Why think step by step? Reasoning emerges from the locality of experience

Ben Prystawski, Michael Y. Li, Noah D. Goodman











Why does reasoning work?

- We can enhance our inferences by working through a series of steps
- But that doesn't give us any new data
- So how does it make our inferences better?





Let's think step by step

• We can get language models to do better on lots of tasks by "chain-ofthought" prompting

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 🗙

(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: Let's think step by step.

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls.

(Kojima et al., 2022)

Standard Prompting	Chain-of-Thought Prompt
Model Input	Model Input
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?	Q: Roger has 5 tennis balls. He buys 2 mor tennis balls. Each can has 3 tennis balls. H tennis balls does he have now?
A: The answer is 11.	A: Roger started with 5 balls. 2 cans of 3 te each is 6 tennis balls. 5 + 6 = 11. The answ
Q: The cafeteria had 23 apples. If they used 20 to	
make lunch and bought 6 more, how many apples	Q: The cafeteria had 23 apples. If they use
do they have?	make lunch and bought 6 more, how many do they have?
Model Output	Model Output
A: The answer is 27. 🗙	A: The cafeteria had 23 apples originally. T 20 to make lunch. So they had 23 - 20 = 3. bought 6 more apples, so they have $3 + 6 = 100$ answer is 9.





The "step to rationality"

• Shepard (2008): thought experiments let us apply internalized intuitive knowledge of principles and symmetries

• Do heavier objects fall faster?



Our hypothesis

Step-by-step reasoning lets (humans/LMs) chain together local inferences between variables they have seen together a lot in order to support longer-distance inferences

Probabilistic inference as language modeling

- We can just write a sample from a Bayes net as a string
- All variables are Booleanvalued

target: X3 X0=1 X1 = 0X2=1 X3 = 1

target: X1 X2=1 X0=1 X3=1 X1 = 0

target: X3 X1=0 X2 = 1X0 = 1

X3=1

6

Estimating using a trained model The direct prediction estimator

Observed Variable X1=0

Target Variable

Get probabilities of 1 and 0 going here, then normalize

target: X2

Reasoning as free generation

- Run the model forward, generating names and values for intermediate variables
- Compute the probability of the target variable when the model generates its name
- Resample intermediate variables / values 10 times, averaging probabilities

Initial prompt target: X5 Observed Variable X1=0

What training conditions lead free generation to outperform direct prediction?

Generating the training data

- We randomly generate Bayes nets
 - 100 nodes
 - 100 random edges
 - Conditional probability tables ~ $\text{Beta}(\frac{1}{5}, \frac{1}{5})$

Holding out pairs

- We select pairs of variables to never co-occur with each other in training
 - Distance at least two
 - High mutual information
- All training conditions use the same set of held-out pairs
- Key metric: how well can a trained language model infer conditional probabilities for held-out pairs? (MSE)

Training data - local neighborhoods

• Each training sample includes only a *local neighborhood* of size $k \sim \text{Geom}(0.5)$

Training data - variable dropout

• We remove a random subset of the variables in the local neighborhood

Training

- Concatenate 1 million samples, showing only the selected variables
- 10 different Bayes nets, separate transformer for each
- 300k gradient steps

The reasoning gap emerges over training

Estimator

- direct prediction
- free generation -

Train condition

- fully observed
- local (geom)

15

An alternative estimator: scaffolded generation

- We generate the smallest set of variables that *d*-separate the observed and target variables
 - Ordered from closest to the observed variable to farthest (in practice they're generally 1-2 variables)

16

Using scaffolds

- Sample the values of the scaffold variables, then get the target variable probability
- Resample 10 scaffolds

Controls

Training Conditions

- pairs)
- the one the samples are drawn from

Estimators

in the scaffold

• Fully-observed: complete samples from the Bayes net (except for held-out

• Wrong local neighborhood: local neighborhoods from a Bayes net other than

• Negative scaffolded generation: reason through random variables that are not

Comparison across all training conditions / estimators

Estimator

direct prediction negative scaffolded free generation scaffolded

19

Falling back on the marginal

Estimator

direct prediction negative scaffolded free generation scaffolded

Conclusions

- structured locally
- between humans and machine learning models

• Reasoning through intermediate variables improves estimation when data is

• Locally structured data + reasoning might help explain the gap in data efficiency

Acknowledgements

Questions? benpry@stanford.edu

Paper:

