UMEÅ UNIVERSITY

Formalization of Dialogues from Movie Corpus using DAMSL Annotation Scheme as Cooperating Distributed Grammar Systems

Maitreyee Tewari

No 21.07/2021

Department of Computing Science

ISSN 0348-0542

Formalization of Dialogues from Movie Corpus using DAMSL Annotation Scheme as Cooperating Distributed Grammar Systems

Maitreyee Tewari

Department of Computing Science
Umeå University
maittewa@cs.umu.se

Abstract. Dialogues are an integral and intuitive communication mode for humans to reach a consensus and for their effective application to human-robot interaction a formal model of their structure is required. Formalization of dialogues enable robots to have constructive interaction and management of dialogue breaches such as ambiguity, sudden topic change and misunderstanding. In line with viewing dialogues as collaborative, a formal model of dialogues is proposed through Cooperative Distributed Grammar Systems (CDGS), which are abstract models based on 'Blackboard problem' in Artificial Intelligence, for building multi-agent systems, through formal grammars. For this research formalisation is done in two stages: first, a corpus of movie scripts is selected, and annotated with the help of an annotation scheme (combination of speech acts, response and social obligation tag). Then the annotated labels are structured through dialogue policies (progressive, binding and co-occurring) proposed in this paper. Finally, formalization of derived structure through CDGS is done.

1 Introduction

Romney [15] defines dialogues as conversations between two or more people, focused on answering questions, suggesting solutions and examining thoughts and actions. Dialogues are intuitive communication mode for humans. Recent research [10,9] in human–robot interaction (HRI) is actively building interfaces for robust and intuitive dialogue interaction.

Spoken dialogue management systems (DMS) facilitate human-like interaction in robots. Some of the challenges faced by current systems are intractability in case of data-driven systems, domain specificity for knowledge driven systems, non-understanding of spoken words and error-prone conversational breach (conflict) resolution. We believe that formalization of dialogue structure could enable robots to have constructive interaction and management of dialogue breaches (for instance ambiguity, sudden topic change and misunderstanding). For dialogue structure formalization, we propose to pre-process the dialogues to build a consistent structure and then formalize them based on CDGS [7].

Grammar based paradigms have been widely used for modeling the conversations. D'Ulizia et al. [8] elaborate on the consistency between human–human dialogues and grammar paradigms. A grammar system [14,7] is a set of grammars with a starting symbol called the sentential form, following specific protocols called the production rules to generate one language. Among several different grammar systems, we elaborate on three relevant categories: sequential grammar system or Cooperating Distributed Grammar Systems (CDGS). This grammar system has a shared sentential form and the production rules corresponding to the different grammars, working in turns to generate a language. Parallel Computing Grammar System (PC Grammar System), where grammars works in parallel on their sentential forms. Eco-grammars are a convergence of CDGS and PC Grammar Systems. In this system, grammars perform derivations in a shared environment. Here the environment evolves based on the information state and the grammars evolve based on the environment state.

Aydin et al. [2] successfully modeled dialogues with augment CDGS with memory system, while in [18] the authors extended CDGS with expectation mechanism for managing miscommunications in assistive robotics scenario. Bel et al. [3] proposed an extension of Reproductive Eco Grammar Systems (EREG) for multi-agent dialogue systems. They complemented the grammar system with multi-agent protocol language (MAP). The protocols for the social norms that must be followed by participants to participate in a conversation are determined by MAP. In 2008 Bel et al. [4] proposed Conversational Grammar Systems (CGS), which could be used to model task driven dialogues as inter-action, providing flexibility to accept the new concepts and modification of the protocols. Another grammatical framework of Multimodal Grammar Generation System (MGGS) [8] was presented. The system takes multimodal dialogues, and automatically generates grammar rules through a grammar inference algorithm. The framework is based on Context-Free Grammar (CFG), and in particular the inductive Cocke–Younger–Kasami (CYK) algorithm because of its simpler derivation rules and shorter computation time.

Here we chose a slightly different procedure to formalize. First, dialogues are pre-processed through the speech act annotation scheme and structuring through dialogue policies. The processed dialogues are formalized based on CDGS. This formalization helped in understanding the policies and in deriving the structures in dialogues. The structure of the paper is as follows: The background and preliminary section elaborates on some of the definitions and concepts used. Methodology section illustrates the process of data aggregation, annotation and structuring. The paper is concluded with results and future work.

