Cooking Common Sense:
Personalized Recipe ‘Tweak’ Inference via Common Sense Reasoning

Shuyang Li*, Bodhisattwa Prasad Majumder*, Julian McAuley

shl008@, bmajumde@, jmcauley@eng.ucsd.edu

* These authors contributed equally

Figure 1: Images of food cooked according to tweaked recipes

Motivation

Rarely do we find that ‘one recipe fits all’. Cooks often ‘tweak’ a recipe to make it more palatable and/or easy to cook. It is infeasible for a recipe retrieval system (online aggregators) to include all reasonable variants of a recipe to cater to individual tastes.

Problem Statement

We propose to approach **personalizing recipes via recipe tweak selection**: This is modeled as an entailment task (Zeller 2018): **choosing the best tweak given a recipe and a goal.**

Recipe Dataset

We collect data from two online recipe aggregators: Allrecipes & Food.com (Majumder et. al 2019). We have additionally collected tweaks from Food.com

Source	# Recipes	Avg # Steps	Avg # Ingredients	# Tweaks	Avg Words / Tweak
Food.com	231,637	9.77	9.05	72,052	64
Allrecipes.com	176,833	7.91	7.44	--	--

Table 1: Recipe and tweak statistics for data scraped from Food.com and Allrecipes

Most common ingredients

Allrecipes:
salt, butter, sugar, onion, water, eggs, olive oil, flour, milk, garlic, pepper, brown sugar, all-purpose flour, baking powder, egg, parmesan cheese

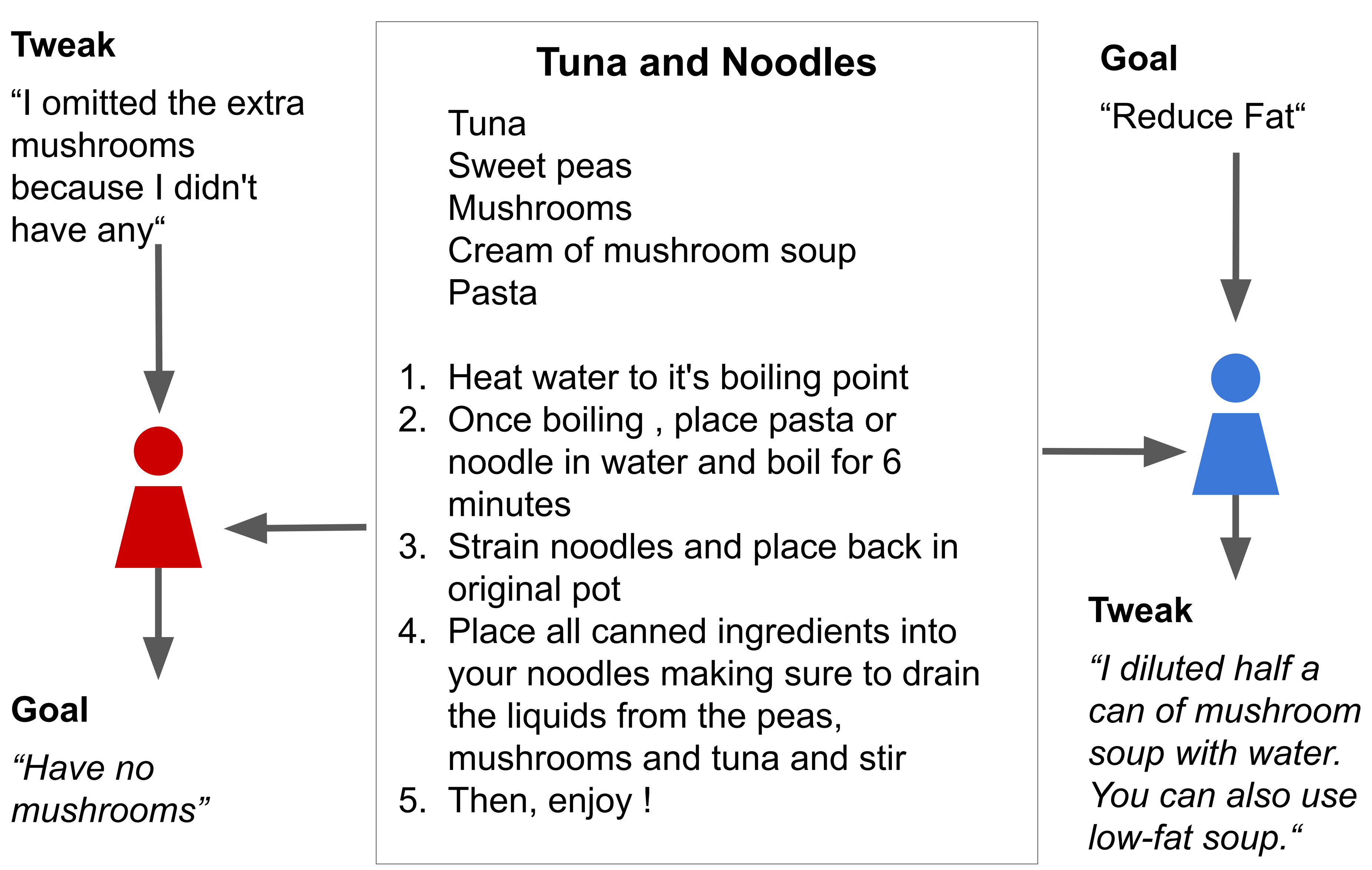
Food.com:
salt, white sugar, butter, eggs, water, garlic, onion, all-purpose flour, black pepper, vanilla extract, olive oil, milk, brown sugar, cinnamon, lemon juice

