Natural Language Processing
Final Project: Recipe Transformer
Joe Blass, David Ryan, & Ted Wu

How We Represent Recipes

Given a URL of a recipe from AllRecipes.com, we split the recipe into Ingredients and Steps (below)

INGREDIENTS

1st position

Each Ingredient

string search: for substrings in ingredient, search master list for substr

Quantity: the number before the measure, if measure, or the first number. Not always present. If given "1 (4.5 ounce) can", will return "4.5" as Quantity and "ounce" as Measure

Name: at each word after measure, make substrings of a few words; if one or more substrings are found in our ingredients master list, takes the longest as the name (this selects "pork chops" over "pork"). If no match found, the entire string after the measure is treated as the ingredient

Description: everything between the measure and the name (if anything)

Preparation: everything after the name (if anything)

2nd position

Measure: extracted from a master list of measures (if present)

measure & quantity are converted to standard units.

Weight: a multiplier of how much of the "standard amount" of the ingredient is present. Used in transformations. Defaults to 1.

weight = recipe measure / standard measure

RECIPE (steps)

Tools: list comprehension of all tools in master list that are found in recipe

Techniques: list comprehension of all cooking terms in master list that are found in recipe

Associated Tools: for particular techniques, add associated tools to tool list even if not mentioned (i.e., "stir" in recipe adds "spoon" to implements.

Individual steps are NOT represented in our formulation.