Learning Analytics: Modeling, Understanding, and Changing Student Behavior

Michael Eagle, Computational and Data Sciences,
George Mason University, meagle@gmu.edu

Michael Eagle

UNC CHARLOTTE

B.S. Computer Science

B.A. International Studies

Minor: Japanese

学習院大学
GAKUSHUIN UNIVERSITY

1.5 Years Abroad

WB | turbine

BLIZZARD ENTERTAINMENT

Data Science Internships

PhD Computer Science

SCS Carnegie Mellon School of Computer Science

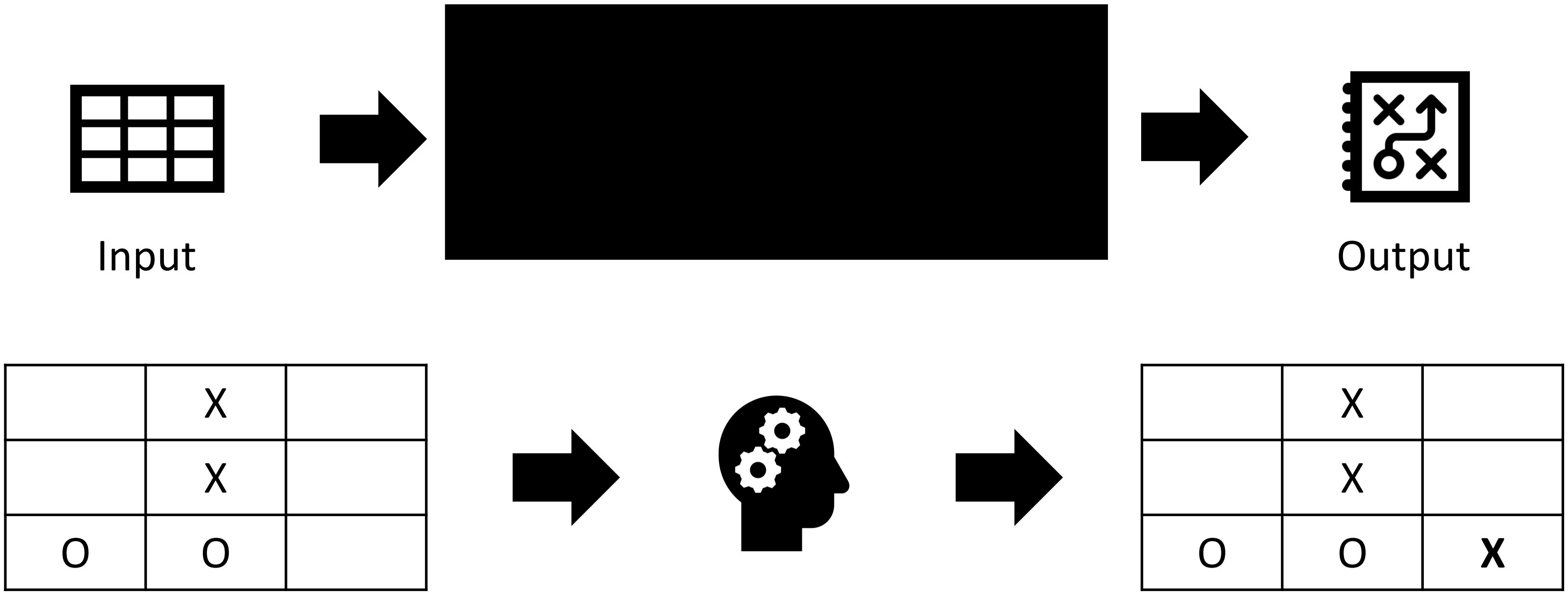
PostDoc

Computational and Data Sciences

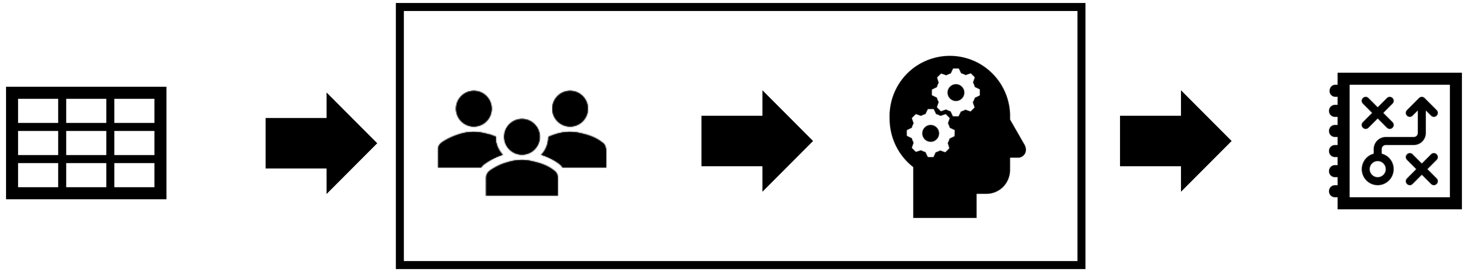
Assistant Professor

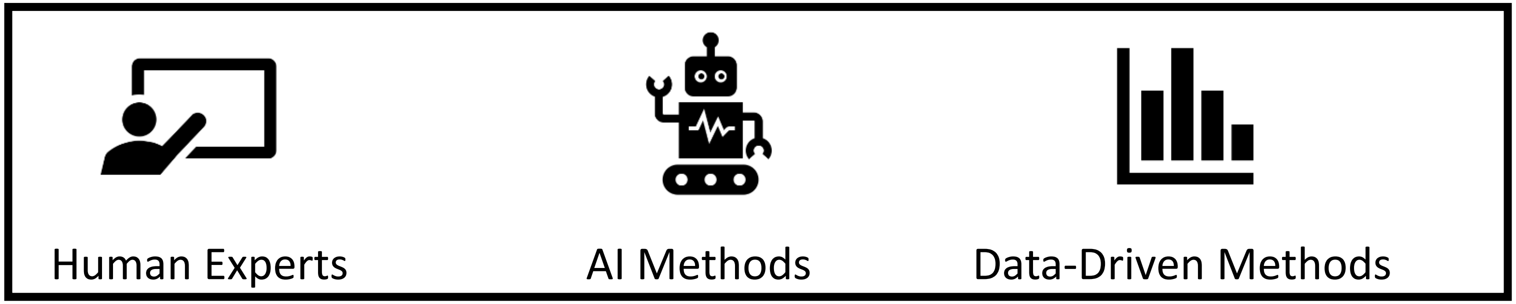
Behavior

Behavior

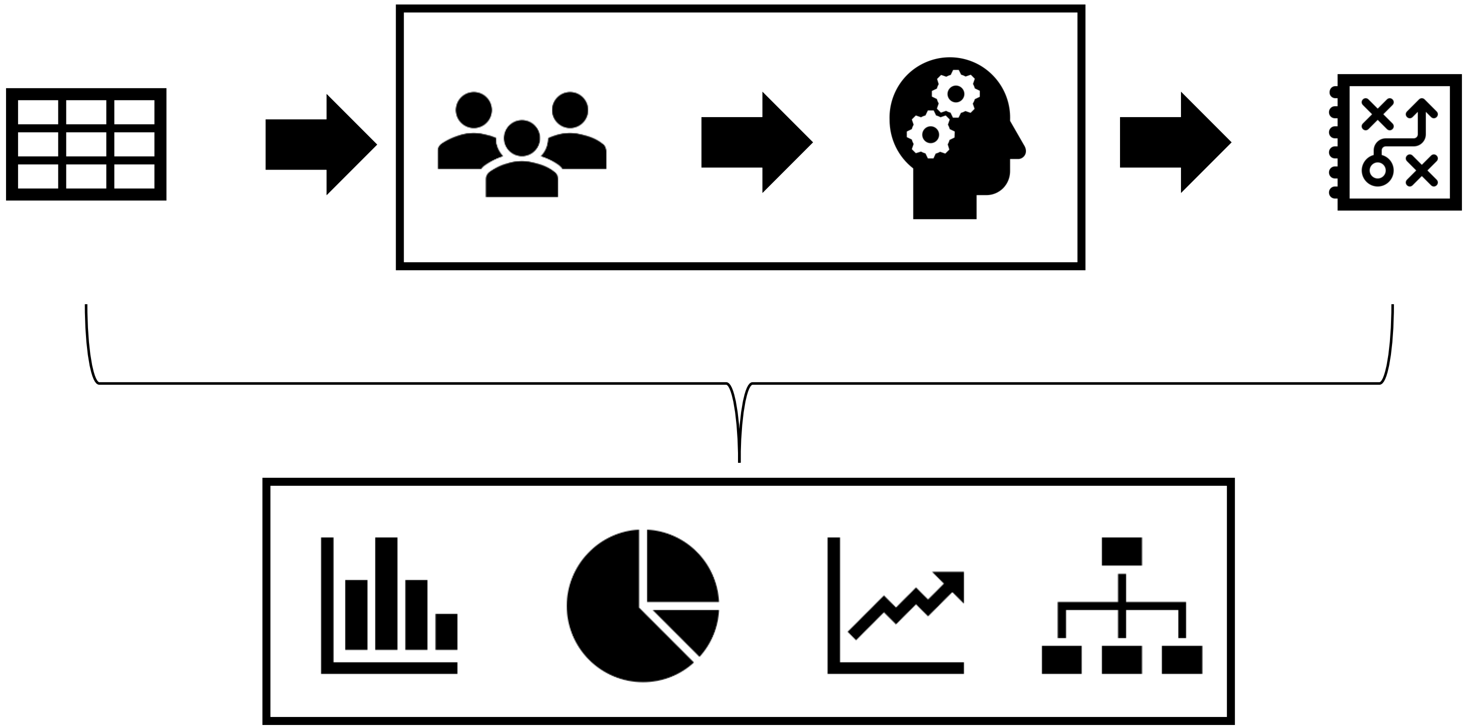


Modeling Behavior

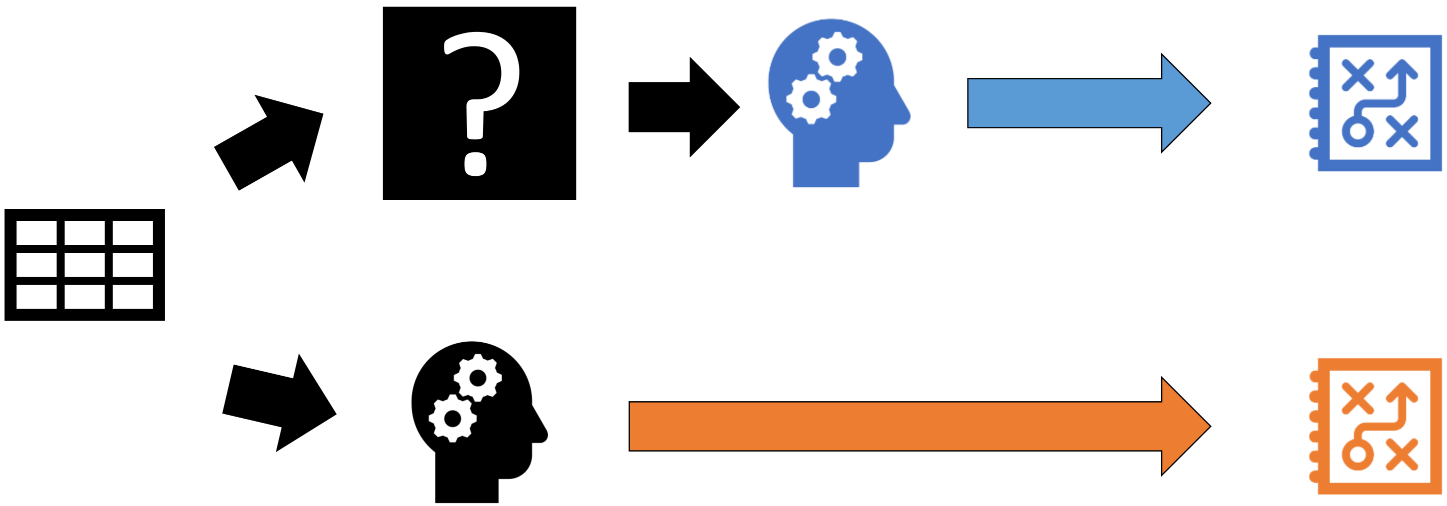




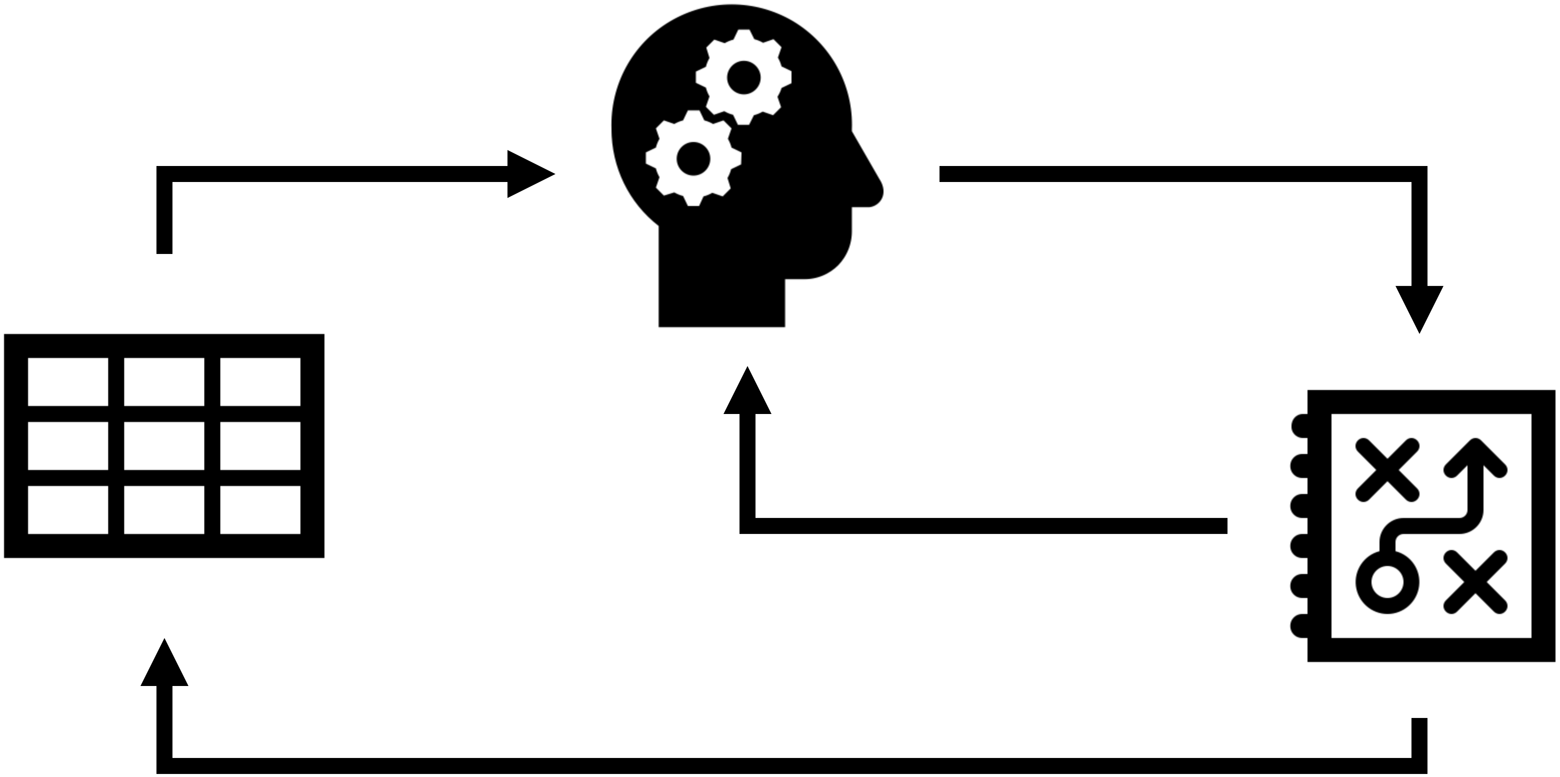
Understanding Behavior



Changing Behavior



Interactive Systems



Data Science Back Ground

Player Centered Data Science

WB | turbine

BLIZZARD ENTERTAINMENT

TERC

Learner Centered Data Science

Carnegie Mellon University
Master of Computational Data Science

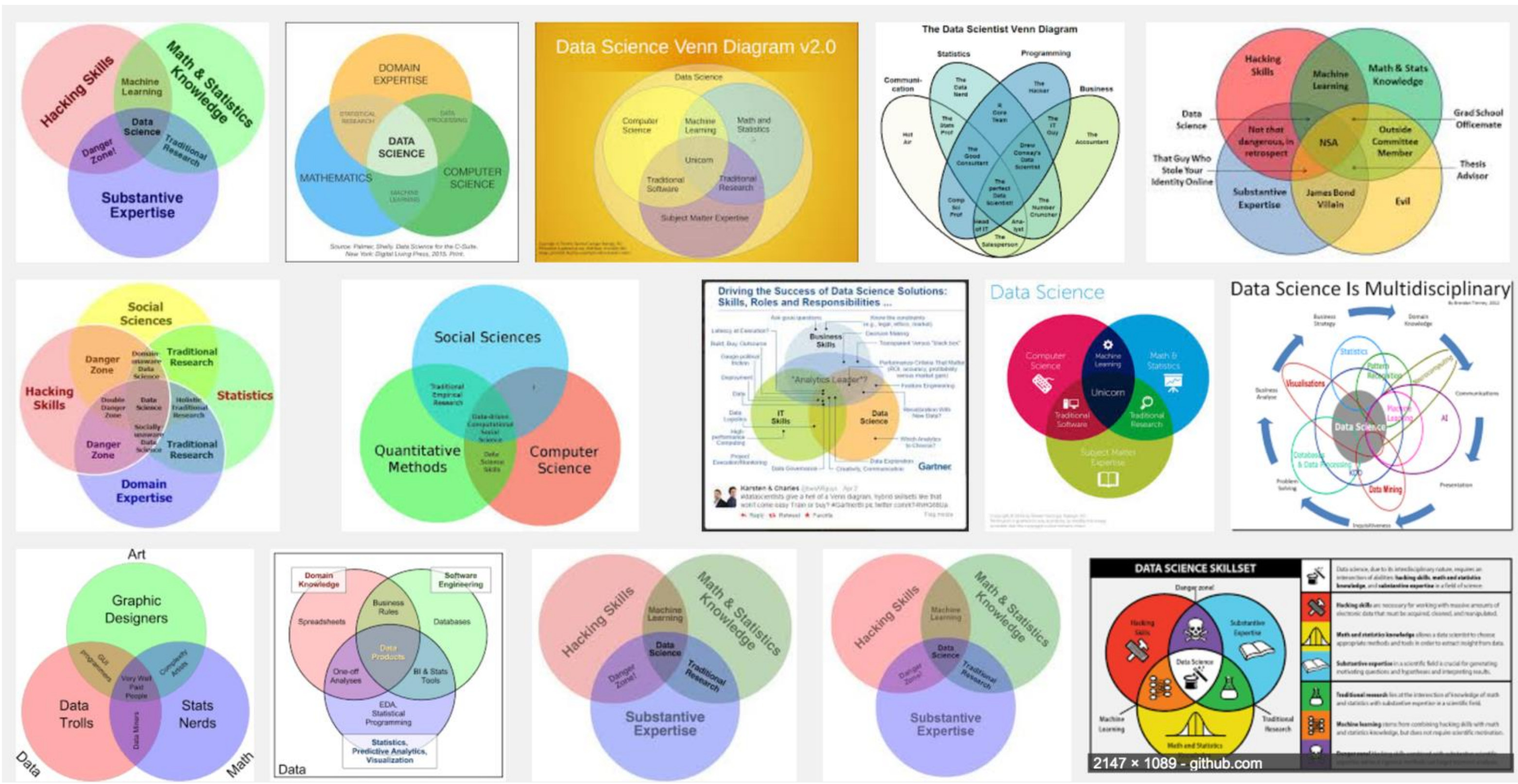
Human-Computer Interaction Institute

TutorGen

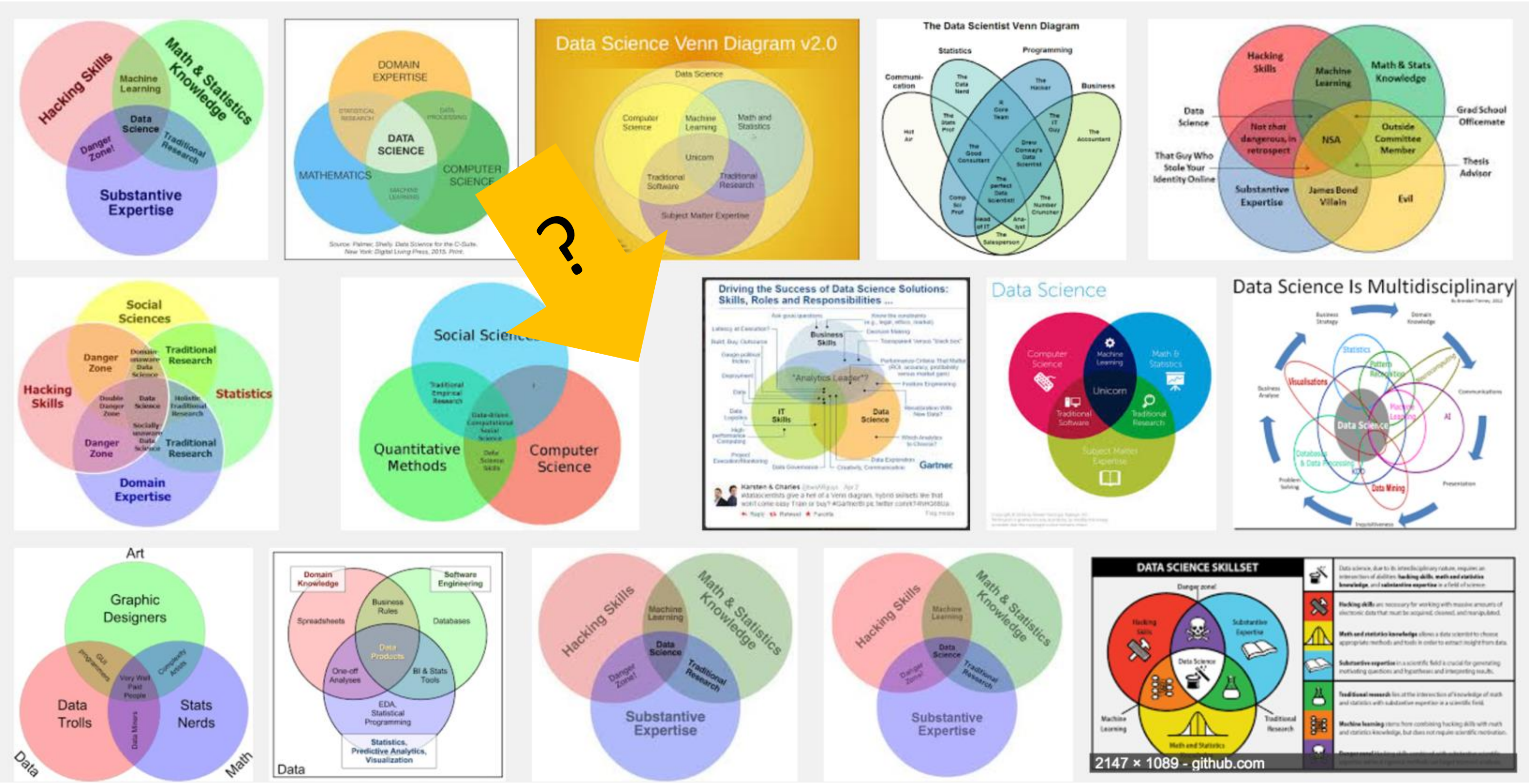
METALS
Masters of Educational Technology & Applied Learning Science

PhD

Data Science



Data Science: Where are you?



Why Human Centered Data Science?

- Intersection of multiple fields
- Research Methods
- Rapid Feedback Loops
- Understanding and modeling users
- Changing behaviors
- Communication with Design Teams
- *Scrappy Science*

Data Science: Hat Balancing

- Balance Between
  - Qualitative
  - Quantitative
  - Machine Learning
  - Rigor vs. Speed
  - **Communication**

Research Methods in Data Science

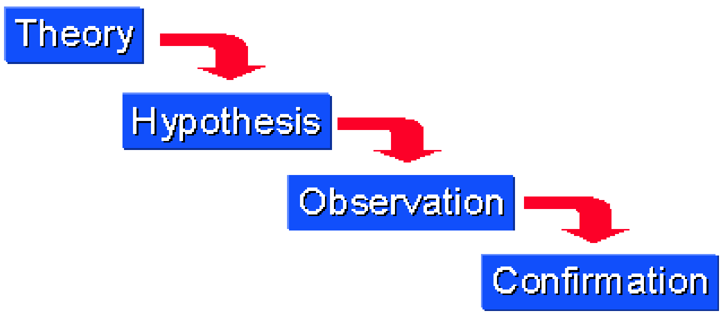
- Most people agree that Statistics and Machine Learning are core components of Data Science
- Less agreement on *Inferential* Statistics and Social Research Methods
- Experimental Testing Expertise is Increasing Desired from Industry (A/B Testing, etc.)

Qualitative Data?

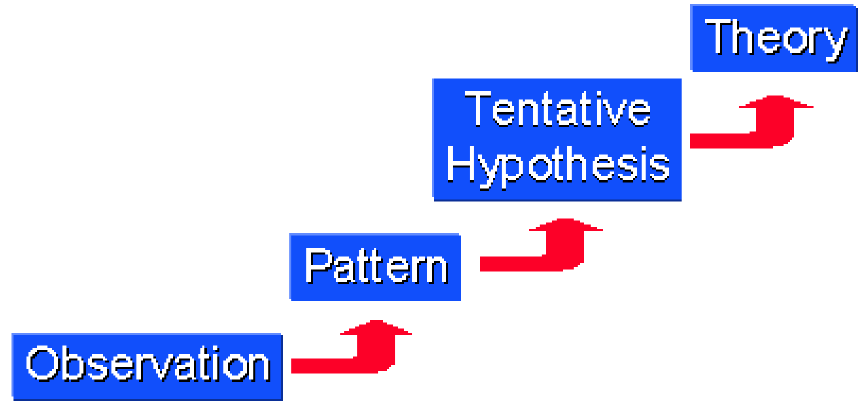
- Connected to domain knowledge
- Understanding new environments
- Guard against missing variables
- Understanding the products
- Get Insight Before Launch

Methodologies

- Deductive



- Inductive



Learning Analytics

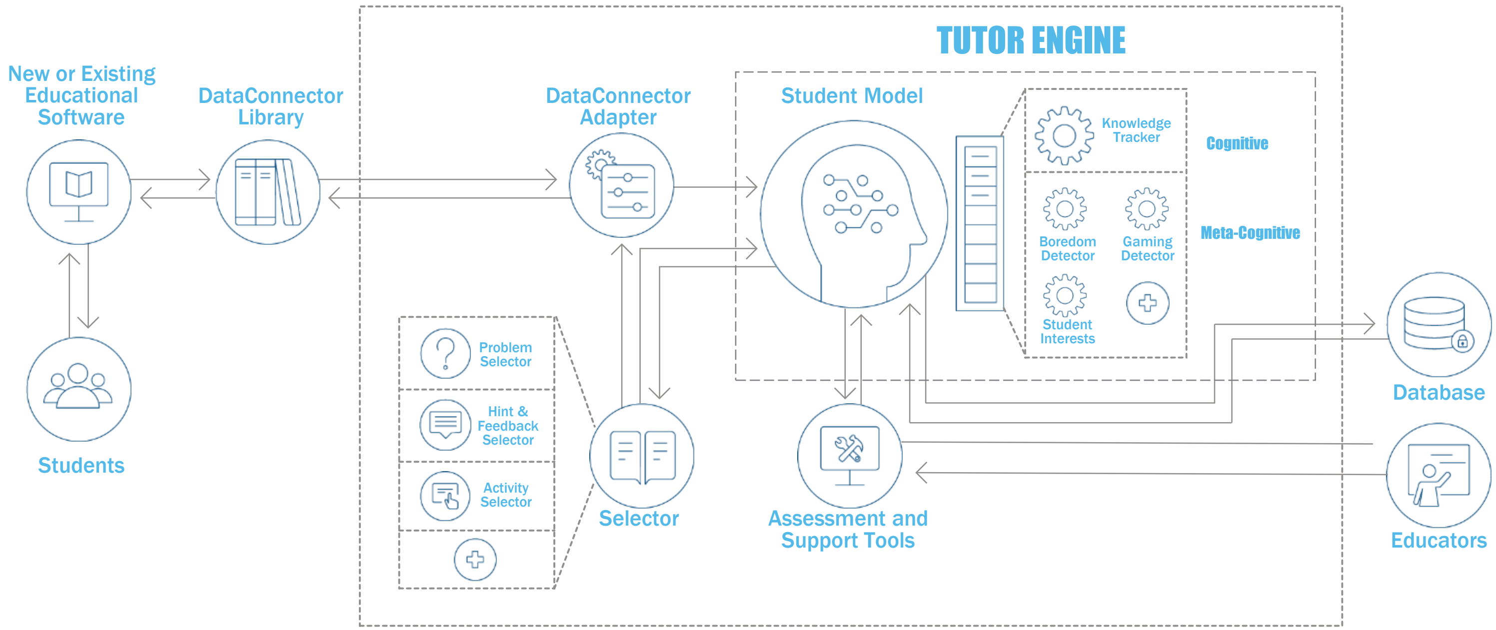
Where we are going

Tools that empower the designer, instructor, and students

- Predictive models that identify dropout
- Decision support for instructors
- Adaptive and personalized education

- Large systems that pull together lots of data

SCALE: STUDENT-CENTERED ADAPTIVE LEARNING ENGINE



SCALE Onboarding Process

ASSESS

TutorGen leads a review of existing client content and data to develop insights. Our process uses existing data with our architecture to reduce cost by not having clients recreate what they already have built.

