An Overview of
Item Response Tree (IRTree) Models

1

Section Learning Objectives

1 An Overview of IRTree Models

Understand the conceptual framework of item response tree (IRTree) models

Specify a tree structure based on a hypothesized internal decision process

Transform observed item responses into pseudo-item responses

Show how to compute the probability of a terminal response based on a specified IRTree model

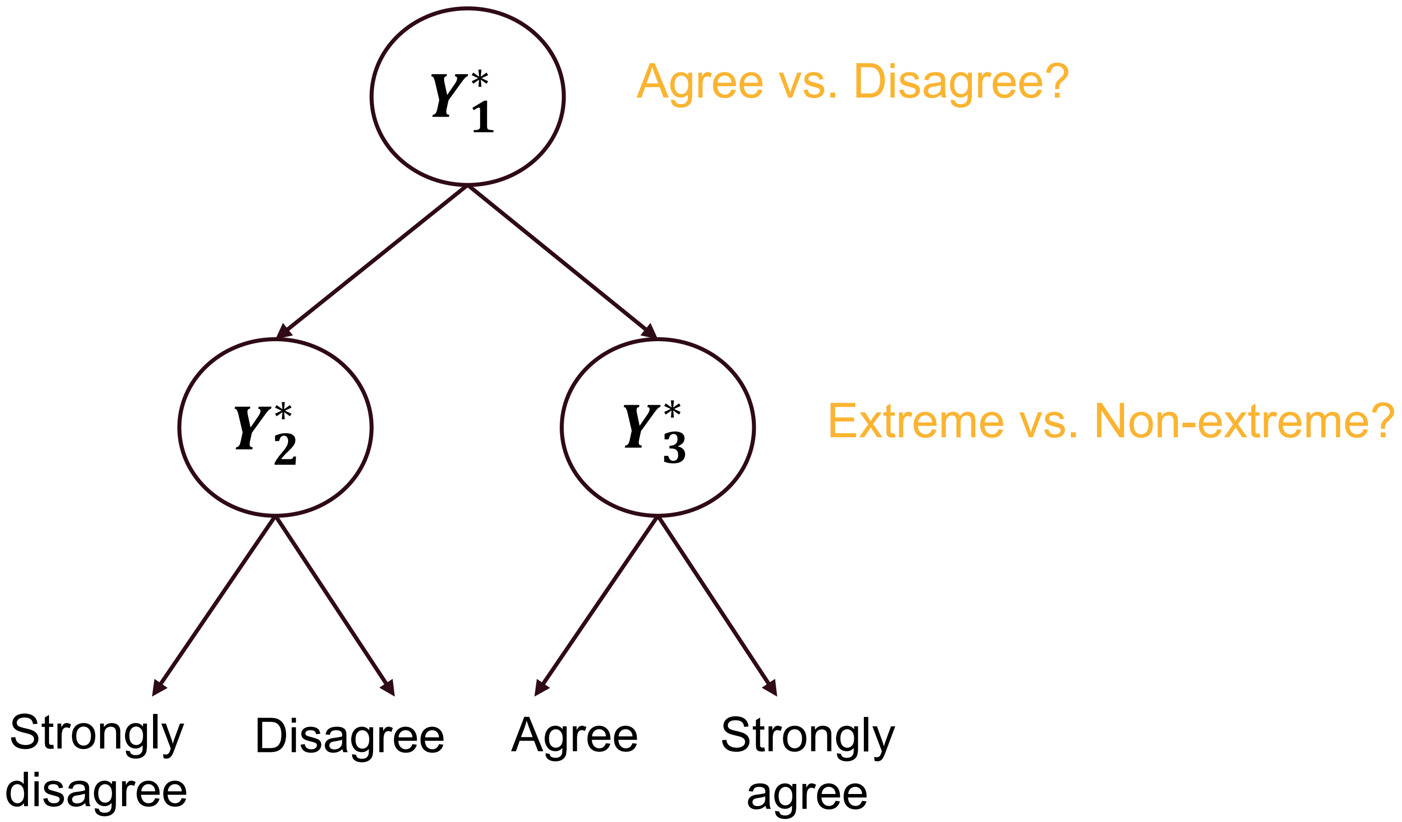
Basic Idea

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. I get stressed out easily.				
2. I am relaxed most of the time.				
3. I feel comfortable around people.				
⋮		⋮		

- Typical item response theory (IRT) models use terminal outcomes (i.e., observed item responses)
- Item response tree (IRTree) models (Böckenholt, 2012; De Boeck & Partchev, 2012; Jeon & De Boeck, 2016) can describe psychological or cognitive processes underlying item responses

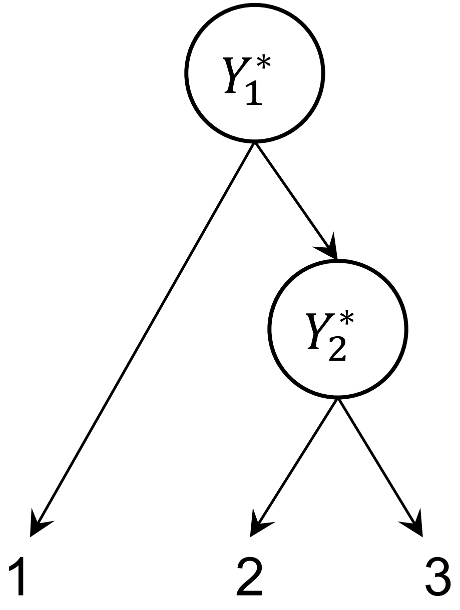
Basic Idea

- IRTree model decomposes a hypothesized response process using a decision tree structure

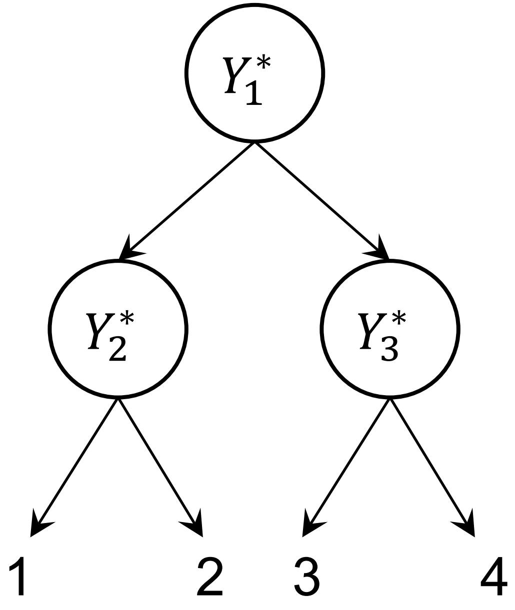


Basic Idea

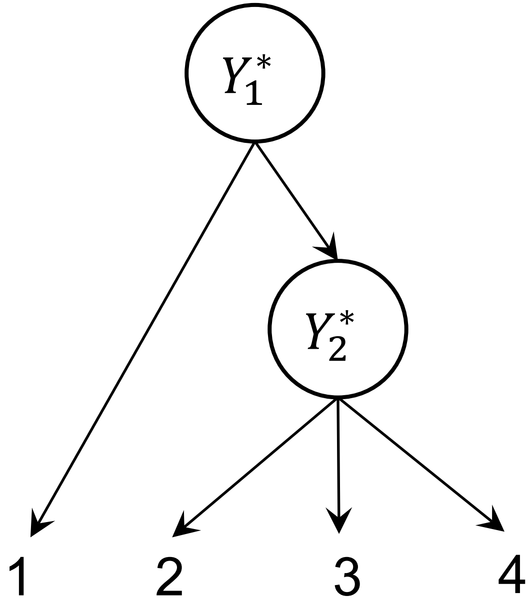
- Example tree structures:



Linear



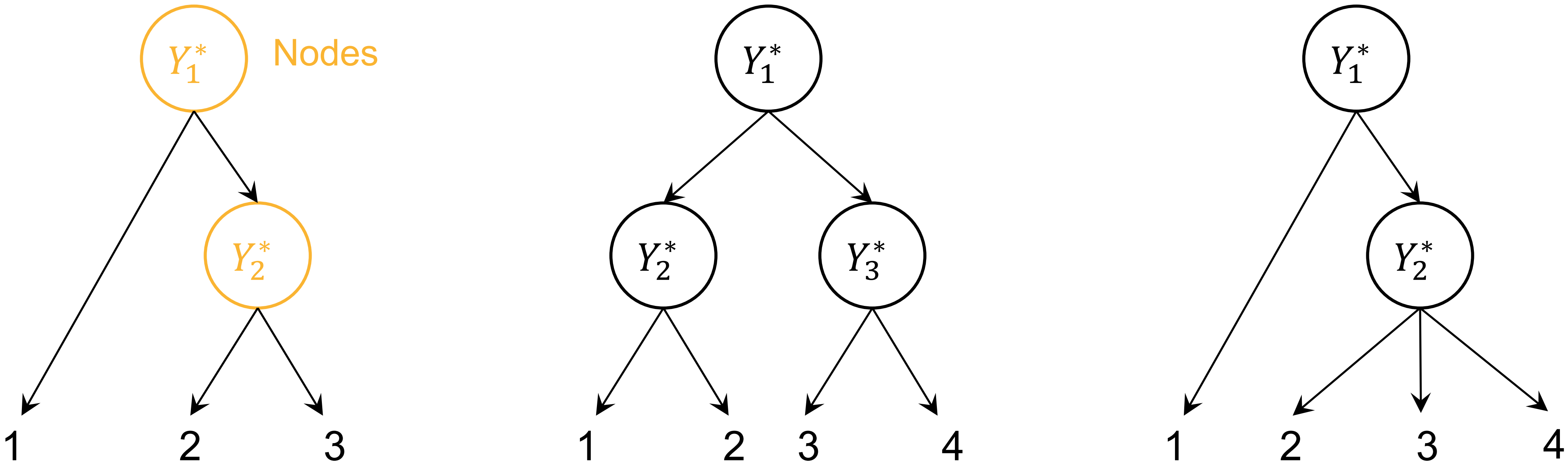
Nested



Polytomous

Basic Idea

- Example tree structures:



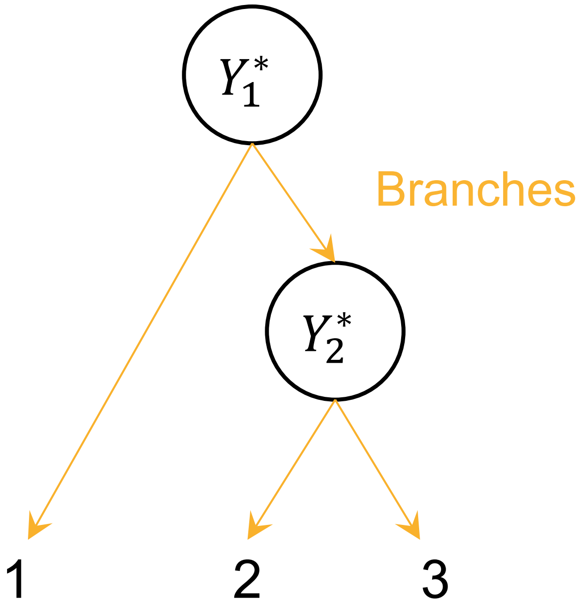
Linear

Nested

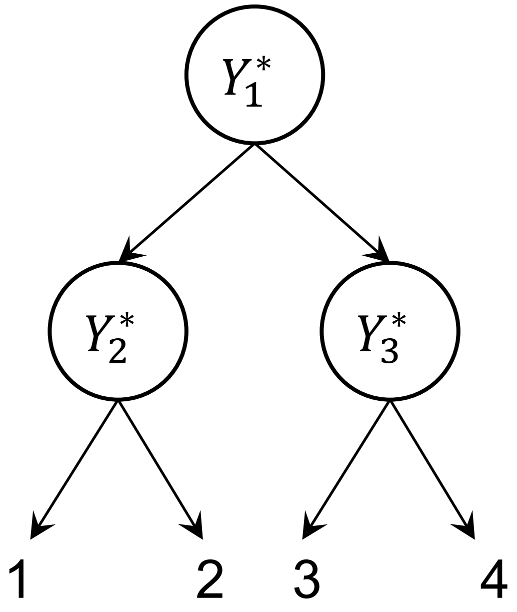
Polytomous

Basic Idea

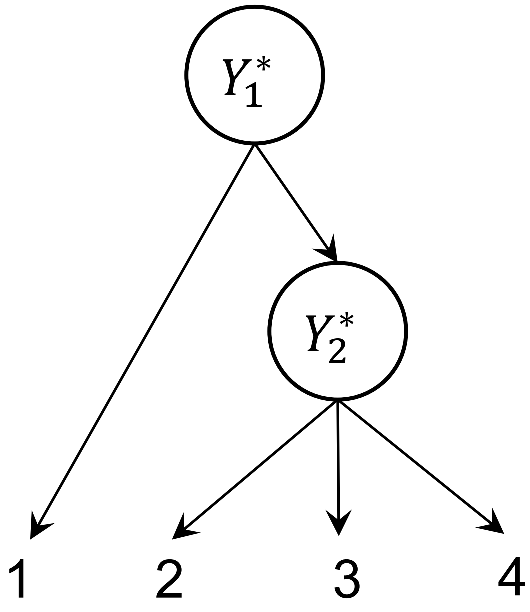
- Example tree structures:



Linear



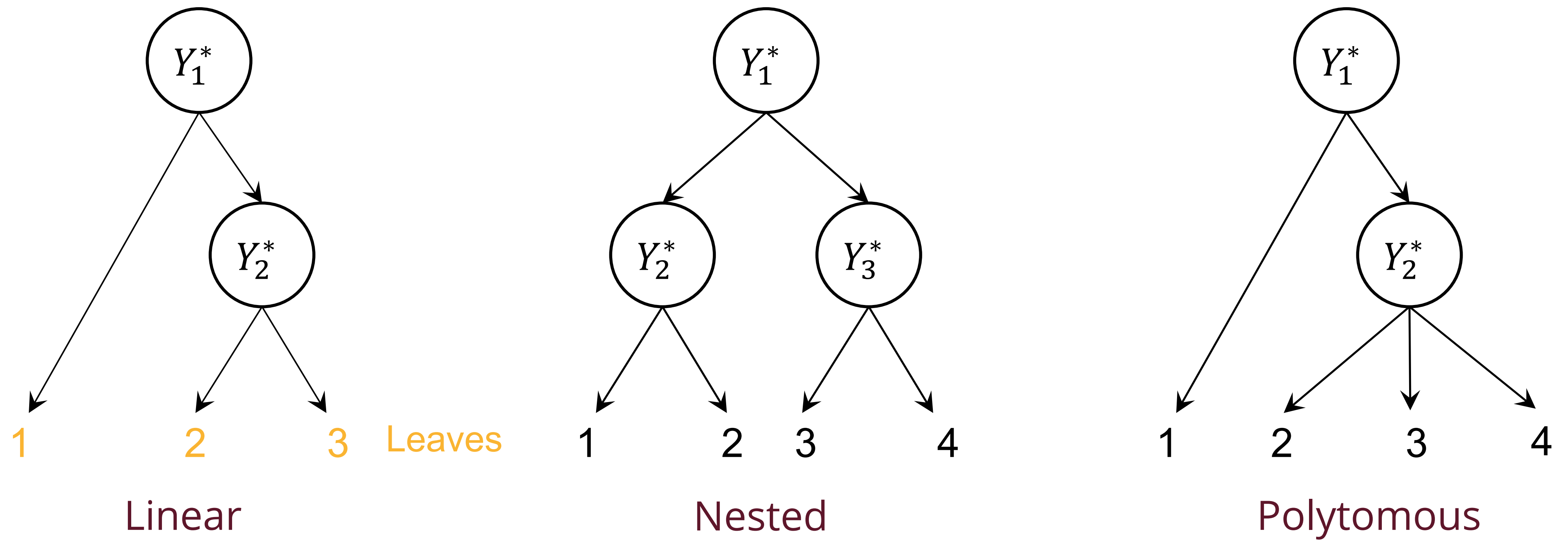
Nested



Polytomous

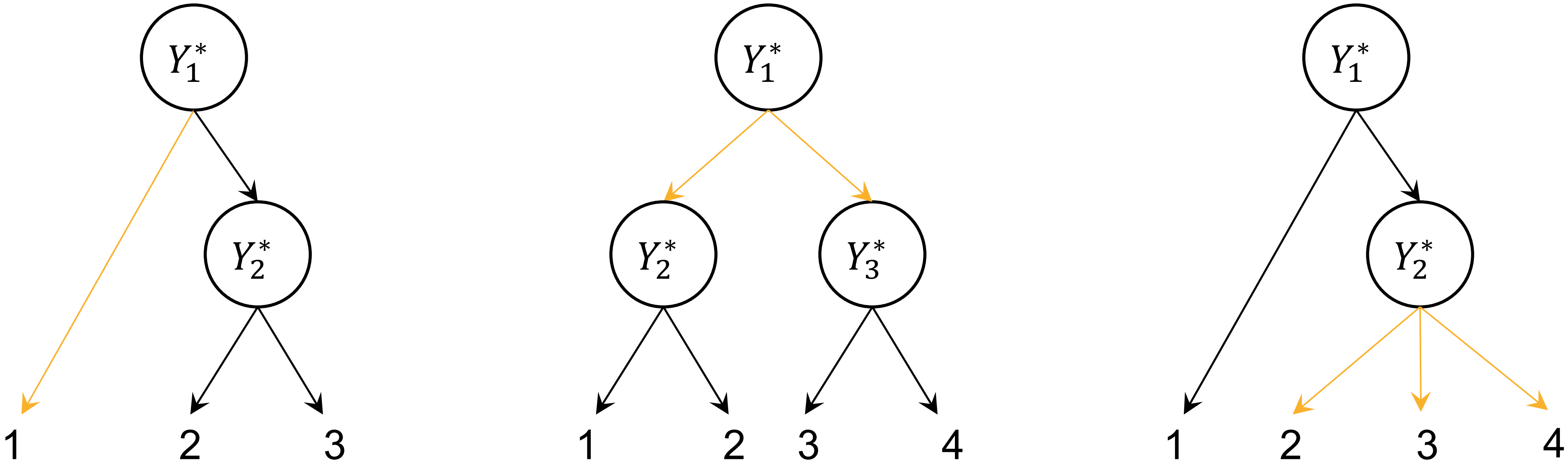
Basic Idea

- Example tree structures:



Basic Idea

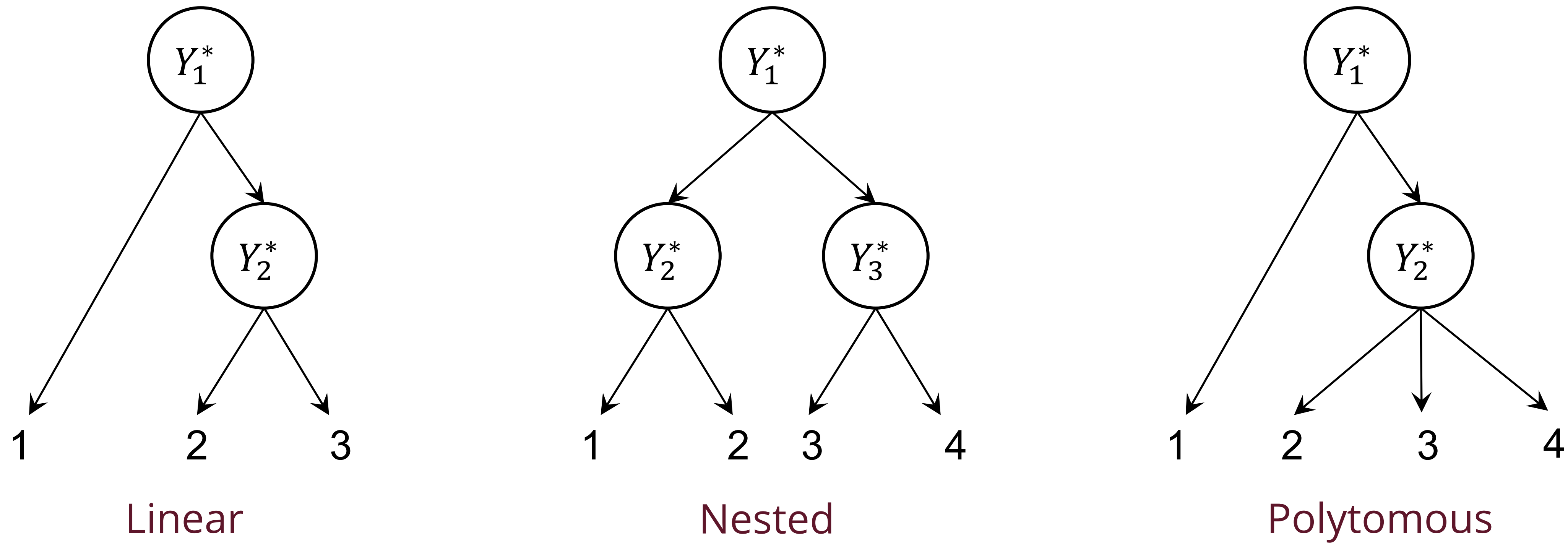
- Example tree structures:



Linear Nested Polytomous

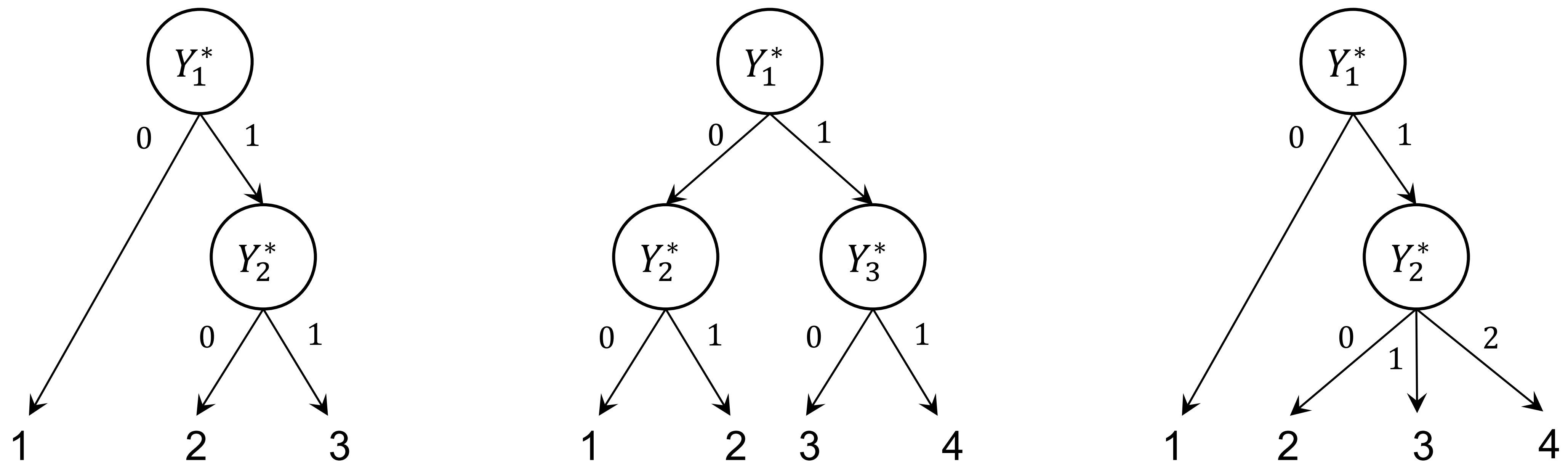
Basic Idea

- We can attach a separate IRT model to each node
- The distinct nodes can be used to statistically separate the influence of multiple traits



