Final Project: Warm-up

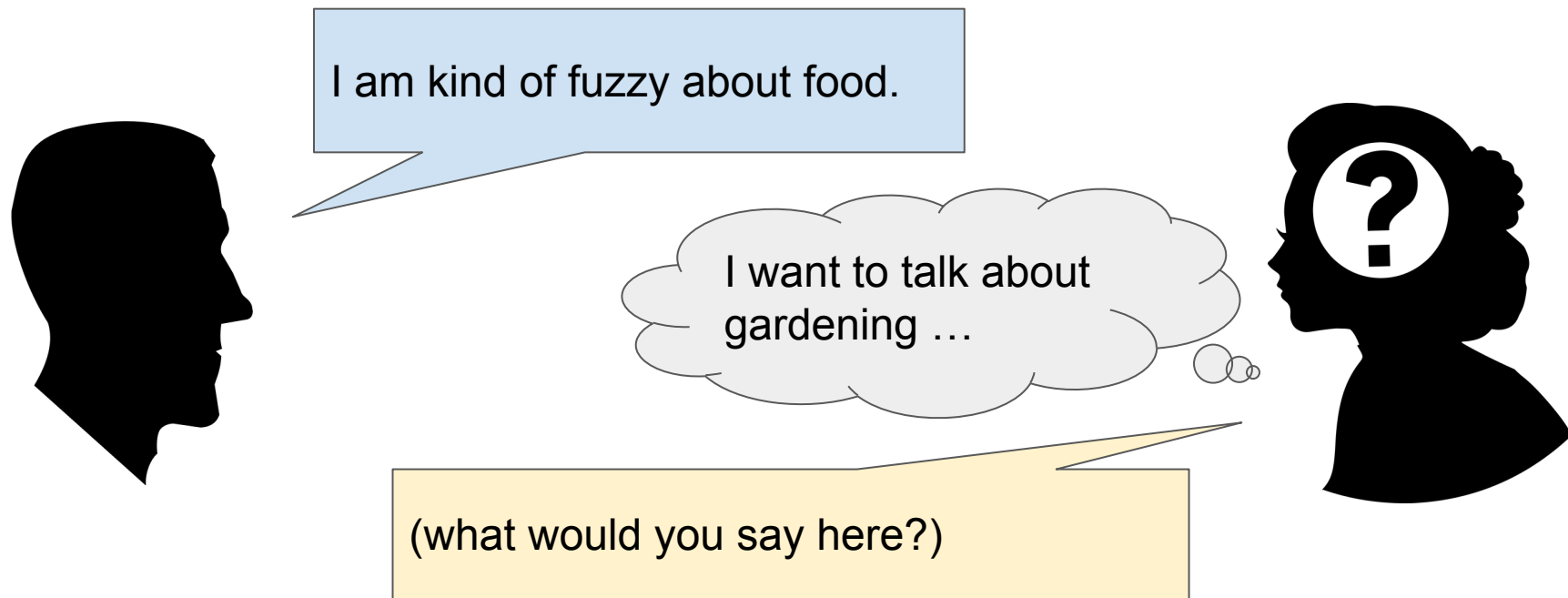Due date: 5/15

IMPORTANT NOTE

❏ ***THIS IS NOT THE FINAL PROJECT !***

❏ The problems would be quite similar
