Contextualizing Customer Complaints by Identifying Latent Beliefs and Tailoring a Chatbot's Dialog through Epistemic Reasoning

C. Anantaram, Amit Sangroya
TCS Innovation Lab, Tata Consultancy Services Limited, ASF Insignia, Gwal Pahari, Gurgaon, India
{c.anantaram, amit.sangroya}@tcs.com

Abstract

A customer's complaint about a particular product or service describes the problem faced by the customer, and is sometimes a complex set of natural language sentences describing the situation that he/she faces. Such a description may also include preconceived beliefs and emotional status of the customer. When a chatbot engages in a dialog with the customer for addressing his/her complaint, the bot interacts based on its underlying model and tries to resolve the problem. However, there may be situations when the bot is unable to interact appropriately due to the latent information in the complaint. We propose that it is important for a chatbot to contextualize its dialog by identifying the latent beliefs in customer complaints and use that to tailor the dialog. We present a mechanism that uses machine-learning to categorize the complaint and use that to identify the latent beliefs held by the customer. The beliefs thus identified together with the facts extracted from the complaint, trigger epistemic reasoning that helps tailor the dialog and make it consistent with the set of beliefs of the customer. Such a mechanism helps carry out meaningful conversations with customers.

1 Introduction

Many-a-times customers describe their complaints about a particular product or service that they use, through a complex set of natural language sentences. Such descriptions may even contain preconceived notions regarding the problem and/or the customer's emotions embedded in it. As a result, dialog based interaction between a customer and a chatbot may become tedious and irrelevant if the chatbot's beliefs are not synchronized with the set of beliefs held by the customer. Beliefs are cognitive representational states that represent the presumed facts or context of the conversation perceived by each agent. In other words, the presumed context that each agent has should match with the evolving context of the dialog for a meaningful conversation to take place.

As an example, let us suppose that a customer interacts with a chatbot and complains about a problem faced in the car that he/she is using through the following description: "*The gears were slipping when I drove and the car jolted suddenly as it went in and out of gear. I immediately took to the dealer only to have them flush the transmission. Later the transmission was 'fixed' by placing an oil jet kit in the car which has done absolutely nothing. Now even transmission wont go into 3rd gear.*" In this complaint, it is evident that customer is rather upset (text: *car jolted suddenly*); car is in a running state (text: *took to the dealer,* and *transmission wont go into 3*rd *gear*); customer has already visited service center and needs priority service (text: *Later the transmission was fixed*). However, if an automated car diagnosis chatbot had responded back asking "*Does the car start?*", then the customer may get irritated and respond by saying "*What do you think? When I said that transmission wont go into 3*rd *gear, I wouldn't start the car?*" and abruptly end the conversation. In such a scenario, the chatbot did not factor in the set of probable latent beliefs of the customer and thus responded mechanically, in turn making the context inconsistent with the customer's beliefs, leading to such a response. An automated car diagnosis chatbot system aiming to resolve the customer issue must account for such beliefs while carrying out a conversation with the customer. The standard method to avoid inconsistent system responses is to train the system with huge dialog corpus covering various scenarios, or have a complex set of handcrafted rules. Both approaches may be impractical in many real-world domains.

In our work, we propose a model that uses machine learning mechanism to categorize the customer complaint, and then uses information extraction with belief-evaluation-rules to identify the latent beliefs held by the customer. Our model then evaluates the beliefs through epistemic rules of the domain, and tailors the dialog to make it consistent with the set of beliefs of the customer. This process helps contextualize the interaction and drives the conversation in a meaningful way.

Our paper is organized as follows: Section 2 provides a brief overview of state of art approaches in conversational systems. Section 3 presents our approach to contextualize the dialog by identifying latent beliefs in a customer's complaint and using epistemic rules to tailor the interaction. In section 4, we discuss the experiments with the proposed system in car-diagnosis domain. Section 5 presents the conclusions and future work.

2 Related Work

A number of mechanisms have been presented that make use of machine-learning to build intelligent dialog systems [Henderson *et al.*, 2014; Weston, 2016]. Deep learning based dialog systems [Miller *et al.*, 2016] use memory networks to learn the underlying dialog structure and carry out goal-oriented dialog. On the other hand, traditional dialog systems are specialized for a domain and rely on slot-filling driven by a knowledge base and a finite-state model [Lemon *et al.*, 2006; Wang and Lemon, 2013].

