NEURAL INFORMATION PROCESSING SYSTEMS

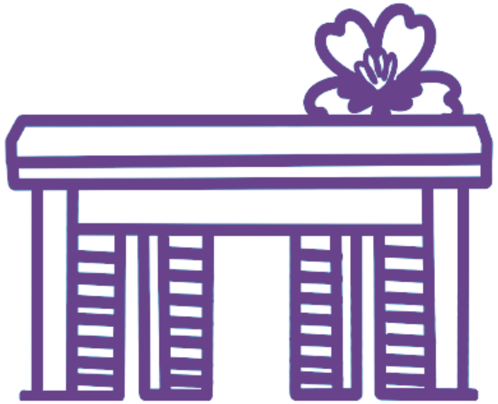
Collaborative Cognitive Diagnosis with Disentangled Representation Learning for Learner Modeling

Weibo Gao, Qi Liu, Linan Yue, Fangzhou Yao, Hao Wang,
Yin Gu, Zheng Zhang

认知智能全国重点实验室
STATE KEY LABORATORY OF COGNITIVE INTELLIGENCE

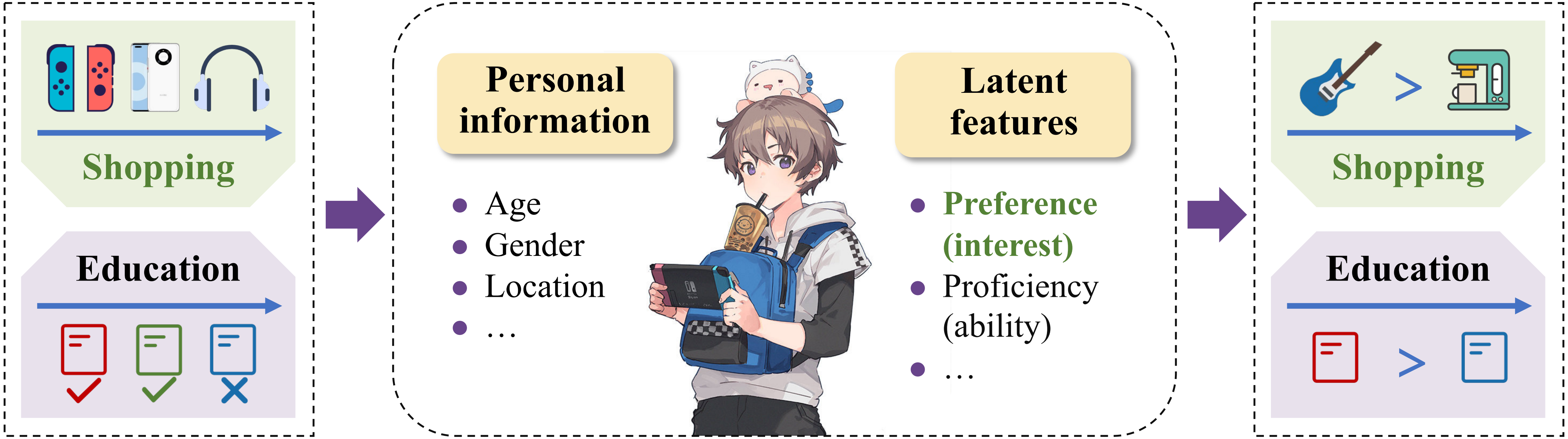
Reporter: Weibo Gao

Outline



Background

- In the applications of AI, it needs to characterize the difference of individuals in both personal information and latent features



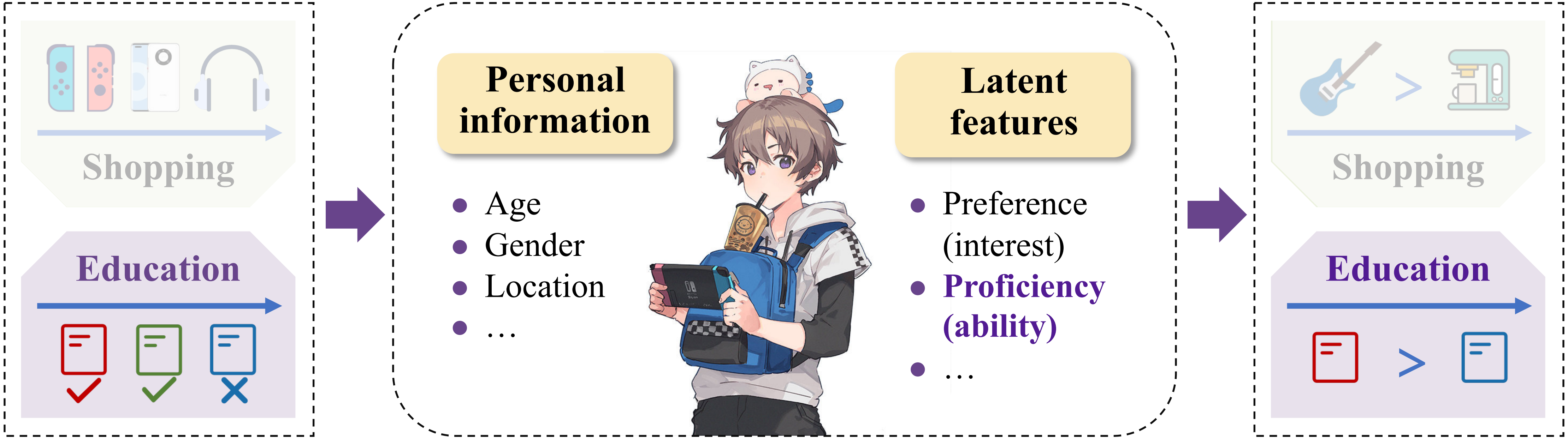
Observed behaviors
- Shopping
- Exercising
- …