2 Background and Preliminaries

CDGS enable dialogues to be represented as cooperative conversation between participants. Here, each dialogue is universally shared as a label and structured through dialogue policies. These policies determine the next dialogue by the participant. This section elaborates on the information required to understand the CDGS. It also introduces the basic terminologies, and annotation scheme required here.

2.1 Notion and Terminology

We assume the reader to be familiar with the notions of formal languages and grammars, however we mention some of the notions that are specific to this work. For more detailed understanding refer to [12]. An alphabet is a finite non-empty set. Let V be an alphabet and λ be an empty string. The set of all strings including the empty string λ generated over the symbols is denoted by the free monoid V^*, and a semigroup generated by V without the empty string is $V^+ = V^* - \{\lambda\}$.

2.2 Definition of CDGS

CDGS are abstract devices representing multi-agent systems. Formally represented by grammar components, that take turns and cooperate together through sets of derivations/production rules, to generate a language [14]. Furthermore, CDGS could be explained through Blackboard Problem, where grammars correspond to participants in a system, with the objective to solve a specific problem on the blackboard. Derivations or production rules correspond to pieces of knowledge that changes the state of the problem on the blackboard. The sentential form or start symbol gives the formal counterpart of the problem on the blackboard. The generated language represents the accepted knowledge piece as solutions/part of solutions.

Definition 1. *CDGS are an (n+2)-tuple [7]*

$$\Gamma = (T, G_1, G_2, ..., G_n, S), \tag{1}$$

where,

- *for* $1 \leq i \leq n$ *, each* $G_i = (N_i, T_i, P_i)$ *is a grammar with the set* N_i *of non-terminal symbols, the set* T_i *of terminal symbols, the set* P_i *of production rules, and without start-symbol* S*,*
- *Terminals* T *is a subset of* $\bigcup_{i=1}^{n} T_i$*,*
- *Start symbol* $S \in \bigcup_{i=1}^{n} N_i = N$*.*

The grammars G_i where $(1 \leq i \leq n)$ are said to be the components of the grammar system Γ. N and T are disjoint alphabets, where elements of N are non-terminals, those of T are terminals. Further we set alphabets $V_i = N_i \cup T_i$ and $V_\Gamma = \bigcup_{i=1}^{n} V_i$.

The grammar system is represented through an $(n+2)$ tuple Γ. The grammar system has components, and each component is a grammar. Component grammar has production rules denoted by $P_i, ..., P_n$. The production rules determine the transition between symbols in the grammar and are over symbols $N \cup T$. Where, N and T are set of non-terminal and terminal symbols respectively. A grammar consist of production rules such as $A \rightarrow a$, with the left-hand and the right-hand side. The non-terminal are expandable symbols present on the left-hand side of production rule and usually denoted by capital letters. Terminal symbols are non-expandable symbols always on the right hand side of the production rule and usually represented by small letters. Grammar generates a language, which is a set of terminal symbols. In the next definition we define the derivation mode, which controls the cooperation among these grammars.

Definition 2. *Let* Γ *be a CDGS system as in Definition 1. Let* $x,y \in V_i^*$*. Then we begin by deriving* $x \Rightarrow_{G_i} y$ *and finally* $x \Rightarrow_{G_i}^{k} y$ *if and only if there are words* $x_1, x_2, ..., x_{k+1}$ *such that*

i) $x = x_1, y = x_{k+1}$*,*

ii) $x_j \Rightarrow_{G_i} x_{j+1}$, i.e. $x_j = x'_j A_j x''_j$, $x_{j+1} = x'_j w_j x''_j$, and there exists $A_j \rightarrow w_j \in P_i$, $1 \leq j \leq k$.

