On a chatbot providing a defeating reply in a customer support domain

Boris Galitsky
Oracle Corp Redwood Shores CA USA

Abstract
We formulate a problem of finding a defeating reply for a chatbot to conclude the user session. Defeating reply is expected to attack the user claims concerning a product usability and interaction with a customer support and provide an authoritative conclusive answer to attempt to satisfy this user. We develop reasoning technique to build a representation of a logical argument from discourse structure and reason about it. Our evaluation also involves a machine learning approach and confirms that a hybrid system assures the best performance finding a defeating answer from a set of search result candidates

Introduction

In spite of the great success with chatbots, their deployment in customer support domain is still not robust enough. As mobile and portable devices become popular and enable a number of new products and services, customer expectations of the quality and availability of customer support have significantly raised. Customers expect not only answers to basic questions but also an assistance with resolving problems such as unsatisfactory product features or issues with a service rendered.
Building controllable task-oriented chatbots capable of providing a defeating reply is an essential milestone in developing chatbots that can solve customer problems rather than just providing recommendation or performing a basic transaction.

In many cases, customers want to take advantage of an organization, are very demanding, or just in a bad mode. Supporting a conversation with such customer, the chatbot needs at some point put this customer in his place by authoritative answer defeating customer claims. The chatbot would need to break the argumentation patterns employed by the customer, explain that the customer is wrong. The chatbot needs to reject customer claims
In this study we explore what kind of discourse representation is required to confirm that a given answer is a good defeating reply to an utterance.
As a result, we will build a filter on top of a generic search engine that would select answers defeating the arguments in customer requests, if appropriate.
In these considerations we will not follow the "customer is always right" paradigm but instead demonstrate how a demanding request can be defeated.
Miss Duncan:
'My dear Mr. Shaw: I beg to remind you that as you have the greatest brain in the world, and I have the most beautiful body, it is our duty to posterity to have a child.'
Whereupon Mr. Shaw replied to Miss Duncan: 'My dear Miss Duncan: I admit that I have the greatest brain in the world and that you have the most beautiful body, but it might happen that our child would have my body and your brain. Therefore, I respectfully decline.'

```
elaboration (LeftToRight)
 same-unit
  TEXT:My dear Mr. Shaw :
  contrast (RightToLeft)
   attribution (RightToLeft)
    TEXT:I beg to remind you
    joint
     attribution (LeftToRight)
      TEXT:that
      TEXT:as you have the greatest brain in the world ,
     TEXT:and I have the most beautiful body ,
   elaboration (LeftToRight)
    TEXT:it is our duty to posterity
    TEXT:to have a child .
 elaboration (LeftToRight)
  elaboration (LeftToRight)
   TEXT:Whereupon Mr. Shaw replied to Miss Duncan :
   elaboration (LeftToRight)
    TEXT:My dear Miss Duncan :
    attribution (RightToLeft)
     TEXT:I admit
     joint
      TEXT:that I have the greatest brain in the world
      contrast
       TEXT:and that you have the most beautiful body ,
       attribution (RightToLeft)
        TEXT:but it might happen
        TEXT:that our child would have my body and your brain .
  TEXT:Therefore , I respectfully decline .
```
Fig. 1 Discourse tree for a pair of utterances *suggestion-denial*

One can see that the main feature of a concise, convincing answer is proper handling of entities. In this case, the reply should characterize **body** and **brain**. To defeat the proposal of the initiator of this conversation, the reply must include opposite sentiments to what was proposed. Hence we have a mapping:

you … brain - I … body

↓ ↓ ↓ ↓

my … body - **your … brain**

https://www.brainyquote.com/authors/oscar_wilde
https://quoteinvestigator.com/2013/04/19/brains-beauty/

Our second example is related to customer support in finance, specifically, to foreign transaction fees:

'*I am an American expat living in the Republic of Panama. I just want you to know that I am thoroughly disgusted with the way you do business, because you charged me an overdraft fee, although I have Overdraft Protection. In spite of that your customer Non-Service wrote to tell me I should watch my balances! Despite of that you charged me several times for late fees on Sunday due dated bills, although they have paid on Monday! Why are you now charging a Foreign Transaction Fee of 3%? I asked about this but the form letter sent to me told me about fees for withdrawing from ATMs. However I have never in my life used an ATM, because I do not want to pay to withdraw my own money!*';

Many banks still charge foreign transaction fees for withdrawing cash at foreign ATMs, even if foreign credit card purchase transactions can occur fee-free. Banks have to convert your money spent into

U.S. dollars so they can charge your account. That conversion costs money, and some card-issuing banks pass that cost along to consumers in the form of foreign transaction fees. Some banks waive certain fees if you withdraw money from partner bank ATMs. For example, Bank of America generally charges 3 percent fee to withdraw cash from a foreign ATM. However, if you use an ATM at one of the company's Global Alliance Partners, the $5 fee is waived.

