Yelp Reviews Annotation Guidelines – Draft

Qishen "Justin" Su, Kelley Lynch, Yuanyuan Ma
 February 23, 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 dialog, while Section 3 concerns with non-dialog discourse. Each section presents various types of connections between text messages with specific examples. Many of the types in Section 2 are adapted from the backward looking functions in the DAMSL framework (Dialog Act Markup in Several Layers), ¹.

When examples of annotation are provided in this document, the *id* of each example text message is marked in **blue**; for example, id="m0001". The *time* field represents the time and date when a text message is sent. The *participant id* is the identifier of a sender; different participants have different ids.

Then, the connection between two text messages are represented using a predicate argument structure, in the form of CONNECTION(from_id=id#1, to_id=id#2, relation=relation_type), which indicates that the text message of which the id is id#1 is connected to the text message with id#2 in certain type of discourse relation.

¹Allen, J. & Core, M. (1997, March 22). Draft of DAMSL: Dialog Act Markup in Several Layers. Retrieved February 1, 2016, from https://www.cs.rochester.edu/research/speech/damsl/RevisedManual/node12.html

2 Dialog

2.1 Agreement

Agreement can be a positive or negative response of a participant to a certain proposal, request, statement or belief. In general, a participant could accept or reject certain proposal. Occasionally, a participant may have an uncertain response to a proposal or statement. Sections 3.1 through 3.3 defines each subcategory and provides corresponding examples.

2.1.1 Acceptance

Acceptance can be positive responses to proposals, requests, or assertions. Common key words of acceptance are "yes", "ok", "alright", etc. For example, CONNECTION(from_id="m0121", to_id="m0122") shows that participant 152255 agrees with the assertion in m0121.

```
1. id="m0119" time="2015-02-03 21:37:05 UTC" participant="152255" I want to take you out and treat you id="m0120" time="2015-02-03 21:37:13 UTC" participant="152212" Hell no id="m0121" time="2015-02-03 21:37:17 UTC" participant="152212" but yes we need to hang out id="m0122" time="2015-02-03 21:37:20 UTC" participant="152255" Yes!

CONNECTION(from_id="m0119", to_id="m0120", relation = Rejection) CONNECTION(from_id="m0120", to_id="m0121", relation = Elaboration) CONNECTION(from_id="m0121", to_id="m0122", relation = Acceptance)
```

2.1.2 Rejection

Rejection are negative responses to proposals, requests, or assertions. Rejects often includes words like "no" or "nah". In example 1 of Section 2.1.1, m0120 is a rejection to the proposal in m0119. In example 2, message m0030 declines the proposal in m0028, which is also a case of rejection.

```
2. id="m0028" time="2015-02-03 21:02:56 UTC" participant="152255"
Omg when does your class end? Maybe we can get a drink?
id="m0029" time="2015-02-03 21:02:56 UTC" participant="152255"
Or is that too late for you?
id="m0030" time="2015-02-03 21:15:23 UTC" participant="152212"
I won't be able to. I am sorry. I am pretty busy through february. I'm going away a lot. But if you'll be at school tomorrow, ill try to get there early enough to chat!
CONNECTION(from_id="m0028", to_id="m0029", relation = Alternative)
CONNECTION(from_id="m0028", to_id="m0030", relation = Rejection)
```

2.1.3 Maybe

Maybe applies to non-committal responses to a proposal or assertion. Some example responses are "maybe", "let me think about it", etc, which may leave a proposal open. In the following example, participant 153417 does not directly accept or reject the request of participant 153419, and leaves the request open; therefore, this type of connection is considered as "maybe".

```
3. id="m0123" time="2015-01-19 16:28:15 UTC" participant="153419"

Can you proofread my thing for me this weekend? Lol i think it sucks. It needs some crystal pizzazz

id="m0124" time="2015-01-19 17:07:12 UTC" participant="153417"

You silly

CONNECTION(from_id="m0123", to_id="m0124", relation = Maybe)
```

2.2 Understanding

This type of responses concerns if a participant understands a statement from a previous message, and the examples of signaling understanding and non-understanding are given in the following subsections.

