Empirical Specification of Dialogue Games for an Interactive Agent

Guillaume Dubuisson Duplessis, Nathalie Chaignaud, Jean-Philippe Kotowicz, Alexandre Pauchet, and Jean-Pierre Pécuchet

LITIS - EA 4108, INSA Rouen, Avenue de l'Université - BP 8, 76801 Saint-Étienne-du-Rouvray Cedex, France
gdd@litislab.eu, {name}@insa-rouen.fr

Abstract. This article addresses the challenge of designing the communicative behaviour of an agent interacting with humans. We present a data-driven methodology based on the production of a matrix representation of a corpus from which we extract dialogue patterns. These patterns reflect the minimal units of interaction which turn out to be very attractive for dialogue modelling. We present a framework to specify dialogue games from these patterns based on the notion of social commitments. We exemplify the specification of dialogue games by implementing all the steps of our methodology on a task-oriented corpus. The produced games are validated by showing that they appropriately describe the patterns appearing in a reference corpus.

1 Introduction

Interactions between humans and software agents become commonplace in heterogeneous multiagent systems and mixed communities. However, designing agents interacting with humans is known to be a difficult task (see issues related to mixed-initiative reasoning [1]). Indeed, humans use robust communication and reasoning processes to which agents must adapt. In particular, one challenge in conceiving human-agent interaction is the *communication issue*. Dialogue management has been spotted as being a key feature [1] because it is a very efficient means of communication for people which requires little or no training to use. Besides, it is the most likely way to achieve true, mixed-initiative, collaborative systems. Nevertheless, dialogue management remains a major deadlock in Embodied Conversational Agents (ECAs) [2]. Most of them only integrate basic dialogue management processes, such as a keyword spotter within a finite-state approach or a frame-based approach.

From our perspective, the design of ECAs can be improved by analysing and modelling human-human and human-agent interactions. To this purpose, we presented a *data-driven* methodology which aims at improving the interaction capabilities of agents interacting with humans [2]. This methodology is based on the *collection* of a corpus of dialogues thanks to a user experiment designed on purpose and depending on the future agent that is being modelled. Dialogues in this corpus are turned into a *matrix representation* through an annotation step.

Y. Demazeau et al. (Eds.): PAAMS 2013, LNAI 7879, pp. 49–60, 2013.

Then, *interaction patterns* are extracted and form the basis of the *interaction model* of the agent. Considering our matrix representation, a *dialogue pattern* is defined as a sequence of annotations which arrangement occurs in several dialogues. In other words, a *dialogue pattern* is an ordered set of utterances that is frequently reoccurring during dialogues (e.g., a question/answer pair).

The main focus of this article is the modelling of dialogue patterns obtained through the methodology into dialogue games [3]. These dialogue games constitute the basic interaction units that should be next integrated into the deliberative process of an agent. To illustrate our purpose, we present the full implementation of our methodology on a corpus from collection to pattern extraction. Then we show how to model dialogue games from dialogue patterns.

Section 2 draws some links with related work, with particular attention on existing connections between dialogue patterns and dialogue management. Section 3 describes the corpus used to illustrate our approach (annotation and extraction steps). Section 4 presents the framework used to model dialogue patterns observed in the corpus and its link to dialogue management. Section 5 shows a validation of the modelled games against the reference corpus. Lastly, section 6 concludes this article.