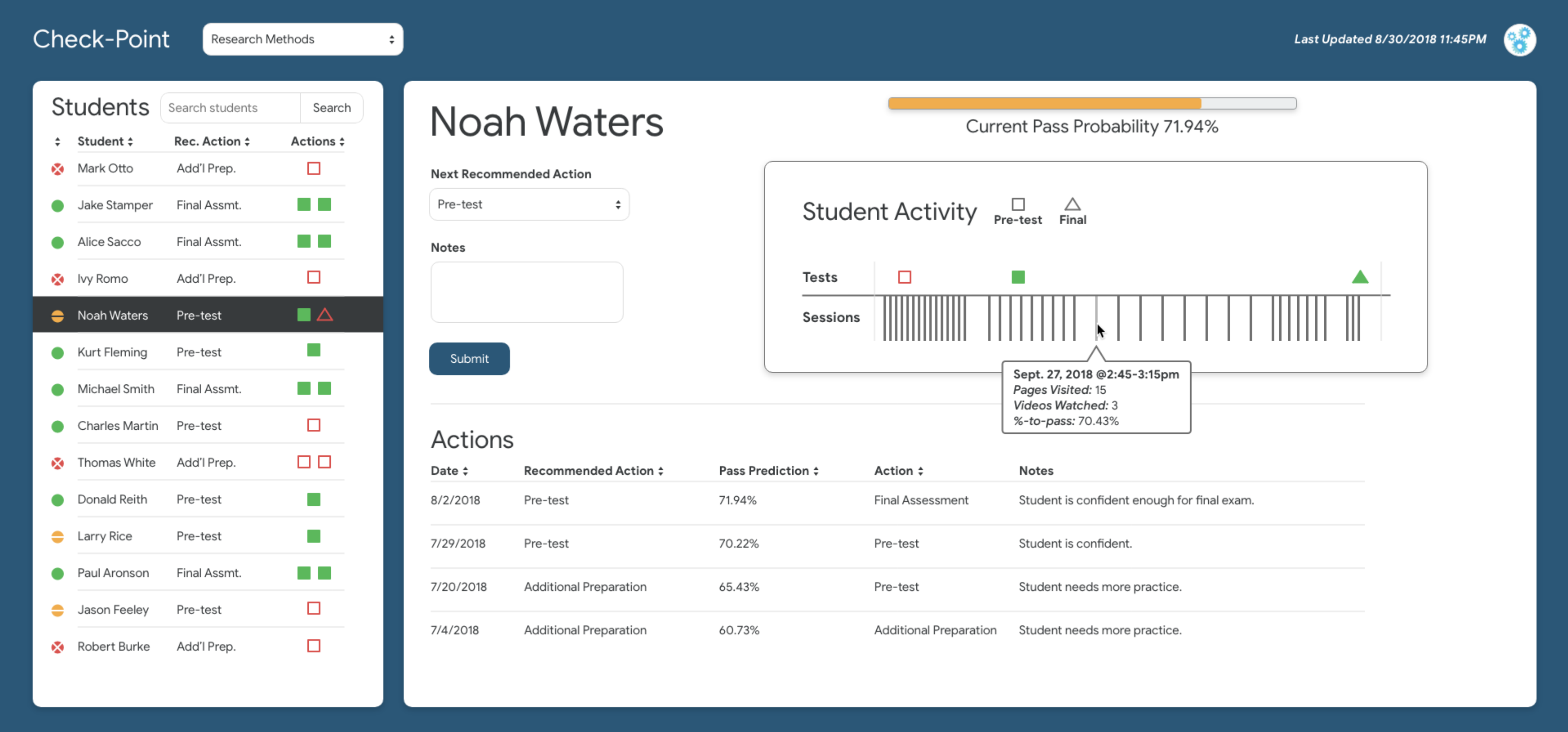
ANALYZE

TutorGen analyzes current data to identify opportunities and gaps within existing content to create an approach tailored to the needs of our clients and their learners.

Adapt

Using a continuous data-driven process, SCALE creates and updates problem and student models, building from subject matter experts, to personalize delivery of existing content to improve learning.

TutorGen®
INSPIRED BY LEARNING. POWERED BY TECHNOLOGY.



CheckPoint Features

Process

Supporting a wide variety of data streams to extract features used to build real-time, predictive student models.

Predict

CheckPoint performs continuous evaluation of available data streams and provides a prediction of student success and specific assessment performance.

Prompt

The CheckPoint dashboard interface recommends student interventions, tracks actual activities, and evaluates student results for review and tuning.

TutorGen
INSPIRED BY LEARNING. POWERED BY TECHNOLOGY.

Tools that empower the designer, instructor, and students

- Predictive models that identify dropout
- Decision support for instructors
- Adaptive and personalized education

- Large systems that pull together lots of data

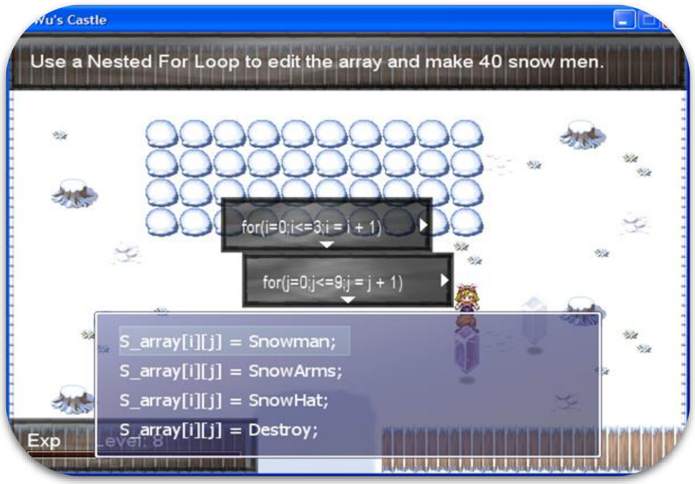
- **Where do we start?**

Background

How a few questions started a massive journey…

First Research: Educational Games

Developed games to teach computer science



Learned about games in education

Wu's Castle: Array

Stratified Sample

- Gender
- Race
- Gaming Ability
- Pretest scores

Switching Replications

- Strong experimental design
- Does not deny the game to some students
- Provides solution for social confounds
  - No one gets jealous of game-playing group
- Provides insight into ordering effects

Study Design

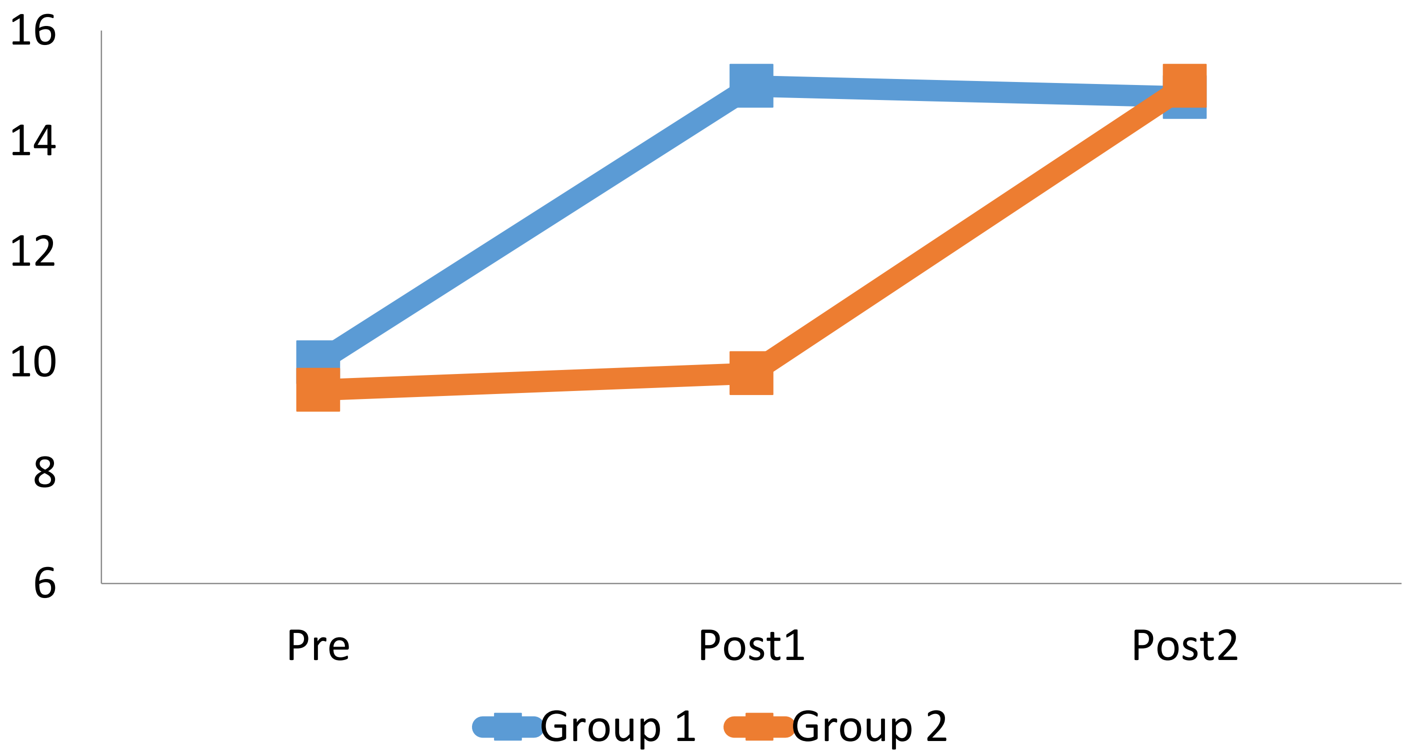
Experimental Group

Pretest
↓
Game
↓
Posttest1
↓
Programming Assignment
↓
Posttest2

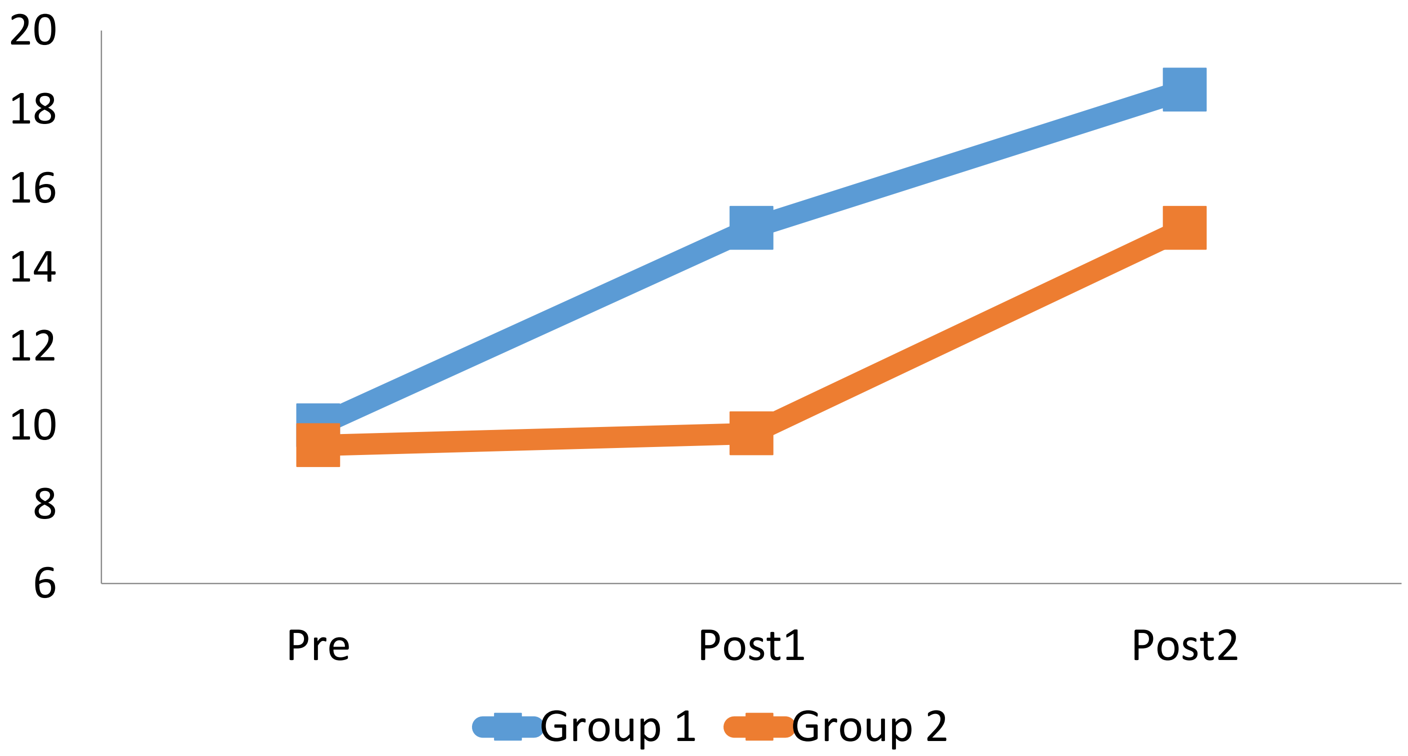
Control Group

Pretest
↓
Programming Assignment
↓
Posttest1
↓
Game
↓
Posttest2

Short-term persistent treatment effect



Long-term continuing treatment effect



Study Design

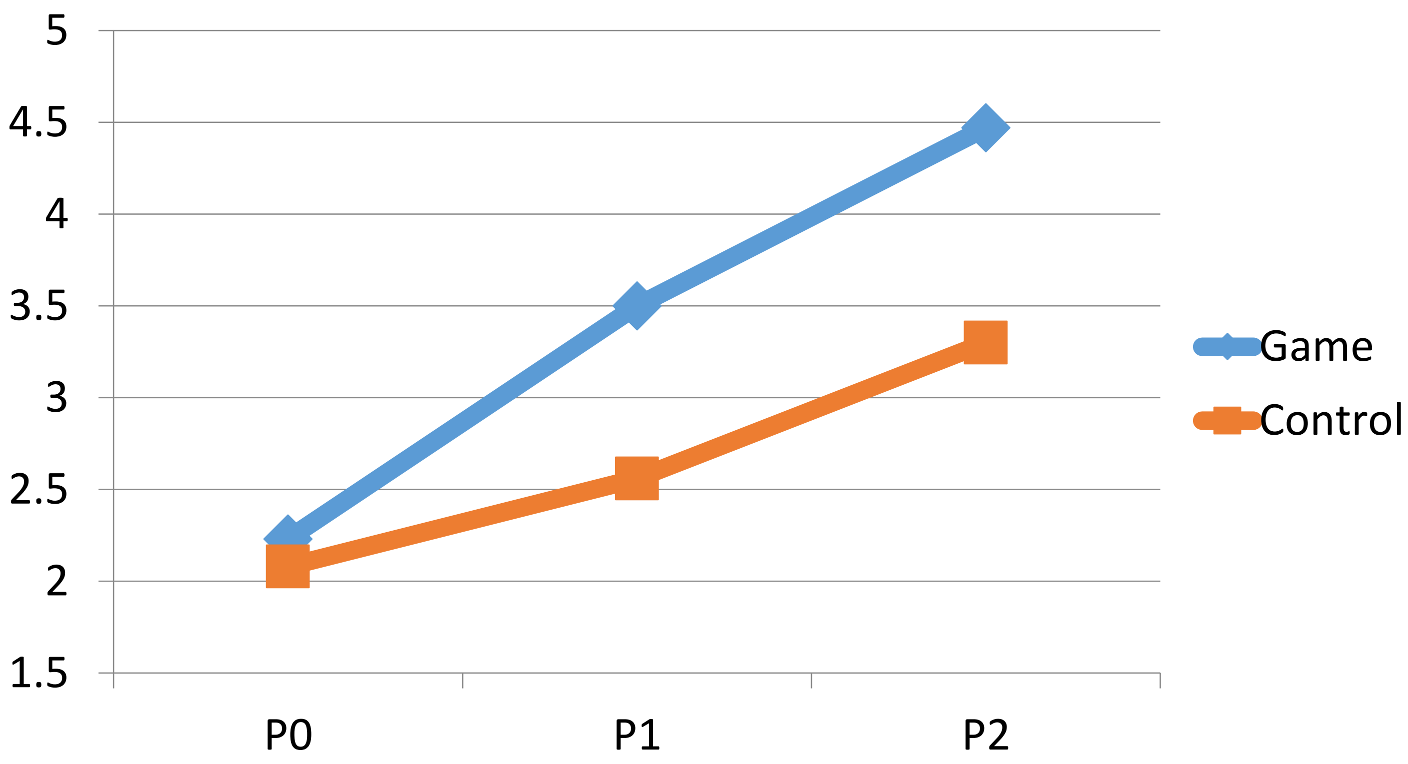
Experimental Group

1. Pretest
2. Game
3. Posttest1
4. Programming Assignment
5. Posttest2

Control Group

1. Pretest
2. Programming Assignment
3. Posttest1
4. Game
5. Posttest2

Long-term continuing treatment effect



Conclusions

- Game resulted in improved learning
  - Learning gains continued after game was played
- Both assignments combined provide better results
  - Game-first ordering is preferred

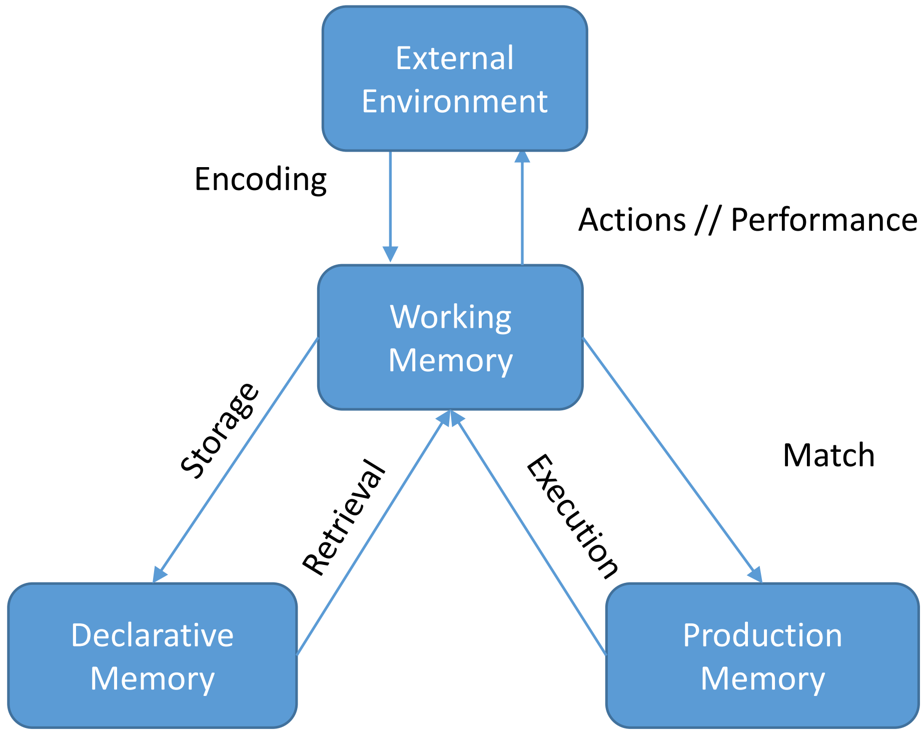
Personal Motivation

The Interesting Stuff is Happening
During the Treatment

Other methods of evaluation

- Techniques from:
  - Intelligent Tutoring Systems (ITS)
    - Systems that provide direct customized instruction or feedback to students without human intervention
  - Educational Data Mining (EDM)
    - Methods for exploring the data that comes from educational settings

ACT-R (1983)


External Environment
Encoding
Actions // Performance
Working Memory
Storage
Retrieval
Execution
Match
Declarative Memory
Production Memory

Declarative Module

- Declarative memory:
  - Facts
    - Washington, D.C. is the capital of the U.S.
    - 2+3=5.
  - Knowledge a person might be expected to have to solve a problem.
  - Called chunks

Procedural Module

- Procedural memory: Knowledge about how to do something.
  - How to type the letter "Q".
  - How to drive.
  - How to perform addition.

Procedural Module

- Made of condition-action data structures called production rules.
- Each production rule takes 50ms to fire.
- Serial bottleneck in this parallel system.

Procedural Module

Pattern Matching

(p name

condition part — Specification of Buffer Tests

delimiter — ==>

action part — Specification of Buffer Transformations

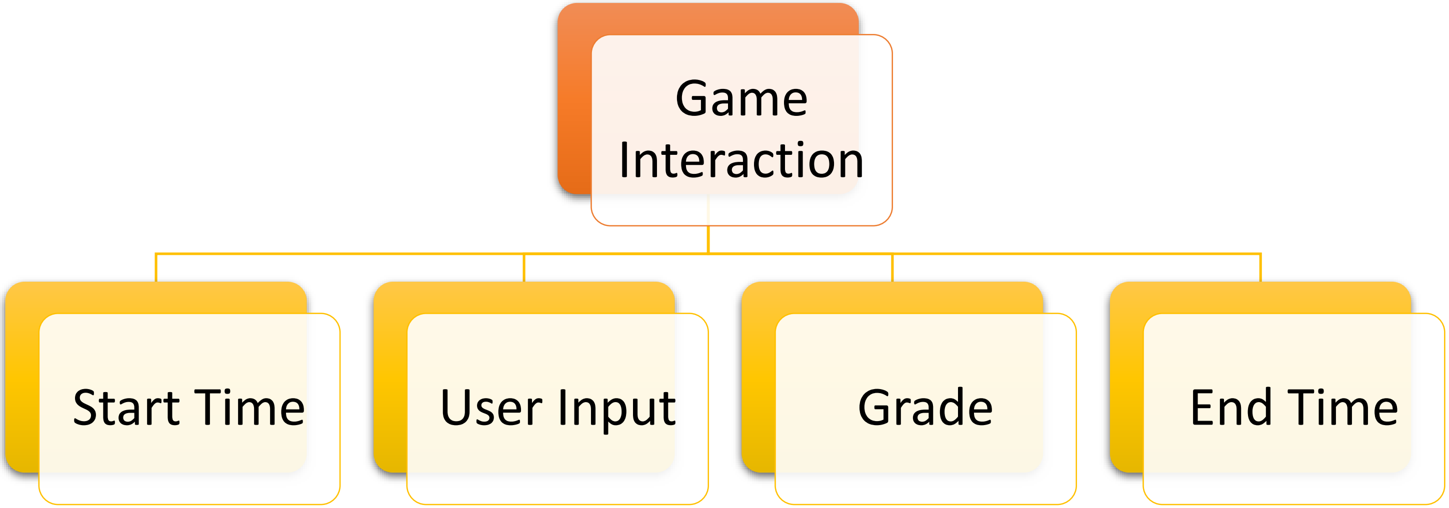
)

Procedural Module

```
( p     example-counting
        =goal>
           isa count
           state counting
           number =num1
        =retrieval>
           isa count-order
           first =num1
           second =num2
   ==>
        =goal>
           number =num2
        +retrieval>
           isa count-order
           first =num2
  )
```

IF the goal is
    to count
    the current state is counting
    there is a number called =num1
    and a chunk has been retrieved
    of type count-order
    where the first number is =num1
    and it is followed by =num2
THEN
    change the goal
    to continue counting from =num2
    and request a retrieval
    of a count-order fact
    for the number that follows
    =num2

Game Logging



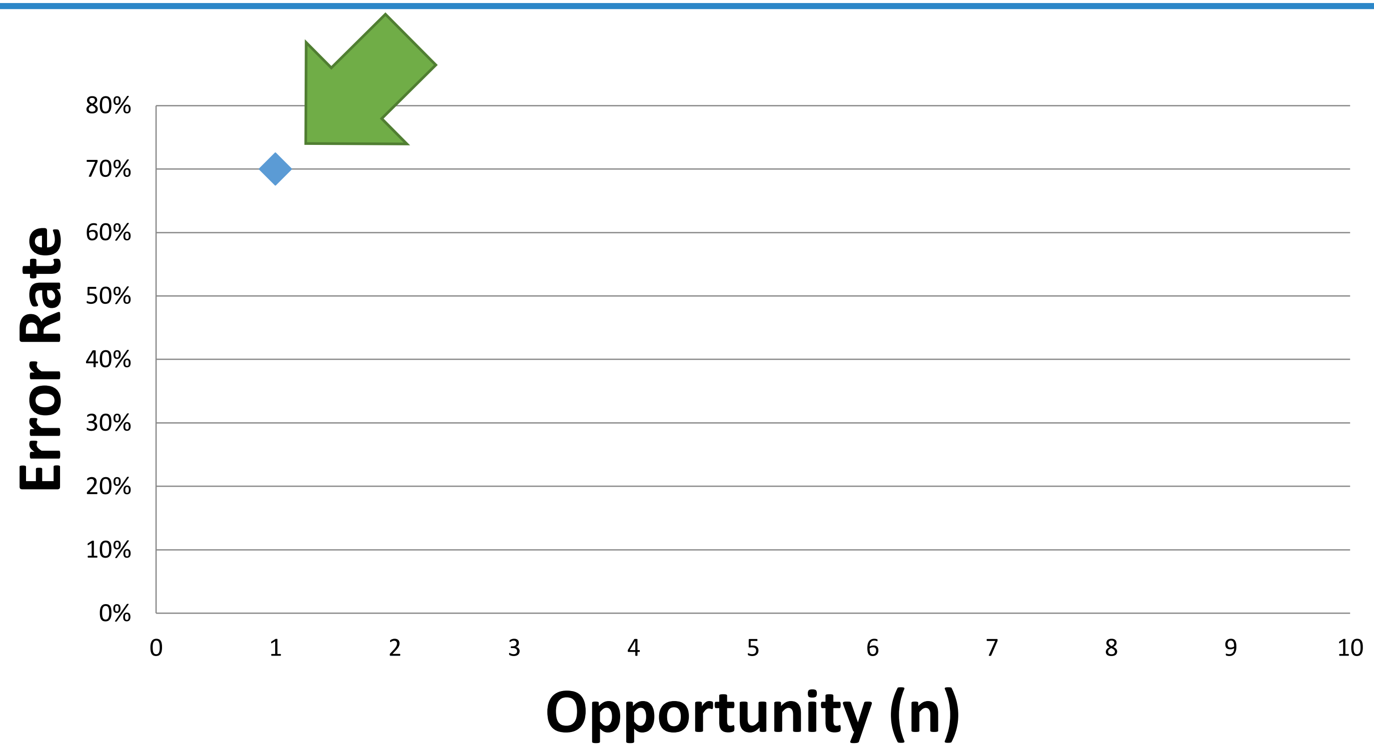
Empirical Learning Curves (ELC)

