Spring 2021 Methodology Applications Series @ the University of Nebraska – Lincoln
February 5, 2021

Modern Psychometric Approaches for Diagnostic Assessment

Ray E. Reichenberg, Ph.D.

Research Assistant Professor, CYFS / MAP Academy

rreichenberg@unl.edu

Agenda

- What is Diagnostic Assessment?
- Where and how is it used?
- What kinds of information do we get from diagnostic assessments?
- What kinds of models are used to estimate attribute profiles?
- What are the basic specifications and characteristics of these models?
- What do these models look like in practice?
- How can these models be extended?
- What software do I need in order to use these models in my own research?
- Where can I find more information?

What is Diagnostic Assessment?

Diagnostic Assessment

- Assessment: A systematic procedure for collecting information that can be used to make inferences about the characteristics of people or objects (*Standards*, 2014).
- Diagnostic Assessment: An assessment procedure whose goal is to make classification-based (i.e., categorical) decisions (diagnoses) about individuals.
- Represents an evolution in psychometrics from test scores (CTT) → individual item responses (IRT) → individual components of items/tasks (de la Torre et al., 2017)

Diagnostic Assessment

- Diagnostic assessments yield (Carragher et al., 2019):
  - Information on multiple skills
  - Grouping of respondents with similar profiles
  - The ability to adapt assessment based on skill profiles
- Whereas traditional assessments (CTT, IRT) are typically focused on scores and rankings, diagnostic assessments are concerned with classifying (e.g., master/non-master, proficient/not proficient, presence/absence, etc.)
- Typically applied to low-stakes assessment scenarios
  - Lack of literature on psychometric properties of models

How are diagnostic assessments used?

Diagnostic Assessment – Applications

- Educational assessment
  - Modeling mastery of fine-grained competencies
  - Using patterns of misconceptions to choose targeted interventions
  - Implementing adaptive testing at a higher resolution than is typical with IRT-based CAT
    - Trade-off: Inferences will be for a much narrower domain
- Clinical assessment
  - Modeling potential diagnoses (profiles) as a function of individual symptoms
  - Choosing interventions based on behavior profiles

Diagnostic Assessment – Applications

- I/O psychology
  - Modeling workplace competencies and then using the resulting profiles to match applicants to positions
  - Identifying employees with particular patterns of deficiencies in order to provide customized professional development

What kinds of information do we get from diagnostic assessments?

Attribute Profiles

- Diagnostic assessment provides estimates of skill status as well as pattern classifications known as attribute profiles.
- Attribute profiles are patterns of classification on the latent variables representing the skills being assessed.
  - For example, you might see [0, 0, 1, 0] used to represent an attribute profile corresponding to non-mastery on skills 1, 2, and 4 but mastery of skill 3.
- These profiles might offer additional information beyond the individual latent skill classifications (i.e., interactions).
- Attribute profiles can represent complex hypotheses about the relationships between the latent skills (e.g., attribute hierarchies).

DiagnOsis scoring report

Student Name: Margo **LanguEdge Reading Comprehension Test 1**

Review Your Answers

Question	1	2	3	4	5	6	7	8	9	10	11	12	13
Your Answer	√	√	√	2	√	1	4	√	√	√	3	2	2
Correct Answer	2	3	2	3	3	3	1	1	1	4	4	3	2,4,6
Difficulty	e	m	e	h	m	m	h	h	m	m	h	m	m

Question	14	15	16	17	18	19	20	21	22	23	24	25
Your Answer	√	√	√	√	2	√	√	√	3	2	3	5
Correct Answer	2	3	2	1	3	2	3	4	2	4	1	1,5,6
Difficulty	e	m	m	m	e	e	e	e	h	m	m	m

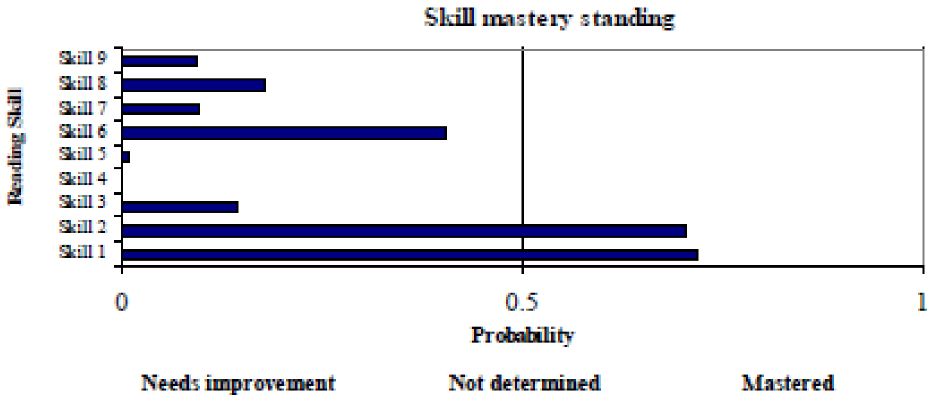
Question	26	27	28	29	30	31	32	33	34	35	36	37
Your Answer	1	√	1	4	4	√	√	√	o	√	3	1,4,6 2,3
Correct Answer	4	2	2	3	3	1	2	2	2	4	1	1,5,6 3,7
Difficulty	m	e	h	m	e	e	e	e	m	h	m	h

Scoring
Correct answer to questions with 4 choices = Plus 1 point
Wrong or omitted answer = No point
Q13 & 25: 3 correct = 2 points, 2 correct=1 point
Q37: 5 correct=3 points, 4 correct=2 points, 3 correct = 1 point

Key
√ Correct
o Omitted
+ Plus partial points
e = Easy, m = Medium, h = Hard
(Difficulty is based on 1372 students' performance on this test)

Score
You earned *20* out of maximum *41* points.
10 points from *12* easy questions
7 points from *17* medium questions
3 points from *8* hard questions
You omitted 1 question.

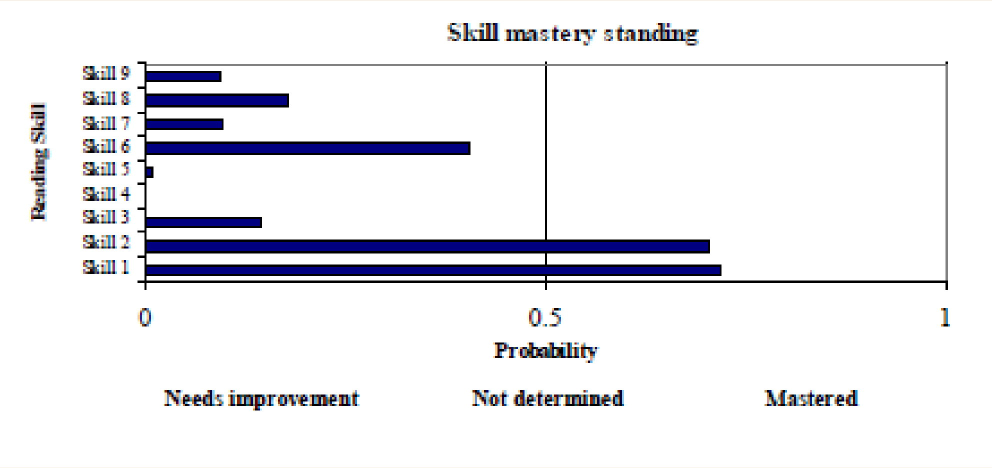
Improve Your Skills



How to Interpret Skill Mastery

- Nine primary reading skills are assessed in this reading comprehension test. Please review skill descriptions and example questions attached to this scoring report.
- The graph on the left side shows your probable mastery standing of each skill.
- The grey region indicates that your probable mastery standing cannot be determined.
- There may be some measurement error associated with the classification.
- This diagnostic information can be more useful when used in combination with your teacher's and your own evaluation of your reading skills.

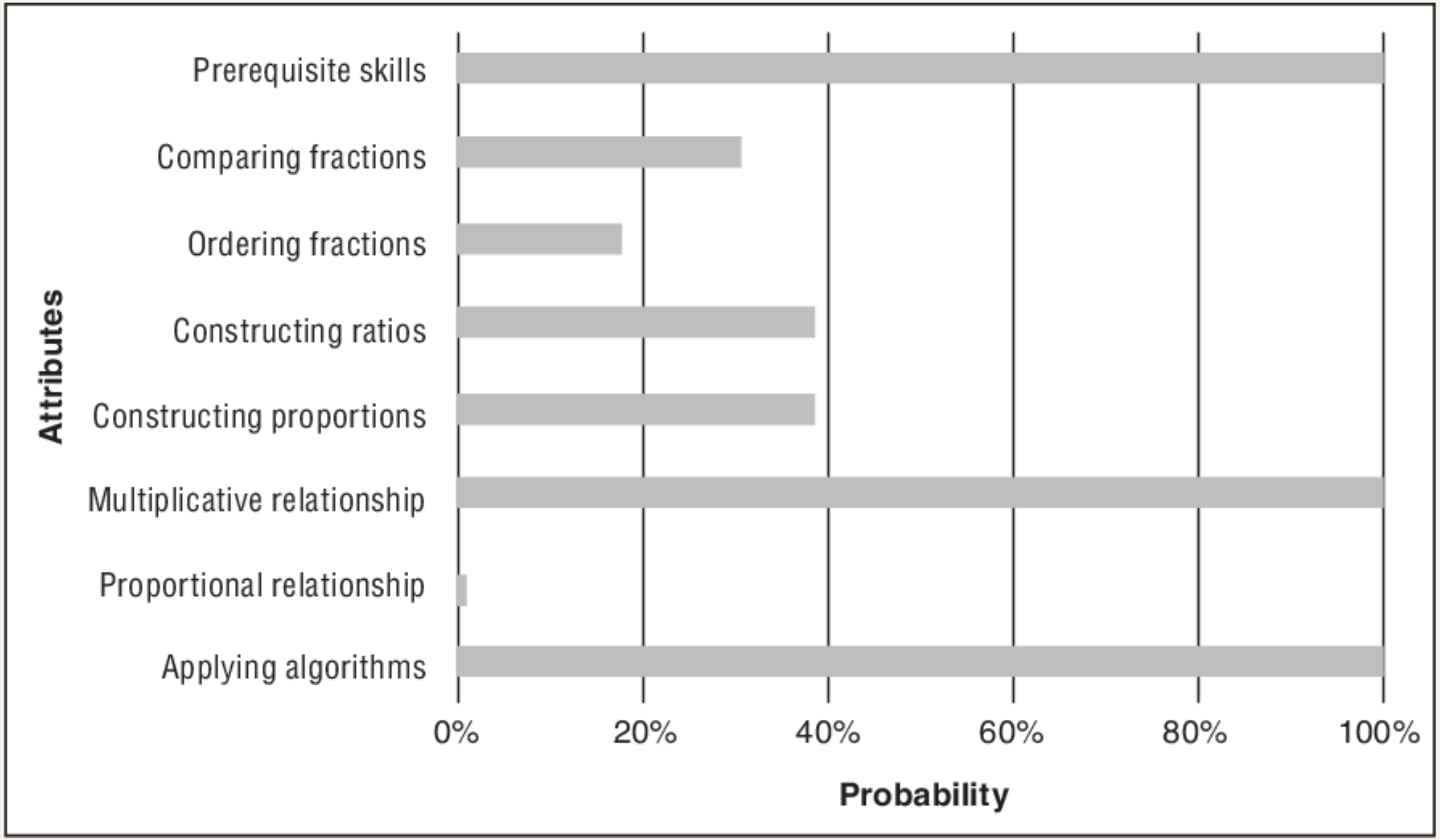
Source: Jang (2005)



We can apply a cut-off rule where Mastery is ≥60%, Non-mastery is ≤40%, values >40% but <60% are *indifferent…*

Skill 1	Skill 2	Skill 3	Skill 4	Skill 5	Skill 6	Skill 7	Skill 8	Skill 9
😀	😀	😞	😞	😞	😐	😞	😞	😞

Source: Jang (2005)



Source: de la Torre et al. (2017)

How do we model attribute profiles?

Modeling Attribute Profiles

- There are two dominant frameworks used for modeling attribute profiles:
  - Diagnostic classification models (DCM)
  - Bayesian networks (BN)
- Really there's just one "dominant" framework but I like Bayes nets, so…
- These models have a lot in common:
  - DCM can, in most cases be considered as a special case of a BN
  - Both tend to be entirely categorical (commonly, binary OVs and categorical LVs)
    - There are exceptions such as the HDCM models (de la Torre & Douglas, 2004; Templin & Bradshaw, 2014).
  - Not that dissimilar from LCA / LTA models in the most basic sense
    - Cross-sectional DCMs are often formulated as something akin to a confirmatory LCA with 2^A latent classes.