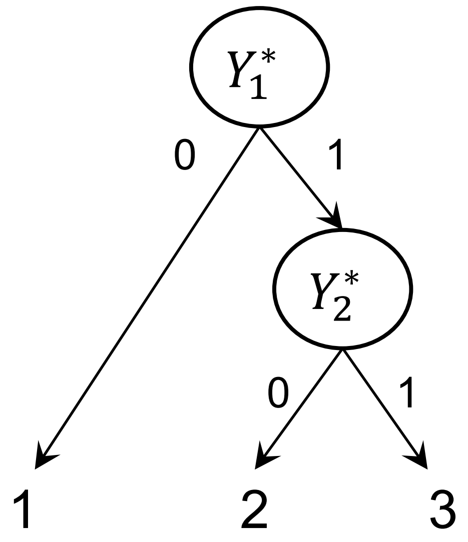
Pseudo-Item Responses

- We transform the observed item responses (Y) into pseudo-item responses (Y^*) based on the specified tree structure
- A binary or categorical decision outcome at each node

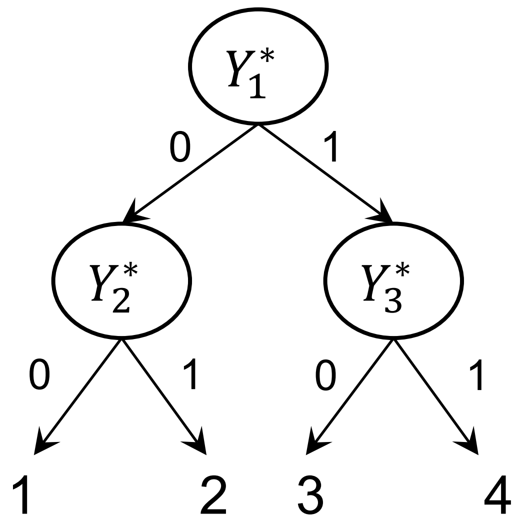


Pseudo-Item Responses

- Examples of transforming observed item responses into pseudo-item responses:



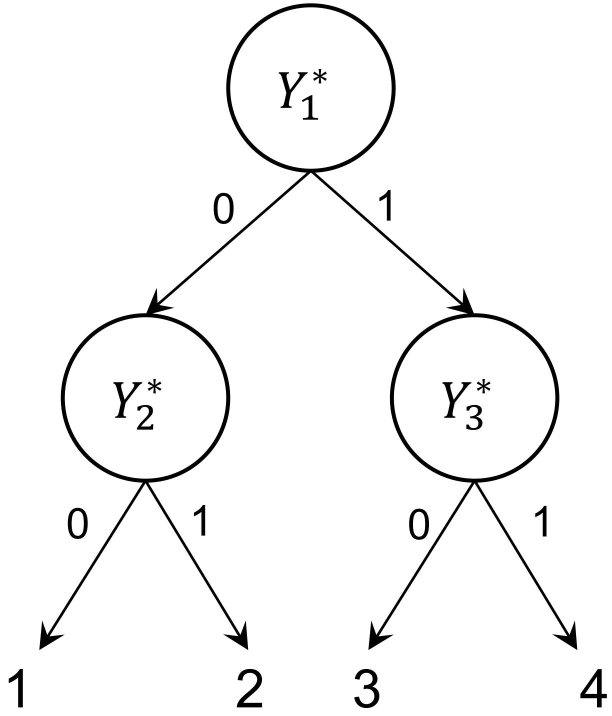
Node	Observed item response (Y)		
	1	2	3
Y_1^*	0	1	1
Y_2^*	NA	0	1



Node	Terminal responses (Y)			
	1	2	3	4
Y_1^*	0	0	1	1
Y_2^*	0	1	NA	NA
Y_3^*	NA	NA	0	1

Model Formulation

- The probability of each pseudo-item response can be specified with an IRT model (e.g., two-parameter logistic (2PL) model)



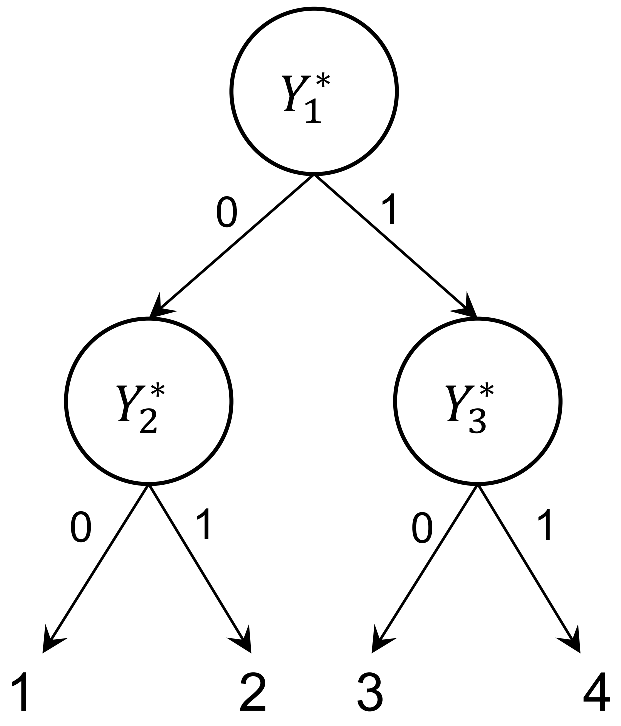
$$P\left(Y_{pi1}^* = 1 \,\middle|\, \theta_{p1}\right) = g^{-1}(a_{i1}\theta_{p1} + d_{i1})$$

$$P\left(Y_{pi2}^* = 1 \,\middle|\, \theta_{p2}\right) = g^{-1}(a_{i2}\theta_{p2} + d_{i2})$$

$$P\left(Y_{pi3}^* = 1 \,\middle|\, \theta_{p3}\right) = g^{-1}(a_{i3}\theta_{p3} + d_{i3})$$

Model Formulation

- The probability of a terminal outcome can be expressed as the product of probabilities of pseudo-item decisions across nodes
  - It is assumed that pseudo-item responses are conditionally independent



$$P\left(Y_{pi}=1\,\middle|\,\boldsymbol{\theta_p}\right)=P\left(Y_{pi1}^*=0\,\middle|\,\theta_{p1}\right)P\left(Y_{pi2}^*=0\,\middle|\,\theta_{p2}\right)$$

$$P\left(Y_{pi}=2\,\middle|\,\boldsymbol{\theta_p}\right)=P\left(Y_{pi1}^*=0\,\middle|\,\theta_{p1}\right)P\left(Y_{pi2}^*=1\,\middle|\,\theta_{p2}\right)$$

$$P\left(Y_{pi}=3\,\middle|\,\boldsymbol{\theta_p}\right)=P\left(Y_{pi1}^*=1\,\middle|\,\theta_{p1}\right)P\left(Y_{pi3}^*=0\,\middle|\,\theta_{p3}\right)$$

$$P\left(Y_{pi}=4\,\middle|\,\boldsymbol{\theta_p}\right)=P\left(Y_{pi1}^*=1\,\middle|\,\theta_{p1}\right)P\left(Y_{pi3}^*=1\,\middle|\,\theta_{p3}\right)$$

Model Formulation

- General formula (Jeon & De Boeck, 2016)

$$P(Y_{pi} = m|\boldsymbol{\theta_p}) = P(Y^*_{pi1} = y^*_{m1}, Y^*_{pi2} = y^*_{m2}, \dots, Y^*_{piK} = y^*_{mK}|\boldsymbol{\theta_p})$$

$$= [\theta_{p1}, \theta_{p2}, \dots, \theta_{pK}]$$

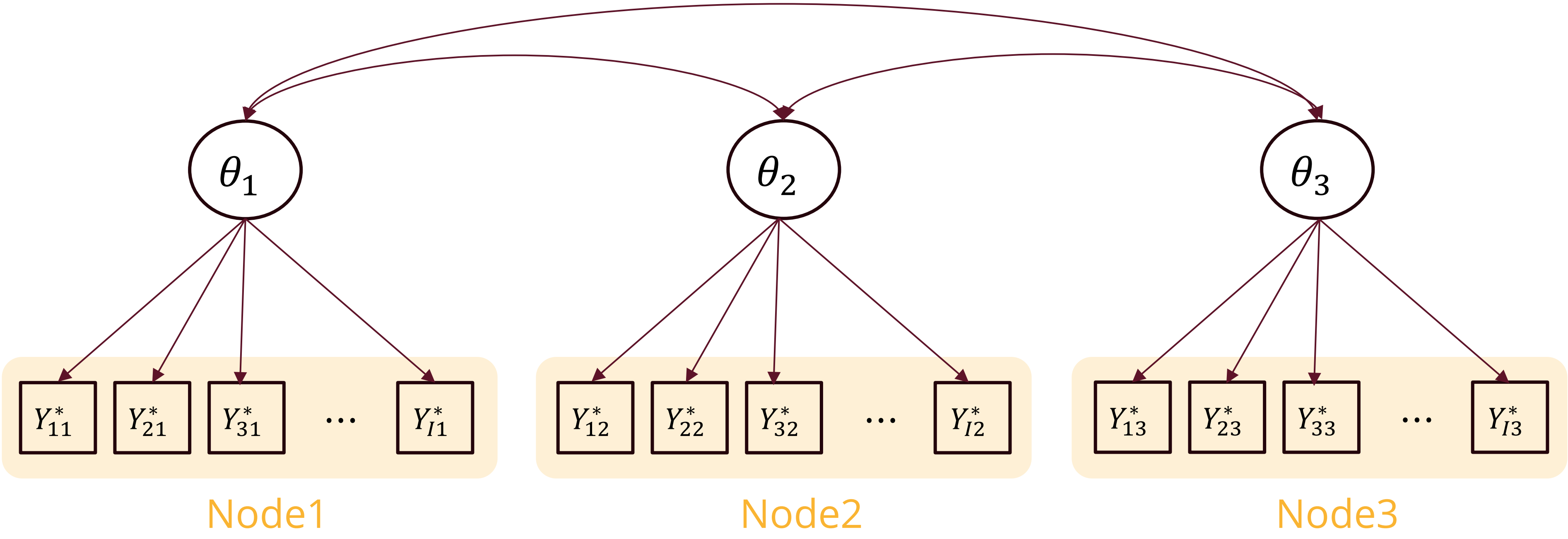
$$= \prod_{k=1}^{K} P(Y^*_{pik} = y^*_{mk}|\theta_{pk})^{t_{mk}}$$

$$= \begin{cases} 0 & \text{if } y^*_{mk} = NA \\ 1 & \text{otherwise} \end{cases}$$

Pseudo-item response on node $k(= 1, \dots, K)$ for terminal outcome m

- Model specification can be viewed as fitting a simple-structure K-dimensional IRT model to pseudo-item responses derived from a specified tree structure

Model Estimation

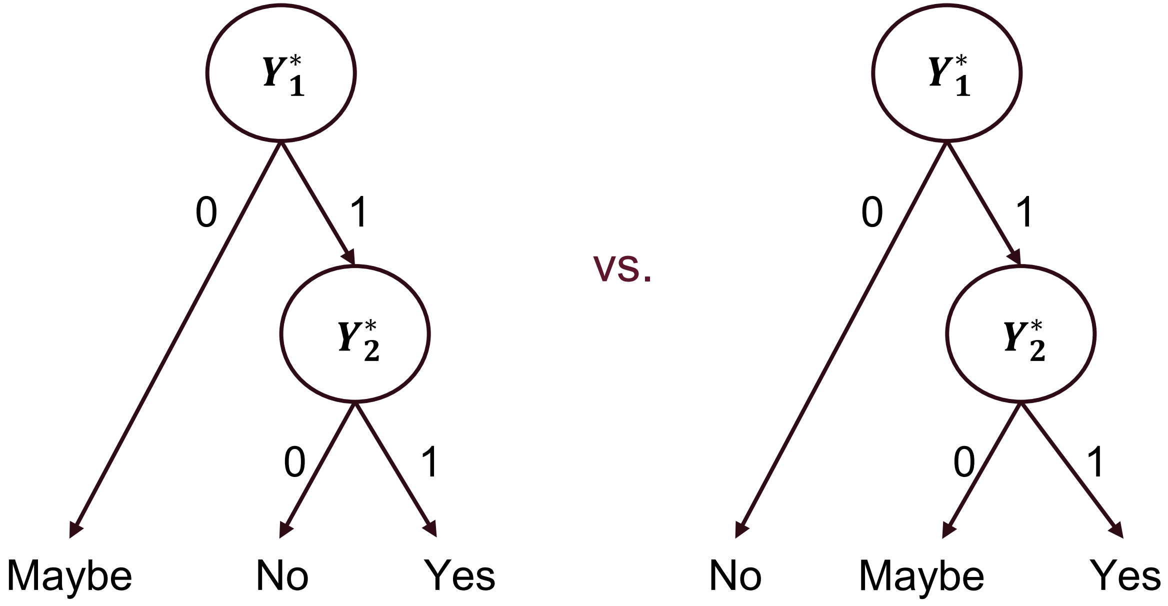


- The approach essentially estimates person and item parameters of a multidimensional IRT model fitted to data with missing responses

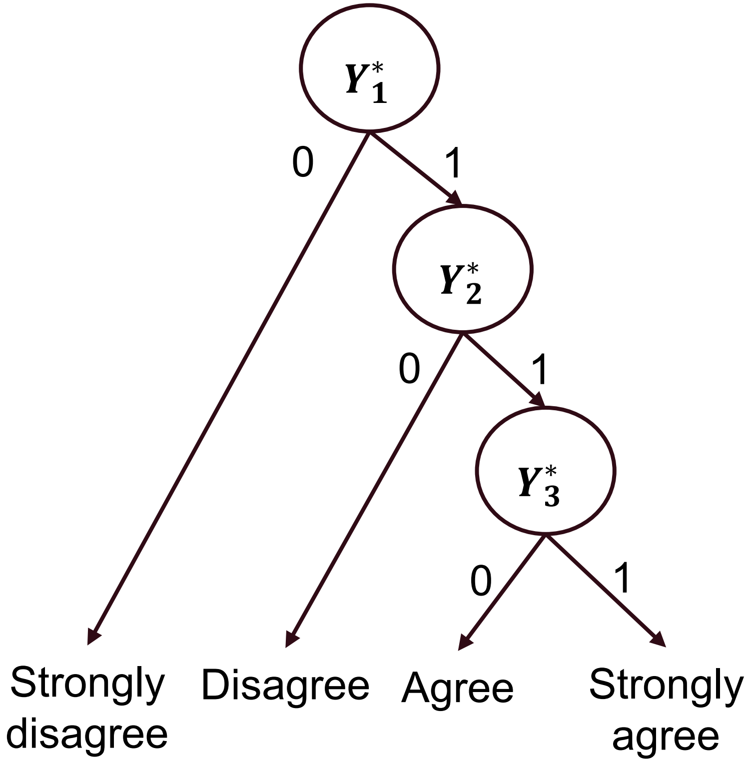
Applications of IRTree Models (1)

- Investigation of the nature of the scale
  (De Boeck & Partchev, 2012; Jeon & De Boeck, 2016)

What would a selection of ‘maybe’ indicate?



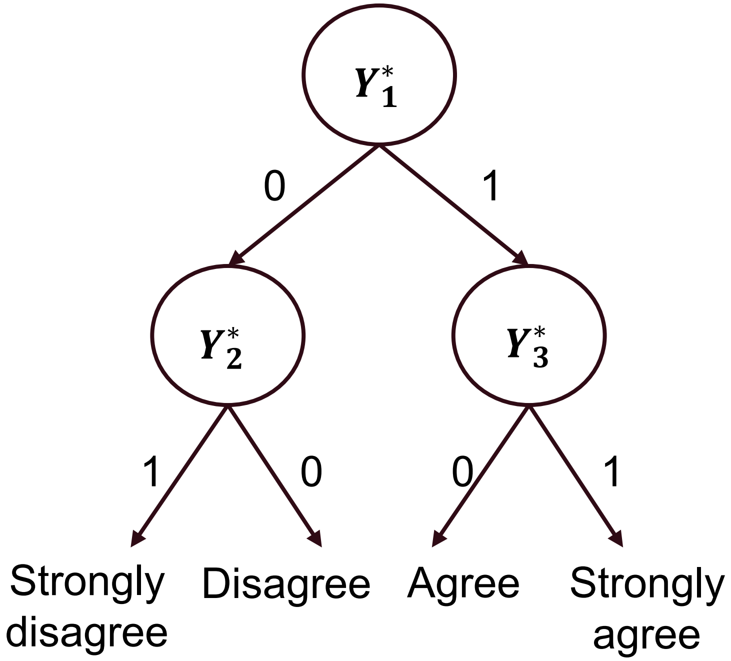
Is this scale ordinal?



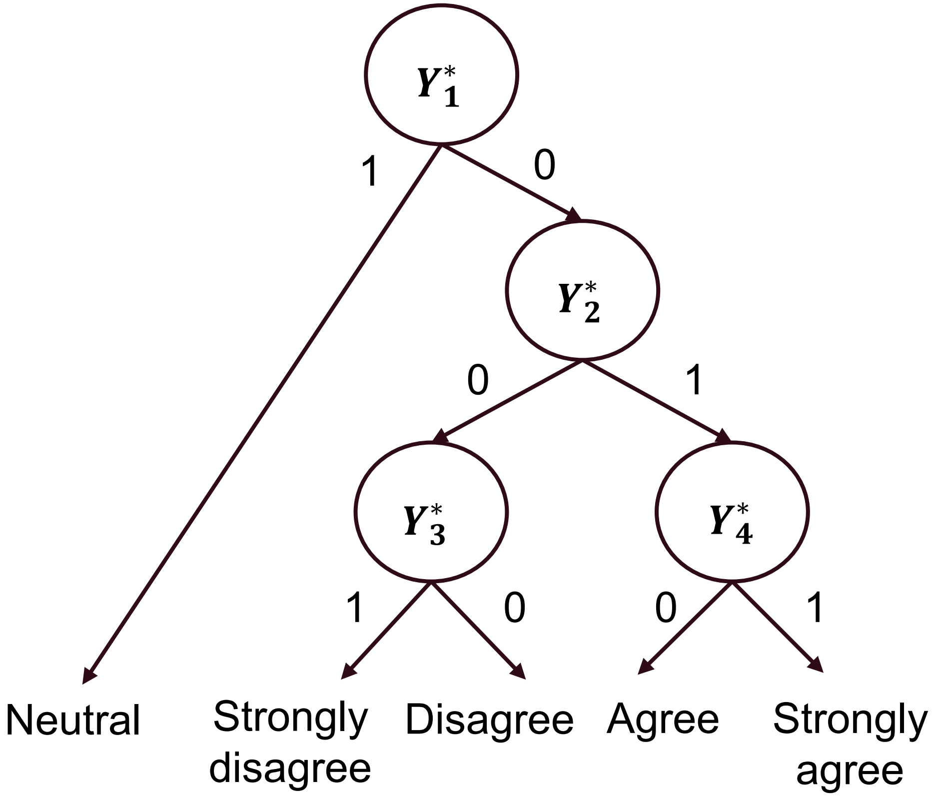
Applications of IRTree Models (2)

- Modeling response styles (Böckenholt & Meiser, 2017; Khorramdel & von Davier, 2014; Plieninger & Meiser, 2014)

Modeling extreme response styles

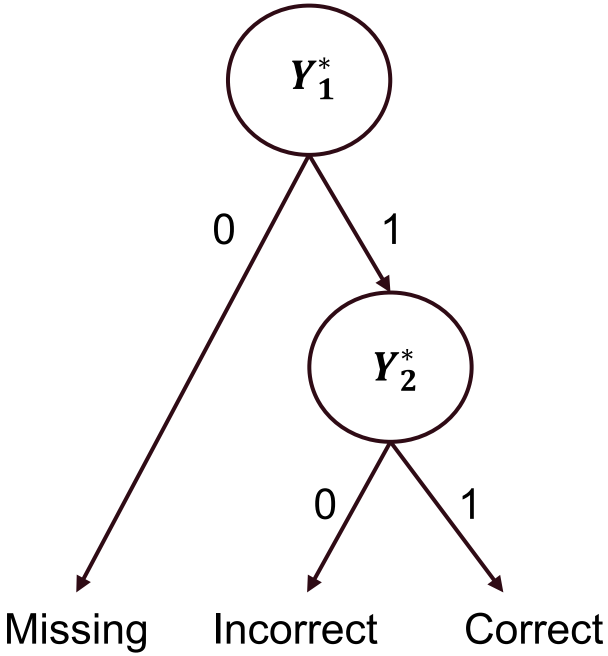


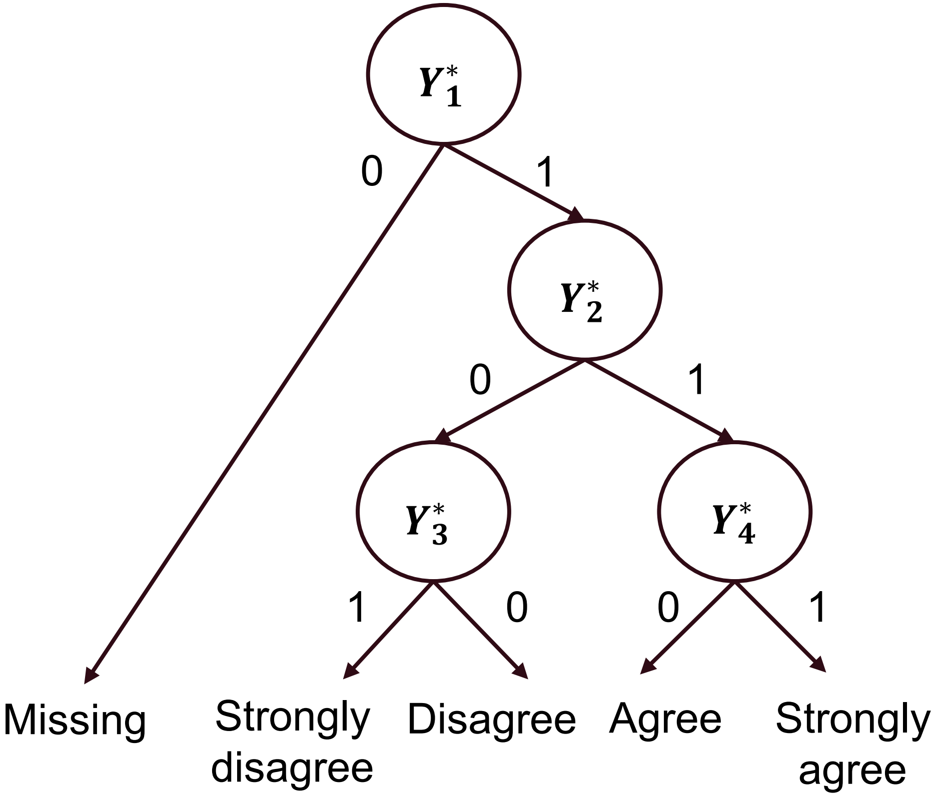
Modeling midpoint and extreme response styles



