

オンライン学習・ アダプティブラーニングへの応用

電気通信大学 情報理工学研究科
堤 瑛美子

学習支援における課題

教師は学習者に教えすぎても、教えなさすぎても
学習者の能力成長は望めない。

基礎的な学習では丁寧に教えて良いが
高度な発達段階では教えすぎてしまうと自分で考える力が伸びない



自力で解決できる最小限の支援をすることが重要

良い教師とは…

与えた問題や支援によって
学習者の力が伸びるかどうかが正しく予測できる

学習支援における課題

良い教師とは
与えた問題や支援によって学習者の力が伸びるかどうか正しく予測できること

➤ 問題・ヒント・学習者の特性をよく理解している教師でなければ難しい

さらに,

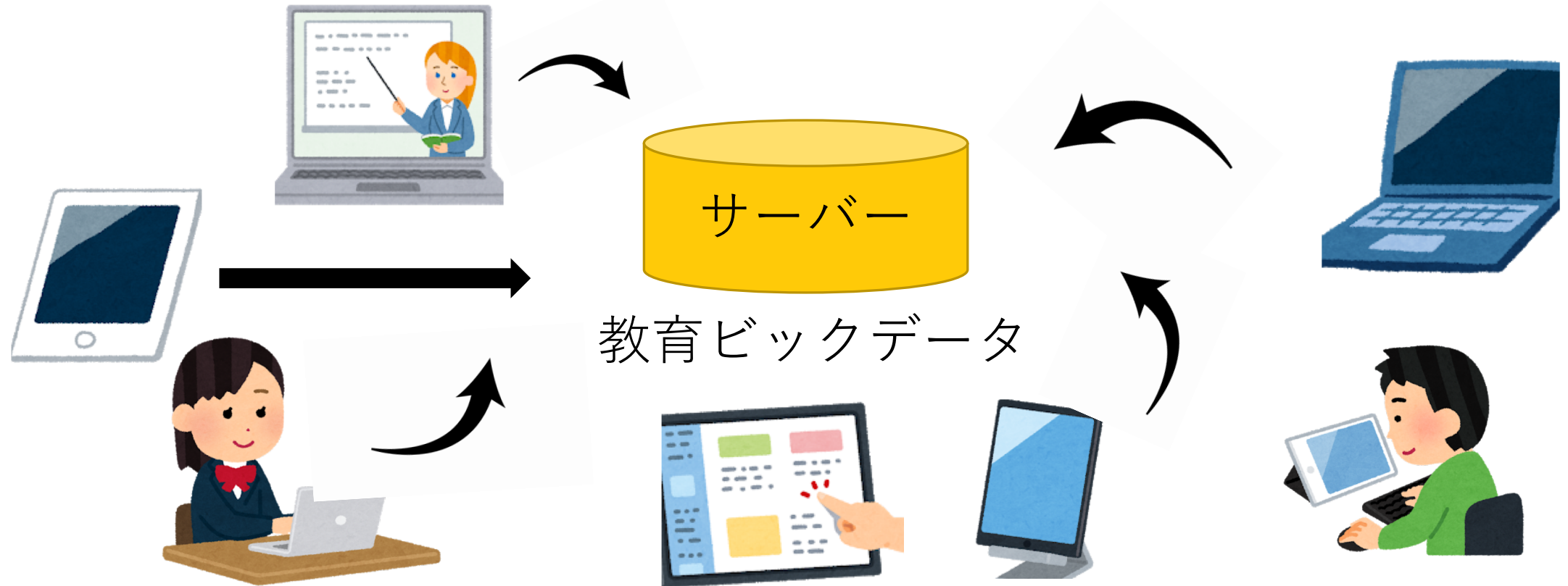
➤ 1人の学習者に合った教え方が他の人にも合うとは限らない



個々の学習者に合った教育を行うため
学習者の能力成長の予測にはAIを用いた方が良い

アダプティブラーニング

教育ビッグデータを分析することにより，学習過程における学習者の知識への習熟度や理解度を把握し，適応的学習環境を実現すること



Knowledge Tracing

学習者の課題への反応から学習者が未習熟な知識やスキルを把握する手法
基礎知識の学習に有効な学習支援

スキル

算数であれば
加算, 減算, 乗算, 除算, … など



- 加算のスキルが不十分な場合…
- 加算の問題を出題する
 - 加算に関する教材を見せる

機械学習手法による
習熟度推定



学習履歴データ

Knowledge Tracing

確率的アプローチ

- Bayesian Knowledge Tracing (BKT)
- Item Response Theory (IRT)

ディープラーニングアプローチ

- Deep Knowledge Tracing (DKT)
- Dynamic Key-Value Memory Network (DKVMN)
- Deep-IRT

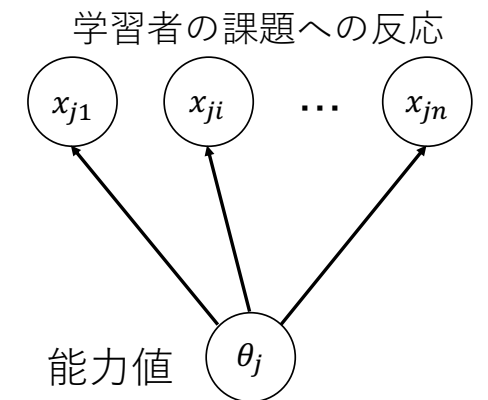
Knowledge Tracing

項目反応理論 Item Response Theory (IRT)

- 学習過程での学習者の習熟度をより詳細に把握するための手法
- 学習データから学習者の能力値を推定すると同時に、各課題への正答確率や反応(正答・誤答)を予測することが可能



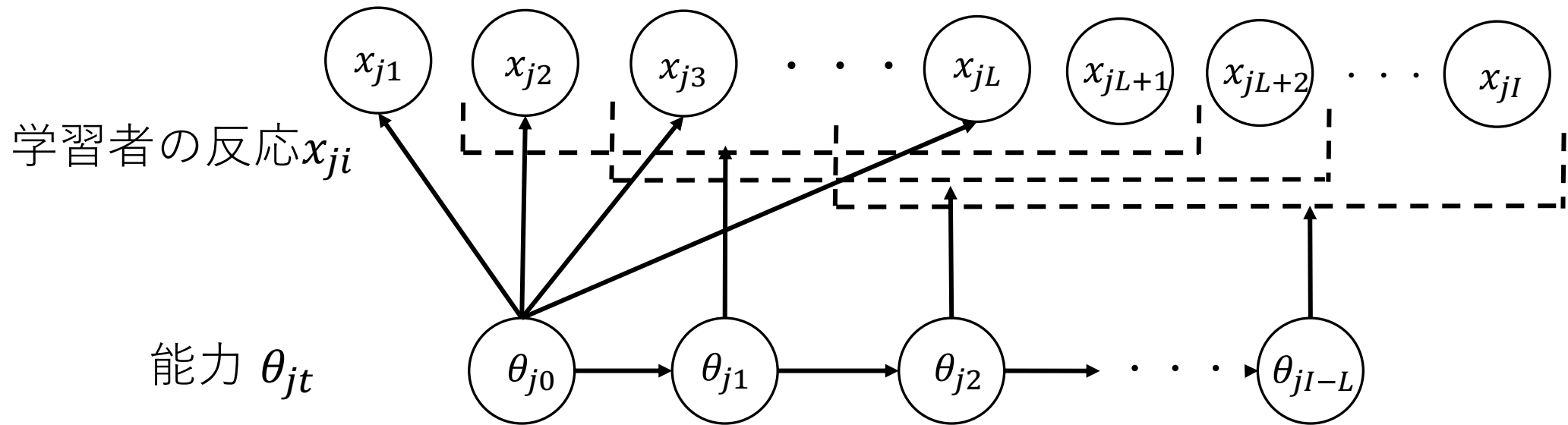
未習熟の知識が含まれる課題を同定
支援の必要性や程度を判断



隠れマルコフIRT(HM-IRT) [堤ら,2020]

