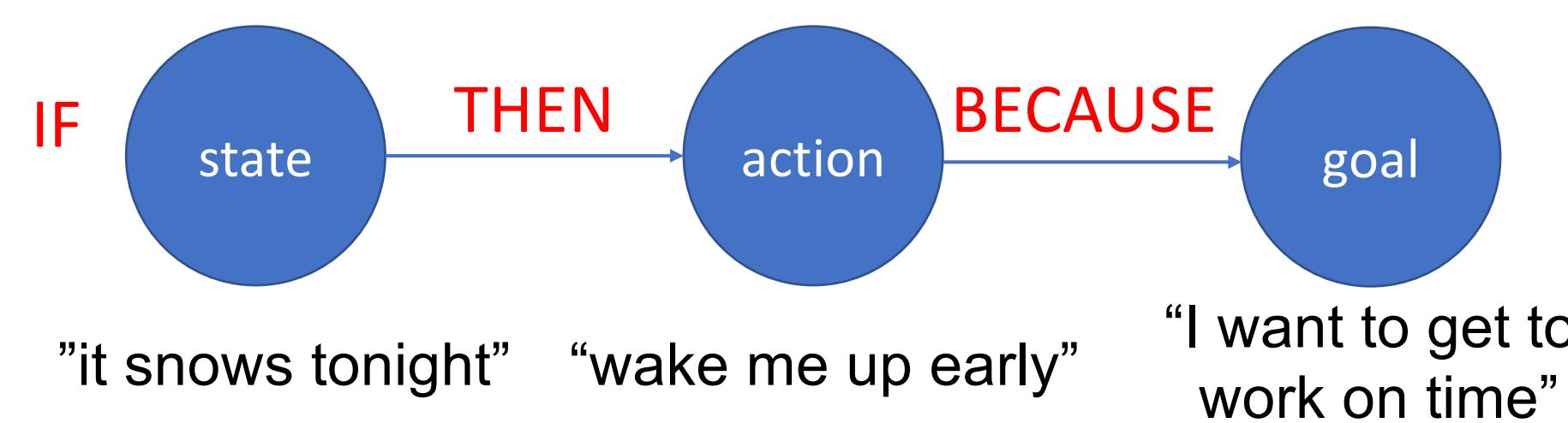


# Reasoning By Instruction Using COMmonsEnse Transformers

Jennifer Lee Senior Thesis Advisor: Tom Mitchell, Forough Arabshahi

## Abstract

LIA, Learning by Instruction Agent, seeks to allow humans to teach the machine a new task through verbal instructions. Though successful at understanding simpler commands, there are reasoning gaps when LIA encounters complex if-then-because commands because the conditional is often underspecified. My goal is to find a path of logic for LIA to complete its knowledge base so that it can reason and mimic common sense.



## Background

COMET[1] is a system that, given an event  $e$ , produces novel commonsense related information.