Moreover, we write

$x \Rightarrow^{\leq k}_{G_i} y$ *if and only if* $x \Rightarrow^{k'}_{G_i} y$ *for some* $k' \leq k$,

$x \Rightarrow^{\geq k}_{G_i} y$ *if and only if* $x \Rightarrow^{k'}_{G_i} y$ *for some* $k' \geq k$,

$x \Rightarrow^{*}_{G_i} y$ *if and only if* $x \Rightarrow^{k}_{G_i} y$ *for some* k,

$x \Rightarrow^{t}_{G_i} y$ *if and only if* $x \Rightarrow^{*}_{G_i} y$ *and there is no* $z \neq y$ *with* $y \Rightarrow^{*}_{G_i} z$.

A derivation step is defined as $x \Rightarrow^{k}_{G_i} y$, where k denotes the number of successive steps an active grammar G_i could perform. In CDGS only one grammar can be active at a time, and when one grammar completes performing its actions another grammar continues. A derivation performed from a symbol x to a symbol y in a grammar G_i, with the derivation mode $\geq k$ denotes that the grammar should perform the actions at-least k times. The derivation mode $\leq k$ puts a constraint on the grammar that it can perform the actions at-most k times. The $*$-mode derivation allows the grammar to perform arbitrary number of actions. The t-mode derivation allows the grammar to perform as long as it has applicable actions.

Definition 3. *Let f be a set of derivation modes, represented as*

$$f \in \{t, *, 1, 2, ..., \leq 1, \leq 2, ..., \geq 1, \geq 2, ...\},$$

and let Γ be a CDGS System. Then the language $L_f(\Gamma)$ generated by Γ is defined as the sets of all words $z \in T^*$ for which there is a derivation

$$S = w_0 \Rightarrow^{f}_{G_{i_1}} w_1 \Rightarrow^{f}_{G_{i_2}} w_2 \Rightarrow^{f}_{G_{i_3}} .... \Rightarrow^{f}_{G_{i_r}} w_r = z.$$

To illustrate the above definitions, we present the following example from [7]

Example 1. Let us consider the CDGS System $\Gamma = (\{a, b, c\}, G_1, G_2, S)$,
with

$$G_1 = (\{A, B\}, \{A', B', a, b, c\}, \{A \rightarrow ab, A \rightarrow aA'b, B \rightarrow c, B \rightarrow cB')\}$$

$$G_2 = (\{S, S', A', B'\}, \{A, B\}, \{S \rightarrow S', S' \rightarrow AB, A' \rightarrow A, B' \rightarrow B\}).$$

Then we obtain

$$L_1(\Gamma) = L_*(\Gamma) = L_{\leq k}(\Gamma) = L_{\geq 1}(\Gamma) = L_t(\Gamma) = \{a^n b^n c^m \mid n \geq 1, m \geq 1\}, k \geq 1,$$

$$L_2(\Gamma) = L_{\geq 2}(\Gamma) = \{a^n b^n c^n \mid n \geq 1\},$$

$$L_k(\Gamma) = L_{\geq k}(\Gamma) = \emptyset \; for \; k \geq 3.$$

We show the relation only for $L_t(\Gamma)$. The derivation starts with $S \Rightarrow S' \Rightarrow AB$ using the rules from P_2. Now we change to P_1 and following cases hold:

- *Case0.* $AB \Rightarrow^2 abc$.
- *Case1.* $AB \Rightarrow^2 abcB'$.
- *Case2.* $AB \Rightarrow^2 aA'bc$.
- *Case3.* $AB \Rightarrow^2 aA'bcB'$.

In Case 0 the derivation gets terminated.

Considering the derivation in t-mode. In Case 1 we can continue only in the following way:

$$abcB' \Rightarrow^t abcB \Rightarrow^t abccB' \Rightarrow^t abccB \Rightarrow^t abc^3B'.....$$

$$..... \Rightarrow^t abc^3B \Rightarrow^t abc^{r-2}B.....$$

$$..... \Rightarrow^t abc^{r-1}B' \Rightarrow^t abc^{r-1}B \Rightarrow^t abc^r$$

In Case 2 only words of the form $a^r b^r c$, $r \geq 2$, can be generated. In Case 3 we apply $A' \rightarrow A, B' \rightarrow B$, which yields $aAbcB$. Using rules of P_1 we obtain $a^2b^2c^2$ or $a^2b^2c^2B'$ or $a^2A'b^2c^2$ or $a^2A'b^2c^2B'$. In all the above cases same structure of word in yielded. Therefore, the language generated is

