

# DEVELOPING A CHATBOT BASED ON OpenNLP

**Boris Galitsky**

OpenNLP / Oracle Mobile Cloud

[boris.galitsky@oracle.com](mailto:boris.galitsky@oracle.com)  
[@bgalitsky](https://twitter.com/bgalitsky)

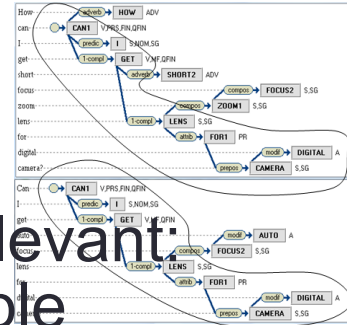
APACHECON North America

Sept. 24-27, 2018



# Which techniques are most critical for the current chatbot state-of-the-art?

- Making personalized, lower volume search more relevant: we cannot just learn from which search results people choose, like web search engines do => We need to get closer to 'understanding' and apply syntactic-level analysis
- We need to systematically learn how intelligent people organize their thoughts: we cannot just learn from available conversations => We need to rely on **discourse-level analysis** in dialogue management



... Not like some customer support agents

# The plan

- Lessons learned from chatbot failures
- How to build a chatbot you can demo
- Apache OpenNLP.Similarity and OpenNLP.Chatbot
- Search Engineering for chatbots: informative and uninformative parts of answers
- Content management for chatbots
- Dialogue management: rhetorical agreement
- Dialogue management: building a dialogue from initial query
- Chatbots tackling argumentation

# Issues with currently available chatbots

# Issues with currently available chatbots

There is a lot of buzz about chatbots but it is hard to find an actual demo

<https://watson-assistant-demo.ng.diuemix.net/> ▼

Quickly build and deploy **chatbots** and virtual agents across a variety of channels, ... In this **demo**, imagine you're in the driver's seat and Watson is your co-pilot.

Visual Chatbot ✓

[demo.visualdialog.org/](https://demo.visualdialog.org/) ▼

Visual **ChatBot**: Lets talk to bot! ... Hi, I am a Visual **Chatbot**, capable of answering a sequence of questions about images. Please upload an image and fire away ...

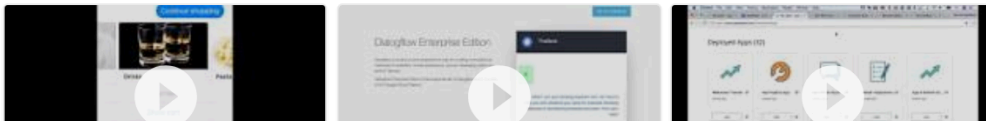
ChatBot and LiveChat Demo | WIX App Market | Wix.com ✓ McAfee SECURE

<https://www.wix.com/app-market/chatbot-and-livechat/demo> ▼

Add a **Chatbot** to your Wix site. Help visitors with instant answers. Generate sales leads. **ChatBot** and LiveChat. Multiple Languages ...

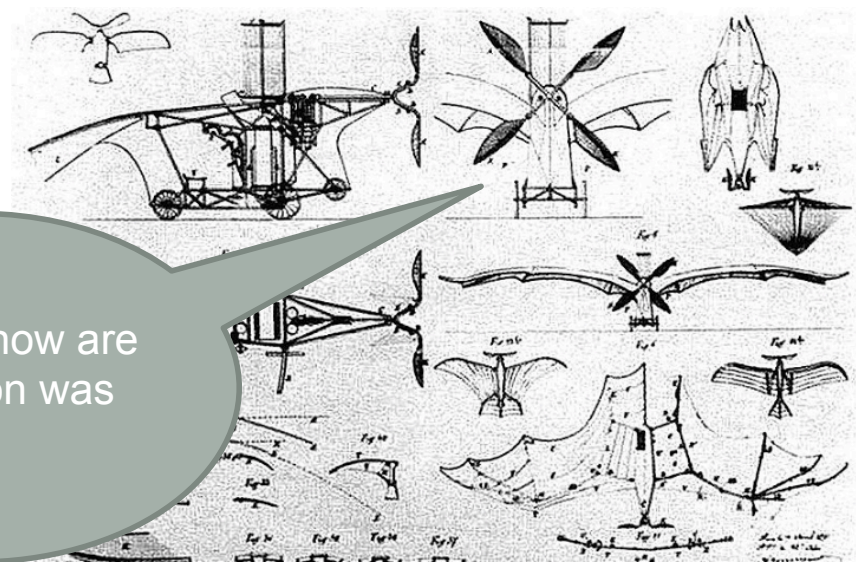
... if you have your favorite chatbot please let me know

## Videos

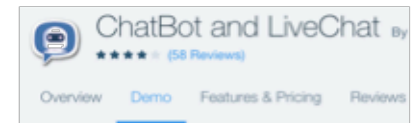
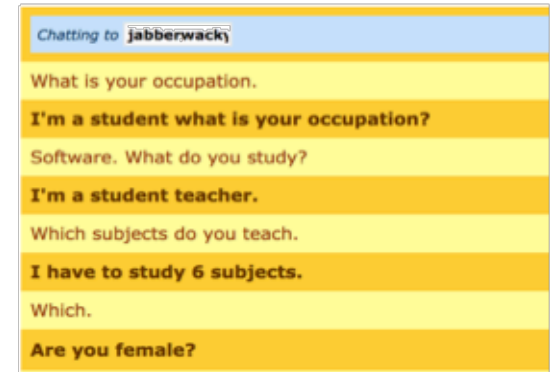
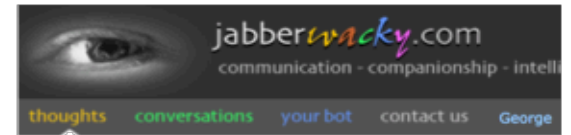


... they are not really available. Vendors are shy to demo their chatbots so show videos instead

Chatbots now are like aviation was in 1900s



Most chatbots available cannot answer basic questions



I am Chomsky, your chatbot friend. Ask me a question!

what is population of Brazil

Powered by Facebook and Twitter

Ask Chomsky

You:

yes

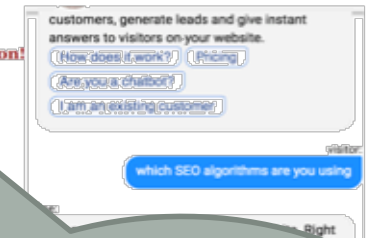
Chomsky:

Brazil is a country in South America.

Neuralconvo  
Chatting with a Deep learning brain

a demo of chatting with a Deep learning bot trained through Neuralconvo, a Torch library that implements Sequence to Sequence learning with Neural Networks (seq2seq).

I was thinking tonight you were in a hotel room while I go



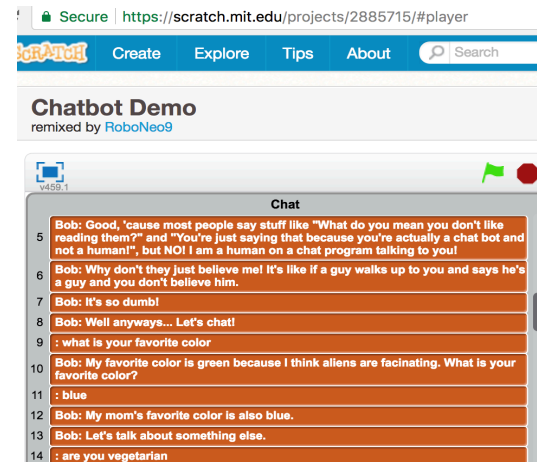
...once they are formulated in a slightly different form (syntactic, semantic) to what was trained / specified

It could not answer questions about tire pressure of a navigator. It does not differentiate between 'turning on' and 'turning off'

# 100 Web Search Results for Online Chatbot Demo

- 5% bots can support some form of conversation so that information cannot be acquired, but there is a feeling of cohesiveness (such as Mitsuku).
- 5% can answer questions like name, color, basic food and occupation but not really support conversation. They avoid understanding more complex questions by asking back
- 15% of deep learning-based bots – hard to find any correlation between questions and answers
- 20% is a movie, not actual demo
- The rest 55% are just websites talking about chatbots

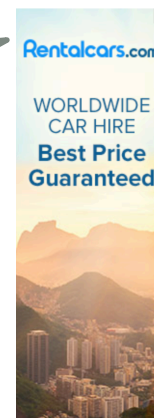
It is not always clear if it is a social bot to keep a user chatting or a task-oriented bot to help this user with something



# And those available ...

- Do not really support conversations
- Answer simple questions like a rudimentary search engine
- Most major vendors such as Microsoft Bot Framework, IBM Bluemix and Google's DialogueFlow do not demo chatbots built by their own teams or built by their customers

The only chatbot capable of supporting some form of conversation (not necessarily meaningful but somewhat concise) was Mitsuku



Type your message to Mitsuku  
 enter

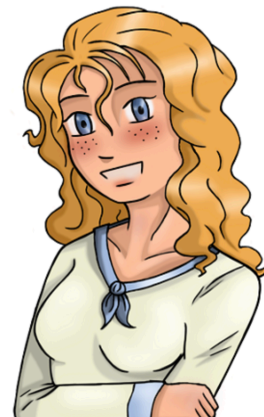
You - who else  
Mitsuku - Mike.

You - who do you see in parks  
Mitsuku - Mousebreaker.

You - what do you do in parks  
Mitsuku - Hmm.. a park can be used to walk in  
I use it to do that.

You - which parks do you go to  
Mitsuku - The big one, I think.

You - what is your favorite outdoor activity  
Mitsuku - Probably walking in the park.





# Example of an unsuccessful attempt at building a virtual assistant

As of March 2018, this is due, in particular, to a lack of dialogue management and also a lack of ontology support for the entities involved.

## Customer service

Do you have questions?

The Virtual Assistant<sup>1</sup> has answers. Ask about passwords, beneficiaries, and other account-related topics.

**Welcome.**  
**How can we help you?**

**Fidelity:** Hi, I'm Fidelity's automated Virtual Assistant. How can I help? Type your question in the box and select Ask.

**You:** rsu distribution schedule

**Fidelity:** You might find this helpful...

- [I need to transfer money](#)

Type your question here

Ask

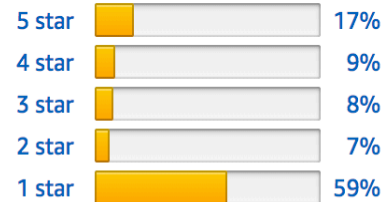
# Customer perception of Alexa Prize Socialbot

Alexa Prize Socialbots › [Customer reviews](#)

## Customer reviews

★★★★☆ 89

2.2 out of 5 stars ▾



## Alexa Prize Socialbots

by Amazon

Write a review

### Top positive review

[See all 23 positive reviews ›](#)

20 people found this helpful

★★★★☆ **Well Done In Higher Education and Technology Advancement**

By Lee Almodovar on October 12, 2017

First of all, the amount of whining from the one-star people. Cut it out. It's a skill that was tied to an Amazon Prize thing that puts out a competition with higher educational institutions to further the research and advancement of human-machine interaction. Without that precedence being set by Amazon and others, we wouldn't have Alexa (or Google's voice assistant, or Microsoft's Cortana), and

### Top critical review

[See all 66 critical reviews ›](#)

12 people found this helpful

★★★☆☆ **I'm not sure how to rate this**

By Amazon Customer on May 19, 2017

The idea behind this is really good - help develop socialbots. The bots themselves are really terrible. I got one that randomly said, "I'm sorry Angela." That's not my name... and I never told it my name... and there was no reason for it to apologize. I got another bot that held... let's just say some uncouth political opinions that it randomly spouted off when it was supposed to be talking about hockey. One said the word YOU about 20 times in a row. I've seen way, way

# How to build a chatbot you can demo

# How to build a chatbot NOT in 10 minutes and NOT relying on an enterprise platform

- Chatbot needs to **differentiate between a user asking it to do something** versus answer a question.
- For a chatbot in a given domain, all **basic methods and components** need to be thoroughly considered and rejected only if there is sufficient evidence and confidence.
- A chatbot needs to **assess relevance** between a user question and an answer to be provided in a meaningful, interpretable way, matching some formal representations of these instead of just computing a similarity in numerical space.
- One needs a systematic, interpretable way to build dialogues. Sometimes, even a detailed initial utterance can give a hint on how to build a dialogue structure, relying on **discourse analysis**



If you rely on an enterprise chatbot dev. platform, the result would not be good for enterprise...

# How to build a chatbot you can demo

- Dialogue management is hard, and conducting a meaningful conversation leading to a task completion is **very hard** in an open domain setting. Therefore at least maintaining a pair of consecutive chatbot utterances with **coreference** will impress the user. A simple sequence of user questions '*which option to chose*' followed by 'how long will it take' should make the chatbot link *option* and *it* to properly answer the second question.
- Frequently, simple dialogue management strategy such as **clarification requests** and pre-built navigation patterns suffice.
- A chatbot needs to **automatically learn** stuff from various available sources such as the web for all kinds of tasks, including entity extraction, understanding what kind of entity it is and identifying **relationships between entities**.
- A **content** for chatbot needs to be well-prepared and organized



If you just learn from customer support logs where banks misrepresent their non-sufficient fund rules ...