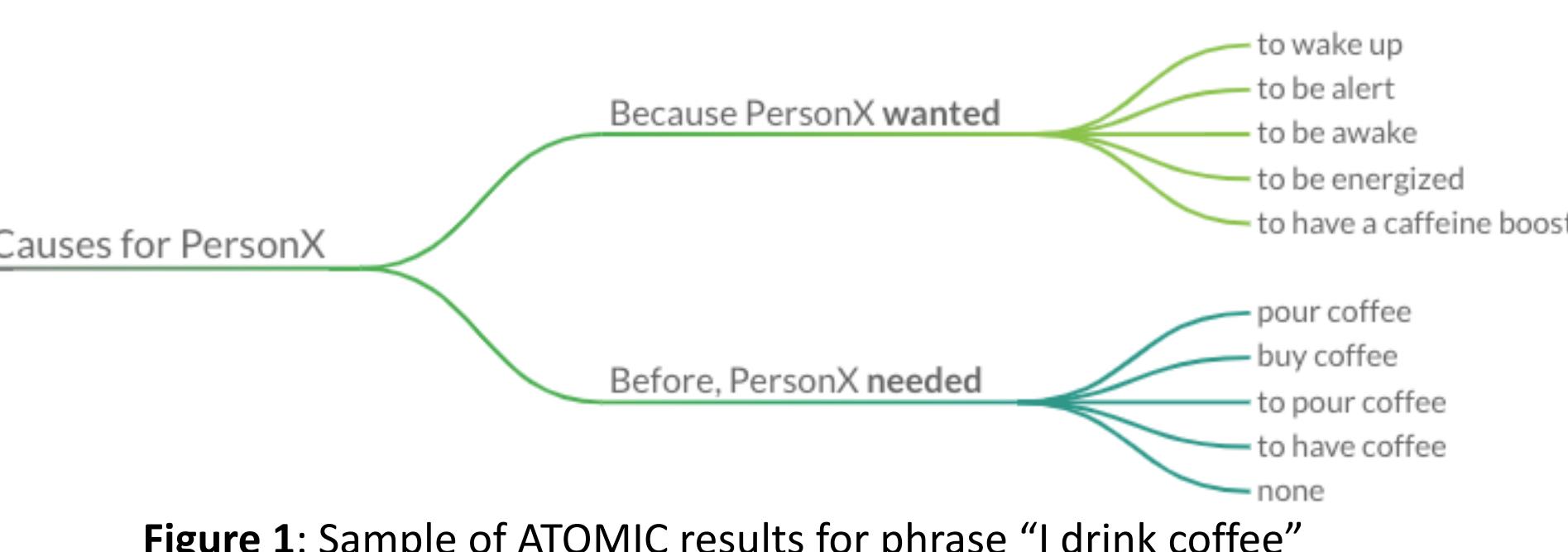


Figure 1: Sample of ATOMIC results for phrase "I drink coffee"