堤瑛美子, 木下涼, 植野 真臣: Knowledge TracingのためのSliding Window隠れマルコフIRT, 電子情報通信学会論文誌 D, Vol.J103, No.12, pp.894-905 (2020)

- 従来のIRTモデルを拡張したウィンドウ型隠れマルコフIRT
- 学習過程における能力の成長を加味して知識の状態を推定できる



隠れマルコフIRT(HM-IRT) [堤ら,2020]

堤瑛美子, 木下涼, 植野 真臣 : Knowledge TracingのためのSliding Window隠れマルコフIRT, 電子情報通信学会論文誌 D, Vol.J103, No.12, pp.894-905 (2020)

$$P(x_{ji} = 1 | \theta_{jt}) = \frac{1}{1 + \exp(-a_i(\theta_{jt} - b_i))}$$

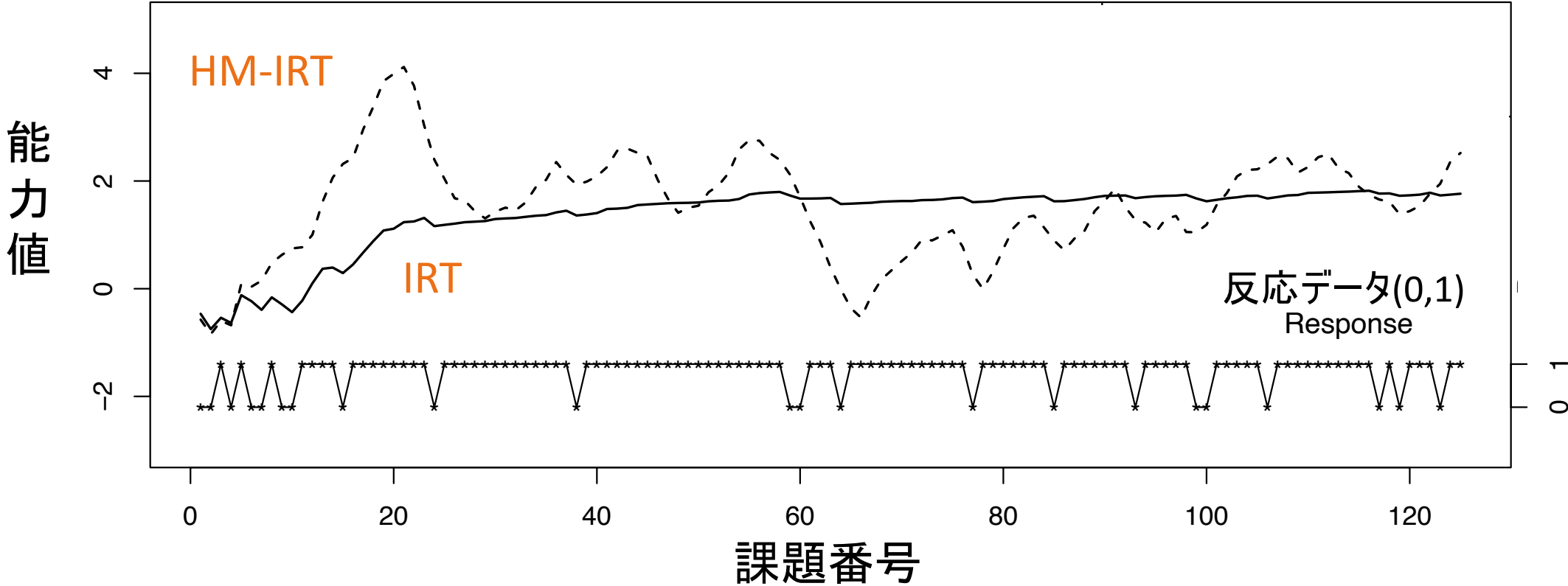
$$\theta_{jt} \sim N(\theta_{jt-1}, \delta)$$

δ : 知識状態の遷移幅
 a : 識別力パラメータ
 b : 難易度パラメータ

パラメータ推定法 : MCMC法

隠れマルコフIRT(HM-IRT) [堤ら,2020]

HM-IRTで推定した学習者の能力値変化



隠れマルコフIRT(HM-IRT) [堤ら,2020]