An answer must address a problem raised in question in a comprehensive, exhaustive manner. An answer cannot just agree to please the user, be a submission to her demand. Instead, to benefit a company, a good answer should deny user demand and instead propose a solution explained to be beneficial for both parties. To do that, one or another premise in user demand needs to be defeated. In our example, instead of proposing a compensation for the incurred fees, the bank representative defeats user claims that fees are unavoidable and unjust and mentions an option to avoid them.

Discourse Trees for this question and answer are shown in Fig.2. Notice that both the user and customer service representative (CSA) used texts with heavy argumentation; in addition, the user tries to amplify her point with strong negative sentiment (shown as [--]). The user relies on rhetorical relations of *Attribution, Explanation* and multiple *Contrasts* to bring her point across: fees should not have been charged. CSA attacks user claims with *Explanation, Attribution* and also multiple *Contrast* relations. Hence the CSA attempts to mimic the discourse of the user claims to defeat them and bring his point across that banks have to charge foreign transaction fees but they can be avoided under certain condition (using certain ATMs).

```
contrast (LeftToRight)
 elaboration (LeftToRight)
  TEXT:' I am an American expat living in the Republic of Panama .
  elaboration (LeftToRight)
   attribution (RightToLeft)
    TEXT:I just want you to know
    elaboration (LeftToRight)
     TEXT:that I am thoroughly disgusted with the way [--]
     explanation (LeftToRight)
      TEXT:you do business ,
      contrast (LeftToRight)
       TEXT:because you charged me an overdraft fee ,
       TEXT:although I have Overdraft Protection .
   elaboration (LeftToRight)
    attribution (RightToLeft)
     TEXT:In spite of that Your customer Non-Service wrote to tell me [--]
     TEXT:I should watch my balances !
    elaboration (LeftToRight)
     contrast (LeftToRight)
      TEXT:Despite of that you charged me several times for late fees on Sunday due dated bills ,
      TEXT:although they have paid on Monday !
     elaboration (LeftToRight)
      TEXT:Why are you now charging a Foreign Transaction Fee of 3% ?
      contrast (RightToLeft)
       TEXT:I asked about this
       elaboration (LeftToRight)
        TEXT:but the form letter told me about fees
        TEXT:for withdrawing from ATMs .
 explanation (LeftToRight)
  TEXT:However I have never in my life used an ATM ,
  TEXT:because I do not want to pay to withdraw my own money !

elaboration (LeftToRight)
 elaboration (LeftToRight)
  explanation (LeftToRight)
   contrast (LeftToRight)
    TEXT:Many banks still charge foreign transaction fees for withdrawing cash at foreign ATMs ,
    TEXT:even if foreign credit card purchase transactions can occur fee-free .
   enablement (LeftToRight)
    TEXT:This is because banks have to convert your money spent into U.S. dollars
    TEXT:so they can charge your account .
  elaboration (LeftToRight)
   joint
    TEXT:That conversion costs money ,
    TEXT:and therefore , some card-issuing banks pass that cost along to consumers in the form of
foreign transaction fees .
   condition (LeftToRight)
    TEXT:Some banks waive certain fees
    TEXT:if you withdraw money from partner bank ATMs .
 contrast (RightToLeft)
  TEXT:For example , Bank of America generally charges 3%fee to withdraw cash from a foreign ATM .
  condition
   TEXT:However ,
   TEXT:if you use an ATM at one of the company ' s GA Partners , then the $ 5 fee is waived .
```
Fig. 2. A pair of discourse trees for Q and A

We show the correspondence between claims of the user and the CSA as a mapping between the phrases in elementary discourse units (EDU) of the Discourse tree pair. A user disagreement with the problem described by the phrase 'charging a Foreign Transaction Fee' is addressed by the CSA phrase 'still charge foreign transaction fees'. For the user, this phrase occurs in the EDU for *Elaboration* (request to answer Why question) so that a Contrast relation follows, and for the CSA attempt to defeat, it occur in the nucleus of *Contrast* relation. The reader can observe that this discourse tree is showing a structure for how the CSA plans his attack on user claims. In the following sections we will explore how to deduce argument representation from a discourse tree structure similar to the one in this example.

Analogously, the verb phrase to pay to withdraw is addressed by two phrases in reply waive certain fees and withdraw money from partner bank ATMs . To provide a defeating reply, the CSA relied on *Explanation-Condition* chain of rhetorical relation to properly handle *Explanation* relation employed by the user.