Applications of IRTree Models (3)

- Modeling omitted/missing responses
  (Debeer et al., 2017; Jeon & De Boeck, 2016)

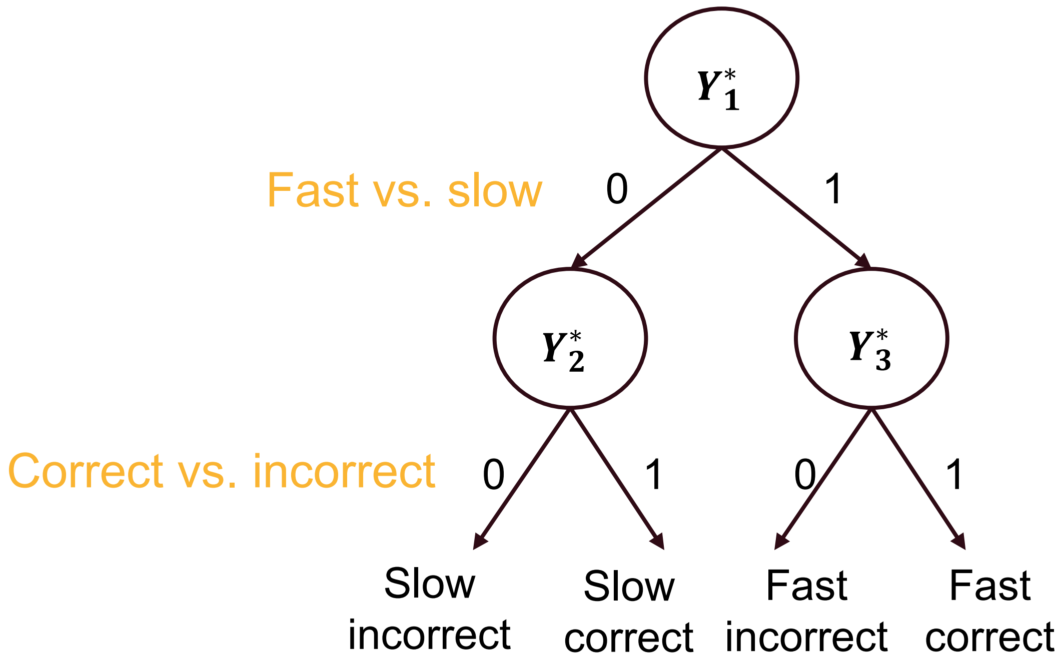




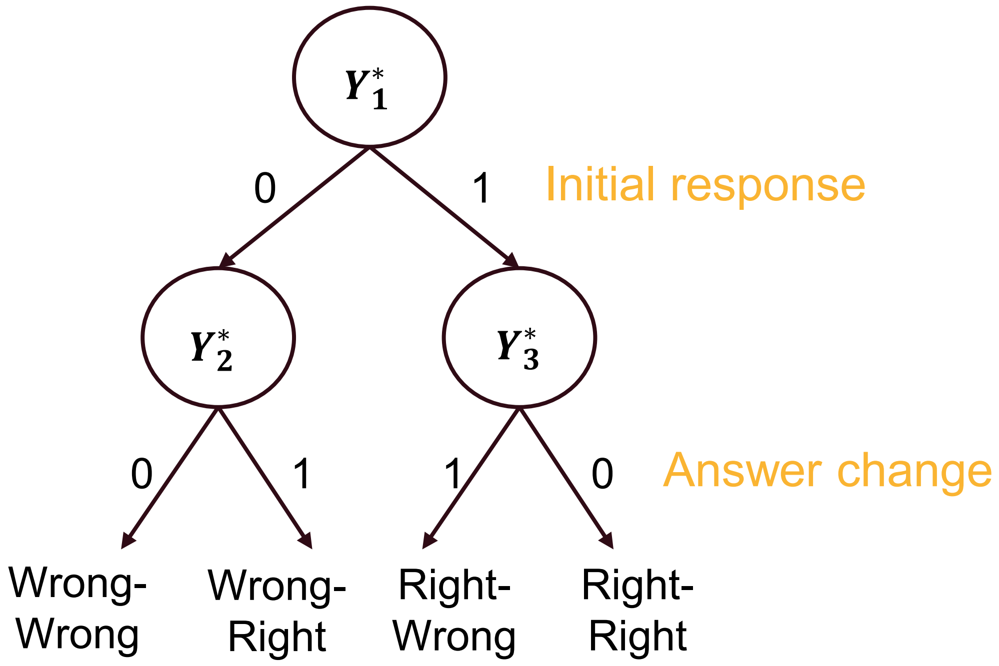
Applications of IRTree Models (4)

- Modeling cognitive processes or test-taking strategy (DiTrapani et al., 2016; Jeon et al., 2017; Partchev & De Boeck, 2012)

Slow vs. fast intelligence?



Modeling answer change behaviors



References

- Böckenholt, U. (2012). Modeling multiple response processes in judgment and choice. *Psychological Methods*, *17*(4), 665–678. https://doi.org/10.1037/a0028111
- Böckenholt, U., & Meiser, T. (2017). Response style analysis with threshold and multi-process IRT models: A review and tutorial. *The British journal of mathematical and statistical psychology*, *70*(1), 159–181. https://doi.org/10.1111/bmsp.12086
- Debeer, D., Janssen, R., & De Boeck, P. (2017), Modeling Skipped and Not-Reached Items Using IRTrees. *Journal of Educational Measurement, 54*, 333-363. https://doi.org/10.1111/jedm.12147
- De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-based item response models of the GLMM family. *Journal of Statistical Software*, *48*(1), 1–28. https://doi.org/10.18637/jss.v048.c01
- DiTrapani, J., Jeon, M., De Boeck, P., & Partchev, I. (2016). Attempting to differentiate fast and slow intelligence: Using generalized item response trees to examine the role of speed on intelligence tests. *Intelligence*, *56*, 82–92. https://doi.org/https://doi.org/10.1016/j.intell.2016.02.012
- Jeon, M. & De Boeck, P. (2016). A generalized item response tree model for psychological assessments. *Behavior Research Method*s, 48(3), 1070–1085. https://doi.org/10.3758/s13428-015-0631-y
- Jeon, M., De Boeck, P., & van der Linden, W. (2017). Modeling Answer Change Behavior: An Application of a Generalized Item Response Tree Model. *Journal of Educational and Behavioral Statistics, 42*(4), 467–490. http://www.jstor.org/stable/26447667
- Khorramdel, L., & von Davier, M. (2014). Measuring response styles across the big five: A multiscale extension of an approach using multinomial processing trees. *Multivariate Behavioral Research, 49*(2), 161–177. https://doi.org/10.1080/00273171.2013.866536
- Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated? *Intelligence, 40*(1), 23–32. https://doi.org/10.1016/j.intell.2011.11.002
- Plieninger, H., & Meiser, T. (2014). Validity of multiprocess IRT models for separating content and response styles. *Educational and Psychological Measurement, 74*(5), 875-899. https://doi.org/10.1177/0013164413514998

Application 1:
Modeling Response Styles

2

Application 1: Modeling Response Styles

Section Learning Objectives

Describe the concept of response styles within the context of noncognitive assessments

Specify a tree structure for modeling different types of response styles

Illustrate statistically the probability of selecting a specific response category based on the IRTree model accounting for response styles

Explain benefits and limitations of the IRTree approach for modeling response styles

Self-Report Rating Scale

- The most common approach to measuring noncognitive constructs is to use a self-report Likert rating scale

How much do you agree with the following statements?

	Strongly Disagree	Disagree	Agree	Strongly Agree
1) I feel supported in my community.	○	○	○	○
2) I feel like I am an important part of my community.	○	○	○	○
3) I feel comfortable being myself in my community.	○	○	○	○

- Respondents select one of the response categories to indicate to what extent they agree or disagree to a given statement
- They may not use or interpret the rating scales in the same way

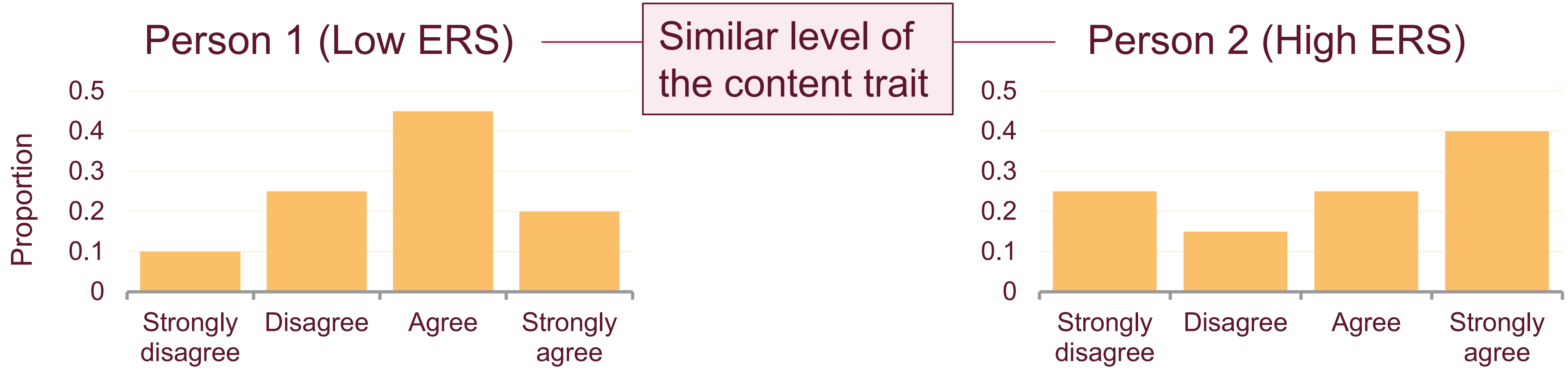
Response Styles

- Some respondents may tend to select certain response categories more often than others
  - For example, some may be inclined to choose extreme or non-extreme response categories
  - Some may consistently overselect or underselect a midpoint response category

Response Styles

Response Styles

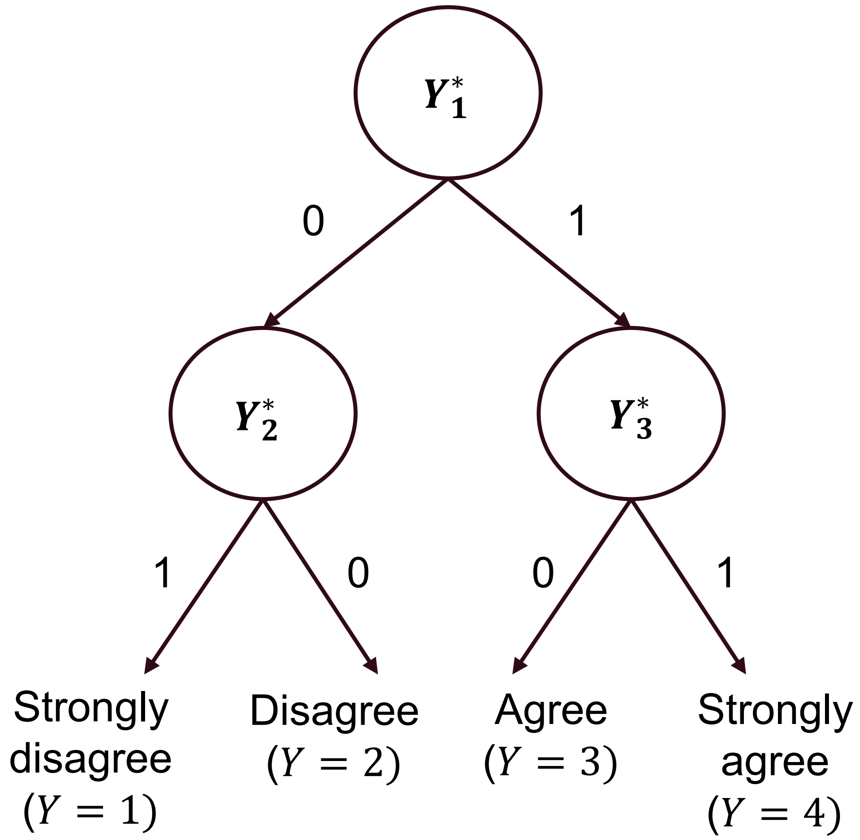
- Response styles (RS) are content-irrelevant tendencies to select certain response categories in rating scale items (Paulhus, 1991)
  - Examples: Extreme response style (ERS), Midpoint response style (MRS)



- Response styles are widely known as a threat to validity
  - They tend to correlate with other person variables such as sociodemographic, personality, and cultural characteristics

IRTree Approach for Modeling ERS

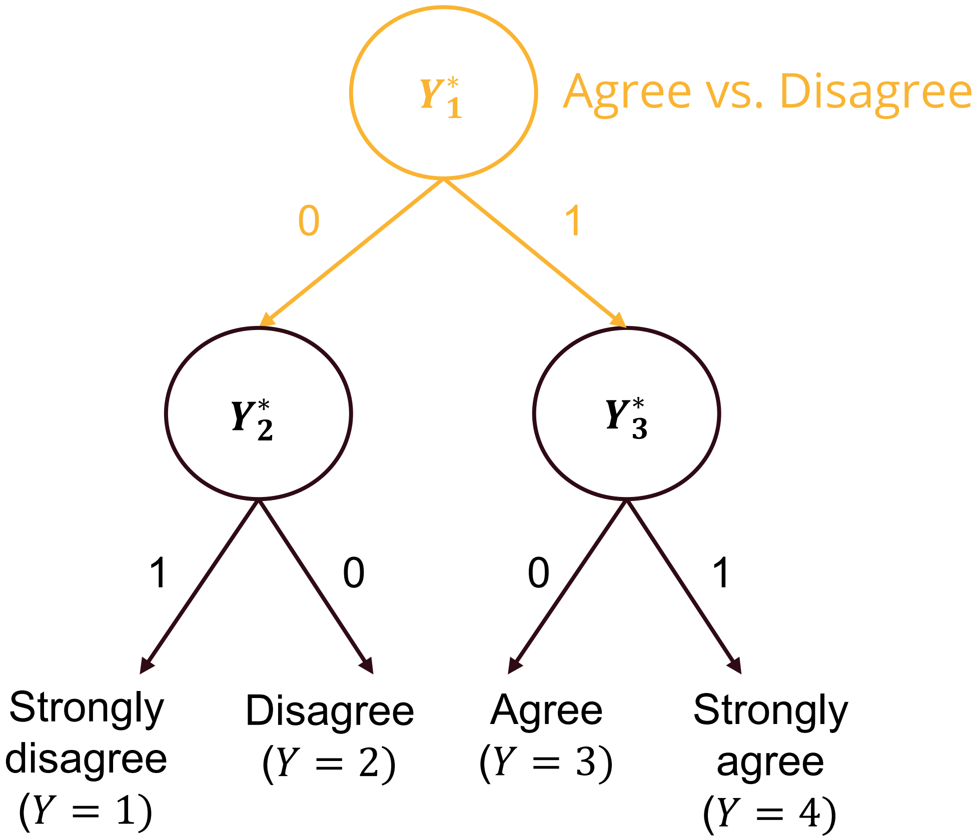
- The figure below illustrates a typical way of specifying a tree structure for modeling ERS



- The tree structure decomposes the response process into a two-stage selection process with three decision nodes

IRTree Approach for Modeling ERS

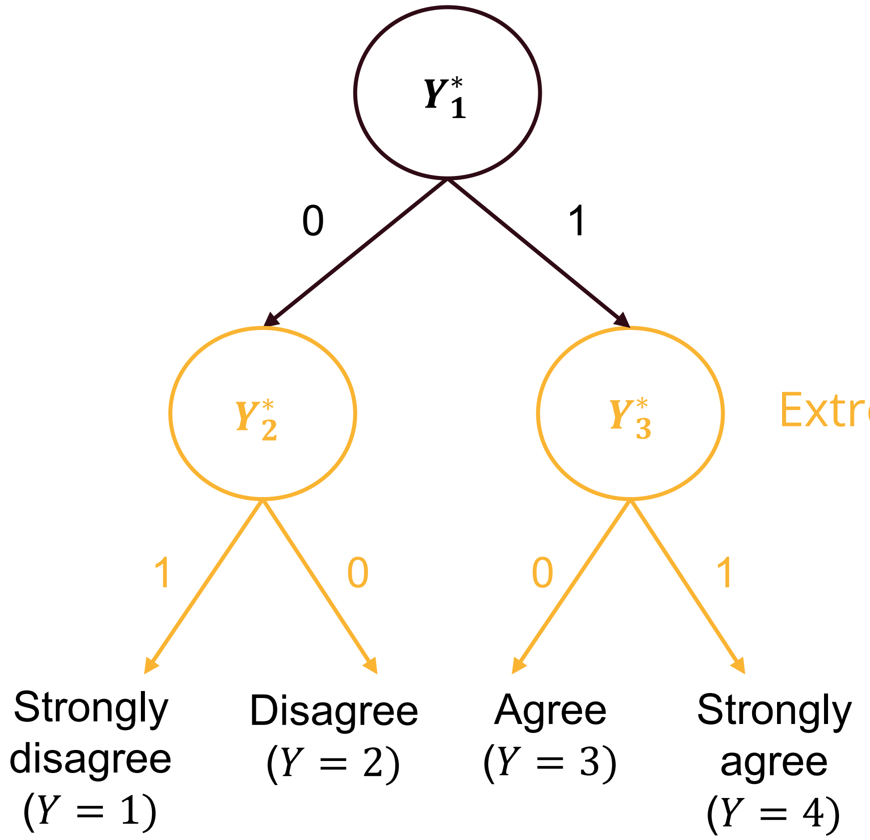
- The figure below illustrates a typical way of specifying a tree structure for modeling ERS



- The tree structure decomposes the response process into a two-stage selection process with three decision nodes

IRTree Approach for Modeling ERS

- The figure below illustrates a typical way of specifying a tree structure for modeling ERS

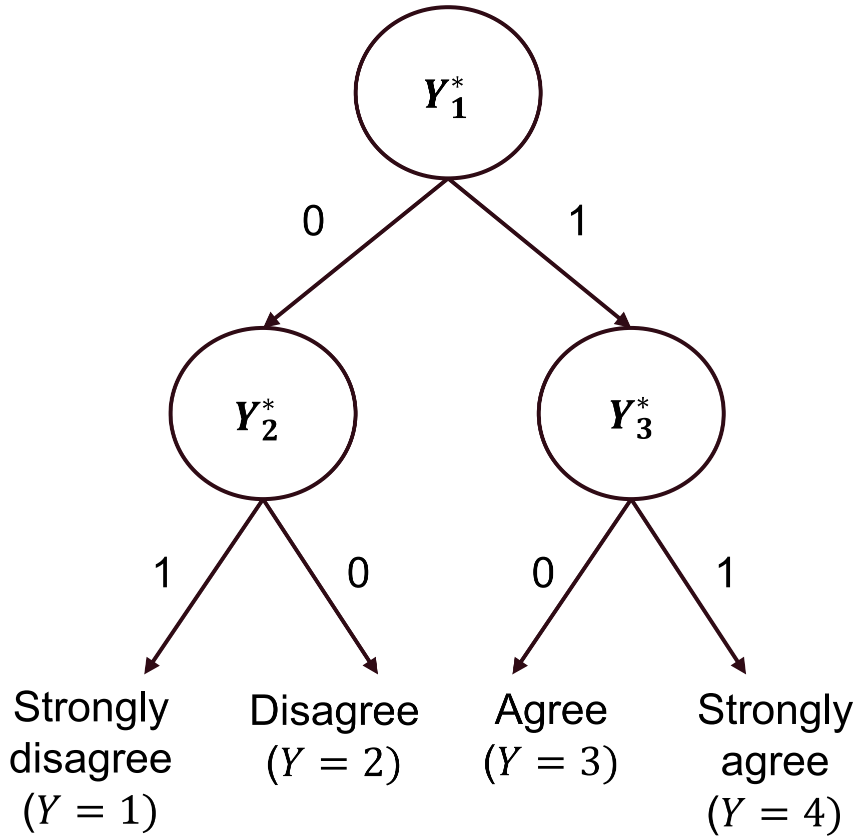


Extreme vs. Non-extreme

- The tree structure decomposes the response process into a two-stage selection process with three decision nodes

IRTree Approach for Modeling ERS

- The figure below illustrates a typical way of specifying a tree structure for modeling ERS

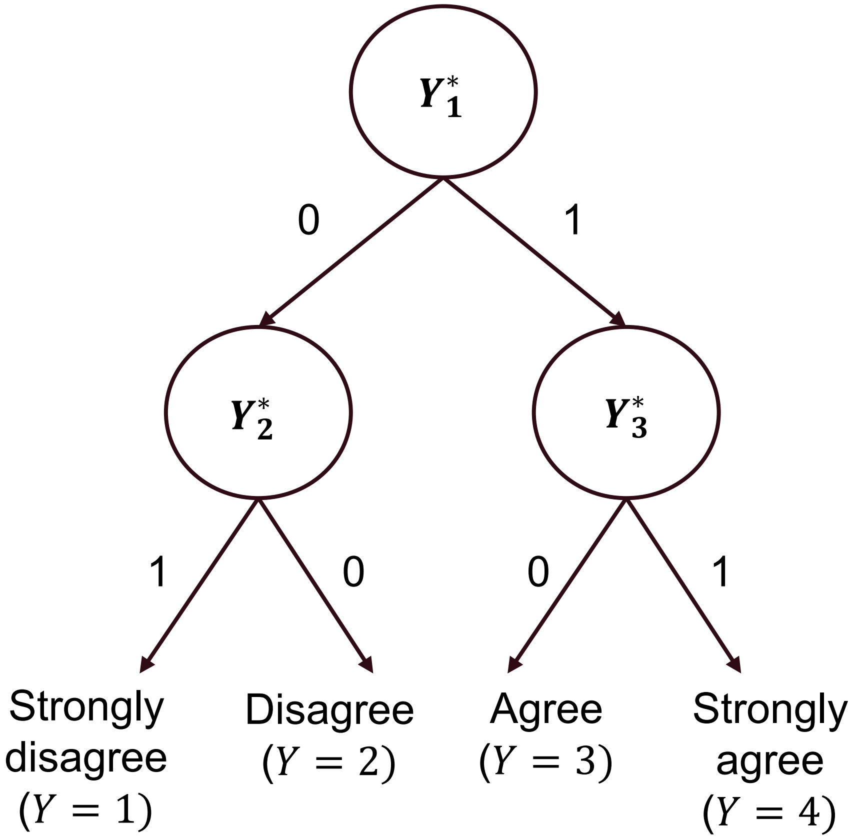


- The tree structure decomposes the response process into a two-stage selection process with three decision nodes
- Pseudo-item responses for each observed item response:

Node	Terminal response (Y)			
	1	2	3	4
Y_1^*	0	0	1	1
Y_2^*	1	0	NA	NA
Y_3^*	NA	NA	0	1

IRTree Approach for Modeling ERS

- We parameterize the probability of a decision at each node using an IRT model (e.g., 2PL model) incorporating a latent trait hypothesized to influence the corresponding response subprocess



Content trait

$$P\left(Y_{pi1}^* = 1 | \theta_p\right) = g^{-1}(a_{i1}\theta_p + d_{i1})$$

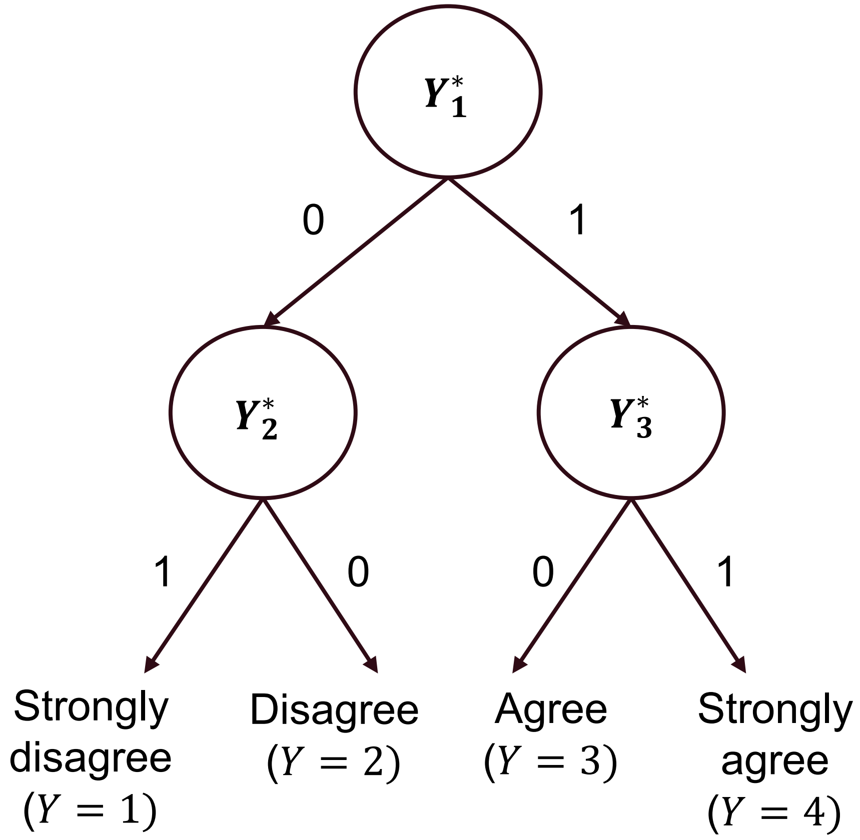
$$P\left(Y_{pi2}^* = 1 | \eta_p\right) = g^{-1}(a_{i2}\eta_p + d_{i2})$$

ERS

$$P\left(Y_{pi3}^* = 1 | \eta_p\right) = g^{-1}(a_{i3}\eta_p + d_{i3})$$

IRTree Approach for Modeling ERS

- We parameterize the probability of a decision at each node using an IRT model (e.g., 2PL model) incorporating a latent trait hypothesized to influence the corresponding response subprocess



Content trait

$$P\left(Y_{pi1}^* = 1|\theta_p\right) = g^{-1}(a_{i1}\theta_p + d_{i1})$$

$$P\left(Y_{pi2}^* = 1|\,\eta_p\right) = g^{-1}(a_{i2}\eta_{p1} + d_{i2})$$

Negative ERS

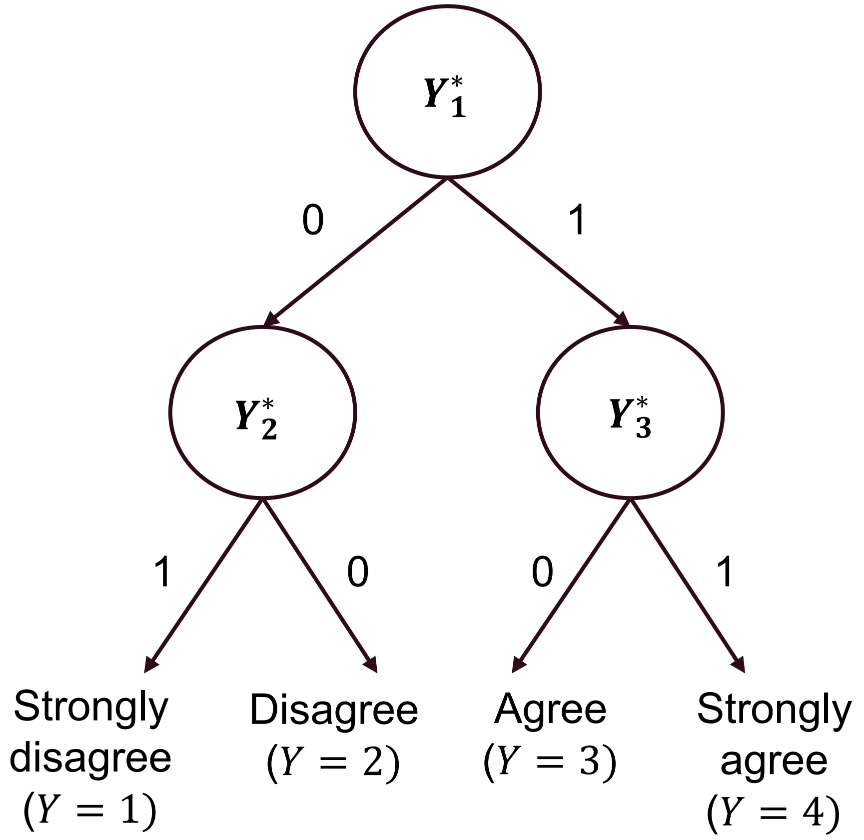
$$P\left(Y_{pi3}^* = 1|\,\eta_p\right) = g^{-1}(a_{i3}\eta_{p2} + d_{i3})$$

Positive ERS

Directional invariance was empirically supported (Jeon & De Boeck, 2019)

IRTree Approach for Modeling ERS

- We parameterize the probability of a decision at each node using an IRT model (e.g., 2PL model) incorporating a latent trait hypothesized to influence the corresponding response subprocess



Content trait

$$P\left(Y_{pi1}^* = 1 | \theta_p\right) = g^{-1}(a_{i1}\theta_p + d_{i1})$$

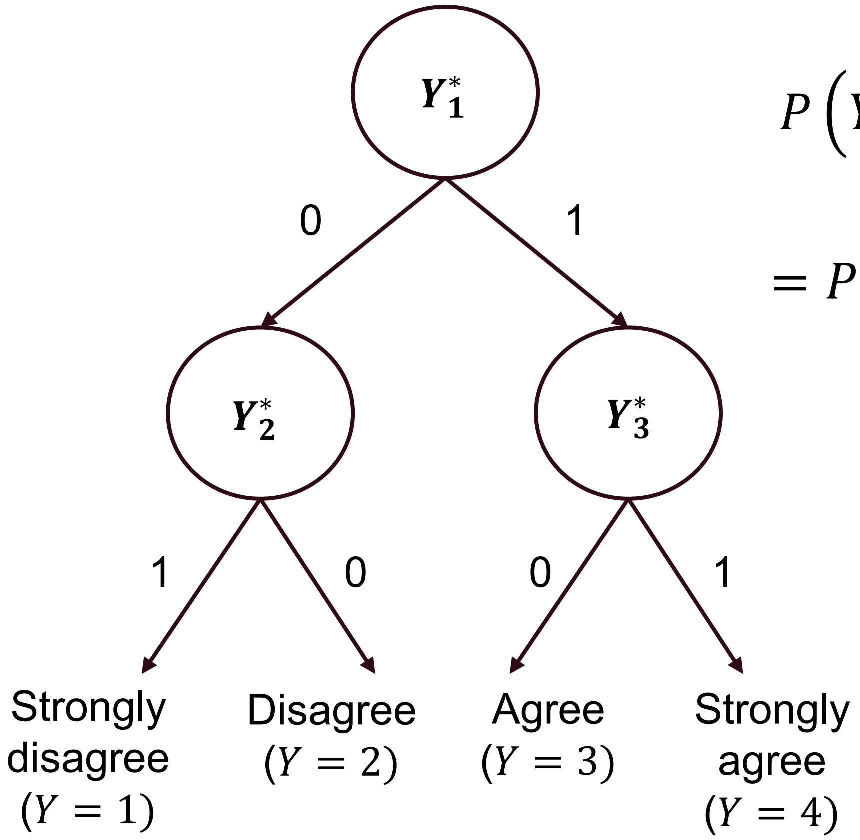
$$P\left(Y_{pi2}^* = 1 | \eta_p\right) = g^{-1}(a_{i2}\eta_p + d_{i2})$$

ERS

$$P\left(Y_{pi3}^* = 1 | \eta_p\right) = g^{-1}(a_{i3}\eta_p + d_{i3})$$

IRTree Approach for Modeling ERS

- The probability of observed item response Y_{pi} can be expressed as the product of probabilities of pseudo-item decisions across three nodes

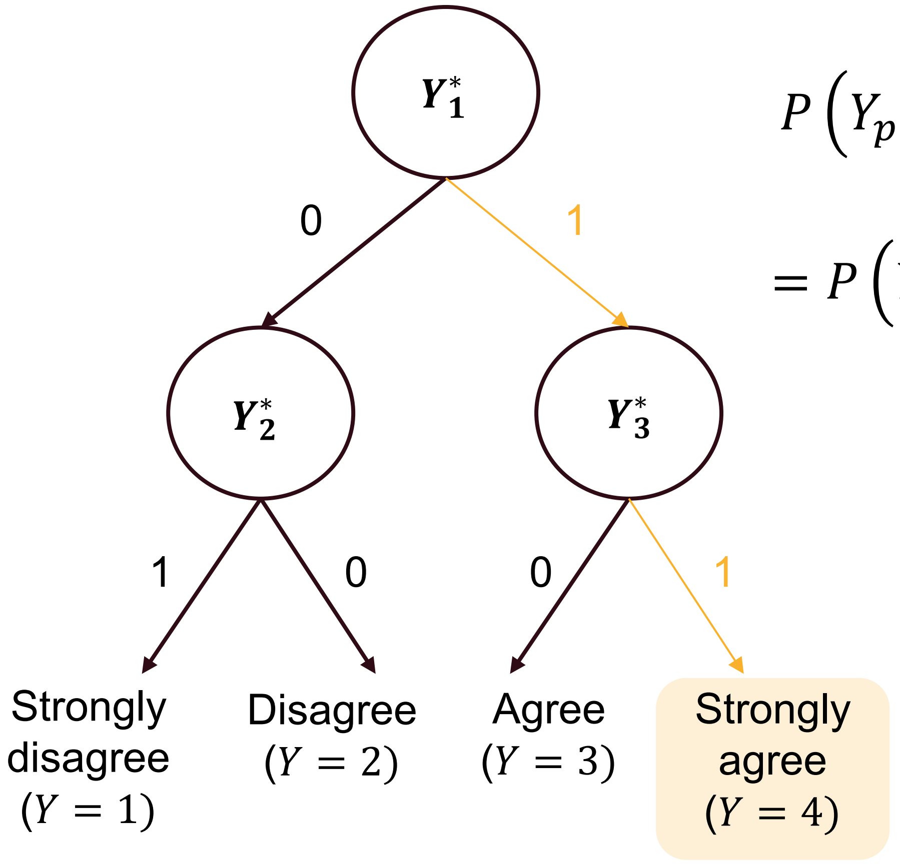


$$P\left(Y_{pi}=m\,\middle|\,\theta_p,\eta_p\right)=P\left(Y_{pi1}^*=y_{m1}^*,Y_{pi2}^*=y_{m2}^*,Y_{pi3}^*=y_{m3}^*\,\middle|\,\theta_p,\eta_p\right)$$

$$=P\left(Y_{pi1}^*=y_{m1}^*\,\middle|\,\theta_p\right)P\left(Y_{pi2}^*=y_{m2}^*\,\middle|\,\eta_p\right)^{1-y_{m1}^*}P\left(Y_{pi3}^*=y_{m3}^*\,\middle|\,\eta_p\right)^{y_{m1}^*}$$

IRTree Approach for Modeling ERS

- The probability of observed item response Y_{pi} can be expressed as the product of probabilities of pseudo-item decisions across three nodes



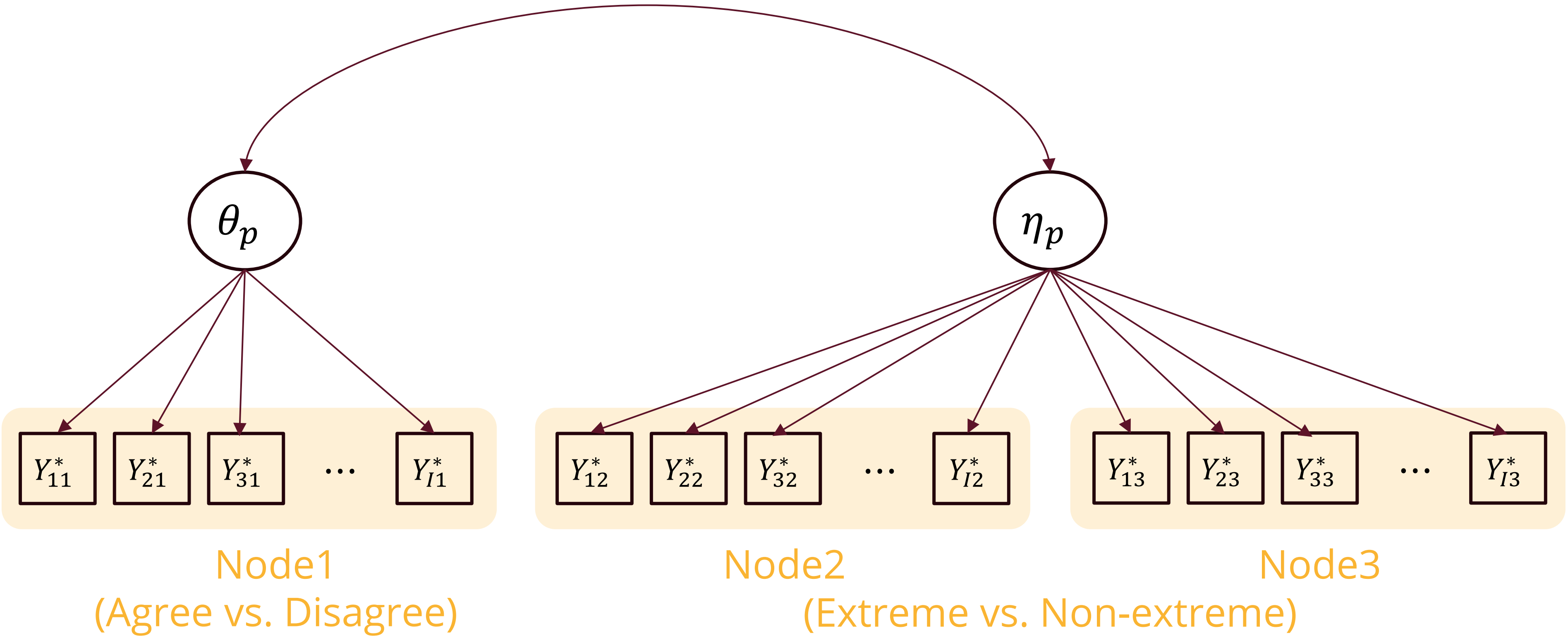
$$P\left(Y_{pi} = m \middle| \theta_p, \eta_p\right) = P\left(Y_{pi1}^* = y_{m1}^*, Y_{pi2}^* = y_{m2}^*, Y_{pi3}^* = y_{m3}^* \middle| \theta_p, \eta_p\right)$$

$$= P\left(Y_{pi1}^* = y_{m1}^* \middle| \theta_p\right) P\left(Y_{pi2}^* = y_{m2}^* \middle| \eta_p\right)^{1-y_{m1}^*} P\left(Y_{pi3}^* = y_{m3}^* \middle| \eta_p\right)^{y_{m1}^*}$$

$$P\left(Y_{pi} = 4 \middle| \theta_p, \eta_p\right) = P\left(Y_{pi1}^* = 1 \middle| \theta_p\right) P\left(Y_{pi3}^* = 1 \middle| \eta_p\right)$$

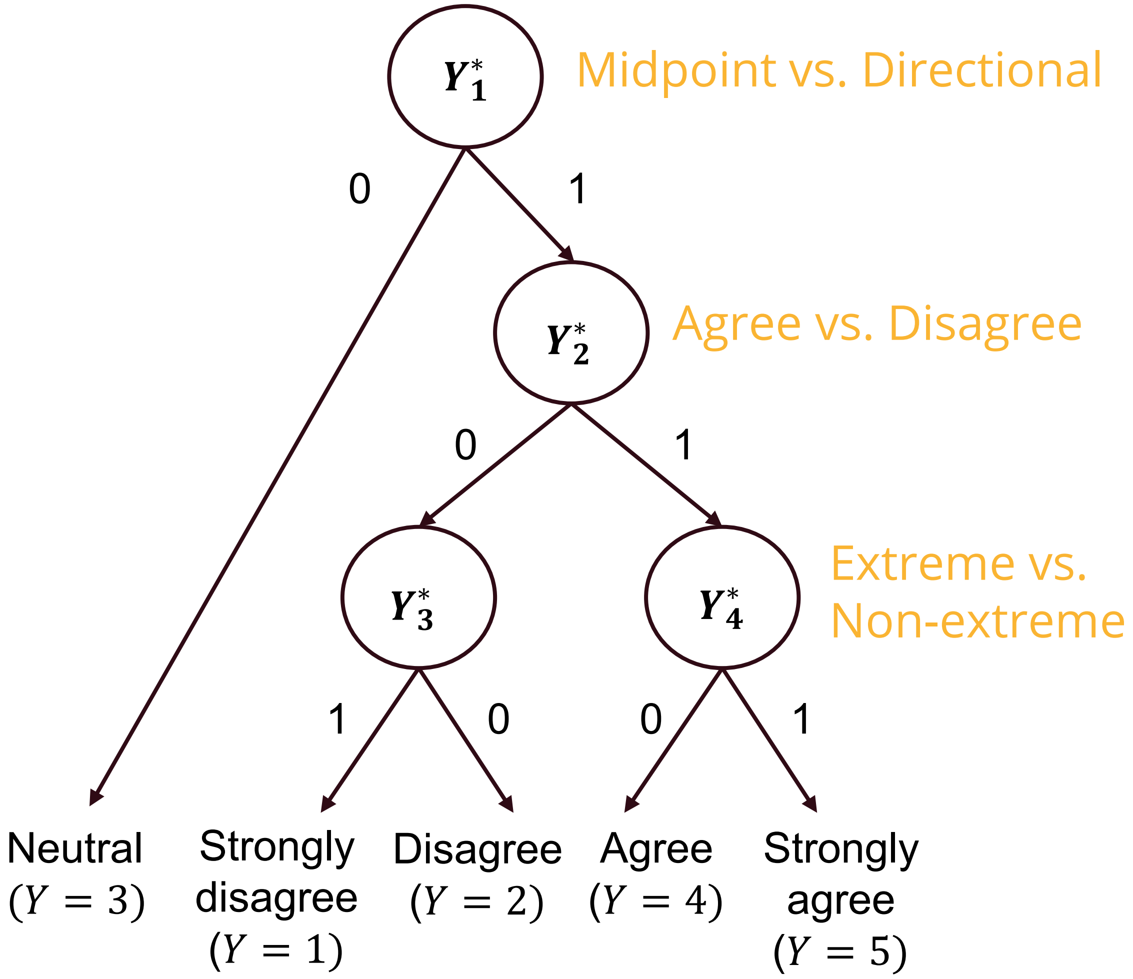
IRTree Approach for Modeling ERS

- A graphical representation of the model fitted to pseudo-item responses:



IRTree Approach for Modeling MRS and ERS

- We can similarly construct an IRTree for modeling both MRS and ERS for a 5-point rating scale



- It is decomposed into a three-stage selection process
- We can estimate the content trait, ERS, and MRS by attaching separate IRT models to different nodes
- Pseudo-item responses for each observed item response:

Node	Terminal response (Y)				
	1	2	3	4	5
Y_1^*	0	0	1	0	0
Y_2^*	0	0	NA	1	1
Y_3^*	1	0	NA	NA	NA
Y_4^*	NA	NA	NA	0	1

Advantages of IRTree Modeling for RS

- IRTree modeling allows us to easily define and differentiate multiple response styles by decomposing the response process into binary internal decisions for selecting specific types of response categories (Böckenholt & Meiser, 2017)
  - It is useful for modeling response styles in a confirmatory way
  - It is easy to statistically separate the influence of the content trait and response styles

Limitations of IRTree Modeling for RS

- The model regards the underlying response process as identical across all individuals (Kim & Bolt, 2021)
- We may lose some information for estimating the content trait (Kim & Bolt, 2021), although the amount of information loss may not be substantial (Plieninger & Meiser, 2014)

References

- Böckenholt, U., & Meiser, T. (2017). Response style analysis with threshold and multi-process IRT models: A review and tutorial. *British Journal of Mathematical and Statistical Psychology, 70*(1), 159-181. https://doi.org/10.1111/bmsp.12086
- Jeon, M. & De Boeck, P. (2019). Evaluation on types of invariance in studying extreme response bias with an IRTree approach. *British Journal of Mathematical and Statistical Psychology, 72*(3), 517-537. https://doi-org.ezp2.lib.umn.edu/10.1111/bmsp.12182
- Kim, N. & Bolt, D. M. (2021). A mixture IRTree model for extreme response style: Accounting for response process uncertainty. *Educational and Psychological Measurement, 81*(1), 131–154. https://doi.org/10.1177/0013164420913915
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes (Vol. 1, pp. 17-59).* Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X
- Plieninger, H., & Meiser, T. (2014). Validity of multiprocess IRT models for separating content and response styles. *Educational and Psychological Measurement, 74*(5), 875-899. https://doi.org/10.1177/0013164413514998