2.2.1 Acknowledgment

Acknowledgments signal a participant's understanding of a previous message. Often, acknowledgments consist of words or phrases like "ok", "I understand", "yes", etc.(see example 4). Sometimes, an acknowledgement can interrupt continuous text messages (see example 5).

```
Im just in a bit of an overwhelmed state (which is my status quo)
id="m0124" time="2015-02-03 21:37:38 UTC" participant="152255"
Yeah I know
CONNECTION(from_id="m0123", to_id="m0124", relation = Acknowledgment)

5. id="m0061" time="2015-02-03 21:28:32 UTC" participant="152212"
you need to trust him
id="m0062" time="2015-02-03 21:28:32 UTC" participant="152255"
Ok
id="m0063" time="2015-02-03 21:28:32 UTC" participant="152212"
or youll drive yourself crazy
CONNECTION(from_id="m0061", to_id="m0062", relation = Acknowledgment)
CONNECTION(from_id="m0061", to_id="m0063", relation = Alternative)
```

4. id="m0123" time=112015-02-03 21:37:32 UTC" participant=11152212"

2.2.2 Non-Understanding

Non-understanding is often signals by clarification questions. Note that not all clarification questions signals non-understanding, and to test if a text message signals non-understanding, it should be able to be phrased as "what did you say" or "what did you mean" (Allen 1997). The following are the examples of non-understanding.

```
6. id="m0209" time="2015-02-04 01:45:39 UTC" participant="152212"
  yeah i think he might be danger of losing his job
  id="m0210" time="2015-02-04 01:45:48 UTC" participant="152252"
  Really??
  CONNECTION(from_id="m0209", to_id="m0210", relation = Non-understanding)
7. id="m0086" time="2015-02-09 20:30:37 UTC" participant="152115"
  And we have a city they were born in.
  id="m0087" time="2015-02-09 20:30:45 UTC" participant="152115"
  John and Katherine
  id="m0088" time="2015-02-09 20:31:35 UTC" participant="152116"
  What
  id="m0089" time="2015-02-09 20:32:04 UTC" participant="152116"
  They were born in the same place?
  CONNECTION(from_id="m0086", to_id="m0087", relation = Clarification)
  CONNECTION(from_id="m0086", to_id="m0088", relation = Non-understanding)
  CONNECTION(from_id="m0086", to_id="m0089", relation = Question)
```

2.3 Directives

This type of discourse relation is used when a participant believes the other participant can perform certain action and wants them to carry out such an action.

2.3.1 Request

The "Request" relation is used when a participant asks another participant to perform certain action.

```
8. id="m0135" time="2015-02-21 23:23:44 UTC" participant="149389"
I obv won't hear from them until this week. You are in a hotel!?
id="m0136" time="2015-02-21 23:25:52 UTC" participant="149389"
Oh! Actually I just found the business card template too!
id="m0137" time="2015-02-21 23:25:52 UTC" participant="152212"
no Im bot
id="m0138" time="2015-02-21 23:25:52 UTC" participant="152212"
call me
CONNECTION(from_id="m0135", to_id="m0136", relation = Change of Subject)
```

```
CONNECTION(from_id="m0135", to_id="m0137", relation = Answer)
CONNECTION(from_id="m0135", to_id="m0138", relation = Request)
```

2.3.2 Suggestion

When a participant provides another participant an idea or plan for consideration of a future action, this type of discourse relation is considered as "Suggestion".

```
9. id="m0099" time="2015-02-03 21:32:48 UTC" participant="152255"
I'm working hard to trust him thoigh. Giving him space.
id="m0100" time="2015-02-03 21:32:48 UTC" participant="152212"
space and trust arent the same thin
id="m0101" time="2015-02-03 21:32:48 UTC" participant="152212"
thing
id="m0102" time="2015-02-03 21:32:48 UTC" participant="152212"
id ask yourself what you need
id="m0103" time="2015-02-03 21:34:56 UTC" participant="152212"
out of the relationship
CONNECTION(from_id="m0099", to_id="m0100", relation = Rejection)
CONNECTION(from_id="m0100", to_id="m0101", relation = Correction)
CONNECTION(from_id="m0100", to_id="m0102", relation = Suggestion)
CONNECTION(from_id="m0102", to_id="m0103", relation = Elaboration)
```