$$L_t(\Gamma) = \{a^n b^n c^m \mid n \geq 1, m \geq 1\}.$$

2.3 Speech as action

Dialogues are studied under different paradigms, some of which are 'Speech act theory' [1,16], 'Conversational game theory' [11] and 'Dialogue interpretation theory' [5]. The notion of viewing utterances or speech as actions to fulfill goals was introduced by Wittgenstein in [17] and modified by Austin [13] and Searle [16]. Speech acts are defined as performative utterances. It consists of locutionary acts: utterances which make sense semantically, contextually and phonologically. Perlocutionary: acts which gives the effect of the utterance. Illocutionary acts: defines the performative force through the utterance. Speech acts[1] have been classified in different ways, for this work we have focused on following four specific categories [13]:

- Representatives: speech acts presenting the information to be true. Such as affirm, believe, deny, conclude, report.
- Commissives: speech acts showing the commitment towards a course of action. Such as guarantee, pledge, swear, promise, vow.
- Directives: speech acts which elicits the person to perform an action. Such as ask, beg, challenge, command, drive.
- Expressives: these speech acts help in expressing an attitude. Such as apologize, appreciate, congratulate, deplore and detest.

Dialogue Act Markup Language (DiAML) proposed by Bunt et al. [6] is an ISO standard dialogue annotation scheme. Bunt explains that dialogues have two major components: semantic content and a communicative function specifying the usage of the semantic content in the dialogue. The communicative function is divided into eight core dimensions namely: task, discourse-structure, allo and auto feedback, turn management, time management, own communication management and social obligation management.

We propose to combine above mentioned speech acts and DiAML dimension 'social obligation' with 'response' tag. We introduced a separate 'response-tag' to generalize all kinds of responses. Dialogue tagging is significant as it brings consistency to dialogue structuring and formalization as elaborated next in the methodology section.

3 Methodology

This section elaborates on the corpora aggregation and pre-processing which includes annotation through speech acts and structuring through dialogue policies. The manual annotation of dialogues through annotation scheme is performed, these tags are then structured with the help of dialogue policies mentioned above.

3.1 Corpora Annotation

The corpora consist of movies from the Natural language and dialogue systems library [20,19]. This dataset was chosen for two reasons: first, movie scripts are analogous to real life scenarios and depict real conversations. Second, the data was pre-processed (punctuation, special characters), so we do not have to spend time in data manipulation. Another corpora that was used is Switchboard-1[2] dialogue corpora with shallow discourse tag-set augmented from Discourse Annotation and Markup System of Labeling (DAMSL). It is a standard corpora used for building dialogue managers for a decade and in our case was used only for manual analysis of dialogue structure.

All the dialogues were pre-processed to remove special characters and non-lexical conversational items were replaced by lexical items conveying the same meaning. Dialogues were then labelled with one or more tags such as: introduction, social-obligation, question, beg. Below is an example of annotated dialogues from a movie. This dialogue has three participants Ruth, Letty and Mrs. Mayer, however the dialogue is mainly between Letty and Mrs. Mayer.

> **RUTH**: *Sorry I'm late.* **(Apologize,Explain)**
> **MRS. MAYER**: *Letty was just getting ready to tell me her good news.* **(Report)**
> **RUTH:** *Tell, tell.* **(Affirm)**
> **LETTY:** *So, I got the district to approve my after-school math program.* **(Declare)**
> **MRS. MAYER:** *How wonderful, darling. What does that mean for you?* **(Appreciate,Question1)**
> **LETTY:** *I'll be running it three days a week, and...* **(Response1)**
> **MRS. MAYER:** *Will you get time off to do that?* **(Question2)**
> **LETTY:** *Not now, but maybe later, if they like the program.* **(Response2,Explain2)**

[1] Allan, Keith. "Meaning and speech acts." Retrieved June 28 (1998): 2004
[2] Godfrey, John, and Edward Holliman. Switchboard Dialog Act Corpus. Web Download. Philadelphia: Linguistic Data Consortium, 2008

If we analyze the above example carefully then we see that sequence of dialogues, beginning and termination is uncertain. Dialogue strings or even annotated dialogues cannot provide uniformity for their formalization. The inconsistency in the structure could be attributed to implicit policies humans follow in a dialogue. Consequently, we introduce dialogue policies connecting the different dialogue . The next sub-section elaborates on the structuring of dialogue tags into dialogue policies and then formalization through CDGS system.

