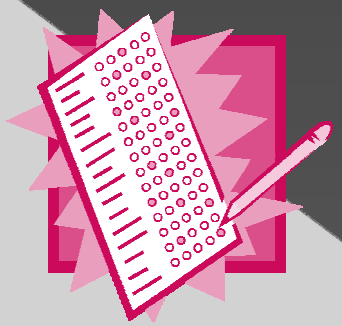
# Graph-based Adaptive Diagnosis

CS 194 Final Project Results

Jessica Wang
Advisor: Amit Sahai

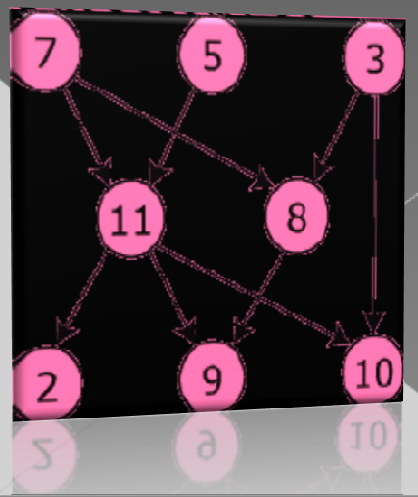


# Problem:

- Create a student evaluation system
  - More effective than a regular test
  - Responds and adapts to the student's answers
- Accurately assess student understanding
- Identify the weak topics/areas from just a few questions

# Proposed Implementation

- Utilization of a directed graph
  - Subtopics are represented by nodes
  - Edges have directions, weight, and type
- Subject: math
- Calculate student's mastery in topic
- Calculate test's confidence of that estimation

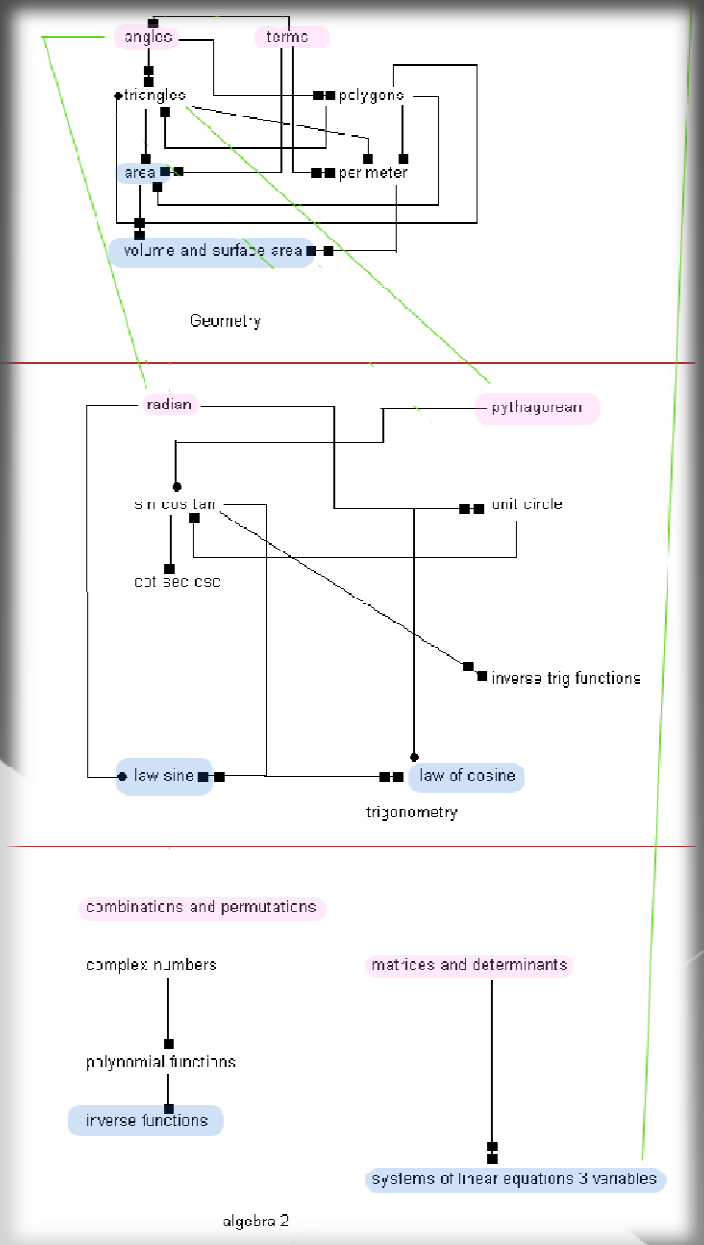


# Actual Implementation

- Graph of nodes and edges
- Edges: traversal done between
  - Critical nodes
  - subtopics that are absolutely necessary for the next topic
- Nodes
  - 2 types of nodes: critical and normal
  - Traversal starts out from critical nodes
- Evaluation algorithm: 2 parts
  - Part 1: region determination
  - Part 2: region evaluation