# How NOT to build a chatbot you can demo

Just throw tons of dialogue data at a machine learning system and believe that it **will figure all the above items automatically** for you and somehow start chatting

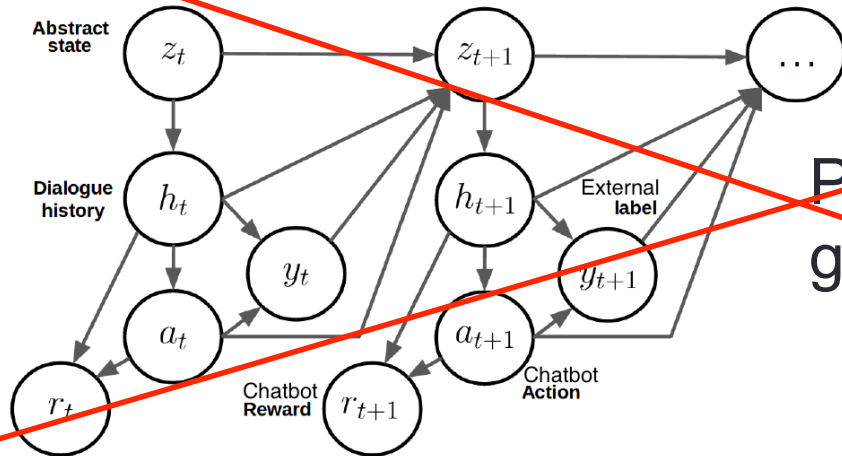


No easy way...  
step-by-state  
feature engineering  
and architectural  
design is required.

**This is not an  
image classification  
problem**

Chatbots work with linguistic structures **interpretable by humans** so you cannot just convert linguistic features into numbers and expect the system to acquire intelligence

# How NOT to build dialogue management for demonstrable chatbot



Probabilistic directed graphical model

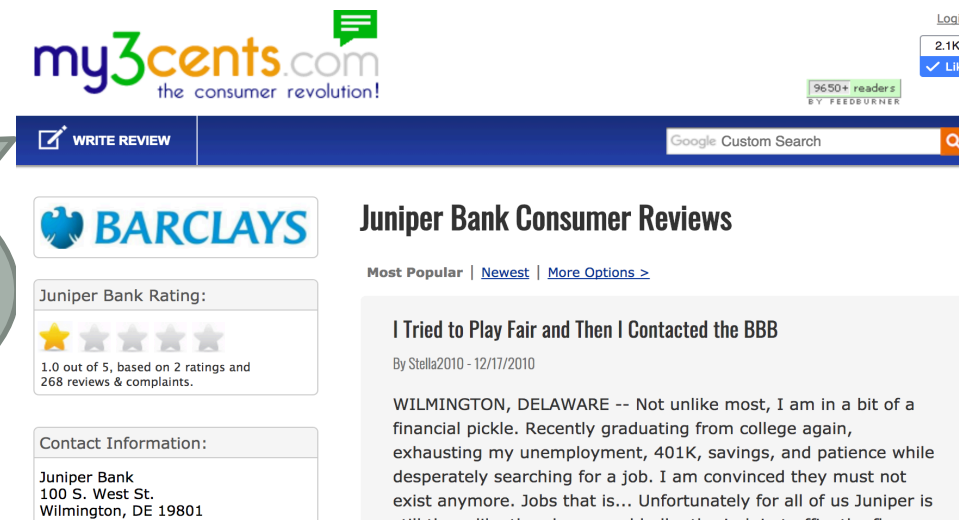
Instead, form **deterministic structures** to guide desired dialogues

Unless we build a garage door opener chatbot

# Communicating difficult topics

- In personal finance, there is a **systematic misunderstanding** how banks charge insufficient fund fees and modify interest rates
- Even **human customer support agents** have difficulties communicating these issues
- We collect a dataset of customer complaints in banking and **learn discourse structures** associated with heated arguments
- Then we **enable our chatbot** to operate with these discourse structures

As a result of conversation,  
we want customer to  
understand how banks can  
reorder transactions to charge  
NSF



The screenshot shows the my3cents.com website, which is described as "the consumer revolution!". The page features a navigation bar with a "WRITE REVIEW" button and a "Google Custom Search" box. The main content area displays the Barclays logo and a section for "Juniper Bank Consumer Reviews". The Juniper Bank Rating is shown as 1.0 out of 5, based on 2 ratings and 268 reviews & complaints. The contact information for Juniper Bank is listed as 100 S. West St., Wilmington, DE 19801. A review titled "I Tried to Play Fair and Then I Contacted the BBB" by Stella2010 is visible, dated 12/17/2010. The review text begins with "WILMINGTON, DELAWARE -- Not unlike most, I am in a bit of a financial pickle. Recently graduating from college again, exhausting my unemployment, 401K, savings, and patience while desperately searching for a job. I am convinced they must not exist anymore. Jobs that is... Unfortunately for all of us Juniper is...".



# OpenNLP.Similarity

# OpenNLP.Similarity

## What is OpenNLP.Similarity?

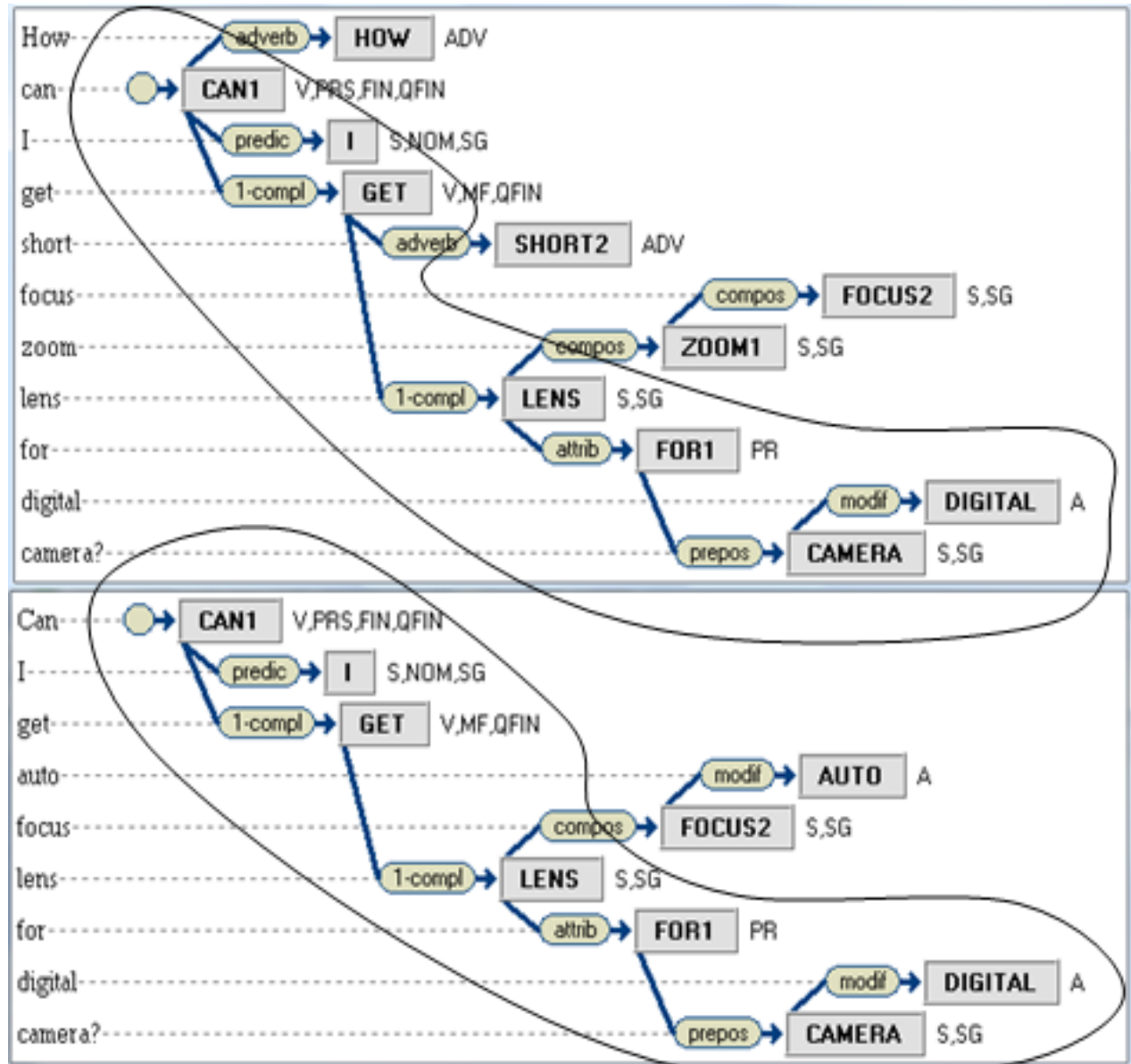
---

OpenNLP.Similarity is an NLP engine which solves a number of text processing and search tasks based on OpenNLP and Stanford NLP parsers. It is designed to be used by a non-linguist software engineer to build linguistically-enabled:

- search engines
- chat bots
- recommendation systems
- dialogue systems
- text analysis and semantic processing engines
- data-loss prevention system
- content & document generation tools
- text writing style, authenticity, sentiment, sensitivity to sharing recognizers
- general-purpose deterministic inductive learner equipped with abductive, deductive and analogical reasoning which also embraces concept learning and tree kernel learning.

OpenNLP similarity provides a series of techniques to support the overall content pipeline, from text collection to cleaning, classification, personalization and distribution. Technology and implementation of content pipeline developed at eBay is described [here](#).

Relevance between a question and an answer is measured via parse trees



# OpenNLP.Similarity.Chatbot

## Creating a simple project

---

Create a project from [MyMatcher.java](#).

## Running a chat bot

---

First you need to set the resource directory. The simplest way is to download and unzip it from [here](#). Then get a chat bot jar from [here](#). To run it (once the resource directory is set): `java -jar cbjar` It will take you to the prompt ">" to type your query. An example session is in [examples] (<https://github.com/bgalitsky/relevance-based-on-parse-trees/tree/master/examples/botSessionExample.txt>). The entry point for chat bot integration is [here](#)

## Engines and Systems of OpenNLP.Similarity

---

### Main relevance assessment function

It takes two texts and returns the cardinality of a maximum common subgraph representations of these texts. This measure is supposed to be much more accurate than keyword statistics, compositional semantic models word2vec because linguistic structure is taken into account, not just co-occurrences of keywords. [Matching class](#) in [matching package] ([https://github.com/bgalitsky/relevance-based-on-parse-trees/tree/master/src/main/java/opennlp/tools/parse\\_thicket/matching](https://github.com/bgalitsky/relevance-based-on-parse-trees/tree/master/src/main/java/opennlp/tools/parse_thicket/matching)) has

```
List<List<ParseTreeChunk>> assessRelevance(String para1, String para2)
```

function which returns the list of [common phrases between these paragraph]s.

# OpenNLP. Similarity processing flow and components

1. Define positive and negative classes of RR pairs:
  - a) Form the positive class from the rhetorically correct RR pairs
  - b) Form the negative class from the relevant but rhetorically foreign RR pairs
2. For each RR pair:
  - a) Parse each sentence  
Stanford NLP Parser, NER, Sentiment module of (Manning et al., 2014, Recasens et al., 2013, Lee et al 2013)
  - b) Obtain VerbNet structure for verbs  
VerbNet, JVerbNet (Kipper et al., 2008, <http://projects.csail.mit.edu/jverbnet/>).
  - c) Obtain coreferences  
*Stanford NLP Parser – Coreference*
  - d) Obtain entity - entity and entity – sub-entity links  
*OpenNLP.Similarity.parse\_thicket*
  - e) Build parse thicket pair for  $PT_{RR}$
  - f) Apply discourse parsing to obtain discourse tree pair  $DT_{RR}$  for RR pair
  - g) Align EDUs of  $DT_{RR}$  with  $PT_{RR}$
  - h) Merge aligned EDUs of  $DT_{RR}$  with  $PT_{RR}$   
*OpenNLP.Similarity.parse\_thicket*
  - i) Obtain  $DT_{RR}$  with VerbNet signatures for CAs
  - j) Obtain parse thicket with enriched RST relations  
*OpenNLP.Similarity.parse\_thicket.rhetic\_structure*
  - k) Build representation for Thicket Kernel learning
  - l) Build representation for Nearest Neighbor learning  
*OpenNLP.Similarity.parse\_thicket*
  - m) Improve text similarity assessment by word2vec model  
*Mikolov et al., 2011, <https://deeplearning4j.org/>*
3. Apply Thicket Kernel learning  
*OpenNLP.Similarity.parse\_thicket.kernel\_interface*  
*Moschitti 2006, <http://disi.unitn.it/moschitti/Tree-Kernel.htm>*
4. Apply Nearest Neighbor learning  
*OpenNLP.Similarity.jsmllearning*  
*OpenNLP.Similarity.parse\_thicket.matching*

# Search engineering for chatbots

Informative and uninformative parts of answers

# Discourse trees and parts of answers

## A good answer for:

*Which camera shoots well in low light*

*How to shoot in low light*

*Low light camera*

*Low light conditions*

## But not for:

*Good shorts at a boat*

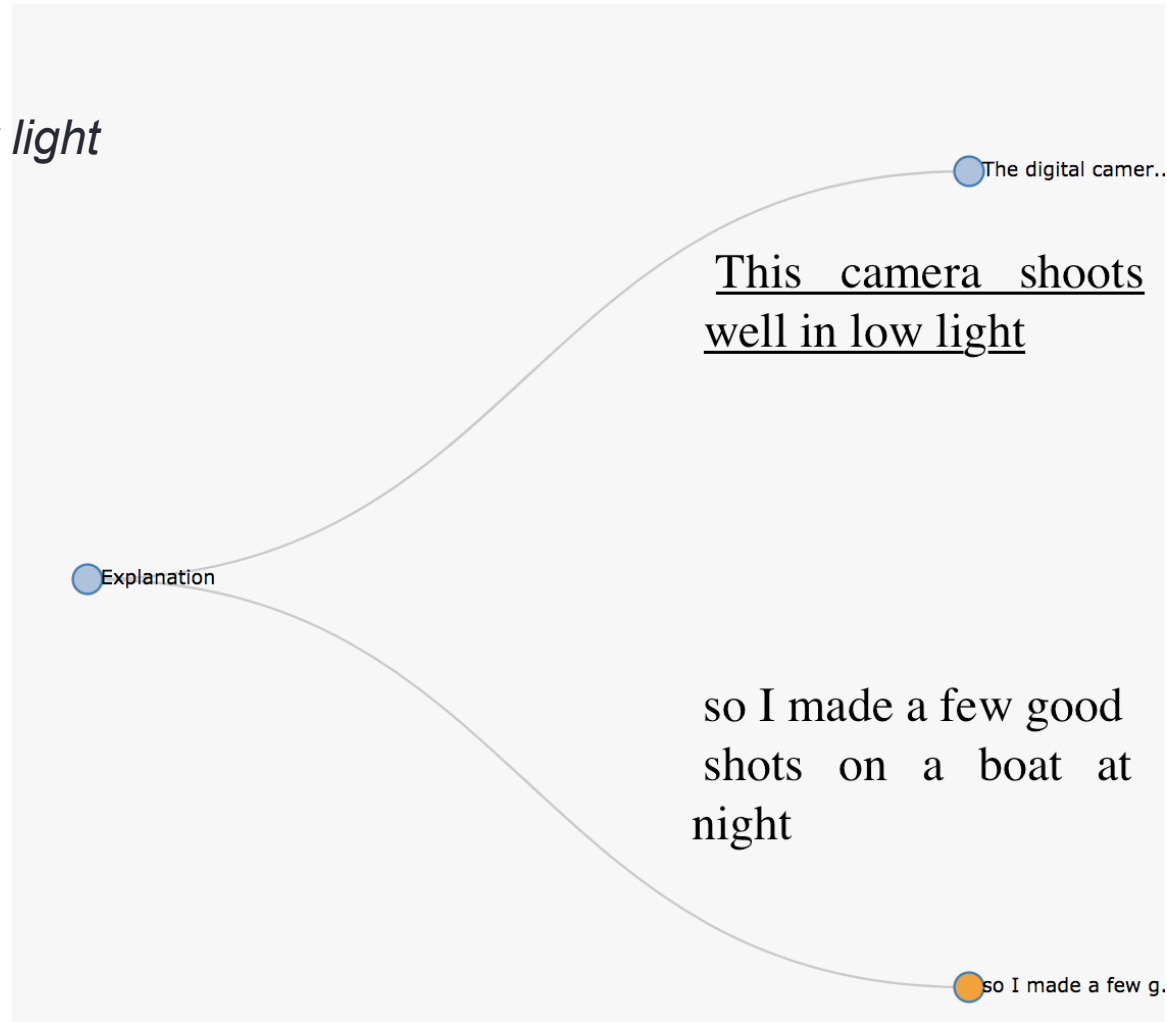
*Night boat*

*Boat at night*

*Good shots*

*Good boat*

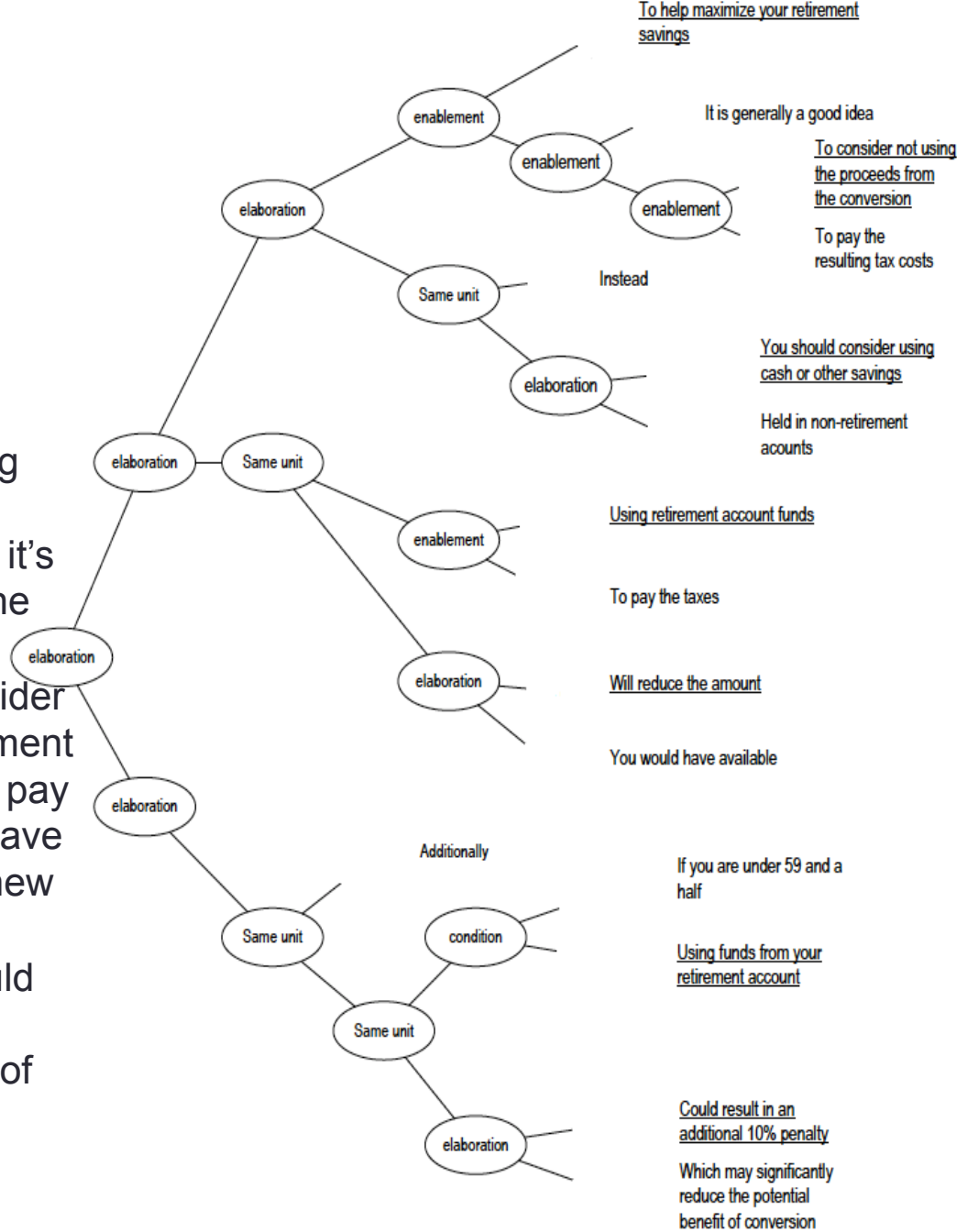
*Good night*



# Informative and uninformative parts of answer

Q: How should I plan to pay for taxes resulting from converting to a Roth IRA?

A: To help maximize your retirement savings, it's generally a good idea to consider not using the proceeds from the conversion to pay the resulting tax costs. Instead, you should consider using cash or other savings held in nonretirement accounts. Using retirement account funds to pay the taxes will reduce the amount you would have available to potentially grow tax-free in your new Roth IRA. Additionally, if you are under 59½, using funds from your retirement account could result in an additional 10% tax penalty, which may significantly reduce the potential benefit of conversion.





# Informative and uninformative phrases in answers (with respect to indexing)

phrases from the **nucleus** EDU:

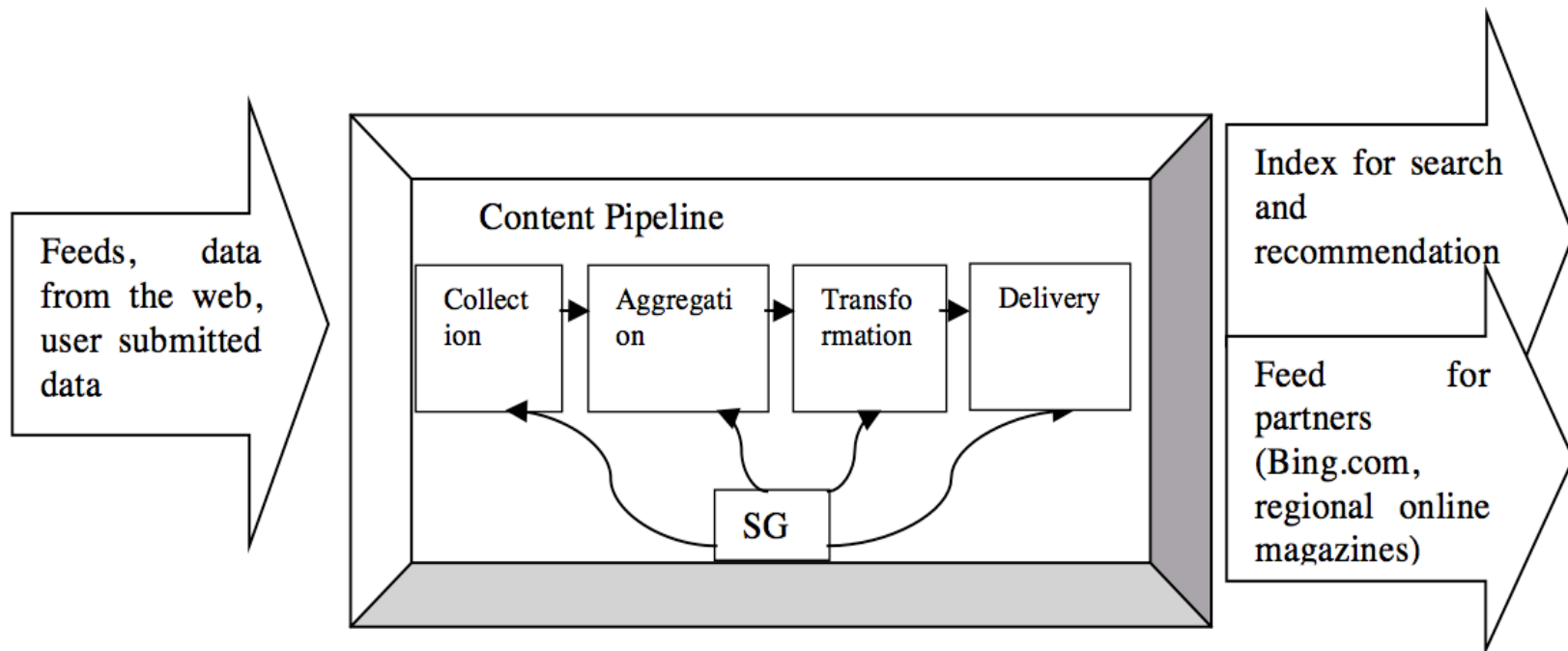
- *help maximize your retirement savings;*
- *proceeds from the conversion;*
- *cash or other savings held in nonretirement accounts;*
- *retirement account funds;*
- *using funds from your retirement account;*
- *result in an additional 10 % tax penalty.*

list of **satellite** EDU expressions:

- *it's generally a good idea* (not related to finance);
- *pay the resulting tax costs* (detached from the context);
- *held in nonretirement accounts* (detached from the context);
- *to pay the taxes will reduce the amount ...*(detached from the context);
- *you would have available to potentially ...* (counterfactual expressions, unlikely to occur in a user question);

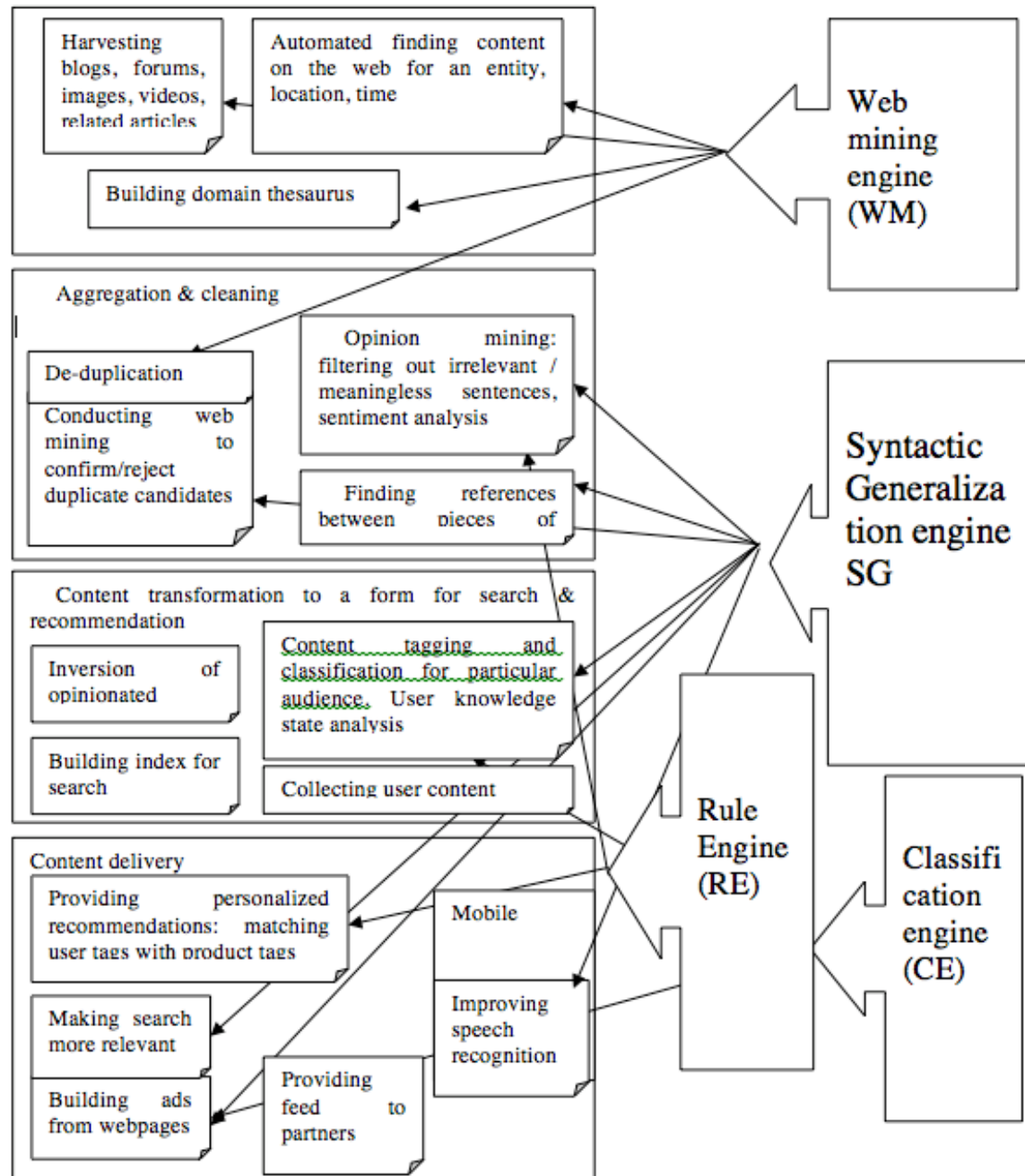
# Content Management System for chatbots

# Input, output and relevance support



- **Content collection** from multiple sources (automated finding content on the web relevant to given topic, feed processing);
- **Content aggregation**, cleaning, enrichment (de-duplication, cleaning, entity extraction, forming links between different pieces of content with the same entity, auto-generation or discovering of missing pieces of content);
- **Content transformation** to a form for search and recommendation (coordinating factual and opinionated content, building index for search and recommendation);
- **Content delivery** (search and recommendation for web and mobile, personalization, search engine marketing, sharing with partners).

# Detailed architecture of chatbot CMS



# Rhetorical agreement

# An answer must match the style of question

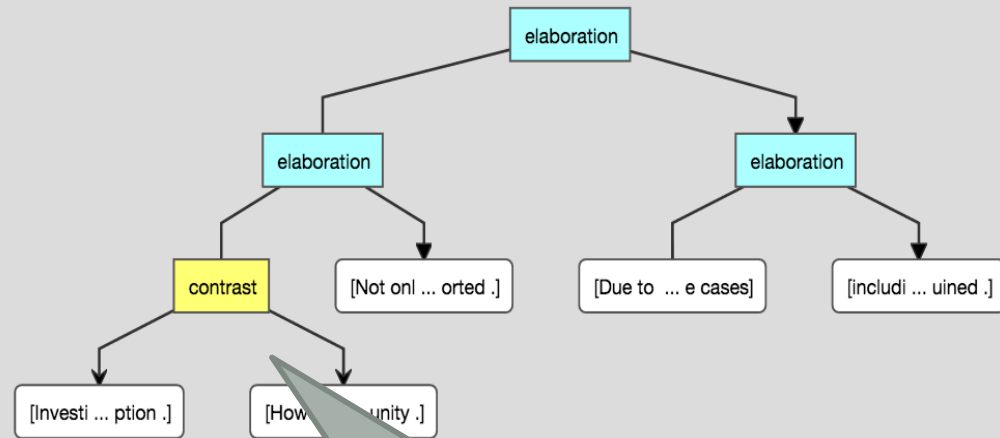
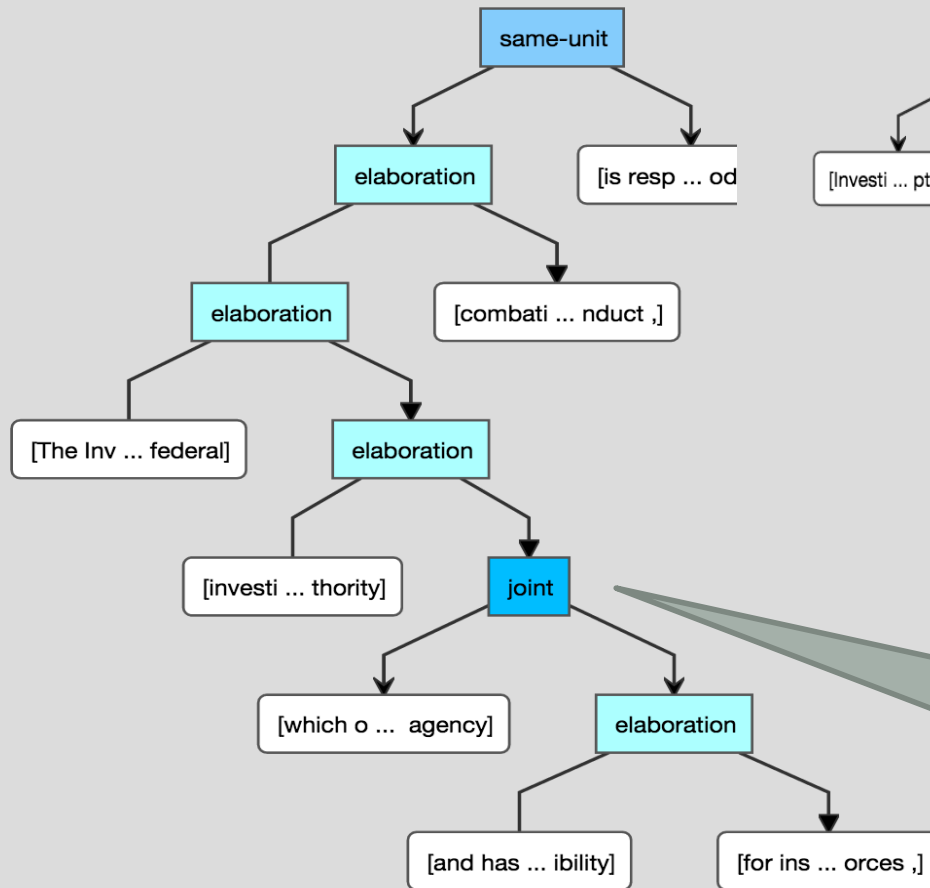
**Q: What does The Investigative Committee of the Russian Federation **do** / **really do****

**A1: Mission statement.** “The Investigative Committee of the Russian Federation is the main federal investigating authority which operates as Russia's Anti-corruption agency and has statutory responsibility for inspecting the police forces, combating police corruption and police misconduct, is responsible for conducting investigations into local authorities and federal governmental bodies.”

**A2: An answer from the web.** “Investigative Committee of the Russian Federation is supposed to fight corruption. However, top-rank officers of the Investigative Committee of the Russian Federation are charged with creation of a criminal community. Not only that, but their involvement in large bribes, money laundering, obstruction of justice, abuse of power, extortion, and racketeering has been reported. Due to the activities of these officers, dozens of high-profile cases including the ones against criminal lords had been ultimately ruined” (CrimeRussia 2016).



# Discourse trees for official and unofficial answers



If a sarcasm can be detected in a question, an answer needs to include contrasting parts

Official answer is just *joints*, *same-units* and *elaborations*

# Question, clarification and answer

My husbands' grandmother gave him his grandfathers truck. She signed the title over but due to my husband having unpaid fines on his license, he was not able to get the truck put in his name. I wanted to put in my name and paid the property tax and got insurance for the truck. By the time it came to sending off the title and getting the tag, I didn't have the money to do so. Now, due to circumstances, I am not going to be able to afford the truck. I went to the insurance place and was refused a refund. Since I am not going to have a tag on this truck, can I get the property tax refunded?

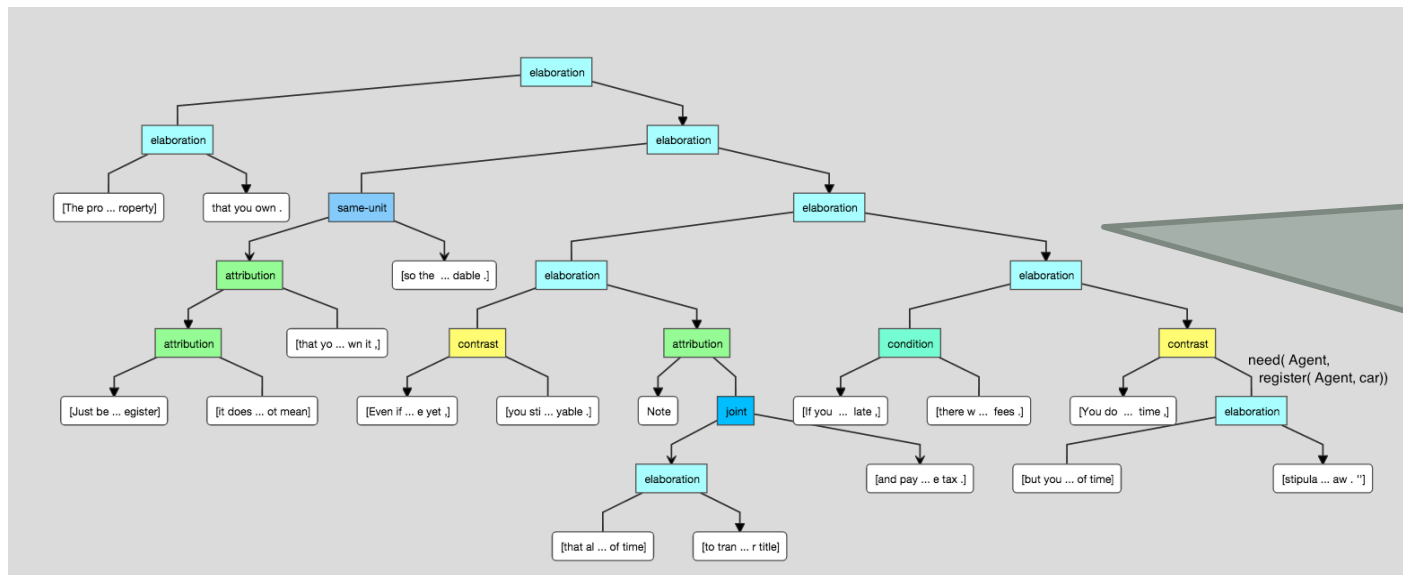
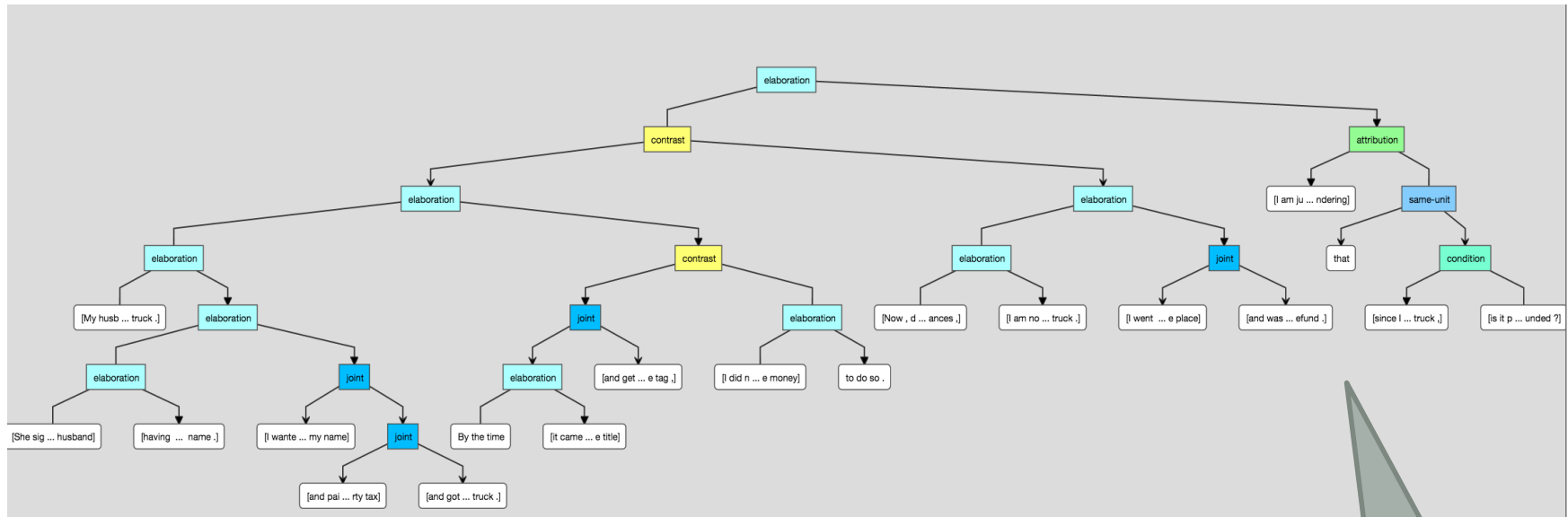
Are  talking about property tax refund, registration or insurance?

Insurance

The property tax is assessed on property that you own. Just because you chose to not register it does not mean that you don't own it, so the tax is not refundable. Even if you have not titled the vehicle yet, you still own it within the boundaries of the tax district, so the tax is payable. Note that all states give you a limited amount of time to transfer title and pay the use tax. If you apply late, there will be penalties on top of the normal taxes and fees. You don't need to register it at the same time, but you absolutely need to title it within the period of time stipulated in state law.

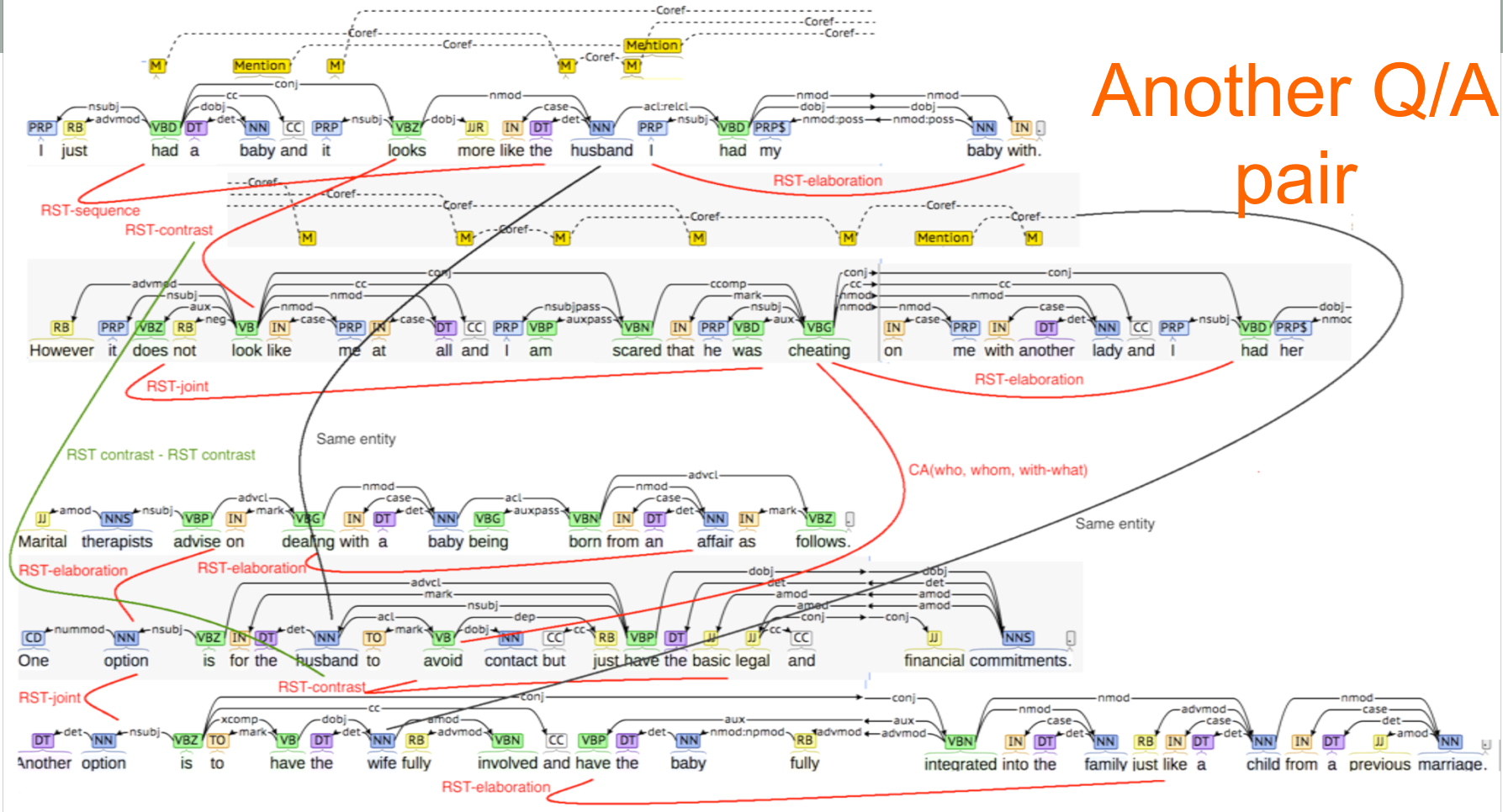


# Agreement between a Q and A



Once relevance is established, coordination of Discourse Trees of Q and A is important to select an answer with suitable style

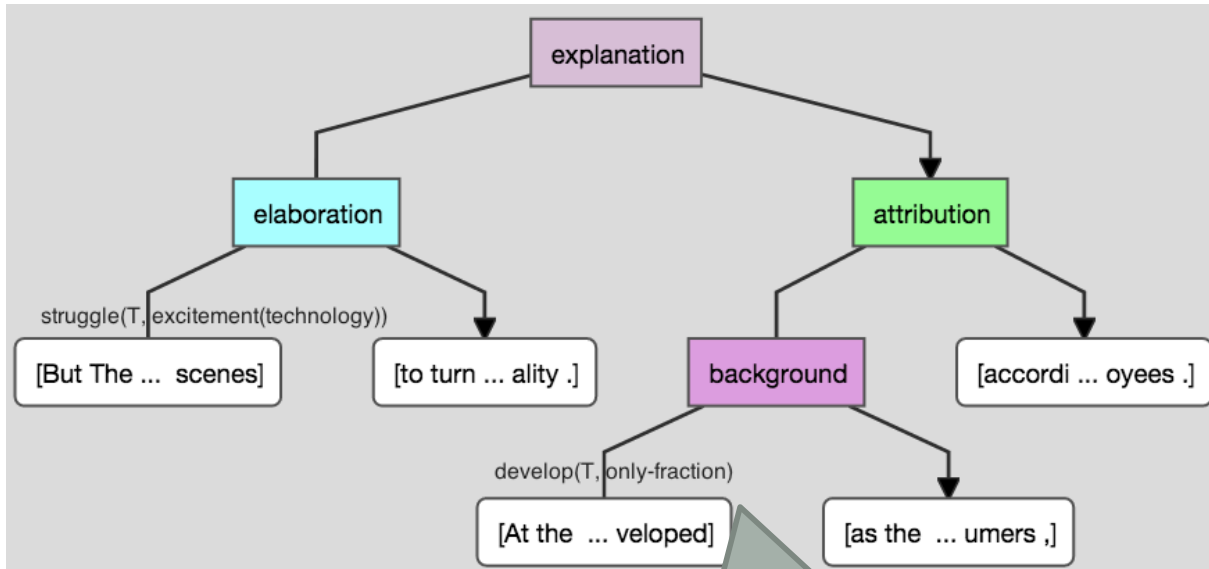
# Another Q/A pair



Q: I just had a baby and it looks more like the husband I had my baby with. However it does not look like me at all and I am scared that he was cheating on me with another lady and I had her kid. This child is the best thing that has ever happened to me and I cannot imagine giving my baby to the real mom.

A: Marital therapists advise on dealing with a child being born from an affair as follows. One option is for the husband to avoid contact but just have the basic legal and financial commitments. Another option is to have the wife fully involved and have the baby fully integrated into the family just like a child from a previous marriage.

# Enabling Discourse Trees with Speech Acts



*“...But Theranos has struggled behind the scenes to turn the excitement over its technology into reality. At the end of 2014, the lab instrument developed as the linchpin of its strategy handled just a small fraction of the tests then sold to consumers, according to four former employees.”*

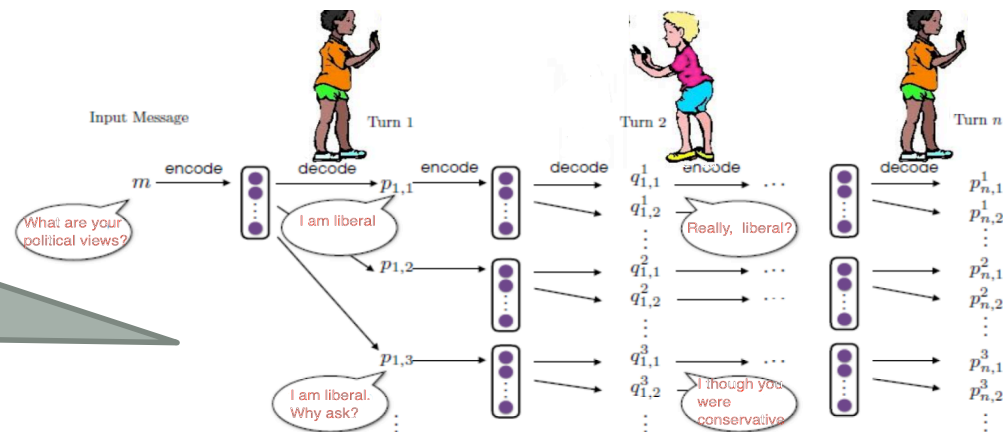
When arbitrary **communicative actions** are attached to a discourse tree as labels of its terminal arcs, it becomes clear that the author is trying to bring her point across and not merely sharing a fact.

# Building a dialogue from initial query

# Building dialogue structure from the first utterance for goal-oriented chatbots

- To build a dialogue scenario, we analyze the **discourse tree** of an initial query of a customer support dialogue that is frequently complex and multi-sentence.
- We then enforce **rhetorical agreement** between DT of the initial query and DTs the answers, requests and responses.
- The chat bot should find answers, which are not only relevant by topic but also **suitable for a given step of a conversation and match the question by style**, argumentation patterns, communication means, experience level and other domain-independent attributes.

Our approach is different from learning turn taking based on a huge conversational corpus



# Customer support dialogue



I cannot see myself on Google Earth, although I keep walking with my laptop outside of my house and waving. I even started to catch my own attention but it is not working. I am confident I am looking at my house since I typed my address.

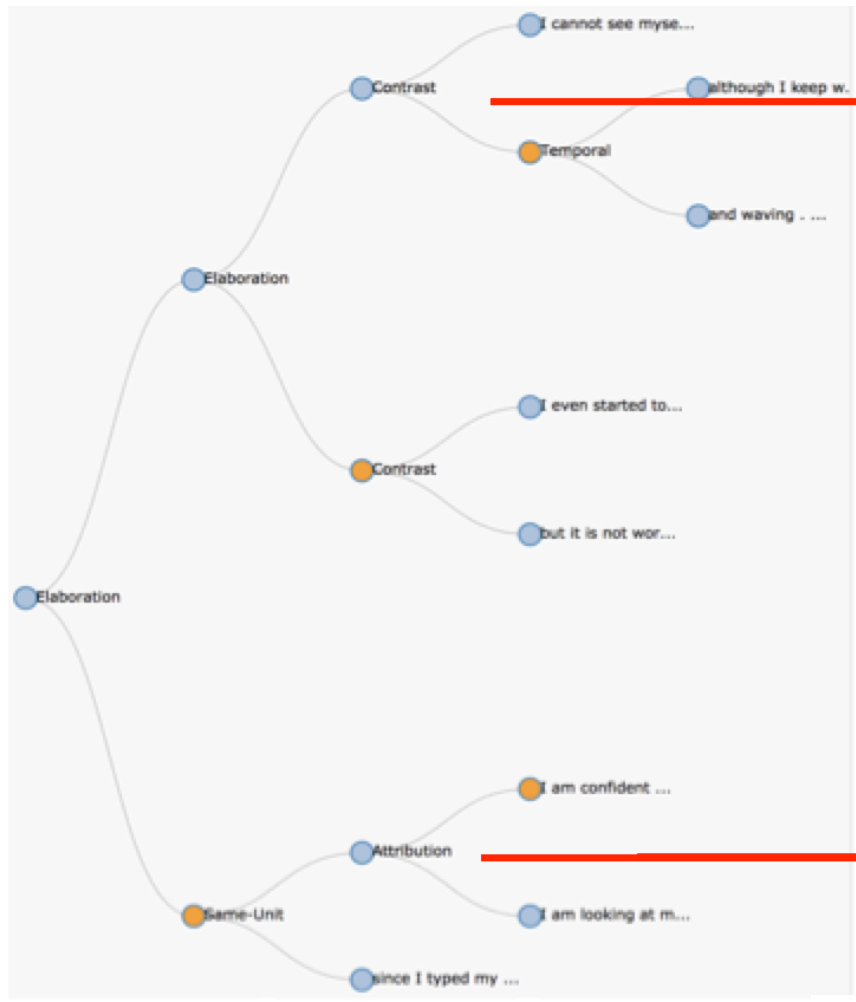
Your business is important to us. Please make sure you are online when you are walking, otherwise you would not be able to connect with Google Maps

Yes I am online, but still cannot see myself.