3.2 Formal structure

The relation of dialogue tags between different participants are explained here through dialogue policies. It is categorized into: 'binding', 'co-occurring', and 'progressive':

- **Binding policy** denotes the relation between dialogue tags where one or more dialogue tags in the participant A's utterance fosters an immediate action on the other participant B(speech act or a response). For instance, a question tag requires an immediate response tag or another question tag, similarly a social-obligation tag demands an immediate social obligation tag. Dialogues cannot be terminated through this policy.
- **Co-occurring policy** signifies the relation between dialogue tags occurring simultaneously in one utterance or a dialogue by a particular participant. During our analysis we found that there are certain speech acts which occur simultaneously on a frequent basis in participant's utterances or dialogues such as (greet,apology), (apology,explanation) and (response,explanation). This policy cannot terminate the dialogue.
- **Progressive policy** gives the relation between dialogue tags which are open to constructive contribution or termination of a dialogue by the participant. This policy allows the participants to choose their response. For an instance if a response tag from participant A comes with the progressive dialogue policy then participant B could terminate the dialogue, ask a question, propose a concept or declare a news.

Figure 1 gives the relation between dialogue tags and the policies illustrated above for the dialogue example in Sub-section 3.1

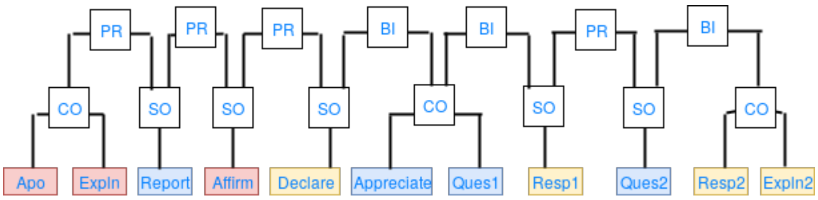

Fig. 1. Dialogue policies for dialogue sample in Section3.1.Dialogue tags are represented linearly at the bottom and different colors represent different participants. Dialogue tags are connected through dialogue relations, defined by dialogue policies. 'CO' are co-occurring policies for dialogue tags. 'SO' are dialogue tags which occur independently. 'BI' are binding policies between two or more dialogues tags.'PR' are turn-taking policies, they allow the participant to either contribute to or to end the dialogue.

For instance the dialogue tags (Resp2,Expln2) co-occur in the utterance so have 'co-occurring' policy. (Ques2) and (Resp2,Expln2) are 'binding' because (Ques2) enforces the contribution from the other participant. (Resp1) followed by (Ques2), share 'progressive' relation, because here a participant 'A' gives the participant 'B' the freedom to contribute through another question, a remark, a conclusion or termination. In the next section, we illustrate through an example, the pre-processing and formalization of dialogues.

4 Results

Based on CDGS formal Definition 1, derivation mode in Definition 2 and CDGS language in Definition 3 the dialogue can be formalized as follows:

Definition 4. *Cooperating Distributed Grammar System (CDGS) of degree n, where n=2 is a 4-tuple,*

$$\Gamma = (T, G_1, G_2, D), \tag{2}$$

- *Terminal symbol T is a subset of $\bigcup_{i=1}^{n} T_i$.*

- *For ease we have replaced the start symbol S in original Definition 1 to D where $D \in N$.*
- *The non-terminal symbols for component grammar G_1 are marked by capital symbol N_i and for component grammar G_2 by $\hat{N}_i$.*

Terminal symbols T are sets of dialogue tags in the dialogue. Non terminal symbols N and $\hat{N}_i$ are the set of all dialogue policies used in a specific dialogue sample. We replace S with D as the start symbol. Formalization is only for two participants with derivation rules denoted by G_1 and G_2 respectively. The t-mode derivation has been chosen, where the agent has to perform the derivations as long as it could contribute to the process.

To illustrate the above formalization for dialogues, below is an example of a short conversation between two participants A and B, where they meet after a long time and update each other about themselves.