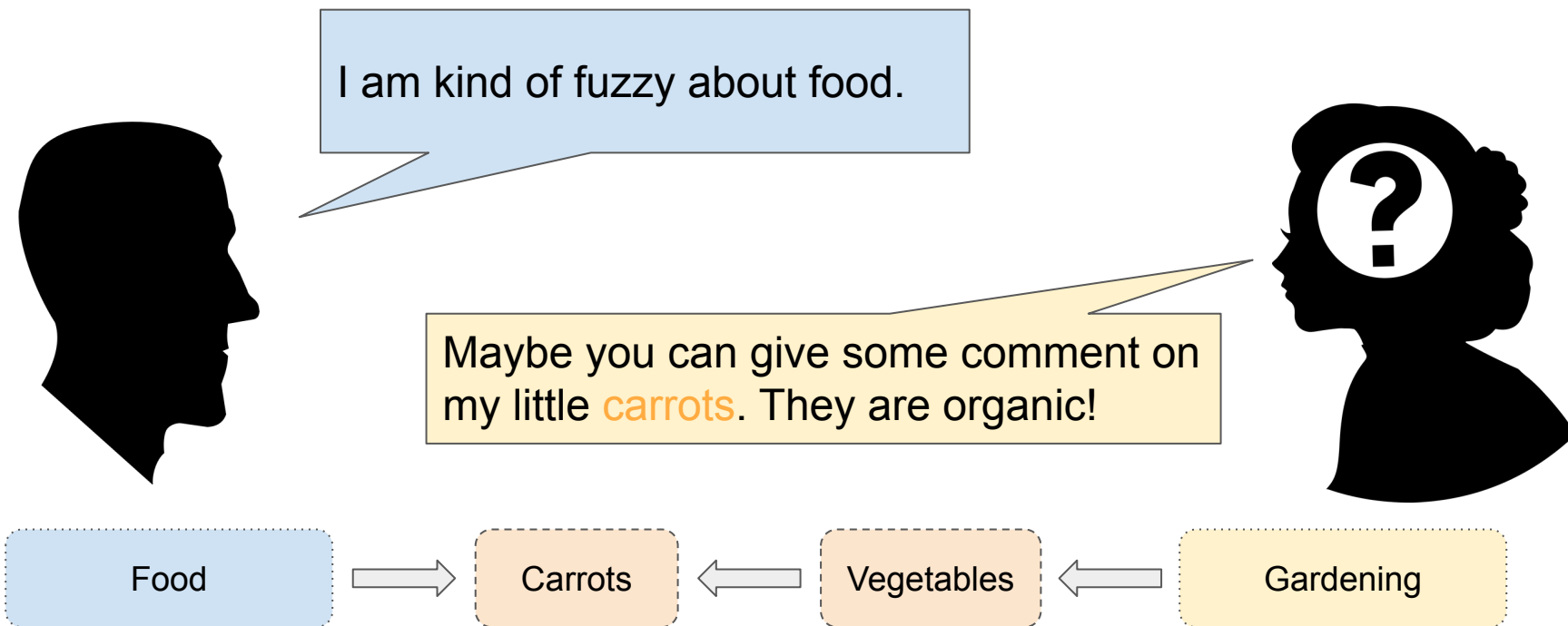
❏ This warm-up aims to get you prepared for the final project
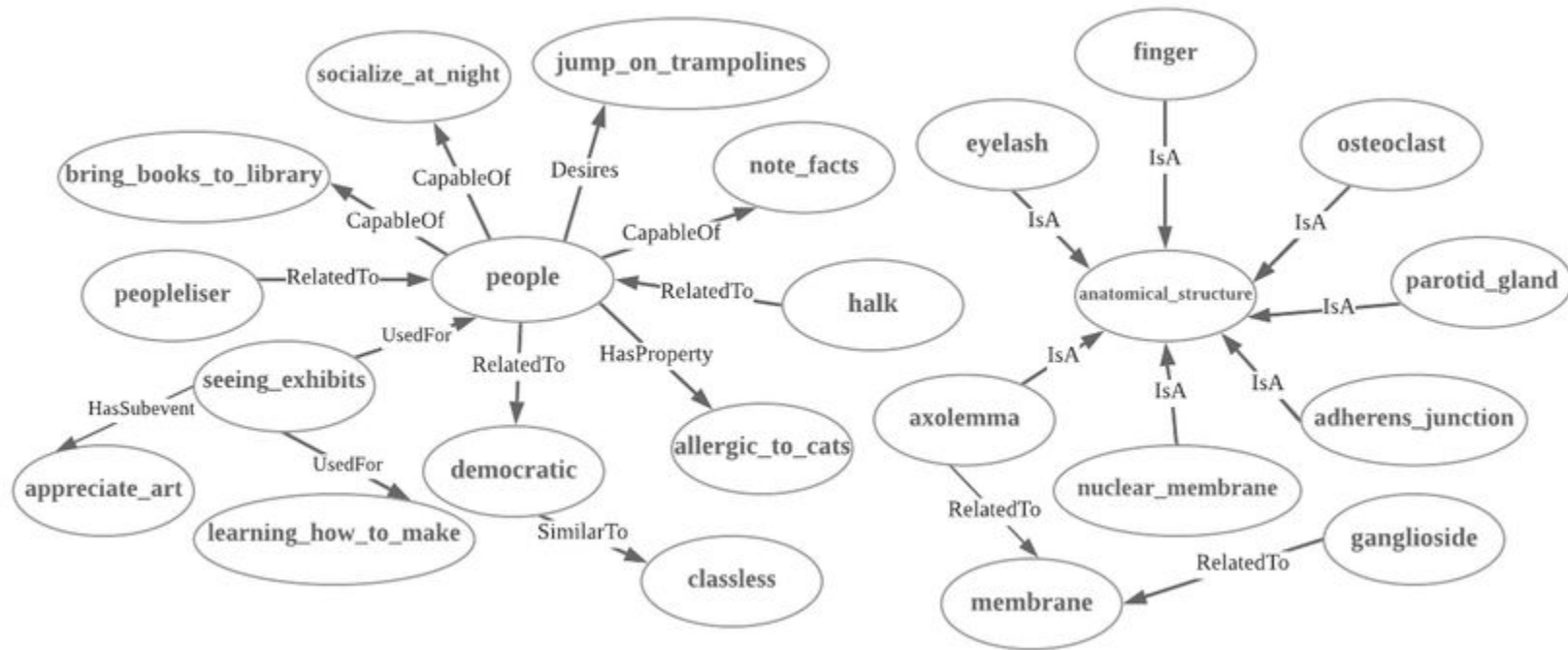
Task Description