2.4 Question

This type of connections concerns with requests of information and clarification. Unlike the clarification questions mentioned previously, this type of questions do not signal non-understanding, but generally request for additional information. Note that in example 11, although the response "Really???" in m0072 is a clarification question, it does not signal that participant 152116 does not understand message m0071.

```
10. id="m0004" time="2015-02-03 20:58:40 UTC" participant="152212" i started seeing a new therapist id="m0006" time="2015-02-03 20:58:40 UTC" participant="152255" How is it going? id="m0007" time="2015-02-03 20:58:40 UTC" participant="152255" Is it a better therapist? Are you on campus? CONNECTION(from_id="m0004", to_id="m0006", relation = Question) CONNECTION(from_id="m0006", to_id="m0007", relation = Question)
```

11. id="m0071" time="2015-02-09 20:15:37 UTC" participant="152115" My mom wanted to call Joe earlier so she's talking to him now.

```
id="m0072" time="2015-02-09 20:16:10 UTC" participant="152116" Really???
id="m0073" time="2015-02-09 20:16:28 UTC" participant="152116" Does she have the questions?
id="m0074" time="2015-02-09 20:16:36 UTC" participant="152116" Was she anxious?
CONNECTION(from_id="m0071", to_id="m0072", relation = Question)
CONNECTION(from_id="m0071", to_id="m0073", relation = Question)
CONNECTION(from_id="m0071", to_id="m0074", relation = Question)
```

2.5 Answer

This kind of connections responses to the request of information of a previous message. An answer can be either complete or partial; or it can a "hold" with which a participant signals their acknowledgment of a question, but does not provide an answer to it.

2.5.1 Answer

An answer provides information to a question in a previous text message. In the following example, m0014 is the answer to the question in m0011.

```
12. id="m0011" time="2015-02-03 20:58:40 UTC" participant="152255"
What are you teaching?
id="m0012" time="2015-02-03 21:00:33 UTC" participant="152212"
good question
id="m0013" time="2015-02-03 21:00:35 UTC" participant="152212"
haha
id="m0014" time="2015-02-03 21:00:48 UTC" participant="152212"
no Im getting paid to go in to a DMI class and discuss some readings
CONNECTION(from_id="m0011", to_id="m0012", relation = Hold)
CONNECTION(from_id="m0012", to_id="m0014", relation = Emotions)
CONNECTION(from_id="m0011", to_id="m0014", relation = Answer)
```

2.5.2 Hold

A participant sometimes signals their acknowledgment of a question, but does not provide an answer to it. In example 12 of Section 2.5.1, message m0012 is considered as a "hold", since it addresses the question in m0011, but does not provide an answer to it. The following example is another case of hold, where the question is not answered.

```
13. id="m0039" time="2015-02-01 19:03:44 UTC" participant="153419" Yayy. Should i bring oliver next week?! Lol id="m0040" time="2015-02-01 19:04:26 UTC" participant="153417" Lol you crazy! CONNECTION(from_id="m0039", to_id="m0040", relation = Hold)
```

2.6 Social Obligations

This type of discourse relation is used when a participant complies with certain social norms or obligations, such as apologies or appreciation.

2.6.1 Apology

This discourse relation is used when one participant apologizes when they realize a social norm is violated, and an apology could be either direct or indirect.

```
14. id="m0078" time="2015-02-09 20:18:56 UTC" participant="152116"
I cant take This you know that
id="m0078" time="2015-02-09 20:19:25 UTC" participant="152115"
I don't know! I can't hear! Haha. Sorry
CONNECTION(from_id="m0078", to_id="m0079", relation = Apology)
```

2.6.2 Gratitude

This relation is used when a participant expresses their gratitude or appreciation to another participant.

```
15. id="m0109" time="2015-02-03 21:34:56 UTC" participant="152212"

your a very special woman, a very beautiful woman. A supportive and understanding partner. hes lucky to have found you.

id="m0110" time="2015-02-03 21:34:56 UTC" participant="152255"

Awwww omg

id="m0111" time="2015-02-03 21:34:56 UTC" participant="152212"

None of the insecurities you are expressing are weird.

id="m0112" time="2015-02-03 21:34:56 UTC" participant="152255"

Thank you! I needed to hear that

CONNECTION(from_id="m0109", to_id="m0110", relation = Emotions)

CONNECTION(from_id="m0109", to_id="m0111", relation = Elaboration)

CONNECTION(from_id="m0109", to_id="m0112", relation = Thanks)
```