A: *Hello, Sorry I'm late.* **(Greet,Apologize)**
B: *Hello, No problem at all.***(Greet,Accept)**
A: *I have been feeling very sick lately.***(Report)**
B: *What a pity. But wanted to tell you that we started the new project.***(Ack,Report)**

The above dialogue tags are structured into dialogue policies. Figure 2 shows the relation between different dialogue tags and policies. We now elaborate the formalization through CDGS.

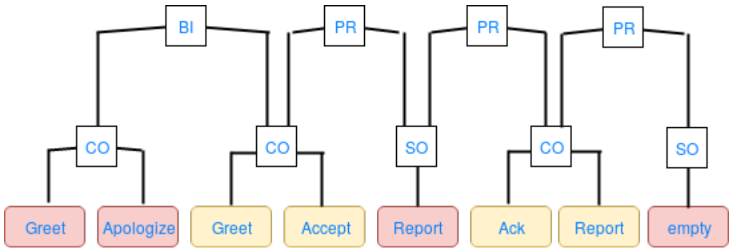

Fig. 2. Dialogue policies for dialogue example in Subsection 3.2, Dialogue tags are represented linearly at the bottom and different colors represent different participants. Dialogue tags are connected through dialogue relations, defined by dialogue policies. 'CO' are co-occurring policies for dialogue tags. 'SO' are dialogue tags which occur independently. 'BI' are binding policies between two or more dialogues tags.'PR' are turn-taking policies, they allow the participant to either contribute or to end the dialogue.

First we define the building blocks of the grammar system: terminal symbols, non-terminal symbols, starting symbol and derivation rules for grammar system. Here λ is the empty string. Consider the following CDGS:

$$\Gamma = (\{Greet, Apologize, Accept, Report, Ack\}, G_1, G_2, D), \tag{3}$$

with

$$G_1 = (\{D, CO, PR, BI\}, \{\hat{BI}, \hat{PR}, \hat{CO}, Greet, Apologize, Report\},$$
$$\{D \to CO,\ CO \to Greet\ Apologize\ \hat{BI},$$
$$PR \to PR',\ PR' \to Report\hat{PR},\ PR \to \lambda,\ BI \to \hat{CO}\}),$$

$$G_2 = (\{\hat{BI}, \hat{CO}, \hat{PR}\}, \{Greet,\ Accept,\ Ack,\ Report, PR\},$$
$$\{\hat{BI} \to \hat{CO}, \hat{CO} \to Greet\ Accept\ PR, \hat{CO} \to Ack\ Report\ PR, ...$$
$$....\hat{PR} \to \hat{CO}, \hat{PR} \to \lambda\}).$$

as production rules for grammar components G_1 and G_2 involved in this specific dialogue in Figure 2. For understanding the derivations below, refer to Figure 3.

$$CO \Rightarrow^t Greet\ Apologize\ \hat{BI}...$$
$$... \Rightarrow^t Greet\ Apologize\ Greet\ Accept\ PR$$
$$\Rightarrow^t Greet\ Apologize\ Greet\ Accept\ Report\ \hat{PR}...$$

$$... \Rightarrow^t \; Greet\ Apologize\ Greet\ Accept\ Report\ Ack\ Report\ PR...$$

$$... \Rightarrow^t \; Greet\ Apologize\ Greet\ Accept\ Report\ Ack\ Report.$$

This Figure is an illustration of two participants A and B. The dialogue formalization is represented through derivation rules in grammar components. Each of these components have their own set of terminal symbols T_i, non-terminal symbols N_i and production rules P_i where $1 \leq i \leq n$. The grammar components also share the terminal symbols T and start symbol D of the grammar system. Participant A is denoted by the grammar component G_1 and B by the grammar component G_2. Grammar components have their production rules P_1 and P_2, respectively, is a consequence of Definition1. In the below Figure 3 Grammar G_1 starts the derivation as it has the start symbol D in its set of production rules as follows ($D \Rightarrow CO \Rightarrow Greet\ Apologize\ \hat{BI}$). Grammar G_1 gives the opportunity to G_2 to contribute through $\hat{BI}$, which enforces G_2 to respond as P_2 contains the derivations for $\hat{BI}$. Similarly, G_2 gives the opportunity to G_1 now to perform derivations through PR, which allows the grammar component to contribute or to terminate. In this way grammar components keep deriving as long as one of grammar component terminates the process through ($PR \Rightarrow \lambda$ *or* $\hat{PR} \Rightarrow \lambda$) production rules.