Diagnostic Classification Models

- DCM are a family of models that classify respondents into classes
  - Classes are mutually exclusive
- Also referred to as: cognitive diagnosis/diagnostic models, latent response models, structured IRT models, cognitive psychometric models, …
- DCMs can be thought of as constrained, or confirmatory latent class models
  - 2^A possible classes (for binary LVs) where *A* is the number of skills contributing to the profile
  - EX: a model with four skills will have $2^4=16$ possible profiles that examinees can be classified into.
- DCMs model responses to categorical (typically dichotomous) items as a function of discrete latent skill variables
- The measurement model is defined by a Q-matrix…

Diagnostic Classification Models

	Skill1	Skill2	Skill3	Skill4
Item1	1	0	0	0
Item2	1	0	0	0
Item3	1	0	0	0
Item4	1	1	0	0
Item5	1	0	1	0
Item6	0	1	0	0
Item7	0	1	0	0
Item8	0	1	0	0
Item9	0	1	1	0
Item10	0	1	0	1
Item11	0	0	1	0
Item12	0	0	1	0
Item13	0	0	1	0
Item14	0	0	1	1
Item15	1	0	1	0
Item16	0	0	0	1
Item17	0	0	0	1
Item18	0	0	0	1
Item19	1	0	0	1
Item20	0	1	0	1

Diagnostic Classification Models

- The probability of a correct response depends on the estimated proficiency for the skills represented by that item (LCDM formulation)

$$P(correct) = intercept + main\ skill\ effect(s) + interactions$$

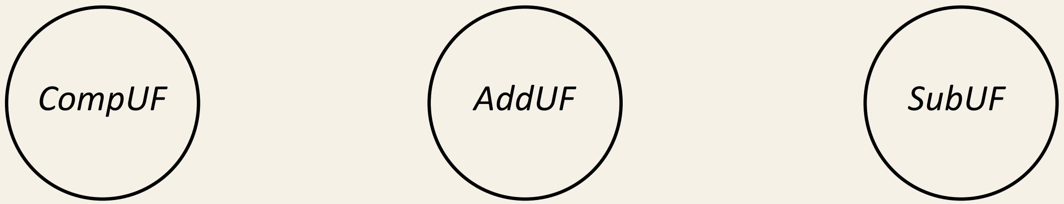
- There are several (and I mean SEVERAL) models in the DCM family which are separated by the assumptions they make about how skills interact to influence task performance
- DINA, DINO, NIDA, NIDO, C-RUM, R-RUM, LCDM, G-DINA, GDM, etc.

Diagnostic Classification Models

- Constructing a DCM typically requires four steps (Rupp, Templin, & Henson, 2010):
  - Identifying the target constructs (latent attributes)
  - Specifying the observed variables (items, tasks)
  - Linking the observed variables to the latent attributes (measurement model)
  - Specifying the relationship among the latent attributes (structural model; e.g., conjunctive [AND], disjunctive [OR], saturated, etc.)

DCM – Basic Example

- Assume we have three skills: comparing, adding, and subtracting unit fractions



DCM – Basic Example

- $J = 9$
- Q-matrix:

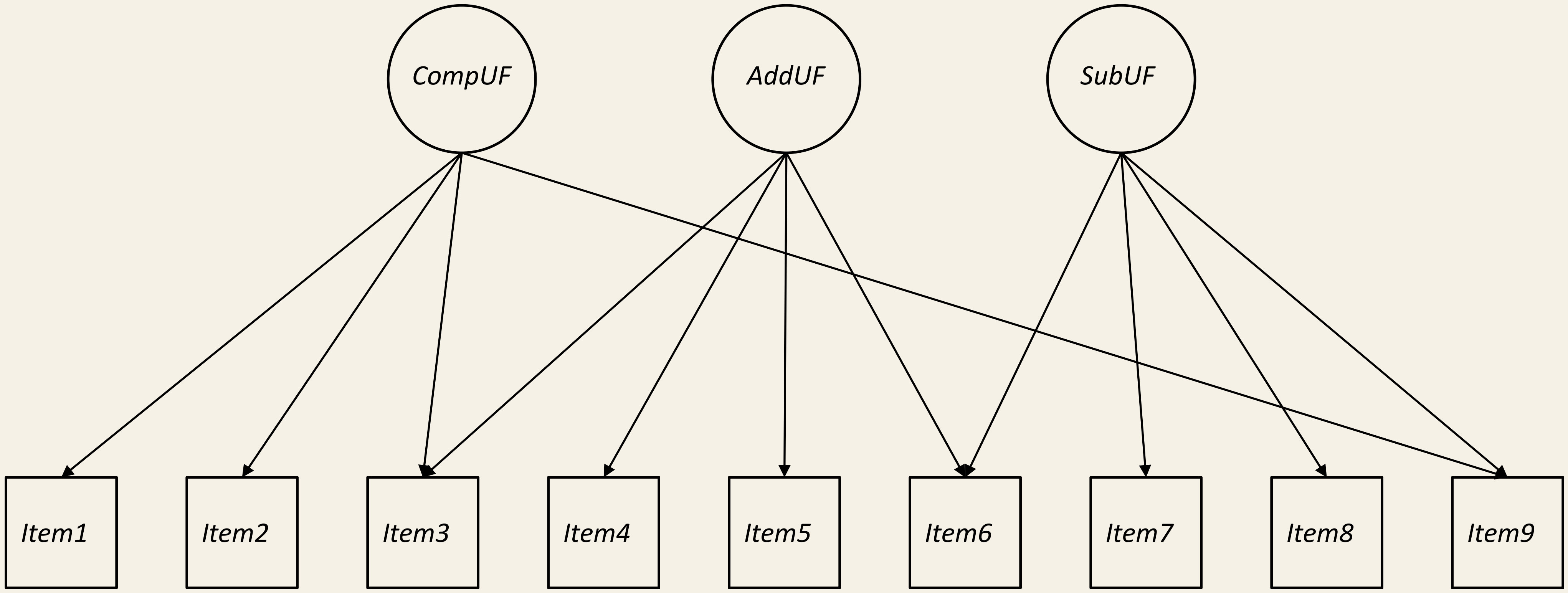
	Comparing UF	Adding UF	Subtracting UF
Item1	1	0	0
Item2	1	0	0
Item3	1	1	0
Item4	0	1	0
Item5	0	1	0
Item6	0	1	1
Item7	0	0	1
Item8	0	0	1
Item9	1	0	1

DCM – Basic Example

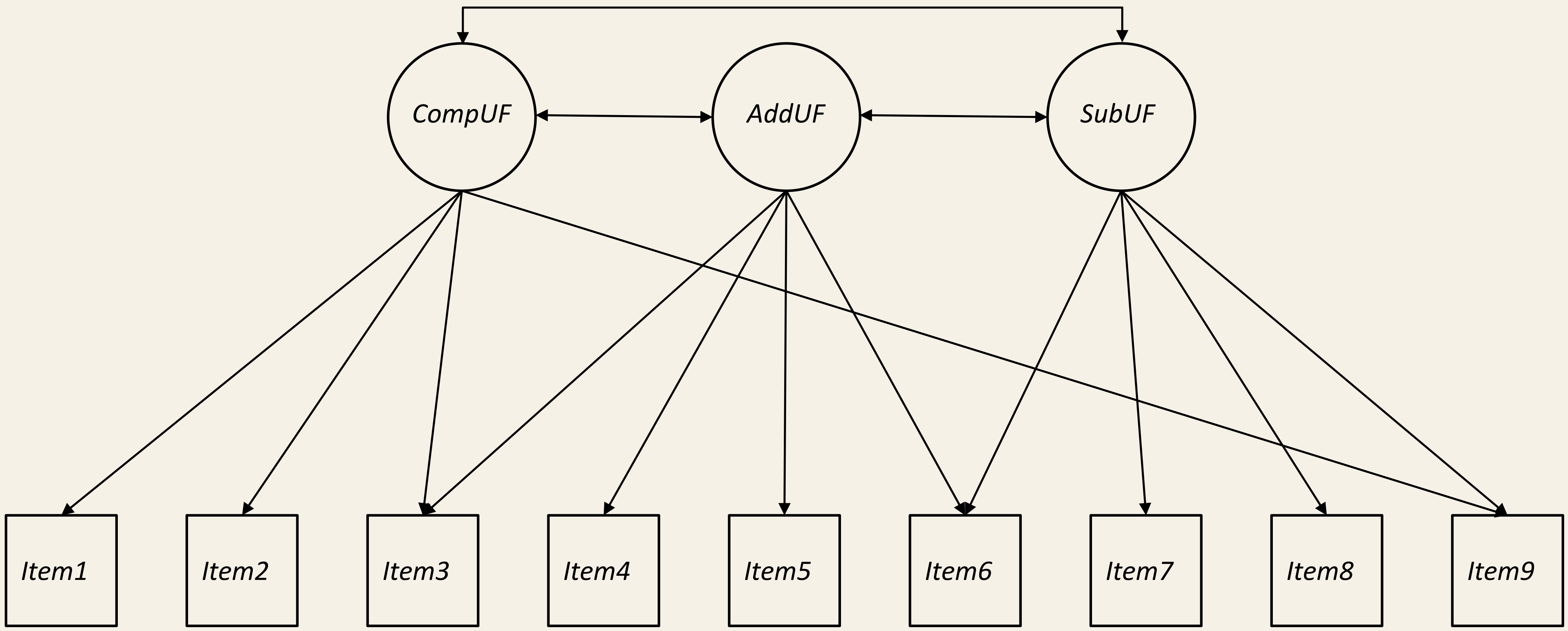
- $2^A = 2^3 = 8$ skill profiles

[0,0,0]	[1,0,0]
[0,0,1]	[1,1,0]
[0,1,0]	[1,0,1]
[0,1,1]	[1,1,1]

DCM – Basic Example



DCM – Basic Example



DCM – Basic Example

- Items 3, 6, and 9 assess two skills
- Loglinear model (LCDM):
  - Success on these items is modeled as a function of the main effect of each skill and the interaction between the two skills
- Disjunctive model (e.g., DINO):
  - Guessing parameter: those who haven't mastered either skill
  - Slipping parameter: those who have mastered at least one skill
- Conjunctive model (e.g., DINA):
  - Guessing parameter: those who have mastered <2 skills
  - Slipping parameter: those who have mastered both skills

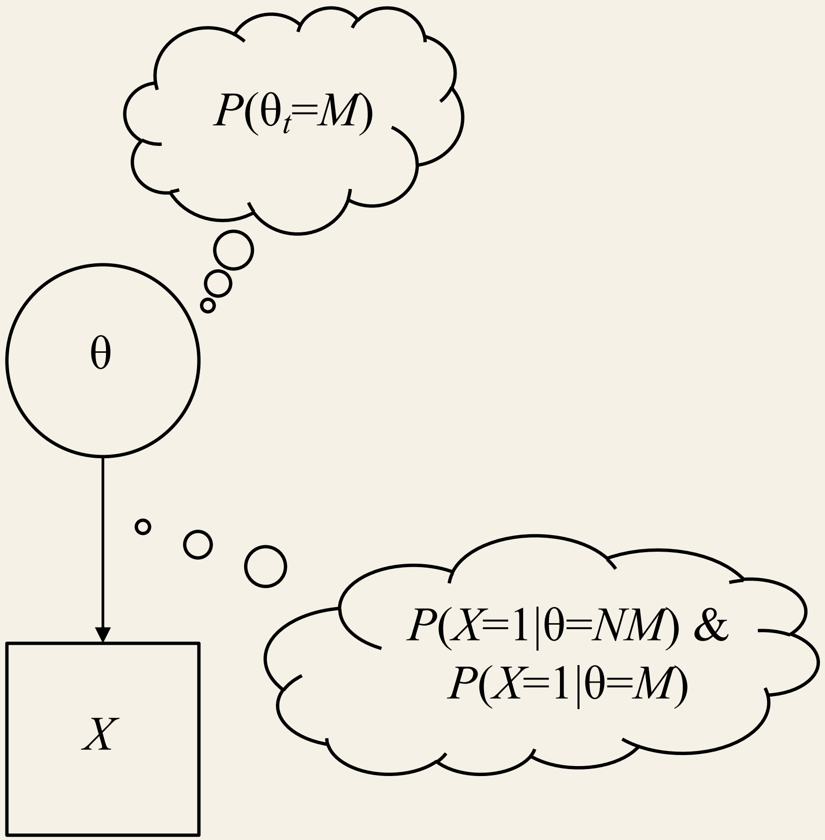
Bayesian Networks

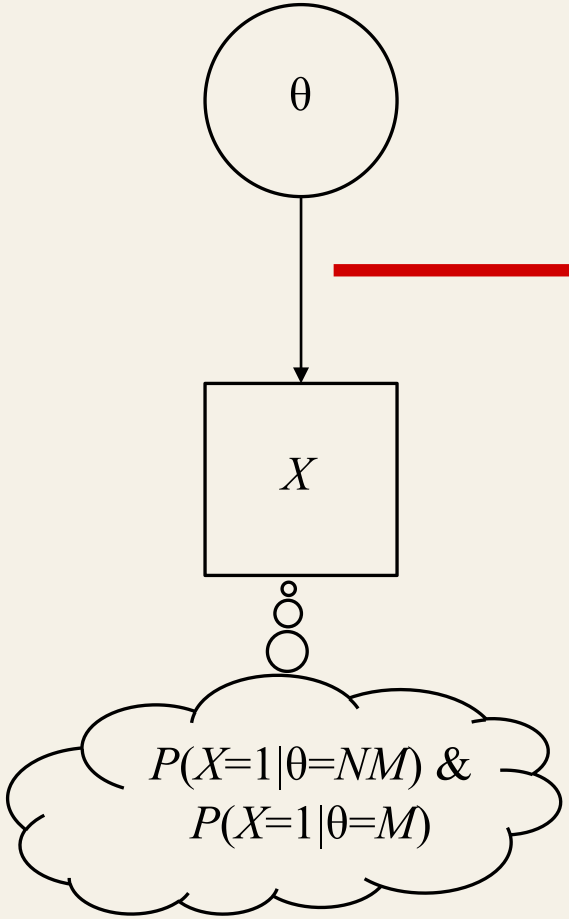
- Bayesian networks (BN) represent a set of conditional dependencies between a collection of random variables.
- More concretely, models the probability of a state conditioned on a set of observed states
- Typically represented as a directed, acyclical graphical model (DAG)
  - Nice, tidy way to represent the joint distribution over the set of variables
- Bayes theorem (Bayes, 1763) provides the mechanism for updating our beliefs as evidence is accumulated