2 Related Work

One striking observation in a human dialogue corpus is the presence of *interaction patterns* [2,3]. In dialogue modelling, this has been analysed both as being evidence of a *plan* from the interlocutors (plan-based approach) and as a manifestation of *conventional* devices used by dialogue participants (conventional approach). The first approach focuses on the *intentional structure* of the dialogue [4]. It lies on the representation and reasoning about the underlying intentions behind dialogue participant utterances. Basically, this approach considers that a speaker utterance conveys an intention that is part of a plan. Then, it is the addressee's task to infer it and respond accordingly to the underlying plan (rather than just to the speaker's utterance). This approach leads to influential results such as the TRAINS system [5] or Collagen [6]. The second approach aims at studying the *interaction patterns* without focusing on the underlying intentions. It comes from the observation that many types of utterances do not seem to be consciously emitted but rather conventionally triggered by the context (e.g., greetings). This leads to the production of rules that describe admissible sequences of utterance types. These reoccurring patterns can be studied either in terms of *dialogue grammars* [7] or *dialogue games* [3]. These two approaches are seen as opposed whereas some researchers strongly argue that they are actually *complementary* [8,9]. This is based on the fact that communication processes can be considered as joint actions between a speaker and hearers [10]. The key characteristic of a joint action is the *coordination* of *participatory actions* by two or more people. However, people cannot deliberate indefinitely in an opportunistic activity such as dialogue. Hence, coordination must stand on devices like *conventions* which are reflected by interaction patterns. Moreover, these authors have spotted some weaknesses of the plan-based approach. Namely, the plan-recognition process remains a very complicated task and is technically speaking difficult to set up [11]. Next, dialogue is an *opportunistic activity* [10]. Consequently, not only are some sentences more likely to be conventionally triggered by the situation but some sequences of acts can not be planned as well [12] (e.g., clarifications). In the light of these results, dialogue is viewed as a shared and dynamic activity that requires both high-level deliberative reasoning and low-level reactive responses.

Here, we focus on dialogue games used to explain human dialogue and to generate artificial dialogues dedicated to humans [3]. A dialogue game is a *bounded* activity with an *entry* and an *exit* where participants play a *role* (initiator or partner). *Rules* of the game specify the expected *moves* for each participant. Participants are expected to play their roles by making the moves expected by the current *stage* of the game. To the best of our knowledge, dialogue games have received few attention for practical applications in the human-computer interaction field. On a theoretical level, they have been seen as initiative-response units [8], and as structures capturing the different commitments created during dialogue [9,13]. On both practical and formal level, dialogue games have been conceived as recursive transition networks [14].

As [8,9], we propose to go towards a hybrid reactive/deliberative architecture where a theory of joint actions may serve as a 'semantics' to the interaction patterns described as dialogue games. Among the existing approaches, that of Maudet seems the most theoretically complete. It is the only approach that explicitly considers the entry and exit phases of a game as well as the multifunctionality of human dialogue [15] by differentiating two kinds of game (dialogue and communication games). This approach is the starting point for our model.

3 Implementation of the Methodology

In this section, we first describe the corpus. Next, we present the DIT^{++} annotation scheme [16] used to annotate our corpus. Then, we insist on the annotation process and its results. Finally, we say a few words about the extraction process.

3.1 Corpus

Our long-term goal is to build a *mixed-initiative assistant* [1] for information retrieval for the CISMeF system [17][1]. We use a formerly constituted corpus [18]. It is composed of dialogues of assistance on a medical information search task between a CISMeF expert and a user. Users are representative of the targeted audience of the future system since they were not medical specialists and they wanted to obtain answers about medical enquiries. Dialogues were recorded during the task where the expert and one user were facing a computer using the advanced search interface. This experimentation produced 18 dialogues between two experts and 18 volunteers. It contains circa 33 000 words and 1054 turns.

[1] CISMeF stands for "catalogue and index of French-language Health Internet resources" and is available at the URL `www.cismef.org`.

3.2 DIT^{++} Annotation Scheme

The use of language can be viewed as the performance of communicative actions since the speech act theory [10]. The DIT^{++} taxonomy takes a *context-change* (or *information-state update*) approach [19] to the interpretation of dialogue acts. Here, context can be viewed as the set of conditions which influence the interpretation or generation of utterances in dialogue [16]. This distinction is essential to take into account the fact that utterances in human dialogues often are *multifunctional* [15] and perform multiple communicative acts contrary to what traditional speech act theory says.

DIT^{++} is a *multidimensional* scheme based on a theoretically grounded notion of dimension [15,16]. Ten *dimensions* are distinguished. Among these, we could point out: the *task* (dialogue act which contributes to advancing the task or activity underlying the dialogue), *social obligations management* (dialogue acts that take care of social conventions such as greetings), *auto-feedback* (dialogue acts that provide information about the speaker's processing of the previous utterance), *time management* (dialogue acts signalling that the speaker needs a little time to formulate his contribution) and *own communication management* (dialogue acts to indicate that the speaker is editing his currently producing contribution). The taxonomy includes 88 *communicative functions* and consists of two parts: *general-purpose functions* itself consisting of three *hierarchies* (information seeking and providing functions, and action discussion functions) and *dimension-specific functions* (e.g., apology, turn grab). A *dialogue act* consists of a dimension and a communicative function. It is a context update operator construed by applying a *communicative function* to a *semantic content*. An utterance is segmented into *functional segments* which are annotated segments with zero or one communicative function per dimension.