- Visualization of a student's performance over time, based on observed data.
- Learning curves:
  - Observations of skill improvement with practice follow power law (Newell 1993)

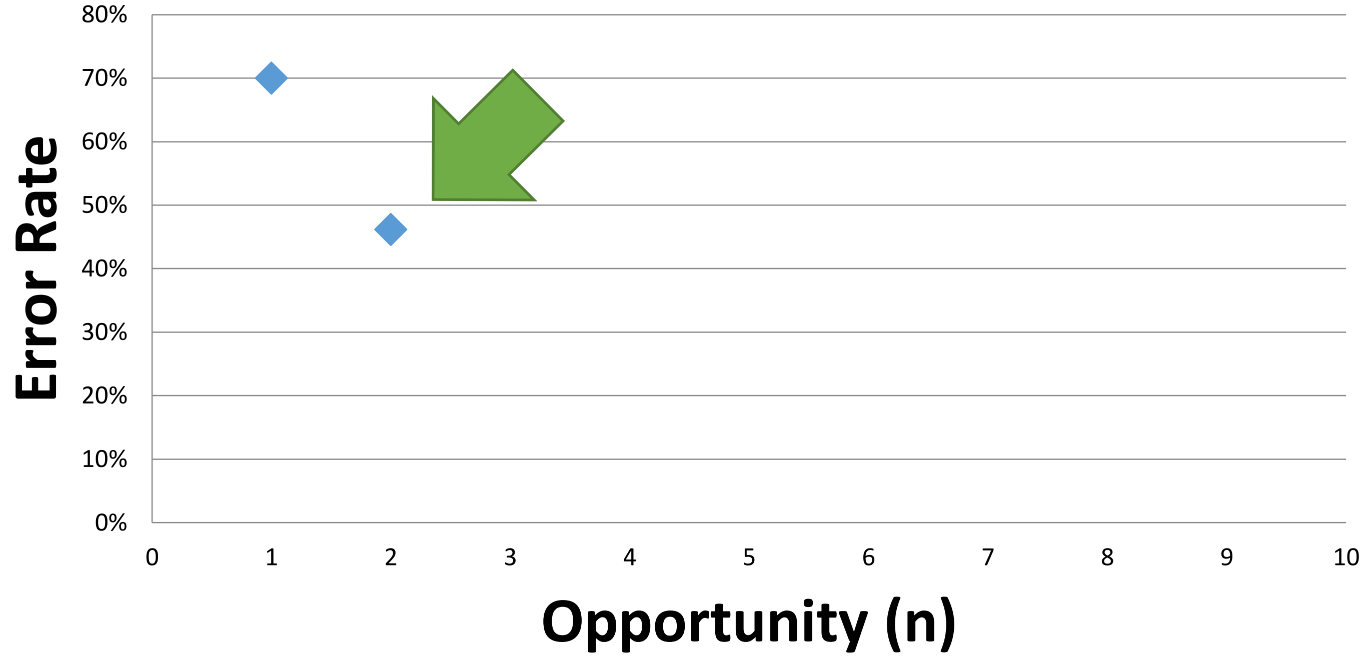
Learning Curves

- $E = Xn^a$
  - E – performance criteria
  - X – performance of first trial
  - n – describes the opportunity to practice a skill
  - a – describes the learning rate