Application 2:
Modeling Test-taking Behaviors

3

3

Application 2: Modeling Test-Taking Behaviors

Section Learning Objectives

Explain the tree structure of IRTree models within the context of test-taking behaviors

Specify the IRTree model for answer change and rapid guessing in the cognitive assessment context

Interpret item parameters in the context of test-taking behaviors

Evaluate the assumptions and advantages of the IRTree models for test-taking behaviors

Test-Taking Behaviors

- Cognitive assessments are designed to assess students' knowledge and skills
- Test scores are used to reflect their test performance
- Understanding test-taking behaviors is important because they can influence and bias test results
- IRTree models have been utilized in cognitive assessment to investigate students' test taking behaviors

Test-Taking Behaviors

- Test-taking behavior can provide insights about:
  - Engagement and Effort (Wise, 2017)
  - Cognitive Processing (Jeon et al., 2017)
  - Score Validity (Rios & Deng, 2021)

Answer Change Behavior

- Students may modify their responses during a test as a result of critical thinking, self-reflection, or test wiseness
- This behavior can have implications for test reliability and score accuracy

Rapid Guessing Behavior

- Students may answer questions too quickly, often without fully reading or considering them
- This may be due to disengagement or test fatigue, impacting the accuracy and fairness of scores

Answer Change (AC)

- Common for multiple-choice format items
- Test-takers may review their initial response
- They may either keep the initial response or change to an alternative answer

- Contradictory ideas/findings on AC and test scores
  - Traditional thought: AC can lower test scores
  - Empirical findings:
    - AC produced more wrong to right answers and improved test scores (Coffey et al., 2024)
    - Test score changes by AC depend on test takers’ characteristics, such as students' ability level (Stylianou-Georgiou & Papanastasiou, 2017) and attitudes (e.g., reflectivity; Friedman & Cook, 1995)

Answer Change (AC)

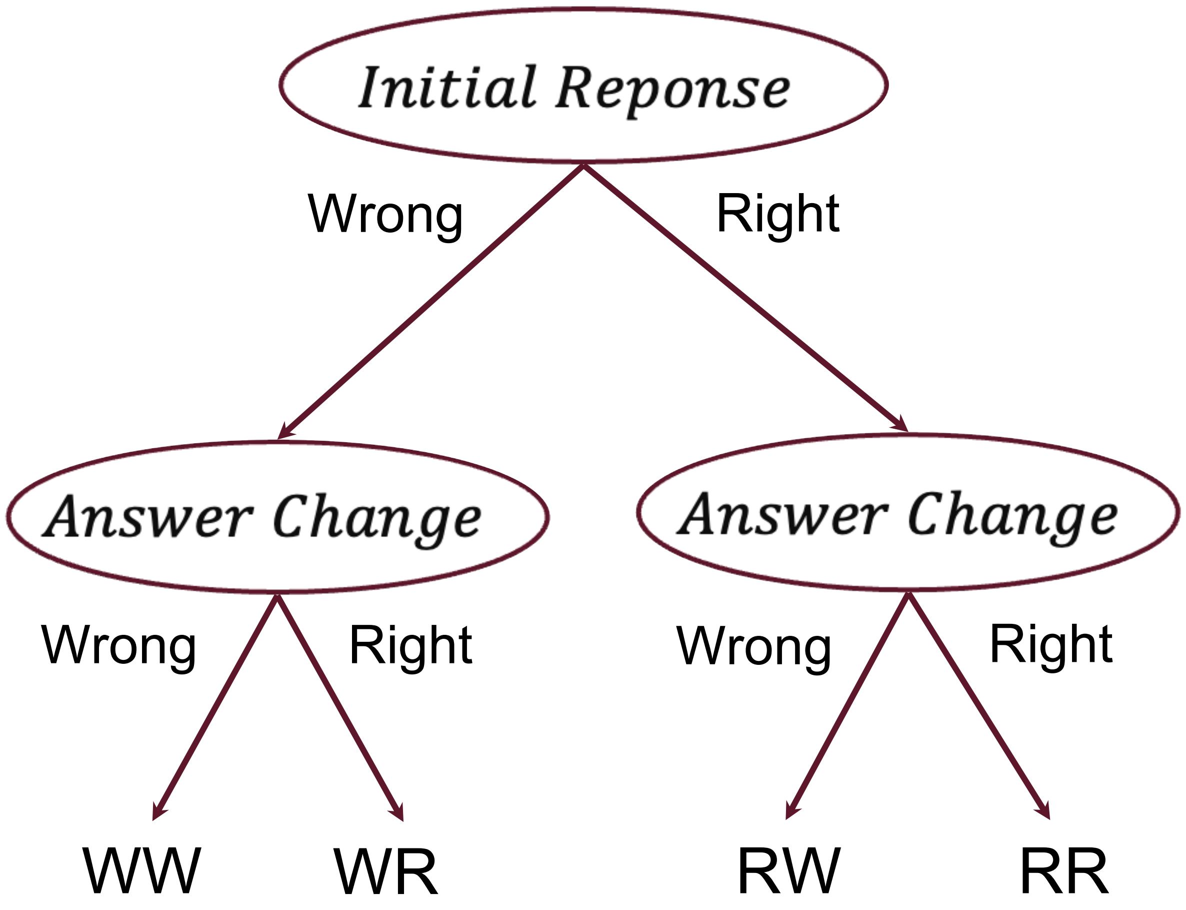
- Previous findings are mostly based on classical test theory (CTT) methods
  - AC outcomes were indicated by counts or proportions
  - Overlooked variability of individuals' ability and item property (e.g., item difficulty)
- Existent IRT-based method
  - A two-stage procedure to model the answer change process (van der Linden et al., 2011)
    - Regular IRT model for the initial response in stage one
    - Fixed-ability logistic regression model for final response after the answer review in stage two
  - Issue: The same latent trait was involved in the two stages

IRTree Model for Answer Change

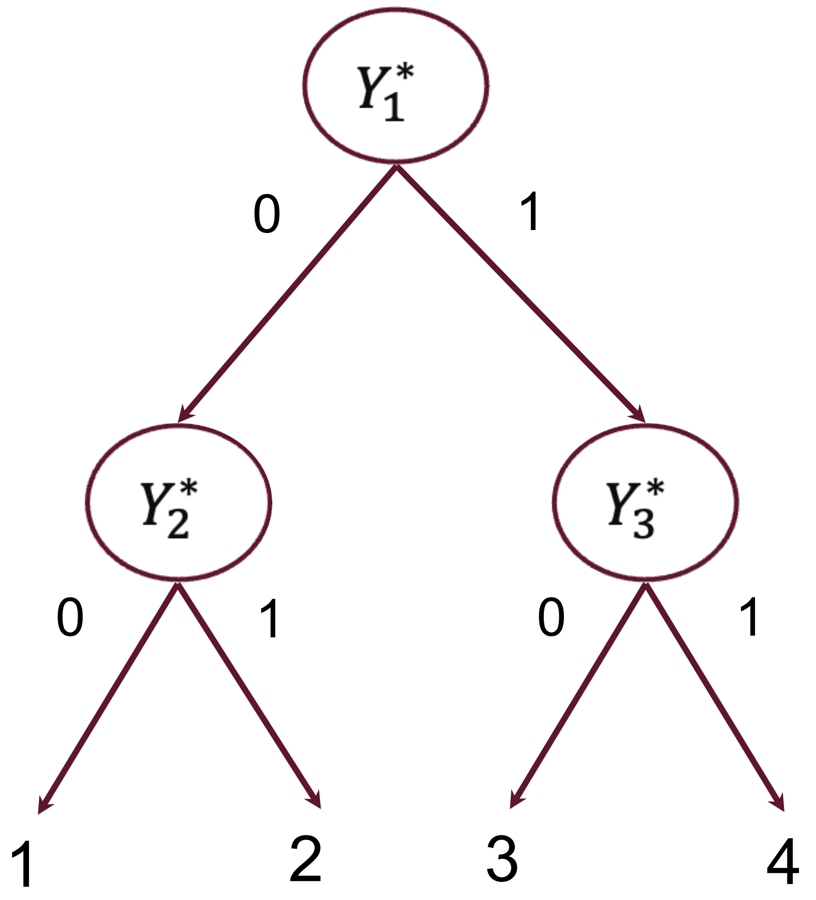
- Initial Response: the first response a test taker provides
- Final Response: the response a test taker submits
- Four terminal responses of answer change behavior
  - A wrong answer to a wrong answer (WW)
  - A wrong answer to a right answer (WR)
  - A right answer to a wrong answer (RW)
  - A right answer to a right answer (RR)

IRTree Model for Answer Change

- Response Process



- Nested Tree Structure



IRTree Model for Answer Change

- Example Data - Key: B for Item 1; A for Item 2

Person	Item1 IR	Item2 IR	…	Item1 FR	Item2 FR	…	Item1 node1	Item2 node1	…	Item1 node2	Item2 node2	…	Item1 node3	Item2 node3	…
1	B	D	…	B	A	…	1	0	…	NA	1	…	1	NA	…
2	A	A	…	A	C	…	0	1	…	0	NA	…	NA	0	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
999	C	B	…	B	D	…	0	0	…	1	0	…	NA	NA	…
1,000	D	A	…	B	A	…	0	1	…	1	NA	…	NA	1	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Note. IR = Initial Response. FR = Final Response. Correct Answers are in red.

IRTree Model for Answer Change

The probability of each pseudo-item response (Jeon et al., 2017):

- The probability of correct responses to item ***i*** for person ***p*** for Node 1 (initial response):

$$P\left(Y_{pi1}^{*}=1\right)=g^{-1}\left(a_{i1}\theta_{p1}+d_{i1}\right)$$

- The probability of correct responses to item ***i*** for person ***p*** for Node 2 and 3 (answer change behavior) conditional on the initial wrong and right response, respectively:

$$P\left(Y_{pi2}^{*}=1\right)=g^{-1}\left(a_{i2}\theta_{p2}+d_{i2}\right)$$
$$P\left(Y_{pi3}^{*}=1\right)=g^{-1}\left(a_{i3}\theta_{p3}+d_{i3}\right)$$

IRTree Model for Answer Change

- The probability of each terminal responses:

Outcome Category	**Terminal Response**	**Pseudo-item 1**	**Pseudo-item 2**	**Pseudo-item 3**
WW	1	0	0	NA
WR	2	0	1	NA
RW	3	1	NA	0
RR	4	1	NA	1

$$P(WW) = P(Y_{pi} = 1) = P(Y^*_{pi1} = 0)\,P(Y^*_{pi2} = 0)$$

$$P(WR) = P(Y_{pi} = 2) = P(Y^*_{pi1} = 0)\,P(Y^*_{pi2} = 1)$$

$$P(RW) = P(Y_{pi} = 3) = P(Y^*_{pi1} = 1)\,P(Y^*_{pi3} = 0)$$

$$P(RR) = P(Y_{pi} = 4) = P(Y^*_{pi1} = 1)\,P(Y^*_{pi3} = 1)$$

IRTree Model for Answer Change

- The probability of the terminal response of 4 (RR):

Outcome Category	**Terminal Response**	**Pseudo-item 1**	**Pseudo-item 2**	**Pseudo-item 3**
WW	1	0	0	NA
WR	2	0	1	NA
RW	3	1	NA	0
RR	**4**	**1**	**NA**	**1**

$$P(RR) = P\left(Y_{pi} = 4\right) = P\left(Y_{pi1}^{*} = 1\right) P\left(Y_{pi3}^{*} = 1\right)$$

$$= \frac{1}{1+\exp\left(-a_{i1}(\theta_{p1}+d_{i1})\right)} \times \frac{1}{1+\exp\left(-a_{i3}(\theta_{p3}+d_{i3})\right)}$$

Rapid Guessing (RG)

- What is RG?
  - Occurs when test-takers quickly respond without adequate consideration of item contents (Wise, 2017)

- Why do we care about RG?
  - Identified as a construct-irrelevant factor
  - Threats to validity by providing inaccurate information about test takers' true proficiency level
  - Can lead to adverse impact on various measurement properties and test-takers' ability estimates (Rios & Deng, 2021)

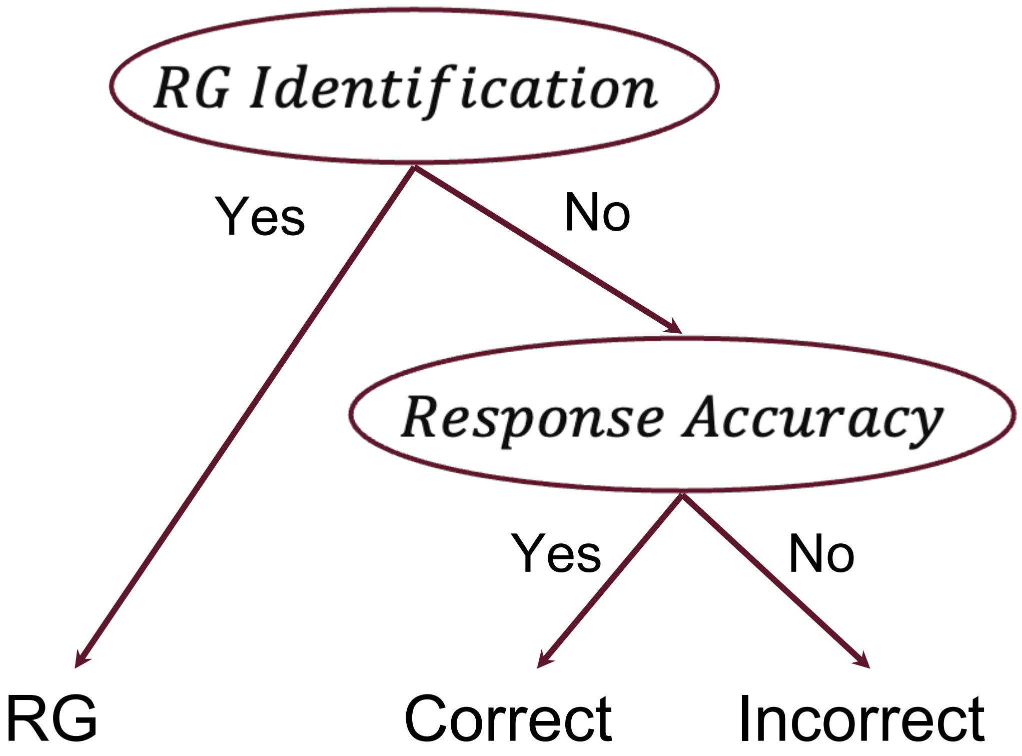
Rapid Guessing (RG)

- How do we identify RG?
  - Global-level Measures (e.g., self-report instruments; Wise & DeMars, 2005)
  - Behavior Indicators (e.g., eye movement tracking; Maddox et al., 2018)
  - **Response Time (RT) Methods** (e.g., Normative Threshold; Wise & Ma, 2012)
    - ➢ Defined as the total time spent on a single item
    - ➢ In computerized tests, response time for each item can be recorded without interrupting individuals' test taking process

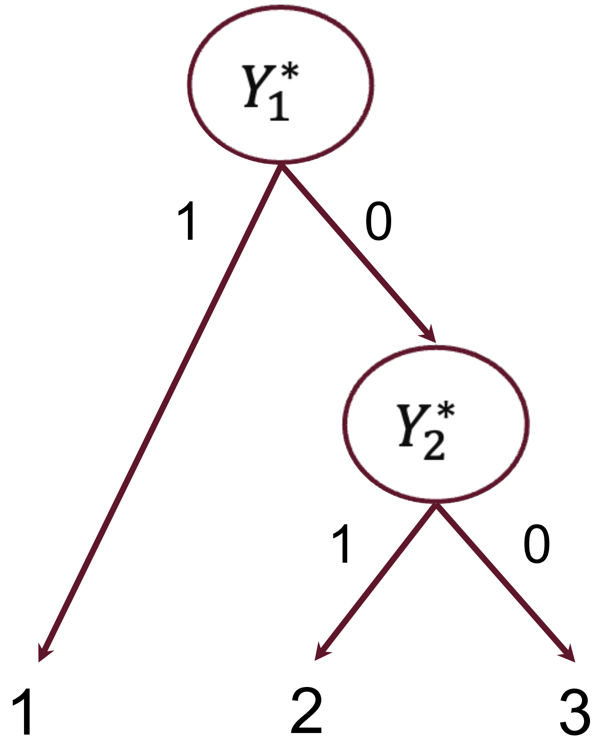
IRTree Model for Rapid Guessing

- Three terminal responses:
  (1) RG response; (2) Incorrect Answer; (3) Correct Answer

- Response Process



- Linear Tree Structure



IRTree Model for Rapid Guessing

- Example Data - Key: B for Item 1; A for Item 2

Person	Item1	Item2	…	Item1 RG	Item2 RG	…	Item1 node1	Item2 node1	…	Item1 node2	Item2 node2	…
1	B	D	…	0	0	…	0	0	…	1	0	…
2	A	A	…	1	1	…	1	1	…	NA	NA	…
…	…	…	…	…	…	…	…	…	…	…	…	…
999	C	B	…	0	1	…	0	1	…	0	NA	…
1,000	D	A	…	0	1	…	0	1	…	0	NA	…
…	…	…	…	…	…	…	…	…	…	…	…	…

Note. Right answers and identified rapid guessing (RG) responses are in red.

IRTree Model for Rapid Guessing

The probability of each pseudo-item response (Node):

- The probability of each rapid guessing response to item *i* for person *p* for Node 1:

$$P\left(Y_{pi1}^{*} = 1\right) = g^{-1}\left(a_{i1}\theta_{pi} + d_{i1}\right)$$

- The probability of correct response to item *i* for person *p* for Node 2 conditional on effortful response:

$$P\left(Y_{pi2}^{*} = 1\right) = g^{-1}\left(a_{i2}\theta_{p2} + d_{i2}\right)$$

IRTree Model for Rapid Guessing

- The probability of each outcome category:

Outcome Category	Terminal Response	Pseudo-item 1	Pseudo-item 2
Rapid Guessing	1	1	NA
Correct Response	2	0	1
Incorrect Response	3	0	0

$$P(RG) = P\left(Y_{pi} = 1\right) = P\left(Y^*_{pi1} = 1\right)$$

$$P(Correct) = P\left(Y_{pi} = 2\right) = P\left(Y^*_{pi1} = 0\right)P\left(Y^*_{pi2} = 1\right)$$

$$P(Incorrect) = P\left(Y_{pi} = 3\right) = P\left(Y^*_{pi1} = 0\right)P\left(Y^*_{pi2} = 0\right)$$

IRTree Model for Test-taking Behaviors

- Advantages
  - Allows for separate continuous latent traits for test-taking behavior and response
    - Answer change: initial response and answer change process
    - Rapid guessing: test-taking effort and response accuracy
  - The two cognitive processes are allowed to be related and modeled simultaneously

IRTree Model for Test-taking Behaviors

- Assumptions
  - The model formulation relies only on the observed responses
    - Answer change: the change between initial responses and final responses
    - Rapid guessing: the identified rapid guessing responses by response time thresholds

Reference

- Coffey, J. S., Maruca, A. T., Polifroni, E. C., & Snyder, M. (2024). Changing test answers: A scoping review. *Nurse Education Today, 133*, 106052.
- Friedman, S. J., & Cook, G. L. (1995). Is an examinee's cognitive style related to the impact of answer changing on multiple-choice tests? *Journal of Experimental Education, 63*, 199–213.
- Jeon, M., De Boeck, P., & van der Linden, W. (2017). Modeling answer change behavior: An application of a generalized item response tree model. *Journal of Educational and Behavioral Statistics, 42*(4), 467-490.
- Maddox, B., Bayliss, A. P., Fleming, P., Engelhardt, P. E., Edwards, S. G., & Borgonovi, F. (2018). Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment. *European Journal of Psychology of Education, 33*, 543-558.
- Rios, J. A., & Deng, J. (2021). Does the choice of response time threshold procedure substantially affect inferences concerning the identification and exclusion of rapid guessing responses? A meta-analysis. *Large-scale Assessments in Education, 9*, 1-25.
- Stylianou-Georgiou, A., & Papanastasiou, E. C. (2017). Answer changing in testing situations: The role of metacognition in deciding which answers to review. *Educational Research and Evaluation, 23*(3-4), 102-118.
- van der Linden, W. J., Jeon, M., & Ferrera, S. (2011). A paradox in the study of the benefits of test-item review. *Journal of Educational Measurement, 48*, 380–398.Wise 2017
- Wise, S.L. (2017), Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications*. Educational Measurement: Issues and Practice* (36), 52-61. https://doi.org/10.1111/emip.12165
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational assessment, 10*(1), 1–17.
- Wise, S. L., & Ma, L. (2012, April 14–16). *Setting response time thresholds for a CAT item pool: The normative threshold method* [Paper presentation]. The annual meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.