The DIT^{++} framework proposes an application-independent taxonomy of functions for the analysis of human dialogue as well as the design of dialogue systems (and especially, the dialogue manager component). It has been shown that a multidimensional approach to dialogue annotation enables a more accurate analysis of human communication. Next, encouraging results were produced concerning the automatic recognition of DIT^{++} communicative functions by machine learning techniques. Eventually, DIT^{++} may be useful for both *interpretation* of verbal communicative behaviour, and *generation* of multifunctional utterance for ECA.

3.3 Annotation Process

Annotation was performed using the DIT^{++} taxonomy by means of the Gate annotation tool [20]. Four annotators worked on this annotation task. Each dialogue was annotated by two persons. One annotator performed the annotation for the whole corpus while the three others annotated one-third of the corpus.

The annotation process consists of two parts: (i) *Segmentation* of utterances into *functional segments* where a functional segment (FS) is "...a minimal stretch of communicative behaviour that has a communicative function. Such stretches do not need to be grammatically well-formed or contiguous, and may

Table 1. IAA for the labelling task for each dimension. R = Recall, P = Precision, F = F-measure

	Strict			Lenient			Average			
	R	P	F	R	P	F	R	P	F	**Proportion**
Allo-Feedback	0.63	0.54	0.58	0.66	0.56	0.61	0.64	0.55	0.59	1.19%
Auto-Feedback	0.77	0.8	0.79	0.8	0.83	0.81	0.79	0.81	0.8	9.44%
Contact Management	0.67	0.46	0.55	0.89	0.62	0.73	0.78	0.54	0.64	0.35%
Discourse Structuring	0.67	0.57	0.62	0.75	0.64	0.69	0.71	0.61	0.65	0.41%
Own Communication Management	0.43	0.49	0.46	0.47	0.54	0.5	0.45	0.52	0.48	5.31%
Partner Communication Management	0.86	0.91	0.89	0.86	0.91	0.89	0.86	0.91	0.89	1.09%
Social Obligations Management	0.43	0.69	0.53	0.48	0.76	0.59	0.46	0.72	0.56	1.18%
Task	0.84	0.85	0.84	0.86	0.87	0.87	0.85	0.86	0.86	68.30%
Time Management	0.75	0.8	0.77	0.81	0.86	0.83	0.78	0.83	0.8	9.93%
Turn Management	0.37	0.73	0.49	0.41	0.8	0.54	0.39	0.76	0.51	2.76%
Summary	0.77	0.81	0.79	0.81	0.84	0.82	**0.79**	**0.83**	**0.81**	

have more than one communicative function." [15]; (ii) *Labelling* of FS with zero or one communicative function per dimension.

The annotation strategy was "strictly indicator-based" [15]. Annotators were asked to mark communicative functions which are recognisable directly from features of the FS and given the context of the preceding dialogue.

3.4 Results of the Annotation Process

6343 communicative functions were produced during the annotation process with a total number of 5484 functional segments. We obtain similar results to [15] on the average number of communicative functions per FS for the selected segmentation strategy which is 1.16. A turn is in average composed of 2.60 FS thus confirming the multifunctionality hypothesis.

Since our annotation process is twofold (segmentation and labelling of functions), we performed an analysis of inter-annotator agreement (IAA) in terms of precision, recall and F-measure (F_1 score). These results come up in three categories: strict, lenient and average. These categories vary on how they consider overlapping annotations in the computation of precision and recall. The *strict* mode considers overlapping annotations as incorrect whereas the *lenient* mode considers them correct. *Average mode* takes the average of the two previous modes. The IAA for the labelling part does not take into account the taxonomic property of DIT^{++}: a `Check Question` and a `Yes/No Question` are considered as different as a `Thanking` and a `Yes/No Question` (which is obviously not true). Taxonomic metrics have been proposed for DIT^{++} but only take into account the labelling part and not the segmentation [16]. Hence, IAAs presented here are to be taken as the worst case, unless otherwise specified.