2.7 Other

Some types of responses are not formally categorized into any one of the previous types; however, it is important to show some representative examples of certain types of responses for annotation purposes

2.7.1 Emotions

A participant may respond to a text message with emotional words or phrases, like laughing words (such as "haha" and "lol"), surprise or excitement words (such as "omg" or "yay"), appreciation words or phrases (such as "awww" and "thank you"), or emoticons. In example 16, m0110 shows the appreciation of participant 152255 toward the compliments in m0109. Example 17 shows two instances of excitement: connection between m0126 and m0127, and connection between m0128 and m0129. Note that m0126 and m0128 are connected because they are two continuous text messages, and m0128 does not respond to m0127.

```
16. id="m0109" time="2015-02-03 21:34:56 UTC" participant="152212"
   your a very special woman, a very beautiful woman. A supportive and understanding partner. he's
   lucky to have found you.
   id="m0110" time="2015-02-03 21:34:56 UTC" participant="152255"
   Awwww omg
   CONNECTION(from_id="m0109", to_id="m0110", relation = Emotions)
17. id="m0126" time="2015-02-03 21:37:52 UTC" participant="152212"
   nate is also coming
   id="m0127" time="2015-02-03 21:38:05 UTC" participant="152255"
   Omg!!! Yaya
   id="m0128" time="2015-02-03 21:38:05 UTC" participant="152212"
   but i decided i didn't want to do a party again
   id="m0129" time="2015-02-03 21:38:06 UTC" participant="152255"
   Yay
   CONNECTION(from_id="m0126", to_id="m0127", relation = Emotions)
   CONNECTION(from_id="m0128", to_id="m0129", relation = Emotions)
   CONNECTION(from_id="m0126", to_id="m0128", relation = Elaboration)
18. id="m0079" time="2015-02-09 20:19:25 UTC" participant="152115"
   I don't know! I can't hear! Haha. Sorry
   id="m0080" time="2015-02-09 20:20:10 UTC" participant="152116"
   D:
   CONNECTION(from_id="m0079", to_id="m0080", relation = Emotions)
```

3 Non-Dialog

3.1 Contingency

This type of discourse relations is used to indicates the situations described in two or more text messages causally influence each other.

3.1.1 Cause

Causes indicate that the information or situations in two text messages influence each other causally, and they are not in a conditional relation (Penn Discourse Treebank 2007). This type of relation is used when the argument of a previous message is the cause, and that of a latter message is the result or effect.

```
19. id="m0040" time="2015-02-04 00:46:53 UTC" participant="152252" There is no street parking id="m0041" time="2015-02-04 00:46:56 UTC" participant="152252" So I gave to leave work early CONNECTION(from.id="m0040", to.id="m0041", relation = Cause)
```

3.1.2 Result

Similar to the "Cause" relation, results also indicates that two arguments have a causal relation, and that they are not in a conditional relation. "Result" is used when the argument of a previous message is the result caused by the situation of a latter message.

```
20. id="m0282" time="2015-02-04 02:44:16 UTC" participant="152212" its just unclear id="m0283" time="2015-02-04 02:46:24 UTC" participant="152212" bc they only had one class so im not sure what kind of flow they have CONNECTION(from_id="m0282", to_id="m0283", relation = Result)
```

