Contextualized Multi-Step Commonsense Reasoning through Context Extension

Hecong Wang[†]; Erqian Xu[‡]; Pinxin Liu[†]; Zijian Meng[†]; Zhen Bai[†]

[†]Department of Computer Science, University of Rochester;

[‡]Warner School of Education and Human Development, University of Rochester

Common Sense in AI

Common sense is a quintessential human capacity but a fundamental challenge of Artificial Intelligence (AI). Human-like commonsense reasoning lies at the center to enable AIs to seamlessly augment human capability while maintaining trust and transparency.

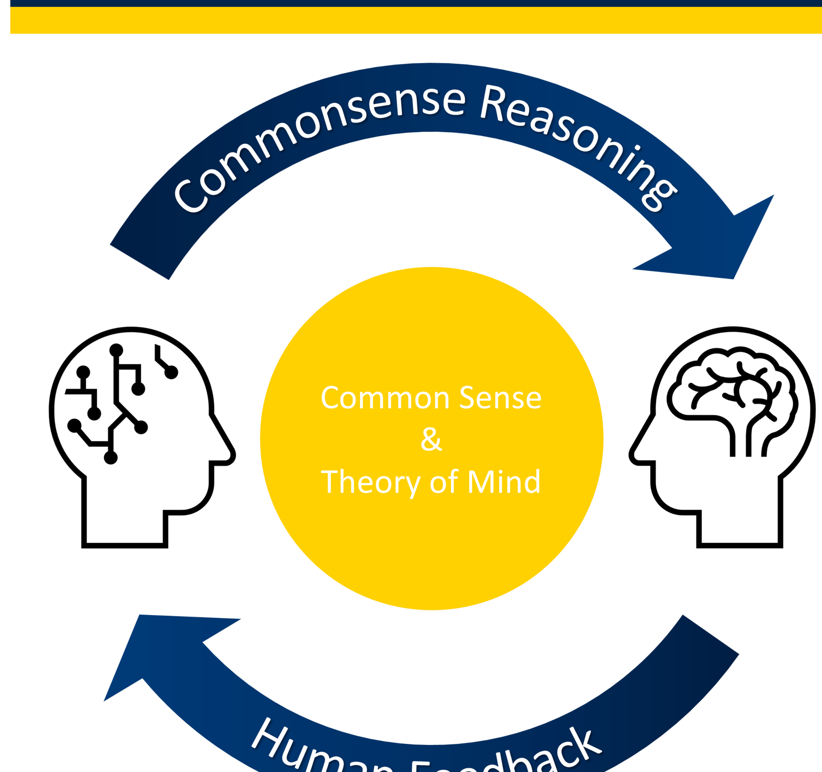
There has been a massive resurgence of interest in commonsense reasoning research in recent years, caused by the rapid advancements in deep learning techniques, the developments of large language models such as OpenAI's GPT-3, and the availability of large-scale commonsense knowledge resources such as ConceptNet.

Multi-step commonsense reasoning is key to achieving human-level performance in AIs. It also enables us to explain, in a human-understandable way, the functioning of AIs. However, making multi-step commonsense reasoning is particularly difficult due to several properties of common sense.

AIs particularly struggle with social commonsense reasoning, which is essential for them to behave as responsible social partners (Sap et al., 2022; Mahowald et al., 2023). The lack of social commonsense reasoning in AIs is largely due to their heavy reliance on static text, which fails to capture the contextual nature of common sense and omits important commonsense knowledge (Sap et al., 2022).

Images generated using OpenAI DALL-E 2 using the prompts "Alex is at the airport for a business trip." (left) and "Zac is at the airport for a Hawaii vacation." (right).