IAA on segmentation is strong since we obtain scores superior or equal to 0.93 in each mode. IAA per dimension for the labelling of communicative functions as well as the percentage of functions per dimension are presented in Table 1. First, we can realise that four dimensions stand out from the set regarding the proportion of functions: the *task* (68.30%), *time management* (9.93%), *auto-feedback* (9.44%) and *own communication management* (OCM, 5.31%) dimensions previously described. In broad outline, we could say that two-thirds of the functions are directed towards advancing the task motivating the dialogue, whereas one-third is directed towards management of the interaction process. IAA for these dimensions (except the OCM) exceeds 0.75 which we could qualify as a *reliable agreement*. The low score in the OCM dimension is due to a difference in the level of specificity of the communicative functions selected by the annotators. Actually, OCM functions form a branch in the taxonomic hierarchy. If we consider them equivalent, we reach a significant agreement of 0.67. IAA in other dimensions indicates a fair agreement that is penalised by a poor recall for the *social obligations management* and *turn management* dimensions. All in all, we obtain a global F-measure of 0.81 indicating a *reliable agreement* for the annotation task.

3.5 Interaction Pattern Extraction Process

Before addressing the extraction process, we established a *reference corpus* by randomly selecting one-third of the corpus (i.e., 6 dialogues out of 18). This reference corpus is kept as a validation basis. The 12 remaining dialogues were used for the extraction process. This process consisted in a *semi-automatic extraction* by one person of dialogue patterns. Patterns have been manually annotated in the corpus and *automatically extracted* by a tool that we design.

We focused on patterns on the *task* dimension for several reasons. This dimension is prevailing over other dimensions in terms of number of communicative functions. Next, its IAA is high (≥ 0.84). Last, functions in the predominant *time* and *own communication management* dimensions are *monologic* in the sense that they are concerned with hesitations and self-corrections from the speaker. They are not likely to be good candidates for the extraction of interaction patterns between two interlocutors.

We found 11 interaction patterns in the task dimension consisting of *general-purpose functions* and splitting into *3 categories*: information-seeking (6 patterns), information-providing (2 patterns) and action-discussion patterns (3 patterns). We mainly observed *initiative-response patterns* [8] consisting of adjacency pairs with preferred and dispreferred second pair parts. For instance, we observed patterns such as `inform/agreement` or `offer/declineOffer`. Each pattern takes the form of an *initiative dialogue act* followed by possible *response dialogue acts* with their observed proportion of occurrence. For instance, the suggestion pattern extracted starts with a `suggestion` act that can either be followed by an `acceptSuggestion` (preferred second pair part, 94.25% of the cases) or a `declineSuggestion` (dispreferred second pair part, 5.57% of the cases). Among the 11 patterns, ten are initiative-response units [8] and one is a three-step pattern, namely the correction pattern (e.g., `inform/correction/agreement`).

We manually extracted 439 occurrences of patterns in the task dimension in which 38.76% come in the information-seeking category, 27.56% in the information-providing category and 33.71% in the action-discussion category.

4 Modelling Interaction Patterns

In this section, we present our game framework based on the notion of *social commitments* as well as examples of games created from interaction patterns.

4.1 Game Framework

Social commitments are to be distinguished from the *private states* of agents such as belief and desire. In fact, social commitments are commitments that bind a speaker to a community [21]. These commitments are *public*. They are stored in a *commitment store* that is part of the public layer of the information state of a dialogue system. Social commitments are distinguished into *propositional* and *action* commitments. The former concerns commitments that do not deal with future action such as when A says "Paris is the capital of France" whereas the latter concerns commitments dealing with future action like "I will come at your place this evening.". We discern *dialogical* commitments from *extra-dialogical* commitments.

