalcin et al. IUI 2022, AIED 2023)



Explaining FUMA Hints

[Conati et al., AI Journal 2021]



Adaptive CSP (ACSP)

FUMA-driven hints
shown to improve
student learning
[Kardan and Conati, CHI 2015]

Evidence that these hints are more effective if the system can **explain why** and **how** they were generated

And that hint **explanations** may be even **more effective** if they are **personalized** to specific **student characteristic**

Conclusions

- ❑ **FUMA**: data-driven framework for user modeling and personalization to support learning with **ELEs**
- ❑ Evaluation with several ELEs show that FUMA can
 - **Identify** clusters with **behaviors** representative of student **performance**
 - **Classify** student **performance** with good accuracy, **early enough** to generate help when needed
 - **Drive** the design of **personalized help** from the detected behaviors
- ❑ Initial evidence that FUMA-driven interventions can help learning
 - And that their effectiveness can be improved with **explanations**

Future Work

- ❑ Apply FUMA to **other OLEs**
- ❑ Experiment with **multimodal data**
- ❑ More evidence that FUMA-driven hints foster learning
- ❑ These hints are **shallow**.
 - How do they compare against **richer, knowledge-based** hints?
- ❑ Consider student **affect** for hint provision
- ❑ Look at **collaborative** activities
- ❑ Continue investigating the value of **personalized explanations** of FUMA-driven hints

Thanks To



Saleema Amershi



Oswald Barral



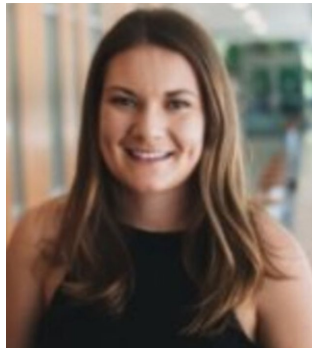
Lauren Fratamico



Sebastien Lalle



Samad Kardan



Vanessa Putnam



Ido Roll



Nilay Yalcin

And to all of you for your kind attention !