AI-Child Collaborative Reasoning

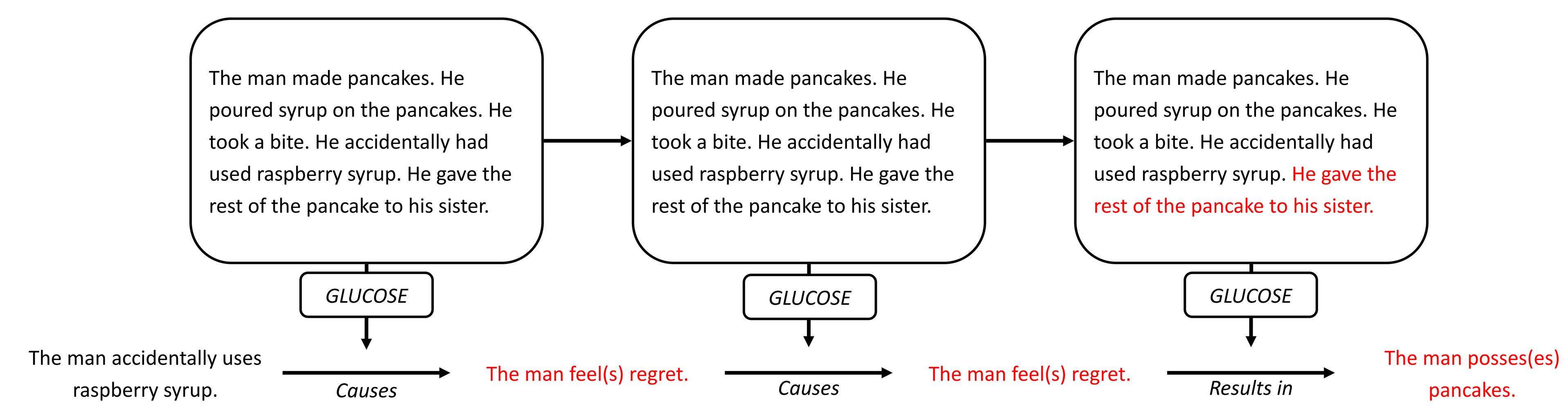


Common sense will allow AIs to better understand and respond to human interactions and enable AIs to explain themselves in a human-understandable way.

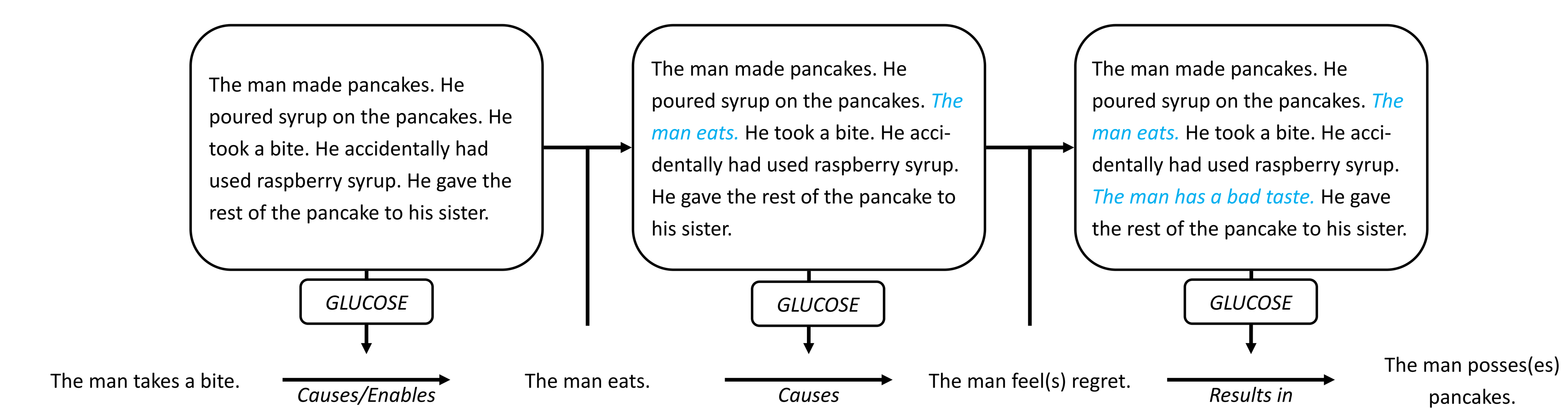
By enabling collaboration and feedback between humans and AIs, AIs can more directly acquire the knowledge and reasoning ability from humans, and humans will also better understand the capabilities and limitations of AIs.

Contextualized Multi-Step Commonsense Reasoning through Context Extension

Reasoning without Context Extension



Reasoning with Context Extension (Random Insertion)



Contextualized Common Sense

Most existing methods for making multi-step commonsense reasoning suffer from two issues:

- **Local contextualization**: Existing commonsense knowledge resources ignored common sense's contextual nature, causing reasoners to apply knowledge inappropriate for the context.
- **Global contextualization**: Existing commonsense reasoners lack the mechanism for effective information sharing, leading to repetitive or incoherent reasoning.

To address the first issue, we adopted GLUCOSE: a contextualized commonsense knowledge resource (Mostafazadeh et al., 2020). For the second issue, we introduced *context extension*: a simple yet effective mechanism for contextualization. In short, the mechanism maintains an informative reasoning context by inserting inferred information into the reasoning context in a coherent manner.

Context extension:

- sustains contextualization over multiple reasoning steps.
- allows a reasoning step to leverage all previous steps.
- can be easily adopted by any search-based reasoner.

References

Davis, Ernest, and Gary Marcus. "Commonsense reasoning and commonsense knowledge in artificial intelligence." *Communications of the ACM* 58.9 (2015): 92-103.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

Mostafazadeh, Nasrin, et al. "GLUCOSE: GeneraLized and COntextualized Story Explanations." *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.

