

JiuZhang: A Chinese Pre-trained Language Model for Mathematical Problem Understanding

Wayne Xin Zhao, Kun Zhou, Zheng Gong, Beichen Zhang, Yuanhang Zhou, Jing Sha, Zhigang Chen, Shijin Wang, Cong Liu, Ji-Rong Wen

Renmin University of China, iFLYTEK Research

❖ Contributions:

- (1) We pre-train the first Chinese pre-trained language model specially for mathematical problem understanding;
- (2) We design a curriculum pre-training method to learn mathematical knowledge and logic, from basic to advanced courses.
- (3) Experimental results on offline evaluation (9 math-related tasks) and online A/B test show the effectiveness of JiuZhang.

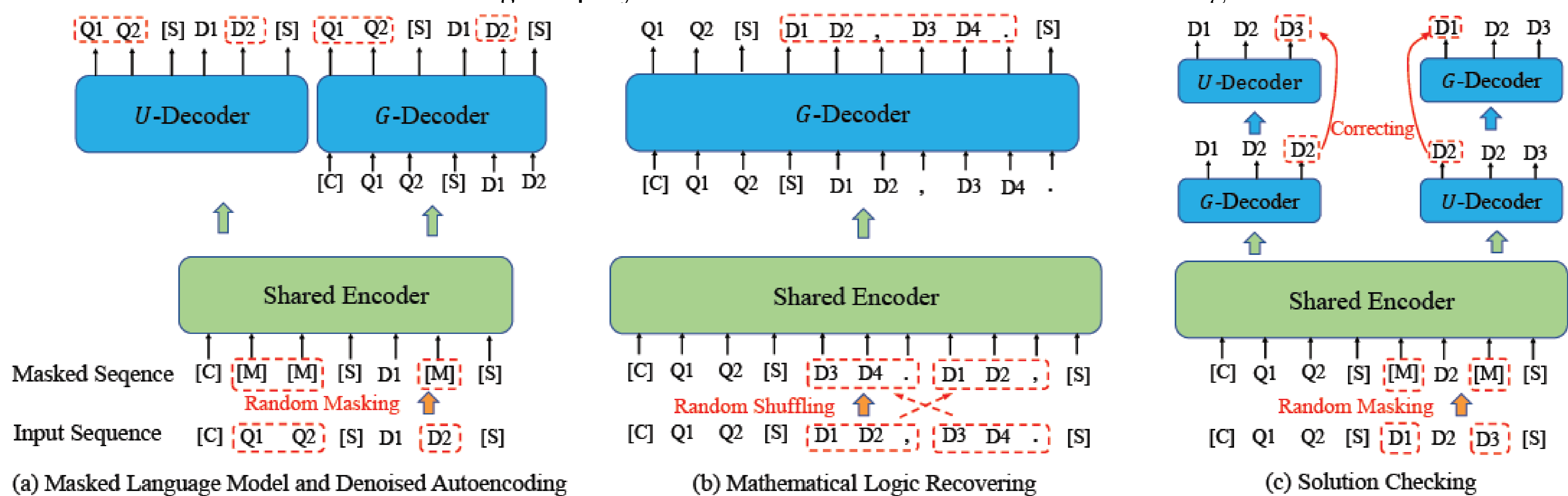
❖ Approach:

➤ 1.Backbone: Unbalanced Encoder-Decoder Transformer

- A Shared Deep Transformer Encoder + Two Shallow Generation- and Understanding-Specific Decoders

➤ 2.Curriculum Pre-training:

- Basic Course: Masked Token Prediction
 - Masked Language Model for U-Decoder & Denoised Auto-Encoding for G-Decoder
 - Position-biased Masking: assigns larger masking weights to words at larger positions
- Advanced Course: Mathematical Logic Recovering
 - Shuffled Sentences Recovering & Shuffled Formulas Recovering
- Advanced Course: Solution Checking
 - Dual-Decoder Solution Checking: employ the two decoders to detect and correct the generated texts from each other