HM-IRTで推定した

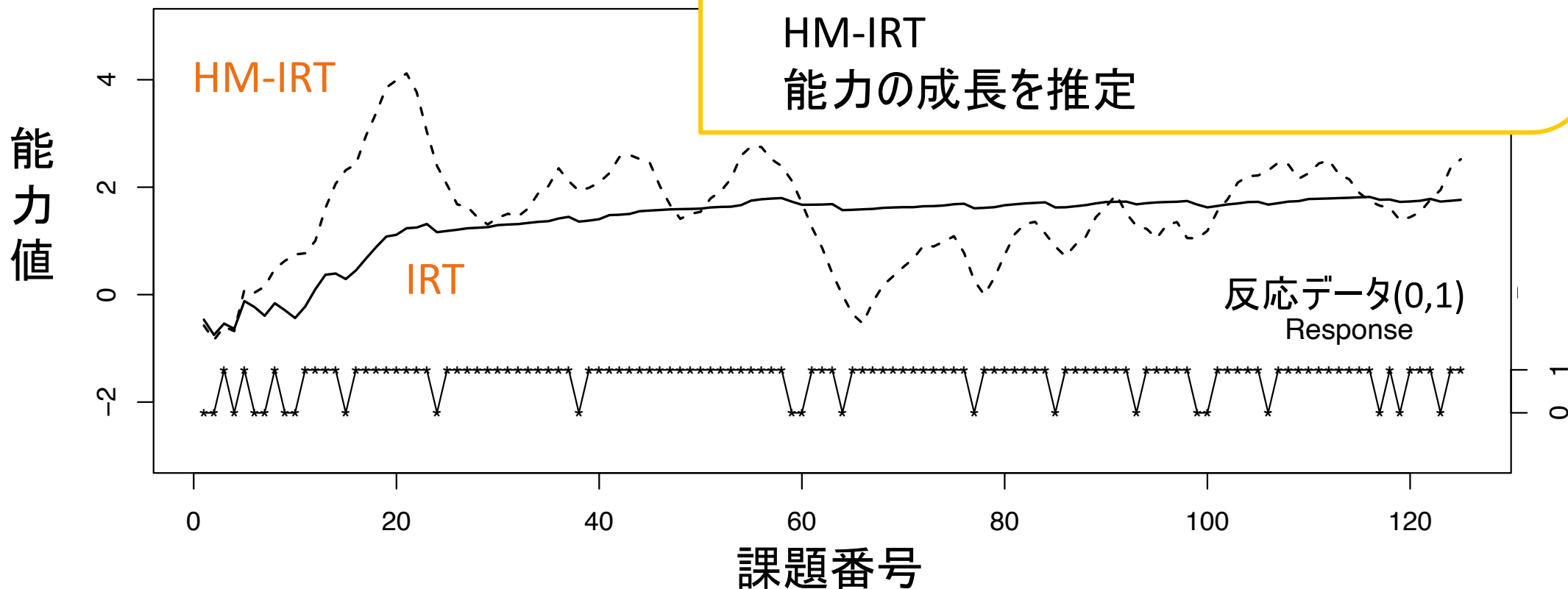
IRT

学習が進むと能力が一定値に収束



HM-IRT

能力の成長を推定



Knowledge Tracing の課題

学習者の未知の項目への反応予測

→未知の項目への正答確率を予測することで
どの問題や知識がわかっていないかを確認する

評価実験

学習者の課題への予測反応と実際の学習データ
の一致割合を予測精度としてモデルの性能比較を行う。

比較手法

IRT	Maomi Ueno, Yoshimitsu Miyazawa: IRT-Based Adaptive Hints to Scaffold Learning in Programming, IEEE Transactions on Learning Technologies, IEEE computer Society, Vol.11, Issue 4, 415-428 (2018)
Martin	A,D,Martin and K,Quinn.:Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S.Supreme Court,1953-1999, Political Analysis,Volume 10, Issue 2,pp.134-153,2002.
DIR	X,Wang, J,O,Berger and D,Burdick. : Bayesian analysis of dynamic item response models in educational testing,The Annals of Applied Statistics 2013, Vol. 7, No. 1, 126-153,2013.
TIRT	K.H.Wilson,Y.Karklin,B.Han,C.Ekanadham,“Back to the basics: Bayesian extensions of IRT outperform neural networks for proficiency estimation”, Educational Data Mining,vol.2, 2016
BKT	A.T. Corbett and J.R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge, "User Modeling and User-Adapted Interaction, vol. 4, no. 4, pp. 253-278(1995)
IBKT	Yudelson.M, Koedinger.K, G.J. Gordon,," Individualized Bayesian Knowledge Tracing Models," <i>Artificial Intelligence in Education 2013</i> , Springer, pp.171-180,(2013)
LHMM	Pelanek,R . : Conceptual Issues in Mastery Criteria: Differentiating Uncertainty and Degrees of Knowledge. <i>Artificial Intelligence in Education</i> , pp. 450-461 (2018).
DKT	C.Piech,J.Bassen,, S.Ganguli,"Deep Knowledge Tracing", Advances in Neural Information Processing Systems 28 (NIPS),2015

学習者の反応予測精度比較 Accuracy

時系列IRT

学習データ	正解率	HMIRT	IRT	Martin	DIR	TIRT	BKT	IBKT	LHMM	DKT
プログラミング基礎1	60.4%	0.747	0.696	0.656	0.688	0.758	0.647	0.641	0.642	0.702
プログラミング基礎2	65.8%	0.773	0.712	0.707	0.725	0.736	0.696	0.712	0.696	0.645
情報社会と情報倫理	81.2%	0.875	0.846	0.854	0.854	0.85	0.836	0.813	0.831	0.833
離散数学	45.3%	0.746	0.732	0.734	0.738	0.682	0.619	0.624	0.624	0.516
ASSISTments (2009-2010)	62.8%	0.753	0.701	0.694	0.65	0.707	0.686	0.688	0.661	0.687

IRT, 時系列IRTの問題点

1. 反応の局所独立性, 学習者の母集団と独立ランダムサンプリング仮定している
2. 同じ項目に繰り返す学習には適応できない
3. 複数のスキル間の関係性 (多次元の能力) を考慮した反応予測ができない

Deep-IRT [堤, 木下, 植野]

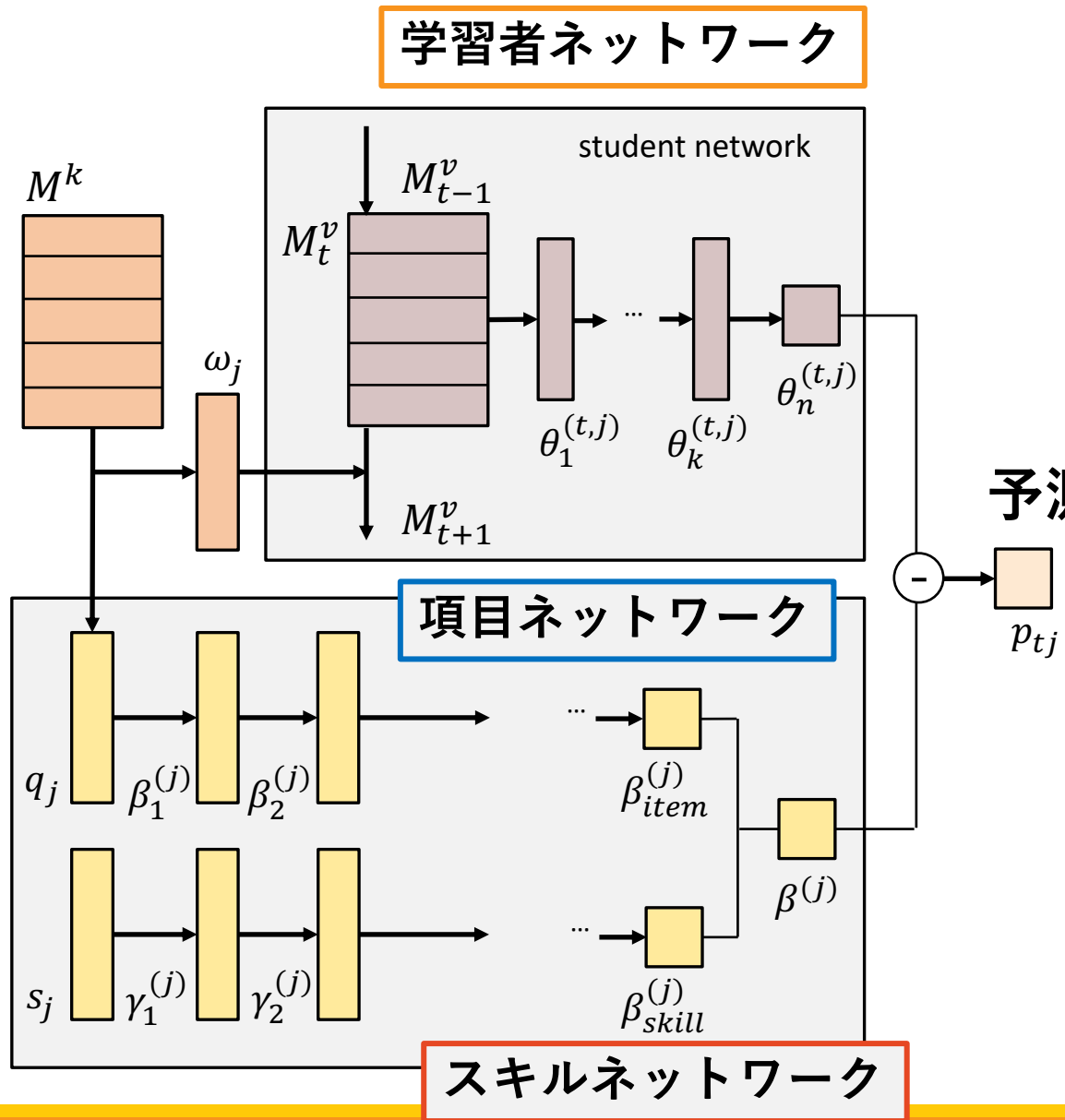
堤 瑛美子 木下 涼 植野真臣：独立な学習者・項目ネットワークをもつパラメータ解釈性向上のためのDeep-IRT, 第23回情報論的学習理論ワークショップ (IBIS2020)

ディープラーニングを用いて学習者の能力値や課題の難易度を推定し, 未知の課題への反応予測を行う.