Dynamic epistemic logic has been used to build formal dialog systems where two agents, an Opponent and a Respondent, engage in an alternating-move dialog to establish the consistency of a proposition [Uckelman et al., 2010]. Sadek et al. proposed a reasoning engine to build effective and generic communicating agents [Sadek et al., 1997]. Rudnicky et al. propose a novel approach for constructing beliefs over concept values in dialog systems by integrating information across multiple turns in the conversation [Bohus and Rudnicky, 2005]. Their approach is data-driven and provide a unified framework for belief updating. Motivated by the need for a data driven framework, Young et al. review a list of statistical driven techniques for spoken dialog system [Young et al., 2013]. However, the above models do not factor in beliefs or trigger epistemic rules to modify the conversation given the customer's context. In our work, we show a mechanism to identify the beliefs held by the customer, use these beliefs to trigger epistemic rules, and use the assertions of the rules to drive the conversation by tailoring the states in a finite-state-machine dialog system [Anantaram and Sangroya, 2017].

3 Identifying Latent Beliefs to Tailor Dialog

We consider beliefs as cognitive representational states that are represented as ground facts extracted from the dialog carried out by each agent. This view is similar to '*beliefs as mental states*' as described in [Gabaldon and Langley, 2015]. Since beliefs are not explicitly stated facts in a conversation, these have to be perceived and extracted, and thus we refer to it as *latent beliefs*. Our method extracts the latent beliefs through a sequence of steps consisting of information extraction, opinion mining and reasoning over the facts explicitly mentioned in the dialog. We describe this process below.

As we pointed out earlier, a customer's complaint description may consist of a set of natural language sentences describing his/her troublesome situation, sometimes including their feelings and emotions. It is fairly obvious that different customers will express their situation in different ways. In such a scenario, we first need to determine the overall category the complaint falls into. We use a recurrent neural network (RNN) to determine the most probable category of the complaint. For example, let us consider the domain of car-complaints wherein customers complain about the problems that they face while using their car. Since such complaints can vary drastically for the various subsystems of a car, a RNN can be used to classify customer complaints into six categories of problems, viz., Transmission, Gear, Windows-Windshield, Engine-failure, Wheels-Hubs and AC-Heater.

Once the most probable category is selected, an associated Frame-slot mechanism extracts the relevant information from the complaint. We use OpenIE [Manning et al., 2014] to extract triples from the complaint. The triples are evaluated with respect to an ontology to identify specific ground facts stated in the complaint. The ground facts are then used to fill the frame slots. For the example in section 1 from the sentence "*Now even transmission wont go into 3rd gear*" the triples extracted by OpenIE are (`transmission, wont, go into 3rd gear`). The ground fact asserted by ontology evaluation of the triple is: "`transmission(not-working), transmission-problem (3rdgear)`". These ground facts help fill the frame-values.

Next, the sentiments and opinions expressed by the customer in the complaint are identified using an opinion miner. A number of techniques are available in the area of sentiment analysis and opinion mining [Chen et al., 2017, Taboada et al., 2011] and we use some of these techniques for extracting opinions from the customer complaints. For the above example, the opinion would be extracted as "`negative(even transmission wont go into 3rd gear)`".

This set of extracted information consisting of `{category, frame-values, opinions}` is then evaluated by a hand-crafted beliefs-knowledge base to assert facts about the probable latent beliefs held by the customer while making the complaint. For the above example the following belief rules asserts that the 'car is difficult to use' and that 'customer is harrased': "`evaluate-transmission-problem:-car(running), gear-shift (slipping), record-fact(belief (transmission(problem))), record-fact( belief( car(difficult-to-use))), evaluate-opinion. evaluate-opinion:- opinions (verynegative), record-fact( belief( customer (harassed))).`"