Some Extensions of IRTree Models

4

Section Learning Objectives

4 Extensions of IRTree models

Discuss some potential limitations of typical IRTree models

Discuss possible extensions of the IRTree model

Learning Objective 3

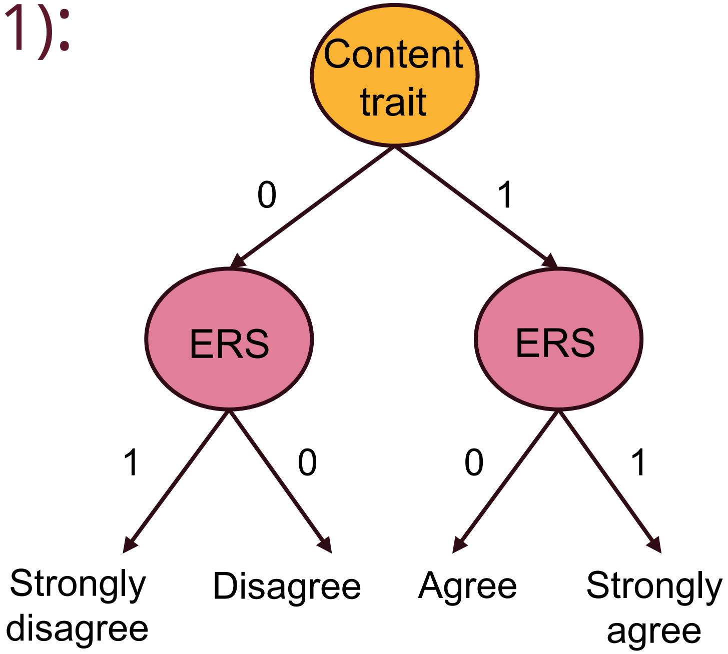
Learning Objective 4

Extension1: Mixture IRTree Models

- One limitation of a typical IRTree model is that it assumes the same underlying response process across all respondents (Kim & Bolt, 2021; Tijmstra, Bolsinova, & Jeon, 2018)
  - The tree structure and underlying latent traits are assumed to be identical for all respondents

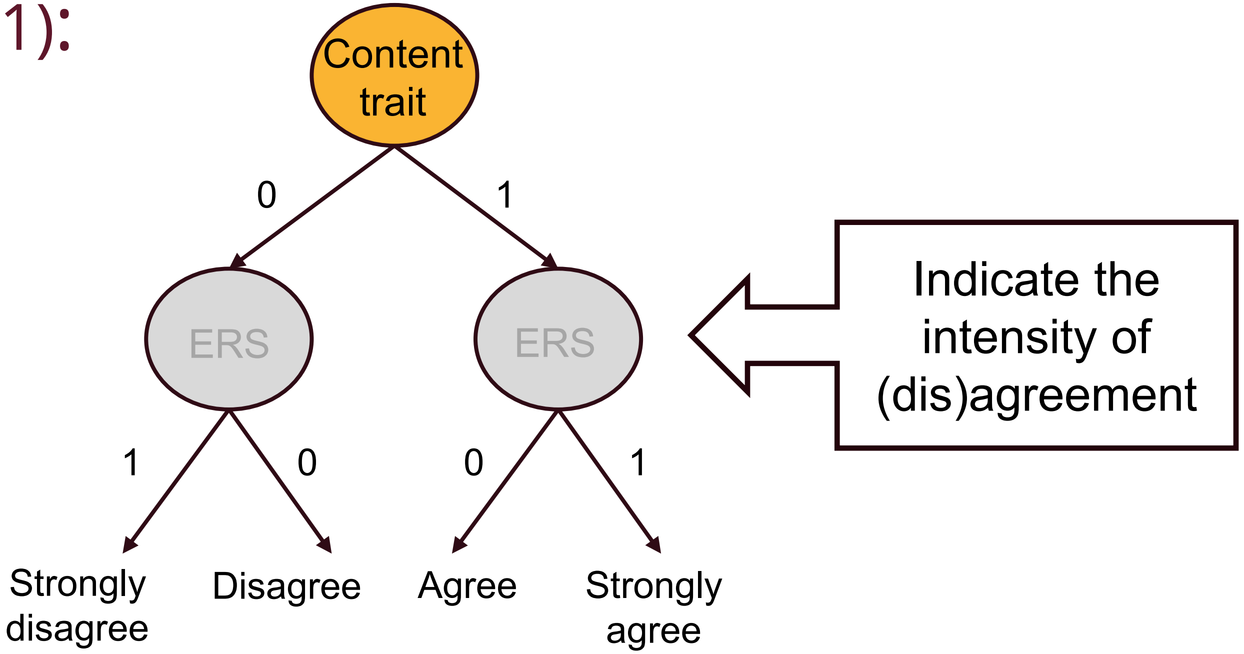
Extension1: Mixture IRTree Models

- One limitation of a typical IRTree model is that it assumes the same underlying response process across all respondents (Kim & Bolt, 2021; Tijmstra et al., 2018)
  - The tree structure and underlying latent traits are assumed to be identical for all respondents
- Example (Kim & Bolt, 2021):



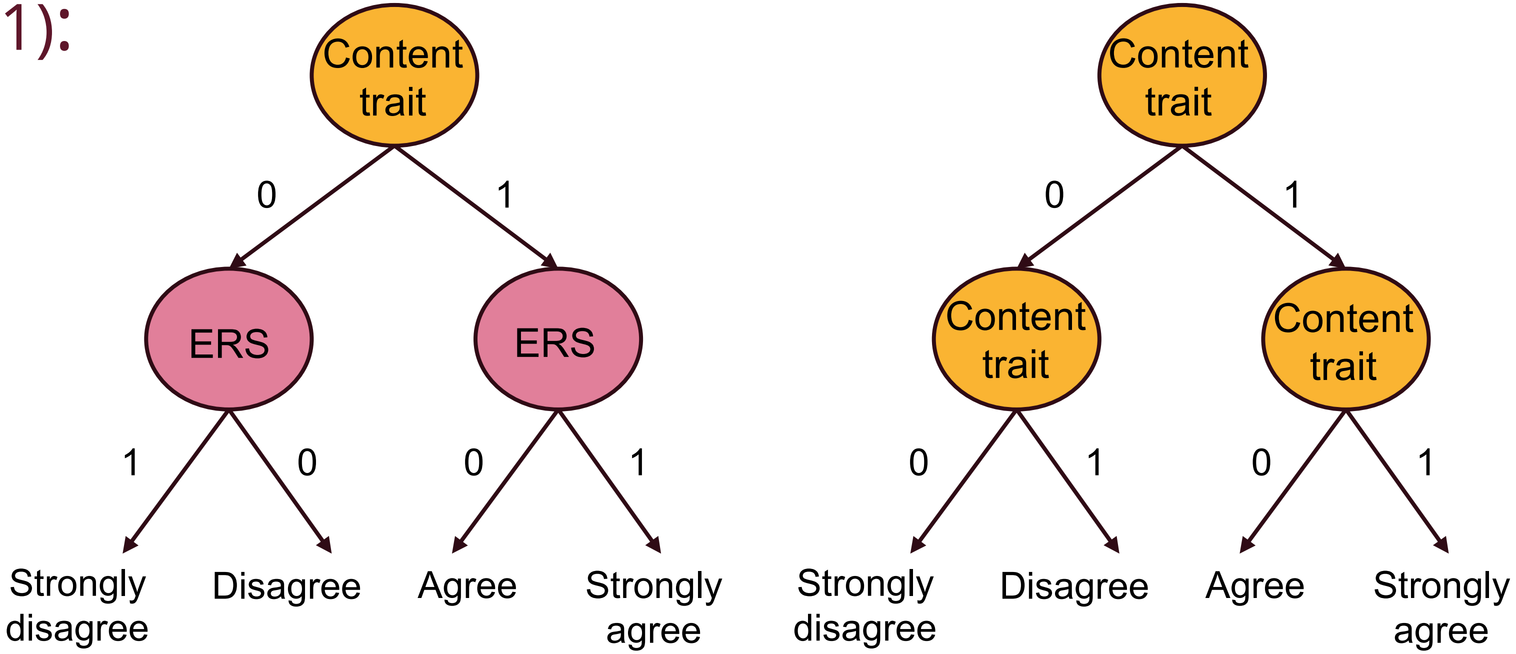
Extension1: Mixture IRTree Models

- One limitation of a typical IRTree model is that it assumes the same underlying response process across all respondents (Kim & Bolt, 2021; Tijmstra et al., 2018)
    - The tree structure and underlying latent traits are assumed to be identical for all respondents
- Example (Kim & Bolt, 2021):



Extension1: Mixture IRTree Models

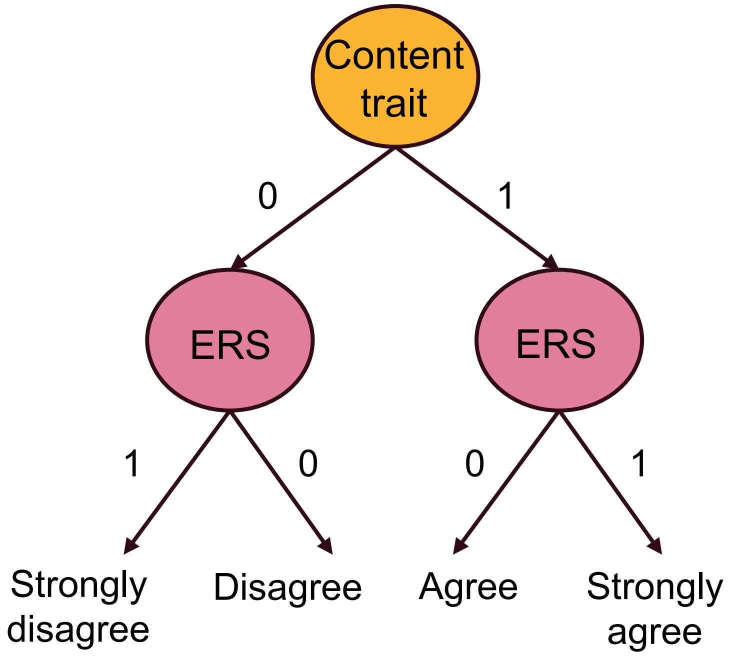
- One limitation of a typical IRTree model is that it assumes the same underlying response process across all respondents (Kim & Bolt, 2021; Tijmstra et al., 2018)
  - The tree structure and underlying latent traits are assumed to be identical for all respondents
- Example (Kim & Bolt, 2021):



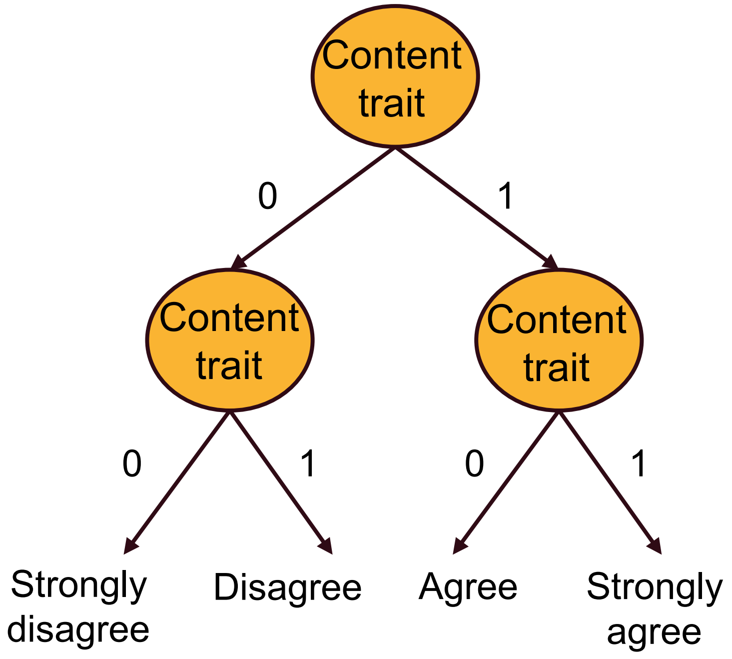
Extension1: Mixture IRTree Models

- We can formulate a mixture model incorporating the two possible IRTree models
- Each respondent is assumed to have a latent membership ($\boldsymbol{z_p}$) in one of the classes

Class 1 $(\boldsymbol{z_p} = \mathbf{0})$

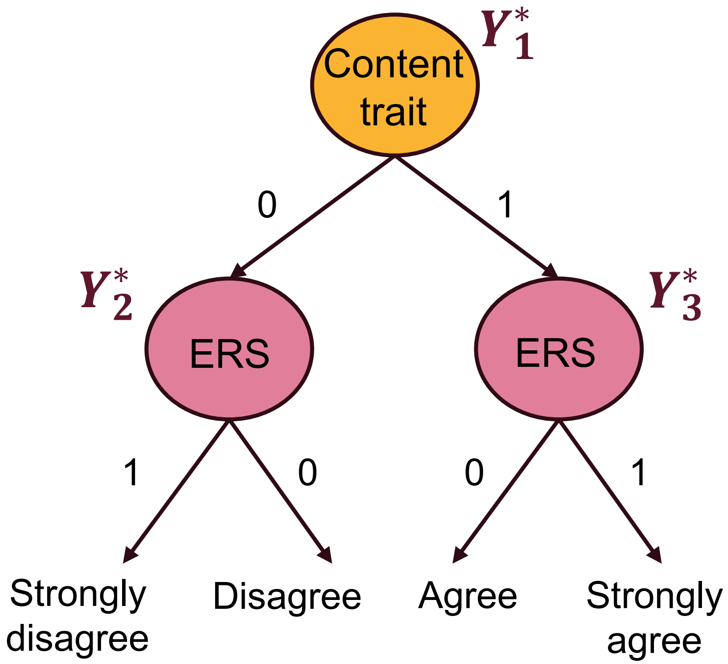


Class 2 $(\boldsymbol{z_p} = \mathbf{1})$



Extension1: Mixture IRTree Models

Class 1 $(z_p = 0)$



Content trait

$$P\left(Y_{pi1}^* = 1|\theta_p\right) = g^{-1}(a_{i1}\theta_p + d_{i1})$$

$$P\left(Y_{pi2}^* = 1|\,\eta_p\right) = g^{-1}(a_{i20}\eta_p + d_{i20})$$

$$P\left(Y_{pi3}^* = 1|\,\eta_p\right) = g^{-1}(a_{i30}\eta_p + d_{i30})$$

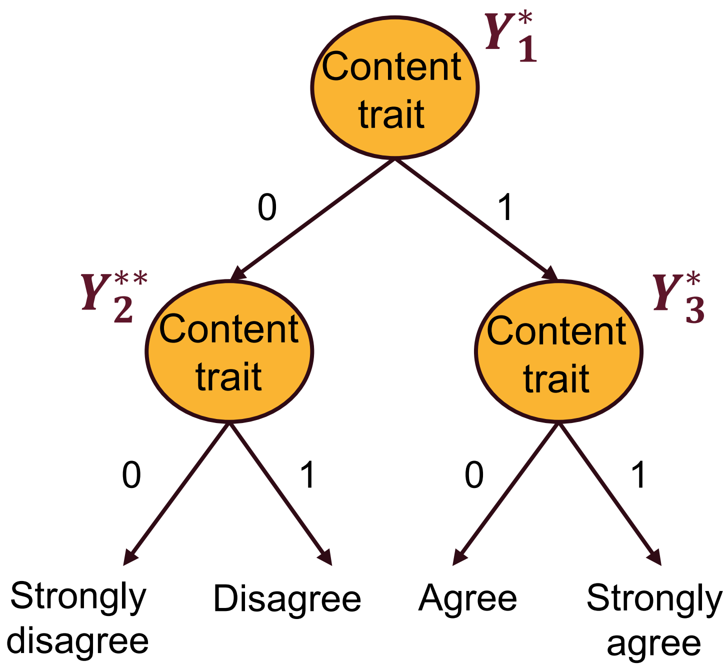
ERS

$$P\left(Y_{pi} = m\,\middle|\,\theta_p, \eta_p\right)$$

$$= P\left(Y_{pi1}^* = y_{m1}^*\,\middle|\,\theta_p\right) P\left(Y_{pi2}^* = y_{m2}^*\,\middle|\,\eta_p\right)^{1-y_{m1}^*} P\left(Y_{pi3}^* = y_{m3}^*\,\middle|\,\eta_p\right)^{y_{m1}^*}$$

Extension1: Mixture IRTree Models

Class 2 $(z_p = 1)$



Content trait

$$P(Y_{pi1}^* = 1|\theta_p) = g^{-1}(a_{i1}\theta_p + d_{i1})$$

$$P(Y_{pi2}^{**} = 1|\theta_p) = g^{-1}(a_{i21}\theta_p + d_{i21})$$

$$P(Y_{pi3}^* = 1|\theta_p) = g^{-1}(a_{i31}\theta_p + d_{i31})$$

$$P\left(Y_{pi} = m \middle| \theta_p\right)$$

$$= P\left(Y_{pi1}^* = y_{m1}^* \middle| \theta_p\right) P\left(Y_{pi2}^{**} = y_{m2}^{**} \middle| \theta_p\right)^{1-y_{m1}^*} P\left(Y_{pi3}^* = y_{m3}^* \middle| \theta_p\right)^{y_{m1}^*}$$

Extension1: Mixture IRTree Models

- Statistical representation of the mixture model:

Latent class membership of respondent p (0 = Class 1, 1 = Class2)

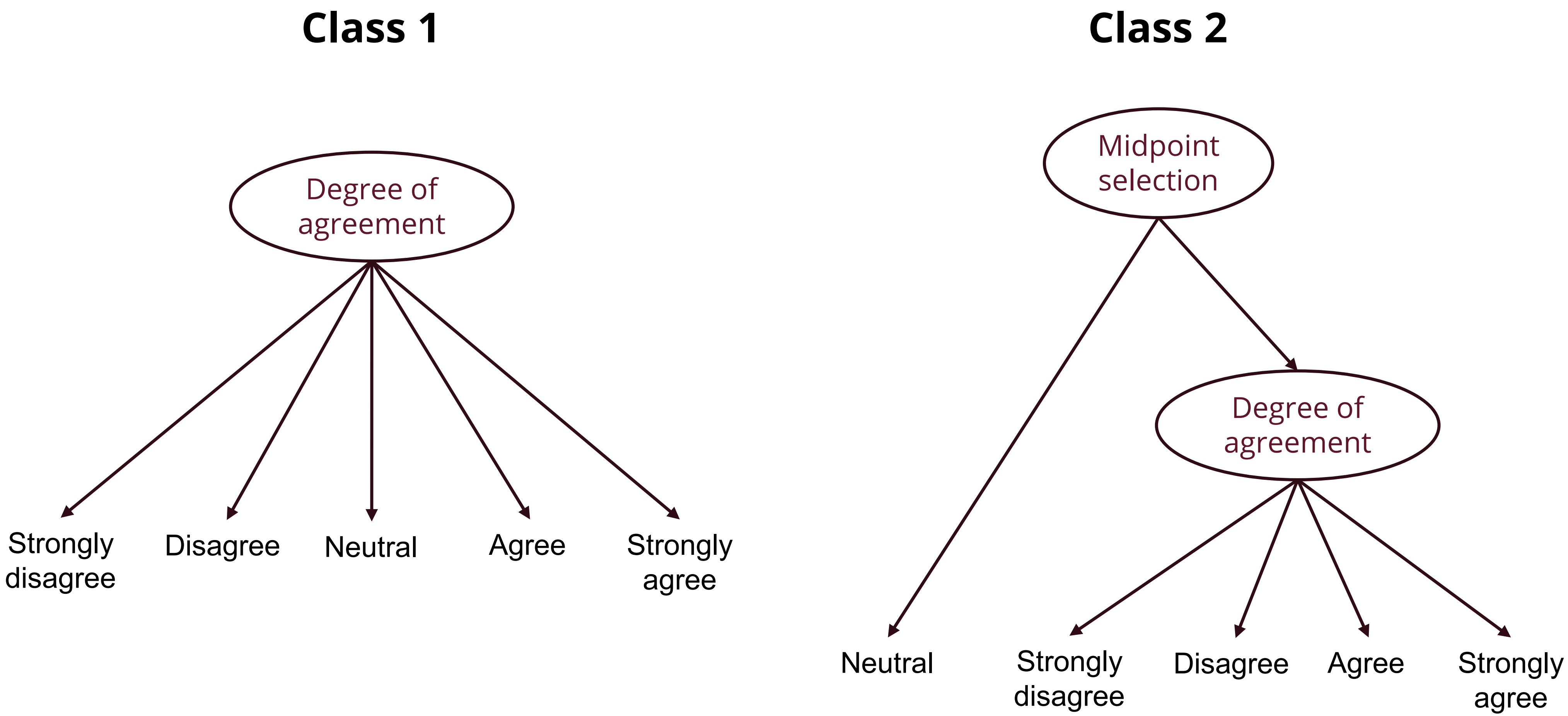
$$P\left(Y_{pi} = m \middle| \theta_p, \eta_p, z_p\right) = \underbrace{P\left(Y_{pi1}^{*} = y_{m1}^{*} \middle| \theta_p\right)}_{\text{Node 1}}$$

$$\times \left\{\underbrace{P\left(Y_{pi2}^{*} = y_{m2}^{*} \middle| \eta_p\right)^{1-y_{m1}^{*}} P\left(Y_{pi3}^{*} = y_{m3}^{*} \middle| \eta_p\right)^{y_{m1}^{*}}}_{\text{Nodes 2 and 3 for Class 1}}\right\}^{1-z_p}$$

$$\times \left\{\underbrace{P\left(Y_{pi2}^{**} = y_{m2}^{**} \middle| \theta_p\right)^{1-y_{m1}^{*}} P\left(Y_{pi3}^{*} = y_{m3}^{*} \middle| \theta_p\right)^{y_{m1}^{*}}}_{\text{Nodes 2 and 3 for Class 2}}\right\}^{z_p}$$

Extension1: Mixture IRTree Models

- Another example (Tijmstra et al., 2018):

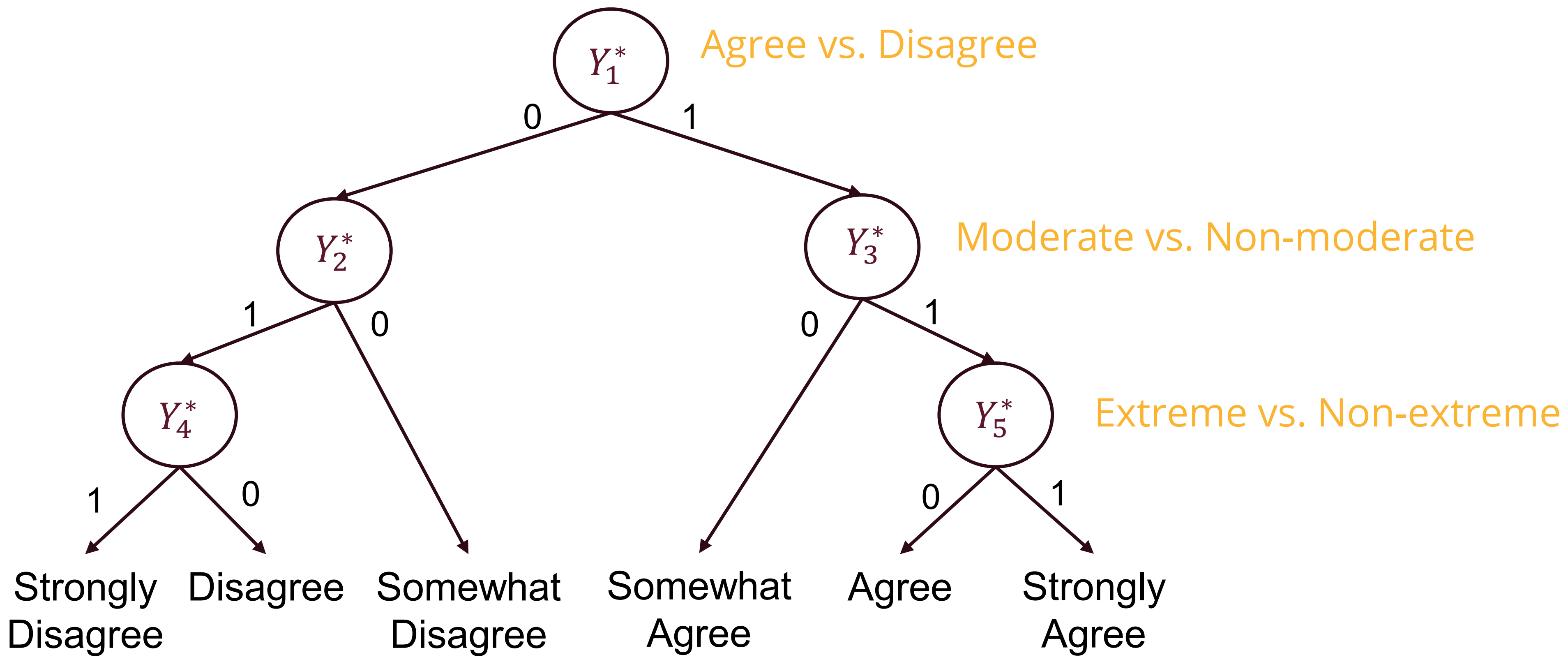


Extension2: Multidimensional IRTree Models

- Typical IRTree models assume a single latent trait for each node (i.e., a unidimensional IRT model is associated with each node)
- We can allow multiple latent traits to be involved for each node