特徴

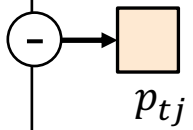
1. 反応の局所独立性, 学習者の母集団と独立ランダムサンプリングを仮定しない
2. 多次元の潜在スキルを考慮した反応予測が可能

Deep-IRT [堤, 木下, 植野]



$\theta_n^{(t,j)}$: 学習者の能力パラメータ
 $\beta^{(j)}$: 項目 j の難易度パラメータ

予測正答確率



$$p_{tj} = \text{sigmoid}(\theta_n^{(t,j)} + \beta^{(j)})$$

Deep-IRT [堤, 木下, 植野]

比較手法

- IRT Maomi Ueno, Yoshimitsu Miyazawa: IRT-Based Adaptive Hints to Scaffold Learning in Programming, IEEE Transactions on Learning Technologies, IEEE computer Society, Vol.11, Issue 4, 415-428 (2018)
- DKT C.Piech,J.Bassen,, S.Ganguli,"Deep Knowledge Tracing", Advances in Neural Information Processing Systems 28 (NIPS),2015
- DKVMN [4]J. Zhang, X. Shi, I. King, and D.-Y. Yeung, "Dynamic key-value memory network for knowledge tracing," Proceedings of the 26th International Conference on World Wide Web, pp.765–774, WWW '17,International World Wide Web Conferences SteeringCommittee, Republic and Canton of Geneva, CHE,2017.
- Yeung [5]C. Yeung, "Deep-irt: Make deep learning based knowledge tracing explainable using item response theory," Proceedings of the 12th International Conference on Educational Data Mining, EDM, 2019.
- IDRT 木下 涼, 植野 真臣：深層学習によるテスト理論：Item Deep Response Theory,電子情報通信学会論文誌 D, Vol.J103, No.4, pp.314-329(2020)

学習者の未知の項目への反応予測精度

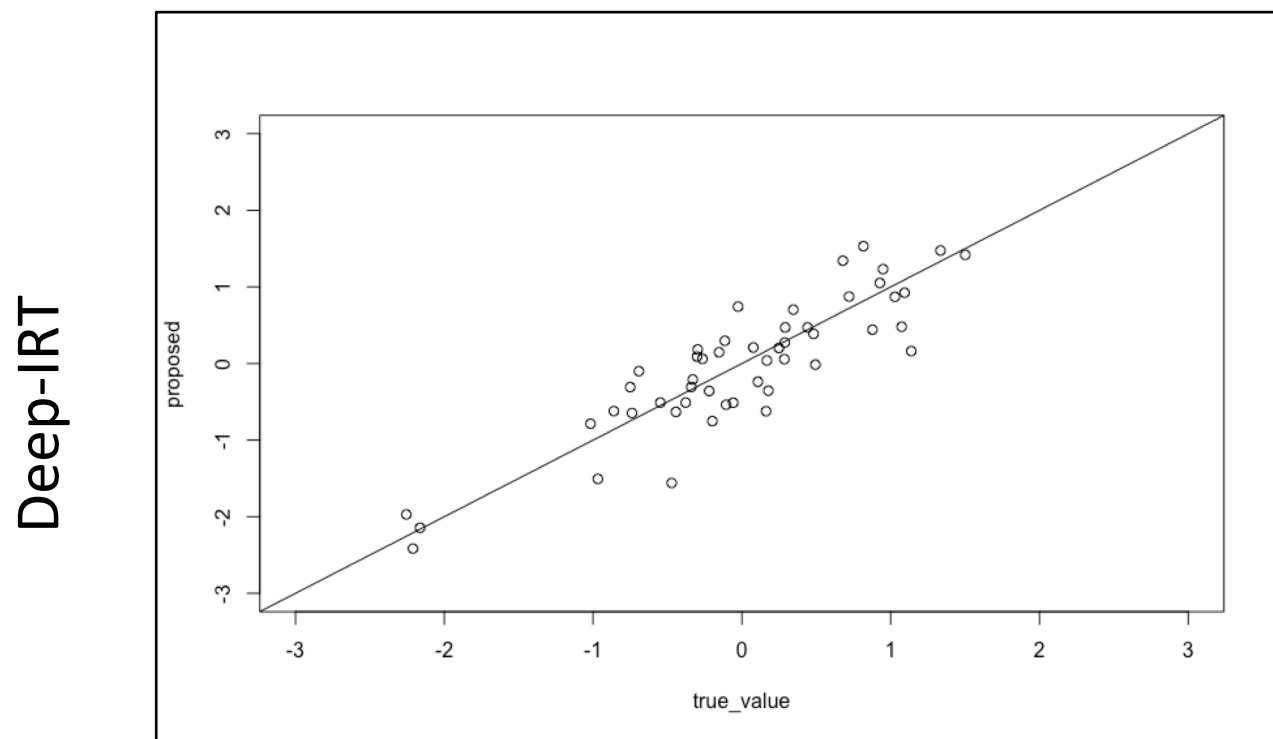
		IRT		DKT		DKVMN		Yeung		IDRT	Deep-IRT	
		item	skill	item	skill	item	skill	item	skill	item	skill	item&skill
ASSIST2009	Acc	0.72	0.765	0.759	0.637	0.763	0.683	0.768	0.71	0.711	0.768	0.765
	AUC	0.785	0.8	0.781	0.659	0.807	0.71	0.806	0.77	0.75	0.818	0.810
	F1	0.636	0.713	0.697	0.602	0.714	0.647	0.718	0.613	0.651	0.725	0.722
ASSIST2015	Acc	N/A	N/A	0.754	N/A	0.732	N/A	0.727	N/A	N/A	0.752	0.752
	AUC	N/A	N/A	0.73	N/A	0.749	N/A	0.747	N/A	N/A	0.751	0.751
	F1	N/A	N/A	0.433	N/A	0.541	N/A	0.54	N/A	N/A	0.543	0.543
Statics2011	Acc	0.816	0.769	0.777	0.805	0.78	0.817	0.787	0.81	0.819	0.789	0.822
	AUC	0.819	0.666	0.652	0.819	0.721	0.822	0.722	0.823	0.821	0.721	0.821
	F1	0.581	0.483	0.461	0.679	0.521	0.681	0.526	0.585	0.679	0.522	0.69
KDDcup	Acc	0.733	0.777	0.784	0.76	0.773	0.779	0.792	0.72	0.78	0.786	0.802
	AUC	0.614	0.549	0.538	0.565	0.594	0.561	0.588	0.61	0.57	0.588	0.601
	F1	0.522	0.439	0.439	0.464	0.439	0.447	0.455	0.501	0.455	0.469	0.478
Average_item	Acc	0.768	0.771	N/A	0.745	N/A	0.765	N/A	0.760	0.773	N/A	0.793
	AUC	0.760	0.707	N/A	0.713	N/A	0.727	N/A	0.753	0.739	N/A	0.761
	F1	0.636	0.598	N/A	0.631	N/A	0.639	N/A	0.625	0.641	N/A	0.668
Average_skill	Acc	N/A	N/A	0.769	N/A	0.762	N/A	0.769	N/A	N/A	0.774	0.785
	AUC	N/A	N/A	0.675	N/A	0.718	N/A	0.716	N/A	N/A	0.720	0.746
	F1	N/A	N/A	0.508	N/A	0.554	N/A	0.560	N/A	N/A	0.565	0.608

Deep-IRT [堤, 木下, 植野]