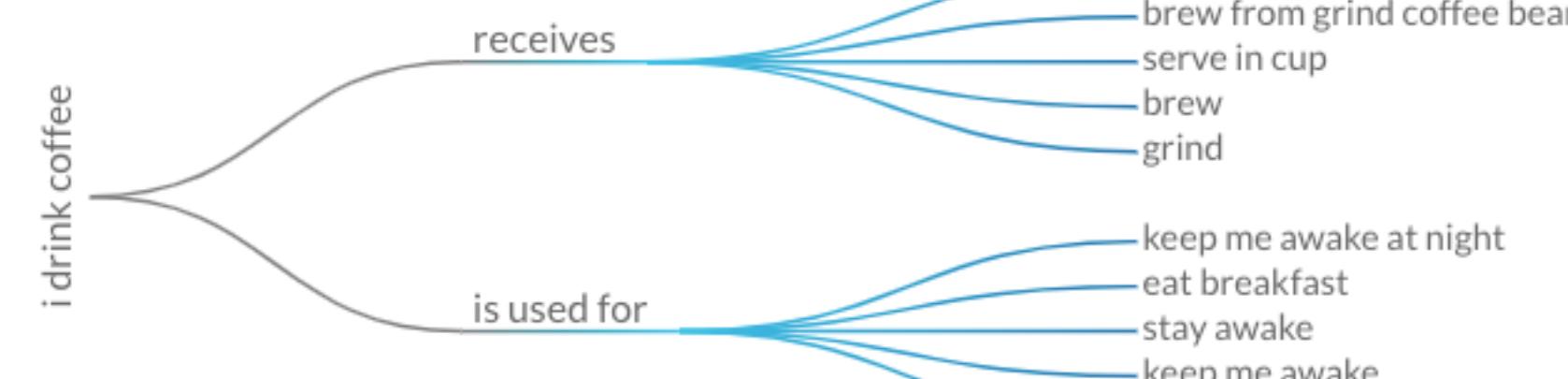
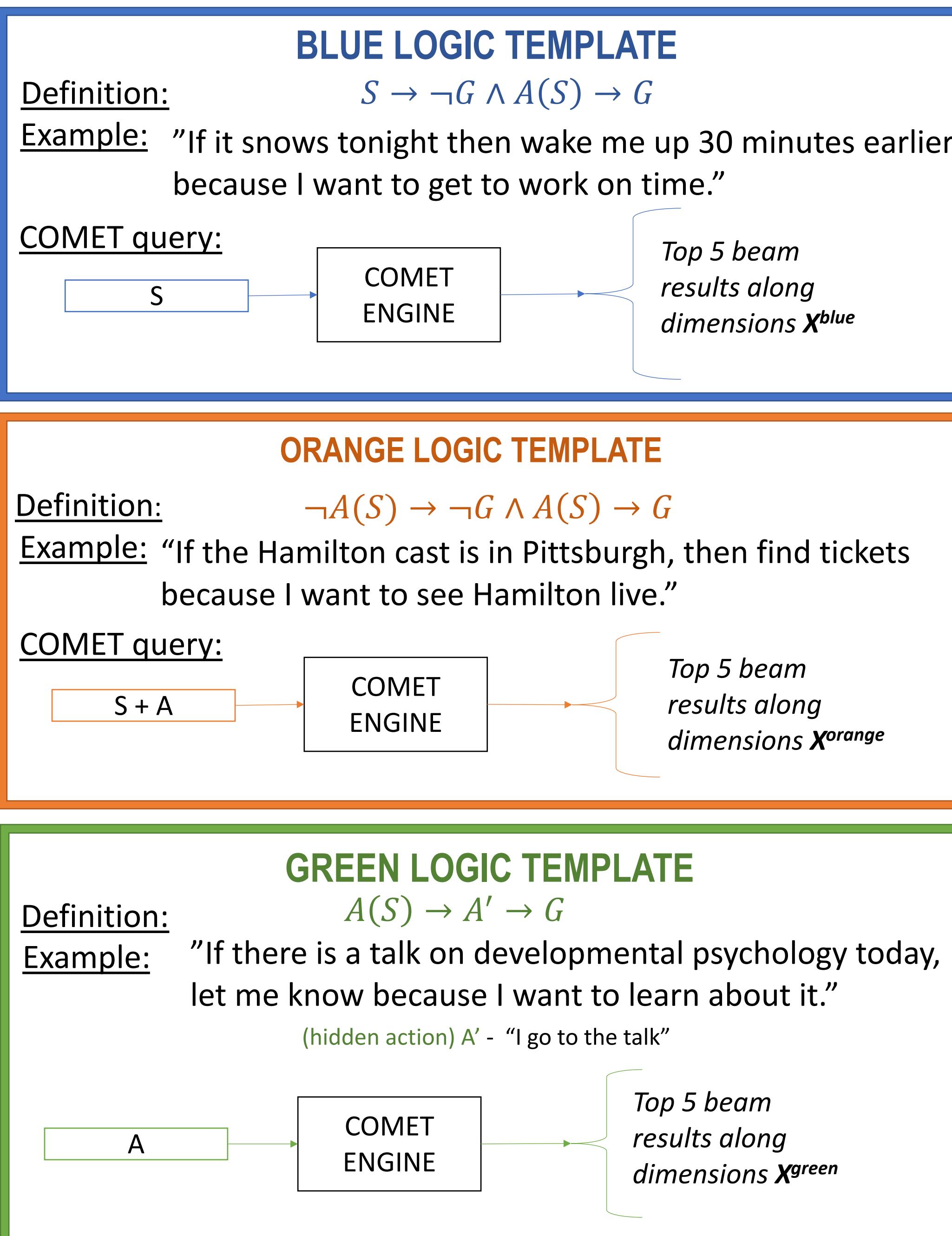


Figure 2: Sample of ConceptNet results for phrase "I drink coffee"

commonsense embeddings - the BERT embeddings re-trained during COMET training; they measure the commonsense semantics of events.

## Methodology

We detected different logic patterns in 160 utterances, which we define as logic templates.