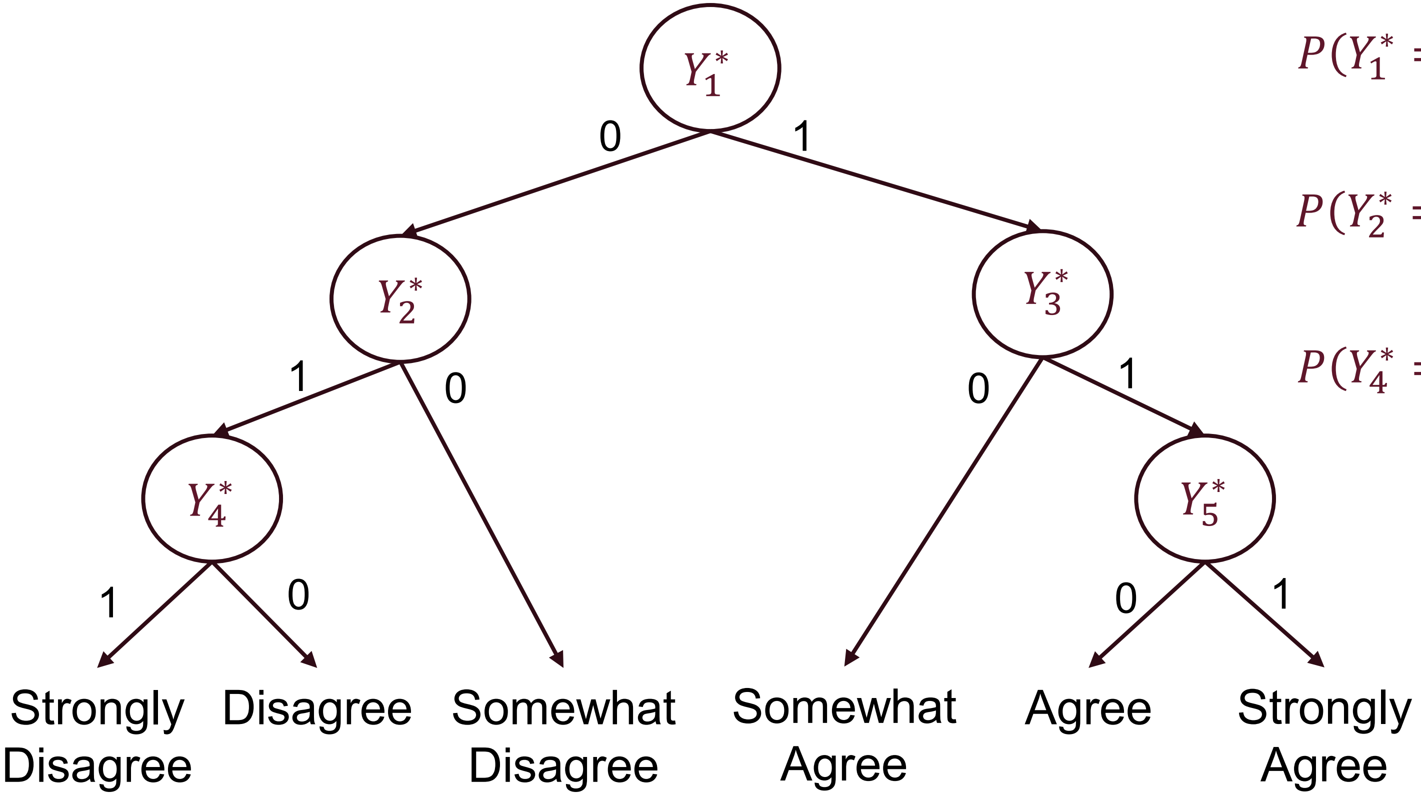
Extension2: Multidimensional IRTree Models

- Example (Meiser et al., 2019):



Extension2: Multidimensional IRTree Models

- Typically, nodes are specified with unidimensional IRT models



Content trait

$$P(Y_1^* = 1) = g^{-1}(\theta_p + d_{i1})$$

Tendency to select non-moderate categories

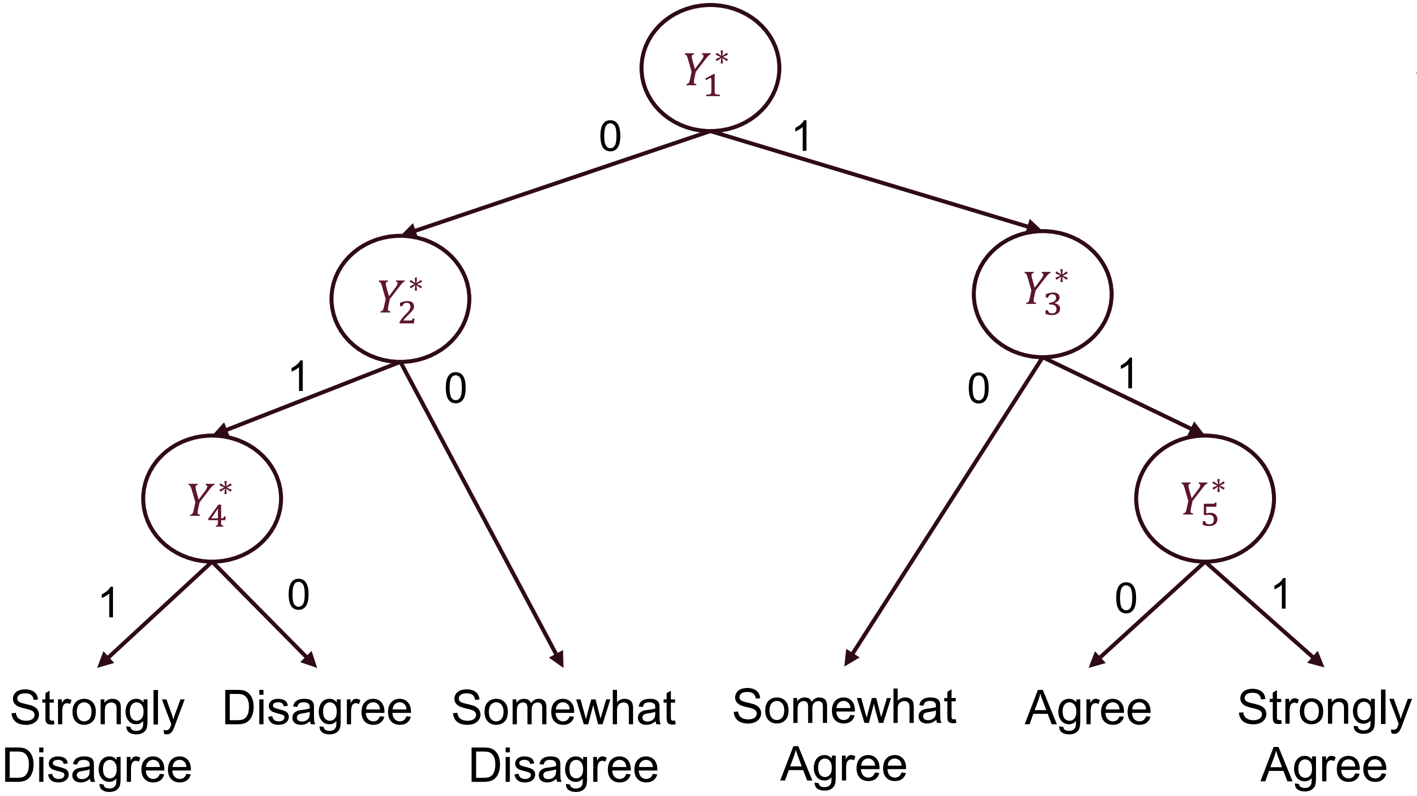
$$P(Y_2^* = 1) = P(Y_3^* = 1) = g^{-1}(\eta_{p1} + d_{i2})$$

$$P(Y_4^* = 1) = P(Y_5^* = 1) = g^{-1}(\eta_{p2} + d_{i3})$$

Tendency to select extreme categories

Extension2: Multidimensional IRTree Models

- We can extend the model by specifying decision nodes with multidimensional IRT models



$$P(Y_1^* = 1) = g^{-1}(\theta_p + d_{i1})$$

$$P(Y_2^* = 1) = g^{-1}(\eta_{p1} - \theta_p + d_{i2})$$

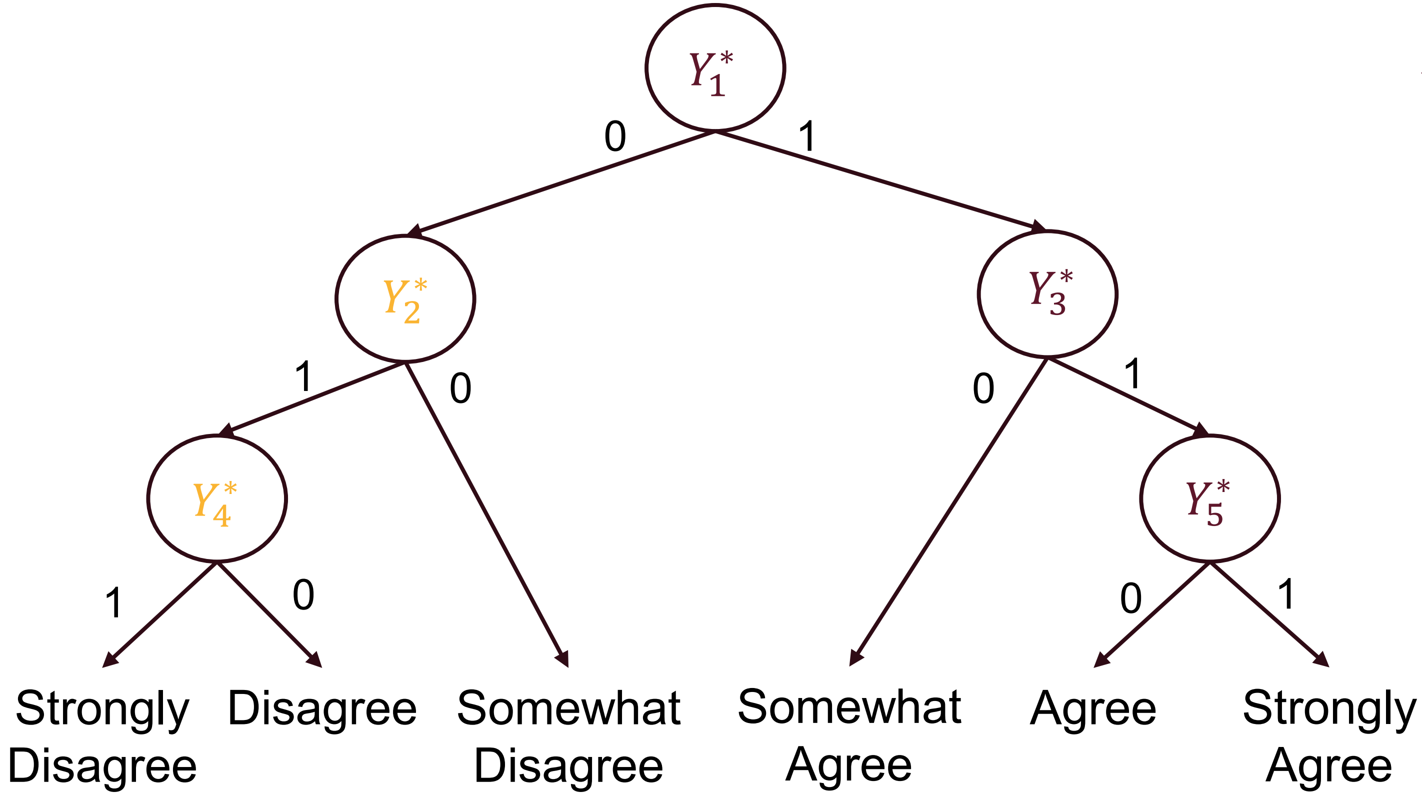
$$P(Y_3^* = 1) = g^{-1}(\eta_{p1} + \theta_p + d_{i2})$$

$$P(Y_4^* = 1) = g^{-1}(\eta_{p2} - \theta_p + d_{i3})$$

$$P(Y_5^* = 1) = g^{-1}(\eta_{p2} + \theta_p + d_{i3})$$

Extension2: Multidimensional IRTree Models

- We can extend the model by specifying decision nodes with multidimensional IRT models



$$P(Y_1^* = 1) = g^{-1}(\theta_p + d_{i1})$$

$$P(Y_2^* = 1) = g^{-1}(\eta_{p1} - \theta_p + d_{i2})$$

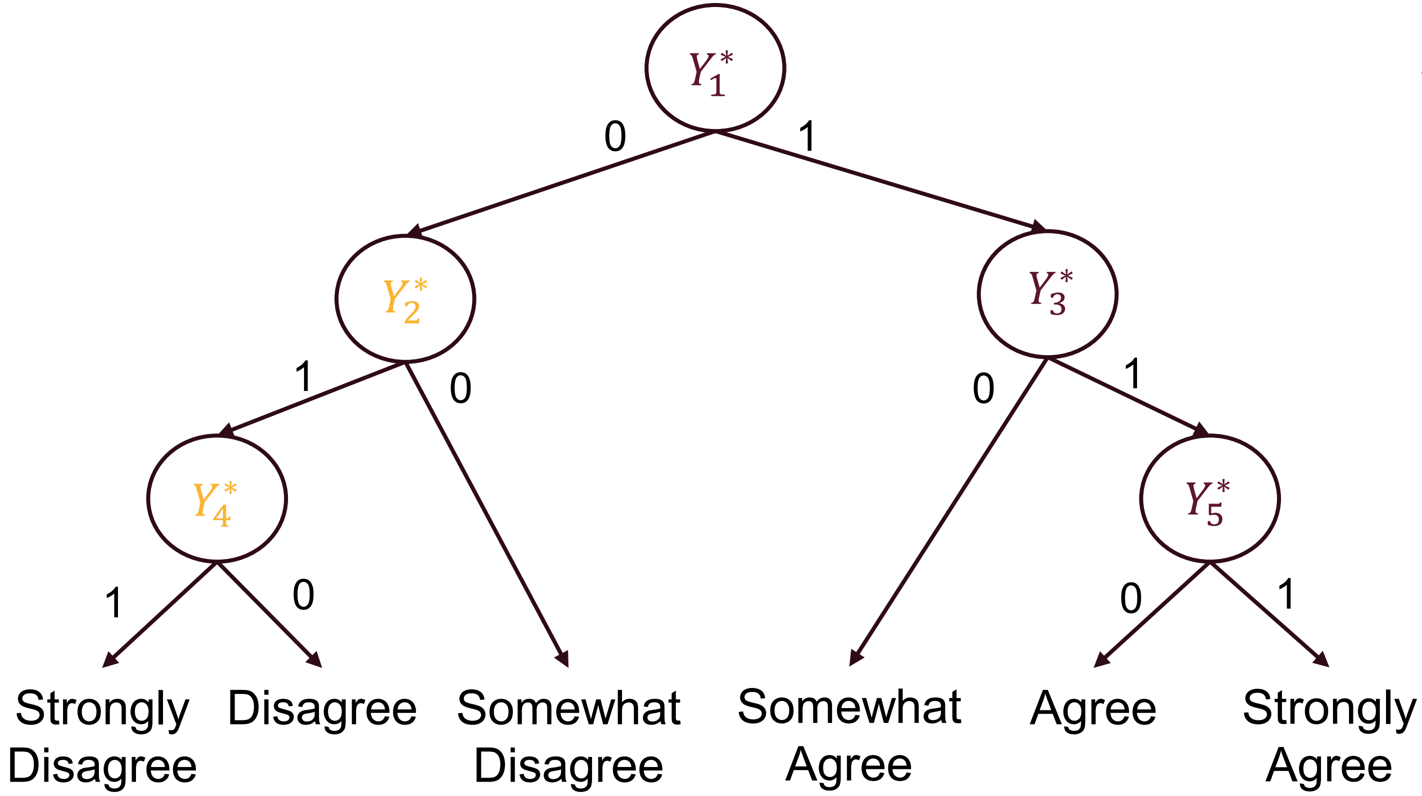
$$P(Y_3^* = 1) = g^{-1}(\eta_{p1} + \theta_p + d_{i2})$$

$$P(Y_4^* = 1) = g^{-1}(\eta_{p2} - \theta_p + d_{i3})$$

$$P(Y_5^* = 1) = g^{-1}(\eta_{p2} + \theta_p + d_{i3})$$

Extension2: Multidimensional IRTree Models

- We can extend the model by specifying decision nodes with multidimensional IRT models



$$P(Y_1^* = 1) = g^{-1}(\theta_p + d_{i1})$$

$$P(Y_2^* = 1) = g^{-1}(\eta_{p1} - \theta_p + d_{i2})$$

$$P(Y_3^* = 1) = g^{-1}(\eta_{p1} + \theta_p + d_{i2})$$

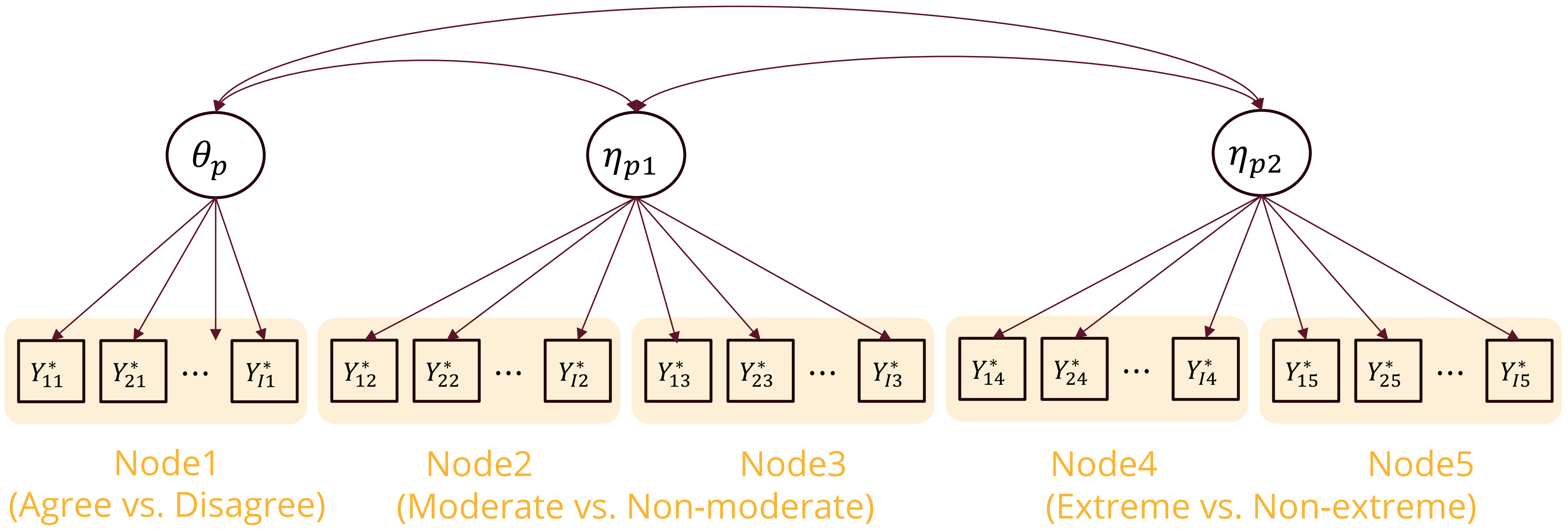
$$P(Y_4^* = 1) = g^{-1}(\eta_{p2} - \theta_p + d_{i3})$$

$$P(Y_5^* = 1) = g^{-1}(\eta_{p2} + \theta_p + d_{i3})$$

θ has negative effects on pseudo-items Y_2^* and Y_4^* (i.e., Higher values of theta decrease the probabilities)

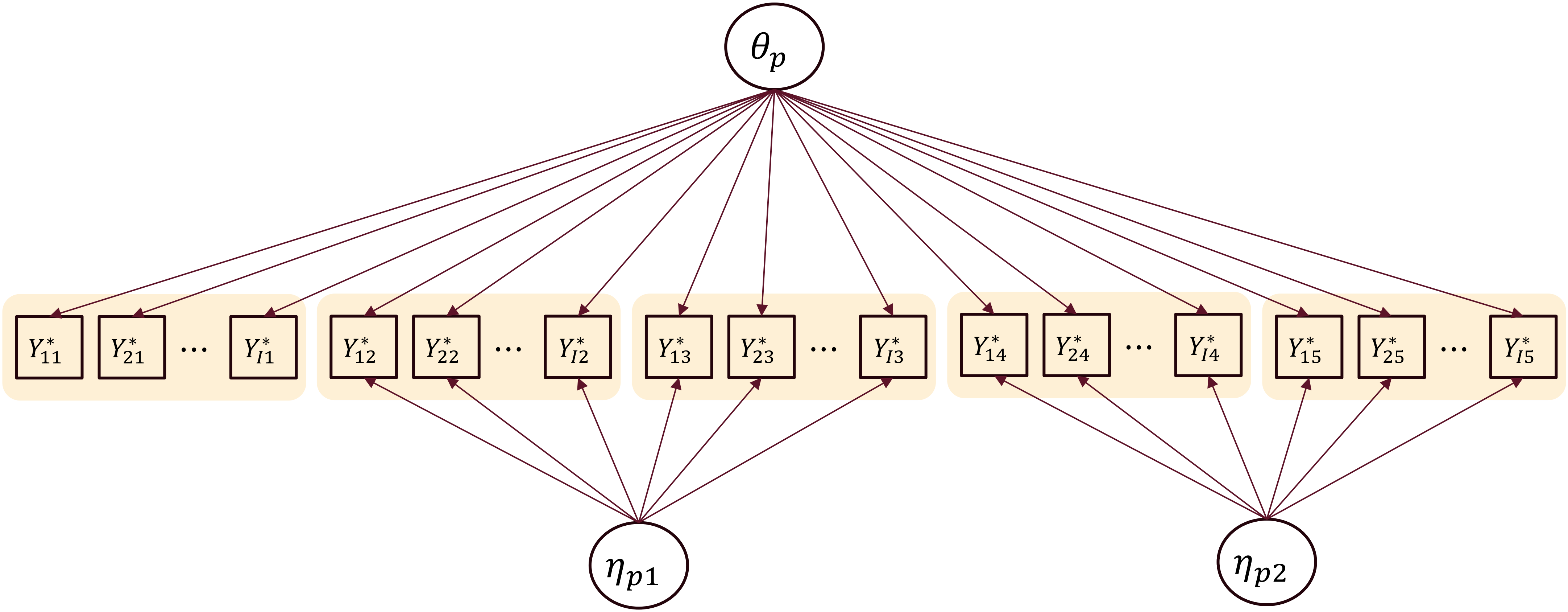
Extension2: Multidimensional IRTree Models

- Graphical representation of the IRTree model with *unidimensional* IRT models specified for internal nodes:



Extension2: Multidimensional IRTree Models

- Graphical representation of the IRTree model with *multidimensional* IRT models specified for internal nodes:



Extension3: Explanatory IRTree Models

- Descriptive IRTree Models
  - Focus on modeling the response process without providing insight into the sources of variability in node outcomes (i.e., pseudo-item responses)
  - Person and item <u>indicators</u> are included as predictors of node outcomes
- Explanatory IRTree Models
  - Extend descriptive IRTree models by incorporating person and/or item covariates to explain the variance in node outcomes within the tree
  - Person and item <u>property variables</u> are included as predictors of node outcomes

Extension3: Explanatory IRTree Models

- Probability of an outcome at node n for person p for item i (Wilson & De Boeck, 2004)

$$P\left(Y_{pin}^{*} = 1\right) = g^{-1}\left(a_{i1}\theta_{pn} + d_{in}\right)$$

$$\theta_{pn} = \sum_{j} \lambda_{jn} Z_{pj} + \eta_{pn}$$

$$\beta_{in} = \sum_{l} \gamma_{ln} X_{pl} + \delta_{in}$$

- Z and X are person and item covariates, respectively
- λ and γ are coefficients for person and item covariates, respectively
- η_{pn} and δ_{in} are person and item random effects, respectively

Extension3: Explanatory IRTree Models

- Advantages of Explanatory IRTree Models
  - Inclusion of person- and item-level covariates
  - Understanding of response processes and latent traits
  - Comparisons across subgroups
- Challenges of Explanatory IRTree Models
  - Model complexity and computational demands
  - Requirement of larger sample sizes
  - Possibility for confounded interpretations of parameters beyond the first node (Lyu et al., 2023)

References

- Kim, N. & Bolt, D. M. (2021). A mixture IRTree model for extreme response style: Accounting for response process uncertainty. *Educational and Psychological Measurement, 81*(1), 131–154. https://doi.org/10.1177/0013164420913915
- Lyu, W., Bolt, D.M. & Westby, S. Exploring the Effects of Item-Specific Factors in Sequential and IRTree Models. *Psychometrika 88*, 745–775 (2023). https://doi.org/10.1007/s11336-023-09912-x
- Meiser, T., Plieninger, H., & Henninger, M. (2019). IRTree models with ordinal and multidimensional decision nodes for response styles and trait-based rating responses. *British Journal of Mathematical and Statistical Psychology, 72*(3), 501-516. https: //doi.org/10.1111/bmsp.12158
- Tijmstra, J., Bolsinova, M., & Jeon, M. (2018). General mixture item response models with different item response structures: Exposition with an application to Likert scales. *Behavior Research Methods, 50*(6), 2325-2344. https://doi.org/10.3758/s13428-017-0997-0
- Wilson, M., & De Boeck, P. (2004). Descriptive and explanatory item response models. In *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 43-74). New York, NY: Springer New York.