An Algorithm for Identifying Answers with Defeating Arguments

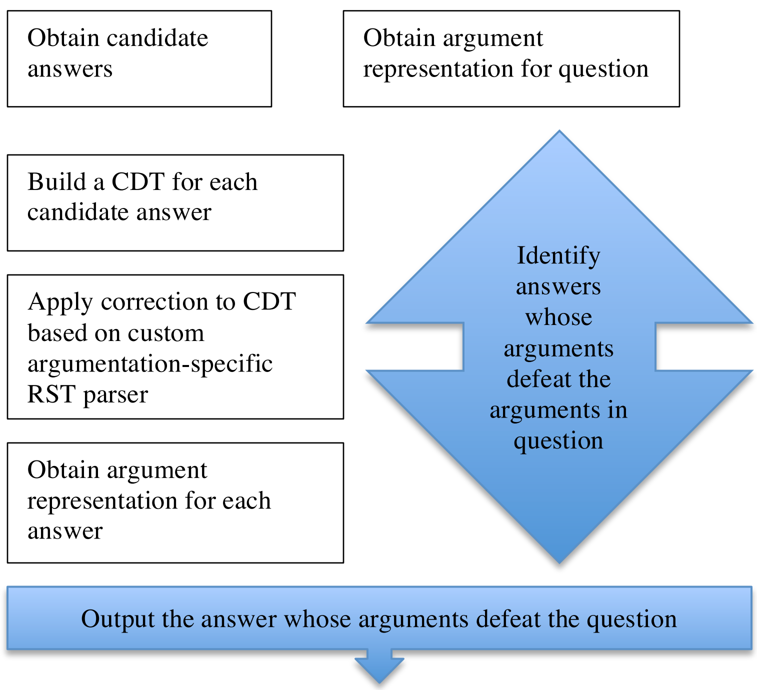

Fig. 3. A high-level architecture of selecting defeating answers

An architecture for selecting a defeating answer is shown in Fig. 3. Candidate answers for a given user utterance are obtained using conventional means to assure relevance; each one is expected to contain the entities and relations between them to properly match the question. Arguments need to be extracted from both the question and each answer, and a correspondence between them needs to be established.

To form an argument representation, we first build CDTs and then improve them by the rules specifically targeting assessing exact rhetorical relations interpretable in terms of arguments. These rules are a basic for an addition rhetorical parser that updates the rhetorical relations established by a conventional RST parser which determines the generic Elaboration relation that needs to be specified to denote *Cause, Reason, Explanation, Conclusion* and others to be properly interpreted as we form a representation for an argument.

We use two classes of approaches, reasoning and learning. The reasoning approach builds a logical representation of a question and attempts to find an answer with such logical representation of argument that it defeats this question. Our machine learning approach learns from good and bad Q/A pairs in terms of how A defeats B. Machine Learning approach does not provide insights how arguments are expressed in text and how they are subjects of reasoning but instead attempts to provide an end-to-end solution of finding a defeating answer.

Representing nested arguments by R-C framework

We first define an argument representation algorithm following Apothéloz () and Amgoud et al(2015).The formalism is built upon a propositional language L with the connectives ¬, ∨, ∘, →, ↔. There are also two operators R(.) and C(.) and an additional negation −. Thus, two negation operators are needed: ¬ for denying propositional formulas (¬x denotes that x is false), and − for denying R(.) and C(.).

An argument is a formula of the form R(y) : (−)C(x). An argument is a reason for concluding a claim. It has two main parts: premises (the reason) and a conclusion. The functions R and C respectively play the roles of giving reason and concluding. Indeed, an argument is interpreted as follows: its conclusion holds because it follows, according to a given notion, from the premises. The notion refers to the nature of the link between them (e.g., the premises imply the conclusion), formally identified by the colon in the definition. However, the contents may be true while the functions do not hold and vice versa. The intuitive reading is as follows:

R(y) : C(x) means that "y is a reason for concluding x"

R(y) : −C(x) means that "y is a reason for not concluding x"

Handling nested argument are important for finding a defeating answer since it is insufficient to handle only object-level or only meta-level layers of argumentation. It is central to handling texts and dialogues: a support for nested arguments and rejections has to be provided. To illustrate some of the expressive richness of our approach, Table … is presented with various forms of arguments and rejections allowed by our definitions (x, y, z, t are propositional formulas to simplify matters). The table is not exhaustive.

It is not shown here how to build a good argument (or a good rejection of an argument). Instead, a representation of arguments (and their rejections) are specified. If an argument or rejection occurs in a text or dialogue, then we want it to be mined, and we want to be able to represent it in our language. A list of arguments below shows that all the forms can be used as a target for natural language. It indicates how to use our language, rather than suggesting that there is a canonical translation of text in to the formal target language. Translating a natural language sentence into R-C logic is shown in Table 1.

Our example argument concern the functionality of a credit card. By default, credit card works (is operational), especially if there is a positive account balance. However, there are exceptions: for whatever reason a bank may decline a transaction.

These examples illustrate that the inner and outer reason R as well as claim C can be potentially identified using argument mining techniques. and then by recursion, the inner reasons and claims can be identified by argument mining techniques. Thus, the nested structure appears better suited as a target language for arguments as they arise in natural language dialogues and texts.