Sap, M., LeBras, R., Fried, D., & Choi, Y. (2022). Neural Theory-of-Mind? On the Limits of Social Intelligence in Large LMs (arXiv:2210.13312). arXiv. https://doi.org/10.48550/arXiv.2210.13312

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2023). Dissociating language and thought in large language models: A cognitive perspective (arXiv:2301.06627). arXiv. https://doi.org/10.48550/arXiv.2301.06627

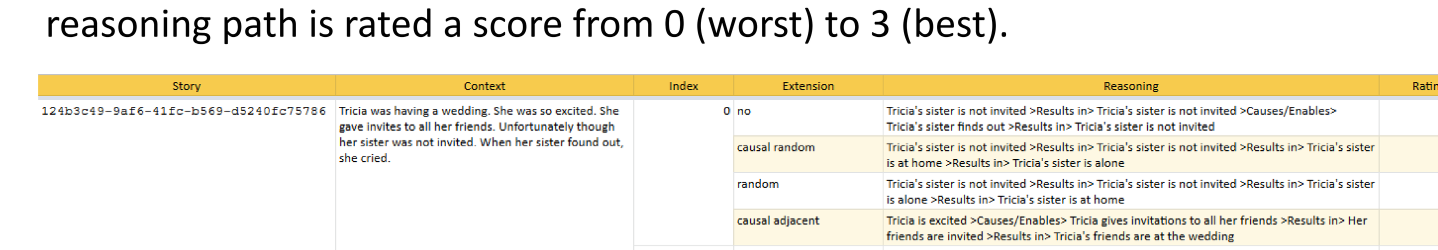
GLUCOSE (Mostafazadeh et al., 2020)

Dimension	**Semi-structured Specific Statement and Inference Rule**: antecedent *connective* consequent
1: Event that directly causes or enables X	A car turned in front of him *Causes/Enables* Gage turned his bike
	Something$_A$ turns in front of Something$_B$ (that is Someone$_A$'s vehicle) *Causes/Enables* Someone$_A$ turns Something$_B$ away from Something$_A$
2: Emotion or basic human drive that motivates X	Gage wants safety *Causes/Enables* Gage turned his bike
	Someone$_A$ wants safety *Causes/Enables* Someone$_A$ moves away from Something$_A$ (that is dangerous)
3: Location state that enables X	Gage was close to a car *Enables* Gage turned his bike away from the car
	Someone$_A$ is close to Something$_A$ *Enables* Someone$_A$ moves away from Something$_A$
4: Possession state that enables X	Gage possesses a bike *Enables* Gage turned his bike
	Someone$_A$ possesses Something$_A$ *Enables* Someone$_A$ moves Something$_A$
5: Other attributes enabling X: N/A (the dimension is not applicable for this example)	
6: Event that X directly causes or enables	Gage turned his bike *Causes/Enables* He fell off his bike
	Someone$_A$ turns Something$_B$ (that is Someone$_A$'s vehicle) *Causes/Enables* Someone$_A$ falls off Something$_B$
7: An emotion that is caused by X: N/A	
8: A change in location that X results in	Gage turned his bike away from the car *Results in* Gage was further from the car
	Someone$_A$ moves away from Something$_A$ *Results in* Someone$_A$ is further from Something$_A$
9: A change of possession that X results in: N/A	
10: Other changes in property that X results in: N/A	

Example entries from GLUCOSE for the story context: "Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee."

Experiments and Results

Four random-walk-based multi-step commonsense reasoners with different context extension mechanisms were implemented and we randomly sampled 50 stories; for each story, five three-step reasoning paths were sampled from each reasoner. Each reasoning path is rated a score from 0 (worst) to 3 (best).



We observed a statistically significant improvement in reasoning quality when context extension is used. Additional experiments also showed that GLUCOSE benefits from the sustained contextualization enabled by context extension. Although context extension is an effective information-carrying mechanism, care should be taken to ensure the coherency of the information.

		Score count					Score distribution (%)				Statistics	
Experiment	Condition	0	1	2	3	Total	0	1	2	3	AVG	STD
Context extension	No extension	66	70	87	27	250	26.4	28.0	34.8	10.8	1.300	0.979
	Random insertion	68	85	64	33	250	27.2	34.0	25.6	13.2	1.248	1.000
	Causal random	52	72	78	48	250	20.8	28.8	31.2	19.2	**1.488**	1.027
	Causal adjacent	40	93	81	36	250	16.0	37.2	32.4	14.4	1.452	0.927
In- and out-of-context	In-context	8	7	9	26	50	16.0	14.0	18.0	52.0	**2.060**	1.150
	Out-of-context	23	6	7	14	50	46.0	12.0	14.0	28.0	1.240	1.302
	Random insertion	29	10	6	5	50	58.0	20.0	12.0	10.0	0.740	1.026
Additional information	Singular, $N=1$	51	92	65	42	250	20.4	36.8	26.0	16.8	1.392	0.993
	Singular, $N=3$	36	87	84	43	250	14.4	34.8	33.6	17.2	1.536	0.940
	Singular, $N=5$	37	82	78	53	250	14.8	32.8	31.2	21.2	**1.588**	0.983
	Multiple, $N=1$	32	56	123	39	250	12.8	22.4	49.2	15.6	**1.676**	0.889
	Multiple, $N=3$	28	73	114	35	250	11.2	29.2	45.6	14.0	1.624	0.861
	Multiple, $N=5$	30	75	118	27	250	12.0	30.0	47.2	10.8	1.568	0.839