Bayesian Networks

- BNs represent a very general, flexible framework
  - LCA, DCM, state-space models (e.g., hidden Markov models, particle filters), etc. are all a special case of a BN
- Advantages of BNs (Almond et al., 2015):
  - Computationally efficient (exploit conditional independence assumptions)
  - Modular
  - Can handle very complex models/systems of variables
  - Can provide real-time feedback/updating
  - Can be used in mixed-methods designs (e.g., incorporate expert knowledge)
  - And many more!

θ : {*master, non-master*}
X : {*correct, incorrect*}





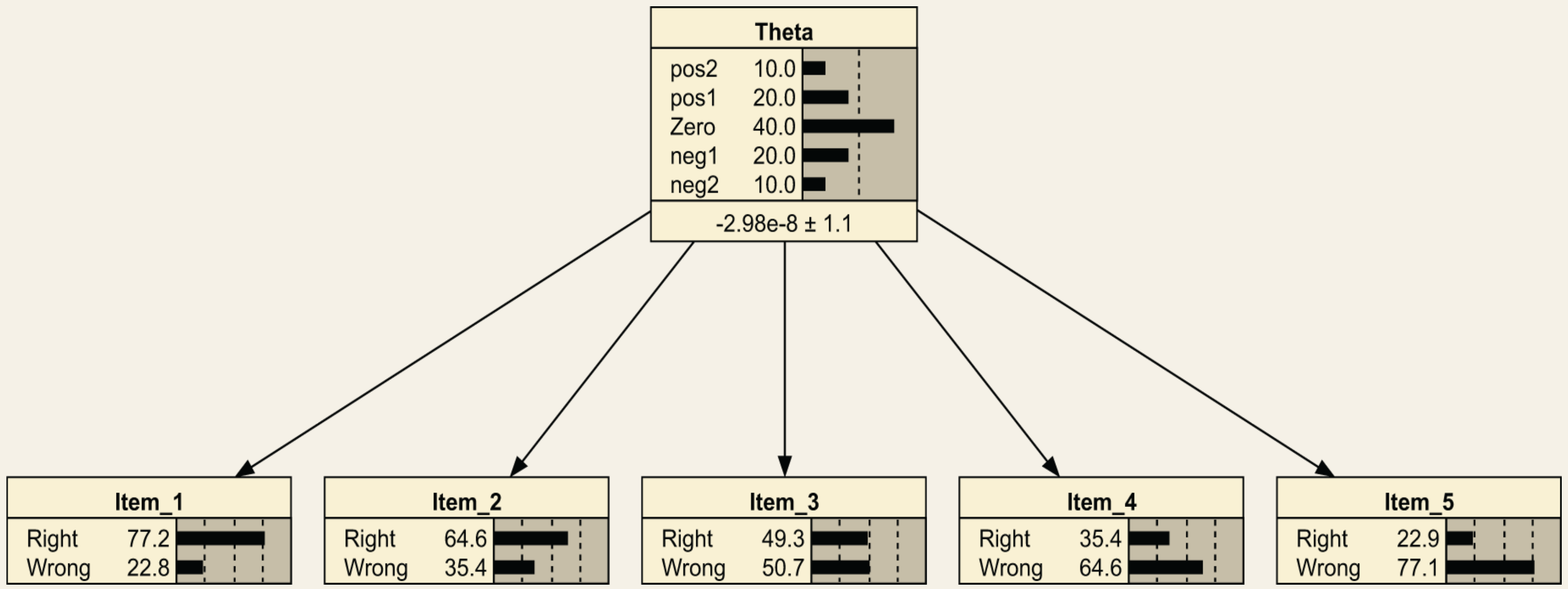
Highly discriminating item; Low uncertainty

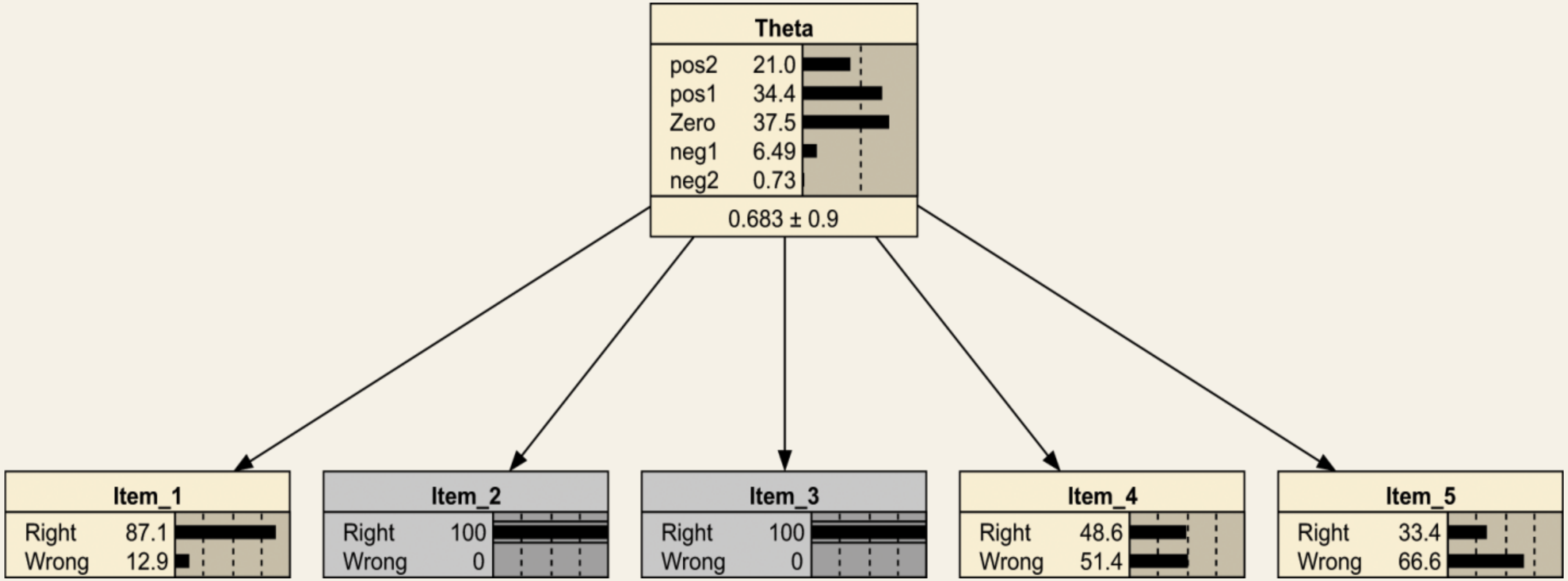
	$\theta = NM$	$\theta = M$
X=0	0.95	0.05
X=1	0.05	0.95

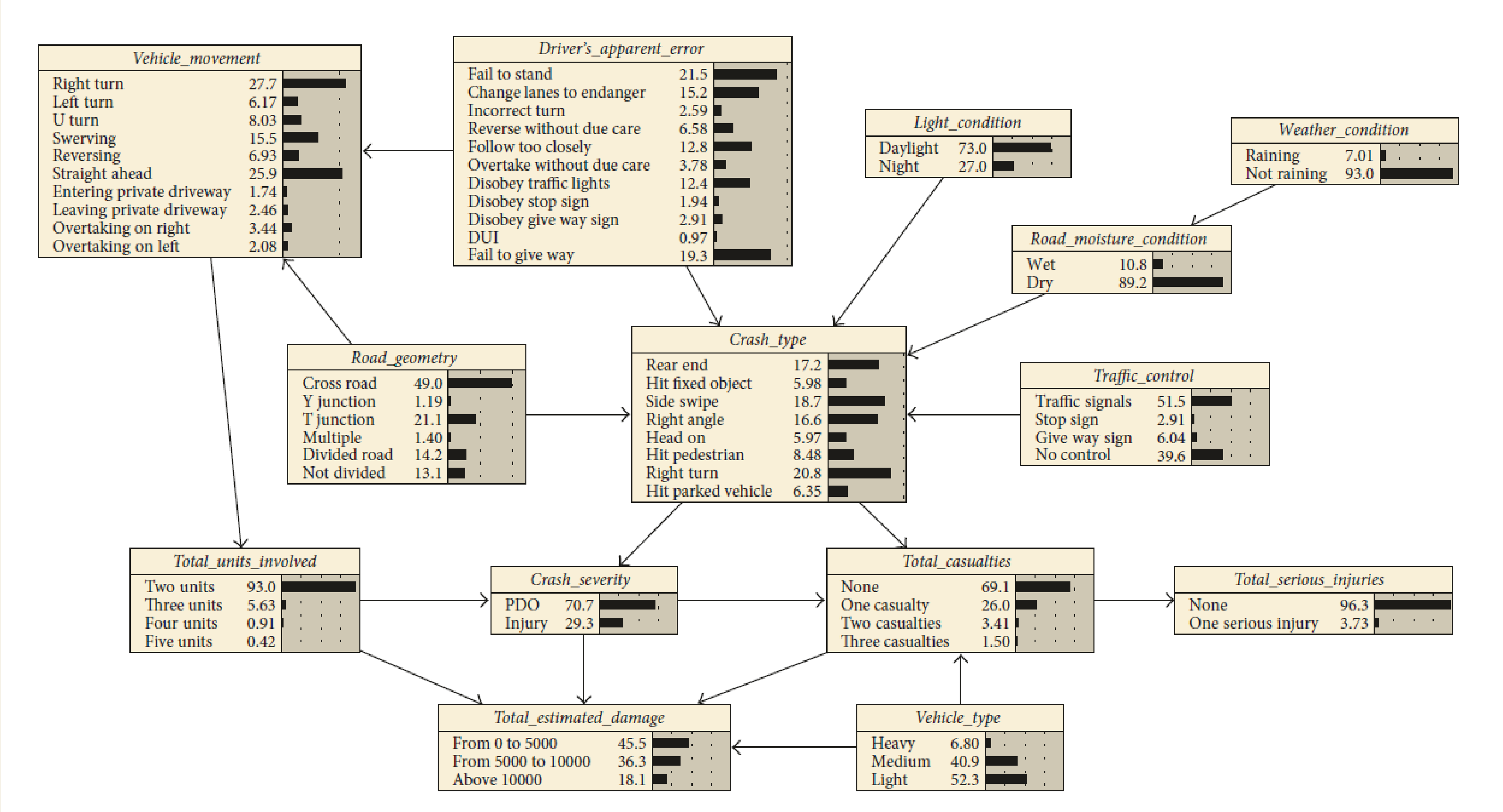
Non-discriminating; maximal uncertainty

	$\theta = NM$	$\theta = M$
X=0	0.50	0.50
X=1	0.50	0.50

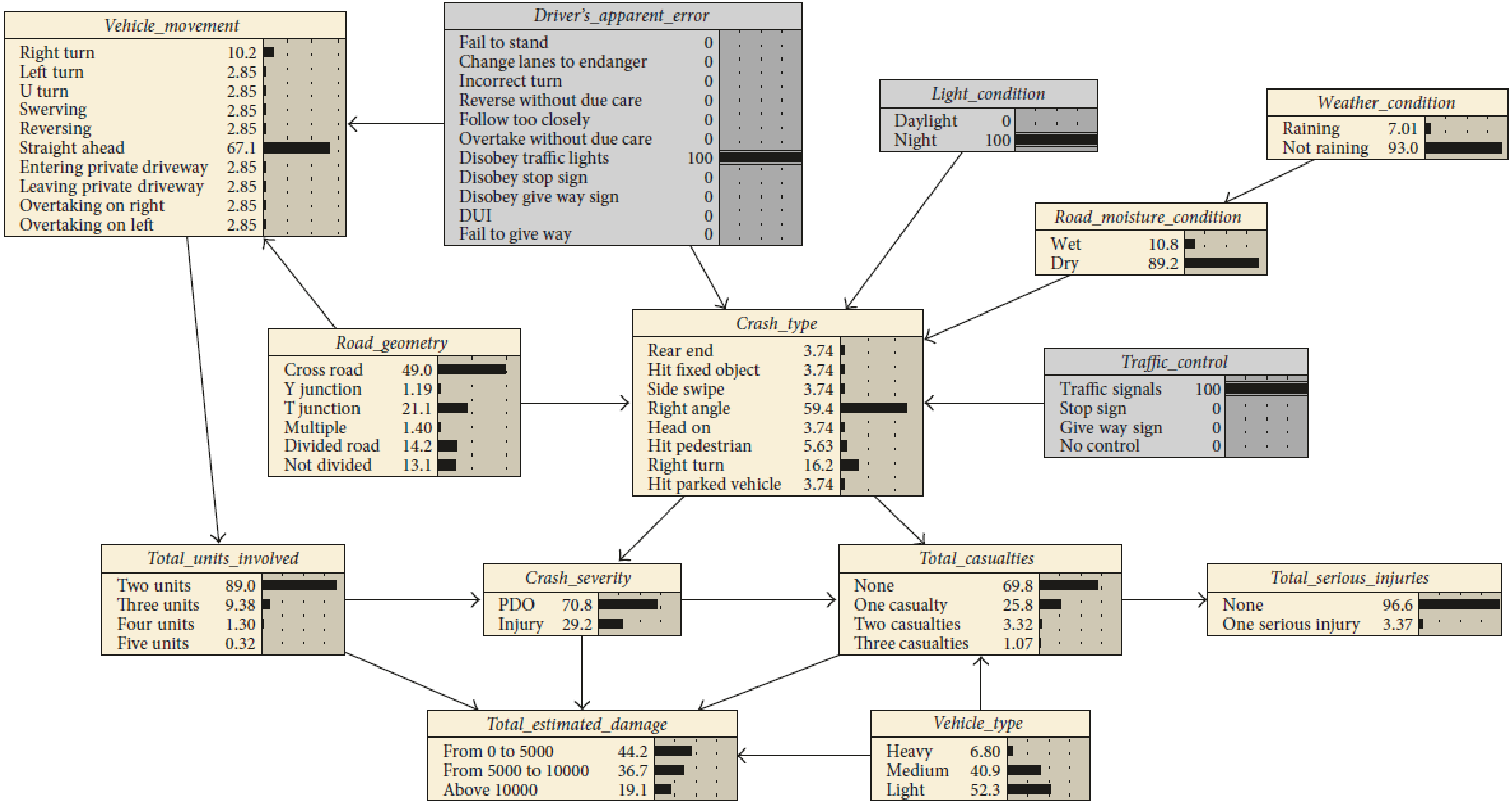
Bayesian Networks - Examples

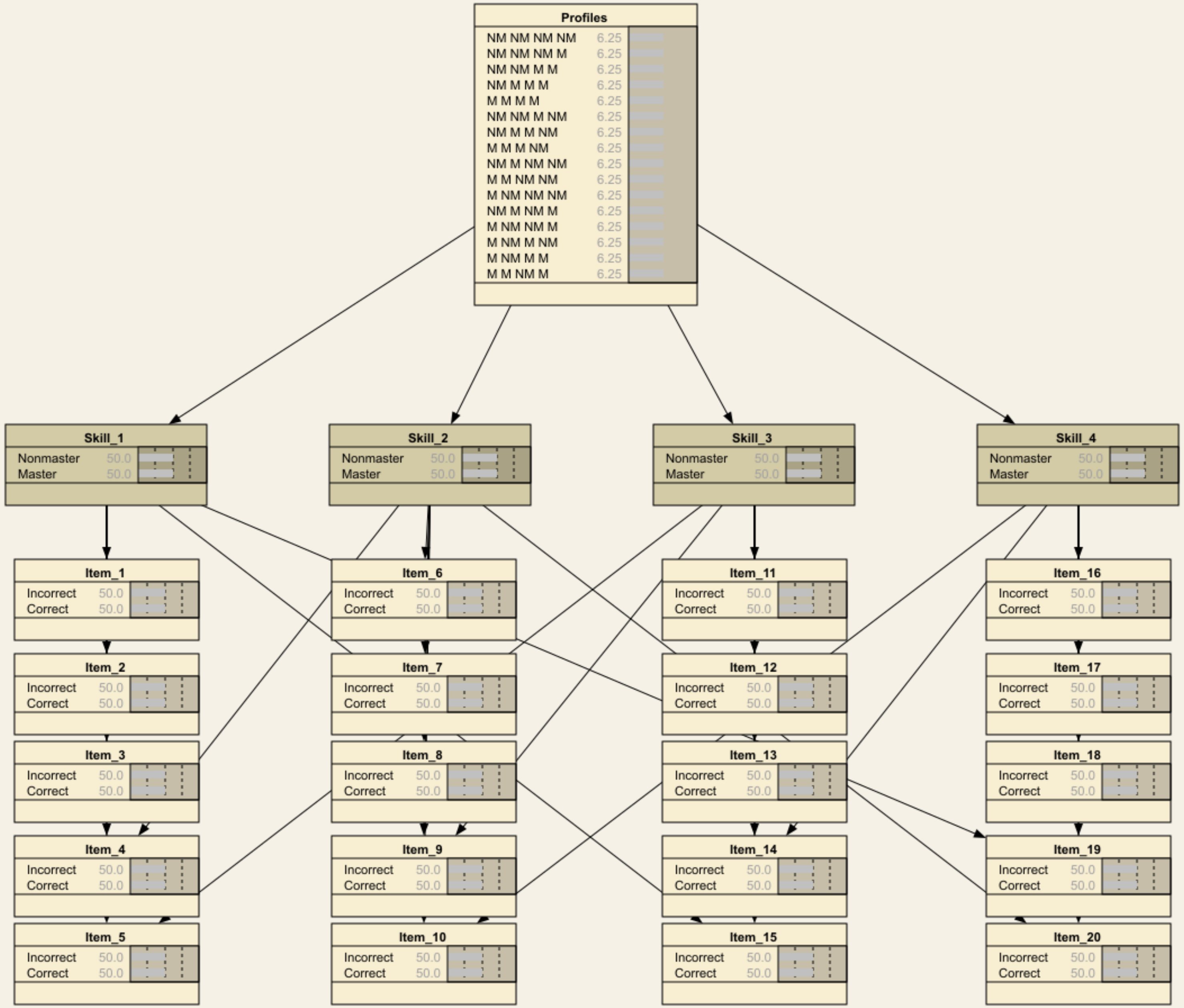



Theta
pos2 21.0
pos1 34.4
Zero 37.5
neg1 6.49
neg2 0.73
0.683 ± 0.9
Item_1
Right 87.1
Wrong 12.9
Item_2
Right 100
Wrong 0
Item_3
Right 100
Wrong 0
Item_4
Right 48.6
Wrong 51.4
Item_5
Right 33.4
Wrong 66.6



Source: Zou & Yue (2017)


Vehicle_movement
Right turn 10.2
Left turn 2.85
U turn 2.85
Swerving 2.85
Reversing 2.85
Straight ahead 67.1
Entering private driveway 2.85
Leaving private driveway 2.85
Overtaking on right 2.85
Overtaking on left 2.85
Driver's_apparent_error
Fail to stand 0
Change lanes to endanger 0
Incorrect turn 0
Reverse without due care 0
Follow too closely 0
Overtake without due care 0
Disobey traffic lights 100
Disobey stop sign 0
Disobey give way sign 0
DUI 0
Fail to give way 0
Light_condition
Daylight 0
Night 100
Weather_condition
Raining 7.01
Not raining 93.0
Road_moisture_condition
Wet 10.8
Dry 89.2
Road_geometry
Cross road 49.0
Y junction 1.19
T junction 21.1
Multiple 1.40
Divided road 14.2
Not divided 13.1
Crash_type
Rear end 3.74
Hit fixed object 3.74
Side swipe 3.74
Right angle 59.4
Head on 3.74
Hit pedestrian 5.63
Right turn 16.2
Hit parked vehicle 3.74
Traffic_control
Traffic signals 100
Stop sign 0
Give way sign 0
No control 0
Total_units_involved
Two units 89.0
Three units 9.38
Four units 1.30
Five units 0.32
Crash_severity
PDO 70.8
Injury 29.2
Total_casualties
None 69.8
One casualty 25.8
Two casualties 3.32
Three casualties 1.07
Total_serious_injuries
None 96.6
One serious injury 3.37
Total_estimated_damage
From 0 to 5000 44.2
From 5000 to 10000 36.7
Above 10000 19.1
Vehicle_type
Heavy 6.80
Medium 40.9
Light 52.3


Profiles
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Skill_1
Nonmaster 50.0
Master 50.0
Skill_2
Nonmaster 50.0
Master 50.0
Skill_3
Nonmaster 50.0
Master 50.0
Skill_4
Nonmaster 50.0
Master 50.0
Item_1
Incorrect 50.0
Correct 50.0
Item_2
Incorrect 50.0
Correct 50.0
Item_3
Incorrect 50.0
Correct 50.0
Item_4
Incorrect 50.0
Correct 50.0
Item_5
Incorrect 50.0
Correct 50.0
Item_6
Incorrect 50.0
Correct 50.0
Item_7
Incorrect 50.0
Correct 50.0
Item_8
Incorrect 50.0
Correct 50.0
Item_9
Incorrect 50.0
Correct 50.0
Item_10
Incorrect 50.0
Correct 50.0
Item_11
Incorrect 50.0
Correct 50.0
Item_12
Incorrect 50.0
Correct 50.0
Item_13
Incorrect 50.0
Correct 50.0
Item_14
Incorrect 50.0
Correct 50.0
Item_15
Incorrect 50.0
Correct 50.0
Item_16
Incorrect 50.0
Correct 50.0
Item_17
Incorrect 50.0
Correct 50.0
Item_18
Incorrect 50.0
Correct 50.0
Item_19
Incorrect 50.0
Correct 50.0
Item_20
Incorrect 50.0
Correct 50.0

DCM vs BN

- DCMs are much more common in practice due to a much longer history in the educational/psychological literature
- BNs are the more flexible and computationally efficient of the two
  - Preferred for very large systems of variables
- There is a rich literature on BN in computer science (e.g., intelligent tutoring systems) and fields such as genomics, but these models are a relatively recent advancement in educational/psychological assessment