Table 1a. Discourse representation or arguments and their rejections

Basic arguments	My credit card is operational o(c). It is not blocked ¬b(c)	*R (¬b(c)) : C(o(c))*
	My credit card has been compromised *m(c)*. It is blocked	*R (b(c)) : C(m(c))*
	Credit card is operational Thus, it is not possible to conclude that a charge can be declined (d(c)	*R(o(c)) : −C(d(c))*
Single-embedding meta-arguments in reason *R*	That debit card can be used *u(c)* because it is operational, is a reason to conclude that the balance is positive (*p(b)*)	*R(R(u(c)) : C(o(c))) : C(p(b))*
	That card is not declined because it has a positive balance is a reason to conclude that it has not been compromised (*m(c)*).	*R(R(¬d(c)) : C(p(b))) : C(¬m(c))*
	Card is operational because its balance is positive, so we cannot conclude that it was	*R(R(p(b)) : C(o(c))) : −C(b(c))*

	blocked	
Single-embedding meta-arguments in conclusion C	The balance on the card is negative. Thus the charge/use attempt will lead to non-sufficient fund fee (*nsf(c)*)	$R(\neg b(c)) : C(u(c) : C(nsf(c)))$
	The fact that a card has been declined in the past is a reason to conclude that having a positive balance is not a sufficient reason for a credit card to always be operational	$R(d(c)) : C(-R(p(b)) : C(o(c)))$
	The fact that all credit cards of team members are operational is a reason for not concluding that a decline charge of a particular high cost transaction *h(c)* is a reason for team credit cards to be compromised.	$R(o(c)) : -C(R(h(c)) : C(m(c)))$
Double embedding of meta-arguments	Bad credit history (*ch(b)*) leads to a decline of a credit card application (*d(a(c))*). Once a user is unable to use credit card (*u(c)*) it is hard to get a loan (*l(u)*)	$R(R(ch(b)) : C(d(a(c)))) : C(R(u(c)) : C(l(u)))$
	Good credit history (*ch(g)*) usually tells us that a credit card application is not declined (*d(a(c))*). However, we cannot imply that successful credit card application leads to a loan approval (other factors play the role as well)	$R(R(ch(g)) : C(d(a(c)))) : -C(R(d(a(c))) : C(l(u)))$

Table 1b. Discourse trees for selected examples

My credit card is operational o(c). It is not blocked ¬b(c)		$R (\neg b(c)) : C(o(c))$
My credit card has been compromised *m(c)*. It is blocked		$R (b(c)) : C(m(c))$
Credit card is operational. Thus, it is not possible to conclude that a charge can be declined (*d(c)*)		$R(o(c)) : -C(d(c))$

That debit card can be used *u(c)* because it is operational, is a reason to conclude that the balance is positive (*p(b)*)	**cause**
	  explanation (LeftToRight)
	    TEXT:That debit card can be used ,
	    TEXT:because it is operational ,
	  cause (LeftToRight)
	    TEXT:is a reason
	    attribution (RightToLeft)
	      TEXT:to conclude
	      TEXT:that the balance is positive
		$R(R(u(c)) : C(o(c))) : C(p(b))$

That card is not declined because it has a positive balance. It is a reason to conclude that it has not been compromised (*m(c)*).		$R(R(\neg d(c)) : C(p(b))) : C(\neg m(c))$

Card is operational because its balance is positive, so we cannot conclude that it was blocked	**conclusion** (LeftToRight) **cause** (LeftToRight) TEXT:Card is operational TEXT:because its balance is positive , **attribution** (RightToLeft) TEXT:so we can not conclude TEXT:that it was blocked	$R(R(p(b)) : C(o(c))) : -C(b(c))$
The balance on the card is negative. Thus the charge or use attempt will lead to non-sufficient fund fee (*nsf(c)*)	elaboration (LeftToRight) **cause** (LeftToRight) TEXT:The balance on the card is negative **cause**(LeftToRight) TEXT:Thus the charge / use attempt will lead to TEXT non-sufficient fund fee	$R(\neg b(c)) : C(u(c) : C(nsf(c)))$
The fact that a card has been declined in the past is a reason to conclude that having a positive balance is not a sufficient reason for a credit card to always be operational	**reason**(LeftToRight) elaboration (LeftToRight) TEXT:The fact TEXT:that a card has been declined in the past is a reason **conclusion**(RightToLeft) TEXT:to conclude **cause**(LeftToRight) TEXT:that having a positive balance is not a sufficient reason TEXT: for a credit card to always be operational	$R(d(c)) : C(-R(p(b)) : C(o(c)))$
The fact that all credit cards of team members are operational is a reason for not concluding that a decline charge of a particular high cost transaction *h(c)* is a reason for team credit cards to be compromised	elaboration (LeftToRight) TEXT:The fact **reason**(LeftToRight) TEXT:that all credit cards of team members are operational is a reason **conclusion**(RightToLeft) TEXT:for not concluding **cause**(LeftToRight) TEXT:that a decline charge of a particular high cost transaction is a reason for team credit cards TEXT:to be compromised	$R(o(c)) : -C(R(h(c)) : C(m(c)))$
Bad credit history (*ch(b)*) leads to a decline of a credit card application (*d(a(c))*). Thus once a user is unable to use credit card (*u(c)*) it is hard to get a loan (*l(u)*)	**cause**(LeftToRight) **cause**(LeftToRight) TEXT:Bad credit history TEXT:leads to a decline of a credit card application . **cause** (LeftToRight) TEXT: Thus once a user is unable to use credit card TEXT:it is hard to get a loan ,	$R(R(ch(b)) : C(d(a(c)))) :$ $C(R(u(c)) : C(l(u)))$
Good credit history (*ch(g)*) usually tells us that a credit card application is not declined (*d(a(c))*). However, we cannot imply that successful	**explanation** (RightToLeft) **cause** (RightToLeft) TEXT:Good credit history usually tells us TEXT:that a credit card application is not declined . **cause** TEXT:However, we can not imply that	$R(R(ch(g)) : C(d(a(c)))) :$ $-C(R(d(a(c))) : C(l(u)))$

