

## Masters Thesis

Part 2 of the notes document.

### Defining Traces and Gen. Graph

From all of the considerations before (found in the previous document), it suggests that we should define the traces we get and the generated graphs resulting from analyzing the traces separately.

*Trace* =  $(V, U, T)$  as defined before, where  $V$  is a set of nodes,  $U$  are sets of actions, and  $T$  are the transition mappings.

We then define what our generated graph would look like:

*Gen. Graph* =  $(N, L, M, W)$

Where  $N$  is a set of traces, represented by nodes.

$L$  is a set of landmarks, represented by nodes.

$M$  is the transition mappings.

$W$  is the mapping function mapping weights to each transition (where the weights represent the length of the trace).

This definition would reflect the image that we have before.

It also suggests to me that the critical-nodes/landmarks are the thing of interest for us in our model. Finding out how to define them, finding out how to think about them in context, finding out how we could derive/calculate/obtain them, how they fit into the graph, etc.

Critical-nodes/landmarks are natural breaks in the conversation, where from our previous definition, no specific player owns the node. The conversation still proceeds normally, and these nodes are likely to be the ones that are densely connected. I guess from a theoretical perspective, the more landmarks an exchange has, the more difficult it will be (landmarks insinuate more opportunities for creation, and creation is not only the defining aspect of learning, it also is the most challenging task). Analyzing and understanding a graph/model based on the

number of landmarks and/or defining a graph based on a certain number of landmarks (maybe with additional properties?) is definitely useful. The paper that introduced the concept of landmarks also suggests that it would logically give us a metric of distance in the graph. Which for us, I am not quite sure how useful it would be? "Distance" in this context would just be the number of utterances/actions before reaching the landmark. I guess we can say that a longer chain of exchanges is more likely to be prone to error/deviation, and therefore we can probably garner some information from that as well.

The notion here being that traces are highly linear, but we can transform the traces into something less linear, and the resulting graph would theoretically provide us with a framework to do something useful. We just have to formally define it I guess.

It also seems like there may not be an algorithmic way (at least not entirely algorithmic) to determine the connections of the nodes (traces to other traces, traces to landmark, etc). It seems like this graph generation comes from taking lots of traces, and running a trained edge detection model. Which makes sense. I guess my worry is that we don't want to spend a lot of effort obtaining this graph only to find that the information we get isn't so useful? So I guess the formal definitions and clear defining inputs/outputs could help us arrive at a conclusion for whether or not the hypothetically obtained graphs would be useful or not.

I guess it would benefit us to draw up an example of this graph -> specifically cases where we determine short/long cycles in the graphs and what that even may mean in context of a conversation, whether or not the graph still becomes linear, etc etc.

There's something to be said about a problem in a conversation of a foreign language, is the issue of the context switching too fast/the flow of logic is not understood (i.e: "I don't understand how one thought connected to the next", "What I expected them to talk about suddenly changed into something else.") Additionally, this may have connections to different modes of expression between individuals, which further complicates this issue.

Can we alleviate this issue by deriving some understanding from the graph, such that we can discuss guarantees such as, you must pass a certain node before you reach this specific node? (LTL variations?, eventually, until, etc...)

Does the existence of landmarks complicate or alleviate this issue?

## **Pre Meeting 2 Prep**

Some tasks from last time:

- ~~— Fix the graphics~~
- ~~— Sample conversation graph~~
- ~~— Consider how the graph is generated~~
- Research problem to explore

## **Fixing the Graphics**

We'll probably just scrap the graphics from last time and focus on a sample conversation, along with additional drawings for whatever concepts from the papers that we want to expand on.

I think these two graphics would be the most useful right now:

- ~~— Sample conversation~~
- ~~— Generation of graphs (trace → graph)~~

## **Research Problem**

Not sure if what we have is enough for a research question. But there is something interesting in taking traces of conversation exchanges, and utilizing them to generate a unique graph that grants us some representation of possible exchanges, as well as a framework for analysis (i.e: reachable states, etc, granting certain factors such as vocabulary, etc).

## **Some Paper Notes**

"Schema-based conversation modeling for agent-oriented manufacturing systems"

- Seems helpful in confirming some notions that we had, such as it stating that "... communication can be better modeled and more easily implemented when a conversation rather than an isolated message is taken as the primary unit of analysis." This confirms the idea that we were starting to see before

regarding “zooming out” and abstracting over the linear conversations, and instead focusing on small exchanges as a unit with “landmarks” where many edges are possible.

- This paper suggests to me that our set up from our sample conversation will reveal a lot about the schema and possible conversation interactions, which can further shed light on interesting points for analysis.

One of the papers that we stashed introduced the concept of disfluency, which is the notion of making mistakes (repeating words, uncertainty pauses, etc) in normal conversation (that I presume would be otherwise considered correct). This perhaps may be something that could be a side-effect of effectively using a modeled graph? Perhaps something to think about in context of using our generated graph for an analytical purpose. (Considering intent, flow of conversation, reachability of states, can we use all of these notions to minimize disfluency or provide support?)