DCM vs BN

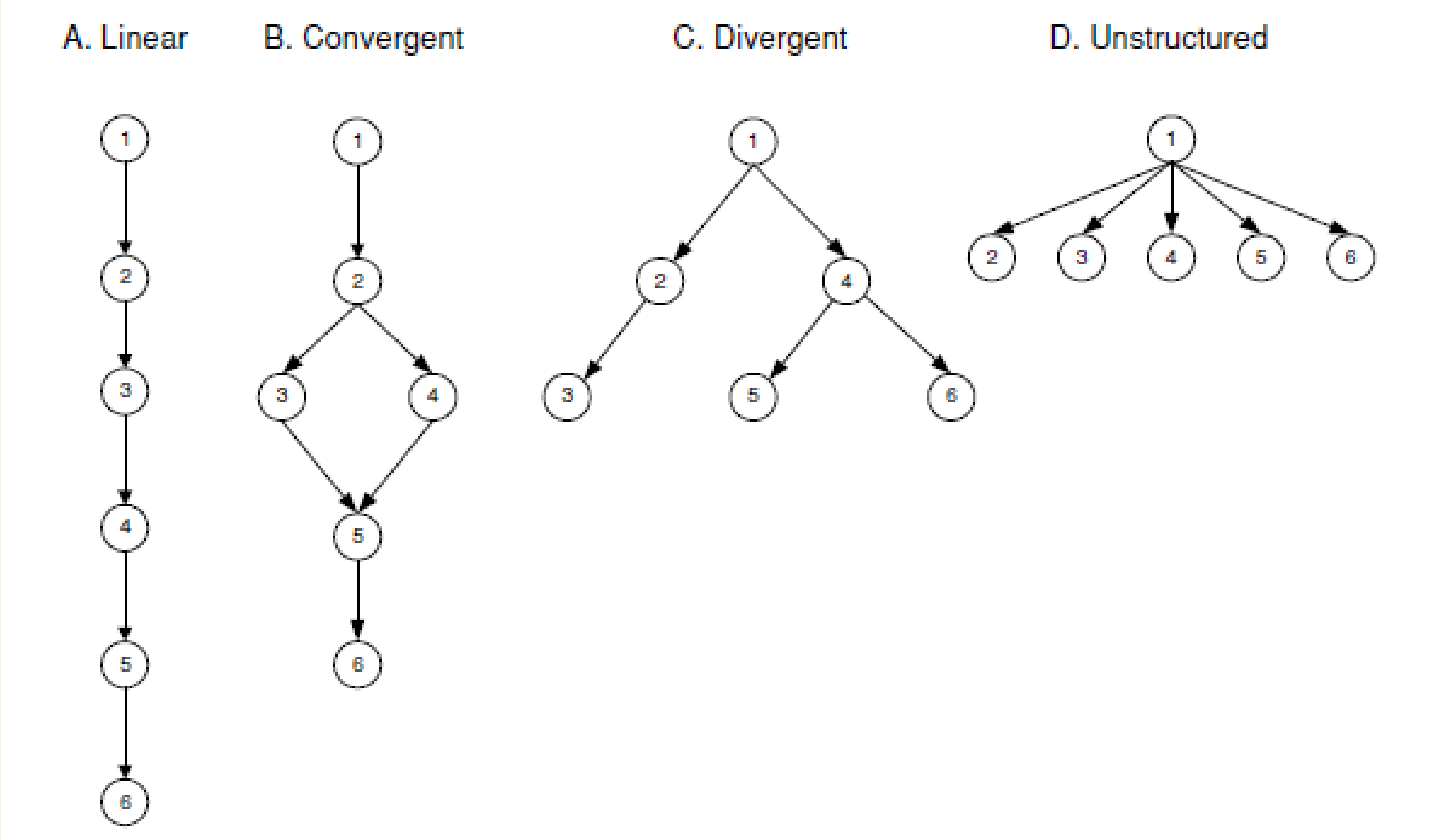
- There may be more opportunity to receive training on DCMs
  - I’m not aware of any graduate courses on BNs being offered in the social sciences
  - Though NCME offers pre-conference workshops on both topics most years
- BNs lend themselves well to graphical representations, which can help to make the resulting inferences more interpretable (Almond et al., 2007)
- For longitudinal applications, the BN framework (i.e., dynamic Bayesian networks; DBN) tends to be preferred*

How can these models be extended?

Modeling Learning Progressions

- "…descriptions of successively more sophisticated ways of thinking about a topic that can follow one another… over time" (National Research Council, 2007)
- There are several ways to operationalize this concept
- Learning progressions are typically represented as attribute hierarchies
- Attribute hierarchies are defined by the relationships between the attributes and specify which attribute profiles should/should not be observed in the population (Rupp, Templin, & Henson, 2010)

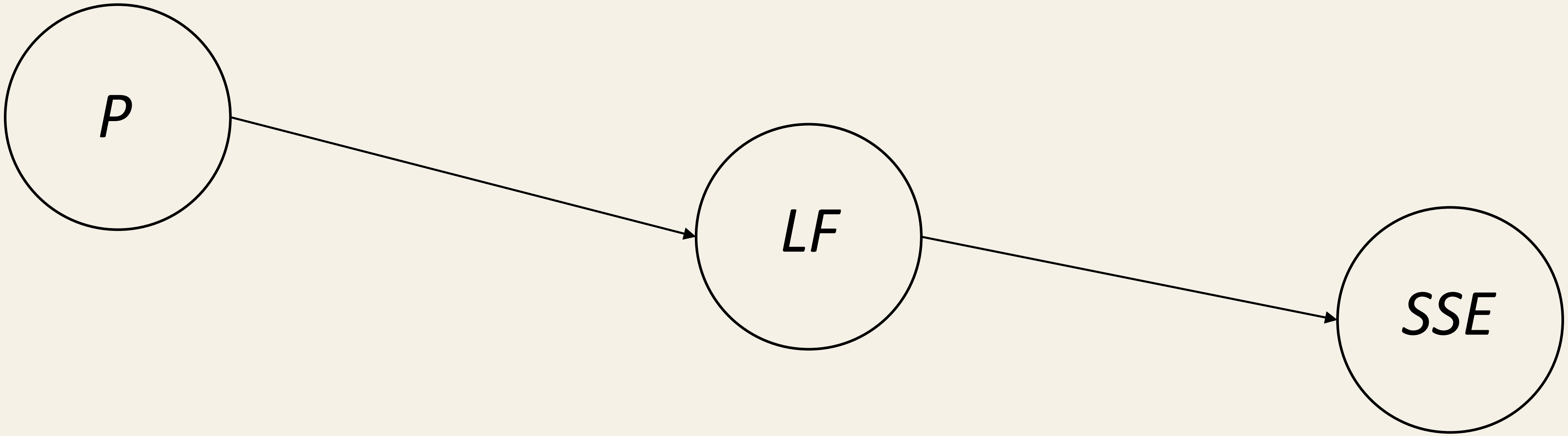
Modeling Learning Progressions



Source: Gierls, Leighton, & Hunka (2007)

Modeling Learning Progressions

- Suppose we have three skills we want to model related to solving linear functions: Prerequisites (*P*), understanding linear functions (*LF*), and solving systems of equations (*SSE*)



Modeling Learning Progressions

- We have $2^3 = 8$ possible profiles:

[0,0,0]	[1,0,0]
[0,0,1]	[1,1,0]
[0,1,0]	[1,0,1]
[0,1,1]	[1,1,1]

Modeling Learning Progressions

- But…

[0,0,0]	[1,0,0]
[0,0,1]	[1,1,0]
[0,1,0]	[1,0,1]
[0,1,1]	[1,1,1]

- We would fix the probability of these profiles to zero

Modeling Growth

- We’re often interested in skill progression, not just skill status.
- Both the DCM and BN frameworks offer longitudinal extensions.
- These extensions model the change in both skill category (e.g., mastery) and attribute profiles.
- Very little research (empirical or methodological) has been done examining these models in psychometric contexts.
  - If you’re a graduate student, this sentence might be important to you.
  - If you think BNs seem exciting and are interested in growth modeling, my office is LPH 269. I have tea available.

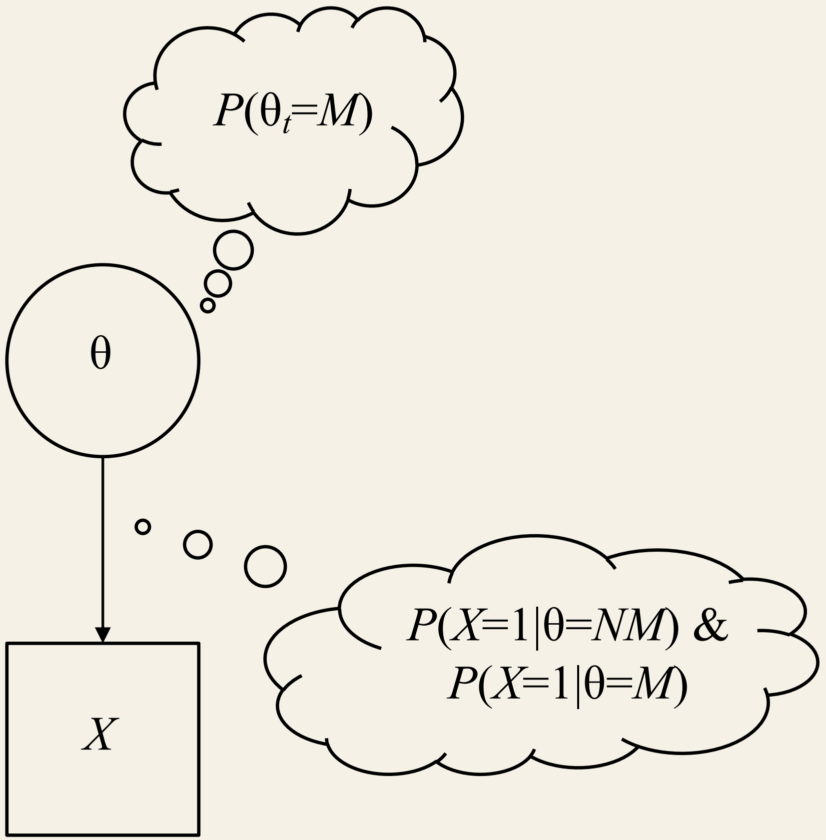
Modeling Growth – L-DCMs

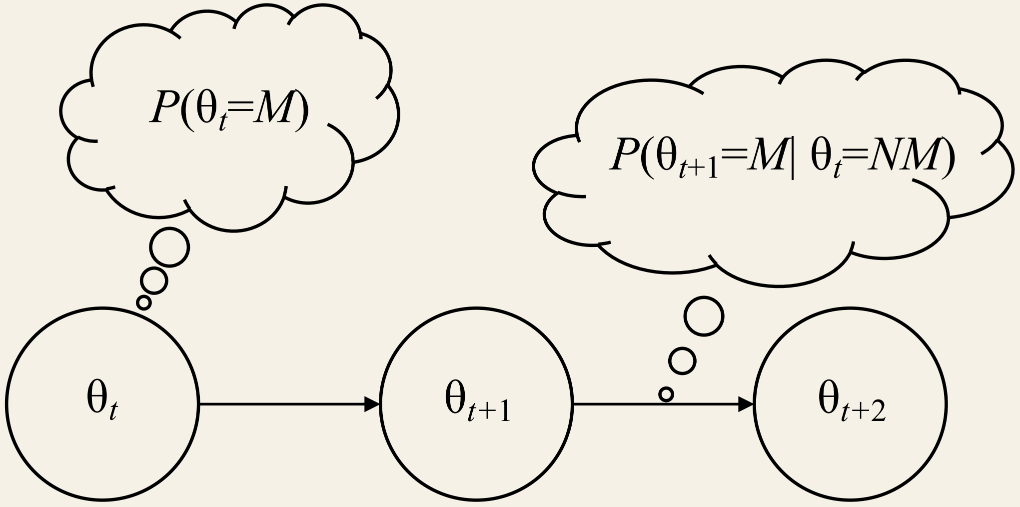
- Latent-transition model-based methods (e.g., Kaya & Liete, 2017; Madison & Bradshaw, 2018)
  - Akin to confirmatory, restricted LTA models
  - Currently only used for pre/post designs (i.e., T=2)
- Higher-order DCMs (e.g., de la Torre & Douglas, 2004; Templin & Bradshaw, 2014)
  - Uses a continuous, higher-order factor
- Multivariate Longitudinal DCM (Pan, Qin, & Kingston, 2020)
  - Loglinear DCM (LDCM) measurement model with a multivariate growth curve component

Modeling Growth – DBNs

- Longitudinal extension of a Bayesian network (BN).
- A series of time-specific BNs connected by a “spine” which bridges the gap between the time slices.
- Can be used to model an individual’s propensity to transition from one profile to another over time.
- Maintain the same advantages as BNs (efficiency, flexibility, etc.; see Reichenberg, 2018)

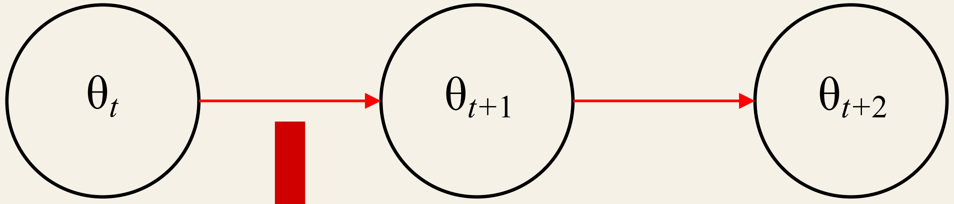
θ : {*master, non-master*}
X : {*correct, incorrect*}