ディープラーニングを用いたモデルは
能力パラメータや難易度パラメータが解釈できないのではないか？

難易度パラメータ解釈性

IRTから生成した各項目の真の難易度と提案手法を用いて推定した困難度パラメータの相関



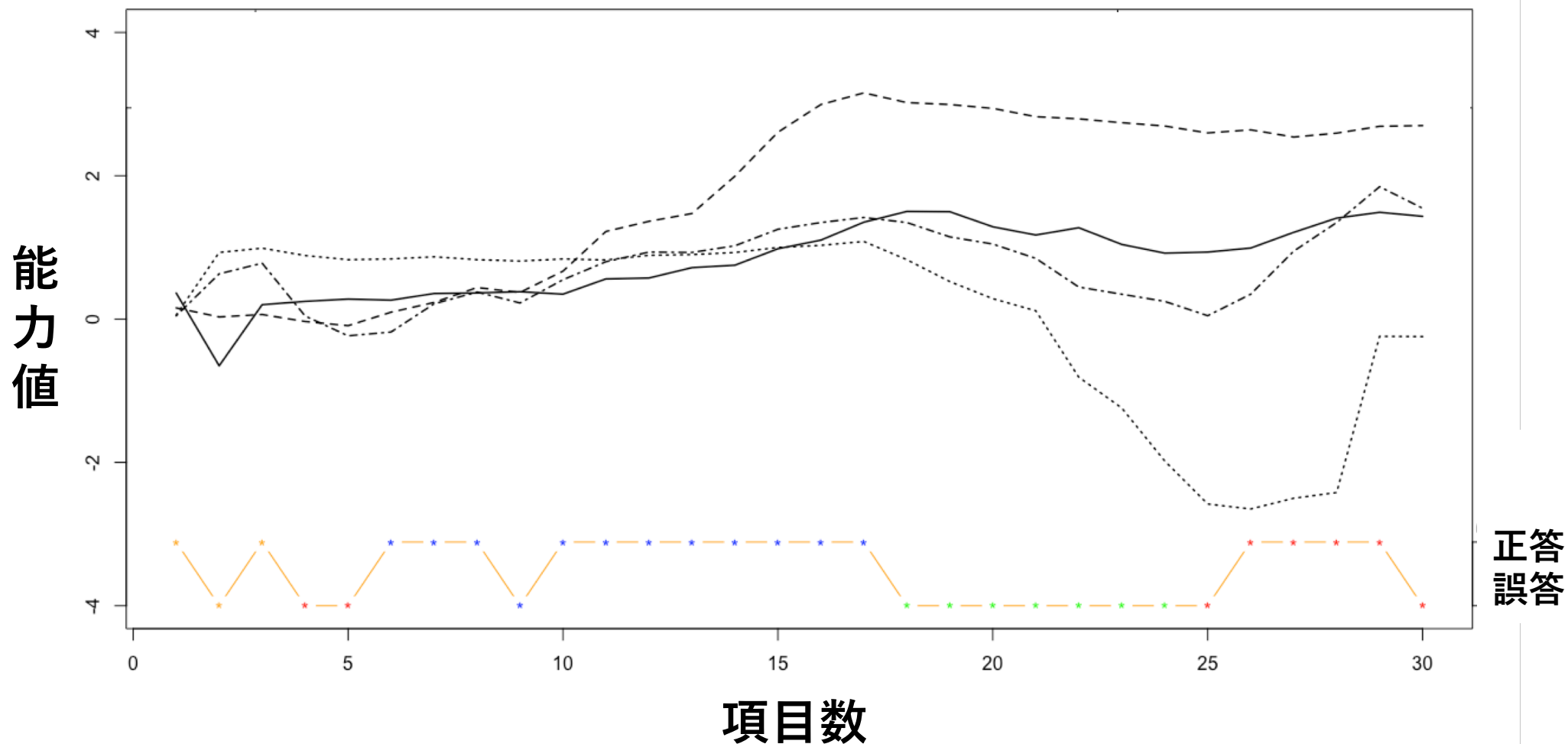
相関係数 $\rho = 0.886$

IRTと同等のパラメータ解釈が可能

真値

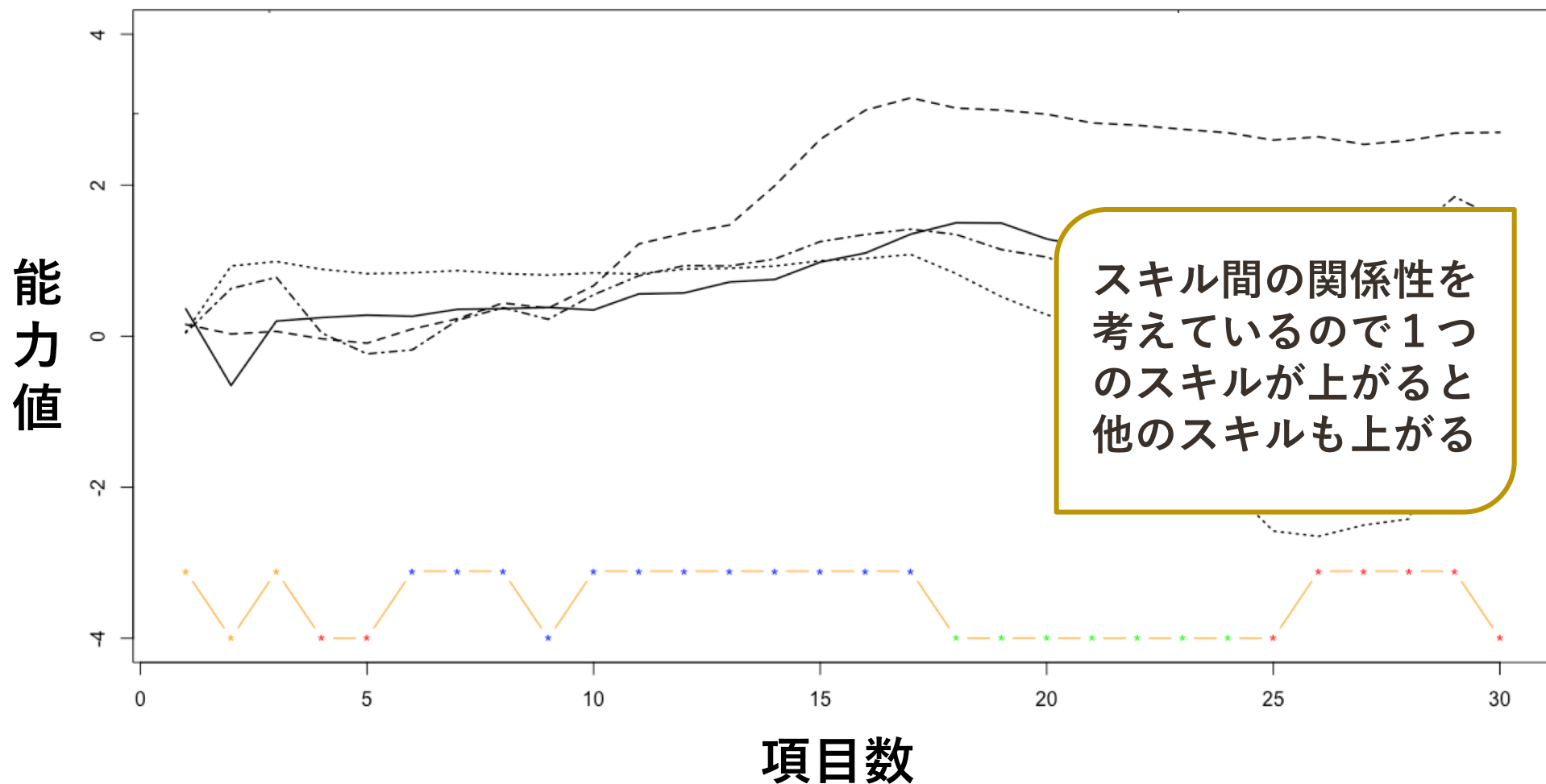
能力パラメータ解釈性

Deep-IRTで推定した学習者の能力値推移



能力パラメータ解釈性

Deep-IRTで推定した学習者の能力値推移



正答
誤答

Knowledge Tracing まとめ

学習者が分かっていない知識・スキルを探し, その知識に関する課題を見つけてくる学習支援

つまり, 基礎知識を丁寧に教える学習には効果的

学習支援における課題

教師は学習者に教えすぎても、教えなさすぎても
学習者の能力成長は望めない。

基礎的な学習では丁寧に教えて良いが
高度な発達段階では教えすぎてしまうと自分で考える力が伸びない



自力で解決できる最小限の支援をすることが重要

良い教師とは…

与えた問題や支援によって
学習者の力が伸びるかどうかが正しく予測できる

Bruner,Jの「足場かけ」

- 学習者の能力より高い難易度の課題に対して
最小限の支援をすること
- 学習者の成長につれて、徐々に支援をなくしていく「フェーディング」
が働くことが望ましい

足場かけの提示として必要な項目

1. 学習者の能力や課題の難易度を正確に判断すること
2. 判断した能力や難易度に応じて適切な支援が行われること

アダプティブラーニング 確率的足場かけシステム[Ueno and Miyazawa 2018]

- IRTを用いて過去の学習履歴データから能力値を推定し，学習効率が高くなる支援の量を予測してヒントを与えるシステム
- 学習効率を最適化する課題への正答確率があると仮定

➤ 学習者の**ヒントごとの正答確率**を正確に予測することが重要



支援は必要か？

ヒント1では予測正答確率30%...