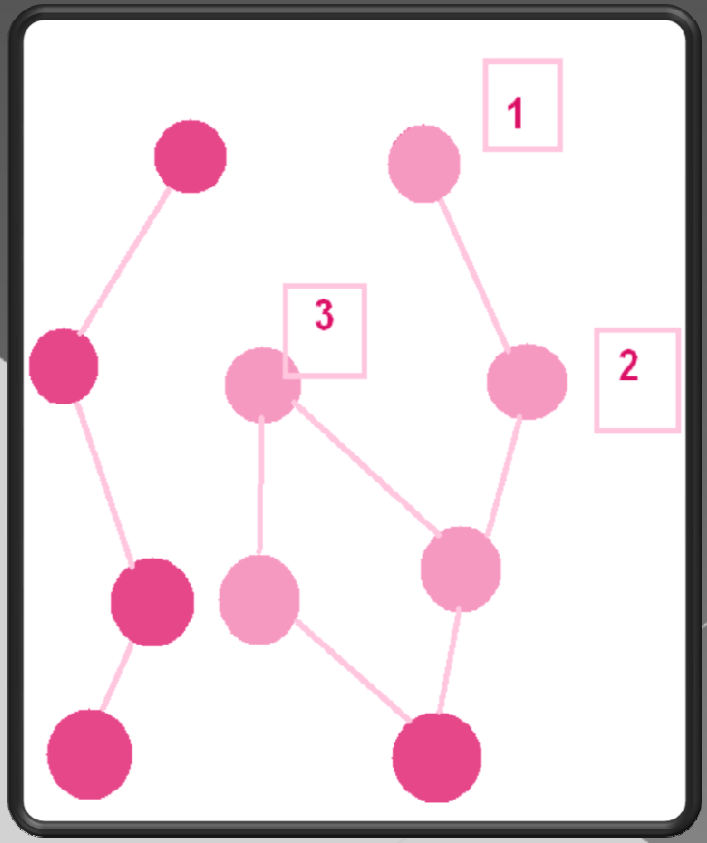
# Region Exploration

- Traversal algorithm between regions:
  - Start from the easiest region
  - Move to higher/lower region depending on student's response
  - Return the region to be explored
  - Return whether this algorithm is conclusive/inconclusive