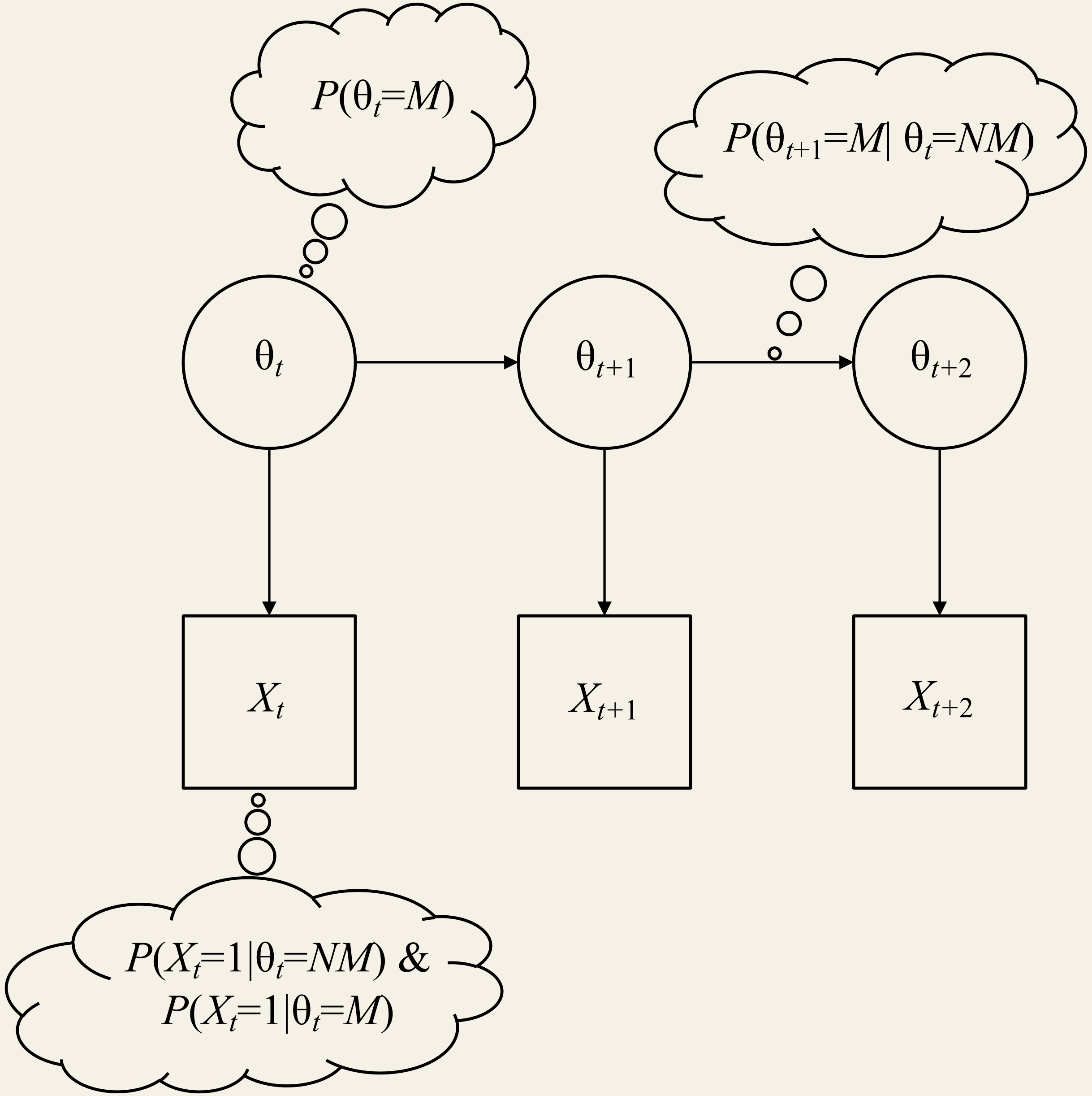
$P(\theta_t = M)$
$P(\theta_{t+1} = M \mid \theta_t = NM)$
θ_t
θ_{t+1}
θ_{t+2}

- The longitudinal aspect of the model is defined by the *transition matrix*, not unlike in a latent transition model.

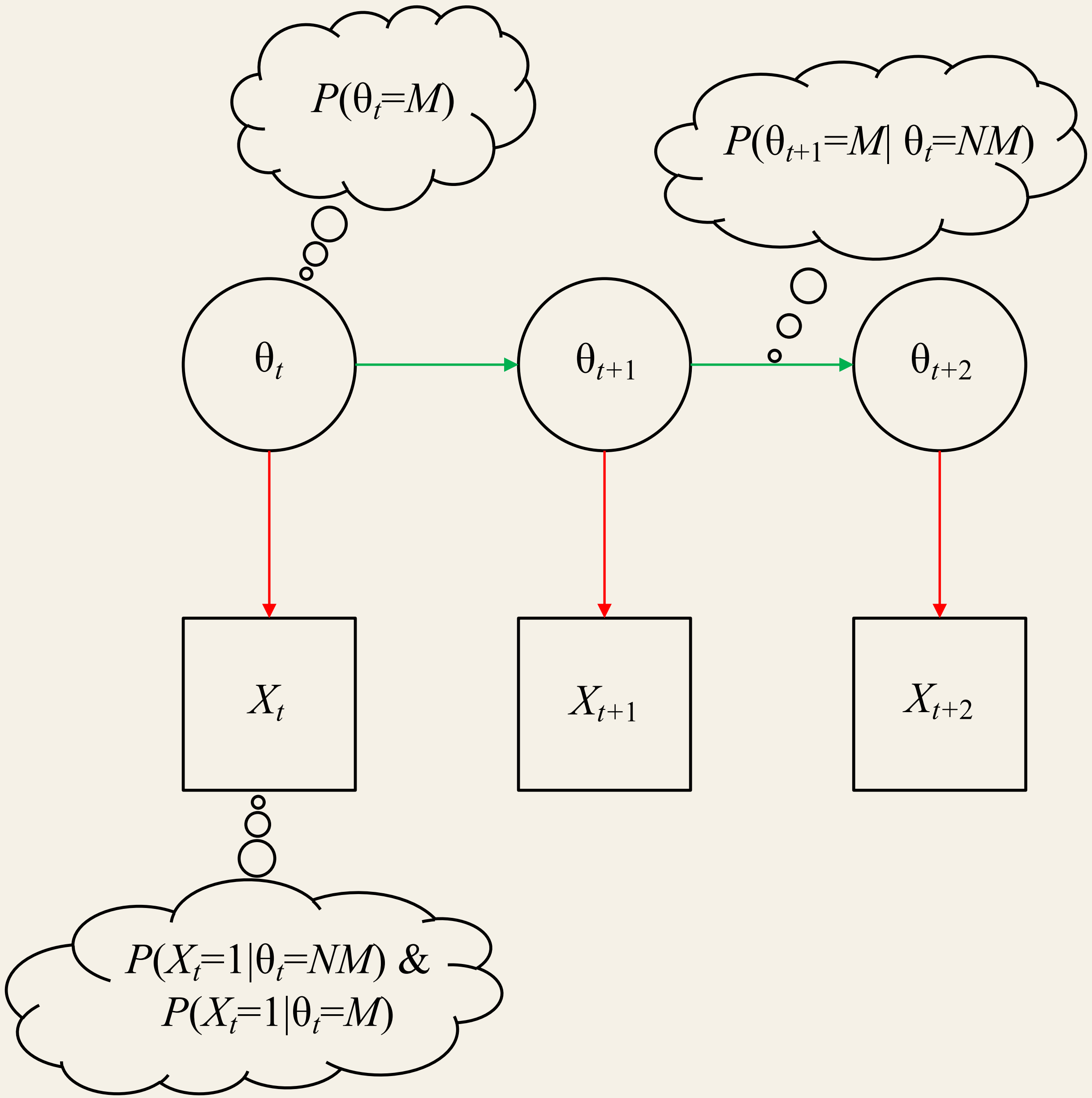


	$\theta_{t+1} = N$	$\theta_{t+1} = M$
$\theta_t = N$	0.80	0.20
$\theta_t = M$	0	1

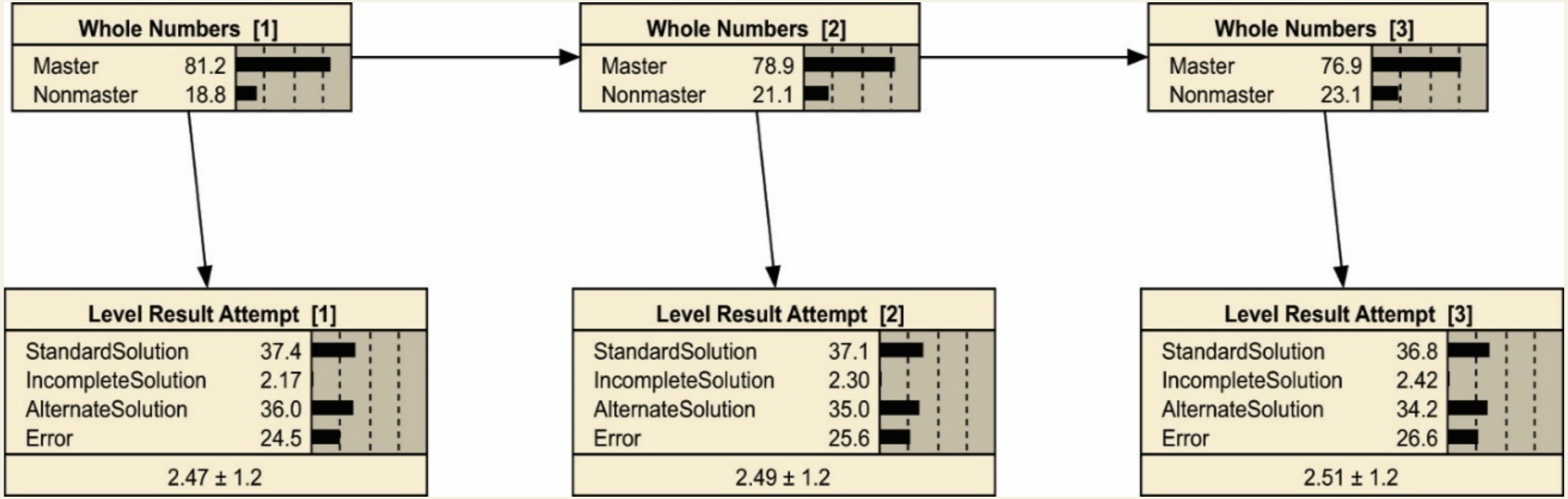
Note. N denotes *non-mastery* while *M* denotes *mastery*.