credit card application leads to a loan approval (other factors play the role as well)	successful credit card application TEXT: leads to a loan approval.	

The templates in Table 1b can be used to extract logical atoms from EDUs, translate rhetorical relations into R-C operators and form a logical representation of arguments.

Reasoning with arguments extracted from text

In this section we follow (Amgoud et al 2015) in describing a reasoning system that takes an argument representation of a question and that of an answer and verifies that the latter defeats the former. We treat a set of arguments and their rejections as a set of formulae which is a subject of a reasoning system application. A consequence operator $|\!\!-\!\!-$ is the least closure of a set of inference rules extended with one meta-rule.

A meta-rule expresses that one can reverse any inference rule

$$\frac{R(y):\ F}{-R(y):G} \quad \text{into} \quad \frac{R(y):G}{-R(y):F}$$

This inference rule reversing process occur whenever negation occurs in front of a leftmost "R" so that, in the general case, an inference rule 1 where i, j ∈ {0, 1}

As to the regular inference rules, we start from consistency:

$$\frac{\mathcal{R}(y):\mathcal{C}(x)}{-\mathcal{R}(y):-\mathcal{C}(x)} \qquad \frac{\mathcal{R}(y):\mathcal{C}(x)}{\mathcal{R}(y):-\mathcal{C}(\neg x)}$$

Reasons are interchangeable. This rules is referred to as mutual support

$$\frac{\mathcal{R}(y):\mathcal{C}(x) \quad \mathcal{R}(x):\mathcal{C}(y) \quad \mathcal{R}(y):\mathcal{C}(z)}{\mathcal{R}(x):\mathcal{C}(z)}$$

The next rule gathers different reasons for the same conclusion within a single argument:

$$\frac{\mathcal{R}(y):\mathcal{C}(x) \qquad \mathcal{R}(z):\mathcal{C}(x)}{\mathcal{R}(y \vee z):\mathcal{C}(x)}$$

Cautious monotonicity means that the reason of an argument can be expanded with any premise it justifies. Cut expresses a form of minimality of the reason of an argument.

$$\frac{\mathcal{R}(y):\mathcal{C}(z) \qquad \mathcal{R}(y):\mathcal{C}(x)}{\mathcal{R}(y \wedge z):\mathcal{C}(x)} \qquad \frac{\mathcal{R}(y \wedge z):\mathcal{C}(x) \qquad \mathcal{R}(y):\mathcal{C}(z)}{\mathcal{R}(y):\mathcal{C}(x)}$$

The two next rules describe nesting of R(.) and C(.). Exportation shows how to simplify meta-arguments and Permutation shows that for some forms of meta-arguments, permutations of reasons are possible

$$\frac{\mathcal{R}(y):\mathcal{C}(\mathcal{R}(z):\mathcal{C}(x))}{\mathcal{R}(y \wedge z):\mathcal{C}(x)} \qquad \frac{\mathcal{R}(y):\mathcal{C}(\mathcal{R}(z):\mathcal{C}(x))}{\mathcal{R}(z):\mathcal{C}(\mathcal{R}(y):\mathcal{C}(x))}$$

When is the smallest inference relation obeying the rules above, reflexivity, monotonicity and cut hold, meaning that with the consequence relation, manipulation of arguments by the inference rules is well-founded (Tarksi 1956). Indeed Let Δ be a set of (rejections of) arguments. Let α, and β be arguments.

Δ α if α ∈ Δ (Reflexivity)

Δ ∪ {α} β if Δ β (Monotonicity)

Δ β if Δ ∪ {α} β and Δ α (Cut)

Also, the consequence relation is paraconsistent in the sense that it is not trivialized by contradiction: not all formulae in language L follow from contradiction.

Adjusting available discourse parsers to argumentation domain

Nowadays, discourse parsers are trained on a fairly limited corpus (Muller et al 2012). Moreover, this is the corpus of news articles where analysis of arguments is not necessarily well represented. They take into account conjunctive adverbs like *however* but do not have enough data to rely on the verb *imply* to figure out the rhetoric relation of *Explanations* vs *Contrast*. Hence results of machine learned discourse parsing need to be overwritten taking into account semantics of verbs in EDUs. This is in addition to using these verbs' signatures for DT edge's labels.