AI applications
- **E-commerce**
- Education
- …

Background

NeurIPS 2024

- In the applications of AI, it needs to characterize the difference of individuals in both personal information and latent features



Observed behaviors
- Shopping
- Exercising
- …

AI applications
- E-commerce
- **Education**
- …

Background

- In the applications of AI, it needs to characterize the difference of individuals in both personal information and latent features

Cognitive Diagnosis

Observed behaviors
- Shopping
- Exercising
- …

AI applications
- E-commerce
- **Education**
- …

NeurIPS 2024

Cognitive Diagnosis (CD)

- **Goal**: Diagnosing the **cognitive states** of each learner (i.e., **proficiency** on specific knowledge concepts) by fitting their practice records (i.e., practice correctness)
- **Components**: Learners, questions and knowledge concepts
- **Applications**: Supporting personalized tutoring services

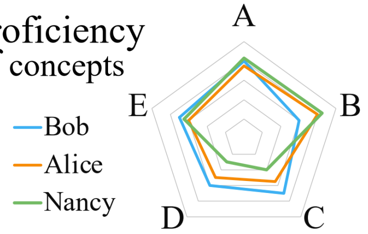
Questions	Knowledge concepts	Bob	Alice	Nancy
q_1	A	✓	✓	✓
q_2	B	✗	✓	✓
q_3	A, B	✓	✓	?
q_4	C, D	✓	✗	✗
q_5	C, E	✓	✓	?

Practice records

Cognitive Diagnosis

Legend
A: Function
B: Derivative
C: Number
D: Cube
E: Cone
✓ / ✗: Correct / wrong logs

Proficiency of concepts



Cognitive states

Applying

- Question/course recommendation
- Adaptive testing
- …

Tutoring services

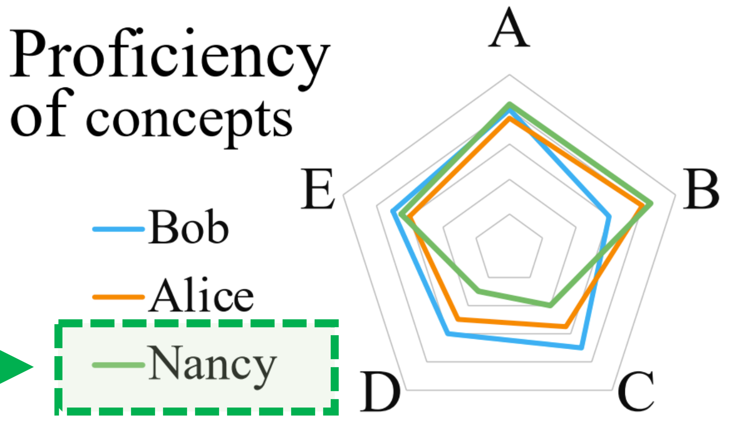
Current Limitation of CD

- Current **CD** models mainly focus **inner-learner** modeling, but ignore **inter-learner** information
  - **Inner-learner: Individual** attributions and practice records
  - **Inter-learner: Collaborative** clues between learners with similar cognitive states

Questions	Knowledge concepts	Bob	Alice	Nancy
q_1	A	✓	✓	✓
q_2	B	✗	✓	✓
q_3	A, B	✓	✓	?
q_4	C, D	✓	✗	✗
q_5	C, E	✓	✓	?

Practice records

Individual practice records



Cognitive states

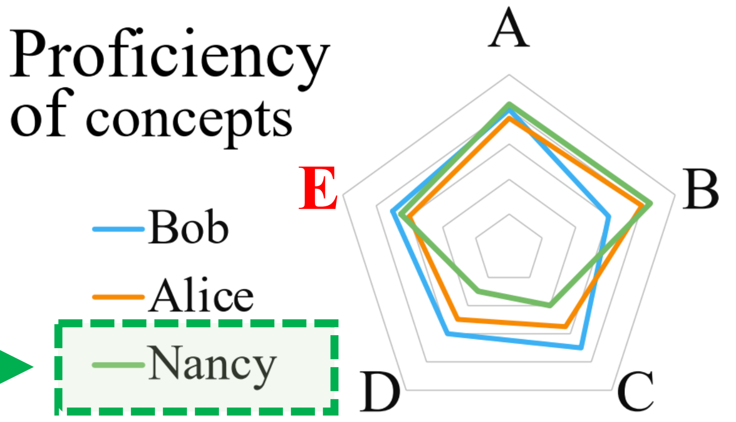
Current Limitation of CD

- Current **CD** models mainly focus **inner-learner** modeling, but ignore **inter-learner** information
  - **Inner-learner: Individual** attributions and practice records
  - **Inter-learner: Collaborative** clues between learners with similar cognitive states