The latent beliefs and the facts extracted trigger the epistemic rules that are hand-coded into an epistemic-knowledge base. The evaluation of the epistemic rules gives us the states in the finite-state-machine (FSM) to skip or focus on, thus tailoring the dialog the chatbot has with the customer. For example, the following epistemic rule asserts that the chatbot should skip 'general chit-chat' conversation when the 'car is difficult to use' and the 'customer is harrased' "`epistemic-rules-transmission-problem:-belief (transmission(problem)), belief (car (difficult-to-use)), belief (customer (harassed)), record-fact( skipstate (general-chit-chat)).`"

We assume that we have a FSM based complaint handling system to process the extracted information and the asserted facts, and carry out the dialog with the customer. However, since the FSM is generic for the domain, it needs to be tailored for handling a particular customer complaint effectively. In order to achieve tailoring of the FSM, the latent beliefs of the customer that were asserted as facts, and the category of the complaint are then evaluated by the epistemic rules encoded in an epistemic-knowledge base for the domain. The rules make assertions about the states in the FSM that need to be skipped and the states that need to be evaluated in order to factor-in the beliefs of the customer. The subsequent dialog is carried out and the next set of beliefs are then evaluated. The cycle then continues.

We believe that it is very important for a customer-interaction-chatbot to identify and process a customer's latent beliefs and use that to tailor its interactions. In our architecture this is achieved through two levels of abstraction: (a) domain-specific situations that indicate implied states; for example beliefs such as "car(difficult-to-use)", and (b) generic mental states of the customer based on his/her opinions expressed in the complaint; for example latent beliefs such as "customer(harrassed)". Such an abstraction facilitates focus on the hidden states that are not explicitly processed in various chatbot architectures. Additionally, one of the advantages of our architecture is that belief rules such as "opinions(verynegative), record-fact(belief(customer (harassed)))" can be generic across domains. Further, some of such belief rules can possibly be machine-learnt from large customer interactions corpora.

4 Experiments

We have conducted experiments in the domain of car complaints, where customers report problems they are experiencing on their cars. Feld et al. describes the automotive ontology to facilitate inference and reasoning [Feld and Müller, 2011]. We show a small ontology in Figure 1.

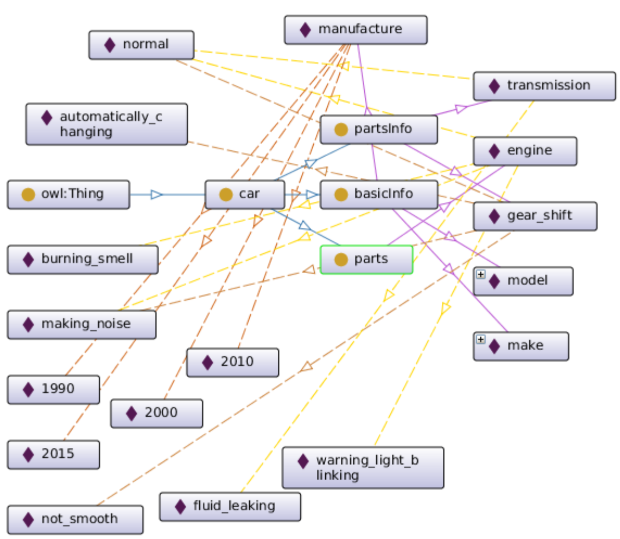

Figure 1: An Example of Car Ontology

In this example, we demonstrate a small part of the ontology such as car's basic information, information about various parts, car's model and year of manufacturing etc. All these entities can take a list of possible real world values. Possible values of *parts* are {transmission, gears, engine, brakes} etc. The ontology along with a rule base assists in driving the conversation for diagnosis. The rule base consists of domain knowledge about various car issues, their symptoms and common diagnosis for these issues.

A few example complaints along with their category, frame-values, opinions and beliefs, epistemic rules, actions states are shown in Table 1. Firstly, as shown in second column we categorize the complaint using RNN based machine learning module. We train RNN using 0.6 million training examples. We validate the model using 0.2 million test examples. We got approximately 70% test accuracy in complaints classification. We also extract customer opinions as shown in the third column. In these examples the opinions of customers are negative and strong negative. In Table 3, we demonstrate the conversation for one of the above examples. The first state is mostly a detail problem description by the customer. Thereafter, using the complaint classification, opinions and belief extraction, the chatbot validates the epistemic rules. As a result of this, it can drive the FSM i.e. it can skip some states like general chit chat, car running experience etc. It initiates the conversation and at each state it tries to fill the slots of the dialog frame and in the end propose a diagnosis to the customer. As shown in Table 2, we can observe that chatbot is able to skip some FSM states as a result of epistemic rules that were derived from information extraction and initial processing of the customer complaint. The second column depicts the belief rules that were evaluated from the complaint. In the third column, examples of epistemic rules are presented that are validated against the set of beliefs. Fourth column shows the example of the FSM: Primarily, the states it has to skip and states it has to ask.