We express commitments as predicates with an arity of 4. An extra-dialogical propositional commitment takes the form C(x,y,p,s) meaning that x is committed towards y about proposition p. s refers to the state of the commitment that we explain below. The previous example produces the following commitments: C(A,y,$capital(france, paris)$,Crt) meaning that A is committed towards y to $capital(france, paris)$. For the sake of readability and since we are only considering dialogue involving 2 partners, we will ignore the second argument which specifies the partner. A *dialogical commitment* is contextualised in a game g and takes the form C_g(x,α,Crt) (meaning that x is committed towards y to do action α in the context of game g). Furthermore, it is possible to compose actions in commitments with the choice ($\alpha|\beta$), the conditional statement ($\alpha \Rightarrow \beta$) meaning that β will occur if α does and the persistent conditional action ($\alpha \overset{*}{\Rightarrow} \beta$) meaning that β will occur each time α does.

Finally, three operations are considered on commitments: *creation*, *satisfaction* and *cancellation*. This leads to the following five states that are possible for a commitment (inspired by [22]): (i) *inactive* (Ina) which is the state by default, (ii) *created* (Crt) which is the state right after the creation of the commitment, (iii) *cancelled* (Cnl) which is the state after a cancellation of a created commitment, (iv) *fulfilled* (Ful) which is the state after a satisfaction of a created commitment, and (v) *failed* (Fal) which is the state if a tentative to socially create the commitment has failed.

Our formalisation of games refines the one proposed by Maudet [9] by the addition of failure conditions, game-specific effects of dialogue acts and coherence constraints on the semantic contents of acts. Games represent *conventions* between interlocutors and are *shared bilateral* structures. They can be divided into two categories: *dialogue games* and *communication games*. Dialogue games are a particular kind of *joint activity* [10] *temporarily* activated during the dialogue for a *specific goal* (e.g., information-seeking game, action-seeking game, etc.).

Communication games are dedicated to more general interaction processes (e.g., mutual understanding, turn-taking, etc.) and are *permanently* activated.

A dialogue game is a 5-tuple $\langle \mathrm{En}_g, \mathrm{S}_g, \mathrm{F}_g, \mathrm{R}_g, \mathrm{Eff}_g \rangle$ where (i) En_g is a pair $\langle \mathrm{En}_g^I, \mathrm{En}_g^P \rangle$ defining the *entry conditions* of the *initiator* (En_g^I) and of the *partner* (En_g^P) which are conditions that must hold at the beginning of the game, expressed in terms of extra-dialogical commitments, (ii) S_g is a pair $\langle \mathrm{S}_g^I, \mathrm{S}_g^P \rangle$ defining the *success conditions* of the *initiator* (S_g^I) and of the *partner* (S_g^P) which are conditions that define a state of success in terms of extra-dialogical commitments, (iii) F_g is a pair $\langle \mathrm{F}_g^I, \mathrm{F}_g^P \rangle$ defining the *failure conditions* of the *initiator* (F_g^I) and of the *partner* (F_g^P) which are conditions that define a state of failure in terms of extra-dialogical commitments, (iv) R_g is a pair $\langle \mathrm{R}_g^I, \mathrm{R}_g^P \rangle$ defining the *rules* of the *initiator* (R_g^I) and of the *partner* (R_g^P) which are specifications of dialogue rules expressed in terms of dialogical commitments where constraints on the semantic contents of dialogue acts could be specified (as in [8]), and (v) Eff_g is a pair $\langle \mathrm{Eff}_g^I, \mathrm{Eff}_g^P \rangle$ defining the *contextualised effects of dialogue acts* in terms of the creation of *extra-dialogical* commitments for the *initiator* (Eff_g^I) and for the *partner* (Eff_g^P). S_g and F_g are conditions that motivate an exit of the game. As for communication games, their structures come down to *dialogue rules*.

As previously mentioned, game playing can be seen as a joint action [8]. Hence, game has an *entry*, a *body* and an *exit* [3,10]. Therefore, dialogue games need to be established. The communication game of *contextualisation* is dedicated to this task: we currently consider a simple version of this game consisting of two proposition phases (one for the entry and one for the exit) that must be explicitly accepted by the partner as in [13].

4.2 Examples of Games from Dialogue Patterns

We now present one dialogue game and one communication game created from the patterns extracted during the extraction process. For the sake of readability, we skip the create operation in conditional statements ($\alpha \Rightarrow \mathrm{create}(x,\mathrm{C_g}(y,\beta,\mathrm{Crt}))$ is equivalent to $\alpha \Rightarrow \mathrm{C_g}(y,\beta,\mathrm{Crt})$). Furthermore, dialogue acts take the form: $f\,(s,c)$ where f is the communicative function of the dialogue act, s the speaker that produces this act and c the semantic content. Hence we do not specify the dimension since we focus on the task one.