❖ Experiment

(1) Classification: KPC, QRC, QAM; (2) Retrieval: SQR, QAR; (3) QA Task: MCQ, BFQ; (4) Generation: CAG, BAG

Tasks	KPC		QAM		QRC		SQR		QAR	
Metrics	Accuracy	F1-macro	Accuracy	F1-macro	Accuracy	F1-macro	HR@3	NDCG@3	HR@3	NDCG@3
MathBert	68.9	47.1	98.9	98.9	85.3	69.8	0.652	0.656	0.705	0.545
COMUS	71.0	63.3	99.0	99.0	88.0	73.3	0.661	0.664	0.724	0.561
DAPT-CPT	72.0	58.0	99.1	99.1	88.8	76.7	0.664	0.668	0.723	0.556
Ours	73.3	59.4	99.4	99.4	89.4	79.2	0.667	0.672	0.724	0.556

Tasks	MCQ	BFQ	CAG			BAG				
Metrics	Accuracy	Accuracy	BLEU-4	ROUGE-2	ROUGE-L	Accuracy	BLEU-4	ROUGE-2	ROUGE-L	Accuracy
CPT	37.90	46.31	39.56	50.07	68.20	43.34	37.20	47.99	67.25	34.26
DAPT-CPT	46.26	53.41	40.46	50.84	68.87	46.52	38.39	49.28	68.04	41.82
Ours	47.73	54.60	40.81	51.09	69.45	48.51	39.28	49.62	68.37	44.03

Math Problem	已知 $\sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$, 则 $\cos(\frac{\pi}{3} - \alpha) = ()$ It is known that $\sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$, then $\cos(\frac{\pi}{3} - \alpha) = ()$ A. $\frac{\sqrt{5}}{5}$ B. $-\frac{\sqrt{5}}{5}$ C. $\frac{2\sqrt{5}}{5}$ D. $-\frac{2\sqrt{5}}{5}$
Ground-Truth Analysis	$\because \cos(\frac{\pi}{3} - \alpha) = \cos[\frac{\pi}{2} - (\alpha + \frac{\pi}{6})] = \sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$. 故选C $\because \cos(\frac{\pi}{3} - \alpha) = \cos[\frac{\pi}{2} - (\alpha + \frac{\pi}{6})] = \sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$. Therefore, choose C
Ours	$\cos(\frac{\pi}{3} - \alpha) = \{\cos[\pi + \frac{\pi}{3} - \alpha] = \sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$ 故选C $\cos(\frac{\pi}{3} - \alpha) = \{\cos[\pi + \frac{\pi}{3} - \alpha] = \sin(\alpha + \frac{\pi}{6}) = \frac{2\sqrt{5}}{5}$. Therefore, choose C
Math Problem	若复数 $z = \frac{2i}{1-i}$ (i 是虚数单位), 则 z 的共轭复数 $\bar{z} = ()$ If the complex number $z = \frac{2i}{1-i}$ (i is the imaginary unit), then the complex conjugate $\bar{z} = ()$ A. $1 + i$ B. $1 - i$ C. $-1 + i$ D. $-1 - i$
Ground-Truth Analysis	复数 $z = \frac{2i}{1-i} = \frac{2i(1+i)}{(1-i)(1+i)} = -1 + i$, 则 z 的共轭复数 $\bar{z} = -1 - i$. 故选D Complex number $z = \frac{2i}{1-i} = \frac{2i(1+i)}{(1-i)(1+i)} = -1 + i$, then the complex conjugate $\bar{z} = -1 - i$. Therefore, choose D
Ours	复数 $z = \frac{2i}{1-i} = \frac{(2i)(1+i)}{(1-i)(1+i)} = 1 + i$, 则复数的共轭复数 $\bar{z} = -1 - i$. 故选D Complex number $z = \frac{2i}{1-i} = \frac{(2i)(1+i)}{(1-i)(1+i)} = 1 + i$ then the complex conjugate $\bar{z} = -1 - i$. Therefore, choose D