## Results

1. Successfully incorporate COMET into the KB of LIA, thus improving its ability to reason and generate one-hop inferences
2. Reduce the amount of user interaction which improves the user experience
3. Complete proof traces for simpler utterances without using LIA's logic engine

### MAKE ONE-HOP INFERENCESES

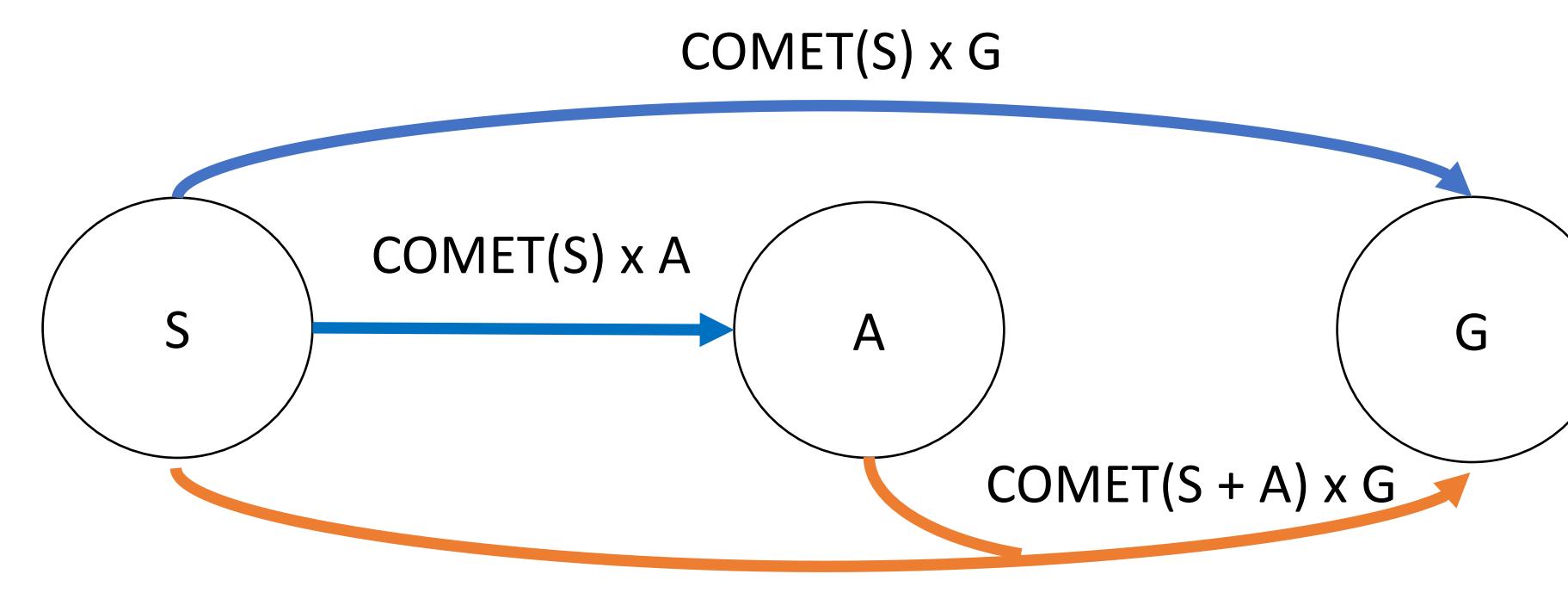


Figure 3: one hop inference simulations for orange and blue logic templates

Let's use the blue template as an example.

1. Generate COMET results for  $S$
2. Create matrix  $M$  of the *normalized* commonsense embeddings  $\forall p \in beam_s$
3. Create a vector of the *normalized* commonsense embeddings of the goal,  $e_G$
4. Calculate  $M \times e_G = scores$
5. if  $scores[i] \geq 0.7$ , add  $G : - beam_s[i]$  to knowledge base  $KB$
- Create multi-hop inferences using a series of connected one-hop inferences.

