Analysis of Learning Map Structure for a Dynamic Assessment

Feng Chen, Amy K. Clark, Russell Swinburne Romine
The University of Kansas

Abstract

This study examines the structure of a learning map used for dynamic assessment in grades three through high school. Students were assessed on items measuring 307 nodes, a dramatically larger number of skills when compared with previous diagnostic assessment analyses. Student response data were used to model mastery probabilities at the node level, and these values were used to make recommendations regarding node-to-node connections in the learning map as well as node granularity. These findings were shared with content teams to serve as supporting evidence in their decision-making process for map revisions pertaining to the order and size of cognitive skills.

Research Purpose

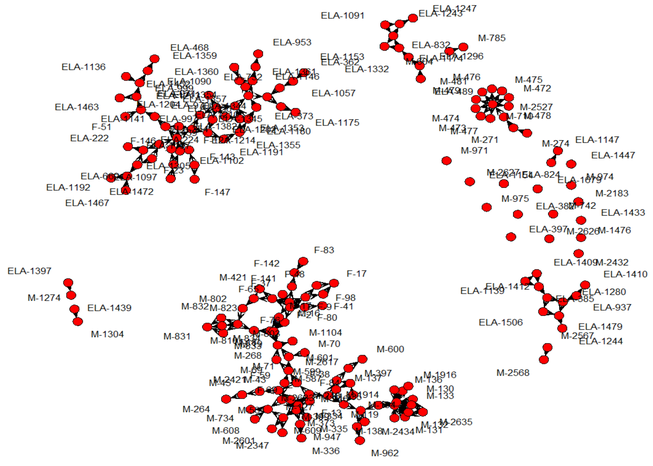
The current study covers three critical analyses in the evaluation of the underlying structure of a learning map used in a dynamic assessment environment.

1. Analysis of the connections between nodes in the learning map
2. Analysis of node granularity in the learning map
3. Content review of statistical recommendations

Methodology

Data

- Data came from 22,733 students in grades 3 through 12 across seventeen participating states assessed from spring 2014 to spring 2015
- 1,744 English language arts items
  - Only included items with sample size ≥100
- 532 English language arts testlets
  - Each testlet consisted of 3 - 8 items
- 214 nodes measured English language arts



Modeling

- Bayesian Inference Networks were employed to make causal inferences using conditional probabilities of multiple observations
- Loglinear cognitive diagnosis model (LCDM; Henson, Templin, & Willse, 2009), testlet effect included
- The Metropolis-Hastings algorithm, a Markov Chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution was employed

Results

Node-Level Example Results:

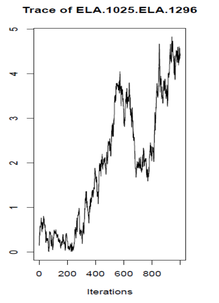
Map parameters	Logit	Credible Intervals	Geweke Z	Heidelberger p-value
ELA-1296	-.91	[-1.23, -.47	5.94	.08
ELA-1025.ELA-1296	1.97	[.09, 4.48]	-5.52	.08
ELA-1025	.37	[-.95, 1.30]	5.38	.13

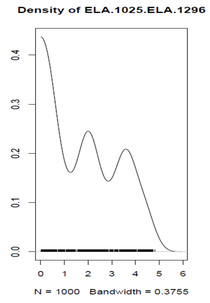
$$P(\alpha_{n_1} = 1) = \frac{\exp(\mu_{n_1,0})}{1+\exp(\mu_{n_1,0})}$$ Node: ELA-1296

Logit for Master of Node	Master	Non-Master
$\mu_{ELA-1205,0=} - 0.911$	.29	.71

$$P(\alpha_{n_2} = 1|\alpha_{n_1}) = \frac{\exp(\mu_{n_2,0}+\mu_{n_2,1,(n_1)}\alpha_{sn_1})}{1+\exp(\mu_{n_2,0}+\mu_{n_2,1,(n_1)}\alpha_{sn_1})}$$ Node: ELA-1025

Node ELA-1296	Logit for Master of Node ELA-1025	Master	Non-Master
Non-Master	0.37	.59	.41
Master	0.37 + 0.29	.91	.09

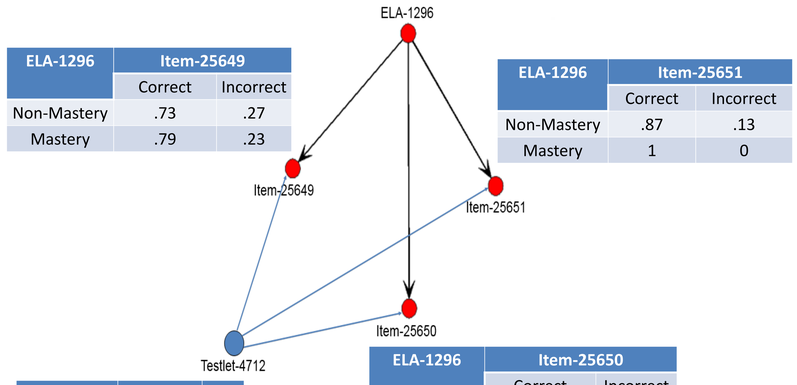




Item-Level Example Results:

ELA - 1296	
Non-Mastery	Mastery
.29	.71

$$P(Y_{si(t)} = 1|\boldsymbol{\alpha}_s, \gamma_t) = \frac{\exp(\gamma_t + \lambda_{i(t),1,a}\alpha_{sa})}{1 + \exp(\gamma_t + \lambda_{i(t),1,a}\alpha_{sa})}$$



ELA-1296	Item-25649	
	Correct	Incorrect
Non-Mastery	.73	.27
Mastery	.79	.23

ELA-1296	Item-25651	
	Correct	Incorrect
Non-Mastery	.87	.13
Mastery	1	0

Testlet	Variance	SD
Testlet-4712	.57	.15

ELA-1296	Item-25650	
	Correct	Incorrect
Non-Mastery	.56	.44
Mastery	1	0

Results Continued

Overall convergence rate: 84%

Node-Level Results:

- Nodes convergence rate: 90%
- Node reversal: nodes where non-masters have a high chance of being master on subsequent nodes
  - Includes intercept value of the successor node greater than zero
  - 46 reversal nodes out of 214 English language arts nodes
- Node overspecification: where two nodes are not distinguishable from one another
  - Includes intercept value less than -4, and main effect value greater than 8
  - No such nodes found satisfying two of the conditions simultaneously

Item-Level Results:

- Good items: items discriminate well between masters and non-masters of the node
- Non-informative items: items not discriminating, and thus to be flagged
  - 299 out of 1,744 items flagged

Conclusions

- Based on the results presented, map connections and causal relationships hold in general. There are 46 out of 214 nodes were flagged as potential reversals, which may jeopardize the node connections. Such results could be caused by non-informative items within the nodes, testlet effects in the model, or length of chains being burnt.
- In terms of node granularity, no overspecified nodes were detected among the 214 nodes assessed by items with a sample size of ≥100, meaning that all the examined nodes were reasonably distinct from their precursors and did not need to be collapsed.
- In addition to node-level estimation, item-level examination for informing test construction, scoring, quality control, and other features of test development. Items are expected to discriminate well between masters and non-masters of the node. Items that fail to provide such information are to be flagged and reviewed by content experts.
- Statistical evidence provides one source of information for evaluating the structure of the map. Content experts review flags in context of testlet, test specifications, and underlying cognitive processes to make final item quality determinations.

References

Gierl, M. J. (2007). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. Journal of Educational Measurement, 44, 325-340. doi: 10.1111/j.1745-3984.2007.00042.x

Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74,* 191-210.

Huff, K., & Goodman, D. P. (2007). The demand for cognitive diagnostic assessment. In J. P. Leighton & M. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications.* (pp. 19-60). Cambridge: Cambridge University Press.

Contacts

Feng Chen — chenfeng27@ku.edu — Graduate Research Assistant
Amy K. Clark — akclark@ku.edu — Psychometrician
Russell Swinburne Romine — swin0030@ku.edu — ELA Research Lead

Dynamic Learning Maps | Achievement and Assessment Institute

This study was developed under grand 84.373X100001 from the U.S. Department of Education, Office of Special Education Programs. The views expressed herein are solely those of the author(s), and no official endorsement by the U.S. Department should be inferred.