Yelp Reviews Annotation Guidelines – Draft

Qishen "Justin" Su, Kelley Lynch, Yuanyuan Ma

February 25, 2016

Contents

1 Introduction

This document provides annotation guidelines for the Yelp restaurant reviews annotation task. The Yelp restaurant reviews annotation task involves (1) annotating the relation between a specific dish mentioned in a review and its described quality or characteristics, (2) annotating the relation between a dish and its ingredients or its parts, and (3) annotating the anaphors of dishes. These guidelines provide details on how to annotate Yelp restaurant reviews.

This document is organized in the following manner: Section 2 is focused on annotating extent tags, i.e. FOOD, QUALITY and ANAPHORA, while Section 3 is focusing on annotating link tags, i.e. PART_OF, OPINION, and COREFERENCE.

When examples of annotation are provided in this document, texts with different extent tags are marked in different colors: blue is for FOOD, red is for QUALITY, and yellow is for ANAPHORA. A QUALITY tag contains three types: *positive*, *negative*, and *neutral*, and in this document, a type of QUALITY will be presented in subscript. Then, each link tag is represented using a predicate argument structure given in Table 1.

Table 1: Link Tag Predicate Argument

Link Tag	Predicate Argument Structure
PART_OF	PART_OF(from=text#1, to=text#2)
OPINION	OPINION(from=text#1, to=text#2, relation=opinion_type)
COREFERENCE	COREFERENCE(from=text#1, to=text#2)

The following is an example of how the format is utilized: For example:

1. I had a **Shrimp dish** that was **out of this world**$_{positive}$ and **it** was served over **amazing**$_{positive}$ **fresh**$_{positive}$ **veggies**.
   PART_OF(from=**veggies**, to=**it**)
   COREFERENCE(from=**it**, to=**Shrimp dish**)
   OPINION(from=**out of this world**, to=**Shrimp dish** , relation="positive")
   OPINION(from=**amazing**, to=**veggies** , relation="positive")
   OPINION(from=**fresh**, to=**veggies** , relation="positive")

2 Extent Tags

2.1 FOOD

The FOOD tag concerns with specific food or beverage names, i.e. corns, beans, burritos, street tacos, margaritas, etc. When annotating a food name, articles (i.e. *a* and *the*) should be excluded. For example:

2. The **chips** and **salsa** were **great**$_{positive}$.
   OPINION(from=**great**, to=**chips** , relation="positive")
   OPINION(from=**great**, to=**salsa** , relation="positive")

One exception where a specific food name is not annotated is when there is no description for that dish or beverage, in terms of ingredients or quality. The following example is a complete review of a customer. In this review, a specific dish, *Beach Burger*, is mentioned. It is not annotated with a FOOD tag, because there is no description for this dish.

3. Phew, this is a GREAT place for sure. The reviews are what got me here and no doubt I WILL be back. Had the Beach Burger.

When the generic term "food" is used, it should not be annotated. In the following examples, the terms "food" and "meal" are not annotated with a FOOD tag, because they are not specific food names.

4. The food here is great!

5. We stopped for our first meal in Phoenix and loved it.

2.2 QUALITY

2.3 ANAPHORA

3 Link Tags

3.1 PART_OF

The PART_OF link tag describes a part-whole relation between an ingredient and a dish, or between a dish and a combo. The predicative argument structure PART_OF(from=text#1, to=text#2) indicates, *text#1 is part of text#2.* The following is an example of PART_OF relation between ingredients and a dish.

6. I just couldn't believe my eyes when I saw all the stuff they squeezed into **both sandwiches**! **Ham**, **pork sirloin**, **chorizo**, **sausage**, **egg**, **avocado**, **jalapenos**, **breaded beef**.

   PART_OF(from=**Ham**, to=**both sandwiches**)

   PART_OF(from=**pork sirloin**, to=**both sandwiches**)

   PART_OF(from=**chorizo**, to=**both sandwiches**)

   PART_OF(from=**sausage**, to=**both sandwiches**)

   PART_OF(from=**egg** to=**both sandwiches**)

   PART_OF(from=**avocado**, to=**both sandwiches**)

   PART_OF(from=**jalapenos**, to=**both sandwiches**)

   PART_OF(from=**breaded beef**, to=**both sandwiches**)

3.2 OPINION

3.3 COREFERENCE