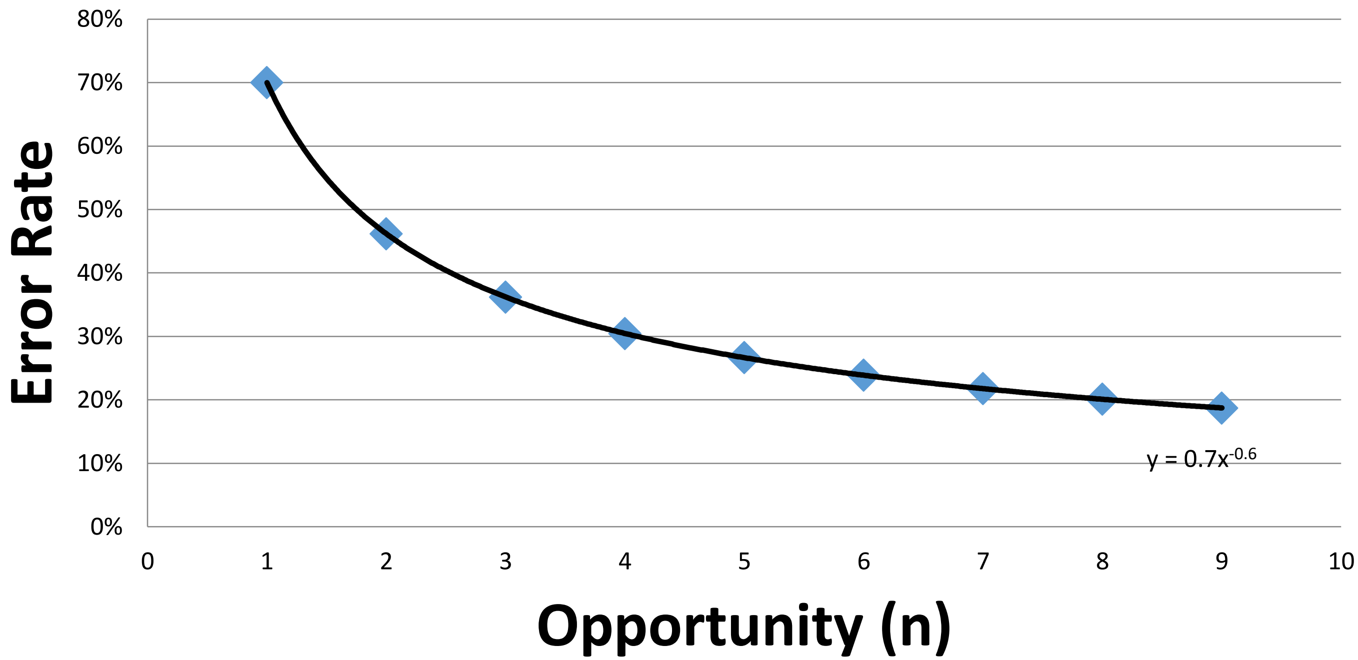
Learning Curve Example



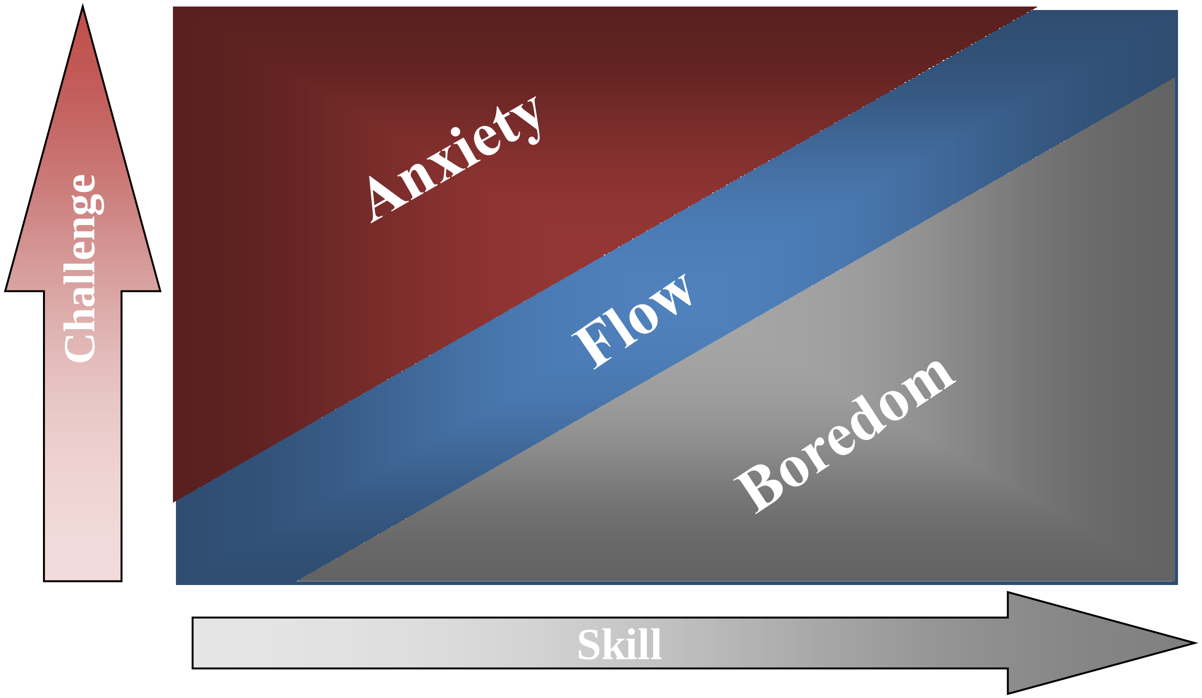
Learning Curve Example



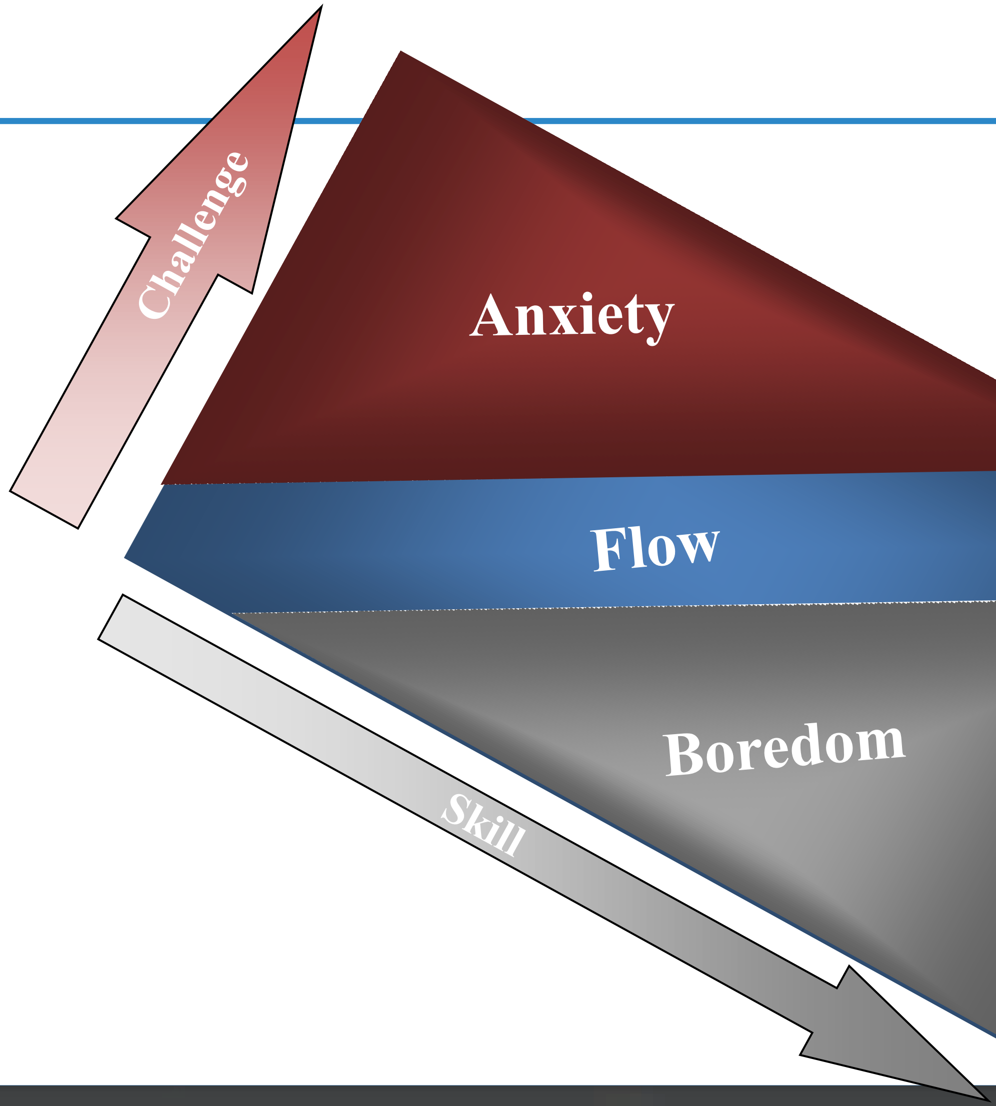
Learning Curve Example



Flow

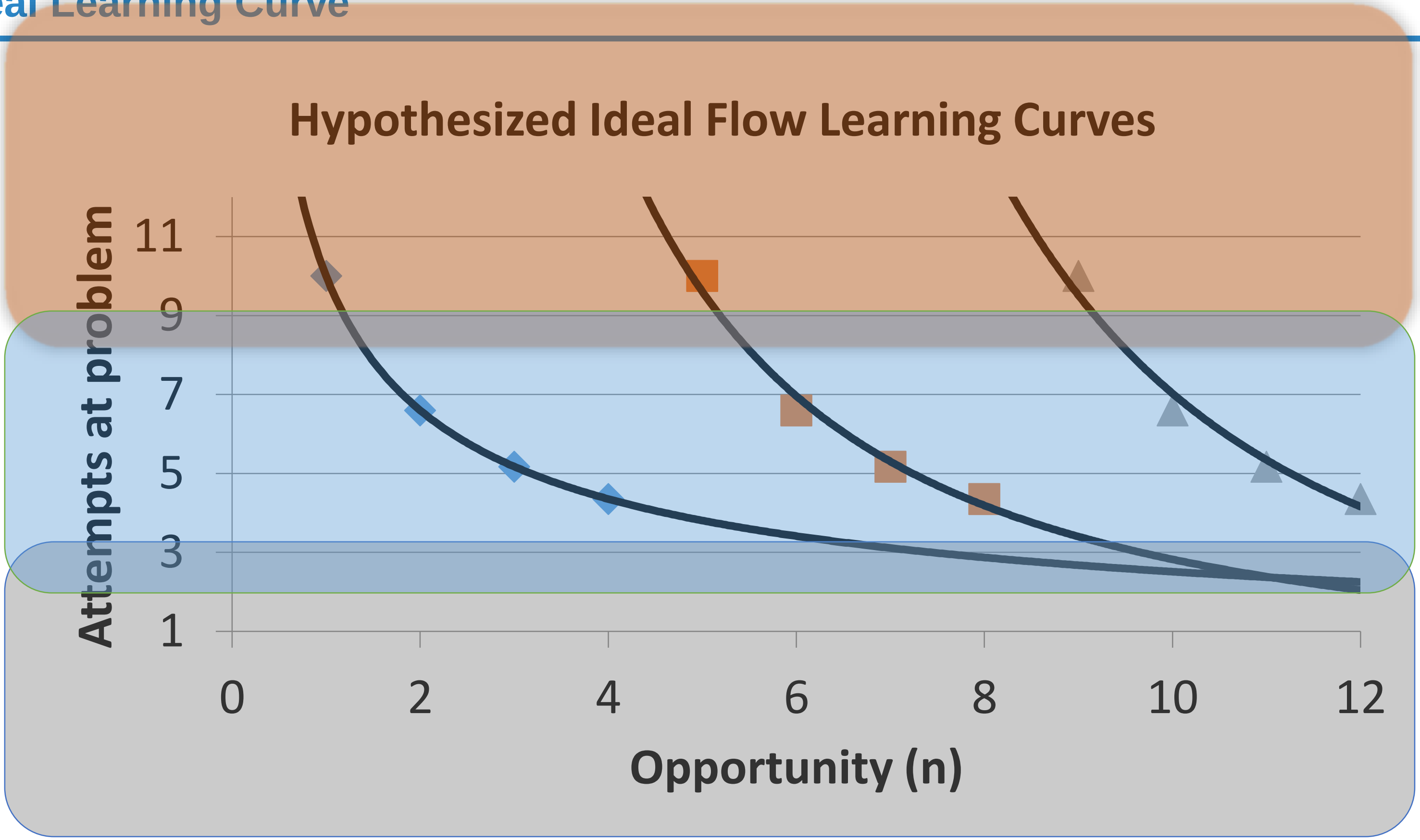


Flow

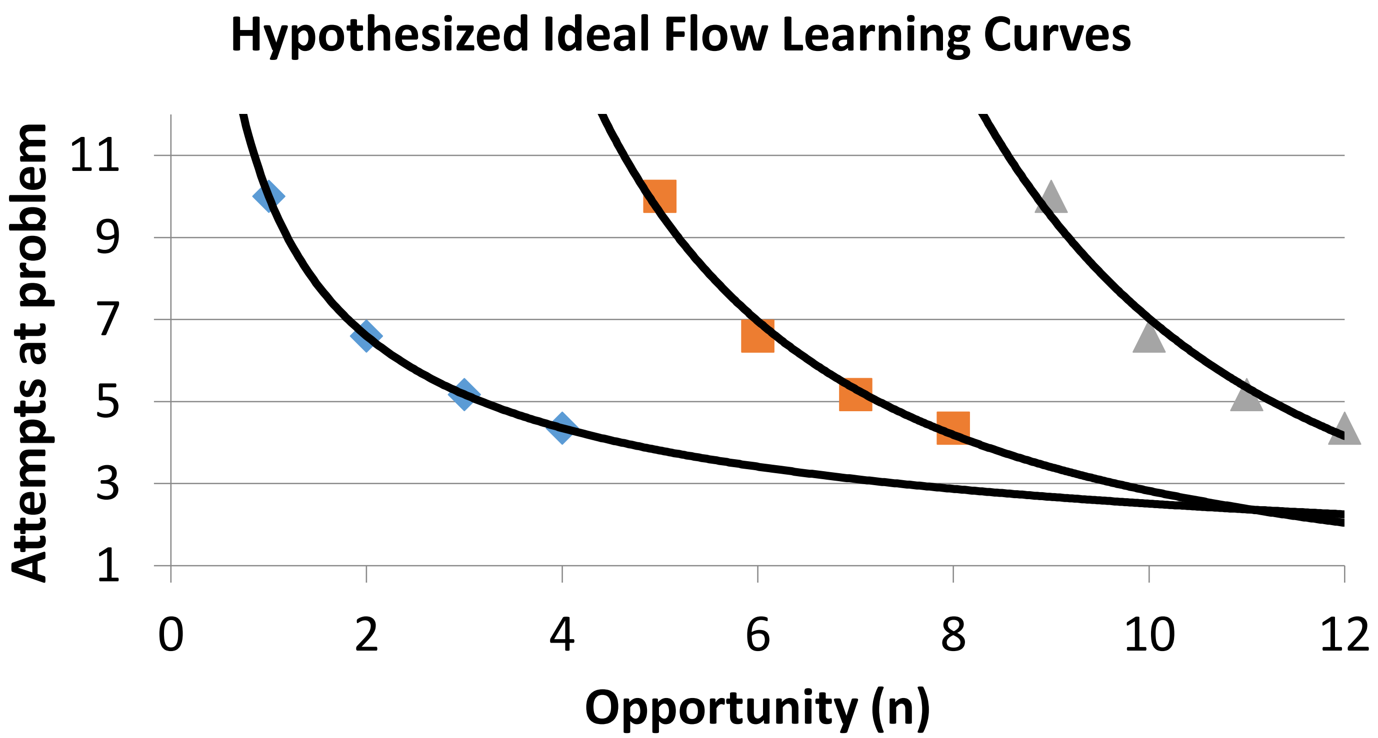


Ideal Learning Curve

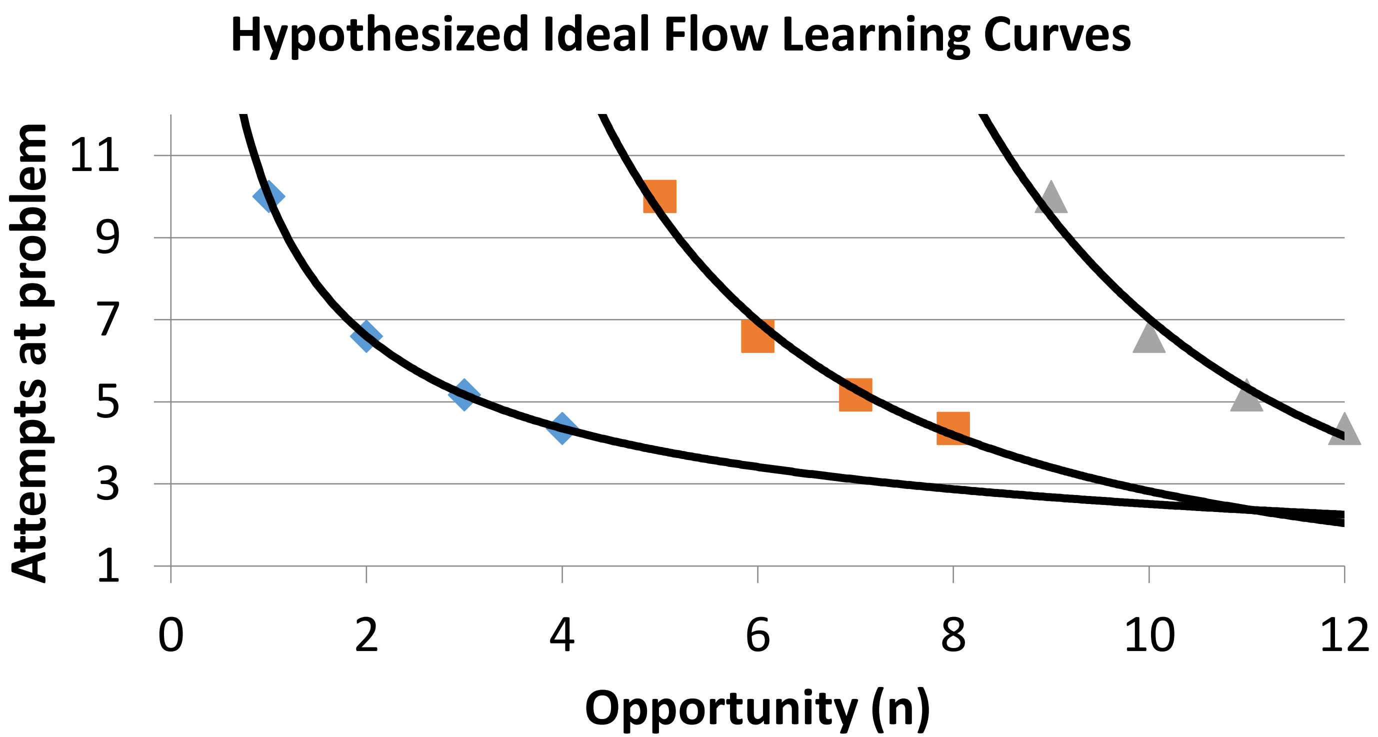
Ideal Learning Curve



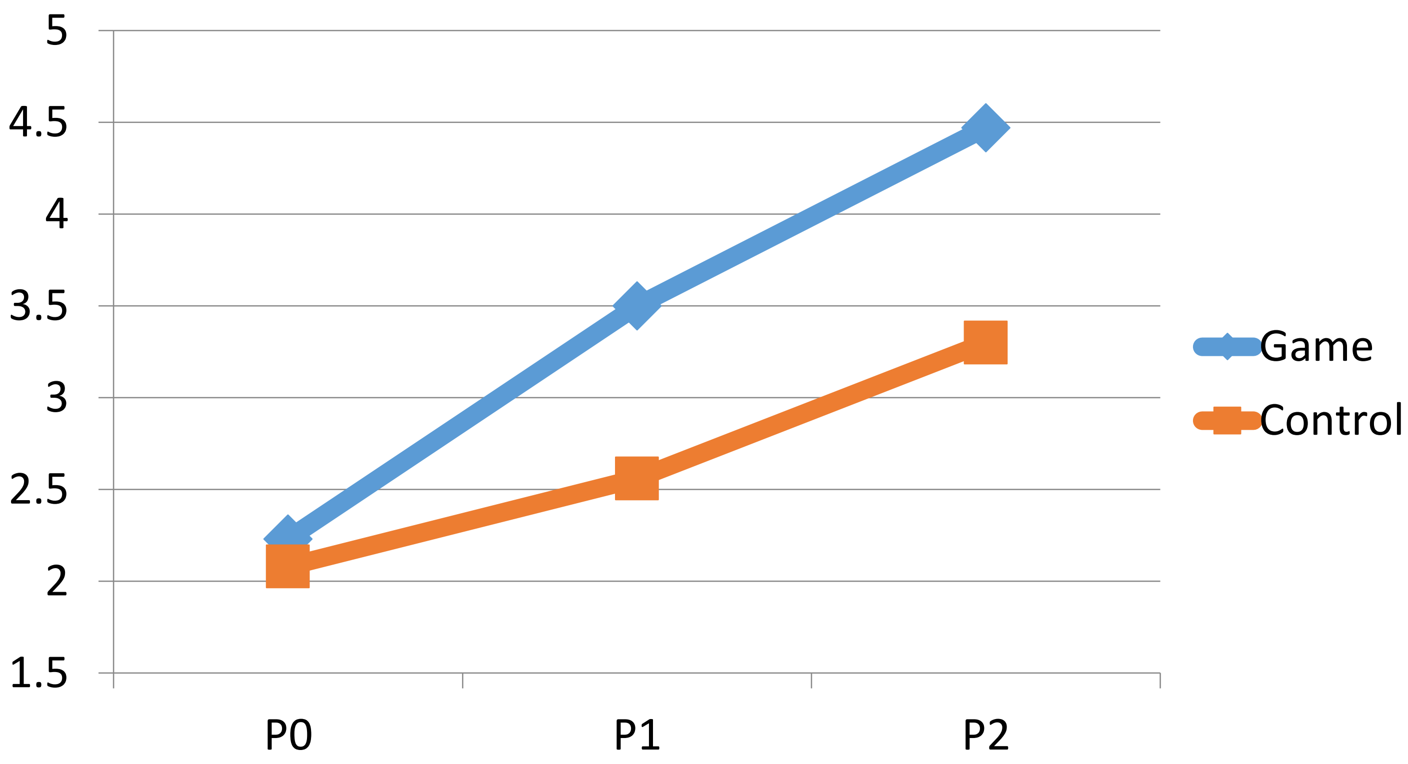
Ideal Learning Curve

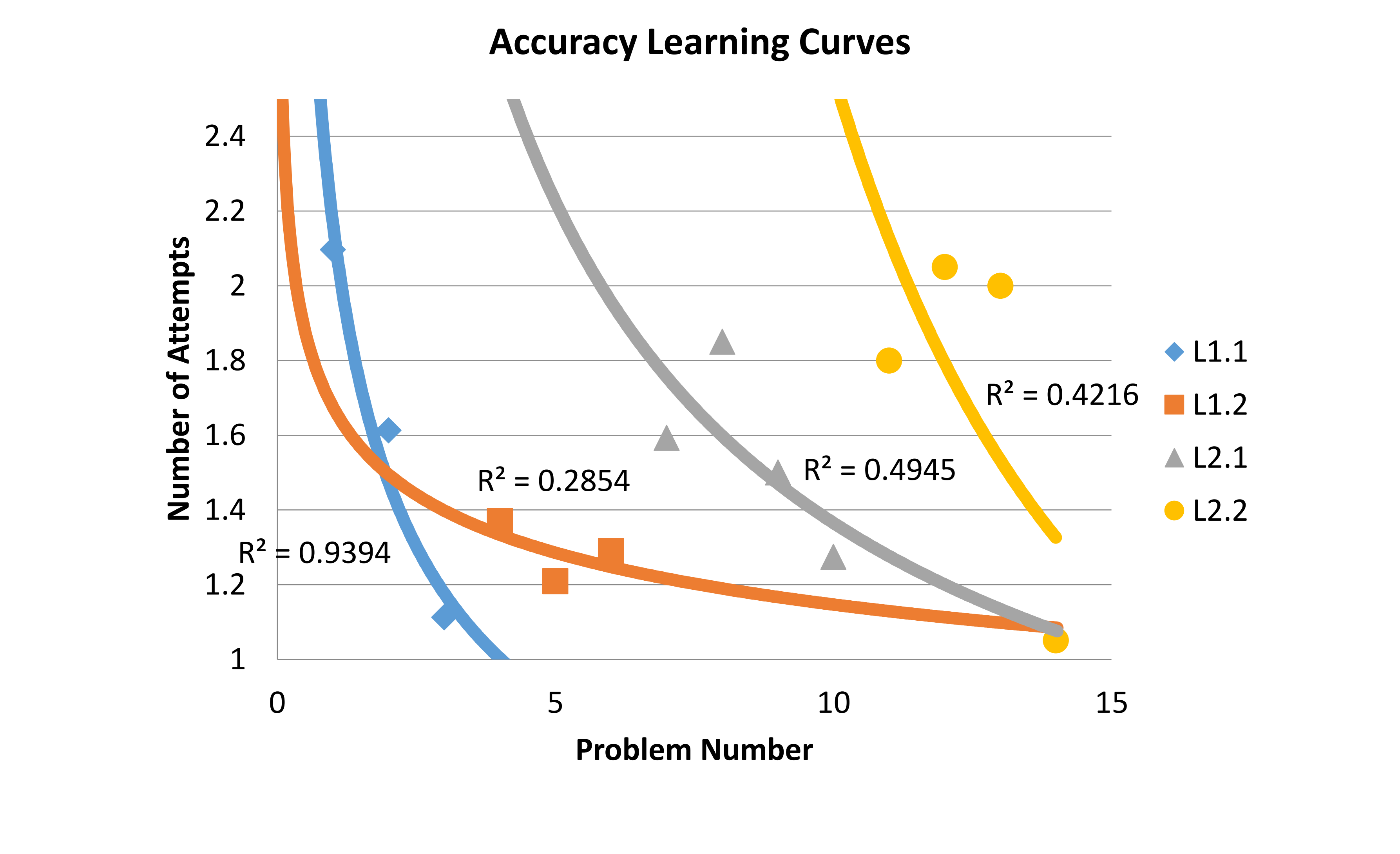


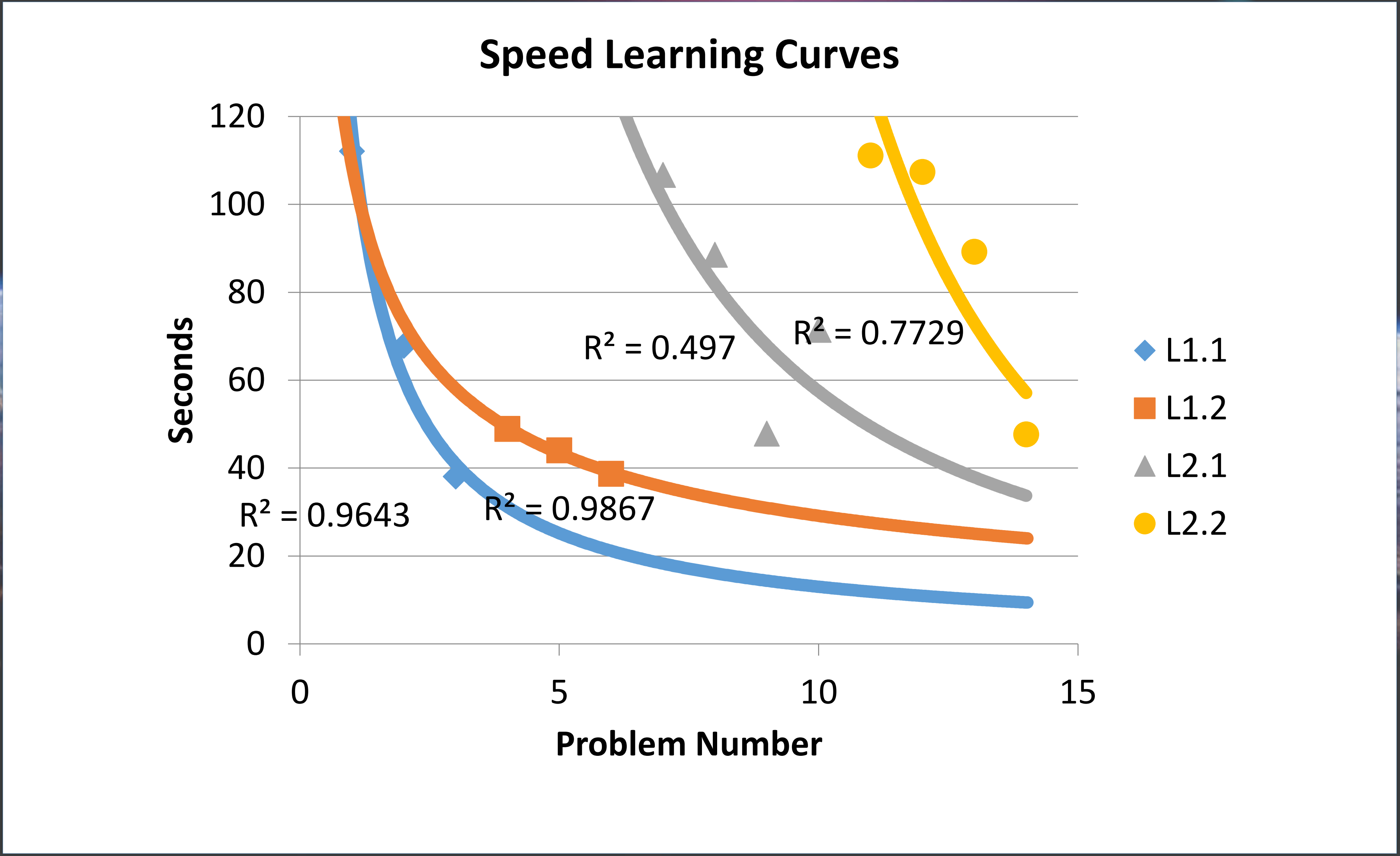
Ideal Learning Curve



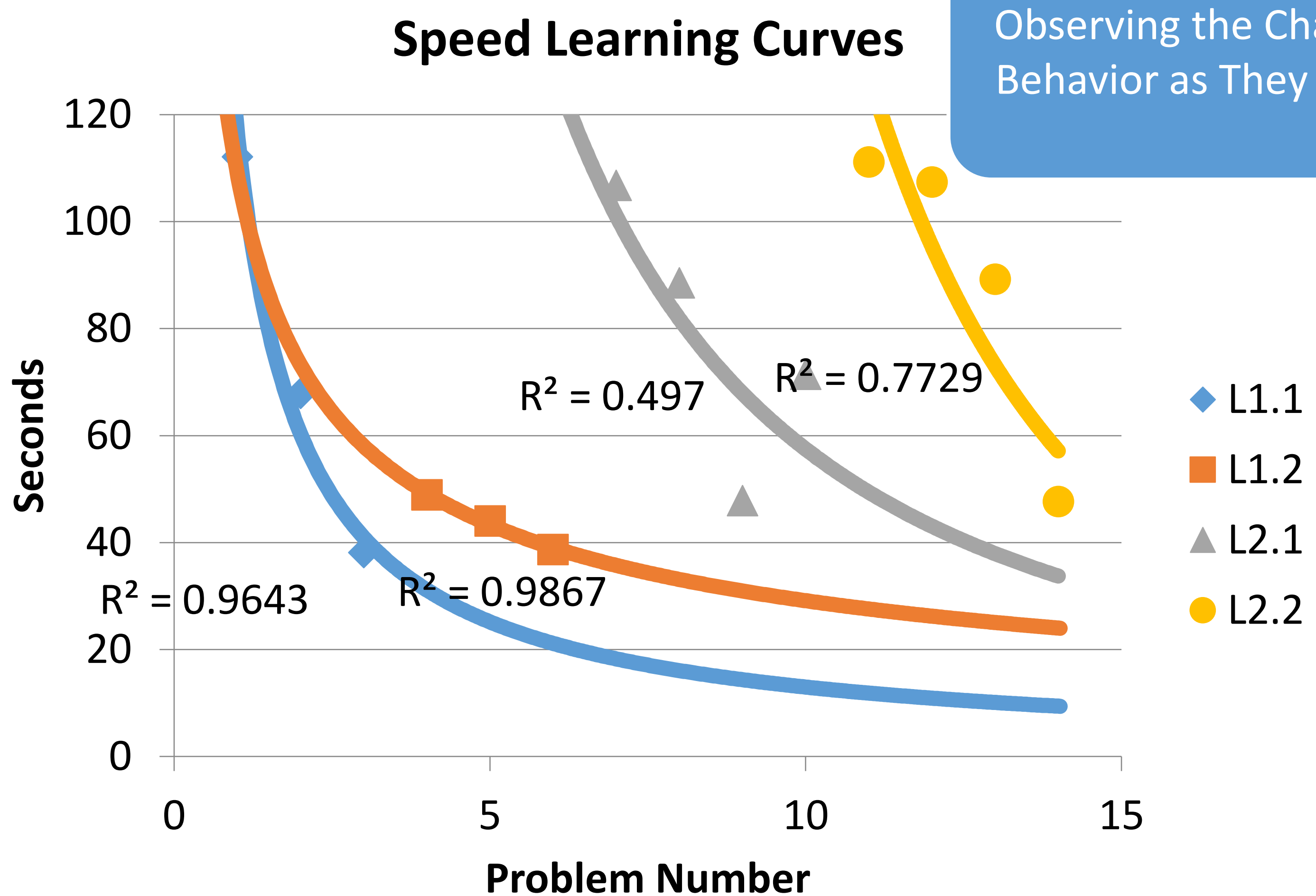
Long-term continuing treatment effect




Accuracy Learning Curves
Number of Attempts
Problem Number
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0
5
10
15
R² = 0.9394
R² = 0.2854
R² = 0.4945
R² = 0.4216
L1.1
L1.2
L2.1
L2.2


Speed Learning Curves
Seconds
120
100
80
60
40
20
0
0
5
10
15
Problem Number
R² = 0.9643
R² = 0.9867
R² = 0.497
R² = 0.7729
L1.1
L1.2
L2.1
L2.2

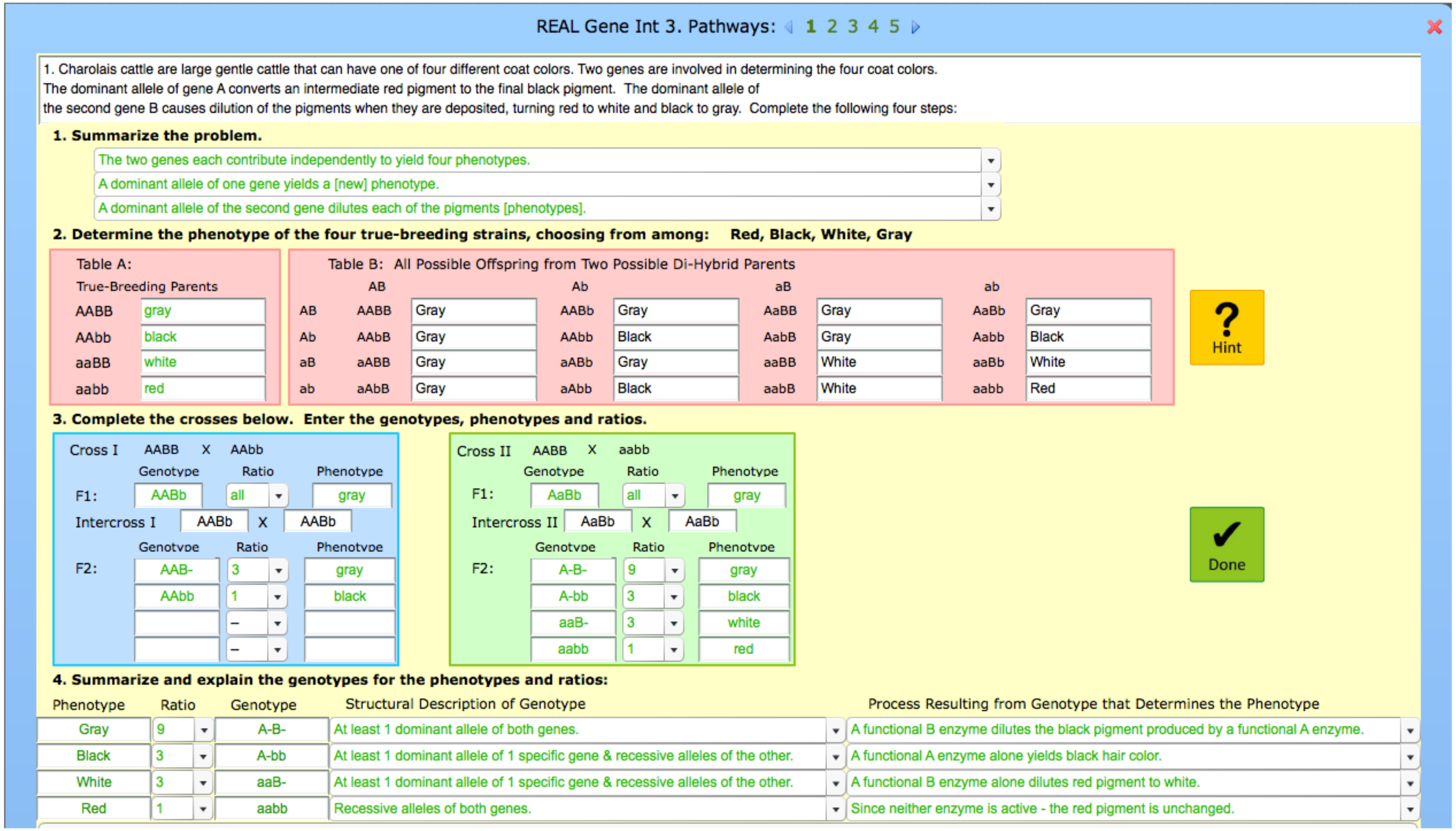
Observing the Changes in Behavior as They Happen



Learning Curve Conclusions

- Learning curves can provide insight into learning from within a game
- Analysis of learning curve data could improve game challenge balancing
- Learning curves provide a non-invasive way to demonstrate learning effects
- Study of learning curves in games can help define a research paradigm for games

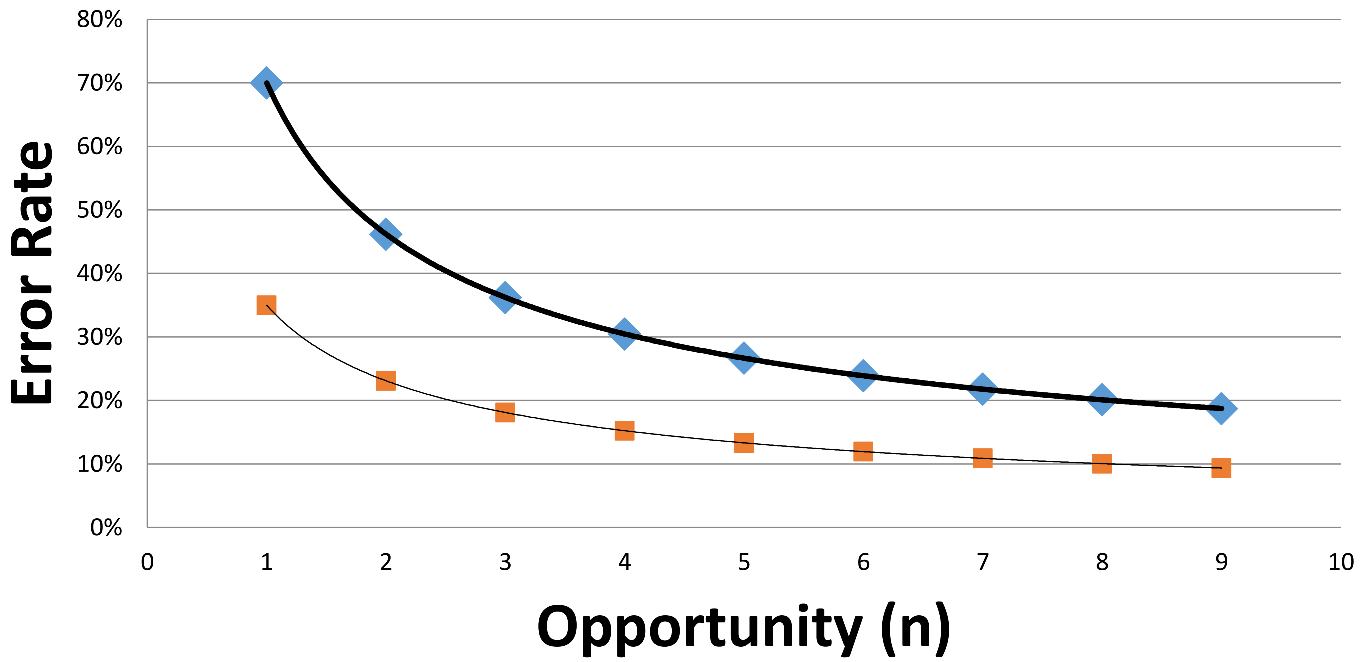
Cognitive Genetics Tutor



New Predictive Features

- Reading
  - L1 & L2 Reading Times (8)
  - L1 & L2 Pages Revisited (4)
- Pretest
  - Accuracy (3)
  - Answer Changes (2)
  - Total Time (1)
- Prev. Lessons
  - L1 & L2 IDWs (8)
  - L1 Performance Features (13)
  - L2 Performance Features (13)

Individualized Learning Curve



Learning Curves: Visualization of Learning as it Happens

Student ID	Knowledge Component	Opportunity	Correct
S1	KC1	1	0
S1	KC1	2	1
S2	KC1	1	0
S2	KC2	1	1
S3	KC3	1	0

KC	Opportunity	Probability
KC1	1	0.7
KC1	2	0.5
KC1	3	0.25
KC1	4	0.15

Learning Curves: Visualization of Learning as it Happens

Student ID	Knowledge Component	Opportunity	Correct
S1	KC1	1	0
S1	KC1	2	1
S2	KC1	1	0
S2	KC2	1	1
S3	KC3	1	0

KC	Opportunity	Probability
KC1	1	0.7
KC1	2	0.5
KC1	3	0.25
KC1	4	0.15

Can be done in Excel

Less constrained environments?

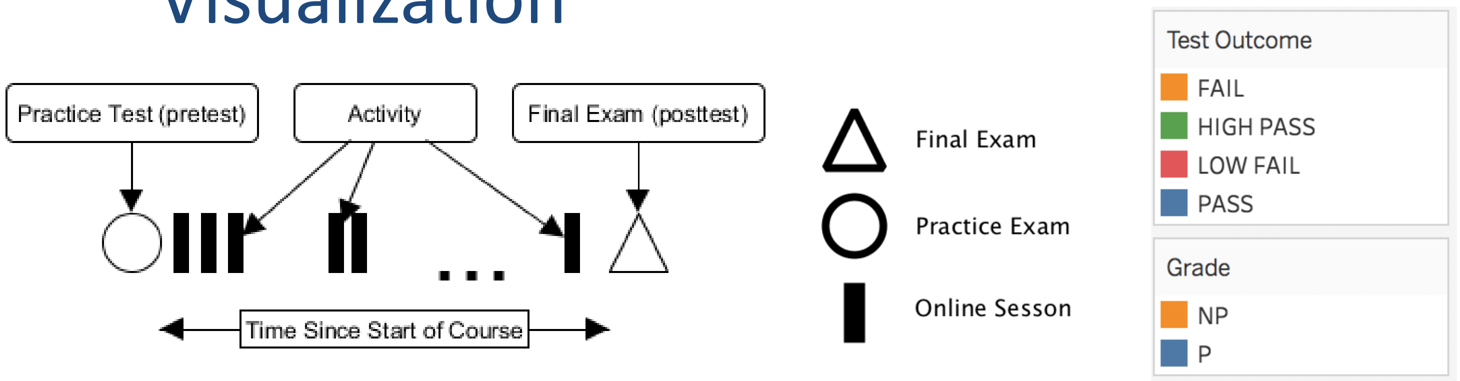
- Cognitive Modeling with BKT works well when the knowledge components are well defined

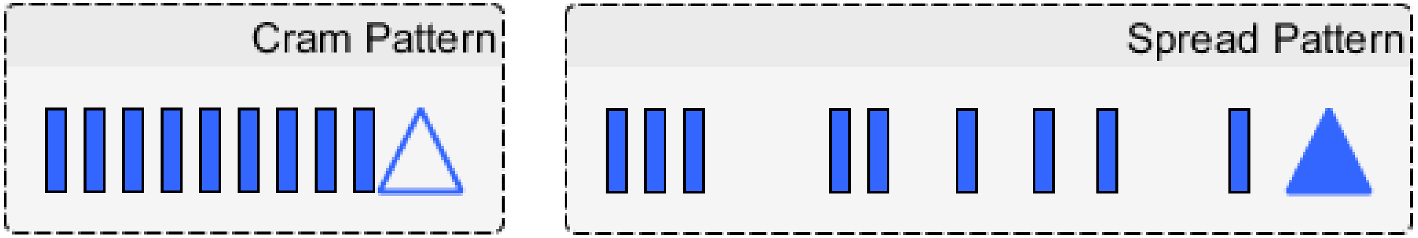
Student ID	Knowledge Component	Opportunity	Correct
S1	KC1	1	0
S1	KC1	2	1
S2	KC1	1	0
S2	KC2	1	1
S3	KC3	1	0

- What about when they are less clear?

How do students interact?

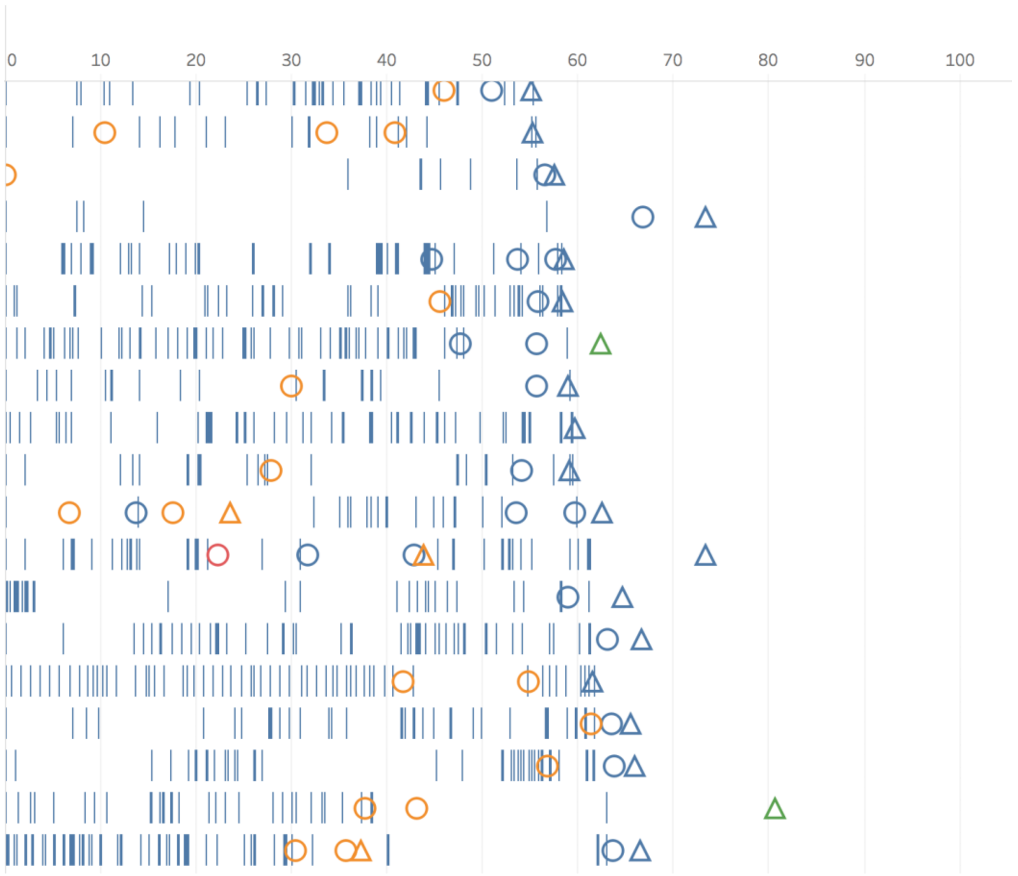
Visualization





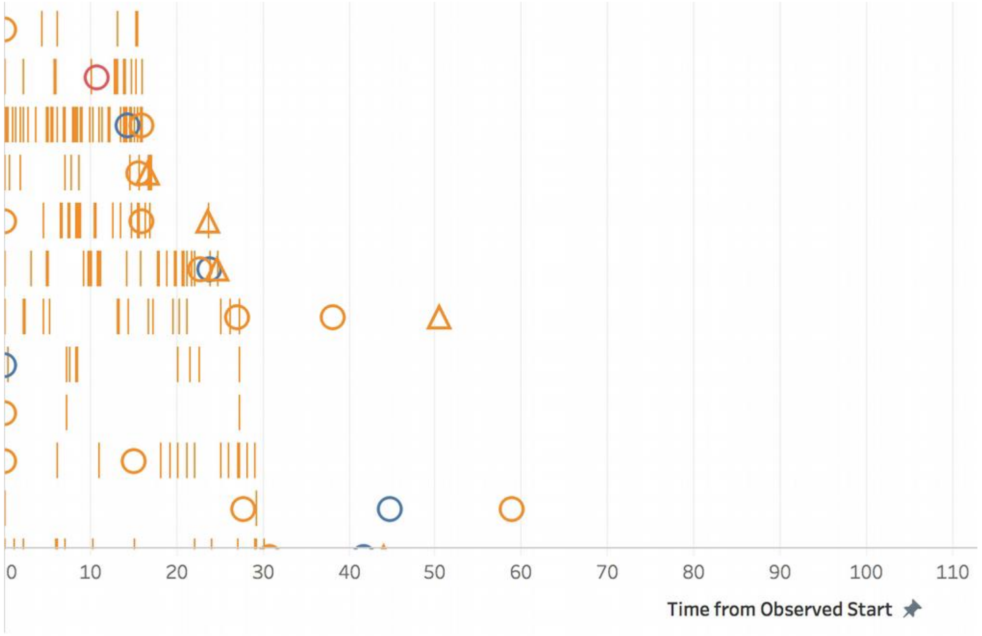
Expected Pass Behaviors

Visualization



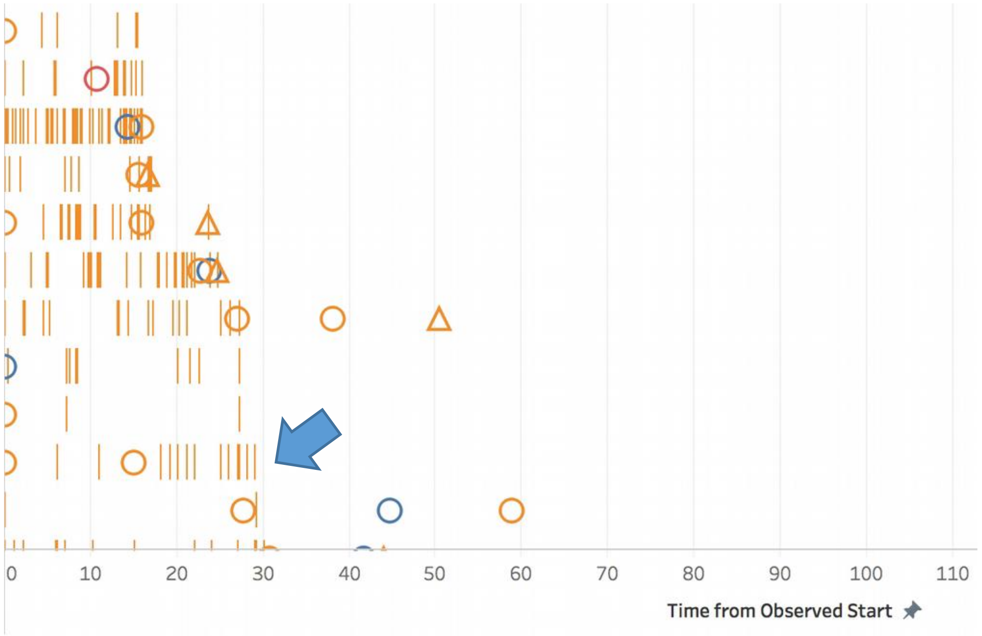
NP Behaviors

Visualization



NP Behaviors

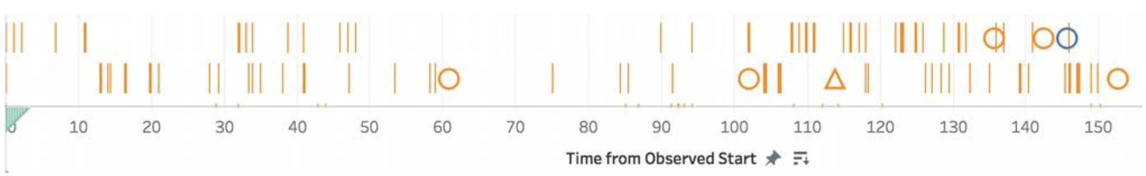
Visualization



What behaviors describe non-passing students?