The choice question dialogue game is a representative example of all aspects of our framework and is presented on Table 2. We took the simplified semantic representation for choice questions from Larsson [23] but other semantics may be applicable (e.g., see [8]). A choice question q is viewed as an alternative question between propositions belonging to a set (e.g., "Would you like to add the keyword paludism, therapeutic or vaccine?") and takes the form: $\{?p_1, ?p_2, \ldots, ?p_n\}$. Three predicates relate to this kind of question. First, the **resolves** (p,q) predicate of arity 2 is true when the proposition p resolves q. Here, a resolving proposition is p_i with $1 \leq i \leq n$ (e.g., "Vaccine!"). Next, the **relevant** (p,q) predicate of arity 2 is true when the proposition p is about q.

Table 2. The Choice Question Game

Game g	Initiator (x)	Partner (y)
En_g		C(y,p,Ina) with **resolves** (p, q)
S_g	C(y,p,Crt) with **resolves** (p, q)	C(y,p,Crt) with **resolves** (p, q)
F_g	C(y,**fail** (q),Crt)	C(y,**fail** (q),Crt)
R_g	choiceQuestion(x, q)	choiceQuestion(x, q) ⇒ C_g(y,answer(y, p)|answer(y, s)|execNegativeAF(y, q),Crt) with **resolves** (p, q), **relevant** (s, q) answer(y, s) ⇒ C_g(y,answer(y, p)|answer(y, r)|execNegativeAF(y, q),Crt) with **resolves** (p, q), **relevant** (s, q), **relevant** (r, q)
Eff_g		answer(y, s) ⇒ C(y,s,Crt) execNegativeAF(y, q) ⇒ C(y,**fail** (q),Crt)

Table 3. The Agreement Communication Game

α	β
f(x,p)	C(y,agreement(y, p)|disagreement(y, p),Crt)
agreement(y, p)	C(y,p,Crt)
disagreement(y, p)	C(y,$\neg p$,Crt)

f ∈ {Inform, Answer, Agreement, Disagreement, Correction, Confirm, Disconfirm}

Here, a relevant but not resolving proposition would be $\neg p_i$ with $1 \leq i \leq n$ (e.g., "Not therapeutic."). Eventually, the **fail** (q) proposition indicates that an answer cannot be found by taking into account the current information state. The same simplifying assumption than [23] is done that the *resolves* and *relevant* relations are shared between interlocutors. Entry conditions specify that the partner must not already be committed on a proposition that resolves the question. Success conditions are reached when the partner is committed to a proposition that resolves the question. Failure conditions establish that a state of failure of the game happens when the partner is committed to the fact that he is not able to find an answer. Rules specify that the initiator is committed to play a `choiceQuestion` act. Once it is done, the partner is committed to play `answer` moves or an `execNegativeAF` move. This latter expresses the fact that a resolving answer can not be found. The rules state that the partner can give as many relevant answers as he can, and only one resolving answer. This enables to cover cases like "Not therapeutic. And yes vaccine !" where the first answer is relevant and the second is resolving. Finally, the effects precise that, *in the context of this game*, playing an `answer` move commits the speaker to its semantic content and playing an `execNegativeAF` move commits the speaker to **fail** (q).

Eventually, we present a communication game: the agreement game (Table 3). Rules take the form: $\alpha \overset{*}{\Rightarrow} \beta$. This game specifies that an addressee of an information-providing dialogue act can play an agreement or disagreement move after receiving this act. If he agrees, he is committed to the semantic content of the information-providing act, else he is committed to its negation. For example: "– We got 115 articles. (Inform) – Exactly (Agreement)"[2].

[2] All examples comes from our corpus and were translated from French to English.