Questions	Knowledge concepts	Bob	Alice	Nancy
q_1	A	✓	✓	✓
q_2	B	✗	✓	✓
q_3	A, B	✓	✓	?
q_4	C, D	✓	✗	✗
q_5	C, E	✓	✓	?

Practice records

Individual practice records



Cognitive states

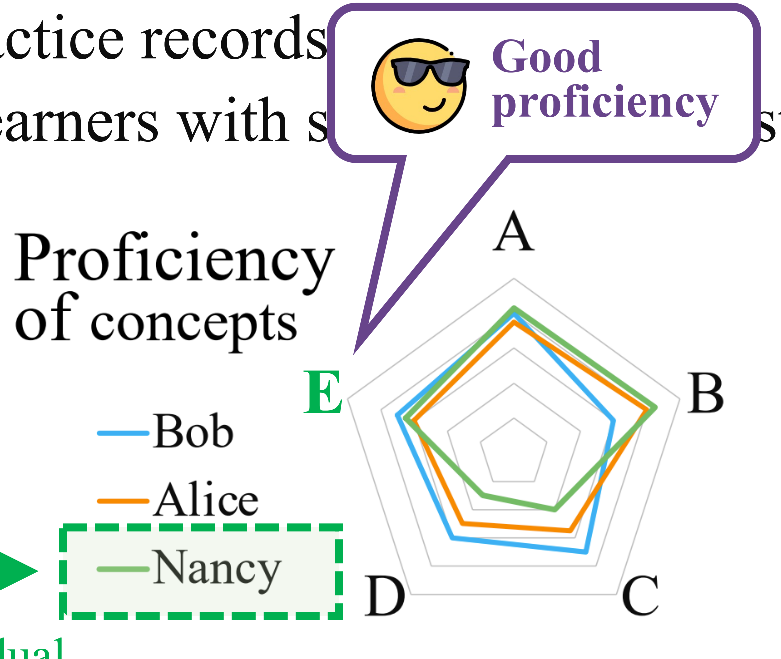
Current Limitation of CD

- Current **CD** models mainly focus **inner-learner** modeling, but ignore **inter-learner** information
  - **Inner-learner: Individual** attributions and practice records
  - **Inter-learner: Collaborative** clues between learners with s states

Questions	Knowledge concepts	Bob	Alice	Nancy
q_1	A	✓	✓	✓
q_2	B	✗	✓	✓
q_3	A, B	✓	✓	?
q_4	C, D	✓	✗	✗
q_5	C, E	✓	✓	?

Collaborative clues

Individual practice records



Cognitive states

Motivation

NeurIPS 2024

Our research goal

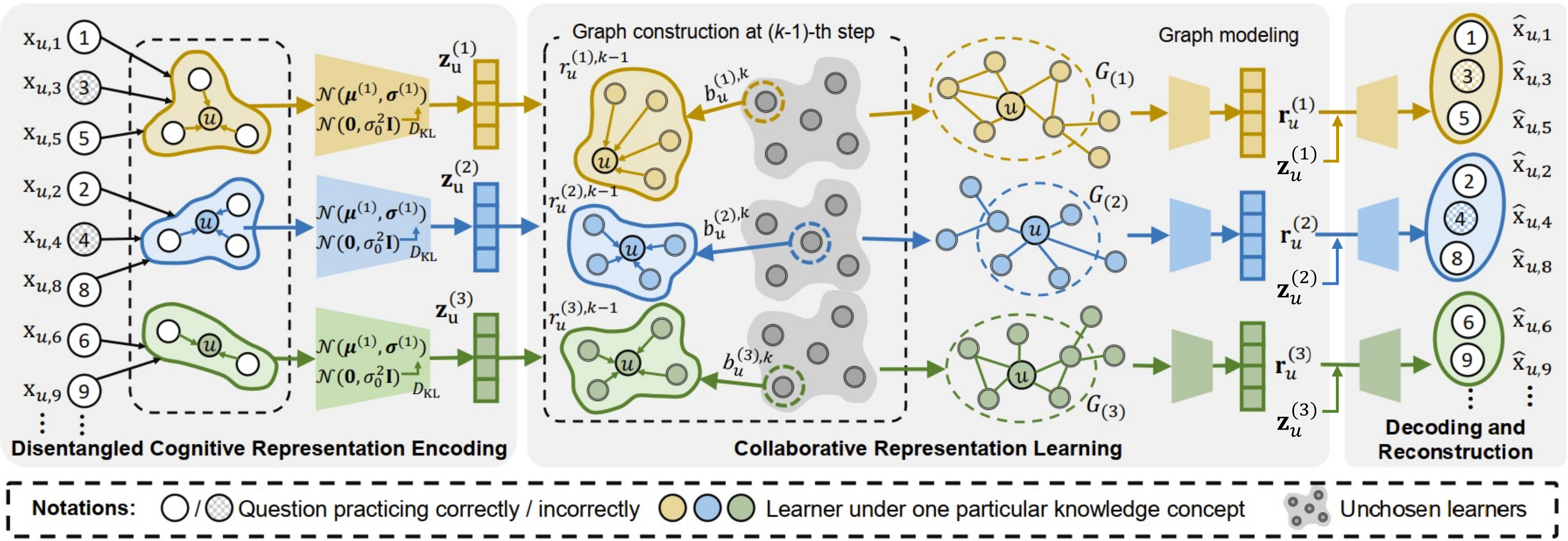
An ideal cognitive diagnosis model should consider both <u>inner-</u> and <u>inter-</u> learner information