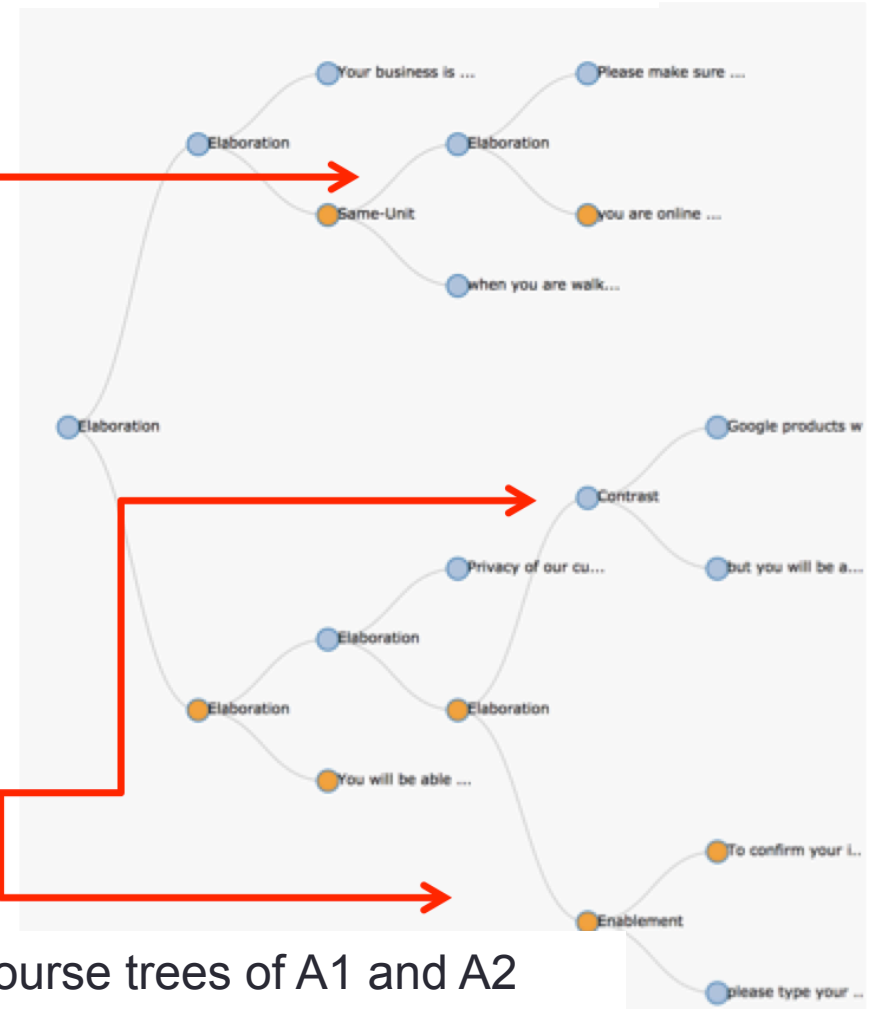
Privacy of our customers is very important to us. Google products will help you to catch your own attention, but you will be anonymized. To confirm your identity, please type your address. You will be able to see yourself in Google Plus

Thank you! I got it. I need to connect Google Maps and Google Plus

# DT-As complement DT-Q



Discourse tree of Q



Discourse trees of A1 and A2



# Algorithm of utterance search

- 1) Build  $DT-Q$ ;
- 2) Split  $DT-Q$  into parts  $Q_1, Q_2, \dots$  to correspond to  $A_1, A_2, \dots$ ;
- 3) Form search query for  $A_1$  from  $Q_1$  in the form *RST-relation [phrase]* ;
- 4) Run the search against the set of dialogue utterances and obtain the list of candidate answers for the first step  $A_1$  candidate;
- 5) Build  $DT-A_{1\text{candidate}}$  for each candidate and approve/reject each based on *complement(DT-Q, DT- $A_{1\text{candidate}}$ )*;
- 6) Respond to the user with the selected  $A_1$  and receive  $C_1$ ;
- 7) Form search query for  $A_2$  from  $Q_1 \& C_1$ ;
- 8) Repeat steps 4) and 5) for  $A_2$ , respond to the user with the selected  $A_2$  and receive  $C_2$ ;
- 9) Conclude the session or switch to a human agent



# Model of a customer support dialogue

We don't have to learn 1000s of conversation types. We just need to support well a limited number of dialogue modes to efficiently resolve problems

**$Q$ :** Formulate the problem and associated issues:

- Present how it happened
- Explain why it happened
- What had been tried
- Believe why unsuccessful

**$A_1$ :** Propose a first option for a solution and address certain associated issues from  $Q$ .  
- Request some clarification

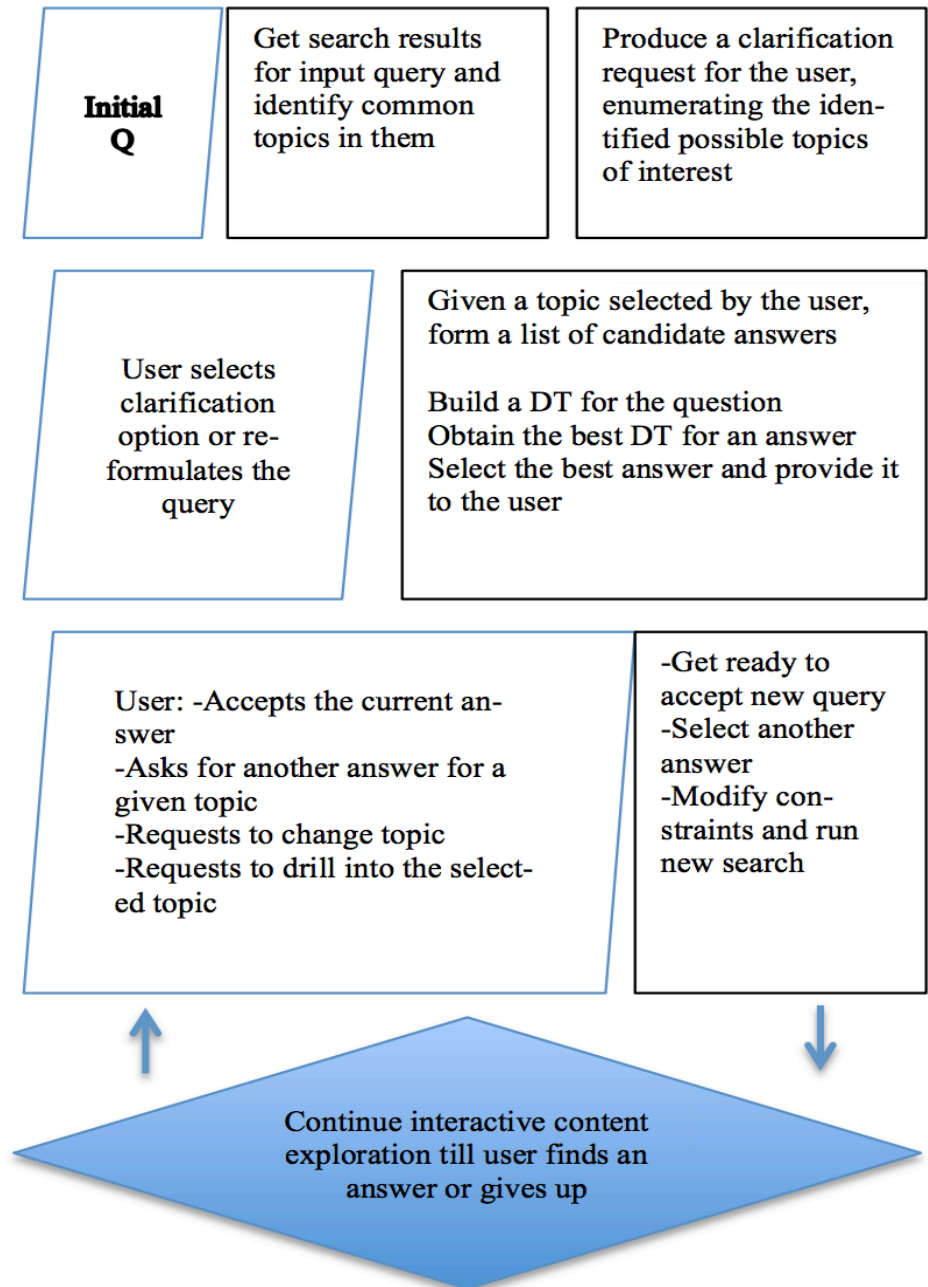
**$C_1$ :** Confirm that solution from  $A_1$  has been tried  
- Share results  
- Present more associated issues  
- Provide clarification

**$A_2$ :** Propose a second solution option based on clarification  $C_1$   
- Request more clarification

**$C_2$ :** Confirm that solution from  $A_2$  has been tried  
- Share results  
- Provide clarification

Conclude the session  
- Summarize outcome

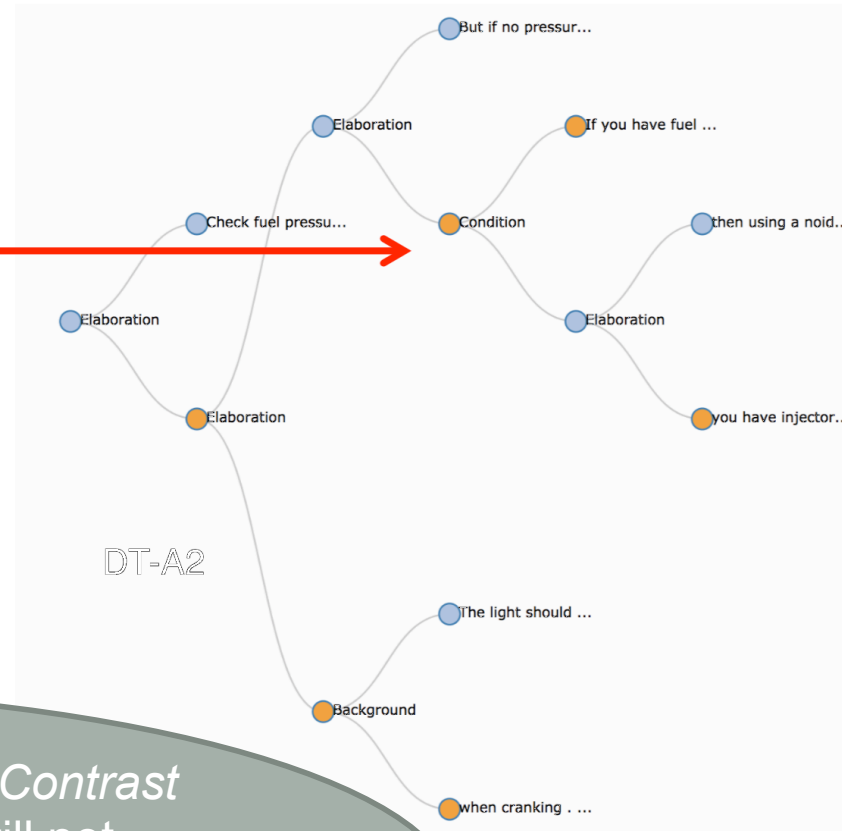
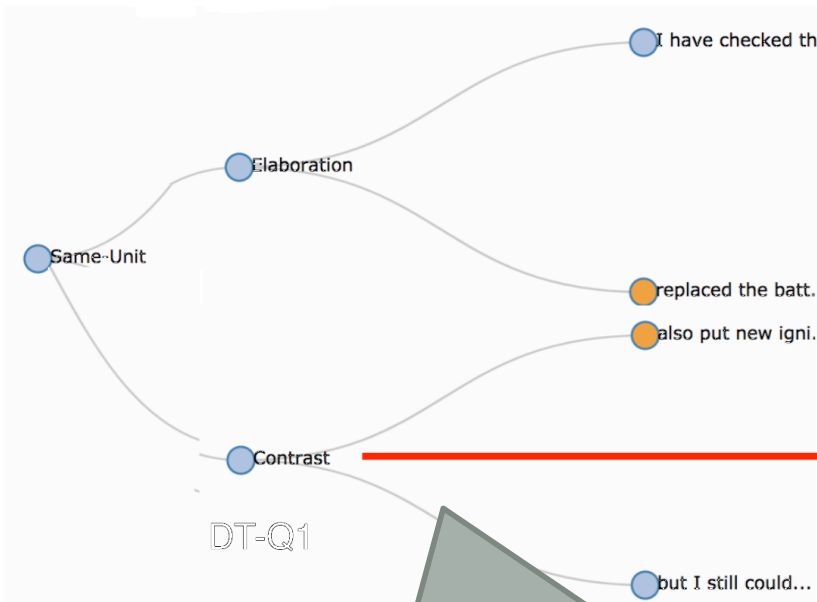
# Clarification mode of the dialogue manager



# DT-Q1 vs DT-A2: Car repair domain

- $Q_1$ : *I have checked the spark plugs, replaced the battery, also replaced ignition coils and the solenoid, but I still could not start my car.*
- $A_1$ : *Is the engine getting spark and fuel? Will it start for a couple seconds using starting fluid?*
- $Q_2$ : *It did at first but not anymore. What should I do?*
- $A_2$ : *Check fuel pressure. But If no pressure check the fuel pump for possible fault. If you have fuel pressure, then using a noid light make sure you have injector pulse. The light should blink when cranking. If light is not blinking does the check engine light come on with key on engine off? If no light then look for a bad main relay or no power to the engine control module.*