ヒント2では**予測正答確率45%**...

ヒント3では予測正答確率65%...



システムの概要

Programming Test

Programming Code

Answer the final values of **a**, **b**, and **c** after the program runs.

```
1 public class Question1_2 {
2 public static void main(String args[]){
3     int a = 2;
4     int b = 4;
5     int c = 6;
6
7     c = a + a * b;
8     b = c / a;
9     a += b;
10    a++;
11 }
12 }
```

Hint2: c = 2 + 2 * 4; c = 2 + 8; c = 10;

Hint3: b = 6 / 2; b = 3;

Hint4: a = a + b; a = 2 + 4; a = 6;

Answer

a	<input type="text" value="1"/>	b	<input type="text" value="1"/>	c	<input type="text" value="1"/>
---	--------------------------------	---	--------------------------------	---	--------------------------------

SUBMIT

Hint1

Variable

= Assignment Statement

Operator

++ Increment

+ = Addition assignment operator

++ Increment

++ : Increment

a ++; \Rightarrow a = a + 1;

Code

```
int a = 5;
a++;
```

```
a = a + 1;
a = 5 + 1;
a = 6;
```

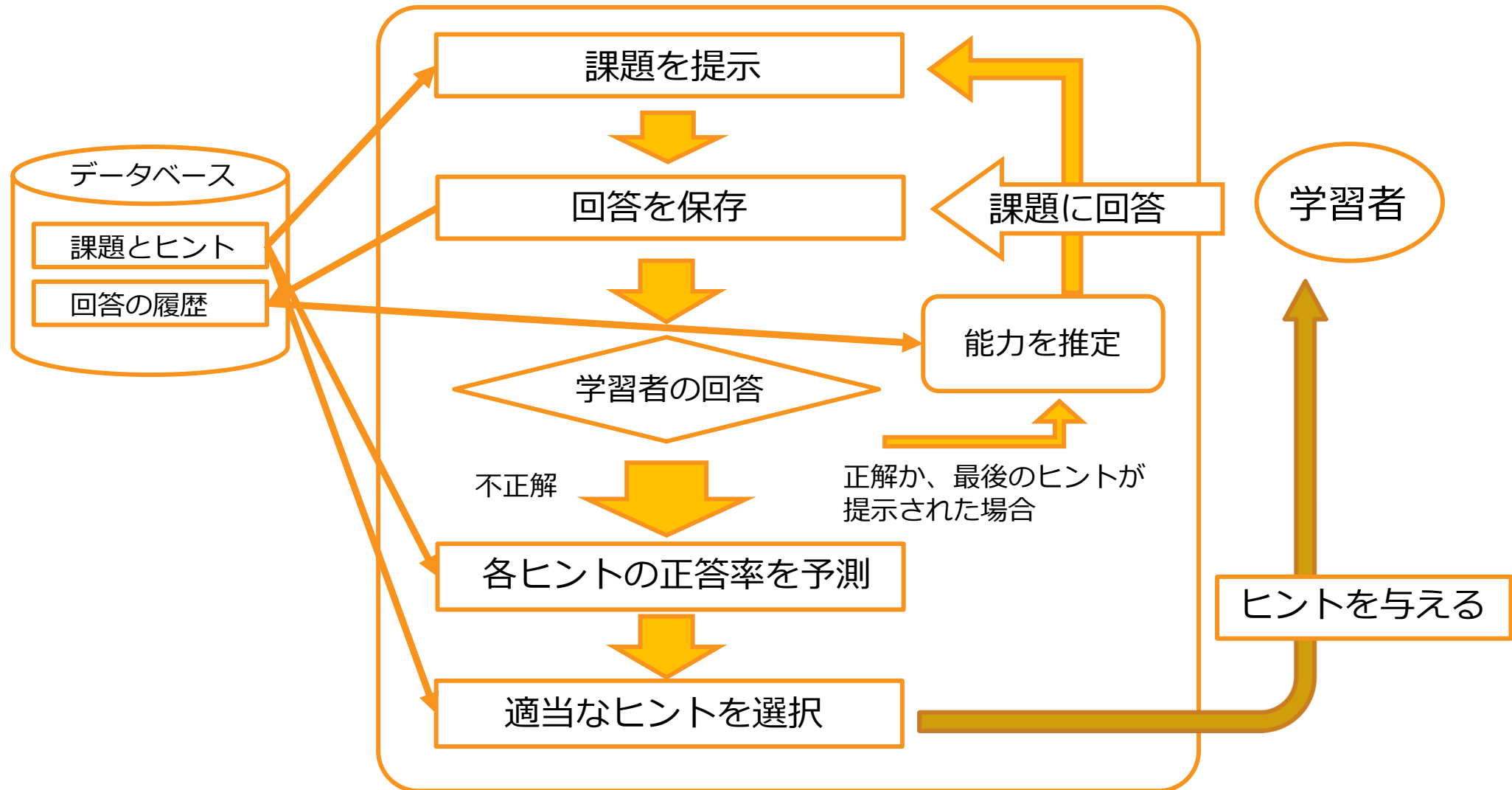
top

段階反応モデル

学習者 i が課題 j でヒント k を与えられたときに正答する確率

$$P(u_j = k | \theta_i) = \frac{1}{1 + \exp(-a_j \theta_i + b_{j(k-1)})} - \frac{1}{1 + \exp(-a_j \theta_i + b_{jk})}$$

システムのアルゴリズム



ヒント提示手法の違いによる学習効果比較

- プログラミングに関する知識は入門レベルの被験者93名
- 被験者が事前テストを受けた後、本システムで学習させる
- ヒント提示手法の違いによる事前事後テストの学習効果の比較を行った