3.1.3 Condition

Two text messages are in a conditional relation when the argument of one message is the condition and that of the other message is the consequence. The order of conditional relation does not matter; in other words, the condition can either precede or succeed the consequence. The following examples show two different orderings of condition and consequence.

```
21. id="m0125" time="2015-02-04 01:12:32 UTC" participant="152212" i actually would suggest you rent a garage for the month id="m0126" time="2015-02-04 01:12:32 UTC" participant="152212" especially if you are saving money on gas
```

```
CONNECTION(from_id="m0125", to_id="m0126", relation = Condition)
```

```
22. id="m0263" time="2015-02-20 13:25:06 UTC" participant="152212" also if you and alec can pose next to a baby carriage id="m0264" time="2015-02-20 13:25:10 UTC" participant="152212" or a moped id="m0265" time="2015-02-20 13:25:12 UTC" participant="152212" anything funny id="m0266" time="2015-02-20 13:25:21 UTC" participant="152212" i promise ill make you both iron on t-shirts

CONNECTION(from_id="m0263", to_id="m0264", relation = Elaboration)

CONNECTION(from_id="m0263", to_id="m0265", relation = Alternative)

CONNECTION(from_id="m0263", to_id="m0266", relation = Condition)
```

3.2 Expansion

This type of relations concerns with the expansion and continuity of discourse. It can be completing sentences, proceeding a conversation, clarification, or supplying additional information.

3.2.1 Elaboration

A text message is considered as elaboration of a previous one, when the current message completes or elaborates on the information that the previous message conveys. In the following example, m0004 is considered as an elaboration of m0003.

```
23. id="m0001" time="2015-02-03 20:54:24 UTC" participant="152212" i can't come tonight. I'm prepping for a class on wednesday night that I'm teaching id="m0002" time="2015-02-03 20:58:27 UTC" participant="152212" but thank you for the invite id="m0003" time="2015-02-03 20:58:37 UTC" participant="152212" I'm a bit crazy if you haven't noticed id="m0004" time="2015-02-03 20:58:40 UTC" participant="152212" i started seeing a new therapist CONNECTION(from_id="m0001", to_id="m0002", relation = Elaboration) CONNECTION(from_id="m0002", to_id="m0003", relation = Change of Subject) CONNECTION(from_id="m0003", to_id="m0004", relation = Elaboration)
24. id="m0037" time="2015-02-03 21:24:16 UTC" participant="152212" i think it is just post graduation. Like i have all these little gigs id="m0038" time="2015-02-03 21:24:16 UTC" participant="152212"
```

```
and i nothing is guaranteed id="m0039" time="2015-02-03 21:24:16 UTC" participant="152212" and I'm constantly needing to prove myself id="m0040" time="2015-02-03 21:24:16 UTC" participant="152255" Yeah I've always heard the first two years are the toughest CONNECTION(from_id="m0037", to_id="m0038", relation = Elaboration) CONNECTION(from_id="m0038", to_id="m0039", relation = Elaboration) CONNECTION(from_id="m0037", to_id="m0040", relation = Agreement)
```

3.2.2 Clarification

The "Clarification" relation is used when a text message clarifies or supply additional information for a previous message to avoid confusion. In example 25, m0087 clarifies to whom "they" in m0086 refers.

```
25. id="m0086" time="2015-02-09 20:30:37 UTC" participant="152115"

And we have a city they were born in.
id="m0087" time="2015-02-09 20:30:45 UTC" participant="152115"

John and Katherine

CONNECTION(from_id="m0086", to_id="m0087", relation = Clarification)
```

3.3 Correction

Corrections generally concern with correcting wrong information from a previous text message, such as typos. Any related following messages should connect to the message with typos. In the example below, m0101 corrects the typo in m0100; thus, there is a connection between these two messages. Then, m0100 should connect to m0103.

```
26. id="m0099" time="2015-02-03 21:32:48 UTC" participant="152255"

I'm working hard to trust him thoigh. Giving him space.

id="m0100" time="2015-02-03 21:32:48 UTC" participant="152212"

space and trust arent the same thin

id="m0101" time="2015-02-03 21:32:48 UTC" participant="152212"

thing

id="m0102" time="2015-02-03 21:34:56 UTC" participant="152212"

id ask yourself what you need

id="m0103" time="2015-02-03 21:34:56 UTC" participant="152212"

out of the relationship

CONNECTION(from_id="m0099", to_id="m0100", relation = Rejection)

CONNECTION(from_id="m0100", to_id="m0101", relation = Correction)

CONNECTION(from_id="m0100", to_id="m0102", relation = Suggestion)

CONNECTION(from_id="m0100", to_id="m0103", relation = Elaboration)
```