Outline

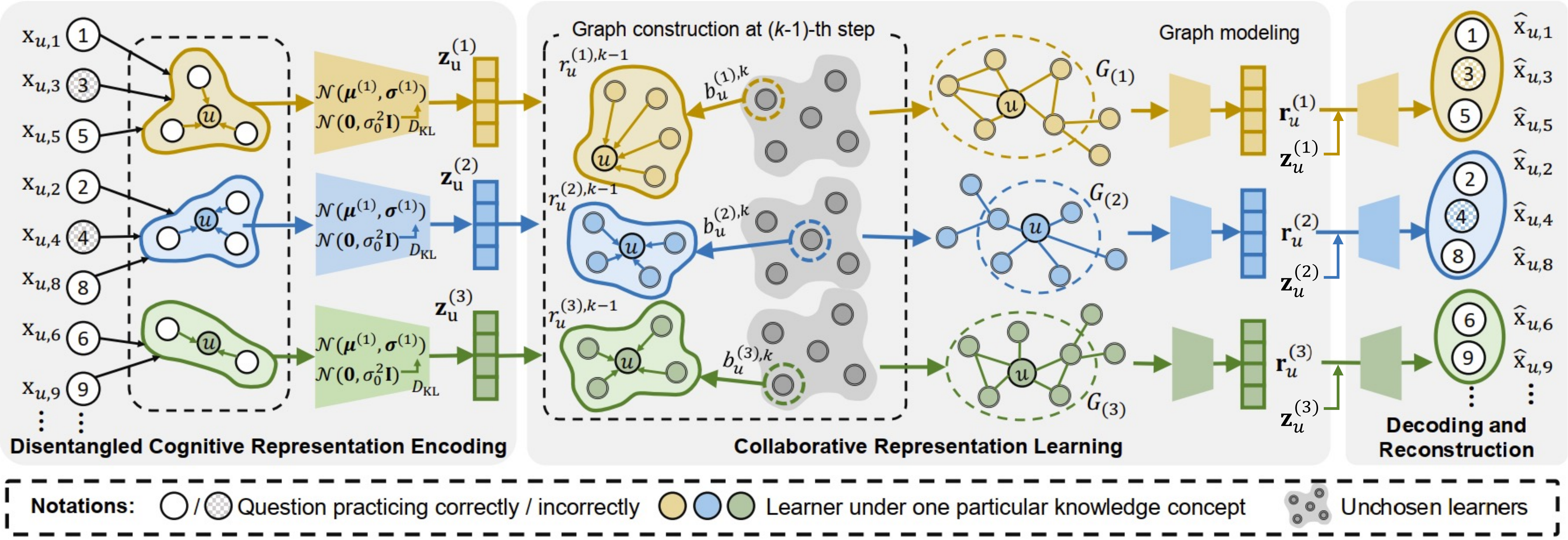
Coral Model for Cognitive Diagnosis

NeurIPS 2024

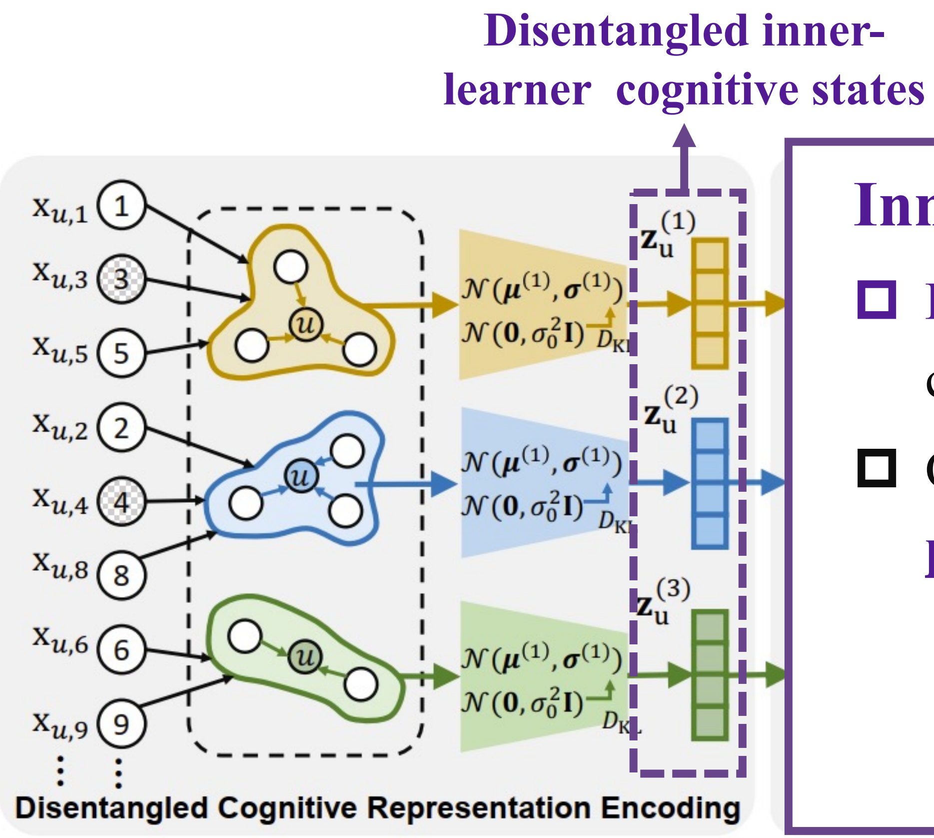
We propose **Coral**, a **Co**llabo**ra**tive cognitive diagnosis model with disentang**l**ed representation Learning for both **inner-** and **inter-** learner information Modeling



Coral Model for Cognitive Diagnosis



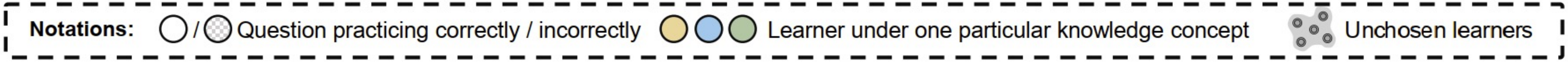
Coral Model for Cognitive Diagnosis



Inner-learner modeling:

- **Disentangle** the learner's cognitive state into multiple components
- Optimize disentangled cognitive states by **fitting their practice performance**