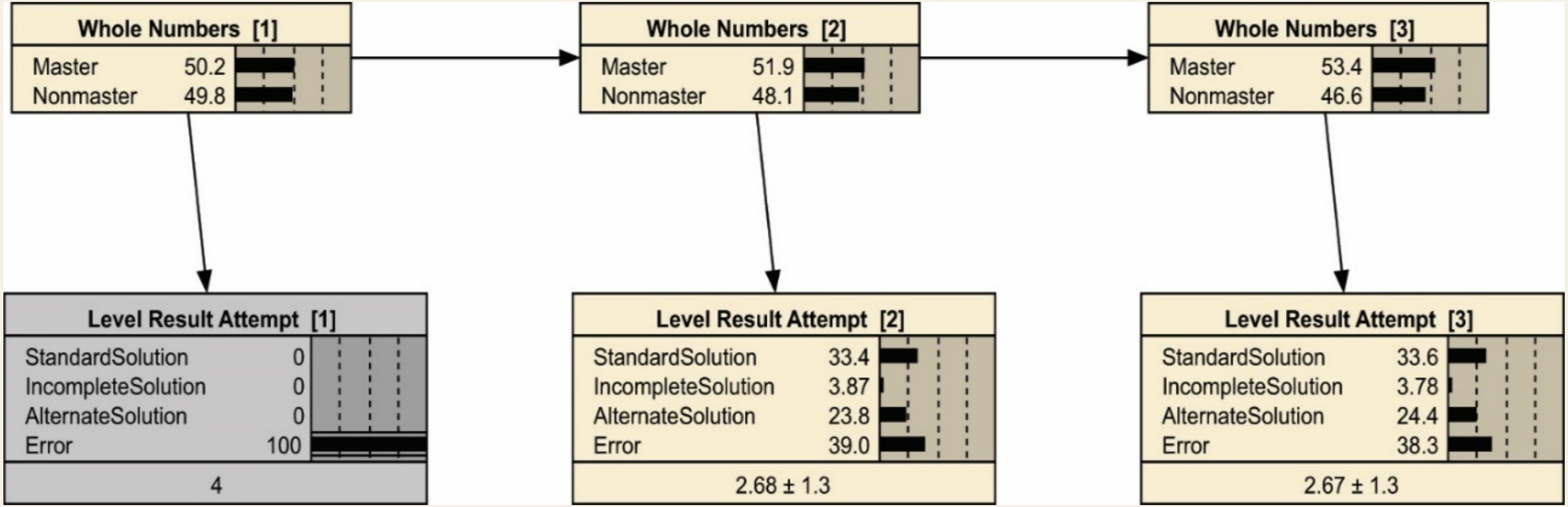


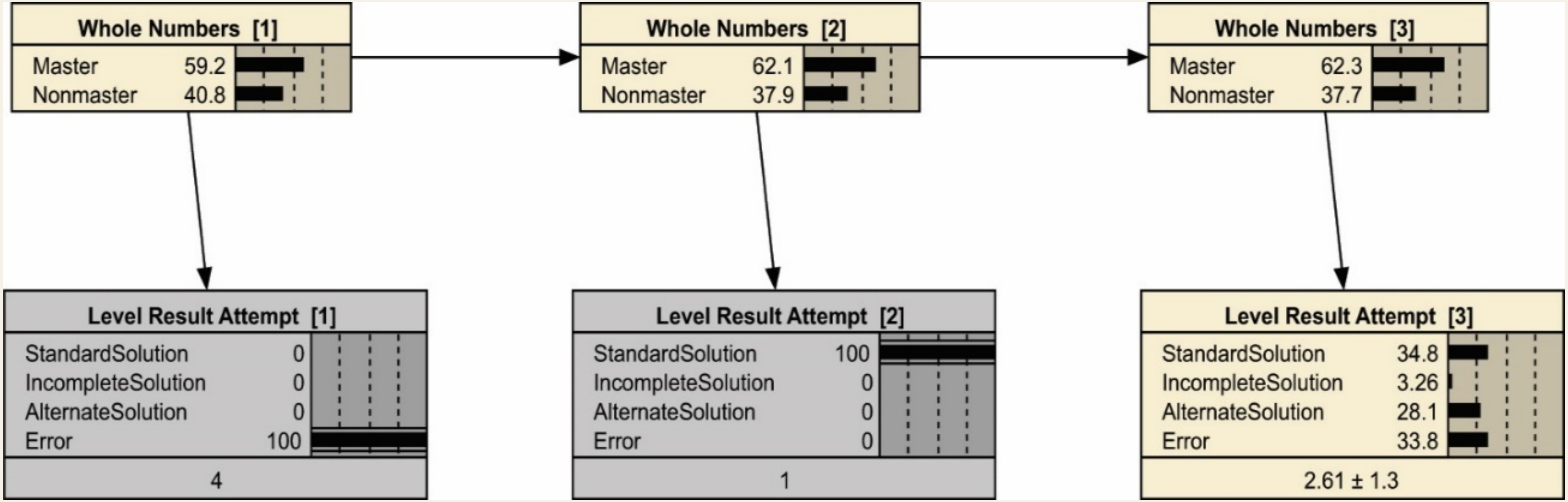
$P(\theta_t\!=\!M)$
$P(\theta_{t+1}\!=\!M|\ \theta_t\!=\!NM)$
θ_t
θ_{t+1}
θ_{t+2}
X_t
X_{t+1}
X_{t+2}
$P(X_t\!=\!1|\theta_t\!=\!NM)$ &
$P(X_t\!=\!1|\theta_t\!=\!M)$

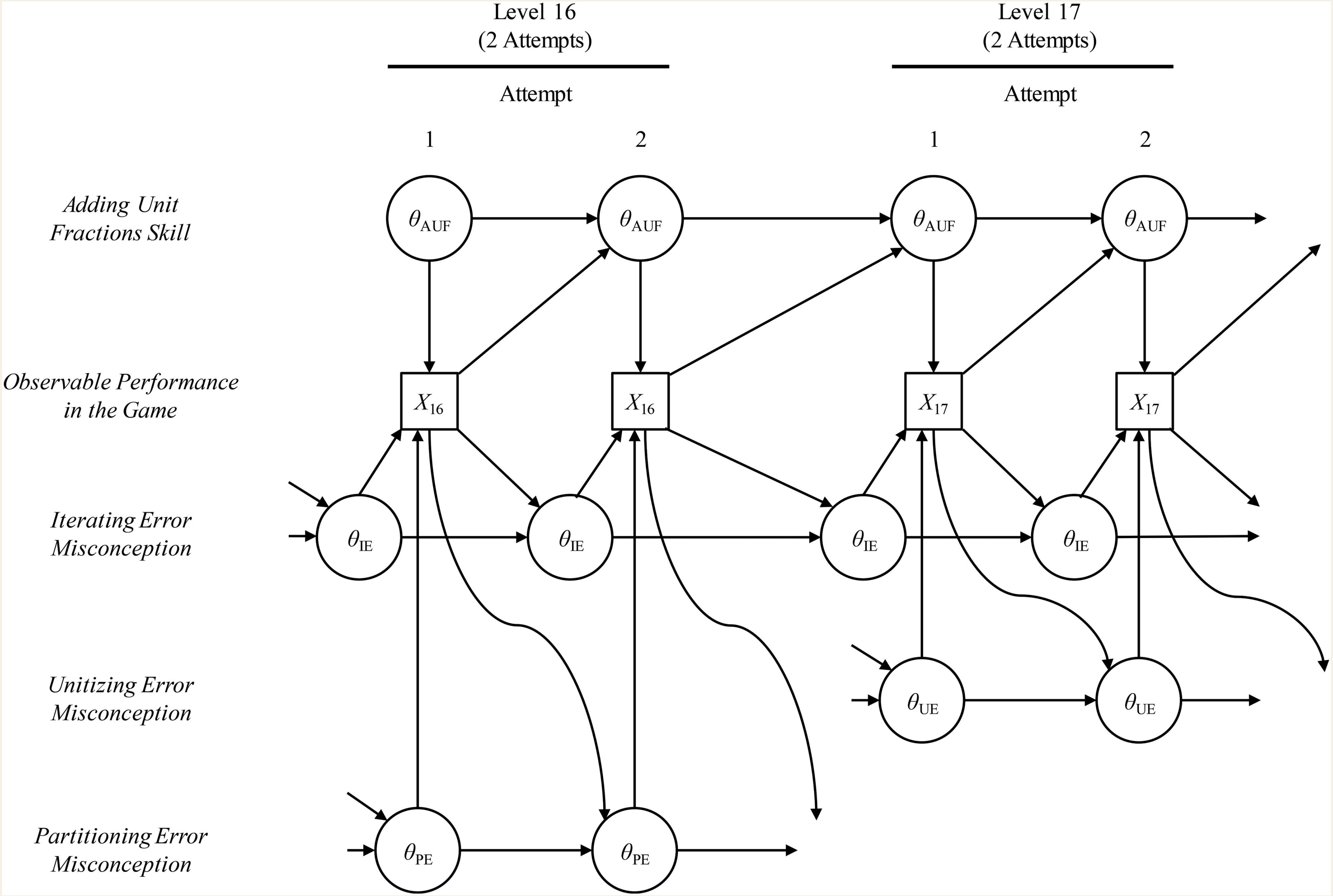

P(θt=M)
P(θt+1=M| θt=NM)
θt
θt+1
θt+2
Xt
Xt+1
Xt+2
P(Xt=1|θt=NM) &
P(Xt=1|θt=M)

Example DBN

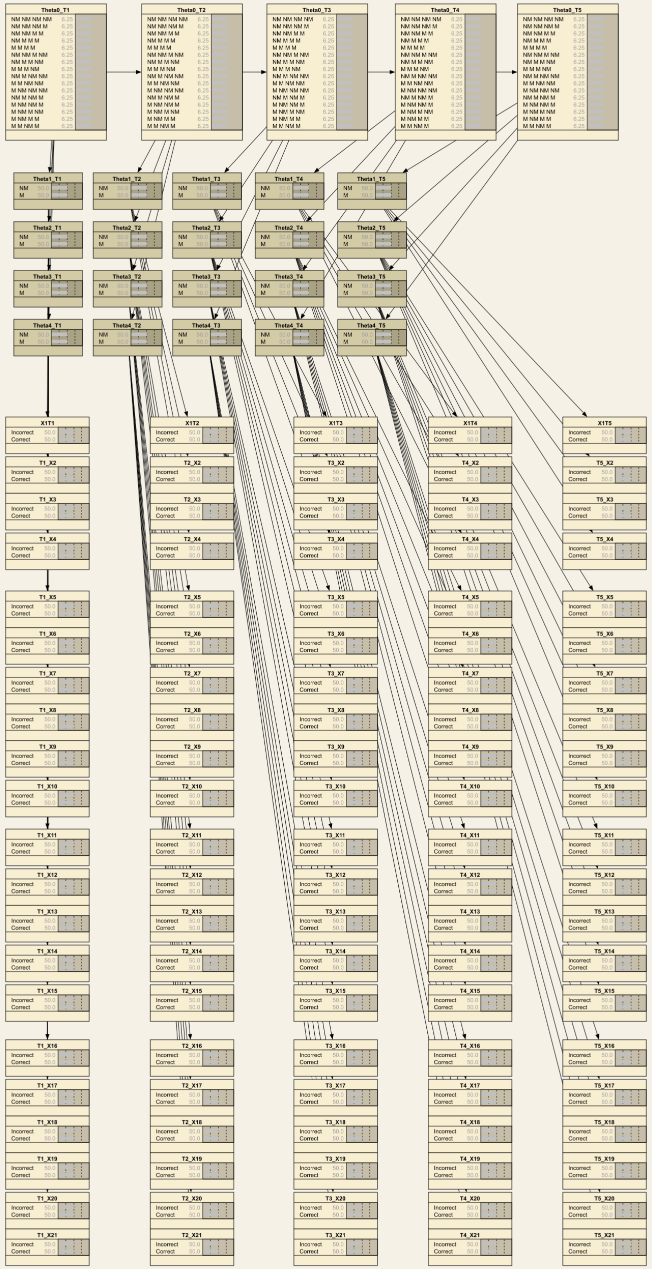

Whole Numbers [1]
Master 81.2
Nonmaster 18.8
Whole Numbers [2]
Master 78.9
Nonmaster 21.1
Whole Numbers [3]
Master 76.9
Nonmaster 23.1
Level Result Attempt [1]
StandardSolution 37.4
IncompleteSolution 2.17
AlternateSolution 36.0
Error 24.5
2.47 ± 1.2
Level Result Attempt [2]
StandardSolution 37.1
IncompleteSolution 2.30
AlternateSolution 35.0
Error 25.6
2.49 ± 1.2
Level Result Attempt [3]
StandardSolution 36.8
IncompleteSolution 2.42
AlternateSolution 34.2
Error 26.6
2.51 ± 1.2