# DT-Q1 vs DT-A2

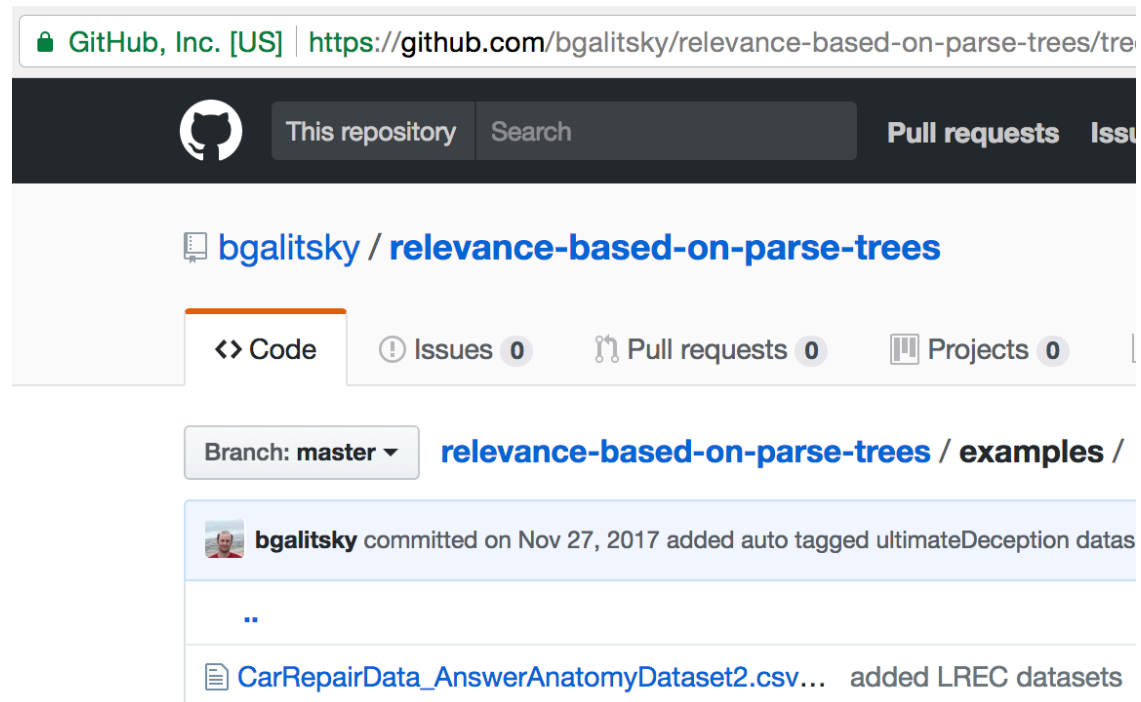


Since  $Q_1$  includes rhetoric relation of *Contrast* (something is done correctly but still not working), a detailed answer should include *Condition*: it is either done correctly or incorrectly, so both cases need to be addressed

# Data source for evaluation of dialogue construction

- Scraped from [www.2carpros.com](http://www.2carpros.com)  
Contains the real-life car problem diagnostics scenarios
- First utterance usually include multiple sentences, what has been attempted, what has been observed
- Reasonable grammar for syntactic and rhetoric parsing

We used customer complaints domain for discovering rhetoric structures



The screenshot shows a GitHub repository page for 'bgalitsky / relevance-based-on-parse-trees'. The page includes a navigation bar with 'This repository' and 'Search' buttons, and 'Pull requests' and 'Issues' links. Below the navigation bar, the repository name is displayed. There are tabs for 'Code', 'Issues 0', 'Pull requests 0', and 'Projects 0'. The current branch is 'master'. A commit history is shown, with the most recent commit by 'bgalitsky' on Nov 27, 2017, titled 'added auto tagged ultimateDeception datas'. Below this, a file named 'CarRepairData\_AnswerAnatomyDataset2.csv...' is listed with the description 'added LREC datasets'.

# Evaluation results

We considered the dialogue to be built *correctly* if an *actual* dialogue from the dataset is formed, given the first Q as a seed.

Otherwise, if the sequence of utterances does not occur in the dataset, we consider it to be *incorrect*.

Dialogue type	Q-A	Q-A1-C	Q-A1-C-A2	Q-A1-C1-A2-C2-A3
Baseline 1	62.3±4.5	60.2±5.6	58.2±5.0	52.5±5.7
Baseline 2	67.0±4.8	63.8±4.8	57.3±5.3	55.6±5.9
DT-Q dialogue formation	72.3±5.6	70.3±4.9	65.1±5.5	65.9±5.7

Limitations of our evaluation approach: since some actual dialogs are illogical and some synthetic dialogues built from distinct ones can be plausible, but it allows avoiding a manual assessment.

# Argumentation in Chatbots

# Argumentation in Chatbots

To answer questions about controversial topics chatbot needs:

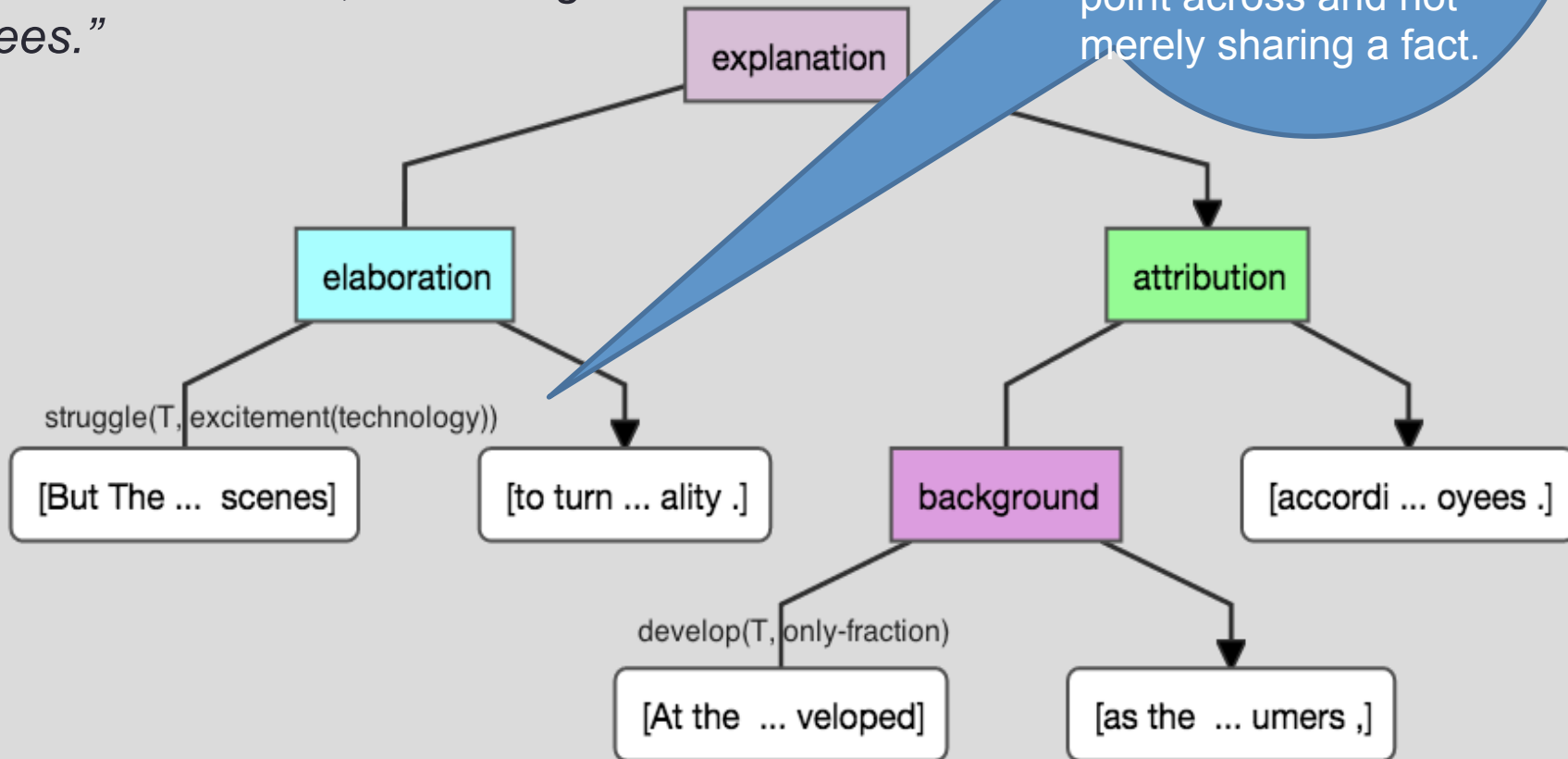
- to detect argumentation in user utterance
- to back up its claim with proper argumentation patterns
- recognize content with proper argumentation patterns



# A claim and its CDT

*“...But Theranos has **struggled** behind the scenes to turn the excitement over its technology into reality. At the end of 2014, the lab instrument developed as the linchpin of its strategy handled just a small fraction of the tests then sold to consumers, according to four former employees.”*

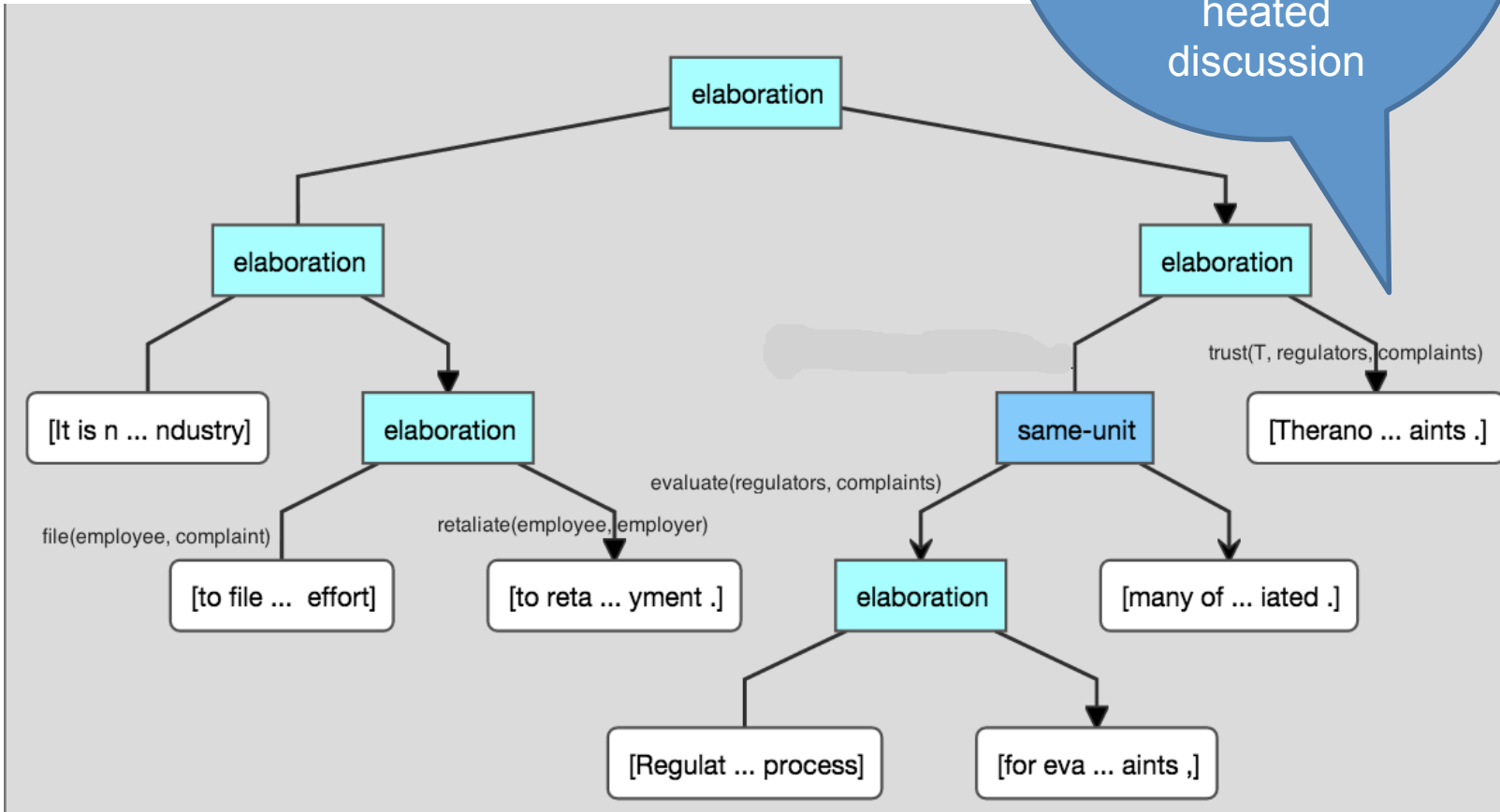
When arbitrary communicative actions are attached to DT as labels of its terminal arcs, it becomes clear that the author is trying to bring her point across and not merely sharing a fact.