Visualization

Reason	Proportion
Quit Early, low use of resources	13%
Has Activity, Complete Testing Avoidance	4%
Has Activity, Predicted Pass w/o Test	35%
Ran out of time in term (Started very late)	8%
Low Activity	22%
Not Explained	18%



Interaction Networks

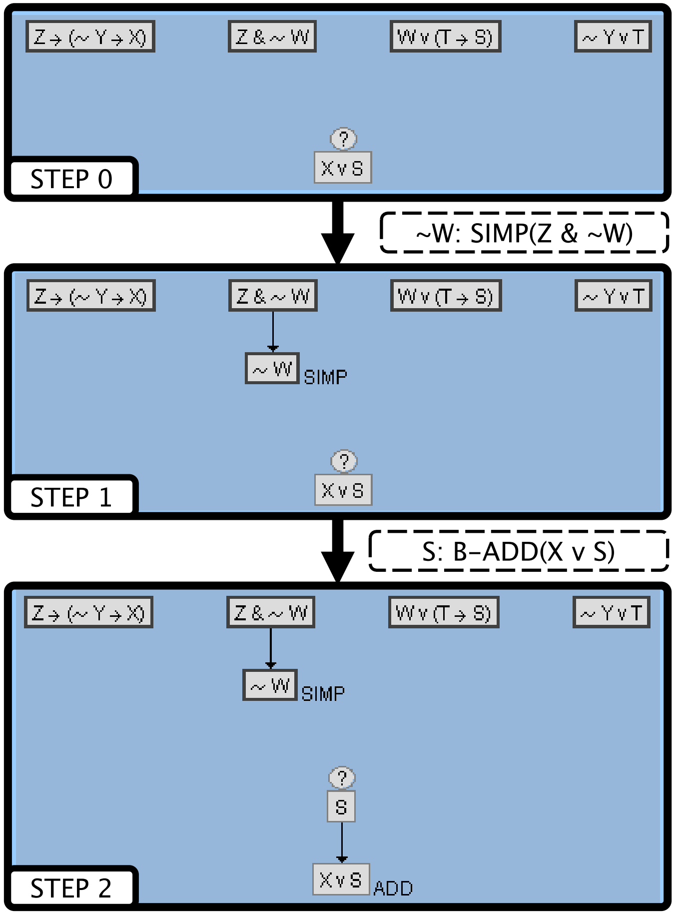
A Data-Driven Method for Understanding, Evaluating, and Changing Behavior in Educational Problem Solving

My High Level Work Flow

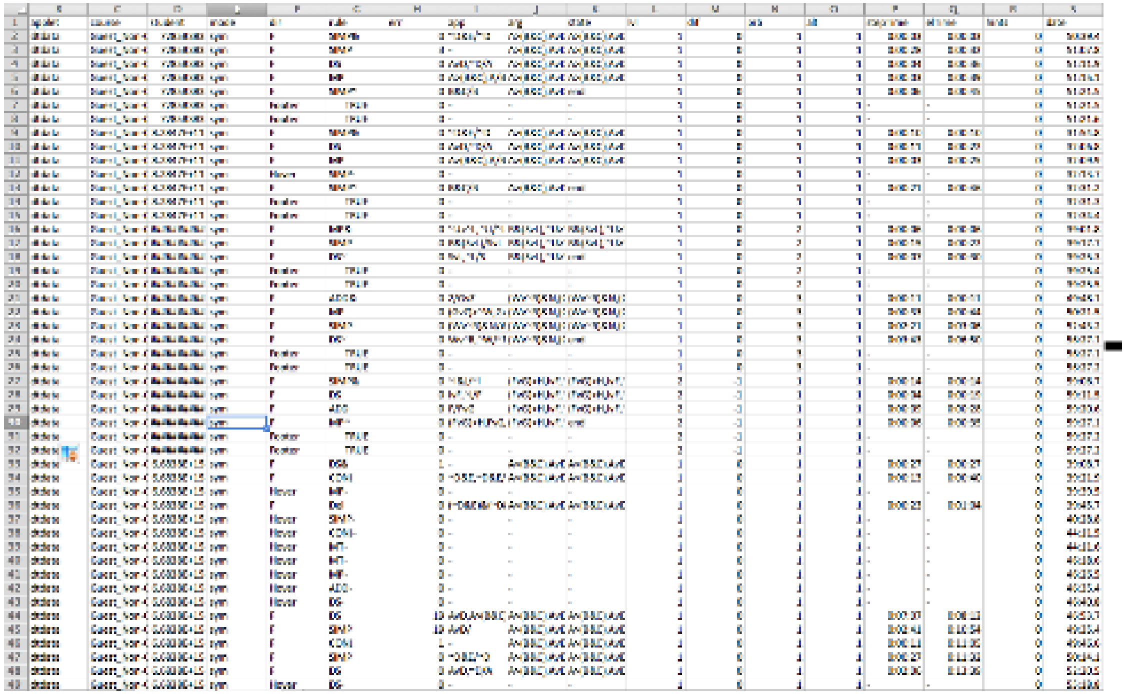
Short Version

Example Interactions

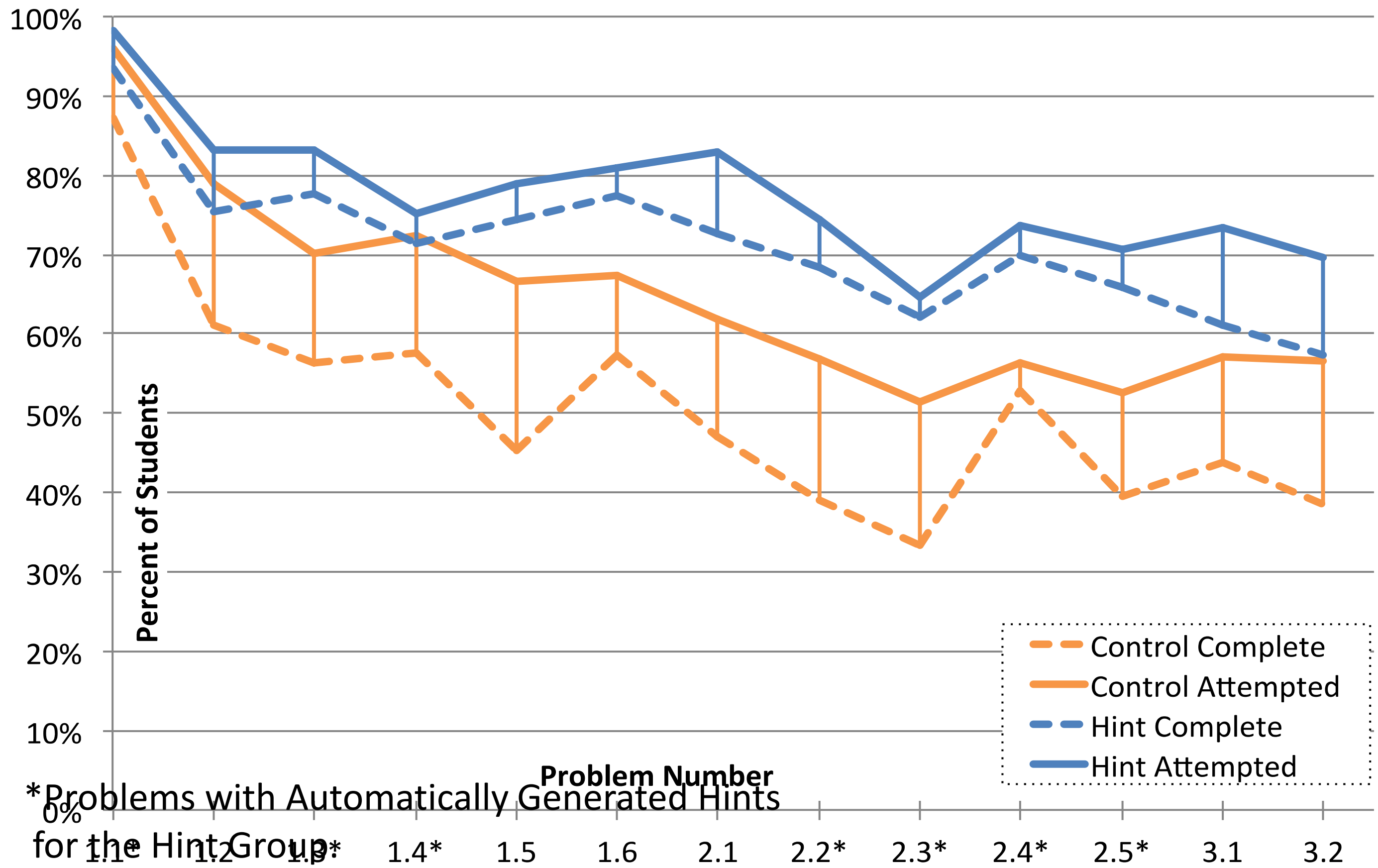
- Interaction 1: User moves from Step 0 to Step 1 via the SIMP rule
- Interaction 2: User moves from Step 1 to Step 2 via the B-ADD rule



Interaction Logs

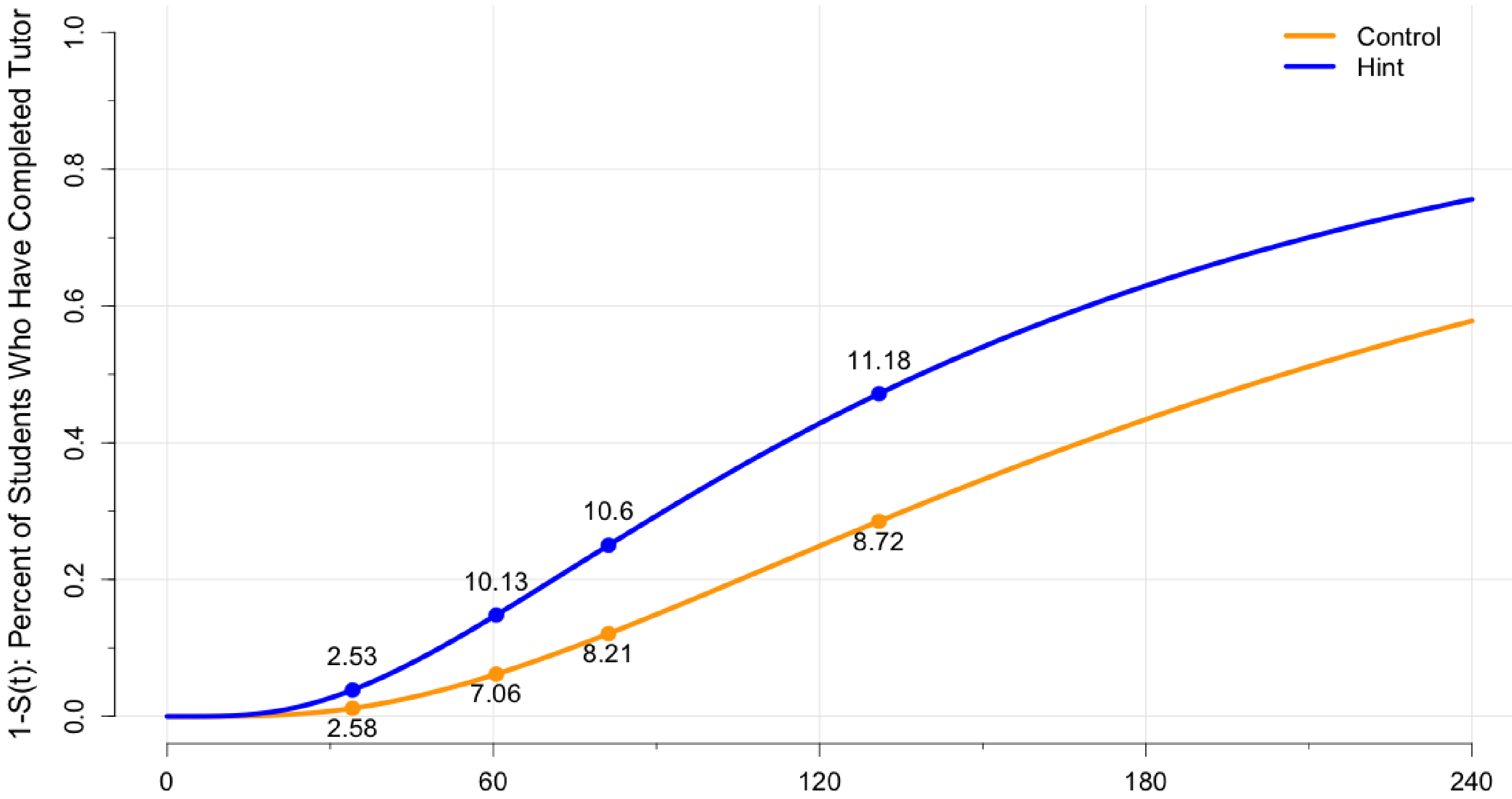


Performance Analysis

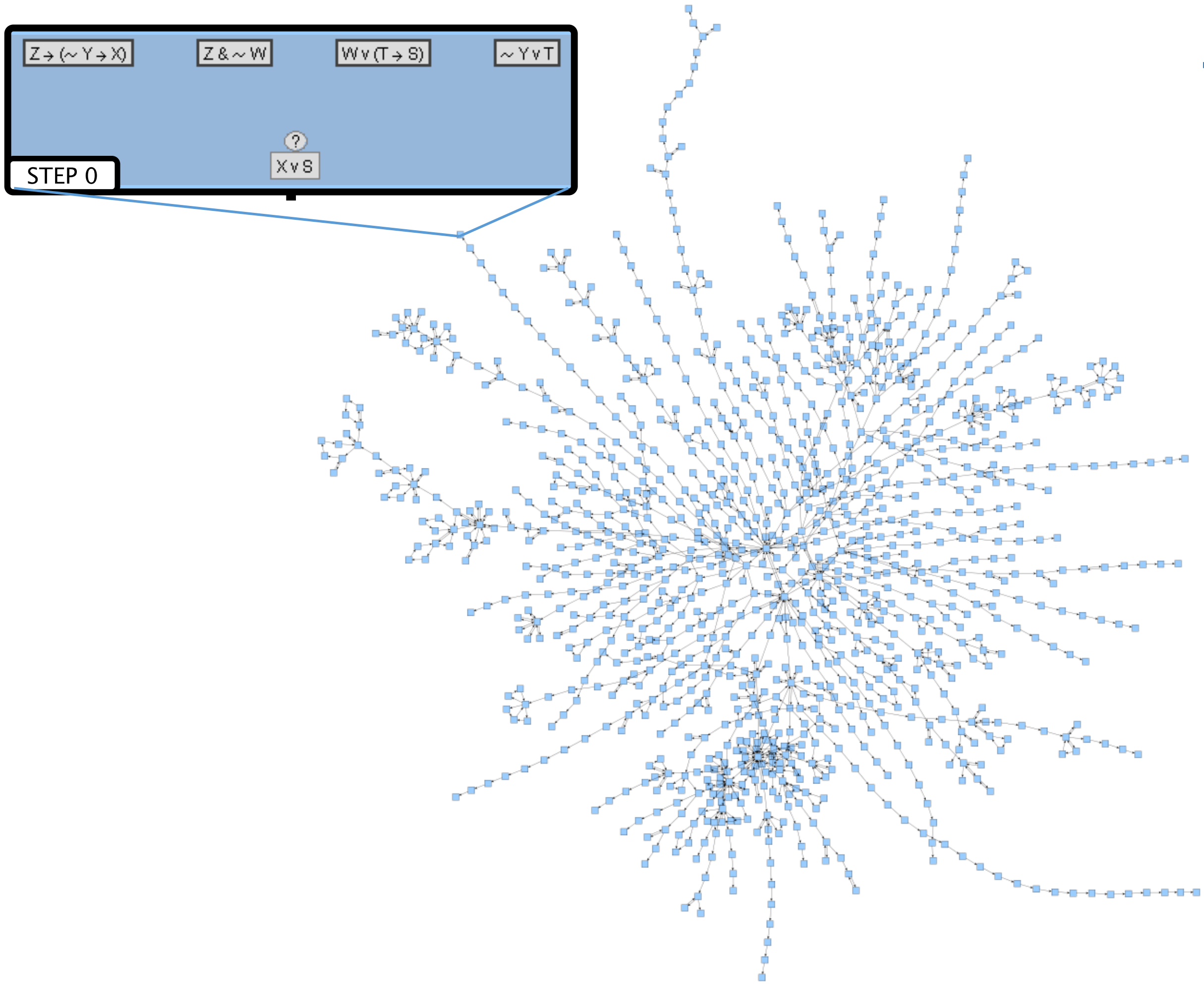


*Problems with Automatically Generated Hints for the Hint Group.