The tables below show the results of our experiments.

Customer Complaint *Cx*	**Category and Frame-values**	**Opinions**
C1: The gears were slipping when I drove and the car jolted suddenly as it went in and out of gear. I immediately took to the dealer only to have them flush the transmission. Later the transmission was "fixed" by placing an oil jet kit in the car which has done absolutely nothing. Now even transmission wont go into 3rd gear.	**Category through ML:** Transmission **Frame-values:** gearshift(slipping) car(running) transmission-lastfix(flushed) transmission(not-working) transmission-problem(3rdgear)	**strongneg**(car jolted suddenly) **strongneg**(immediately took to the dealer) **neg**(even transmission wont go into 3rd gear)
C2: my car just died on me. No warning no check engine. Car just out of extended warranty all maintenance up to date. Had issues with charcoal	**Category through ML:** Engine-failure **Frame-values:** car(notrunning), last-fix(shift-lever), last-	**neg**(No warning no check engine)

canister, and shift lever. Engine croaked without overheating, no warning, no check engine.	fix(charcoal-canister),engine(dead),overheating(no),warninglights(no),checkenginelight(no),extendedwarranty(over),scheduledmaintenance(done)	

Table 1: Complaint, its category, extracted values, opinions

Cx	Belief rules	Epistemic rules	FSM states
C1	evaluate-transmission-problem :- car(running), gear-shift(slipping), record-fact(belief (transmission(problem))), record-fact(belief(car(difficult-to-use))), evaluate-opinion. evaluate-opinion:- opinions(verynegative), record-fact(belief(customer (harassed))).	epistemic-rules-transmission-problem :- belief(transmission(problem)), belief(car(difficult-to-use)), belief(customer(harassed)), record-fact(skipstate(general-chit-chat)), record-fact(skipstate(car-movement-experience)), record-fact(skipstate(basic-maintenance-questions)), record-fact(askstate(advanced-maintenance-questions)).	skipstate (general-chit-chat)), skipstate (car-movement-experience)), skipstate (basic-maintenance-questions)), askstate(advanced-maintenance-questions))
C2	evaluate-engine-failure :- car(not-running), scheduledmaintenance(done), engine(dead), overheating(no), warninglights(no),checkenginelight(no), record-fact(belief(engine-failure(sudden))), record-fact(belief(car-was-running-before)),record-fact(belief(basic-car-care-done)), evaluate-opinion. evaluate-opinion:- opinions(negative), record-fact(belief (customer(very-upset))).	epistemic-rules-engine-failure :- belief(engine-failure(sudden)), belief(car-was-running-before), belief(basic-car-care-done), belief (customer(very-upset)), record-fact (skipstate(car-starting)), record-fact (skipstate(basic-maintenance-questions)), record-fact(skipstate(general-chit-chat)).	skipstate(car-starting) skipstate (basic-maintenance-questions) skipstate (general-chit-chat).

Table 2: Beliefs and Epistemic rules for the complaints

As shown in Table 2, we can observe that chatbot is able to skip some FSM states as a result of epistemic rules that were derived from information extraction and initial processing of the customer complaint. In Table 3 we demonstrate the conversation for one of the above examples.

Customer	The gears were slipping when I drove and the car jolted suddenly as it went in and out of gear. I immediately took to the dealer only to have them flush the transmission. Later the transmission was "fixed" by placing an oil jet kit in the car which has done absolutely nothing. Now even transmission wont go into 3rd gear.
Bot (skipstate: general-chit-chat) ; (skipstate: car-movement-experience); (askstate: advance-maintainence-questions)	Okay. Any burning smell coming from your car?
Customer	Yes, there is a burning smell
Bot (askstate: advance-maintainence-questions)	Have you ever got the car clutches checked?
Customer	Not yet
Bot	Oh Okay. Well then it could probably be a clutch issue. You need to get the clutches checked.
Customer	Well, okay

Table 3: Sample Output of Dialog System

We have done a number of experiments on various real-world complaints in the domain of car-complaints. We validate the system across six categories of car complaints. We got a 20% increase in efficiency as a result of the proposed framework. Overall results of our experiments evaluation are summarized in Table 4.