# CDT for an attempt by Theranos to acquit itself

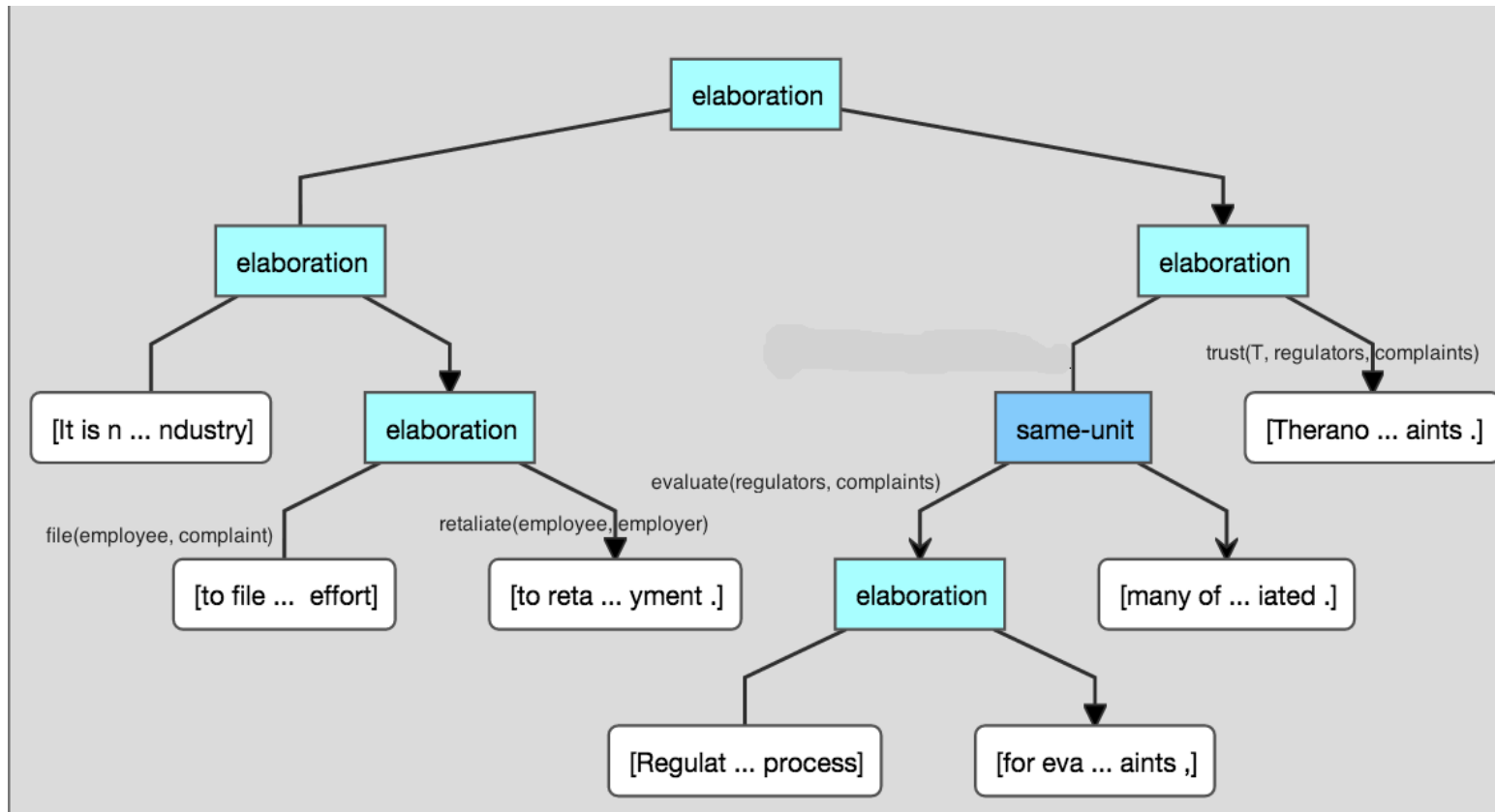
*“It is not unusual for disgruntled and terminated employees in the heavily regulated health care industry to **file** complaints in an effort to **retaliate** against employers for termination of employment. Regulatory agencies have a process for **evaluating** complaints, many of which are not substantiated. Theranos **trusts** its regulators to properly investigate any complaints.”*

Communicative actions as labels for rhetoric relations helps to identify a text which contains a heated discussion

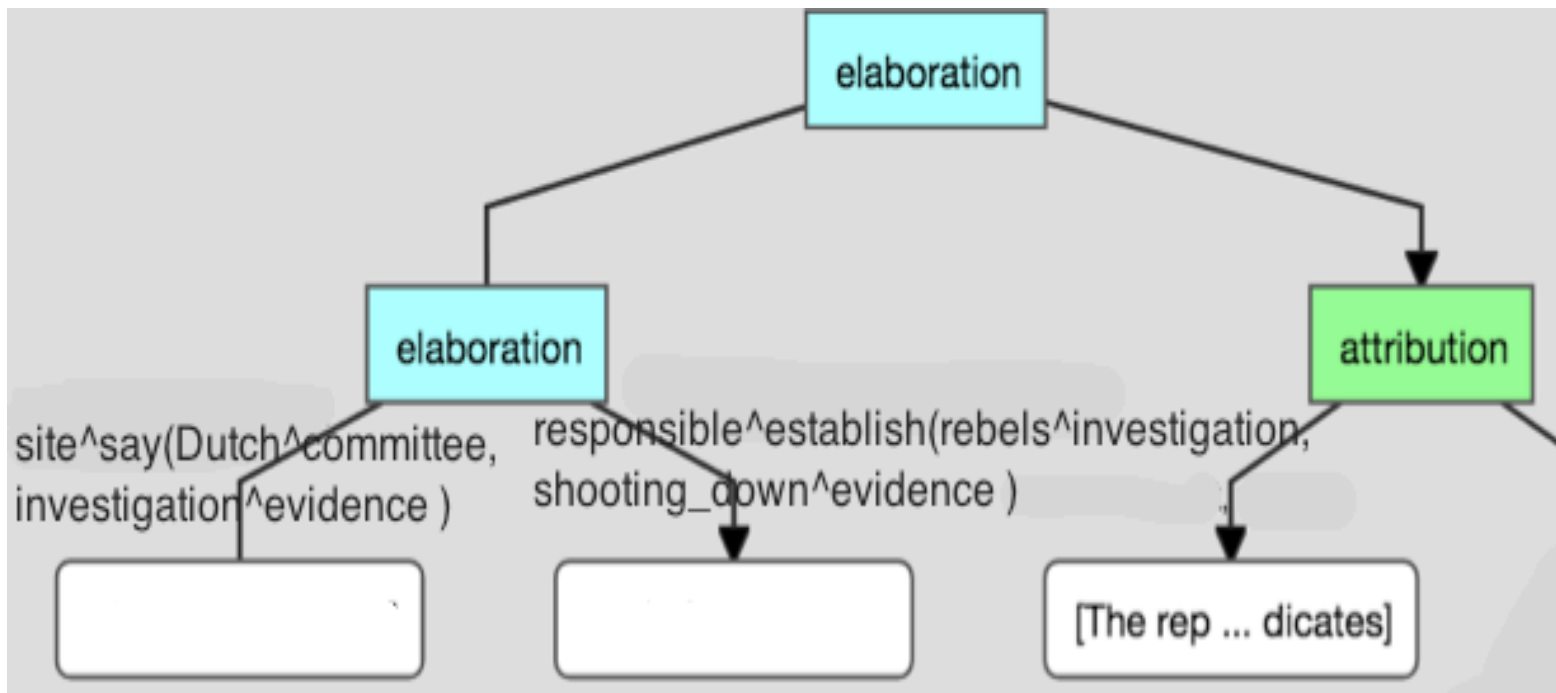


# Another CDT for an attempt to make even stronger counter-claim

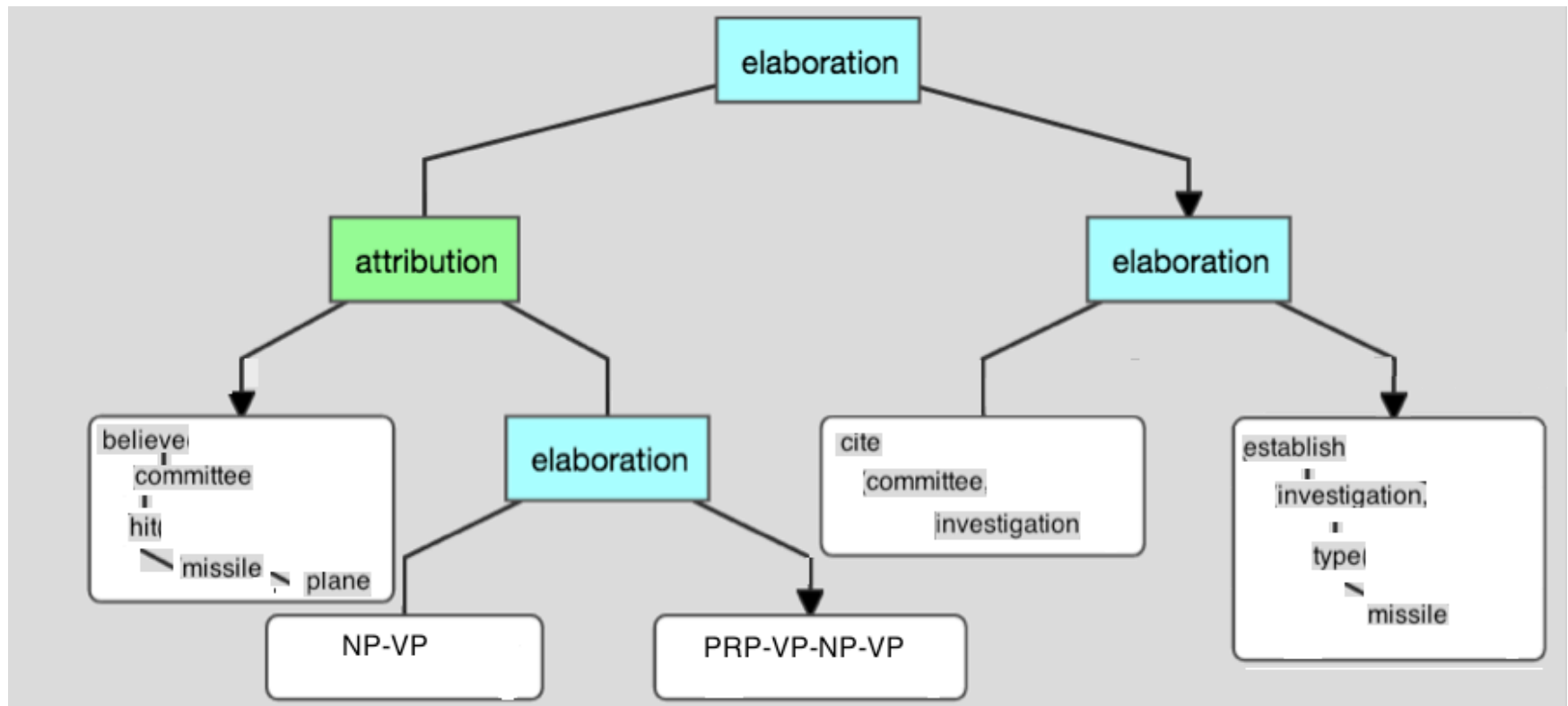
*“Theranos remains actively engaged with its regulators, including CMS and the FDA, and no one, including the Wall Street Journal, has provided Theranos a copy of the alleged complaints to those agencies. Because Theranos has not seen these alleged complaints, it has no basis on which to evaluate the purported complaints.”*



# Similarity between CDTs: maximal common sub-CDTs with generalized labels



# Numerical similarity between CDTs: representation for kernel model



# How to detect argumentation reliably

Method	Sources	P	R	F1
<b>Bag-of-words</b>	Keywords	57.2	53.1	55.07
<b>Naïve Bayes</b>	Keywords	59.4	55.0	57.12
<b>Graph-based kNN</b>	Reduced Communicative Discourse Tree (DT only)	65.6	60.4	62.89
<b>Graph-based kNN</b>	Reduced Communicative Discourse Tree (Communicative actions only)	62.3	59.5	60.87
<b>Graph-based kNN</b>	Full Communicative Discourse Tree	<b>83.1</b>	<b>75.8</b>	<b>79.28</b>
<b>SVM Tree Kernel</b>	Reduced Communicative Discourse Tree (DT only)	63.6	62.8	63.20
<b>SVM Tree Kernel</b>	Full Communicative Discourse Tree	<b>82.4</b>	<b>77.0</b>	<b>79.61</b>

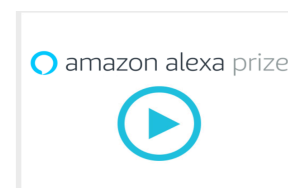
# Autistic and Schizophrenic chatbots

- A rule-based ontology-enabled chatbot displays **autistic** behavior: it answers adequately questions it "knows" and performs actions which "sound familiar" to it. Once an unknown domain or unanticipated kinds of conversation is encountered, a rule-based chatbot tends to refuse answering questions and is reluctant to interact in a new way. [task-oriented dialogues]
- Conversely, a deep/statistical learning-based chatbot instead displays a **schizophrenic** behavior. It energetically answers all kinds of questions, supports all kinds of conversations, displaying a disorganized thinking . The whole conversation usually does not make sense, it is spontaneous and irresponsible, with random topics. [non-task-oriented dialogues]

# Conclusions 1

- We looked at how chatbots provide **efficient access** to information which is hard to find via a search engine
- No “chit-chatting”. Resolve a customer problem by **finding correct solutions** if possible, not just providing relevant info.
- Trial and error are OK
- Systematically explore how dialogues can be **coherent in terms of their logical structure** (expressed by means of discourse analysis)

Building bots, most groups tend to learn answer relevance and dialogue management together



- [Amazon Alexa Prize - Conversational AI: The Science Behind The Alexa Prize](#) +
- [Carnegie Mellon University - Building CMU Magnus from User Feedback](#) +
- [Carnegie Mellon University - RubyStar: A Non-Task-Oriented Mixture Model Dialog System](#) +
- [Czech Technical University - Alquist: The Alexa Prize Socialbot](#) +
- [Emory University - Emersonbot: Information-Focused Conversational AI Emory University at the Alexa Prize 2017 Challenge](#) +
- [Heriot-Watt University - Alana: Social Dialogue using an Ensemble Model and a Ranker trained on User Feedback](#) +



# Conclusions 2

- We discovered that a dialogue structure could **be built from the discourse tree of an initial utterance** if it attempts to outline the whole problem
- This structure is built on top of the **default conversational structure** implementing such features as clarification, personalization or recommendation.
- If clarification scenario type is chosen, topics are automatically formed by the chat bot and are presented for a user to choose.
- Discourse analysis of dialogues **support dialogue scenario management in a universal way**, for a broad range of available text fragments and previously accumulated responses.

Video showing chatbot helping to solve a difficult problem for a user

<https://drive.google.com/file/d/0B-TymkYCBPsfV3JQSGU3TE9mRVk/view>