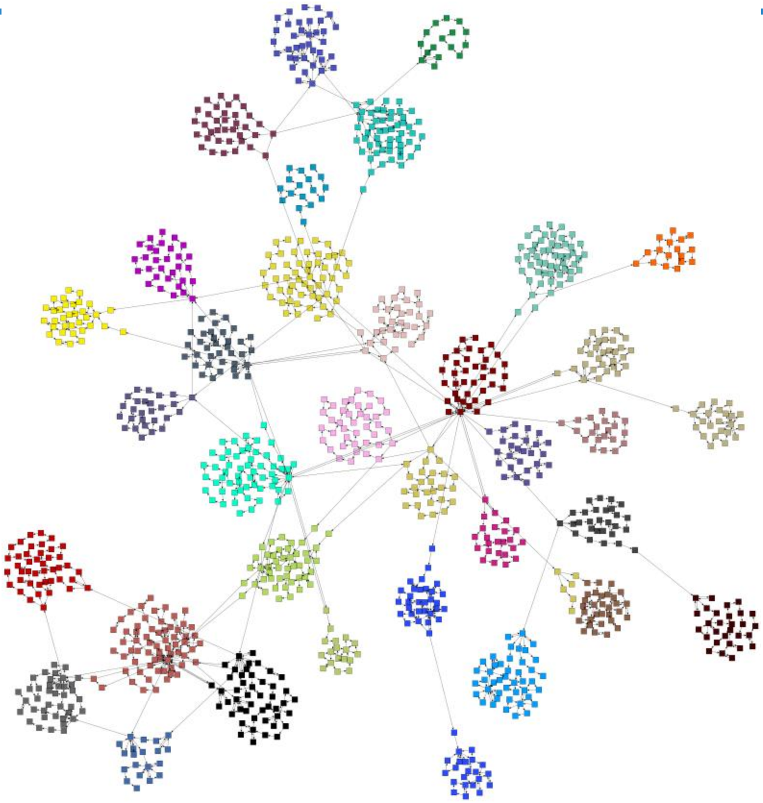
Transactional Analysis



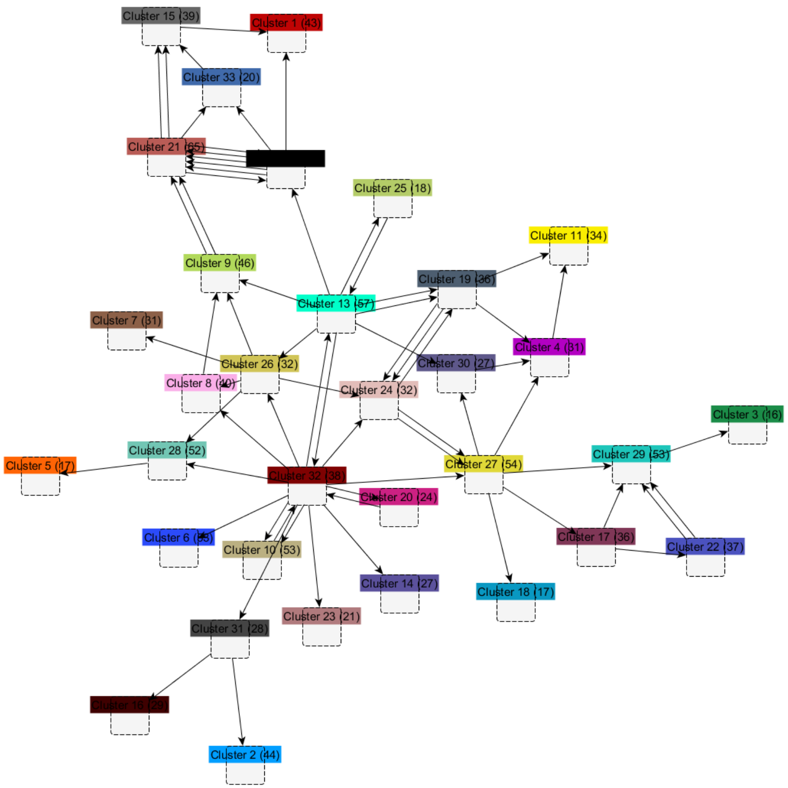
Interaction Network



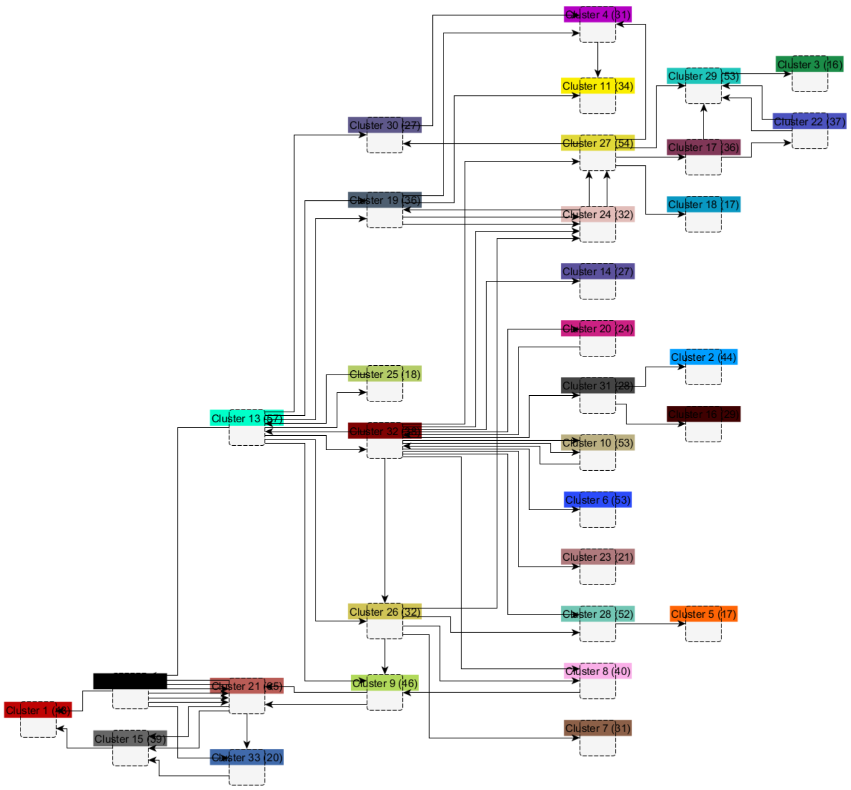
Edge-Betweenness Regions



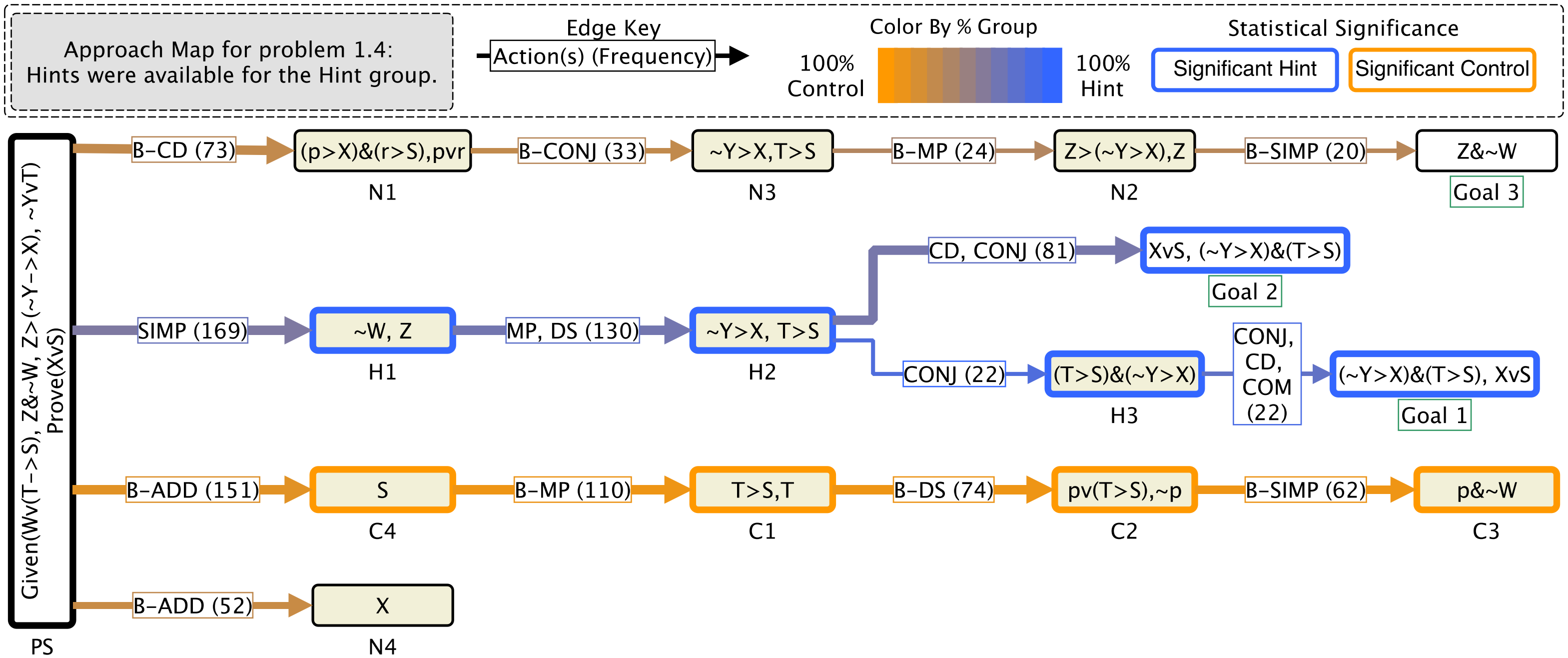
Reduction to Regions



Different Layout

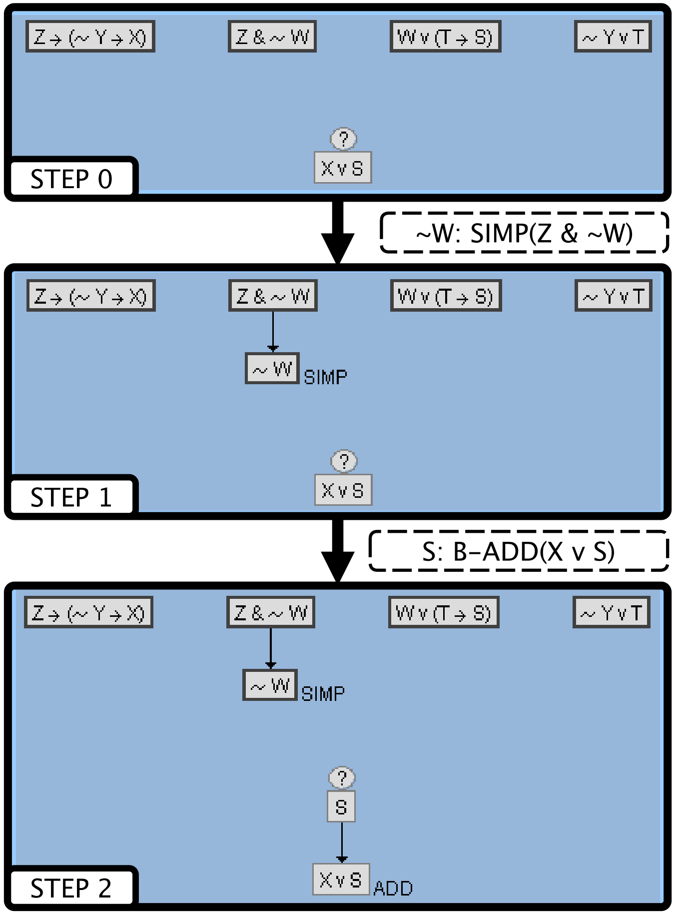


Approach Maps

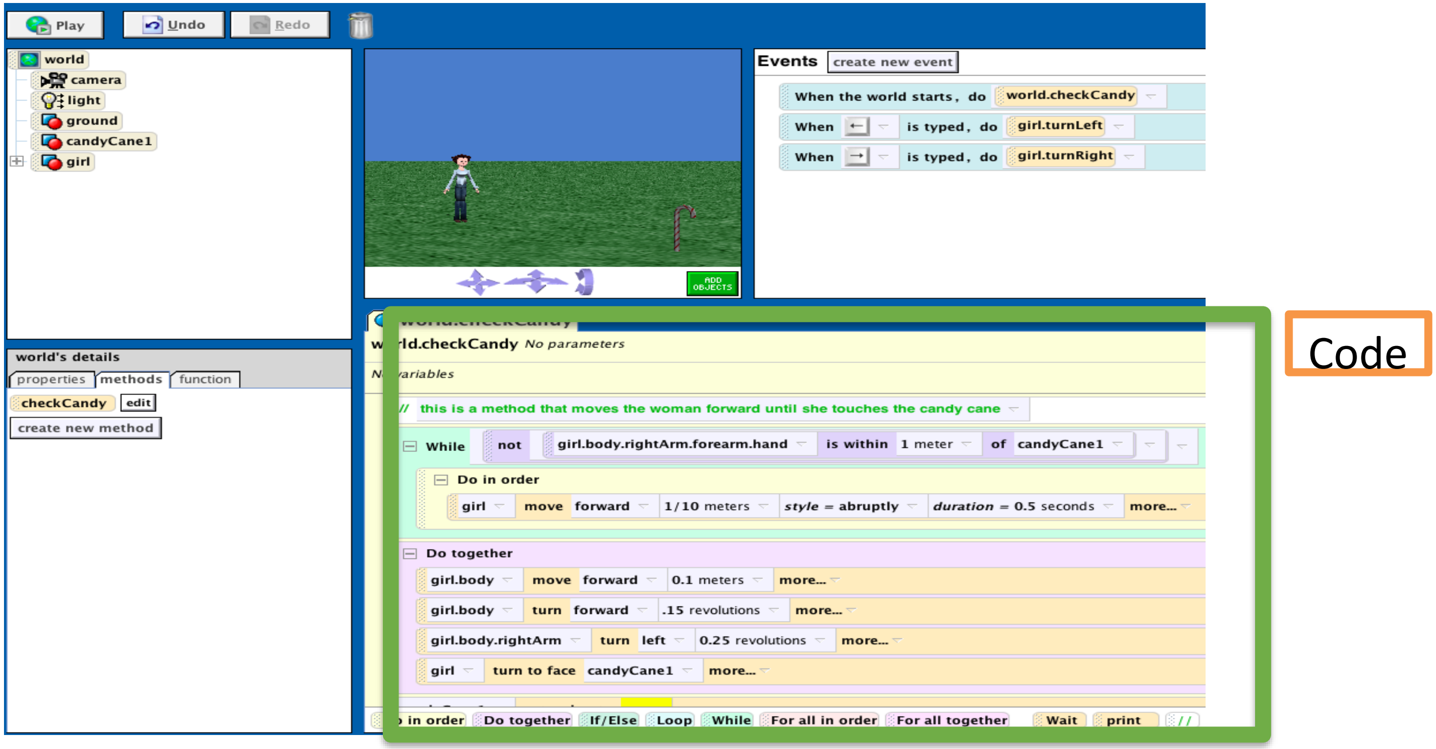


Open-ended Environments

- Body of my dissertation research
- Focus is not on KCs
- Use State-to-State data to produce ITS like feedback

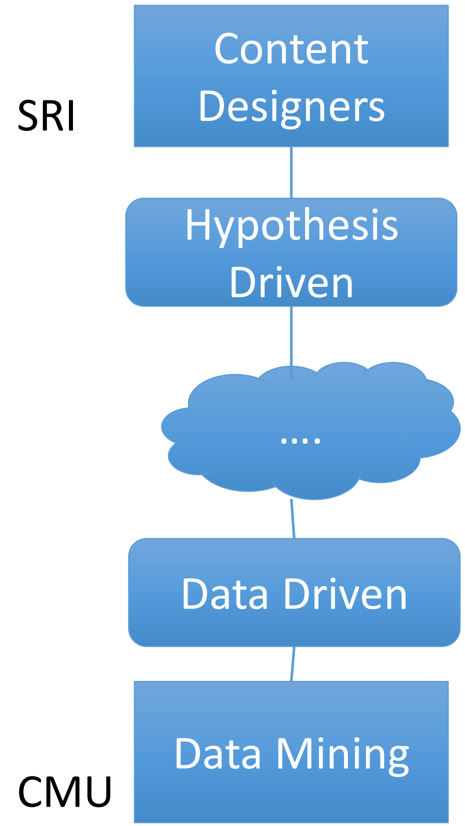


Alice Environment

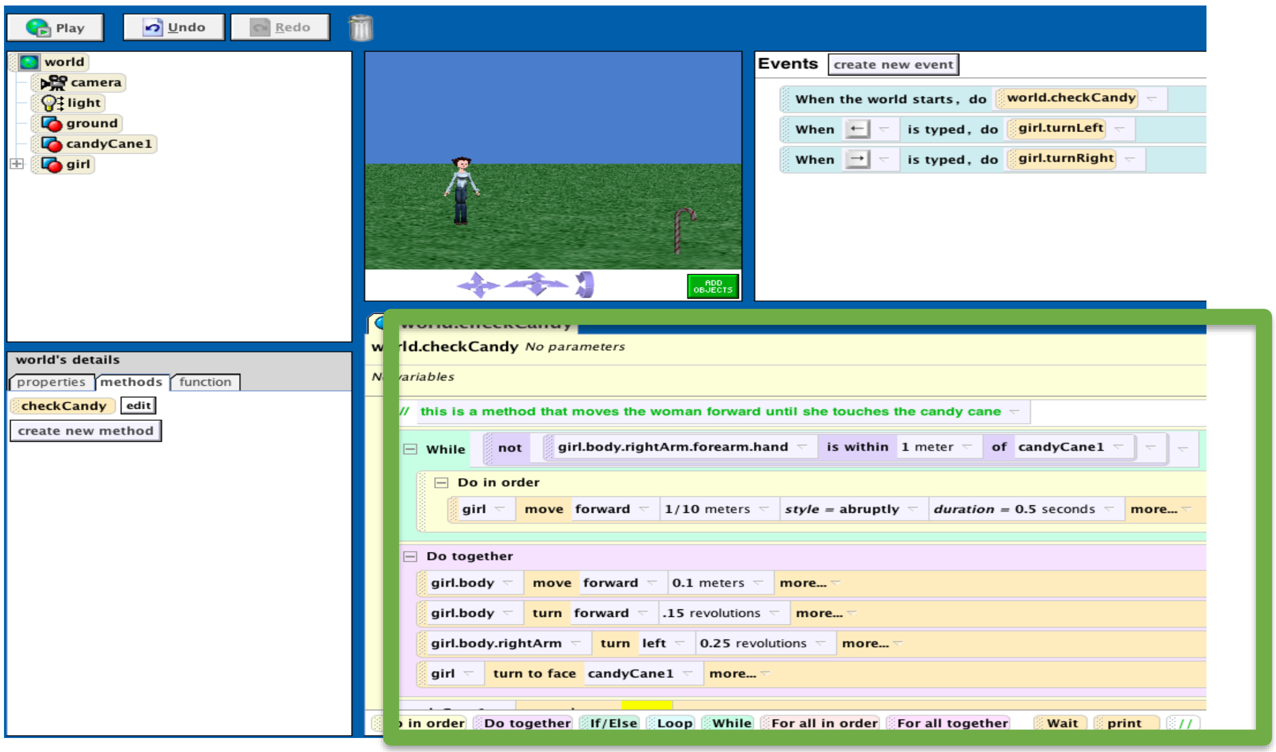


Data-Driven Learning Analytics for Assessing Computational Thinking

- Collaboration with SRI
- Attempting to derive meaningful CT concepts from low-level data

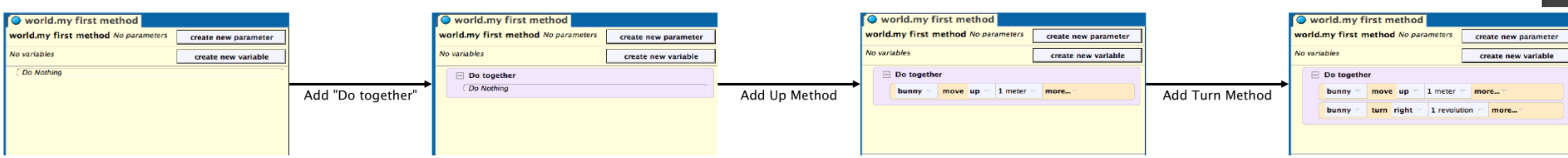


Alice Environment

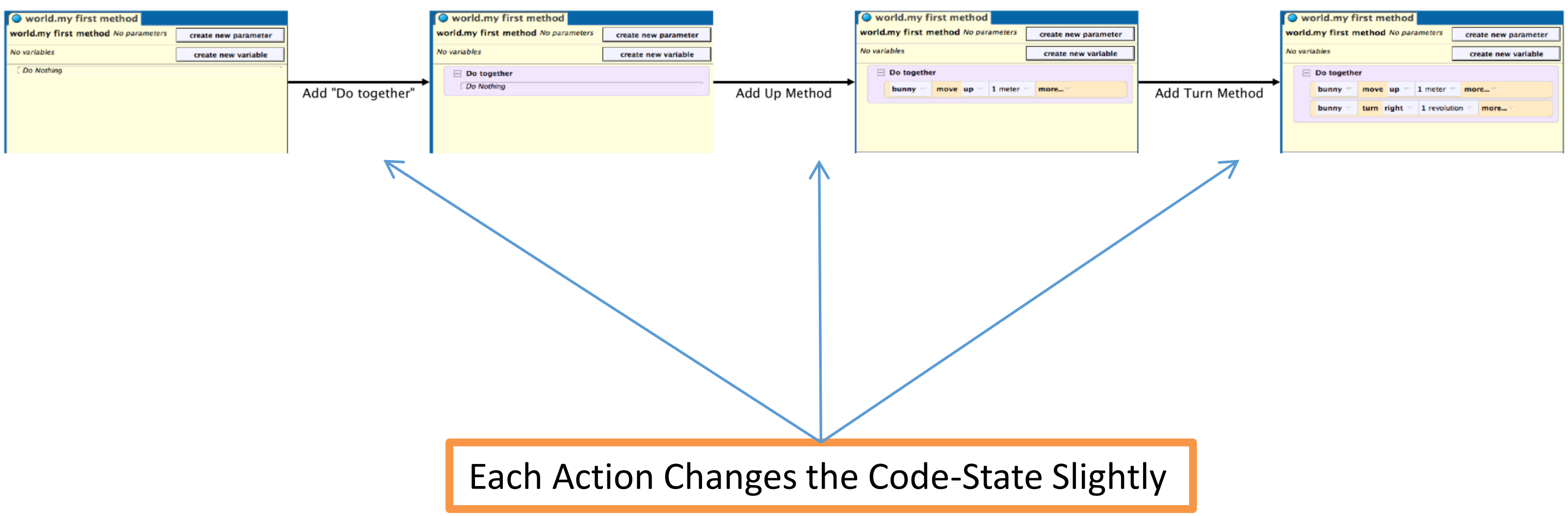


Code

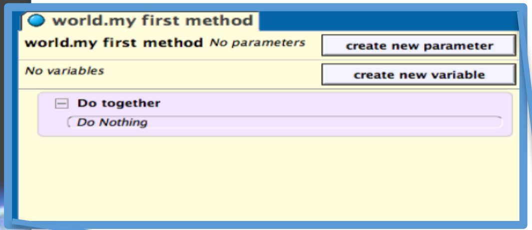
Track the Code-State After Each Student-Alice Interaction

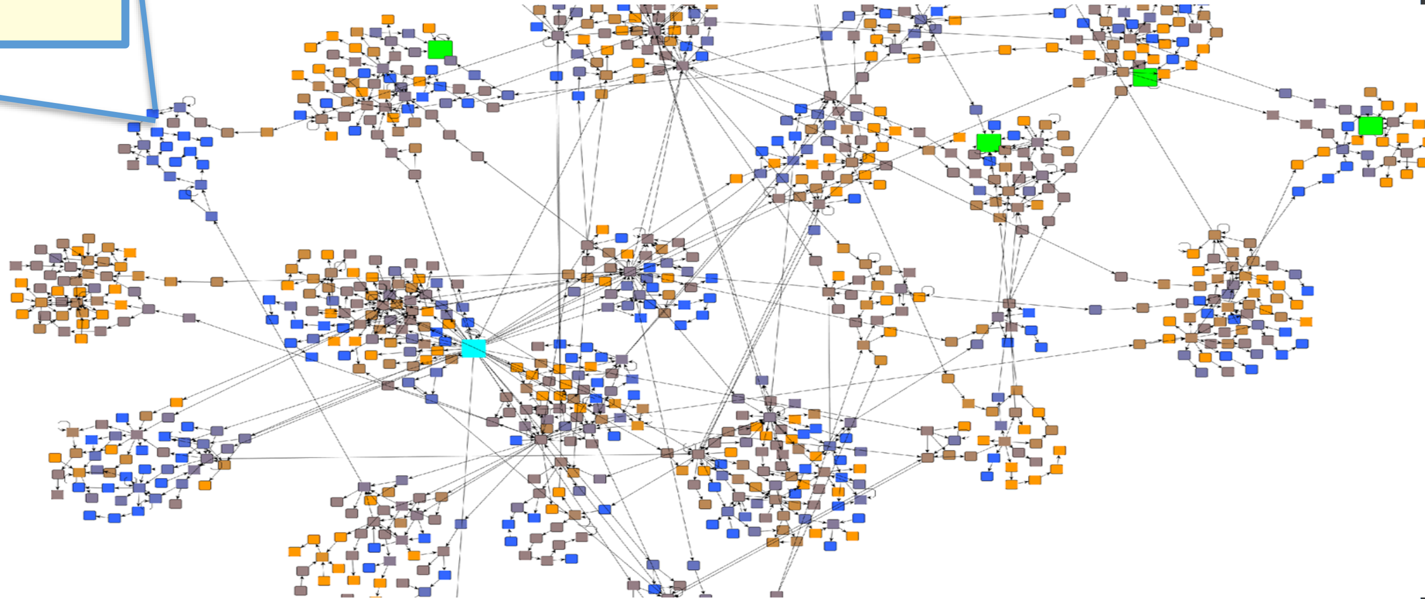


Track the Code-State After Each Student-Alice Interaction



Each Code-State is Represented as a Vertex in the Network





Data from Fairy Assessment Task (From ETR)

For **~315** students

- Final .a2w files and corresponding .txt log files
- Manually graded projects per rubrics

Task1

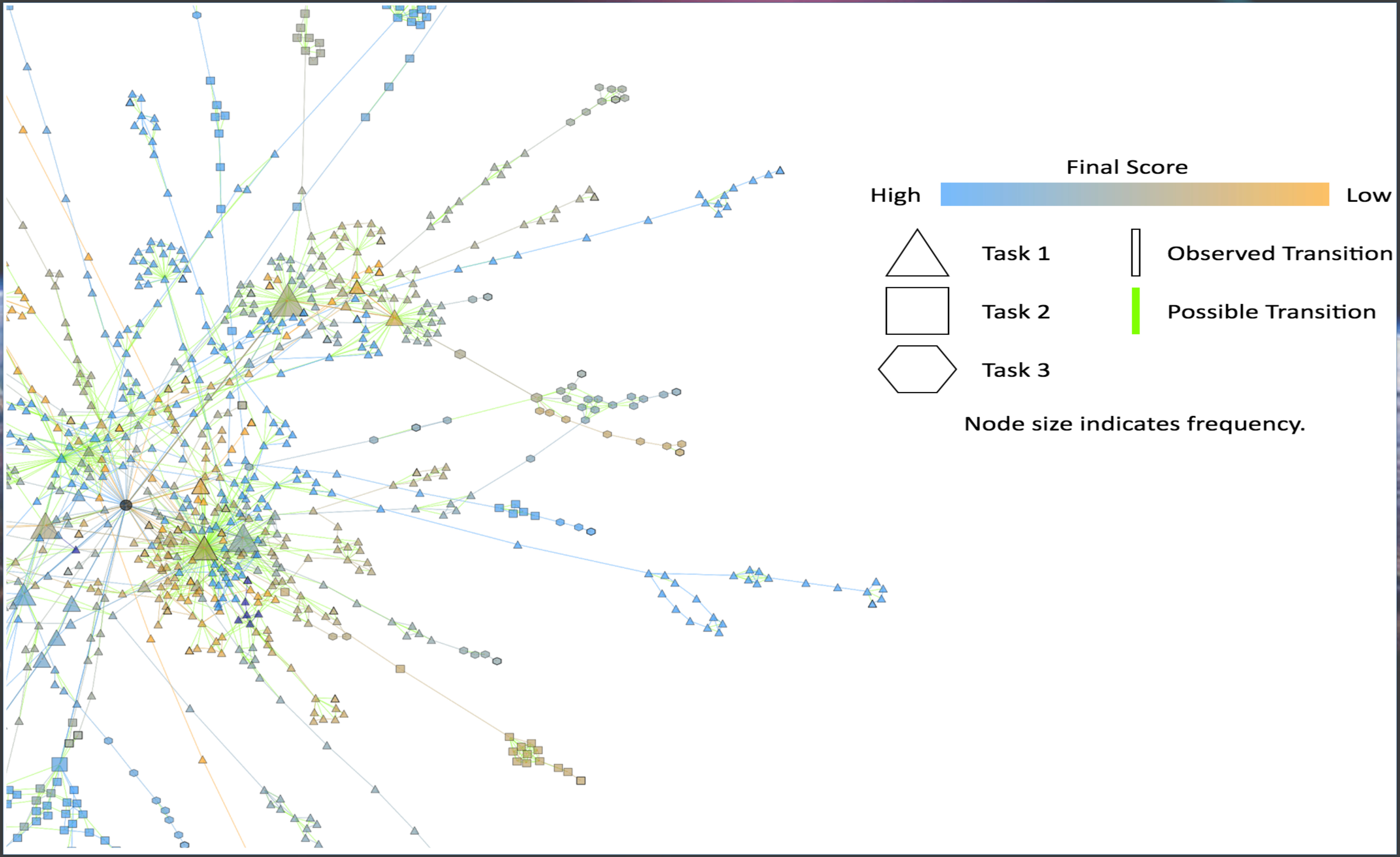
	facing		doTogether					duration			
StudentID	FacingDouglas	SomeTurn	doTogether w/ walk	doTogether no walk	walk, turn, walk	turn next to walk	walk not to forest	duration = (3-5, 11-13)	duration shows	Other	Total Task 1
3003	4				4						8
3004	4		4					2			10
3005	4			2							6
3005	urns to face LeafFlameGlimmer when LeafFlameGlimmer says "Help me Halo										
3007	4				4						8
3008	4		4					2			10
3009	4	2	4	2	4	2	-1	2	1		
able to open											
3010	4		4					2			10
3011	4		4					2			10
3013	4		4					2			10
3014	4		4					2			10
3014					duration is 15						
3015	4		4								8
3017	4					2					6
3019	4		4					2			10
3020	4		4								8

Task 2

	Event					No Event			
StudentID	returnToSize fixed	resize 2 used	resize (not 2) used	no automatic change	UP arrow connected	resize 2 or returnToSize fixed and	returnToSize fixed, not called	Other	Total Task 2
3003	6			2	2				10
3004									0
3005	6			2	2				10
3007									0
3008	6			2	2				10
3009	6	6	2	2	2	6	4		
3010	6			2	2				10
3011		6		2	2				10
3013									0
3014									0
3015									0
3017									0
3019									0
3020	6			2	2				10

Task 3

	event	fly			connect			
StudentID	LeafFlame mouse click	flyAway 1 or greater	flyAway less than 1	move/ walk out	event connected	Other	Total Task 3	Total
3003	4	4			2		10	28
3004	4	4			2		10	20
3005	4	4			2		10	26
3007							0	8
3008	4			2	2		8	28
	ed fly which moves LeafFlame out of the forest							
3009	4	4	3	2	2			
3010	4	4			2		10	30
3011	4			2	2		8	28
	Moves to the left of Halo							
3013	4		3		2		9	19
	lame walks away and the flyAway does not gc							
3014				2			2	12
	LeafFlame walks out of the forest							
3015							0	8
3017							0	6
3019	4	4			2		10	20
3020	4		3		2		9	27



Mission-focused, Employee-centric, Teamwork, Integrity, Speed

Results

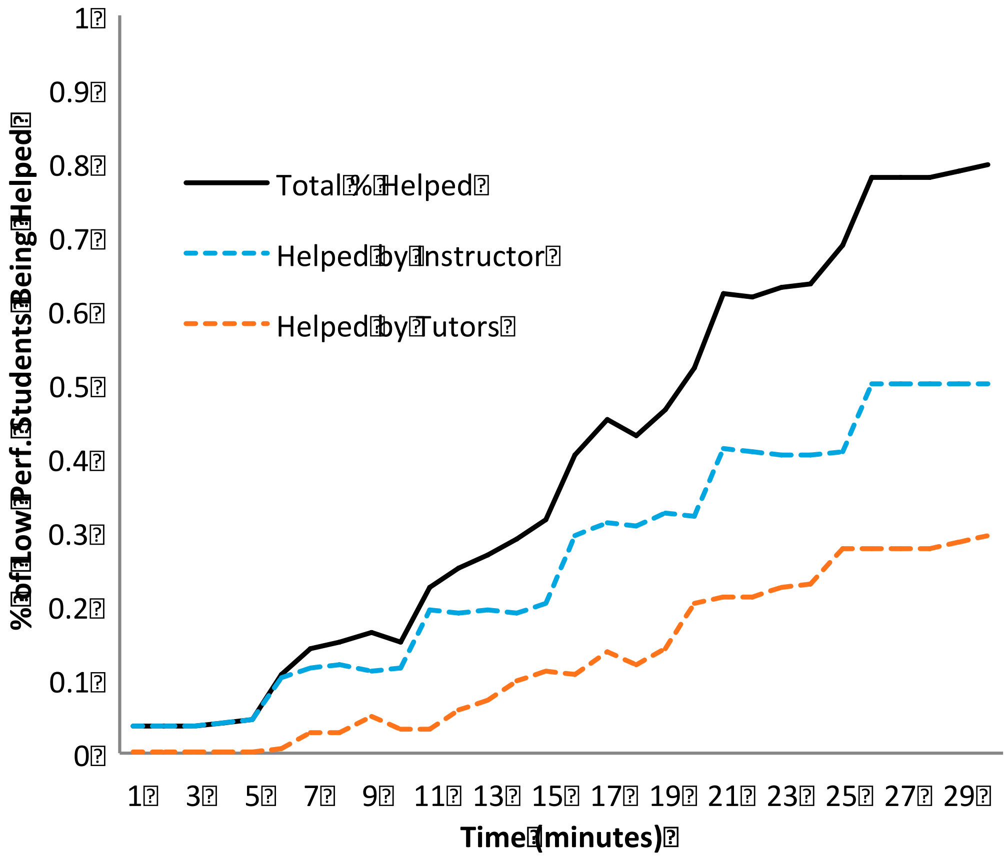
- Can predict the overall success of students based on their solution paths
- We can identify important code-chunks that math the human defined rubric

Is this result useful?