$$p_\Theta(\mathbf{x}_u) = \mathbb{E}_{p(\mathbf{C})}\left[\int p_\Theta(\mathbf{x}_u \mid \mathbf{z}_u, \mathbf{C})\, p_\Theta(\mathbf{z}_u)\, d\mathbf{z}_u\right]$$

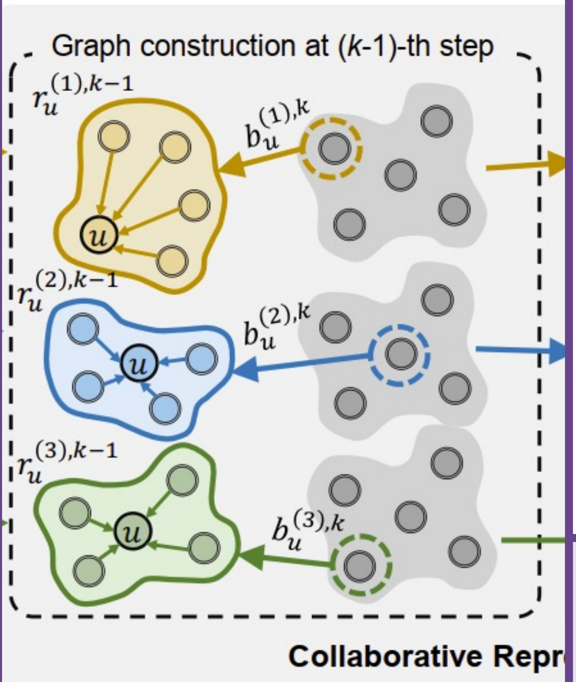


Coral Model for Cognitive Diagnosis

NeurIPS 2024

Inter-learner modeling (1/2):

- Find K collaborative neighbors with similar cognitive states for each learner
- **Theoretically** derive the optimal condition ($\max \log p_{\Theta}(G \mid V, \mathbf{Z})$) for **building the collaborative graph of learners**



Notations: ◯ / ◯ Question practicing correctly / incorrectly ◯◯◯ Learner under o

Property 3.