I guess the theme is that we're trying to apply notions of formal modeling to provide some sort of framework for analysis in order to solve issues that the field traditionally applies language models and machine learning in order to “numerically” solve

## Post Meeting 2 Notes

We have some new directions:

- Move starting from the conversation, to derive something like a sub-conversation within the conversation, away from a graph.
- If we start from a set of labels that have high abstraction (open.question, closed.question, response, check.reception, clarification, etc). But Dimitri said having a higher abstraction is better.
- Create a graphic -> two full conversations, where every node is labeled, we have some way to define a sub-conversation from the conversations, and then we can see if we can relate the two conversations based on sub-conversations.

Next meeting is Monday, 05 October.

### **Sample Conversation - Subconversation**

With the goal of seeing if we can formulate a sub-conversation from a trace of a conversation, and then relate the two sub-conversations resulting from two different traces, we start with defining a set of labels.

It was suggested that we keep the labels broad first.

*open.question*

*closed.question*

*respond.agree*

*respond.deny*

*display.reflection*

*give.opinion*

*deflection \*\**

*use.social.convention*

*relax.atmosphere*

**\*\***We define deflection as a followup that is neither *respond.agree* nor *respond.deny*, but is categorically attached as a response to the original question.

This first set of labels seems reasonable. There are some slight flexibility, such as *give.opinion* labels for sentences that are not opinions but simple declarations, or *respond.agree* to mean both "responded to the question with an actual answer" and "responded to the question in the affirmative" and the same but in the negative for *respond.deny*. (Note that responding by explicitly saying you wish not to answer a question is not deflection, and is often rare in non-hostile environments)

With regards to deflection, unsure how often it comes up in conversation, but it may be interesting to see (because it is a unique category), although I guess we could also group it with *respond.deny*, but I feel they are different enough (it's value neutral rather than a response in the negative or denial to answer, whereas I feel answering a question at all is a response in the affirmative).

## Trace 1 - International Law

- U100 欸，你也是從台灣來的, closed.question
- U101 也是？你的flag是中國的, deflection
- U102 我不知道他們有台灣的國旗, deflection
- U103 當然有，如果你是台灣就因該貼台灣的, give.opinion
- U104 不然現在人家會以為你有corona, relax.atmosphere
- U105 哈哈, relax.atmosphere
- U106 trueeeee, use.social.convention
- U124 你是台灣的哪裏人？, open.question
- U125 臺北，respond.agree
- U126 欸，我也是，respond.agree
- U127 好巧哦，小世界，use.social.convention
- U128 好久沒有回去台灣了，display.reflection
- U129 現在想回去也沒辦法，display.reflection
- U130 欸，可是他們抵抗corona還很好哦，respond.deny
- U131 對，可是去了話可能回不來，respond.agree
- U132 哈哈, relax.atmosphere
- U133 也有一點不想回來，就回台灣好了, display.reflection
- U134 書也不想念了，美國政府也沒好事，display.reflection
- U135 哈哈, relax.atmosphere
- U136 美國真的是，不知道他們在幹什麼。。。, respond.agree
- U137 好，這裏念完書回台灣，give.opinion
- U138 哈哈, relax.atmosphere
- U139 哈哈, relax.atmosphere
- U107 你是台灣來念書的嗎？closed.question
- U108 來念書，可是其實從美國來的, respond.agree
- U109 我台灣出生的，可是在美國長大的, respond.agree
- U110 你是在這裏工作嘛？, closed.question
- U111 對，我是個律師, respond.agree
- U112 哦~ 什麼律師？, open.question
- U113 international law , respond.agree
- U114 哇，那日內瓦一定很好, give.opinion
- U115 還好啦，respond.agree

U123 要怎麼說, "*international law is the diminishing point of law?*", open.question  
U116 國際法是法律的遞減點?, give.opinion (\*\*elicit.task.information ??)  
U117 不確定, 可是聽起來是對的, respond.agree  
U118 哈哈, relax.atmosphere  
U119 哈哈, relax.atmosphere  
U120 意思明白就好了, 國際法真的有一點是個diminishing point, respond.agree  
U121 不同國家的法律永遠, 很少, 會同意., give.opinion  
U122 對! 就是這個意思, respond.agree  
U123 你現在念碩士吧?, closed.question  
U124 對, 現在在寫thesis, respond.agree  
U125 你thesis題目是什麼?, open.question  
U126 我喜歡語言交換, 所以在看用電腦模範(mo2fan3)語言交換有關的,  
respond.agree  
U127 哇, 聽起來很複雜, give.opinion  
U128 哈哈, 有一點, respond.agree  
U129 但是我覺得是應為電腦係裏現在做跟語言有關係的都是在做機器翻譯, 很少研究用  
機器幫學習語言的方法, give.opinion  
U130 欸, 那你這種研究是哪裏不一樣?, open.question  
U131 有點難解釋, 因為我也還沒有完全想通細節, 可是現在電腦係裏很火的是機器學習,  
然後機器學習係裏跟語言有關的都大大是跟自動翻譯有關的, 像把語言變成數字去解決這  
個問題, 我比較喜歡學語言的理論(li3lun4), 所以想研究一些比較把語言留著的分析,  
respond.agree  
U132 啊, 那這種研究的數學會比較複雜嗎? 跟機器學習比, open.question  
U133 mmm, 對我來說複雜一點, 可是它們兩個用的數學不太一樣, 機器學習用  
optimization跟derivatives, 這種語言交換模範的數學比較像theoretical mathmatics,  
很多跟graph theory有關的, respond.agree  
U134 但是其實我覺得我的Theoretical math(純粹chun2cui4數學)比我的applied math(  
應用ying4yong4數學)差一點, give.opinion