We perform an additional classification of rhetorical relations based on communicative actions verbs and phrases such '*as be a reason for'* in nucleus and/or satellite. In the first example above, *nucleus[no CA] -> satellite [conclude]* gives *Conclusion*. In the second example, *nucleus[tell] -> satellite [imply]* gives *Explanation*.

We first show how rhetorical relations can be determined by connectives, and show substantial ambiguity preventing one from properly determining these relations based on connectives only. Notice that in a few thousand sentence – sized training dataset it is possible to generalize such connectives but not necessarily other determining phrases: a significantly larger training dataset for rhetorical parsing is required. Therefore, we intend to explore the determining features of rhetorical relations in the context of argumentation in a rule-based manner.

Table 2: Connectives providing cues for rhetorical relations

	cause	conclusion	explanation	contrast	condition	enablement	elaboration	joint	attribution	restatement	comparison	temporal
however				1								
for example			1				1	1		1		
and	1	1						1				
meanwhile							1					
therefore	1											
hence	1											
finally	1			1								
nevertheless				1								
instead				1								
moreover												
then	1	1	1		1							
on the other hand		1		1			1					
in particular							1			1		
indeed		1					1		1			
overall		1	1							1		
in other words		1								1		
rather											1	
by contrast												
by then												
otherwise				1	1							
thus		1										
yet				1			1					
since	1		1		1							
to						1						
but				1								
if	1				1							
as a result	1	1										
because	1		1									
by									1			
due									1			
when												

Given this table which reflects our observation on how connectives determine rhetorical relations, we apply Formal Concept analysis to visualize the interrelationships between rhetorical relations in terms of their discourse cues (Fig. 4).

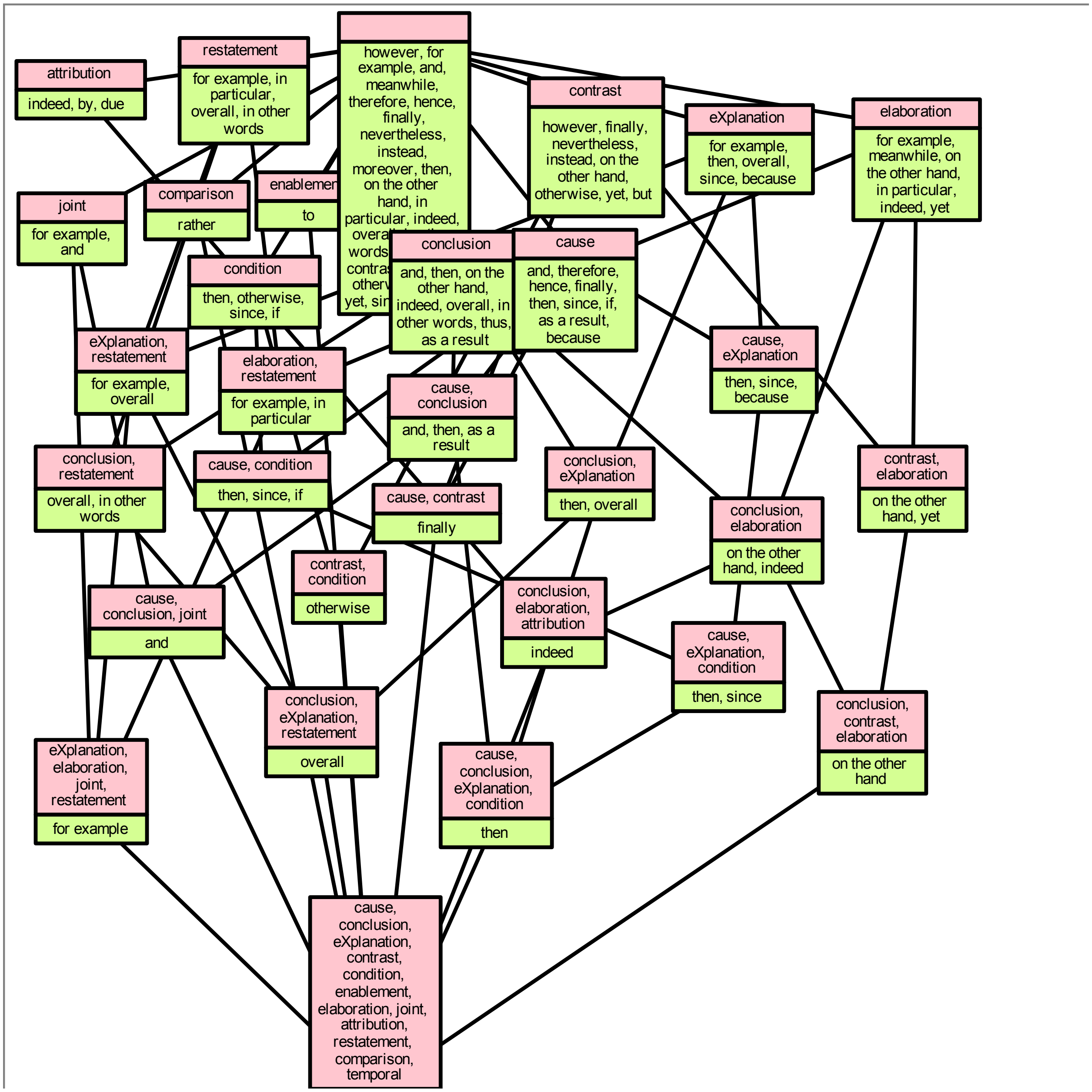

Fig.4 Visualization of a lattice for connectives as attributes of rhetorical relations

We now proceed to more specific treatment of connectives and show how classes of verbs in a nucleus and satellite determine the rhetorical relations.

Table 3: Syntactic patterns for rhetorical relations adjusting to argumentation domain

Verb or phrase	Adverbs in Nucleus	Conjunctive Adverbs in Satellite	Nucleus	Satellite	Relation	
		However, so,	Is a reason		cause	
			Is caused		cause	
			to <verb>		enablement	
		that	Consequences are		cause	
		so	<action>	PRP can	enablement	
<Nucleus>… so they can … <Nucleus>… to do something <verb that has completion state>						
		Hence, so	Summarize, conclude, sum up, result, lead to		conclusion	
		Because, so			explanation	
	If				condition	
		that	Imply, lead to, cause, brings upon		cause	
			Imply, lead to, cause, brings upon		cause	
Thus <phrase> lead to <phrase>						
		That is why			explanation	
	but		inform		contrast	
ask				but	contrast	
I asked but the form did						
		Although, in spite of the fact, despite the fact	Charge, apply, demand, require	although	contrast	
You charged me an overdraft fee, although I have Overdraft Protection.						
I do, however, I use …						

More detailed linguistic sentence-level patterns determining particular rhetorical relation is shown in Table 3. We specify connectives for both nucleus and satellite, as well as verb classes with the focus on communicative actions, associated with particular rhetorical relation. In some cases, below the pattern row, we show a sentence example or a generalized phrase to indicate a source of a given pattern.

Evaluation

In this study we conduct three-step evaluation:

1) Manual evaluation of communicative discourse tree (CDT) construction, R-C representation and reasoning;
2) Automated evaluation of overall recognition accuracy for defeating answers;
3) Assessment of how learned feature of defeating answers matches intuition of search users in terms of how they score these answers in a social search environment.

We first evaluate the argumentation extraction component (Table 4). Whereas CDT are built and corrected reasonably well, R-C representation accuracy is almost 10% lower since there is an ambiguity transitioning from CDT to R-C representation mapping rhetorical relation into either R or C. Adequate inference is achievable in almost 60%: further five percent are adequately represented but inadequately being reasoned about.

Table 4: Resultant accuracies for each step of argument representation algorithm

	Correctly represented CDT of a question	Correctly represented CDT of a defeating reply	Correctly represented logical argument of a question	Correctly represented logical argument of a defeating reply	Sound inference matching arguments for Q and A
Customer complaints	75.4	73.4	66.0	65.4	58.9
Auto Repair	79.2	78.1	69.9	71.1	64.2
Financial Recommendation	69.0	72.5	67.6	68.6	62.8
Yahoo! Answers	82.7	77.8	75.2	73.2	66.7

Table 1: Results for argument extraction, improving rhetorical parsing and converting a discourse tree into R-C representation

Now we proceed to evaluation of the overall defeating argument selection system. For that we form hypothetical dialogues from customer complaints and select the final, defeating reply by a company representative. We split the complaint text into utterances based on indirect speech indicators and communicative actions ('I said' … 'they replied'). The last utterance is frequently the reason complaint arise, so these company replies should have managed to bring their points across and upsetting customer at the same time. From these utterances, we want to learn the real-world rhetorical and argumentative structure used by customer support representatives.

To assess how we can classify an answer as defeating, given a question or an arbitrary utterance, we represent complaint texts as unordered sets of question/answer pairs concealing the actual sequence for testing. We classify each company response as final or not final, assuming that the final response is defeating: the customer gave up on further communicating with the opponent company and resorted to other means to fix his problem.

We also apply similar considerations to the auto repair dataset. The final response usually either solves the problem or convinces the user that something else needs to be done and it is reasonable to leave the auto repair conversational thread. In this respect, the last utterance in an auto repair thread is also a defeating answer since a user is convinced not to continue the thread for whatever reason. In both these datasets, random classifier achieves about 33% accuracy: there are 3.3 utterance pairs for customer complaint and 2.8 utterance pairs for auto repair.

The results of the end-to-end evaluation for both reasoning and ML system are shown in Table 5.

Table 5. Evaluation of the stand-alone and hybrid defeating answer recognition system

	P argument	R argument	F1 argument	P ML	R ML	F1 ML	F1 hybrid
Customer complaints	73	74	73.6	67	71	69.8	77.2
Auto Repair	80	78	79.4	79	82	81.0	84.6
Yahoo! Answers	75	72	73.5	77	75	76.2	81.5

We also explore if defeating replies are rated highly by readers of a conversation or answer. We rely on Yahoo! Answers dataset to assess if defeating replies are rated higher than non-defeated, based on human assessment and based on our model trained and verified on Complaints and Car Repair datasets.

Percentages of most defeating answers from the list of user answers which have the highest rating are shown in Table 6. These percentages estimate our intuition that defeating answers are frequently wanted by users who want to solve the problem, ready to get to the final solution or appeal to the last instance.

Table 6. Discovering correlation between a defeating answer and the one with the highest rating.

	As determined by logical argumentation component	As determined by ML component	As determined by the hybrid system
Business	21.7	19.6	23.8
Job-related	12.6	14.0	15.4
Travel and entertainment	27.3	22.4	29.4
Personal life	16.8	21.0	24.5
Sports	22.4	23.8	27.3
Shopping	23.8	21.7	25.9

One can see that in different domains the users of Yahoo!Answers have different expectations concerning how an answer should defeat a point of a novice user who initiates a thread, being not knowledgeable. In highly opinionated travel, entertainment and shopping domains readers accept that a point raised by an initial question is defeated. At the same time, in less opinionated domains the answers with highest rating do not defeat the claim or opinion of a thread initiator but instead support it and provide useful information without trying to make the thread initiator look as someone possessing limited knowledge. \

Discussion and Conclusions

Proper recognition of rhetorical relations in a specific domain such as argumentation is associated with the task of predicting discourse connectives (Malmi et al 2018). The authors believe that a dialog system might assemble a long and informative answer by sampling passages extracted from different documents retrieved from various sources. In this study, on the contrary, we demonstrated how a dialog can be driven in terms if its genre to a defeating answer completing this dialogue and attempting to convince a user with authoritative answer.

Certain people behavior forms are associated with question answering activities on sites such as Yahoo! Answers. A number of studies have looked at the structure of the community and the interaction between askers and responders. Studies of user typology on the site have revealed that some user category (specialists) answer from personal knowledge, and others prefer to use external sources to construct answers. Observing a social network of Yahoo! Answer users, it turns out that it is possible to distinguish "answer people" from "discussion people" with the former found in specialist categories for factual information, such as mathematics and the latter more common in general interest categories, such as relationship and travel. They also show that answer length is a good predictor of "best answer" choice (Adamic et al. 2008). Looking at the comments given by users on choosing best answers, one can observe that the most significant criteria (Kim and Oh 2009) are as follows:

1) content completeness,
2) solution feasibility and
3) personal agreement/confirmation.

What we assessed in this study is the first item.

There are multiple strategies people use to defeat their opponents, such as what us referred to as Straw man approach. Sometimes it helps to misrepresent an argument so that one can more easily defeat it. Just as a straw man is easier to knock down than a real man, so a distorted version of an argument is easier to defeat than the actual argument. If an argument is over-generalized, then it is easier to find a counter-argument for it:

'My wife recently told me I should take out the trashcan. I responded, "Why do I have to do everything? If I spent my entire weekend doing housework, I would not have any time to work on my book"'

This is like a straw man fallacy because the original claim (that I should do something (i.e. take out the trash)) was taken and over-generalized and misrepresented towards the statement that I should "do everything."

We demonstrated that answers defeating users' claims can be filtered out, if available, relying on hybrid reasoning + ML approach. Here we focused on the former components and evaluated both of them, confirming that they complement each other.

References

Apothéloz, D (1993 The function of negation in argumentation. J. of Pragmatics , 23–38
Leila Amgoud, Philippe Besnard, Anthony Hunter. Representing and reasoning about arguments mined from texts and dialogues. 13th European Conference, ECSQARU 2015, Jul 2015, Compiègne, France. pp. 60-71.
Adamic, L.; Jun Zhang, Eytan Bakshy, Mark S. Ackerman (2008). Knowledge Sharing and Yahoo Answers: Everyone Knows Something". WWW 2008 / Refereed Track: Social Networks & Web 2.0 - Analysis of Social Networks & Online Interaction. Retrieved December 5, 2017.
Kim, S and Oh, S. (2009) Users' relevance criteria for evaluating answers in a social Q&A site". Journal of the American Society for Information Science and Technology. 60 (4): 716
Constrained decoding for text-level discourse parsing
Philippe Muller, Stergos Afantenos, Pascal Denis, Nicholas Asher (2012) COLING, 1883–1900, Mumbai, India
Tarski, A (1956) Logic, Semantics, Metamathematics. J. H. Woodger (ed), Oxford U. Press.
Tatiana Makhalova, Dmitry Ilvovsky and Boris Galitsky (2019) Navigate and Refine: IR Chatbot based on Conceptual Models. ICCS 2019.
Eric Malmi, Daniele Pighin, Sebastian Krause and Mike Kozhevnikov (2018) Automatic prediction of discourse connectives. https://arxiv.org/pdf/1702.00992.pdf. Proceedings of LREC.