Table 4. Results of the evaluation process (percentage of exchanges that fit a game)

Suggestion	Request	Offer	Agreement	Check Q	Posi-check	Nega-Check	Y/N Q	Choice Q	Set Q	Correction
84.85%	72.73%	78.13%	100.00%	68.92%	100.00%	66.67%	66.67%	87.50%	88.57%	100.00%

4.3 Discussion

An evident limitation of our work is that we only defined minimal dialogue games. The issue of how bigger interaction pattern emerges from these minimal games is left unaddressed for the moment. However, they may result of *compositions* of minimal joint action [8,9,10,13] such as *embedding*, *chaining* and *pre-sequencing* [10]. This problem is tightly connected to the idea of a *contextualisation game*. This is an area left for future work. However, we could point out that the minimal nature of the dialogue games that we defined associated with composition rules make the dialogue game formalism *flexible* and *reusable*.

The issue of the integration of dialogue games in the deliberative process of an agent is open. The model that we propose is a *normative* one. It can then be viewed as an *independent module* regulating the dialogical capabilities of an agent on both *interpretation* and *generation* of dialogue acts (thus simplifying the process of intention recognition). On the *interpretation* level, dialogue games make it possible to produce a set of conventionally expected dialogue acts whereas, on the generation level, they give a conventional motivation for the production of acts. The *independence* of the module is reinforced by the declarative nature of the formalism which is disconnected from the private states of the dialogue participants.

5 Validation

We conducted a study to validate the *structural property* of our model based on the reference corpus. Two voluntary persons of our laboratory were given the whole game definitions and were asked to annotate each *exchange* (i.e., a sequence of two (or more) functional segments produced by different speakers) with a game *if possible*. It was generated for each dialogue and for each game a ratio between exchanges that fit the game and those that do not.

Results of the validation process are presented in Table 4. The global conclusion is that 83.43% of the 350 exchanges of the reference corpus match a game that we defined. If we go into details, we see that scores are all beyond 66% indicating a reliable adequacy between dialogues being modelled and our games.

In addition, we investigated the 16.57% of mismatch cases and we identified 2 main categories: those related to the *inter-game structure* and those related to the *intra-game structure*. Mismatches related to the first category are twofold. On the one hand, the partner deliberately ignores the utterance of the initiator in approximately 10% of the cases. It is what Clark calls a *withdrawal from the joint project* [10] and is exemplified by: "– Would you like me to open this document ? (Offer) – Document 13 was interesting. (Inform)". On the other hand, the partner opens an *embedded game* that obsoletes the *parent game* in around

45% of the cases. It is illustrated by this example: "– Is there any documents that would suit you? – Well, is there the keyword 'prevention'? – No, we did not add it.". Mismatches related to the second category can also be subdivided into two parts. First, the predominant case (approx. 35%) is when the partner seems to take a shortcut in a bigger interaction patterns. It includes cases that has been called *indirect speech act* and are mostly appearing with Request game and CheckQuestion game (for instance: "– Can you formulate your information need or not? (CheckQuestion) – 'What is known about the evolution of current treatments of migraine' (Inform)"). Last cases appear when the partner alters its response from what is expected by the pattern to something that he is able and willing to comply with (approx. 10%), called an *alteration of the joint project* [10]. This can be illustrated by the following example: "– Then, you do not know the equivalent of Zomig? (CheckQuestion) – This is a product family.".

6 Conclusion and Discussion

We presented a data-driven methodology based on the study of human interactions to address the challenge of designing efficient human-agent interaction. This methodology is based on the constitution of a *matrix representation* of a corpus in several steps (collection, annotation, pattern extraction) that makes it possible to extract *interaction patterns*. The issue of the interest of dialogue patterns for dialogue management was raised. It turns out that minimal dialogue patterns can be viewed as the minimal unit of interaction, and therefore are very attractive to dialogue modelling. We then presented a framework to model such dialogue patterns based on theoretical work about *dialogue games*. This framework stands on the notion of *social commitments* that permits to specify dialogue games independently of the architecture of the agent. Hence, it makes it possible to envision the implementation of this model as a separate module with a normative role. We exemplified the specification of dialogue games from dialogue patterns by implementing all the steps of our methodology on a task-oriented corpus. We validate the games that we modelled against a reference corpus. The modelled games were able to describe appropriately the exchanges of the reference corpus.