Recipe	Tweak	Goal
Cherry Pound Cake	I'm not a fan of maraschino cherries so I substituted 2 cans of good quality cherry pie filling.	Don't like maraschino cherries
Seafood Casserole	I used 2% milk in place of both the evaporated and whole milk (it was still very rich).	Less rich
Peppermint Patties	I substituted spearmint for the peppermint and they were very good	Prefer spearmint to peppermint
Fantastic Fish Pie	As I have to cut out fats, I used an unsweetened fat free Greek yoghurt instead of double cream	Reduce fat
Creamy Parmesan Risotto	Switched chicken broth to vegetable broth	Vegetarian

Table 2: Sample tweaks with hand-annotated goals

Tweak Triplet Generation

Tweaks are submitted *ex post facto* by users who have made the recipe in question. We wish to obtain a user’s initial intent, in order to infer an appropriate tweak to the target recipe. We thus pose this as a crowdsourced task in two parts:



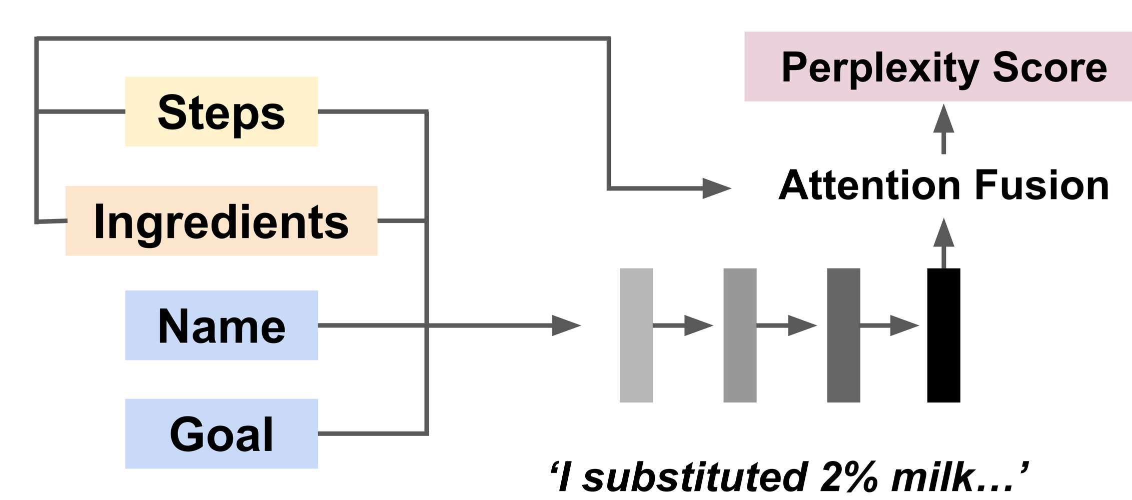


Figure 2: Encoder-Decoder framework with attention over recipe steps and ingredients incorporated in an attention fusion layer

Inference Model

We propose an Encoder-Decoder sequence-to-sequence model encoding a recipe via its name, steps, and ingredients

We additionally apply attention (Vaswani et al. 2017) over steps and ingredients to draw dependencies on specific elements of the recipe to be tweaked.

We frame the multiple-choice as a ranking problem (Radford et al. 2018) and select the most appropriate tweak to be the most likely output sequence as measured by our language model perplexity:

Incorporating Personalization

Drawing from methods in Ni et al. 2017 and Majumder et al. 2019, we propose an additional attention mechanism over previous user interactions (recipes made and reviewed) to incorporate user context in our model. This has been shown to improve the personalization of generated recipes.

Tweak Generation

While we frame the tweak inference task as a ranking task of choosing the ‘most appropriate’ tweak from a list of candidates, our perplexity ranking metric (Mao et al. 2019) allows us to pivot to a generative task: generating natural text tweaks for recipes given the recipe and a goal. We can accomplish this by sampling from the decoder module.

References

Bodhisattwa Prasad Majumder, Shuyang Li, Jianmo Ni, Julian McAuley. 2019. *Generating Personalized Recipes from Historical User Preferences*. In EMNLP.

Huanru Henry Mao, Bodhisattwa Prasad Majumder, Julian McAuley, Garrison W. Cottrell. 2019. *Improving Neural Story Generation by Targeted Common Sense Grounding*. In EMNLP.

Jianmo Ni, Zachary C. Lipton, Sharad Vikram, and Julian McAuley. 2017. *Estimating reactions and recommending products with generative models of reviews*. In IJCNLP.

Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. *Language models are unsupervised multitask learners.*

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. *Attention is all you need*. In NIPS.

Rowan Zellers, Yonatan Bisk, Roy Schwartz, and Yejin Choi. 2018. *SWAG: A large-scale adversarial dataset for grounded commonsense inference*. In EMNLP.