3.4 Concession

This type of discourse relation is used to highlight prominent differences between the arguments in two text messages. More specifically, "the highlighted differences are related to expectations raised by one argument which are then denied by the other" (Penn Discourse TreeBank 2007).

```
27. id="m0157" time="2015-02-22 01:32:55 UTC" participant="149389"
I should have sent you pics!
id="m0158" time="2015-02-22 01:33:04 UTC" participant="149389"
Though I guess it is nothing compared to boston
CONNECTION(from_id="m0157", to_id="m0158", relation = Concession)
```

3.5 Alternative

This discourse relation is used when two text messages describe alternative situations. Words like "or", "instead" and "otherwise" may be helpful to identify an "Alternative" relation.

```
28. id="m0157" time="2015-02-04 01:33:21 UTC" participant="152212" i wouldnt mind putting 200$ on a cc for a month id="m0158" time="2015-02-04 01:33:23 UTC" participant="152212" to put the car away id="m0159" time="2015-02-04 01:33:25 UTC" participant="152212" or two weeks id="m0160" time="2015-02-04 01:33:28 UTC" participant="152212" or whatever id="m0161" time="2015-02-04 01:33:34 UTC" participant="152212" bc it bothers you so much CONNECTION(from_id="m0157", to_id="m0158", relation = Elaboration) CONNECTION(from_id="m0157", to_id="m0159", relation = Alternative) CONNECTION(from_id="m0157", to_id="m0160", relation = Alternative) CONNECTION(from_id="m0157", to_id="m0161", relation = Result)
```

3.6 Other

Some types of responses are not formally categorized into any one of the previous types; however, it is important to show some representative examples of certain types of responses for annotation purposes.

3.6.1 Change of Subject

It is considered a change of subject when the current text message does not relate to a previous message directly in terms of the information that both messages convey, regardless of time lapse between

two messages. In the first following example, CONNECTION(from_id="m0002", to_id="m0003") shows message m0002 is connected to message m0003, because of the short time span between them. In the second example, message m0145 does not relate to message m0146, because the time difference is large enough to consider that message m0146 initiates a new conversion. Therefore, there is no connection between message m0145 and message m0146.

```
29. id="m0001" time="2015-02-03 20:54:24 UTC" participant="152212"
   i can't come tonight. I'm prepping for a class on wednesday night that I'm teaching
   id="m0002" time="2015-02-03 20:58:27 UTC" participant="152212"
   but thank you for the invite
   id="m0003" time="2015-02-03 20:58:37 UTC" participant="152212"
   I'm a bit crazy if you haven't noticed
   id="m0004" time="2015-02-03 20:58:40 UTC" participant="152212"
   i started seeing a new therapist
    {\tt CONNECTION}(from\_id="m0001",\ to\_id="m0002",\ relation={\tt Elaboration})
    CONNECTION(from.id="m0002", to.id="m0003", relation = Change of Subject)
    CONNECTION(from_id="m0003", to_id="m0004", relation = Elaboration)
30. id="m0145" time="2015-02-03 21:43:00 UTC" participant="152212"
   its silly stuff
   id="m0146" time="2015-02-04 17:35:50 UTC" participant="152255"
    Oh man I got drunk last night and just bombarded Jeff with weird texts about the potluck I was
   having. I apologized this morning and no response.
    CONNECTION(from_id="m0145", to_id="m0146", relation = Change of Subject)
```

3.6.2 Punctuation

Sometimes, a final punctuation of a sentence might be missing, and then supplied in a following message. It is necessary to connect these two messages to retain the semantics of the first sentence. In the following example, the connection between m0131 and m0132 is important to show that m0131 is actually a question, rather than a statement.

```
31. id="m0131" time="2015-02-09 21:21:46 UTC" participant="152116"

And nothing on Katherine
id="m0132" time="2015-02-09 21:21:51 UTC" participant="152116"
?
id="m0133" time="2015-02-09 21:22:19 UTC" participant="152115"
Nope
CONNECTION(from_id="m0131", to_id="m0132", relation = Punctuation)
CONNECTION(from_id="m0131", to_id="m0133", relation = Answer)
```