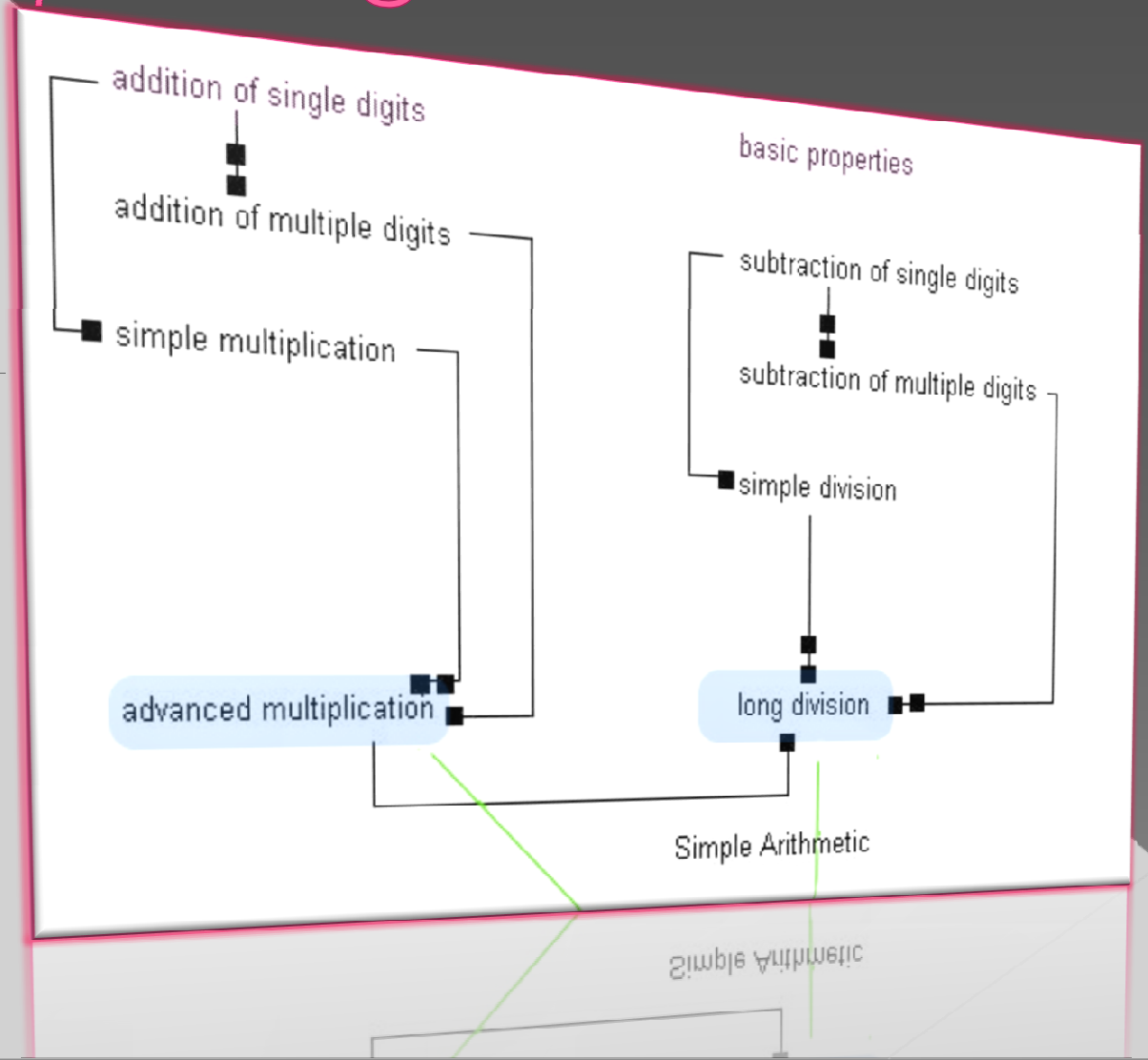


# Subtopics determination - Algorithm

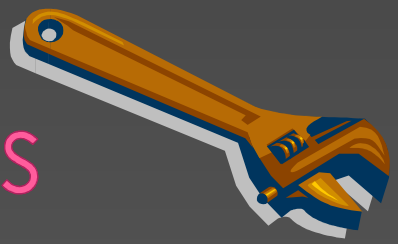
- Algorithm to explore region
  - Weird marker
  - Trickle-down estimation of mastery
  - Parallel traversal of routes: pick the next highest topic to ask



# Subtopics determination – Sample Region



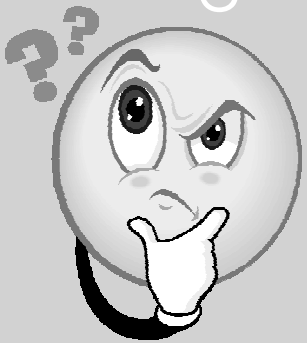
# Issues



- Ambiguous results: did student really master the region/topic?
- How many questions to ask?
  - Per topic
  - Per test
- How many points to deduct?
  - Are points too specific?
  - Use a more general marker

# Results

- Determines the correct region most of the time
  - Algorithm sometimes underestimates student ability
  - Correctly takes into consideration student guessing
    - When answered incorrectly 20% of the time
    - Less accurate when student misses a higher percentage (33% of questions wrong)



# Results

Like your normal math test!

- Efficiency
  - Most efficient when everything is right
  - Worst case: around 30 questions
  - Average case: 15-20 questions (short questions)
- Confidence of estimation
  - Does mastery marker match skill indication?
  - Did weird marker go off?

# Conclusion/Future Work

- Evaluation system
  - Identifies the boundary between regions mastered and not yet mastered
  - If boundary is not clear, identifies a boundary within unclear region
  - Works relatively well: should have better confidence gauge
- Give the test to actual middle/high school students
- Implement graphic representation of the results