$\max \log p_{\Theta}(G \mid V, \mathbf{Z})$ *is bounded as follows:*

$$\max \log p_{\Theta}(G \mid V, \mathbf{Z}) \geq -\sum_{c=1}^{C}\sum_{u=1}^{M}\sum_{k=1}^{K}\mathcal{L}_u^{(c),k}$$

$$\boxed{\mathcal{L}_u^{(c),k} = -\frac{\exp\left(f_{(c)}\left(b_u^{(c),k}; r_u^{(c),k-1}\right)\right)}{\sum_{v \in V_u^{(c)}} \exp\left(f_{(c)}\left(v; r_u^{(c),k-1}\right)\right)}}$$

From steps 1 to K, **iteratively** search for the ***k*-th** collaborative neighbor with **the most similar cognitive state** to the current **context** (existing **(*k*-1)-th** neighbors).

Coral Model for Cognitive Diagnosis

NeurIPS 2024

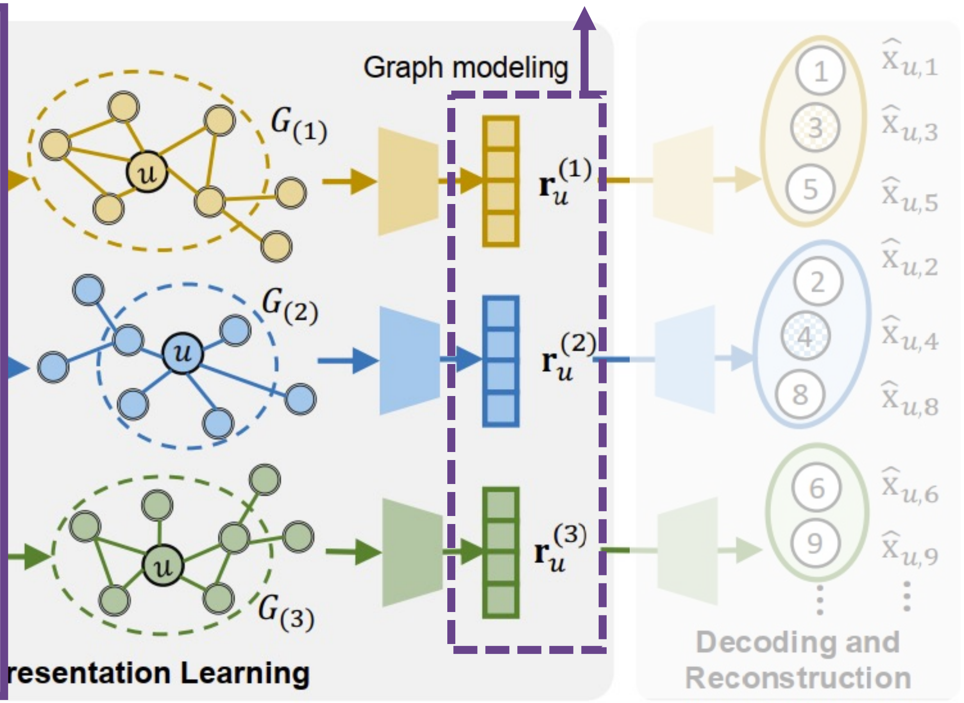
Inter-learner modeling (2/2):

- Based on learner collaborative graphs at each disentangled component, we design a **context-aware GCN** to **learn collaborative learner cognitive states**

$$\mathbf{r}_u^{(c)} = \frac{1}{|\mathcal{N}_u^{(c)}|}\sum_{v\in\mathcal{N}_u^{(c)}} s_{u,v}^{(c)} \cdot \hat{\mathbf{z}}_v^{(c)}$$

$$s_{u,v}^{(c)} = \frac{\hat{\mathbf{z}}_u^{(c)\,T}\cdot \hat{\mathbf{z}}_v^{(c)}}{\sum_{j\in\mathcal{N}_u^{(c)}} \hat{\mathbf{z}}_u^{(c)\,T}\cdot\hat{\mathbf{z}}_j^{(c)}} + \frac{f_{(c),v}}{\sum_{k=1}^{K} f_{(c),v_k}}$$

Disentangled inter-learner cognitive states



Notations: ○ / ◌ Question practicing correctly / incorrectly ○○○ Learner under one particular knowledge concept ⁘ Unchosen learners

Coral Model for Cognitive Diagnosis

NeurIPS 2024

Inner- and inter- view alignment

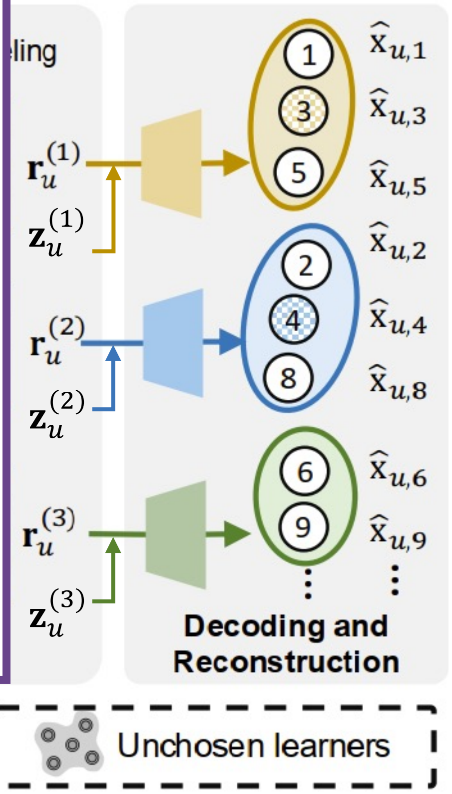
- Achieve the co-disentanglement of inner- and inter- learner cognitive states