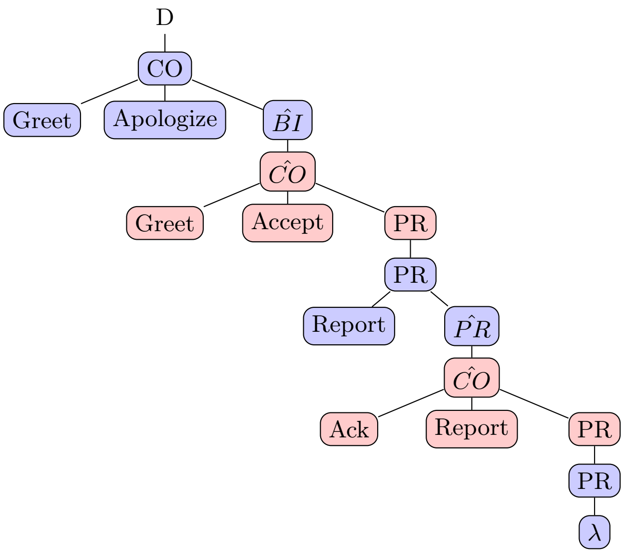

Fig. 3. Derivation tree for the example in Figure.2., Participant A's dialogue are colored blue while B is colored red.

For simplicity we replace above dialogue tags as terminal symbols to x and y in the Equation 2 and all the other annotations for component grammar, non-terminal symbols and start symbol remain the same. The grammar system looks like below:

$$\Gamma = (\{x, y\}, G_1, G_2, D), \quad (4)$$

with

$$G_1 = (\{D, CO, PR, BI\}, \{\hat{BI}, \hat{PR}, \hat{CO}, x, y\},$$
$$\{D \to CO,\ CO \to xy\hat{BI},$$
$$PR \to x\hat{PR},\ PR \to \lambda,\ BI \to \hat{CO}\}),$$

$$G_2 = (\{\hat{BI}, \hat{CO}, \hat{PR}\}, \{x, y, PR\},$$
$$\{\hat{BI} \to \hat{CO}, \hat{CO} \to xyPR$$
$$\hat{PR} \to \hat{CO}, \hat{PR} \to \lambda\}).$$

And following are the derivations in t-mode.

$$CO \Rightarrow^t xy\hat{BI}... \Rightarrow^t (xy)^2PR... \Rightarrow^t (xy)^2(x)^n\hat{PR}... \Rightarrow^t ((xy)^2(x)^n(xy)^n)$$

For Case: $PR \Rightarrow^2 \lambda$

$$CO \Rightarrow^t xy\hat{BI}, ... \Rightarrow^t (xy)^2 PR... \Rightarrow^t (xy)^2$$

For Case: $\hat{PR} \Rightarrow^2 \lambda$

$$CO \Rightarrow^t xy\hat{BI}, ... \Rightarrow^t (xy)^2 PR... \Rightarrow^t (xy)^2 x\hat{PR} \Rightarrow^t (xy)^2 x$$

For the above explained example in Figure 2, grammar system generates the following language,

$$L_t(\Gamma) = \{(xy)^2 v \mid v \; is \; x^n(xy)^n \; or \; v \; is \; x^n(xy)^n x \; for \; n \geq 0\}$$

The language generated above covers all the possible word structures generated in t-mode derivation. Here we showed through a simple example that dialogues could be formalized through CDGS. It is also important to note that pre-processing of dialogues is significant as it gives uniformity and reduces the complexity of the dialogue structure.

5 Conclusion and future work

This paper focused on formalization of dialogues through grammar systems. The dialogues were first, structured with the help of proposed dialogue policies. Then a formalization was done based on CDGS, which was elaborated through a simple example. This work led us to an understanding that dialogues are highly context dependent, speech acts alone cannot be used for modeling dialogues through grammars. A conversation consist of smaller structured dialogues, and when these segments of structure are combined with dialogue policies then they give a comprehensive structure.

