DESIGN AND ANALYZE COMPUTERIZED ADAPTIVE TESTING WITH GRAPH THEORY
Computerized adaptive testing (CAT) improves testing efficiency, yet is a sophisticated system.

- Dynamic administration process
- Complicated item selection algorithm
- A large number of unique test forms

This study uses the graph theory to build a CAT

- Visualize the internal process
- Simplify the administration process
- Control test quality
GRAPH THEORY

IN GENERAL

- Graph is used to model complicated relations between objects
- Graph consists of nodes (objects) and edges (relation)
- Common types of graph:
  - Social (relationship) graph
  - Transportation graph
  - Dependency graph
  - Structural graph

IN THE CASE OF CAT

- Graph is used to model the on-the-fly form-building process in traditional CAT
- Nodes are items and edges are responses
ALGORITHMS

BUILD A GRAPH

- A recursive search algorithm
  - A focal node emits multiple edges, each for a score category (denoted by U)
  - Update $\theta$ on each edge and connect the edge with the optimal item for the updated $\theta$
  - Repeat the process by changing the focal node to the newly connected node
Also, a recursive algorithm

- Find the focal node’s outgoing edges and move along the edge for the observed score
- When no acyclic connections, move backwards to previous node
- Repeat the process until the maximum length or no connections
EXAMPLE OF A 5-ITEM GRAPH

A DETAILED VISUALIZATION WHERE RED AND GREEN EDGES ARE FOR INCORRECT AND CORRECT RESPONSES

A SIMPLIFIED VISUALIZATION WHERE ONLY CONNECTIONS BETWEEN ITEMS ARE SHOWN
A SIMULATION STUDY

- A simulation study was conducted to compare the graph CAT and the regular CAT under 3 conditions:
  - 20, 40, and 60 items
  - Generated a pool of 300 items using 3PL model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1.034</td>
<td>0.213</td>
<td>0.577</td>
<td>1.757</td>
<td>0.872</td>
<td>1.017</td>
<td>1.168</td>
</tr>
<tr>
<td>b</td>
<td>0.059</td>
<td>1.048</td>
<td>-3.321</td>
<td>3.413</td>
<td>-0.62</td>
<td>0.095</td>
<td>0.697</td>
</tr>
<tr>
<td>c</td>
<td>0.1</td>
<td>0.038</td>
<td>0.023</td>
<td>0.263</td>
<td>0.072</td>
<td>0.097</td>
<td>0.127</td>
</tr>
</tbody>
</table>
RESULTS

THE 20-ITEM CONDITION

- Graph CAT
  - Used 78 items (26% of the pool)
  - Corr. = .95, RMSE = .33

- Regular CAT
  - Used 141 items (47% of the pool)
  - Corr. = .97, RMSE = .26
RESULTS

LOWER PRECISION IN GRAPH CAT

MORE UNDER-USED ITEMS IN REGULAR CAT
RESULTS

THE 40-ITEM CONDITION

- Graph CAT
  - Used 137 items (46% of the pool)
  - Corr. = .97, RMSE = .26
- Regular CAT
  - Used 211 items (70% of the pool)
  - Corr. = .98, RMSE = .19
RESULTS

LOWER PRECISION IN GRAPH CAT

MORE UNDER-USED ITEMS IN REGULAR CAT
RESULTS

THE 60-ITEM CONDITION

- Graph CAT
  - Used 177 items (59% of the pool)
  - Corr. = .97, RMSE = .24
- Regular CAT
  - Used 249 items (83% of the pool)
  - Corr. = .99, RMSE = .16
RESULTS

LOWER PRECISION IN GRAPH CAT

MORE UNDER-USED ITEMS IN REGULAR CAT
EVALUATE THE STRUCTURAL QUALITY OF THE GRAPH

- Analyze the graph (in terms of path, connectivity, components, etc.) to evaluate the structural quality of CAT

- Examples:
  - What are most and least busy nodes?
    - In the 20-item graph, node #227 had most outgoing edges (14 edges) and the node #173 and #212 had most incoming edges (12 edges)
  - Is there a path between two specific nodes?
    - In the 5-item graph, both node #26 and node #209 were connected to node #86, whereas node #164 was not connected to node #86.
SLIDES:

HTTPS://GITHUB.COM/XLUO11/2018_NCME

Authors:
1. Xiao Luo, Measured Progress
   luo.xiao@measuredprogress.org

2. Doyoung Kim, NCSBN:
   dkim@ncsbn.org