$$\tilde{\mathbf{z}}_u = \mathbf{z}_u + \mathbf{r}_u$$

Optimization

$$\arg\min \mathcal{L} = \sum_{u=1}^{M} \Big[\sum_{x_{u,i} \in \mathbf{x}_u} \alpha \cdot BCE\left(x_{u,i}, p_\Theta\left(x_{u,i} \mid \mathbf{z}_u, \mathbf{C}\right)\right) - \beta \cdot D_{\mathrm{KL}}^{u} + \sum_{x_{u,i} \in \mathbf{x}_u} BCE\left(x_{u,i}, p_\Theta\left(\hat{x}_{u,i}\right)\right) \Big]$$

$$\text{s.t.} \quad \arg\max \sum^{C} \sum^{K} \mathcal{L}_u^{(c),k}$$



Outline

Experiments

NeurIPS 2024

Dataset

- Each dataset contains learner practice correctness on specific questions

Datasets	ASSIST	Junyi	NeurIPS2020EC
#students	1,256	1,400	1,000
#questions	16,818	674	919
#knowledge concepts	120	40	30
#concepts per exercise	1.21	1	4.02
#records	199,790	70,797	331,187
#records per student	159,07	50.67	331.19
#correct records / #incorrect records	67.08%	77.20%	53.87%

Table 1: The statistics of three datasets.

Experiments

NeurIPS 2024

Evaluation of learner practice performance prediction

Dataset	Method	Metric			
		ACC ↑	AUC ↑	F1-score ↑	RMSE ↓
ASSIST	IRT	69.36	69.81	78.14	45.61
	MIRT	71.26	72.59	79.80	44.50
	PMF	71.34	72.27	80.68	48.67
	NCDM	72.27	74.27	79.97	48.67
	KaNCD	**72.43**	**75.38**	80.22	48.67
	RCD	72.04	73.14	80.60	43.74
	DCD	70.33	73.98	79.09	43.94
	Coral	71.53	74.72	**81.16**	**43.66**
Junyi	IRT	79.26	76.46	87.54	38.38
	MIRT	77.74	74.46	86.05	40.29
	PMF	79.65	77.17	88.18	44.10
	NCDM	79.91	78.91	87.73	38.35
	KaNCD	**81.79**	80.93	89.02	36.11
	RCD	81.02	80.22	88.00	37.23
	DCD	79.29	79.55	87.62	37.83
	Coral	81.15	**80.94**	**89.12**	**36.08**

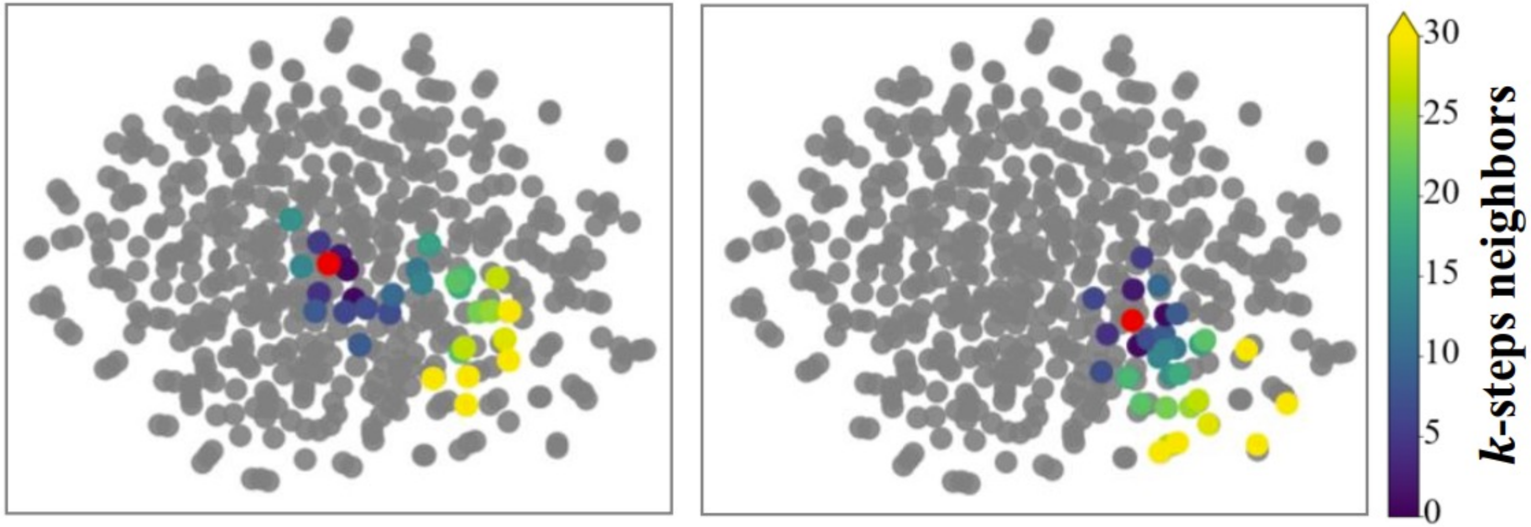
Dataset	Method	Metric			
		ACC ↑	AUC ↑	F1-score ↑	RMSE ↓
NeurIPS2020EC	IRT	70.11	75.60	71.59	44.68
	MIRT	69.95	75.52	71.24	45.51
	PMF	69.85	75.39	72.62	48.33
	NCDM	71.66	78.57	71.36	43.21
	KaNCD	71.28	77.60	72.50	43.71
	RCD	70.43	77.25	72.64	44.01
	DCD	71.53	75.63	71.13	45.60
	Coral	**71.72**	**78.88**	**72.82**	**43.20**