被験者のグループ

A群 正答率 p が0.8 に近くなるようにヒントを提示

B群 正答率 p が0.65 に近くなるようにヒントを提示

C群 正答率 p が0.5 に近くなるようにヒントを提示

D群 ヒントを提示しない

E群 ヒントを難しいものから順次提示

F群 解答と解説を最初から丁寧に提示

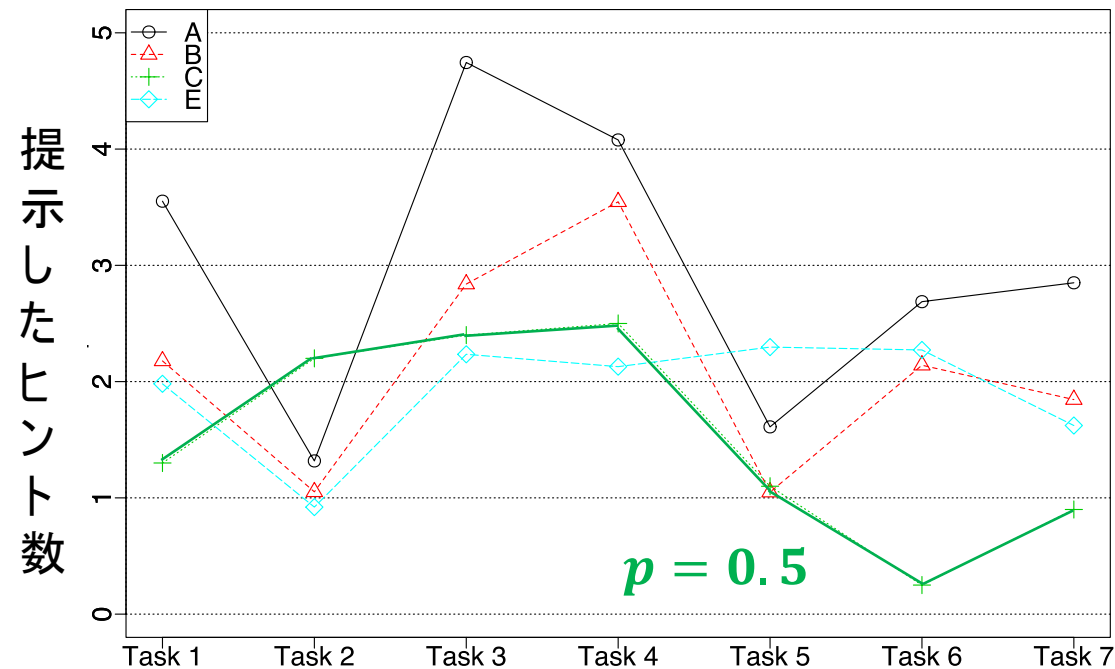
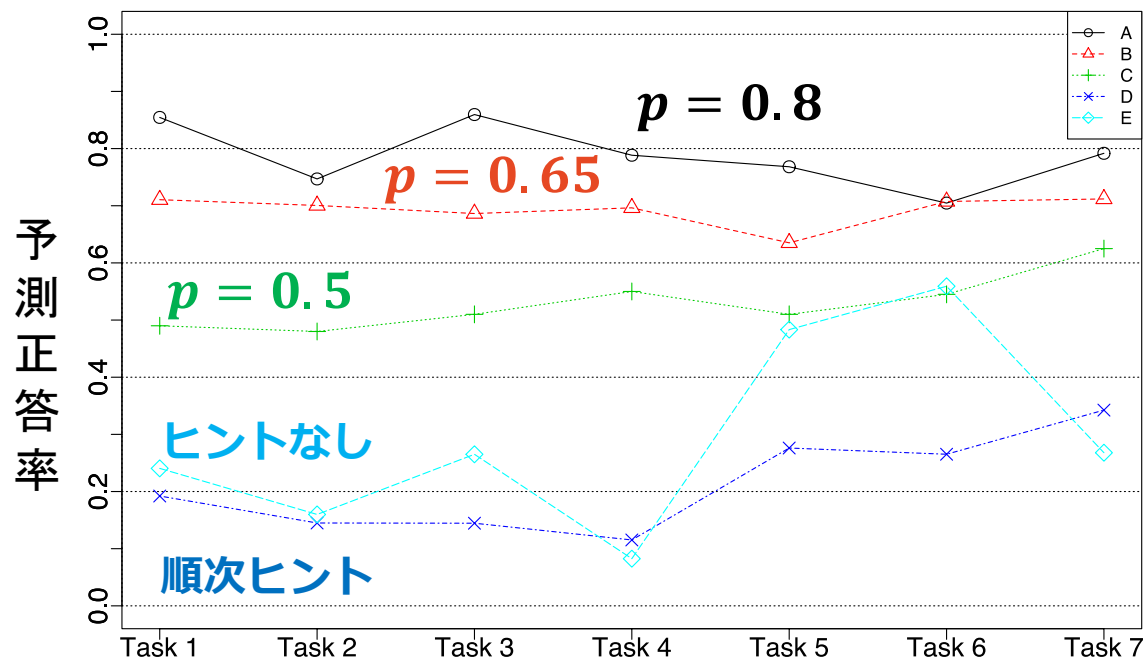
実験結果

	$p=0.8$	$p=0.65$	$p=0.5$	ヒントなし	1つずつ提示	全てのヒント
被験者数	14	16	18	15	12	18
事前テスト	1.14 (1.59)	1.69 (2.44)	1.78 (2.44)	1.33 (1.89)	2.17 (1.40)	2.72 (2.23)
事後テスト	35.4** (2.94)	34.8** (2.13)	<u>40.0</u> (3.15)	36.5* (2.22)	34.8** (2.44)	30.9** (4.92)
記憶保持テスト	20.8 (2.73)	20.8 (2.27)	23.0 (2.18)	20.6 (1.81)	20.6 (1.81)	20.8 (5.41)



予測正答率0.5でヒントを提示したグループが
事後テストの点数が一番高い

ヒント提示後の正答率と与えたヒント数の平均値の推移

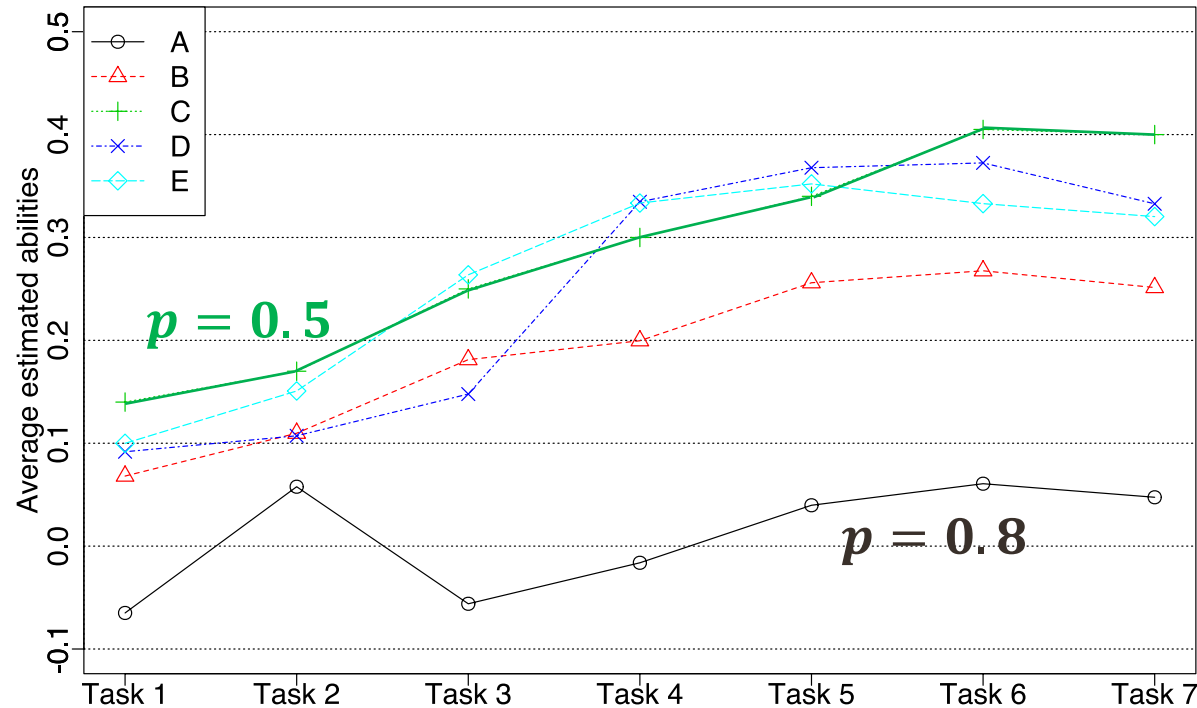


- 予測正答率が0.8, 0.65, 0.5でヒントを提示したグループは予測された正答率に近い値で遷移している

- 課題4以降は予測正答率0.5でヒントを提示したグループのヒント数が全体的に減少している

➡ フェーディングが働いている

推定能力の平均値の推移



- 学習支援後の課題への予測正答確率が50%に近づくように**ヒントの量を調節**することが、学習者の成長に最も効果的である。
- $p = 0.8$ グループの推定能力値が一番低いことから、学習者への過剰な支援は学習者の能力向上に適切でないことが分かる

確率的足場がけシステム[Ueno and Miyazawa 2018]の問題

- 学習者の時系列の能力変化を加味されていない

隠れマルコフIRT(HMIRT) [堤ら,2019]を用いたヒント予測

堤瑛美子, 木下涼, 植野 真臣: ダイナミックアセスメントのための隠れマルコフIRTモデル, 電子情報通信学会論文誌 D, Vol.J102-D, No.2, pp.79-92 (2019)