- Simulated classroom dashboard to explore how useful the information would be
- Subsampling and time alignment to simulate a lab session

Classroom Replay in an Instructor Dashboard

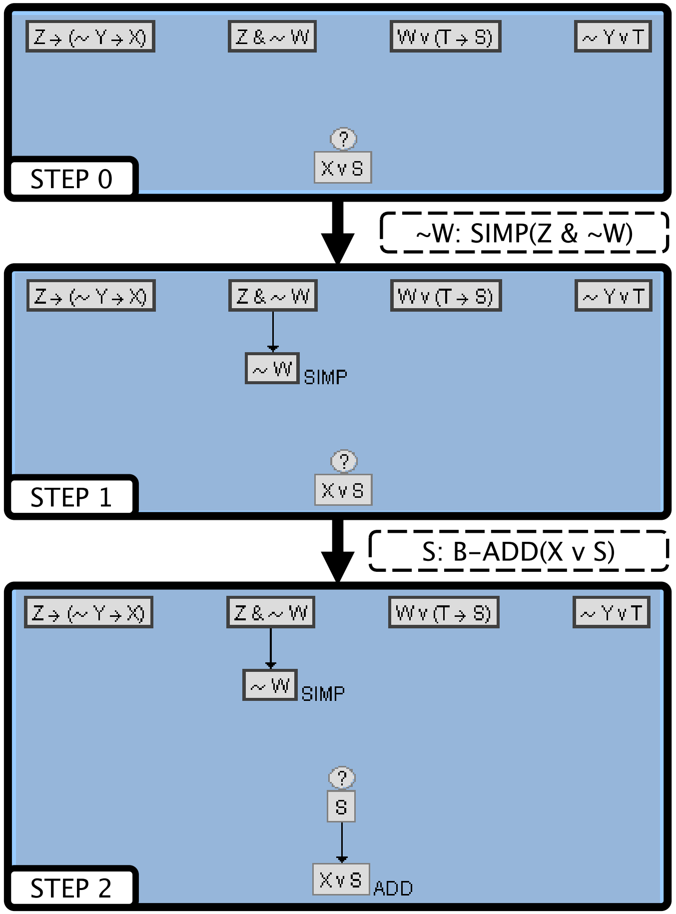
Matching students with peer tutors can increase the amount of help



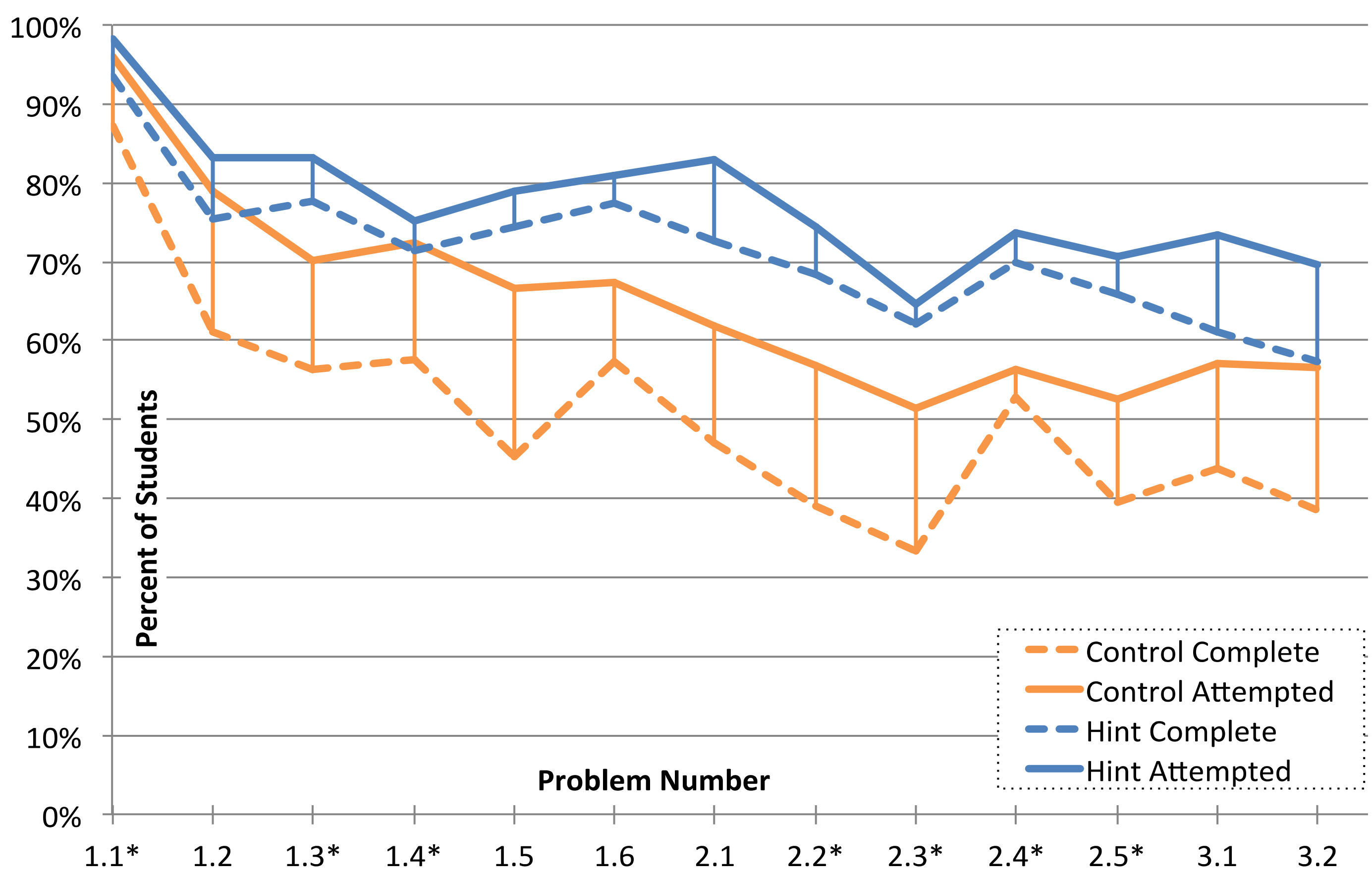
Questions?

Deep Thought Logic Tutor

- The goal is defined, but the process to the goal is not
- There are a finite number of actions used to manipulate the environment
- Success is determined by relevant or useful application of rules



Hint Effects



Hint Factory Primary Results

- Control group odds of dropping after first six problems was 3.6 times that of the experimental group
- Hint group out performed the control group by about half of a standard deviation in terms of tutor performance
- Hint group had a higher overall course score by one third of a standard deviation when compared to the control group

Reason for Dropout Differences?

- In what way did the Hints prevent student dropout?
- Did Lack of Hints resulted in Control Students becoming frustrated and quitting.
- Did students in each group spend different amounts of time in tutor?
    - Turns out to be a complicated question.

Questions to Answer

- "How long would it take students to complete the tutor"
  - Vs. "How long did it take to complete the tutor (for those that managed.)"
- "Were there time differences between-groups?"
  - Vs. "Were there time differences between-groups (for students that managed to finish the tutor"

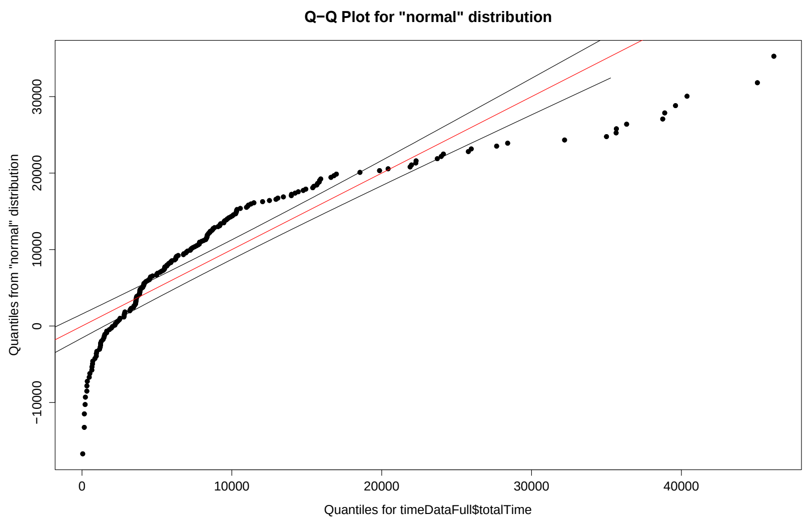
Tutor Efficiency

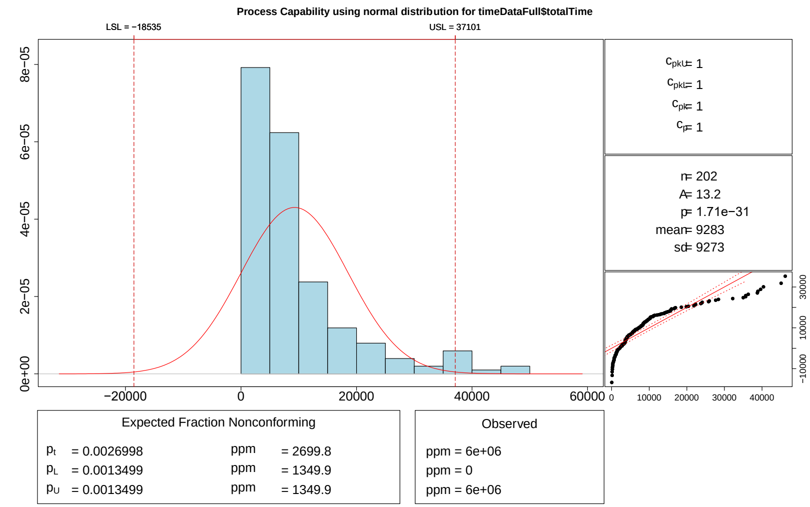
- "How long does it take students to complete the tutor."
- Three Elements
  - Performance – tutor completion percentage
  - Duration – total time spent interacting with the tutor
  - Dropout – whether stopped before completion

Dropout Bias

- Different dropout/attrition rates between experimental groups causes attrition bias
  - The students are self-selected due to achievement levels
- Is a threat to the study generalizability

Duration data is not Normally Distributed



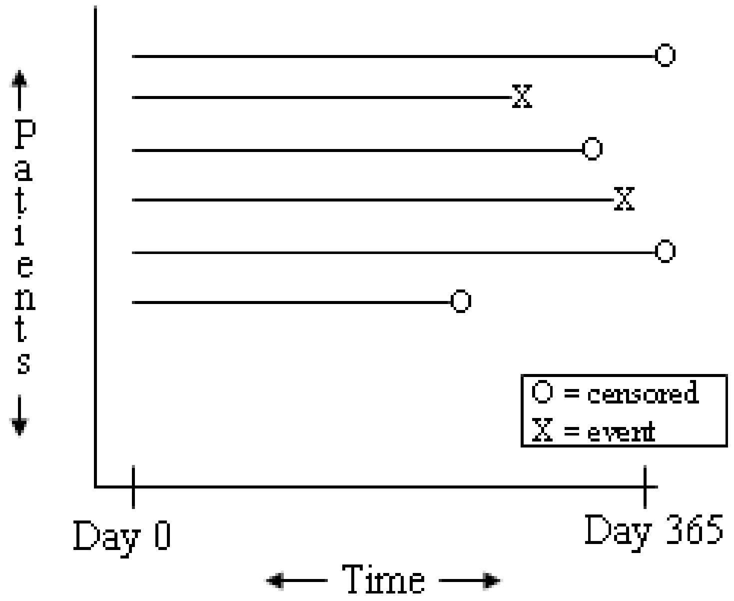


Survival Analysis

- Survival Function
  - $S(t) = Pr(E > t) = 1 - F(t)$
  - t – the time in question
  - E – the time of the Event (tutor completion)
  - F(t) – the duration distribution
- Gives the probability that the time of the tutor completion event, E, is later than t.
  - "The probability that the student has not completed the tutor."

Right Censoring

- We know the *tutor-start time* but do not always know the *tutor-complete time*



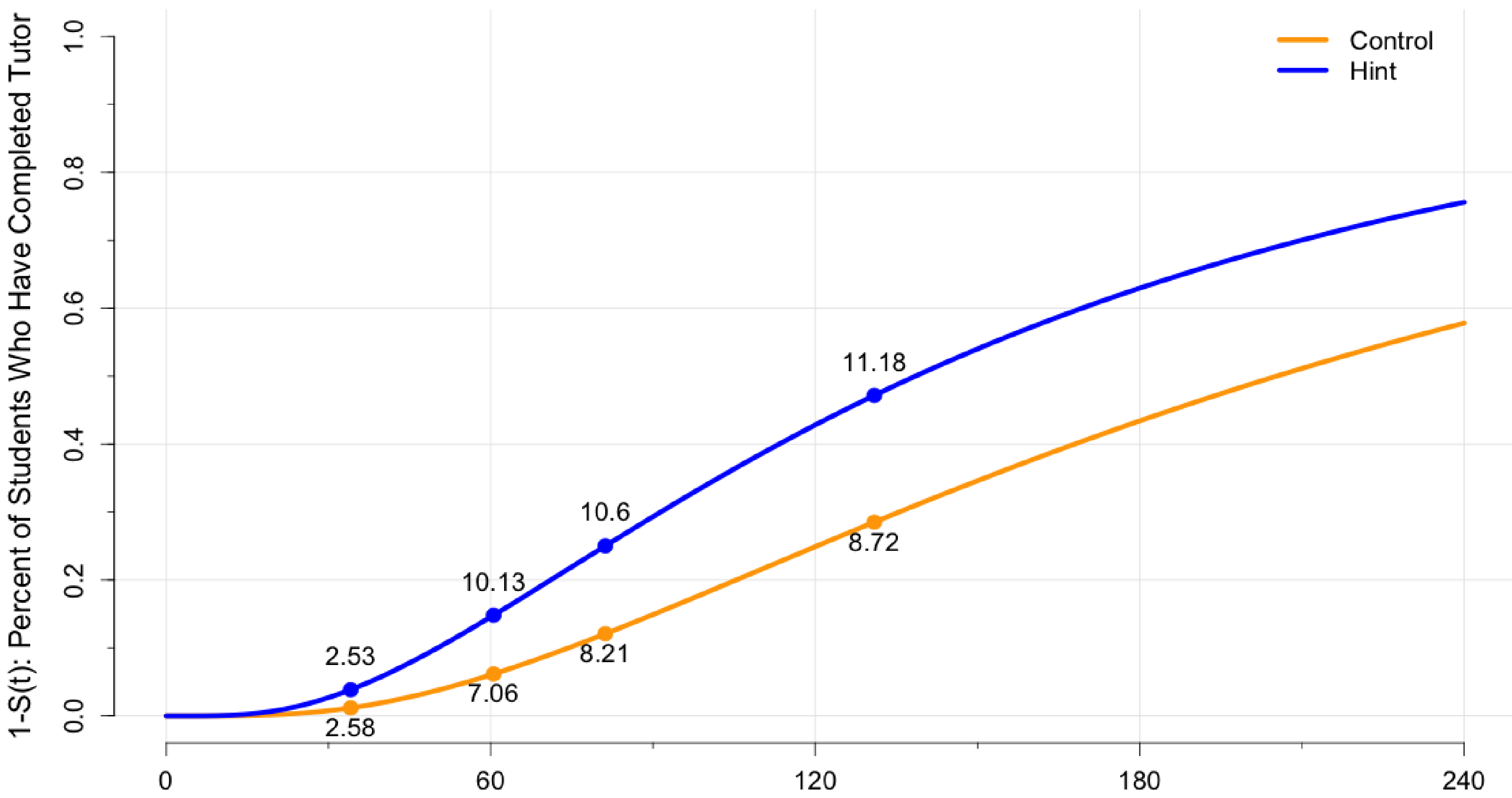
Accelerated Failure Time (AFT)

- Assumes that the effect results in one group having a shortened or lengthened lifetime
- $S(t|\theta) = S(\theta t)$
  - Θ – is a direct modifier to the tutor completion time

Results

- Intercept (mean) = 5.655
- Θ = -0.599
  - Hint group takes 55% of the time it takes the control group to complete the tutor
- Control Median
  - $E^{5.65} = 284.29$
- Hint Median
  - $E^{5.65-.599} = 156.18$

Survival Curve



Estimate Time Needed

- "How much time is needed so that 50% of the students can complete the tutor."
- Using the Survival function S(t) = .5, we can estimate that:
  - Control group needs 4.76 hours
  - Hint group needs just 2.61 hours

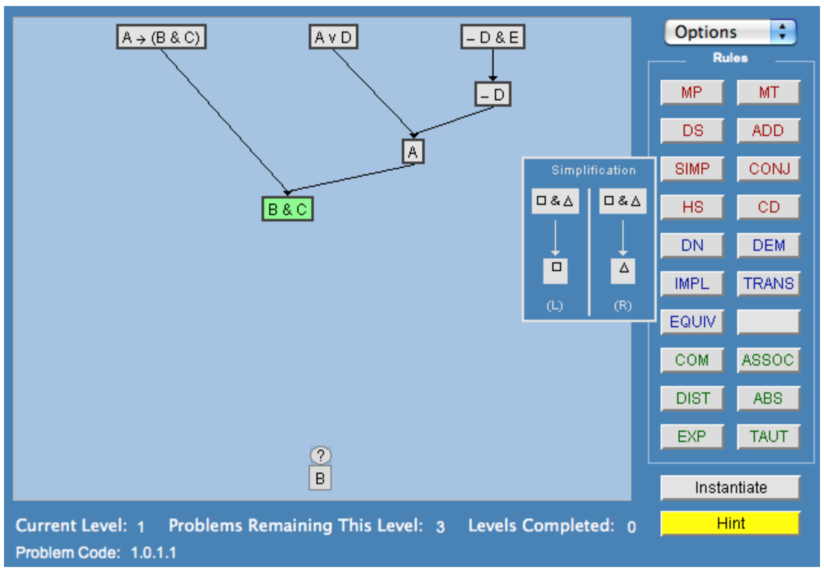
New Understanding

- We know about the observed performance differences
- We understand dropout’s connection to learning efficiency
- But, where is the control group’s time going?

Other Domains?

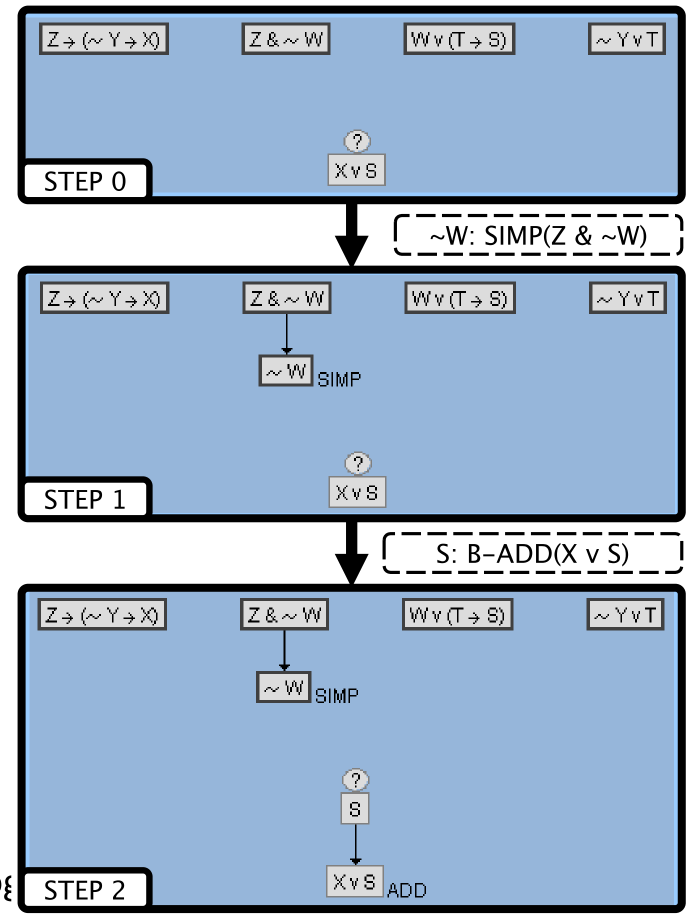
Interaction networks have worked in a number of environments

Propositional Logic Tutor

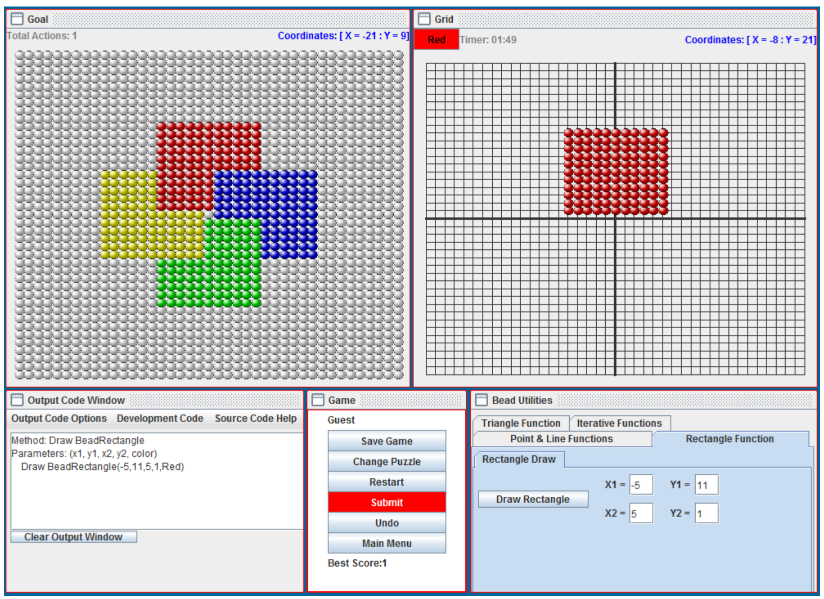


Chapter 4: Network Features

Deep Thought Log



Cartesian coordinate puzzle game

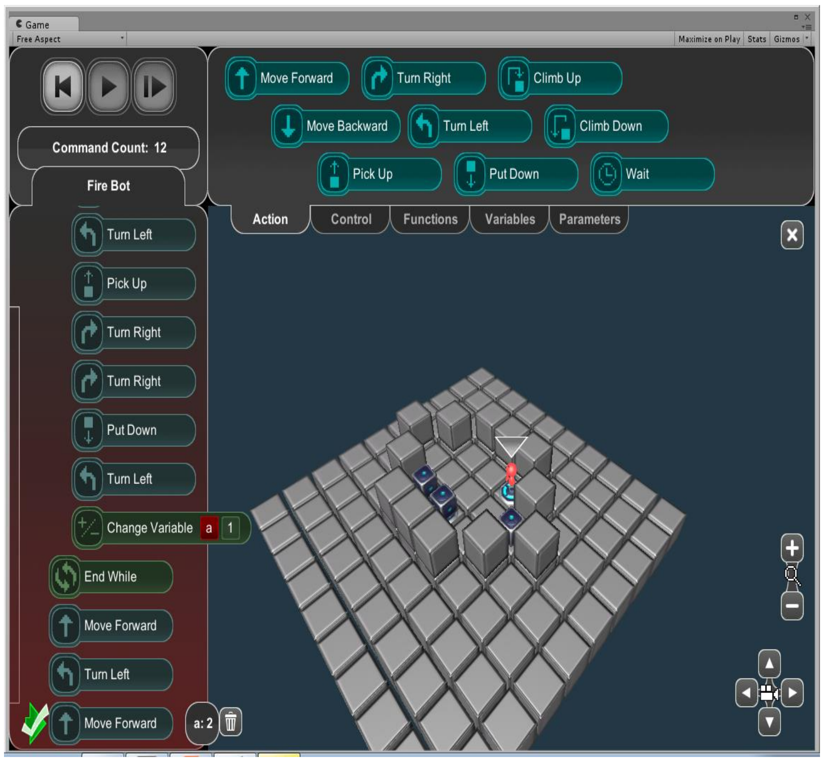


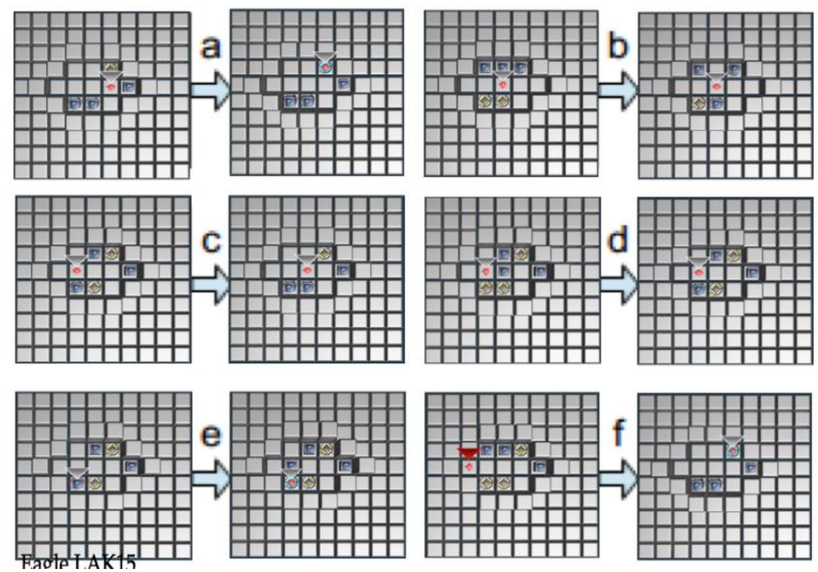
Start-State	Action	Result-State
	rectangle(-5 1 5 11)	
	rectangle(1 -5 11 5)	
	rectangle(-5 -11 5 -1)	
	rectangle(-11 -5 -6 5)	
	rectangle(-5 -5 -1 0)	