The Coral model nearly outperforms all baseline models across three datasets

Experiments

◻ Visualization of iteratively searching collaborative neighbors

- ◻ Randomly select two learners as examples, and **visualize their collaborative neighbors searching processes**



Coral organizes neighbors according to cognitive states and exemplifying a compelling strategy for neighbor selection that takes into account cognitive similarity

Experiments

NeurIPS 2024

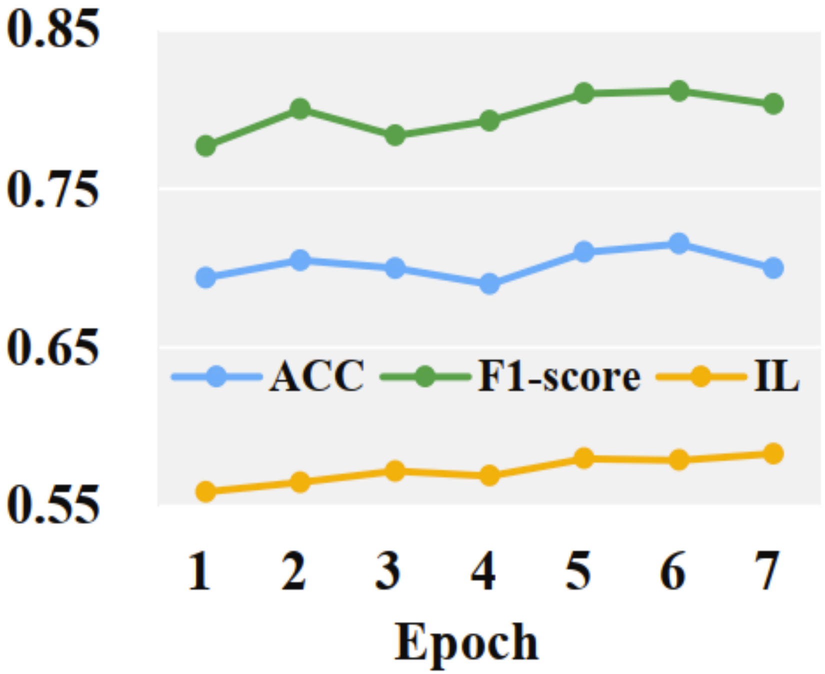
Evaluation of the disentanglement effectiveness

- **Metric**:

$$IL = \sum_{u=1}^{M} IL(u)$$

$$IL(u) = \frac{1}{C}\sum_{c=1}^{C} \frac{2}{d(d-1)} \sum_{1\le i,j\le d} \left|z_u^{(c)}[i] - z_u^{(c)}[j]\right|$$

- The higher the *IL*, the higher the degree of disentanglement



- **Coral gradually achieves a high degree of disentanglement during the training process**
- **Model performances generally exhibit a positive correlation with the degree of disentanglement**

Outline

NeurIPS
2024

Conclusion

NeurIPS 2024

- Introduce **collaborative modeling** into **cognitive diagnosis** (CD)
  - An ideal cognitive diagnosis model should consider both **inner-** and **inter-** learner information
- Propose the **Coral** model
  - **Inner-view: Disentangle** learners' cognitive states into multiple components
  - **Inter-view:** (1) **Iteratively** construct learner collaborative graphs at each disentangled components; (2) Design the **context-aware GCN** to model collaborative clues
  - **Co-disentanglement: Fuse** both **inner-** and **inter-** learner cognitive states
- Experiments
  - Proving Coral's effectiveness of **collaborative modeling** and **disentanglement**
  - Project homepage: https://github.com/bigdata-ustc/Coral

Thank you

Q & A

weibogao@mail.ustc.edu.cn