## Trace 2 - Meeting Kevin

U200 欸, hello, use.social.convention  
U201 hey, use.social.convention

U202 用中文沒問題吧？, closed.question  
U203 沒問題, respond.agree  
U204 你跟你的前一個partner都是用中文嗎？, closed.question  
U205 差不多, 有時候會換英文, respond.agree  
U206 那我們也盡量用中文, give.opinion  
U207 yea, 挺好的, respond.agree  
U208 那你現在都在做些什麼？, open.question  
U209 想在在日内瓦念masters, 寫thesis, respond.agree  
U210 你在加州念書吧？, closed.question  
U211 第幾年了？, closed.question  
U212 對, 我在USC第四年了, respond.agree  
U213 啊, 那快畢業了, 現在一定很亂, give.opinion  
U214 yea, 很累, respond.agree  
U215 其實畢業以後想去紐約找工作, give.opinion  
U216 哦, 紐約不錯, 你是做finance, right?, closed.question  
U217 yea, respond.agree  
U218 紐約的finance很好。但是現在哪裏找工作都很不容易, give.opinion  
U219 你可以再念一年書嗎？還是沒辦法？, closed.question  
U220 可以, 我也是想說如果找工作不容易可以再念個CS minor, respond.agree  
U221 哦, finance念CS minor很好, 一定會用到。respond.agree \*\* (or is it give.opinion here?)  
U222 嗯, 所以還有很多選擇, give.opinion  
U223 啊, 挺好的, respond.agree

### **Draw and Correlate**

So the next step here is to provide a graphic for both traces of conversations including their proper labeling, and then trying to correlate between the two traces to see if we can define a "sub-conversation."

Additionally, determine what direction we can take these sub-directions in and use them to what end.

I guess from the perspective of doing the correlating and sub-conversation definition, we should start by outlining instead of jumping straight to just doing



preliminary algorithms. Might also be worth first looking into the “inductive miners” algorithm etc that’s in the literature to see what we’ve got as groundwork. Additionally, I do think we need to substantially increase the length of the conversation trace to get a good sense of the normal shifts (since our topics that we immediately remember are all introduction based, and that will add a bias).

We should also keep in mind that there might be other metrics that we want to keep in mind when we’re trying to create a sub-conversation (i.e: the length of an utterance based on word count, the length of an utterance based on sentence count, etc) I think these will assist us in defining sub-conversations since they provide some indirect semantic information about an utterance (i.e: sometimes a long utterance with long word count and sentence count with a respond.agree after an open.question tag can imply an “explanation” while a respond.agree with a short word count after an open.question could be closer to a deflection than an explanation. Or something like, long respond.agree with a long give.opinion from the same speaker afterwards means preempting by giving a clarification or expanding, and something like different speaker give.opinion might be displaying active listening via confirming information.

We have some samples of possible sub-conversations from trace 1, they’re based on simple rules of the labels

CLARIFICATION: respond.agree -> long utterance with give.opinion from the same speaker

((‘意思明白就好了，國際法真的有一點是個diminishing point’, ‘respond.agree’, ‘1’), (‘不同國家的法律永遠，很少，會同意。’, ‘give.opinion’, ‘1’))

((‘哈哈，有一點’, ‘respond.agree’, ‘1’), (‘但是我覺得是應為電腦係裏現在做跟語言有關係的都是在做機器翻譯，很少研究用機器幫學習語言的方法’, ‘give.opinion’, ‘1’))

((‘mmm, 對我來說複雜一點，可是它們兩個用的數學不太一樣，機器學習用optimization跟derivatives，這種語言交換模範的數學比較像theoretical mathmatics，很多跟graph theory有關的’, ‘respond.agree’, ‘1’), (‘但是其實我覺得我的Theoretical math(純粹chun2cui4數學)比我的applied math(應用ying4yong4數學)差一點’, ‘give.opinion’, ‘1’))

ACTIVE.LISTENING: respond.agree -> long utterance with give.opinion from different speaker

(( '美國真的是，不知道他們在幹什麼。。。', 'respond.agree', '2'), ('好，這裏念完書回台灣', 'give.opinion', '1'))

(( 'international law', 'respond.agree', '2'), ('哇，那日内瓦一定很好', 'give.opinion', '1'))

(( '我喜歡語言交換，所以在看用電腦模範語言交換有關的', 'respond.agree', '1'), ('哇，聽起來很複雜', 'give.opinion', '2'))