R demonstration

5

Section Learning Objectives

5

R Demonstration

Recode observed item responses into pseudo-item responses using R

Fit a specified IRTree model using R

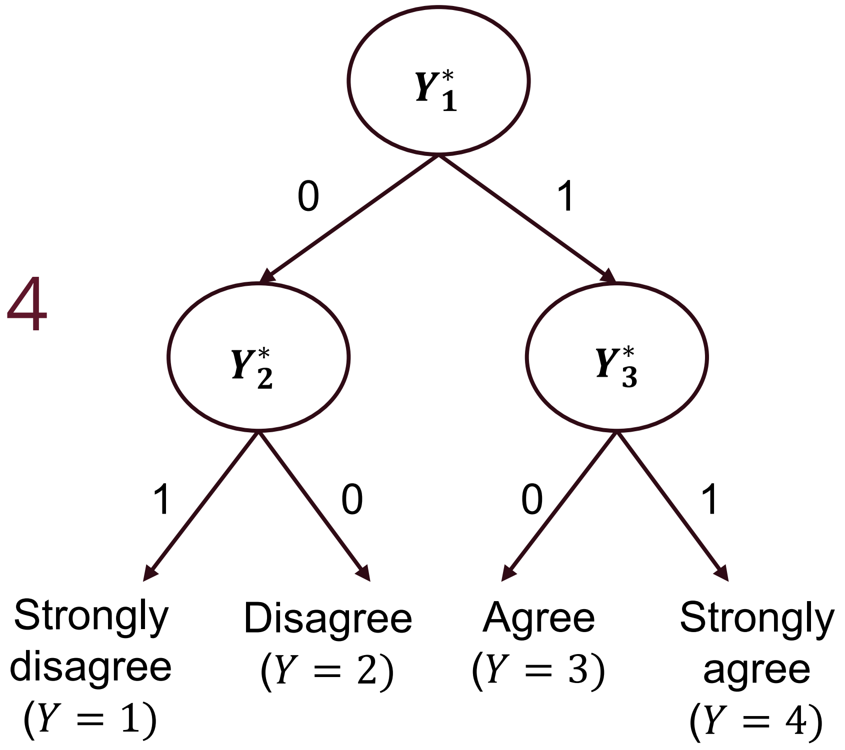
Interpret the R output obtained by fitting an IRTree model

Evaluate the model fit of an IRTree model

Data Preparation

Step 1 – Import dataset to analyze

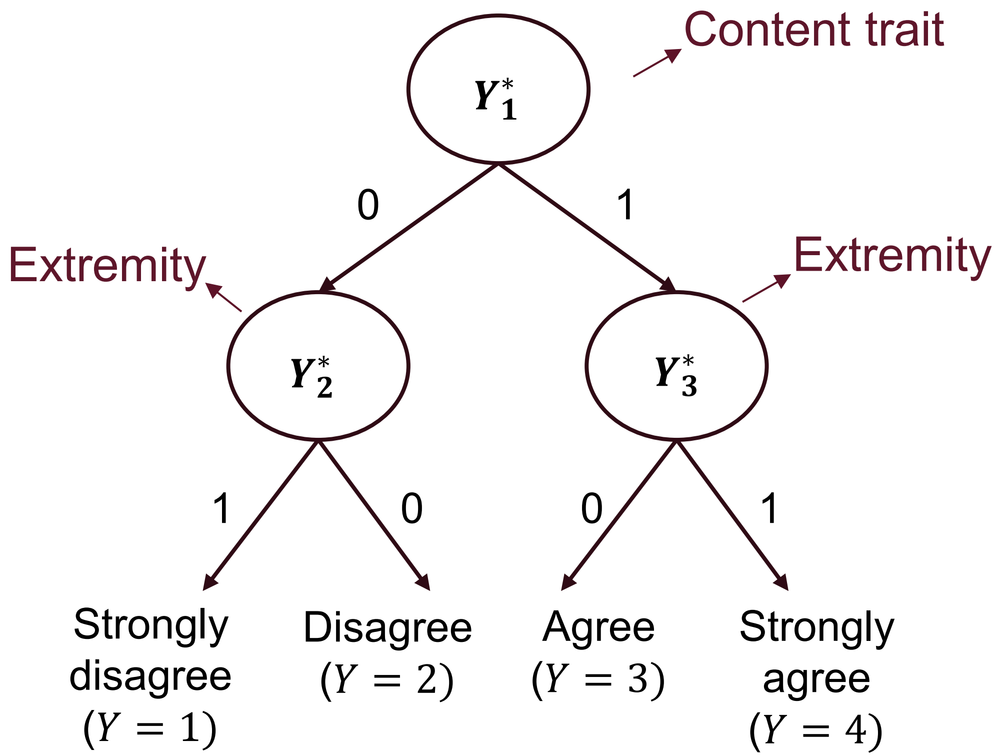
- Simulated item responses of 1,000 respondents to 15 rating scale items
- Response categories ranging from 1 to 4



Data Preparation

Step 2 – Transform item responses into pseudo-item responses

- Fit the same IRTree model to data for demonstration



Node	Terminal responses (Y)			
	1	2	3	4
Y_1^*	0	0	1	1
Y_2^*	1	0	NA	NA
Y_3^*	NA	NA	0	1

Data Preparation

Original dataset

ID	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11	Item12	Item13	Item14	Item15
1	3	1	1	1	4	1	3	2	4	4	4	1	1	2	1
2	3	3	1	2	3	1	3	2	3	4	4	3	1	2	2
3	3	3	1	2	3	1	1	2	4	1	4	4	1	2	2
4	3	4	1	1	4	1	3	2	4	2	1	4	1	1	2
5	3	3	1	2	3	2	2	4	3	3	4	3	1	2	2
6	3	3	1	1	3	1	4	2	3	4	2	3	2	1	2
7	3	4	1	2	3	1	3	2	1	2	3	2	2	2	2
8	4	1	1	1	3	1	1	1	1	1	4	1	1	2	1
9	3	3	1	2	3	2	3	2	1	1	2	3	2	2	2
10	4	3	3	1	1	4	1	1	4	4	2	1	4	4	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

IRTree dataset

ID	Item1.node1	Item2.node1	Item3.node1	Item4.node1	Item5.node1	...	Item1.node2	Item2.node2	Item3.node2	Item4.node2	Item5.node2	...	Item1.node3	Item2.node3	Item3.node3	Item4.node3	Item5.node3
1	1	1	0	0	1	...	NA	NA	1	0	NA	...	0	0	NA	NA	0
2	1	1	0	0	1	...	NA	NA	1	0	NA	...	0	0	NA	NA	0
3	1	1	0	0	1	...	NA	NA	1	1	NA	...	0	1	NA	NA	1
4	1	1	0	0	1	...	NA	NA	1	0	NA	...	0	0	NA	NA	0
5	1	1	0	0	1	...	NA	NA	1	1	NA	...	0	0	NA	NA	0
6	1	1	0	0	1	...	NA	NA	1	0	NA	...	0	1	NA	NA	0
7	1	0	0	0	1	...	NA	1	1	1	NA	...	1	NA	NA	NA	0
8	1	1	0	0	1	...	NA	NA	1	0	NA	...	0	0	NA	NA	0
9	1	1	1	0	0	...	NA	NA	NA	1	1	...	1	0	0	NA	NA
10	1	1	0	1	1	...	NA	NA	1	NA	NA	...	0	1	NA	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Data Preparation in R

```
> ## 1. Load packages
> library(mirt)
> # Install 'flirt' package by downloading the source file from the link:
> # https://sites.google.com/site/arbormj/sofware/flirt
> # install.packages("flirt_1.15.tar.gz", type="source", repos=NULL, header = F, quiet = T)
> library(flirt)
>
> ## 2. Read in simulated data
> data <- read.csv("simdata.csv", header = F)
> names(data) <- paste("Item", 1:ncol(data), sep = "")
> head(data)
  Item1 Item2 Item3 Item4 Item5 Item6 Item7 Item8 Item9 Item10 Item11 Item12 Item13 Item14 Item15
1     1     3     4     2     3     1     3     2     1      1      3      4      1      1      1
2     2     2     2     2     4     2     3     4     2      1      3      4      2      1      1
3     1     4     1     1     3     2     3     3     1      1      3      4      1      1      1
4     1     4     3     1     3     2     1     1     2      1      3      4      3      1      1
5     2     4     4     1     3     3     1     3     4      4      2      4      2      2      4
6     2     3     3     2     3     3     3     3     2      1      3      4      3      1      3
>
> ## 3. Recode item responses
> # Define mapping matrix based on the specified tree structure
> mapping <- matrix(c(0, 0, 1, 1,
+                     1, 0, NA, NA,
+                     NA, NA, 0, 1), 4, 3)
> mapping
     [,1] [,2] [,3]
[1,]    0    1   NA
[2,]    0    0   NA
[3,]    1   NA    0
[4,]    1   NA    1
```

Install and load ‘mirt’ and ‘flirt’ packages

Load simulated data

The first six rows of the simulated data

Define mapping matrix

Rows represent observed response categories and columns represent pseudo-items (decision nodes of the tree structure)

Data Preparation in R

```
> # Transform observed item responses into pseudo-item responses
> pseudo <- dendrify2(data, mapping, wide=T)[-1] #remove the first column containing person indicator
> pseudo[1:2, ]
  value.i01:node1 value.i02:node1 value.i03:node1 value.i04:node1 value.i05:node1 value.i06:node1 value.i07:node1
1               0               1               1               0               1               0               1
2               0               0               0               0               1               0               1
  value.i08:node1 value.i09:node1 value.i10:node1 value.i11:node1 value.i12:node1 value.i13:node1 value.i14:node1
1               0               0               0               1               1               0               0
2               1               0               0               1               1               0               0
  value.i15:node1 value.i01:node2 value.i02:node2 value.i03:node2 value.i04:node2 value.i05:node2 value.i06:node2
1               0               1              NA              NA               0              NA               1
2               0               0               0               0               0              NA               0
  value.i07:node2 value.i08:node2 value.i09:node2 value.i10:node2 value.i11:node2 value.i12:node2 value.i13:node2
1              NA               0               1               1              NA              NA               1
2              NA              NA               0               1              NA              NA               0
  value.i14:node2 value.i15:node2 value.i01:node3 value.i02:node3 value.i03:node3 value.i04:node3 value.i05:node3
1               1               1              NA               0               1              NA               0
2               1               1              NA              NA              NA              NA               1
  value.i06:node3 value.i07:node3 value.i08:node3 value.i09:node3 value.i10:node3 value.i11:node3 value.i12:node3
1              NA               0              NA              NA              NA               0               1
2              NA               0               1              NA              NA               0               1
  value.i13:node3 value.i14:node3 value.i15:node3
1              NA              NA              NA
2              NA              NA              NA
```

Transform observed item responses into pseudo-item responses based on the defined mapping matrix

First two rows of the recoded data

Fitting IRTree models in R

```
> ## 4. Fit an IRTree model
> # Specify a model
> erstree <- mirt.model('F1 = 1-15
+                        ERS = 16-45')
>
> # Fit the specified MIRT model to pseudo-item response data
> erstree.fit <- mirt(pseudo, erstree, itemtype = "2PL", method = "MHRM", verbose = F)
> erstree.fit

Call:
mirt(data = pseudo, model = erstree, itemtype = "2PL", method = "MHRM",
    verbose = F)

Full-information item factor analysis with 2 factor(s).
Converged within 0.001 tolerance after 363 MHRM iterations.
mirt version: 1.37.1
M-step optimizer: NR1
Latent density type: Gaussian
Average MH acceptance ratio(s): 0.337

Log-likelihood = -13272.65, SE = 0.042
Estimated parameters: 90
AIC = 26725.3
BIC = 27166.99; SABIC = 26881.15
```

Specify the MIRT model.
In this model, F1 influences pseudo-items 1-15 and ERS factor influences pseudo-items 16-45

Fit the specified model to the pseudo-item response data using 'mirt()' function

Fitting IRTree models in R

```
> ## 5. Print and interpret output
> # Extract item parameter estimates
> coef(erstree.fit, simplify = T)
$items
                      a1    a2      d g u
value.i01:node1    1.618 0.000 -2.751 0 1
value.i02:node1    0.417 0.000  2.585 0 1
value.i03:node1    1.405 0.000  0.601 0 1
                     ⋮
value.i01:node2    0.000 0.890  0.476 0 1
value.i02:node2    0.000 3.985 -2.505 0 1
value.i03:node2    0.000 0.625 -1.333 0 1
                     ⋮
value.i01:node3    0.000 0.302  2.690 0 1
value.i02:node3    0.000 1.689  0.199 0 1
value.i03:node3    0.000 1.726 -1.349 0 1
                     ⋮

$means
 F1  ERS
  0    0

$cov
     F1  ERS
F1    1    0
ERS   0    1
```

Extract item parameter estimates for each item and node

The item parameter estimates for the first node are the item discriminations and easinesses for disagree versus agree decisions

The item parameter estimates for the second and third nodes are item discriminations and easinesses for extreme versus nonextreme response selections

Fitting IRTree models in R

```
> # Extract person parameter estimates
> fscores(erstree.fit)
              F1          ERS
 [1,] -0.334658929  0.5515306222
 [2,] -0.518829750 -0.2550071676
 [3,] -0.419892391  0.8891145005
 [4,] -0.549804942  0.2738073426
 [5,]  0.457177994  0.2344404716
 [6,]  0.707501264 -1.0554917740
 [7,] -1.444349484  0.4916299264
 [8,] -0.669722334  0.3034440409
 [9,] -1.905848880 -0.7688419705
[10,] -0.394133607  0.0486740162
[11,]  0.537653860 -0.0750325511
[12,]  0.209769267  0.5499973694
[13,] -0.212443243  0.9786005782
[14,]  0.947774231  0.4009591562
[15,]  1.053479875  0.7962604576
[16,]  0.544283787 -0.3088906274
[17,] -1.144823983  1.1296401966
[18,] -0.772237124 -0.4474516599
[19,]  0.681032644 -0.0727261493
[20,]  0.168252669 -0.7736385308
[21,] -1.466507184  0.0332824171
[22,] -1.905848880  0.2969688898
[23,] -0.728652123  0.7233197261

                  ⋮
```

Slightly low level of content trait

Slightly high level of extreme response style

Extract the estimates for person parameters

F1 associated with the first nodes represents individuals' trait/attitude being measured (information given by agree versus disagree decisions)

ERS associated with the second and third nodes represent individuals' extreme response styles (information given by extreme versus non-extreme response selections)

Fitting IRTree models in R

```
> ## 6. Model comparison
> # Specify another model assuming distinct ERS factors for the second and third nodes
> erstree2 <- mirt.model('F1 = 1-15
+                         ERS1 = 16-30
+                         ERS2 = 31-45
+                         COV = ERS1*ERS2')

> # Fit the specified MIRT model to pseudo-item response data
> erstree2.fit <- mirt(pseudo, erstree2, itemtype = "2PL", method = "MHRM", verbose = F)
> erstree2.fit

Call:
mirt(data = pseudo, model = erstree2, itemtype = "2PL", method = "MHRM",
    verbose = F)

Full-information item factor analysis with 3 factor(s).
Converged within 0.001 tolerance after 606 MHRM iterations.
mirt version: 1.37.1
M-step optimizer: NR1
Latent density type: Gaussian
Average MH acceptance ratio(s): 0.358

Log-likelihood = -13220.8, SE = 0.042
Estimated parameters: 91
AIC = 26623.59
BIC = 27070.2; SABIC = 26781.18
```

Let’s fit another model that assumes two distinct ERS factors for ‘agree’ and ‘disagree’ directions.

In this model, F1 influences pseudo-items 1-15 and ERS1 influences pseudo-items 16-30, and ERS2 influences pseudo-items 31-45

Fitting IRTree models in R

```
> # Extract item parameter estimates
> coef(erstree2.fit,  simplify = T)
$items
                     a1    a2    a3      d g u
value.i01:node1 1.627 0.000 0.000 -2.766 0 1
value.i02:node1 0.427 0.000 0.000  2.586 0 1
value.i03:node1 1.410 0.000 0.000  0.593 0 1

                          ⋮

value.i01:node2 0.000 0.989 0.000  0.484 0 1
value.i02:node2 0.000 9.583 0.000 -5.769 0 1
value.i03:node2 0.000 0.666 0.000 -1.352 0 1
                          ⋮
value.i01:node3 0.000 0.000 0.337  2.694 0 1
value.i02:node3 0.000 0.000 2.068  0.237 0 1
value.i03:node3 0.000 0.000 1.915 -1.420 0 1

                          ⋮

$means
  F1 ERS1 ERS2
   0    0    0

$cov
     F1 ERS1 ERS2
F1    1 0.00 0.00
ERS1  0 1.00 0.68
ERS2  0 0.68 1.00
```

Extract item parameter estimates for each item and node

Fitting IRTree models in R

Slightly low level of content trait

```
> # Extract person parameter estimates
> fscores(erstree2.fit)
               F1          ERS1          ERS2
 [1,] -0.321434728  0.7427341920  0.2581136138
 [2,] -0.515152844 -0.5439136852  0.4405966706
 [3,] -0.409550273  1.0904632858  0.5861067185
 [4,] -0.550150240  0.3966347235  0.1277569490
 [5,]  0.458872673 -0.3537361108  0.5889437766
 [6,]  0.722150813 -0.9503548875 -0.9807148447
 [7,] -1.437271782  0.4311806353  0.4338349483
 [8,] -0.667624348  0.2543352124  0.3047659460
 [9,] -1.899808372 -0.5910752504 -1.2244392606
[10,] -0.385533710 -0.5889749657  0.6441684843
[11,]  0.552669404 -0.1055991206 -0.0219547601
[12,]  0.209534817  0.4757477773  0.5491643289
[13,] -0.200639206  0.3993428322  1.2705164159
[14,]  0.948592689  0.1694353505  0.4469360593
[15,]  1.049795948  0.8768873785  0.7387307385
[16,]  0.557459716 -0.1147982593 -0.3327881550
[17,] -1.137889648  1.1639564960  0.9944221484
[18,] -0.770838289 -0.3331329288 -0.8529970881
[19,]  0.695612484  0.2377235848 -0.2840171074
[20,]  0.162663989 -0.9406082067 -0.5256668802
[21,] -1.461537098 -0.0661735485  0.3090372910
[22,] -1.899808372  0.1835163130  0.4797128946
[23,] -0.725762984  0.9315873650  0.4870389237
[24,] -1.437271782 -0.7996530915 -0.0525644966
[25,] -1.410802621  0.2297390996 -0.2491400870
```

⋮

Extract person parameter estimates for each node

Slightly high levels of extreme response style for agree and disagree directions, respectively.

Note that ERS trait for “agree” direction is a bit higher than that for “disagree” direction for this person.

Model Comparison

```
> # Model fit comparison
> anova(erstree.fit, erstree2.fit)
                  AIC     SABIC        HQ       BIC    logLik      X2 df p
erstree.fit  26725.29 26881.15 26893.17 27166.99 -13272.65
erstree2.fit 26623.59 26781.18 26793.33 27070.20 -13220.80 103.703  1 0
```

Compare the two fitted models

The second model appears to fit better than the first model, suggesting two distinct ERS factors for agree and disagree directions for this data (rather than a single ERS for both directions of responses)

References

- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of statistical Software,* 48, 1-29.
- Jeon, M., & Rijmen, F. (2016). A modular approach for item response theory modeling with the R package flirt. *Behavior Research Methods,* 48, 742-755.