Many interesting perspectives are possible. We focus our pattern study on one dimension. Future work involves the extraction and modelling of patterns on other dimensions. The model would include *multidimensional dialogue games* that could help producing multifunctional utterances. As stressed in this article, more work needs to be done on composition and contextualisation of dialogue games as well as implicit phenomena that appear in human dialogue.

References

1. Tecuci, G., Boicu, M., Cox, M.T.: Seven aspects of mixed-initiative reasoning: An introduction to this special issue on mixed-initiative assistants. AI Magazine 28(2), 11–18 (2007)
2. Ales, Z., Dubuisson Duplessis, G., Serban, O., Pauchet, A.: A methodology to design human-like embodied conversational agents. In: AAMAS 2012 Workshop, Human-Agent Interaction Design and Models (2012)

3. Mann, W.C.: Dialogue games: Conventions of human interaction. Argumentation 2(4), 511–532 (1988)
4. Grosz, B., Sidner, C.: Attention, intentions, and the structure of discourse. Computational Linguistics 12(3), 175–204 (1986)
5. Allen, J., Ferguson, G., Miller, B.W., Ringger, E.K., Sikorski-Zollo, T.: Dialogue systems: From theory to practice in Trains-96. In: Handbook of Natural Language Processing, pp. 347–376 (2000)
6. Rich, C., Sidner, C.: COLLAGEN: a collaboration manager for software interface agents. User Modeling and User-Adapted Interaction 8(3), 315–350 (1998)
7. Polanyi, L., Scha, R.: A syntactic approach to discourse semantics. In: Proceedings of the 10th International Conference on Computational Linguistics, pp. 413–419 (1984)
8. Hulstijn, J.: Dialogue models for inquiry and transaction. PhD thesis, University of Twente (2000)
9. Maudet, N.: Modéliser les conventions des interactions langagières: la contribution des jeux de dialogue. PhD thesis, Toulouse 3 (2001)
10. Clark, H.H.: Using language, vol. 4. Cambridge University Press (1996)
11. Cohen, P.: Survey of the state of the art in human language technology, pp. 204–210. Cambridge University Press, New York (1997)
12. Pulman, S.: The TRINDI project: Some preliminary themes. In: Proceedings of the Twente Workshop on Language Technology (1998)
13. Maudet, N., Chaib-draa, B.: Commitment-based and dialogue-game based protocols: new trends in agent communication languages. The Knowledge Engineering Review 17(2), 157–179 (2002)
14. Lewin, I.: A formal model of conversational game theory. In: 4th Workshop on the Semantics & Pragmantics of Dialogue (2000)
15. Bunt, H.: Multifunctionality in dialogue. Computer Speech and Language 25(2), 222–245 (2011)
16. Bunt, H.: The DIT++ taxonomy for functional dialogue markup. In: AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts, pp. 13–24 (2009)
17. Darmoni, S., Leroy, J., Baudic, F., Douyre, M., Piot, J., Thirion, B.: Cismef: a structured health resource guide. Methods Inf. Med. 39, 30–35 (2000)
18. Loisel, A., Dubuisson Duplessis, G., Chaignaud, N., Kotowicz, J.P.: A conversational agent for information retrieval based on a study of human dialogues. In: Filipe, J., Fred, A. (eds.) Proceedings of ICAART, vol. 1, pp. 312–317 (2012)
19. Larsson, S., Traum, D.R.: Information state and dialogue management in the TRINDI dialogue move engine toolkit. Natural Language Engineering 6(3&4), 323–340 (2000)
20. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Aswani, N., Roberts, I., Gorrell, G., Funk, A., Roberts, A., Damljanovic, D., Heitz, T., Greenwood, M.A., Saggion, H., Petrak, J., Li, Y., Peters, W.: Text Processing with GATE (2011)
21. Singh, M.P.: Social and psychological commitments in multiagent systems. In: AAAI Fall Symposium on Knowledge and Action at Social and Organizational Levels, pp. 104–106 (1991)
22. Chaib-Draa, B., Labrie, M.A., Bergeron, M., Pasquier, P.: Diagal: An agent communication language based on dialogue games and sustained by social commitments. Autonomous Agents and Multi-Agent Systems 13(1), 61–95 (2006)
23. Larsson, S.: Issue-based dialogue management. PhD thesis, Department of Linguistics, Göteborg University (2002)