- 学習者の時系列の能力変化を加味したヒント付きの課題への反応をモデル化したIRT

学習者 i が課題 j でヒント k を与えられたときに正答する確率

$$P(u_j = k | \theta_i) = \frac{1}{1 + \exp(-a_j \theta_{it} + b_{j(k-1)})} - \frac{1}{1 + \exp(-a_j \theta_{it} + b_{jk})}$$

$$\theta_{jt} \sim N(\theta_{jt-1}, \delta)$$

隠れマルコフIRT(HMIRT) [堤ら,2019]を用いたヒント予測

堤瑛美子, 木下涼, 植野 真臣: ダイナミックアセスメントのための隠れマルコフIRTモデル, 電子情報通信学会論文誌 D, Vol.J102-D, No.2, pp.79-92 (2019)

- 学習者が課題に正答するまでに利用するヒント数を予測

	課題 2	課題 3	課題 4	課題 5	課題 6	課題 7	課題 8	課題 9	課題 10
HM-IRT	52.7%	63.9%	46.9%	50.9%	68.8%	55.7%	62.7%	48.3%	57.2%
IRT	40.5%	59.9%	45.5%	44.9%	66.1%	52.4%	52.5%	46.9%	54.7%

	課題 11	課題 12	課題 13	課題 14	課題 15	課題 16	課題 17	課題 18	平均
HM-IRT	49.3%	79.7%	83.5%	87.6%	50.5%	96.0%	91.6%	69.2%	65.6%
IRT	40.3%	67.5%	82.9%	87.6%	49.7%	84.9%	90.7%	69.1%	61.0%

IRT, 時系列IRTの問題点

1. 反応の局所独立性, 学習者の母集団と独立ランダムサンプリング仮定している
2. 同じ項目に繰り返す学習には適応できない
3. 複数のスキル間の関係性 (多次元的能力) を考慮した反応予測ができない
4. ヒントの順序が固定されている

Deep-IRT [堤, 木下, 植野]を用いたヒント予測

- ディープラーニングを用いて学習者の能力値や課題・ヒントの難易度を推定し, 未知の課題への反応を予測する
- 確率モデルでの局所独立性の問題を解決
- 多次元の潜在スキルを考慮した反応予測が可能

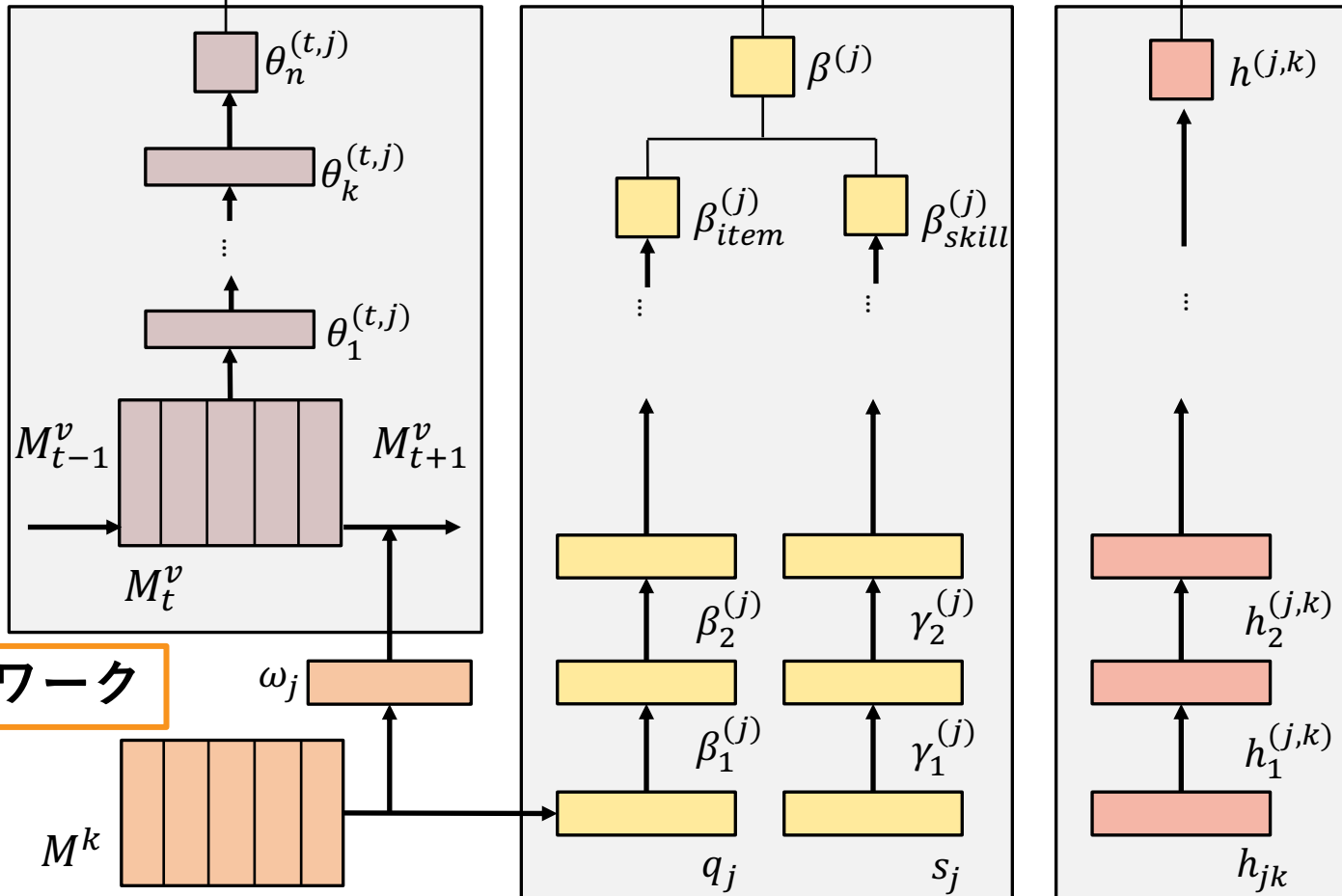
- ヒントの提示順序が制限されない

p_{tj}

予測正答確率

$p_{tj} =$

$$\text{sigmoid}(\theta_n^{(t,j)} + \beta^{(j)} - h^{(j,k)})$$



学習者ネットワーク

項目・スキルネットワーク

ヒントネットワーク

$\theta_n^{(t,j)}$: 学習者の能力パラメータ
 $\beta^{(j)}$: 項目 j の難易度パラメータ
 $h^{(j,k)}$: ヒント k の難易度パラメータ

- 学習者が課題に正答するまでに利用するヒント数を予測

	課題1	課題2	課題3	課題4	課題5	課題6	課題7	課題8
Deep-IRT	37.5%	75.0%	25.0%	6.2%	37.5%	37.5%	31.2%	31.2%
IRTモデル	37.5%	56.2%	43.7%	18.7%	25.0%	37.5%	18.7%	31.2%
	課題9	課題10	課題11	課題12	課題13	課題14	課題15	課題16
Deep-IRT	56.2%	50.0%	25.0%	75.0%	75.0%	25.0%	37.5%	50.0%
IRT	56.2%	50.0%	31.2%	18.7%	68.7%	18.7%	25.0%	31.2%
	課題17	課題18	課題19	課題20	課題21	課題22	平均	
Deep-IRT	68.7%	62.5%	43.7%	56.2%	56.2%	12.5%	43.1%	
IRT	37.5%	25.0%	56.2%	31.2%	56.2%	43.7%	37.2%	

アダプティブヒント予測のまとめ

高度な発達段階では教えすぎてしまうと自分で考える力が伸びない

自力で解決できる最小限の支援をすることが重要



学習者の能力や課題の難易度、**ヒントごとの正答確率**を正確に予測する手法が必要