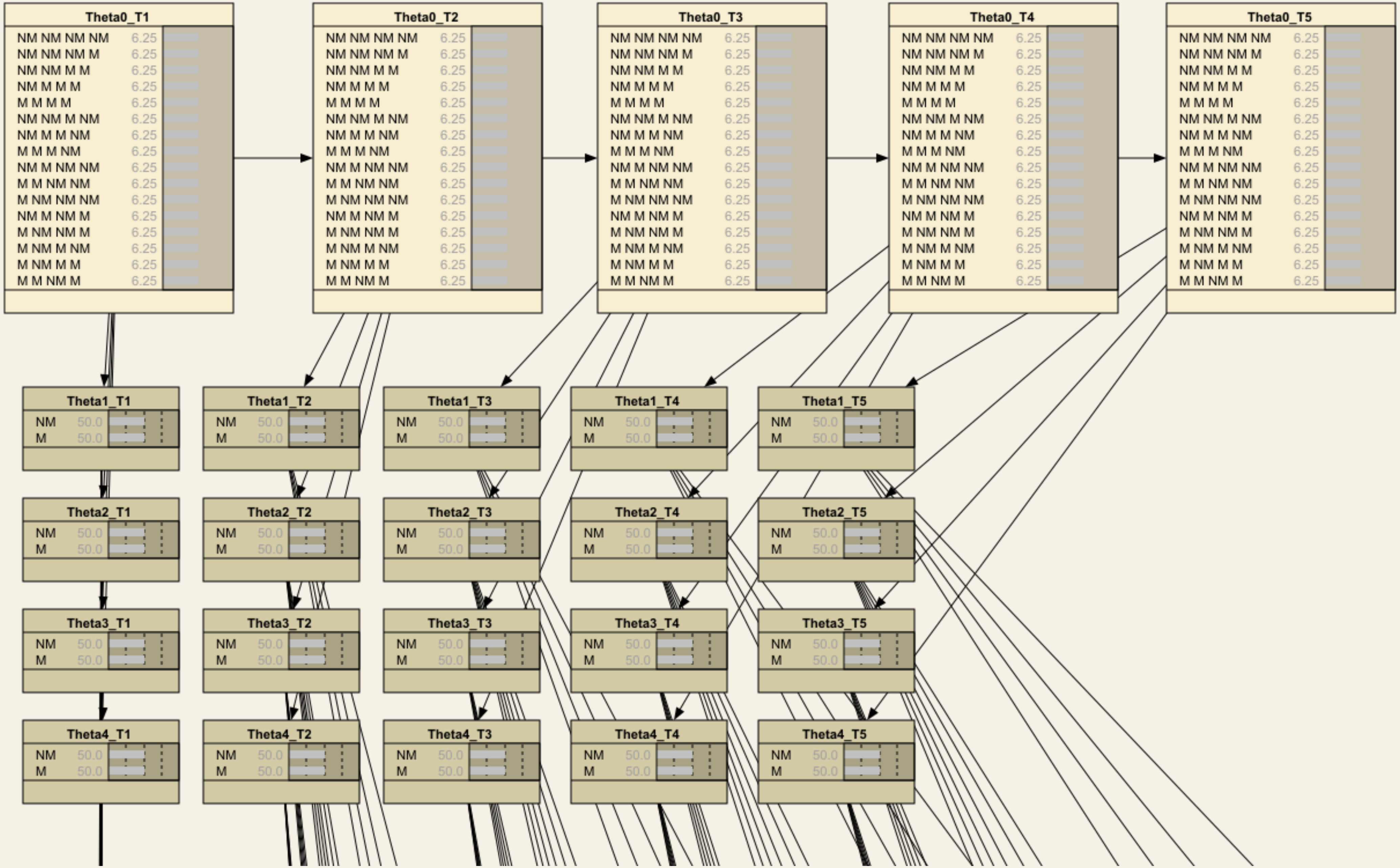

Whole Numbers [1]
Master 50.2
Nonmaster 49.8
Whole Numbers [2]
Master 51.9
Nonmaster 48.1
Whole Numbers [3]
Master 53.4
Nonmaster 46.6
Level Result Attempt [1]
StandardSolution 0
IncompleteSolution 0
AlternateSolution 0
Error 100
4
Level Result Attempt [2]
StandardSolution 33.4
IncompleteSolution 3.87
AlternateSolution 23.8
Error 39.0
2.68 ± 1.3
Level Result Attempt [3]
StandardSolution 33.6
IncompleteSolution 3.78
AlternateSolution 24.4
Error 38.3
2.67 ± 1.3

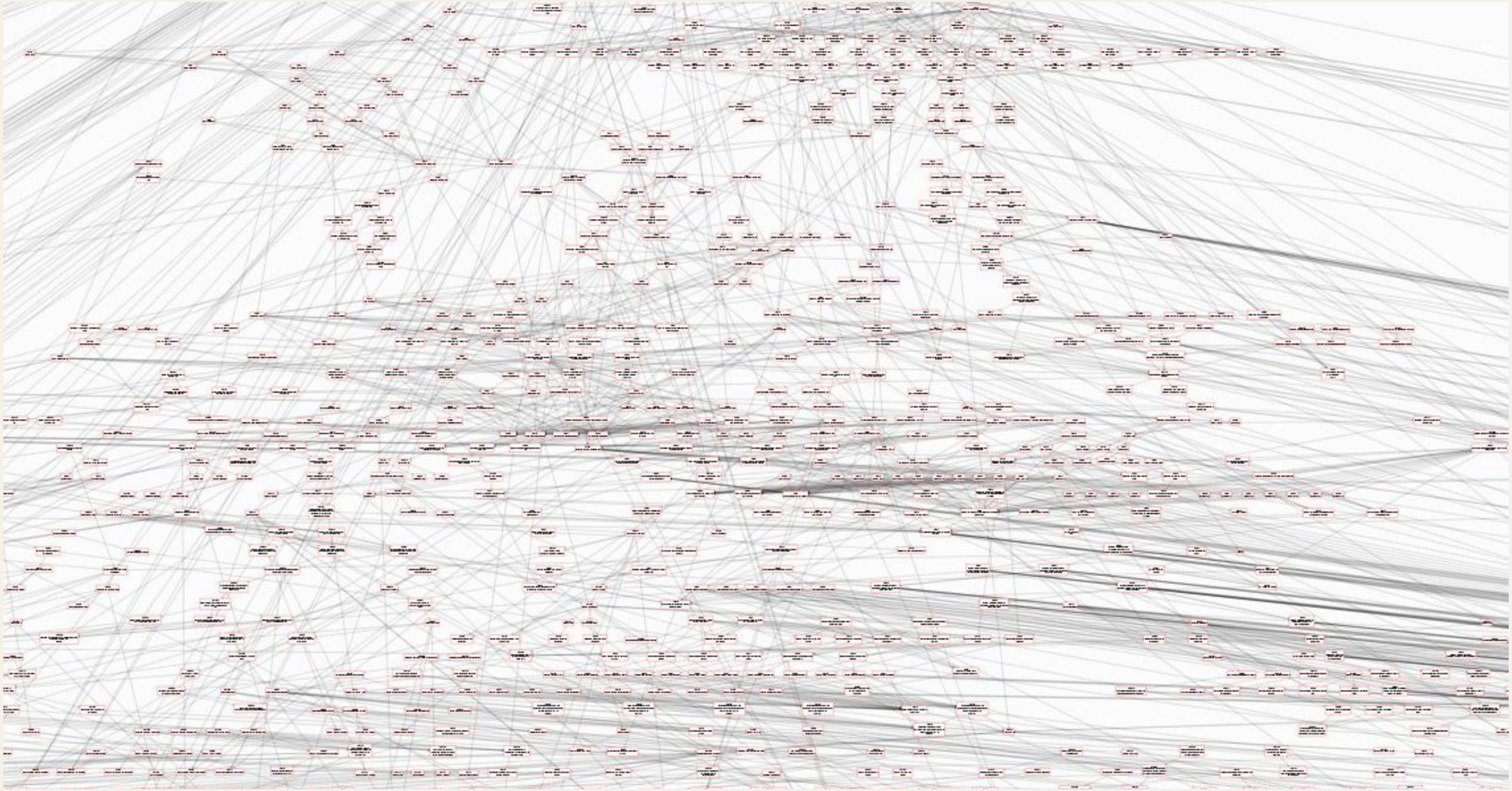

Whole Numbers [1]
Master 59.2
Nonmaster 40.8
Whole Numbers [2]
Master 62.1
Nonmaster 37.9
Whole Numbers [3]
Master 62.3
Nonmaster 37.7
Level Result Attempt [1]
StandardSolution 0
IncompleteSolution 0
AlternateSolution 0
Error 100
4
Level Result Attempt [2]
StandardSolution 100
IncompleteSolution 0
AlternateSolution 0
Error 0
1
Level Result Attempt [3]
StandardSolution 34.8
IncompleteSolution 3.26
AlternateSolution 28.1
Error 33.8
2.61 ± 1.3



Source: Levy (2019)




Theta0_T1
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Theta0_T2
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Theta0_T3
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Theta0_T4
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Theta0_T5
NM NM NM NM 6.25
NM NM NM M 6.25
NM NM M M 6.25
NM M M M 6.25
M M M M 6.25
NM NM M NM 6.25
NM M M NM 6.25
M M M NM 6.25
NM M NM NM 6.25
M M NM NM 6.25
M NM NM NM 6.25
NM M NM M 6.25
M NM NM M 6.25
M NM M NM 6.25
M NM M M 6.25
M M NM M 6.25
Theta1_T1
NM 50.0
M 50.0
Theta1_T2
NM 50.0
M 50.0
Theta1_T3
NM 50.0
M 50.0
Theta1_T4
NM 50.0
M 50.0
Theta1_T5
NM 50.0
M 50.0
Theta2_T1
NM 50.0
M 50.0
Theta2_T2
NM 50.0
M 50.0
Theta2_T3
NM 50.0
M 50.0
Theta2_T4
NM 50.0
M 50.0
Theta2_T5
NM 50.0
M 50.0
Theta3_T1
NM 50.0
M 50.0
Theta3_T2
NM 50.0
M 50.0
Theta3_T3
NM 50.0
M 50.0
Theta3_T4
NM 50.0
M 50.0
Theta3_T5
NM 50.0
M 50.0
Theta4_T1
NM 50.0
M 50.0
Theta4_T2
NM 50.0
M 50.0
Theta4_T3
NM 50.0
M 50.0
Theta4_T4
NM 50.0
M 50.0
Theta4_T5
NM 50.0
M 50.0



Source: www.dynamiclearningmaps.org

Software

Software Options - DCMs

- Mplus (see Templin & Hoffman, 2013)
- flexMIRT
- R
  - *CDM* package
  - *GDINA* package
- von Davier & Lee (2019) offers chapters specific to each of these options

Software Options - BNs

- Netica ($)
- BayesiaLab ($)
- Hugin ($)
- GeNIe
- BayesServer ($)
- MSNBx (no longer supported)

Software Options - BNs

- R
  - bnlearn
  - GeNIe has an R interface (rSMILE) for use with their SMILE API
  - Netica can interface with R using the rNETICA package (limited support)
  - Other API (e.g., BayesiaLab) can be used by calling Java/C from with R
- Python
  - BayesPy
- Probabilistic Programming Languages
  - JAGS
  - WinBUGS/OpenBUGS
  - Stan

Where can I find more information?

Resources

- Textbooks:

  - Rupp, A.A., Templin, J. & Henson, R.A. (2010). *Diagnostic Measurement: Theory, Methods, and Applications*. Guilford.
  - von Davier, M. & Lee, Y.S. (Eds.). (2019). *Handbook of Diagnostic Classification Models: Models and Model Extensions, Applications, Software Packages*. Springer.
  - Almond, R.G., Mislevy, R.J., Steinberg, L.S., & Yan. D. (2015). *Bayesian Networks in Educational Assessment*. Springer.
  - Nielsen, T.D. & Jensen, F.V. (2009). *Bayesian Networks and Decision Graphs*. Springer.

- Diagnostic measurement public database
- BN examples
- Mplus (.inp) / R (.R) code for simple DCM and Netica (.dne) / GeNIe (.xdsl) file for equivalent BN
  - Sample data for model fitting (.csv)

References

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian Networks in Educational Assessment*. Springer.

Almond, R. G., Shute, V. J., Underwood, J. S., & Zapata-Rivera, J.-D. (2009). Bayesian networks: A teacher's view. *International Journal of Approximate Reasoning, 50*(3), 450–460.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. Philosophical Transactions of the Royal Society of London, 53, 370–418.

Carragher, N., Templin, J., Jones, P., Shulruf, B., & Velan, G. (2019). Digital Module 04: Diagnostic Measurement: Modeling Checklists for Practitioners *Educational Measurement: Issues and Practice, 38*(1), 89–91.

National Research Council (2007). Taking science to school: Learning and teaching science in grades K-8. National Academies Press.

de la Torre, J., Carmona, G., Kieftenbeld, V., Tjoe, H., & Lima, C. (2017). Diagnostic classification models and mathematics education research: Opportunities and challenges. *Psychometric Methods in Mathematics Education: Opportunities, Challenges, and Interdisciplinary Collaborations*.

References

De La Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*(3), 333–353.

Gierl, M. J., Leighton, J. P., & Hunka, S. (2007). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills. *Cognitive Diagnostic Assessment for Education: Theory and Applications*, 242–274.

Jang, E. E. (2005). A validity narrative: Effects of reading skills diagnosis on teaching and learning in the context of NG TOEFL. University of Illinois at Urbana-Champaign Illinois, USA.

Kaya, Y., & Leite, W. L. (2017). Assessing change in latent skills across time with longitudinal cognitive diagnosis modeling: An evaluation of model performance. *Educational and Psychological Measurement, 77*(3), 369–388.

Levy, R. (2019). Dynamic Bayesian network modeling of game-based diagnostic assessments. *Multivariate Behavioral Research, 54*(6), 771–794.

Madison, M. J., & Bradshaw, L. P. (2018). Assessing growth in a diagnostic classification model framework. *Psychometrika, 83*(4), 963–990.

Pan, Q., Qin, L., & Kingston, N. (2020). Growth Modeling in a Diagnostic Classification Model (DCM) Framework–A Multivariate Longitudinal Diagnostic Classification Model. *Frontiers in Psychology, 11*, 1714.

References

Reichenberg, R. (2018). Dynamic Bayesian networks in educational measurement: Reviewing and advancing the state of the field. *Applied Measurement in Education, 31*(4), 335–350.

Templin, J., & Bradshaw, L. (2014). Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies. *Psychometrika, 79*(2), 317–339.

Templin, J., & Henson, R. A. (2010). *Diagnostic Measurement: Theory, Methods, and Applications*. Guilford Press.

Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice, 32*(2), 37–50.

von Davier, M., & Lee, Y.-S. (2019). *Handbook of Diagnostic Classification Models*. Springer.

Welcome to DLM | DLM. (n.d.). Retrieved February 9, 2021, from https://dynamiclearningmaps.org/

Zou, X., & Yue, W. L. (2017). A Bayesian network approach to causation analysis of road accidents using Netica. *Journal of Advanced Transportation*, 2017.