Task: Topic Transition

Task: Topic Transition

ConceptNet

ConceptNet

en carrot — RelatedTo ⟶ en food
Weight: 0.45

Source: Verbosity players

en carrot $^{(n)}$ — IsA ⟶ en vegetable $^{(n)}$
Weight: 1.0

Source: OpenCyc 2012

en garden — RelatedTo ⟶ en vegetable
Weight: 3.78

Source: Verbosity players

("gardening" is one word form of "garden")

Task Description

❏ Given

    ❏ Utterance from the other party

    ❏ Target Topic

❏ Output

    ❏ A proper transition sentence so that the conversation turn towards target topic

Dataset

- OTTers: https://github.com/karinseve/OTTers

  - Download data by cloning the repo

NOTE: In this warm-up, target topic is a sentence.

  - Ex. "I like gardening." in the example above

Baseline Models on OTTers
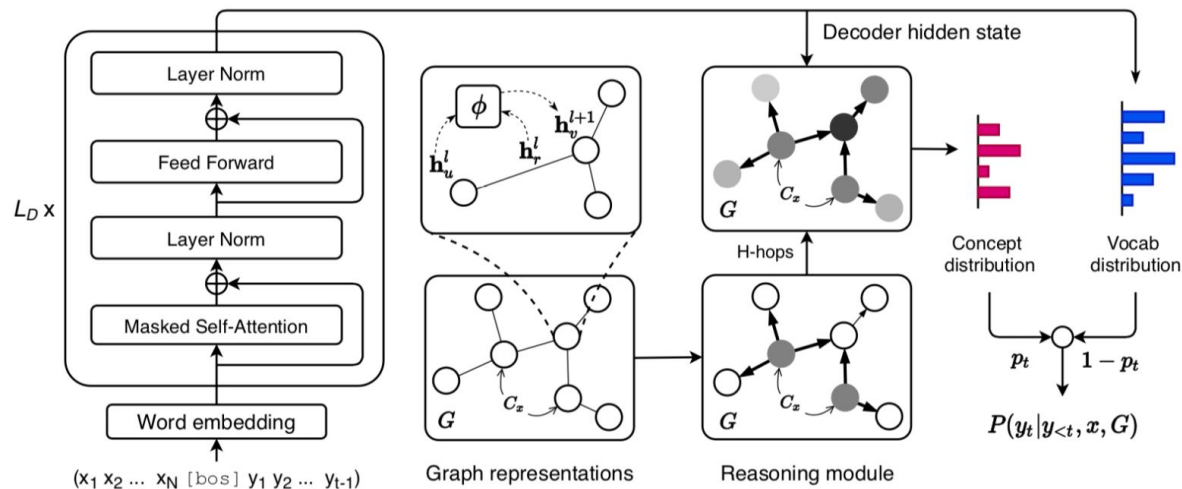
1. Reproduce Multigen