## USER INTERACTION

LIA also uses COMET results to help user provide the best explanation during user interaction

user> if I have an early morning meeting then wake me up early because I want to get to work on time.  
LIA > How do I know if "I am on time."?  
user> if i get to work before 8  
LIA > What does "i have an early morning meeting" cause that makes you want to "wake up early"?  
1. to be on time  
2. personx is late  
3. to be punctual  
4. to go to work  
5. to get ready  
6. None of the above  
Please provide number and explanation:  
user> 2 if I have an early morning meeting and i wake up at the same time as I usually do, then I will be late to my meeting.

Figure 4: Terminal snippet of user interaction with multiple choice prompt

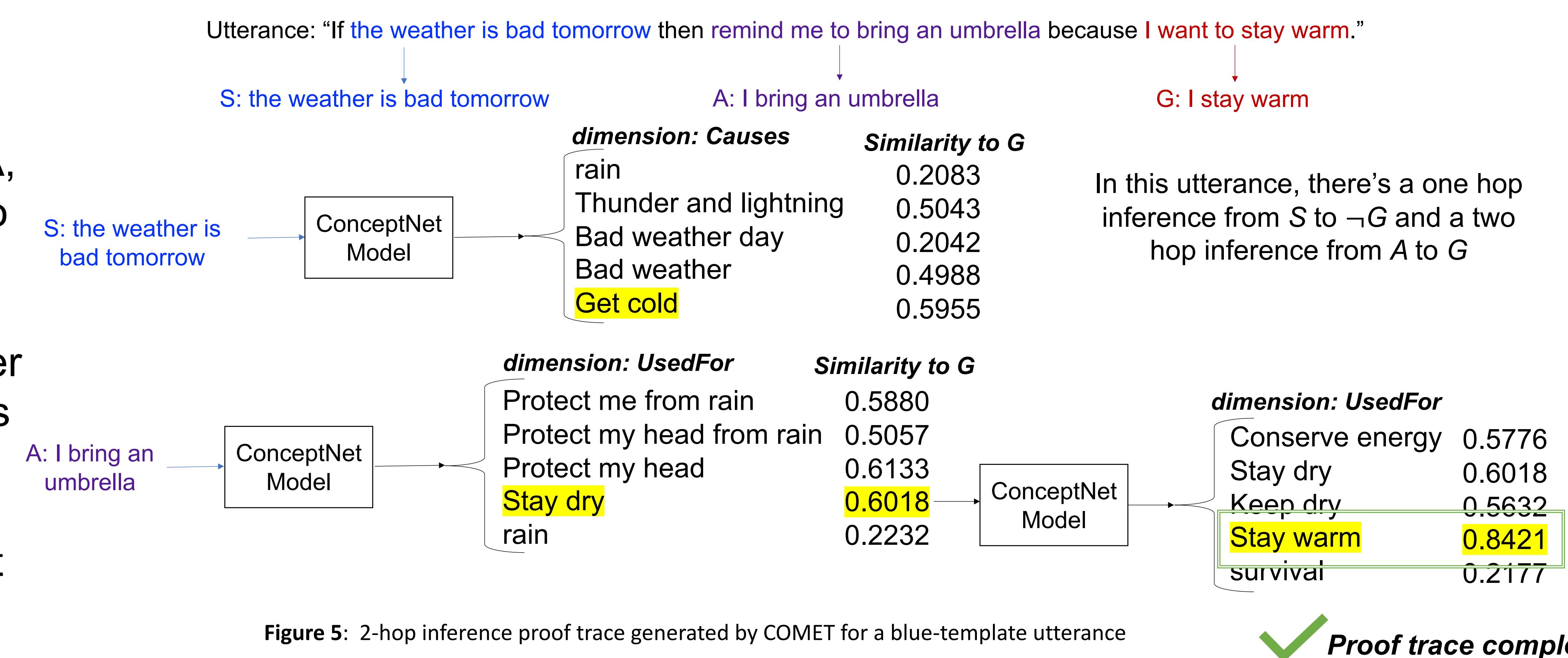


Figure 5: 2-hop inference proof trace generated by COMET for a blue-template utterance