Virtual BeadLoom Game

Chapter 4: Network Features

Programming Game

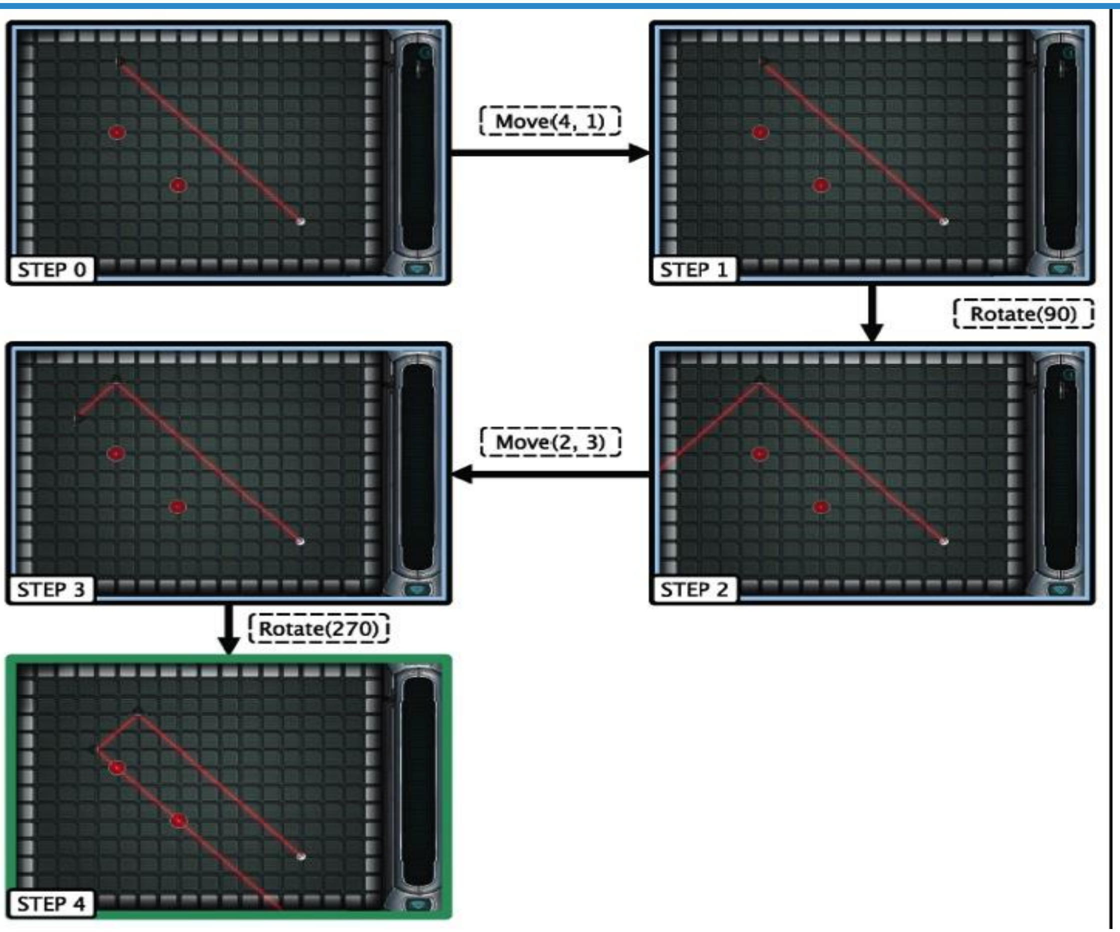




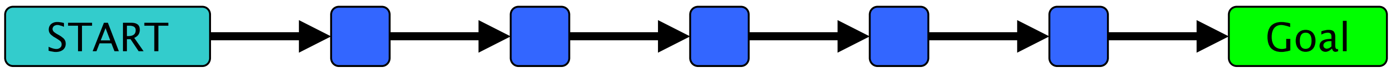
Chapter 4: Network Features

BOTS

Optics Game

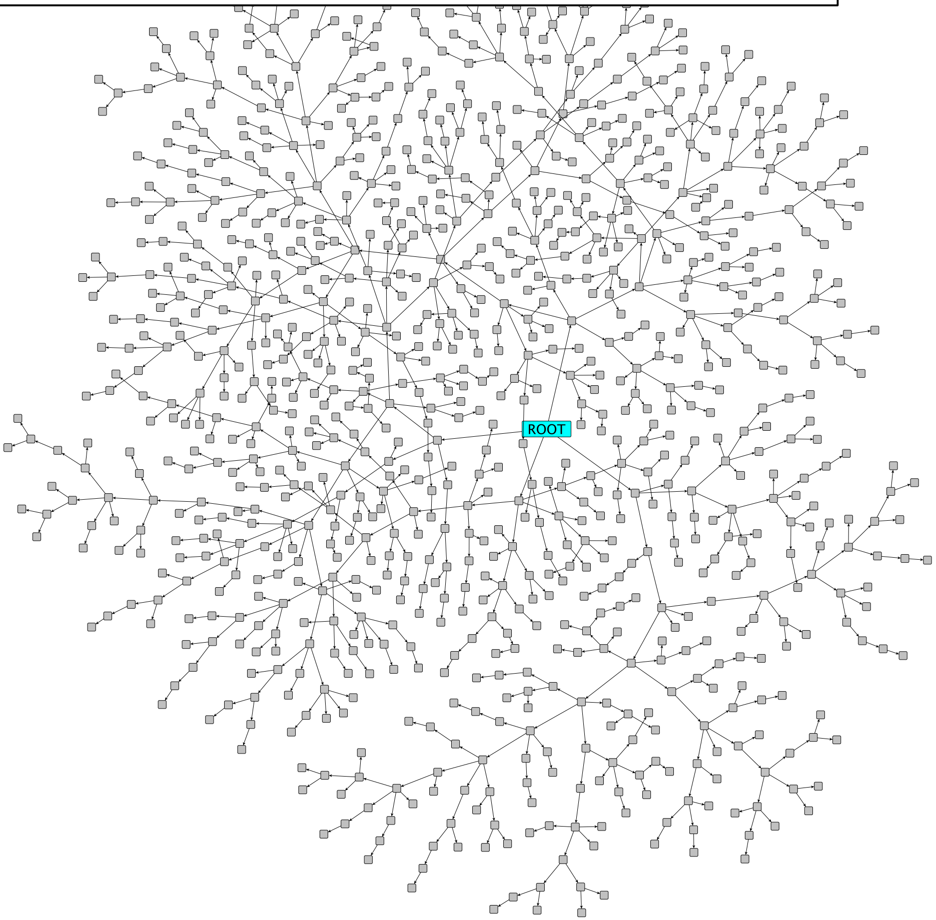

STEP 0
Move(4, 1)
STEP 1
Rotate(90)
STEP 2
Move(2, 3)
STEP 3
Rotate(270)
STEP 4

Interaction Network as a Construct

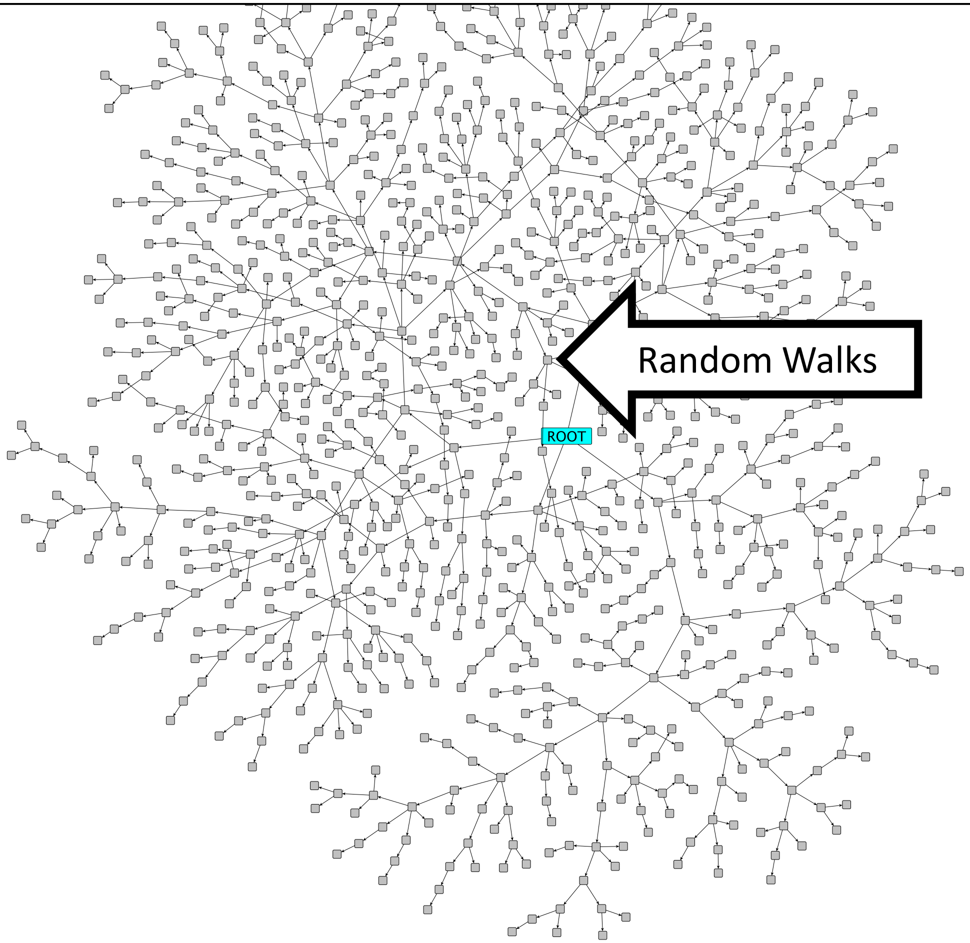


- Internal (unobservable) cognitive processes happen between steps
- Application of rules results in changes in the external environment (observable)
- If rule selection is governed by cognitive biases, then the observable output will reflect this
- That is:
  - A random walker will perform different than a bias walker
  - Walkers with different biases will perform differently

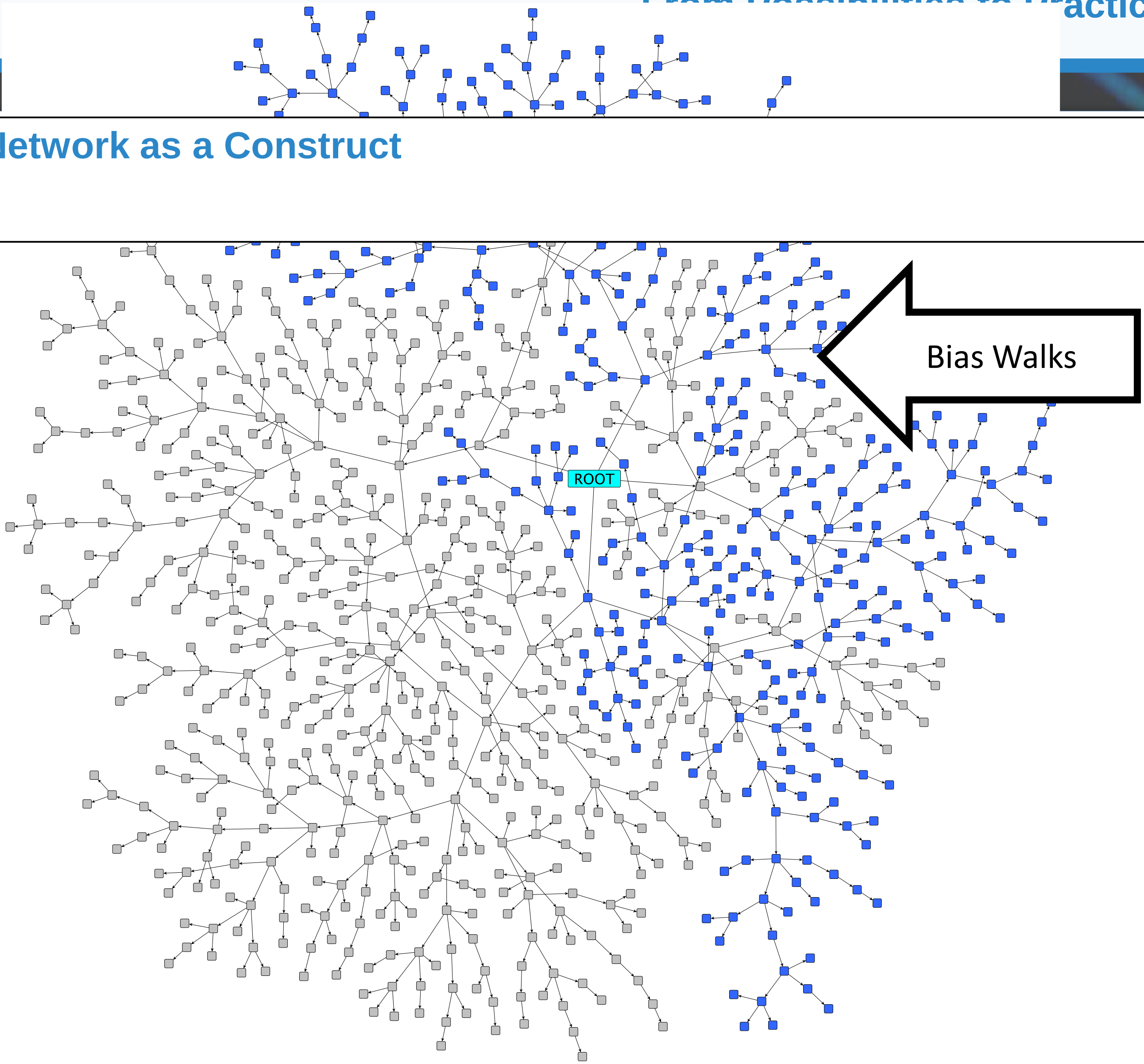
State Space of a Problem



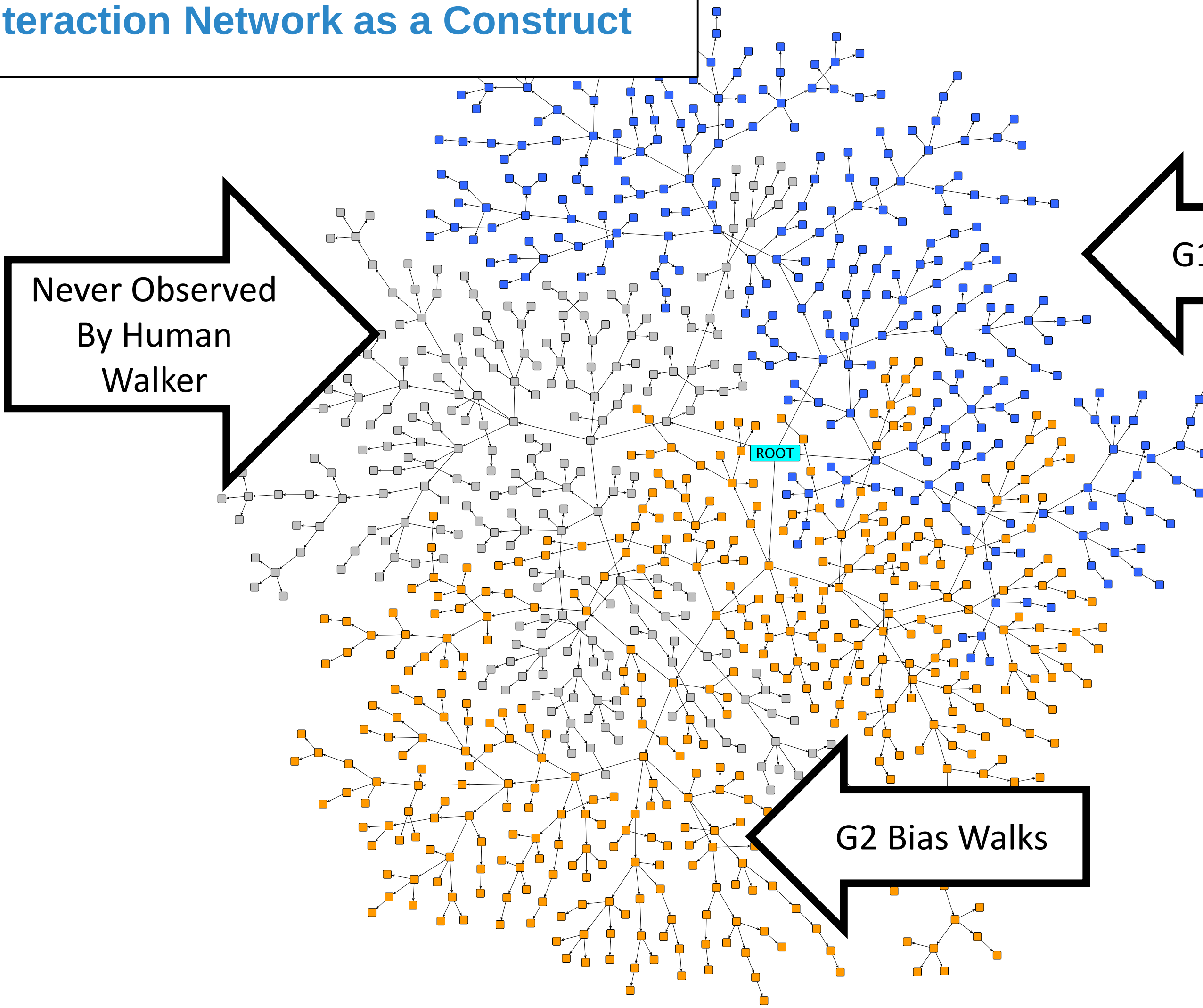
Interaction Network as a Construct



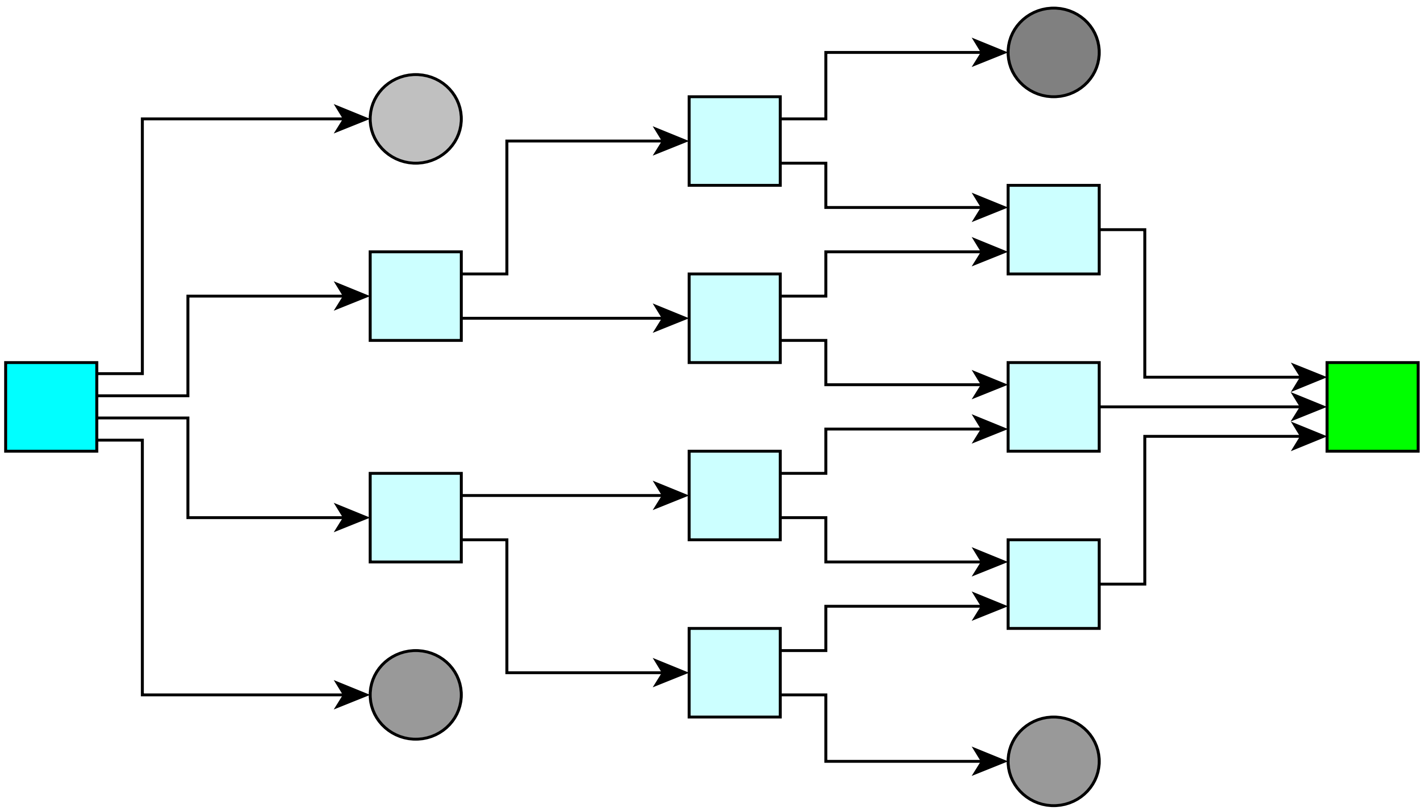
Interaction Network as a Construct



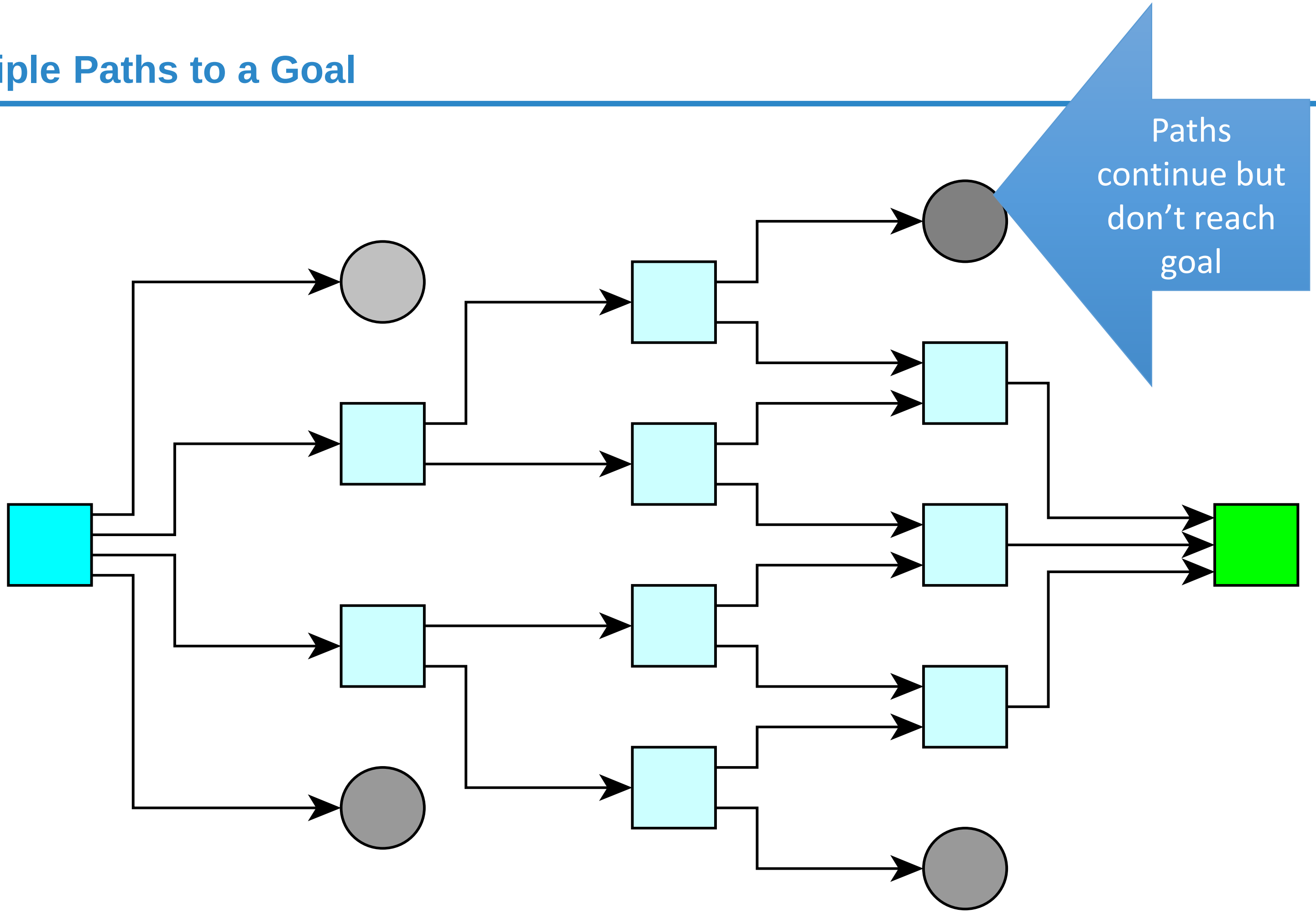
Interaction Network as a Construct



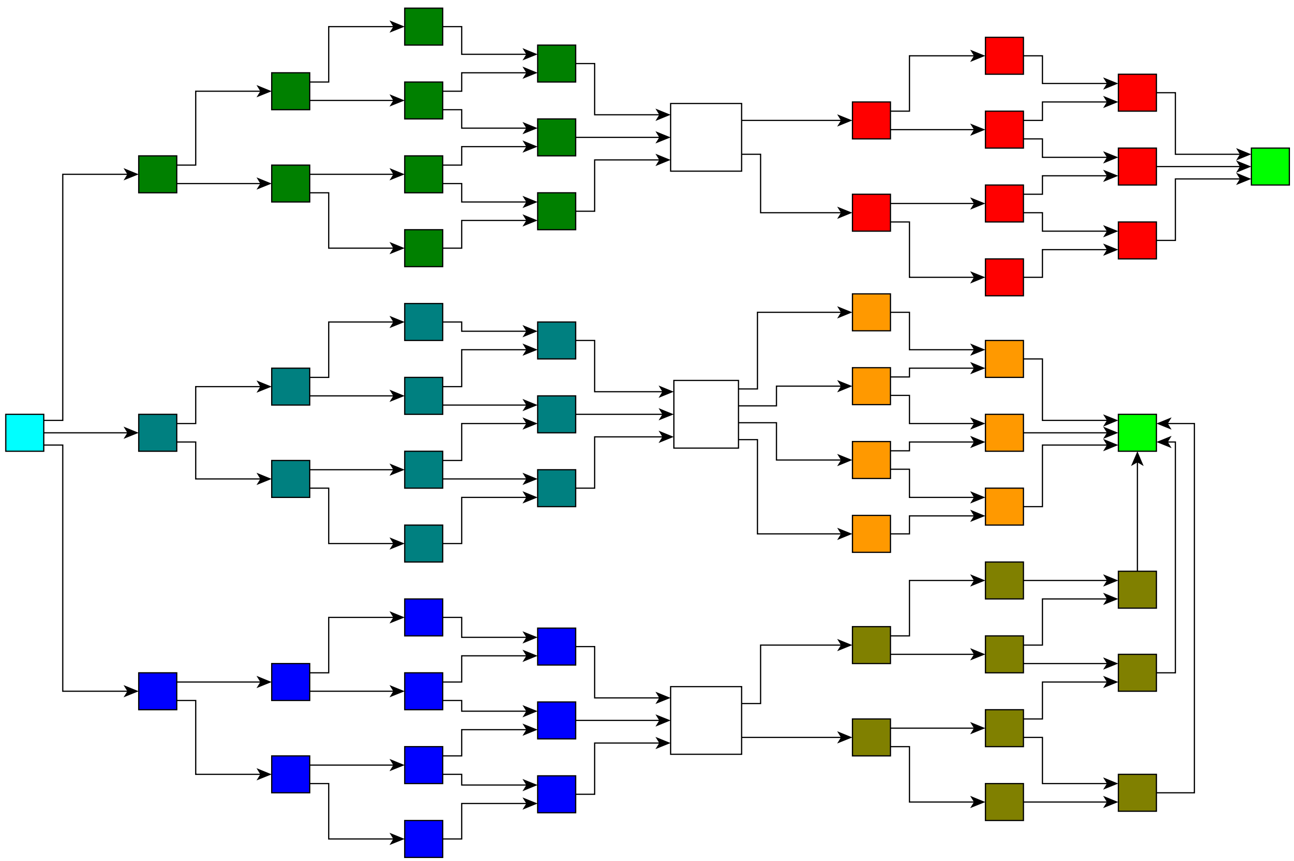
Rule-using Problem: Multiple Paths to a Goal



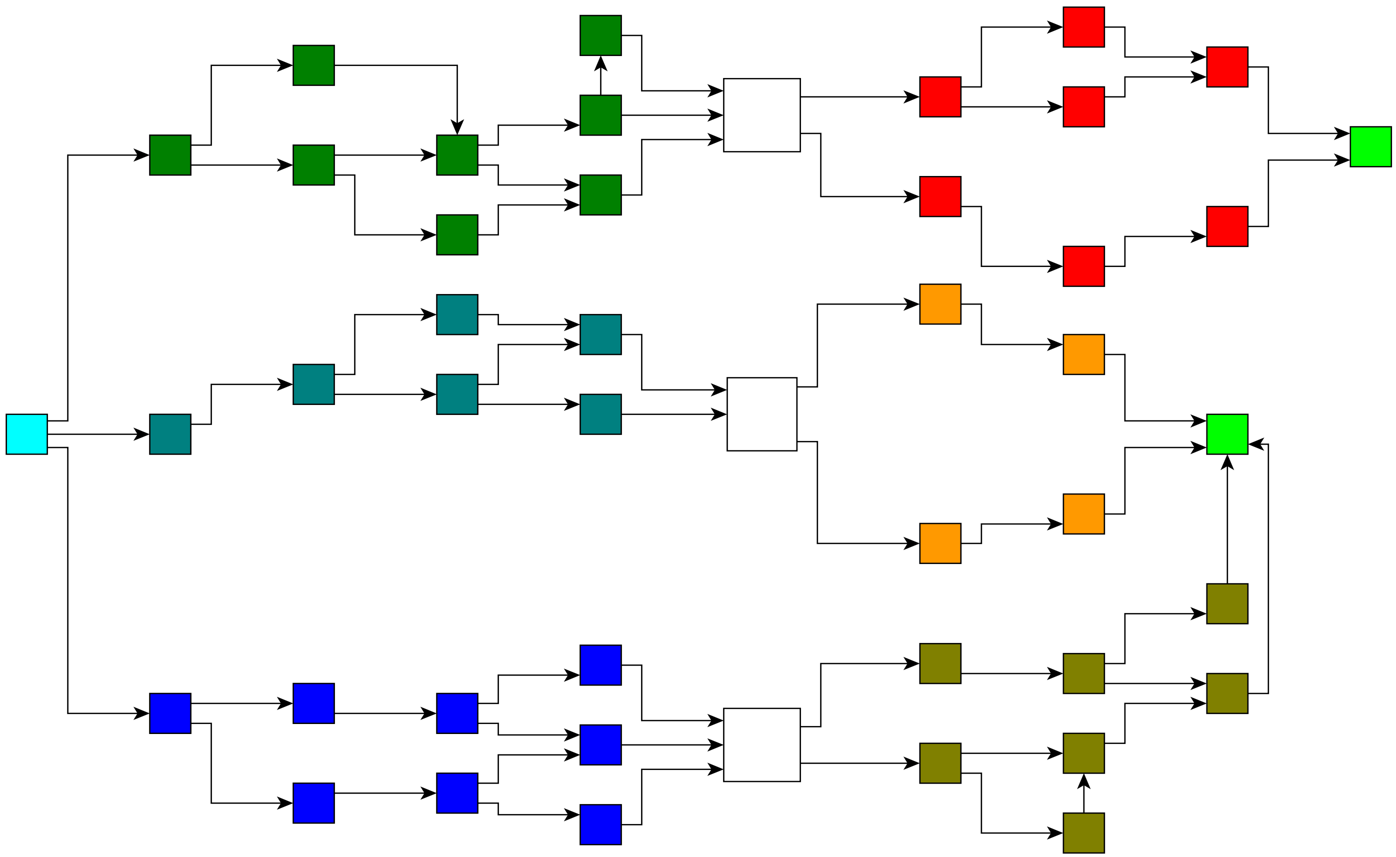
Multiple Paths to a Goal



Multiple Paths and Multiple Goals



Not all paths observed



These States: higher degree