Following this work we want to refine the annotation scheme and increase the number of dialogue policies to accommodate norms. It is important to conduct more experiments with CDGS and dialogues. Next, we would investigate conversational norms and their violations in human–human conversations (such as sudden-topic-change, misunderstanding and non-understanding) in order to build models for norms and conflicts in dialogues and then combine with dialogue modelling proposed here.

Acknowledgments This work received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721619.
I would also like to acknowledge Suna Bensch, Angelika Schwarz, Xuan-Son Vu, Michele Persiani, Chanh Nguyen, and Frank Drewes for their valuable comments to improving the article.

References

1. J. Austin. *How to do things with words.* Harvard University Press, Cambridge,Harvard, UP, 1962.
2. S. Aydin, H. Jürgensen, and L. E. Robbins. Dialogues as co-operating grammars. *Journal of Automata, Languages and Combinatorics*, 6(4):395–410, 2001.
3. G. Bel-Enguix, M. Grando, and M. Jiménez-López. A grammatical framework for modelling multi-agent dialogues. *Agent Computing and Multi-Agent Systems*, pages 10–21, 2006.
4. G. Bel-Enguix and M. Jiménez-López. Modelling dialogue as inter-action. *International Journal of Speech Technology*, 11(3):209–221, 2008.
5. H. Bunt. Dynamic interpretation and dialogue theory. *The structure of multimodal dialogue*, 2:1–8, 1999.
6. H. Bunt, J. Alexandersson, J. Carletta, J. Choe, A. Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary, C. Soria, and D. Traum. Towards an iso standard for dialogue act annotation. In N. C. C. Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, and D. Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, may 2010. European Language Resources Association (ELRA).
7. E. Csuhaj-Varju, J. Kelemen, G. Paun, and J. Dassow, editors. *Grammar Systems: A Grammatical Approach to Distribution and Cooperation.* Gordon and Breach Science Publishers, Inc., Newark, NJ, USA, 1st edition, 1994.
8. A. D'Ulizia, F. Ferri, and P. Grifoni. Generating multimodal grammars for multimodal dialogue processing. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(6):1130–1145, 2010.
9. P. Ioannis and L. Oliver. Combining chat and task-based multimodal dialogue for more engaging hri: A scalable method using reinforcement learning. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, pages 365–366, New York, NY, USA, 2017. ACM.
10. N. Jekaterina, D. Christian, P. Ioannis, and L. Oliver. Sympathy begins with a smile, intelligence begins with a word: Use of multimodal features in spoken human-robot interaction. *CoRR*, abs/1706.02757, 2017.
11. J. Kowtko, S. Isard, and G. Doherty. Conversational games within dialogue. In *University of Edinburgh*, 1993.

12. P. Linz. *An Introduction to Formal Languages and Automata*. Jones & Bartlett Learning, 2011.
13. E. Oishi. Austin's speech act theory and the speech situation. *Esercizi Filosofici*, 1(2006):1–14, 2006.
14. G. PĂUN. *Grammar systems: A grammatical approach to distribution and cooperation*, pages 429–443. Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.
15. P. Romney. The art of dialogue. *www.animatingdemocracy.org*, 2005.
16. J. Searle, F. Kiefer, and M. Bierwisch. *Speech act theory and pragmatics*, volume 10. Springer, 1980.
17. J. R. Searle. *Speech acts: An essay in the philosophy of language*, volume 626. Cambridge university press, 1969.
18. M. Tewari and S. Bensch. Natural language communication with social robots for assisted living. In *Robots for Assisted Living-IROS'2018 Workshop, 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018), October, 1-5, 2018 Madrid, Spain*, 2018.
19. M. Walker, R. Grant, J. Sawyer, G. Lin, N. Wardrip-Fruin, and M. Buell. Perceived or not perceived: Film character models for expressive nlg. *Interactive Storytelling*, pages 109–121, 2011.
20. M. Walker, G. Lin, and J. Sawyer. An annotated corpus of film dialogue for learning and characterizing character style. In *The International Conference on Language Resources and Evaluation*, pages 1373–1378, 2012.