Average Dialog Length with beliefs	12	**Dialog Completion Rate with beliefs**	95%
Average Dialog Length without beliefs	15	**Efficiency Increase**	20%

Table 4: Reduction in dialog turns for the complaints

5 Conclusions

Factoring latent beliefs with epistemic rules help a chatbot tailor its dialog to efficiently handle customer complaints. In our work, we present an approach to contextualize the dialog by identifying latent beliefs in a customer's complaint and using epistemic rules to tailor the interaction. Our experimental results have been promising. The chatbot is able to have one FSM and tailor that appropriately for the customer's situation on hand. This leads to more relevant dialog for the customer.

References

[Anantaram and Sangroya, 2017] C. Anantaram, Amit Sangroya. 'Identifying Latent Beliefs in Customer Complaints to Trigger Epistemic Rules for Relevant Human-Bot Dialog', *IEEE 3rd International Conference on Control, Automation and Robotics*, 2017.

[Bohus and Rudnicky, 2005] Dan Bohus and Alexander I Rudnicky. 'Constructing accurate beliefs in spoken dialog systems'. In *Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on, pages 272–277. IEEE, 2005*.

[Chen et al., 2017] Tao Chen, Ruifeng Xu, Yulan He, and Xuan Wang. Improving sentiment analysis via sentence type classification using bilstm-crf and cnn.(report). *Expert Systems With Applications, 2017-04-15*.

[Feld and Müller, 2011] Michael Feld and Christian Müller. The automotive ontology: Managing knowledge inside the vehicle and sharing it between cars. In *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI '11, pages 79–86, New York, NY, USA, 2011. ACM*.

[Gabaldon and Langley, 2015] Alfredo Gabaldon and Pat Langley: Dialogue Understanding in a Logic of Action and Belief, in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[Henderson *et al*., 2014] M. Henderson, B. Thomson, S. Young, 'Wordbased dialog state tracking with recurrent neural networks' in *Proceedings of the 15th SIGDIAL*, 2014.

[Lemon *et al*., 2006] O. Lemon, X. Liu, D. Shapiro, C. Tollander, 'Hierarchical reinforcement learning of dialogue policies in a development environment for dialogue systems: Reall-dude'. *BRANDIAL*'06.

[Manning et al., 2014] Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60*.

[Miller *et al*., 2016] A. Miller, A. Fisch, J. Dodge, A.H. Karimi, A. Bordes, J. Weston, 'Key-value Memory networks for directly reading documents', *arXiv preprint arXiv:1606.03126,* 2016.

[Sadek et al., 1997] M. D. Sadek, P. Bretier, and F. Panaget, "Artimis: Natural dialogue meets rational agency," in *Proceedings of the Fifteenth IJCAI - Volume 2, 1997*, pp. 1030–1035.

[Taboada et al., 2011] Maite Taboada, Milan Tofiloski, Julian Brooke, Kimberly Voll, and Manfred Stede. 'Lexicon-based methods for sentiment analysis', 2011.

[Uckelman et al., 2010] U. Sara, "Obligationes as formal dialogue systems," in Stairs 2010: *Proceedings of the Fifth Starting AI Researchers' Symposium, vol. 222. IOS Press, 2010, p. 341.*

[Wang and Lemon, 2013] Z. Wang, O. Lemon, 'A simple and generic belief tracking mechanism for the dialog state tracking challenge: On the believability of observed information', *in Proceedings of the 14th SIGDIAL*, 2013.

[Weston, 2016] J. Weston. 'Dialog-based language learning', *CoRR, vol. abs/1604.06045*, 2016.

[Young et al., 2013] S. Young, M. Gai, B. Thomson, and J. D. Williams. 'Pomdp-based statistical spoken dialog systems: A review'. *Proceedings of the IEEE, 101(5):1160–1179, May 2013.*