   You can simply follow the steps in the OTTers github repo

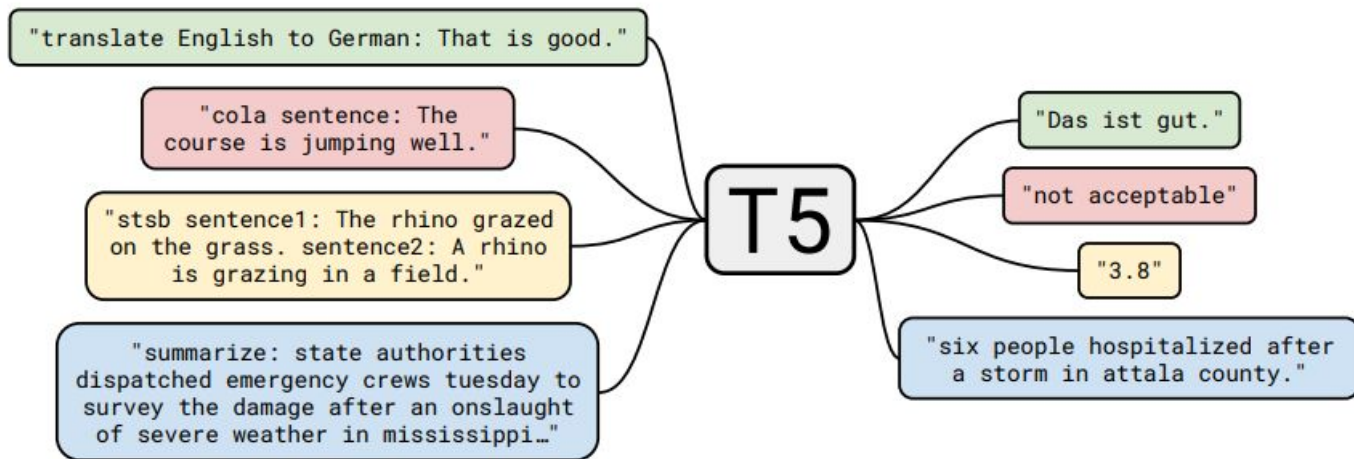2. Fine-tune a T5-small model

   Example script is provided here

Baseline Model 1: Multi-Gen

❏ GPT2 + GCN (consider graph information when decoding)

❏ Input graph is subgraph of *ConceptNet*

Baseline Model 2: T5-small

❏ No more detailed intro!

Evaluation Metrics

❏ N-gram matching based score

    ❏ [sacreBLEU](): overlap across 1-4 ngrams

       ❏ sacrebleu reference.txt -i prediction.txt -b -m bleu -w 3 --lowercase

❏ Fluency score:

    ❏ perplexity measured by GPT2

       ❏ Calculation script also in [here]()

Baseline Metrics

❏ For both in_domain and out_of_domain,

❏ sacreBLEU: around 4 (only your answer)

❏ Perplexity: around 50 (concatenated dialogue)

  ❏ I am kind of fuzzy about food. Maybe you can give some comment on my little carrots. They are organic! I like gardening.

NOTE: These are *suggested* target you should aim for. It would be fine if you do not pass the exact values.

Submission

1. Data & Model

a. Describe what resource you use besides the input utterance and target topic (sentence).
b. Describe your data preprocessing.
c. Describe your model architecture.

2. Metrics and Qualitative Analysis

a. Report your sacrebleu and perplexity score.
b. Pick 5 samples from the test set (from either domain).
   Describe your findings.
   i. Are they good transitions? Why or why not?
   ii. Can you find the hidden concepts on ConceptNet?
   iii. How is the fluency?

Note: No need to cherry pick the best samples. Feel free to discuss any advantages or disadvantages.

Note

- ❏ Write your answer on COOL directly

- ❏ You are expected to get full credits by simply trying both baselines and share one of your results. We hope you can keep your energy on the final project.

- ❏ Again, this warmup is to allow you to understand the task and come up with some models for the target scenarios.