Towards a Conversational Agent Architecture to Favor Knowledge Discovery in Serious Games

Francesco Bellotti
DIBE, University of Genoa
Via Opera Pia 11a
ITALY
16145 Genova
franz@elios.unige.it

Riccardo Berta
DIBE, University of Genoa
Via Opera Pia 11a
ITALY
16145 Genova
berta@elios.unige.it

Alessandro De Gloria
DIBE, University of Genoa
Via Opera Pia 11a
ITALY
16145 Genova
adg@elios.unige.it

Elisa Lavagnino
DIBE, University of Genoa
Via Opera Pia 11a
ITALY
16145 Genova
elilav@elios.unige.it

ABSTRACT

"Conversational Agents" (CAs) are virtual characters controlled by the computer and able to dialogue with users in natural language. CAs are usually employed in virtual world applications, such as for training, gaming or advertising, in order to increase the situation realism and user involvement.

We are implementing a natural language interaction system with a clear focus on instructional dialogues aimed at favoring a player's acquisition of knowledge on specific topics through interaction with specialized NPC. Another key requirement for the system is easy and efficient writing and maintenance of the texts, also by author with no specific expertise on computational language technology.

The system we propose includes a strategy and a tactical level. The former is responsible for managing the high-level aspects of the conversation, while the latter responds to the player's queries by relying on an original combination of an essential syntactical analysis and an ITF-IDF-based procedure.

The technique specifically targets a (educational) gaming application domain. However, we believe that the technique could be extended to meet the requirements of other application domains.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing - *Discourse and language parsing and understanding.*

K.8.0 [**Personal computing**]: General – *Games.*

General Terms

Serious Games, Natural Languge Processing.

Keywords

Serious games, conversational agents, dialogue system, Non Player Characters, natural language interaction

Short presentation, ACE'2011 - Lisbon, Portugal

1. INTRODUCTION

Serious Games (SGs) are getting a growing interest in the technology enhanced education research community. One of the key advantages is their immersive-ness and high realism that allow users to make experiences. A lot of research has still to be done, in particular on how to help users in making effective experiences without getting lost in the learning space. Human-Interaction modalities play an important role in this regard, since natural modalities increase realism and favor access to contents. This paper focuses on speech interaction techniques, that open interesting opportunities for learning and gaming.

The "Conversational Agents" (CA) or automatic Dialog Management Systems (DMS) employ texts, speech and/or images in order to allow a dialogue with users in natural language. CAs are usually employed in automatic answering systems and, now increasingly, in games, with the aim to enhance a situation's realism and user's involvement by allowing an appealing and pleasant access to information. CAs are usually implemented as virtual characters, that take part to a narrative, playing a specific role, providing contents to introduce "back-stories", assigning tasks to players, rewarding performance and, in general, giving information.

Some advanced games and SGs are populated with CA-enabled Non-Player Characters (NPCs) that live autonomously and perform goal-oriented activities, interacting with each other as people do in daily life. These virtual environments are very promising as settings for educational activities, in particular for the cases in which immersion in real-life situations is particularly appropriate, such as for learning a language, exploring an environment (e.g. [1]), interacting with people of different cultures (e.g., [2]) or attitudes (e.g. a criminal gang).

[3] presents Tactical Iraqi, a videogame that is designed to accelerate a learner's acquisition of spoken Arabic to assist in the rapid deployment of soldiers into volatile tactical situations, and Virtual Iraq, a virtual reality simulation intended to lessen the effects of Post-Traumatic Stress Disorder among combat veterans. The authors argue that these Virtual Reality worlds enable participants to develop forms of cultural literacy that soldiers would otherwise be lacking.

US military personnel engaged in civil affairs and peacekeeping operations that are making extensive use of game-base situated cognition courses prior to overseas deployments report that the linguistic and cultural competence that they gained from the courses had a significant impact on the effectiveness of their operations [4]. For instance, a US Marine unit recently

attributed its success (completion of a tour of duty in Iraq without a single combat casualty) to its enhanced ability to develop relationships and establish rapport with the local people. The basis of the approach relies on the idea that "intercultural competence involves awareness of the differences between cultures, knowledge about the beliefs, values, and practices of other cultures, and the skill to apply that knowledge effortlessly and effectively in interpersonal interactions". In their approach – which is common to other military applications [3, 2] -, "videogame technology is used to create virtual worlds populated by non-player characters that speak and understand the target language, and behave in accordance with the norms of the culture. Conversational artificial intelligence technology enables learners to engage in spoken conversations with the non-player characters. Learners must speak the target language, and behave in a culturally proficient manner, in order to succeed at the game" [4].

A literature review of the methods for automatic dialogue management shows that complex sophisticated systems have been developed for interactive narrative. Outstanding examples include Façade [5]. However, it is apparent that these systems are very complex to author, manage and maintain. This probably may explain why they are not widely used in common applications. Moreover, it may be argued that such systems do not perfectly correspond to the usual requirements of serious games, that typically involve very quick interactions among a player's avatar and NPCs. Thus, we have started our work on the basis of the following specifications:

- Every NPC embodies certain well-defined units of knowledge. This is realistic. For instance, a real policeman is able to give street directions, while a art expert may answer questions about artworks and heritage.
- NPCs are aimed at answering a player's questions, in particular aimed at specific knowledge acquisition.
- The context has a clear importance in a serious game, and can be usefully exploited to facilitate the dialogue. Sample elements of the context are the appearance of the NPC (e.g., uniform, cloths, sex, age), its position, role and game level in which it appears.
- Each NPC's knowledge and capabilities should be easy to edit and maintained also by people with no specific expertise in automatic dialogue algorithms.

These specifications are not general, but address a specific target of human-content interactions. In particular, this kind of NPCs may be employed in short dialogues where a player could get knowledge clues about a specific topic. This is a typical need of some game types, such as adventures, and in particular serious games. Thus, such NPCs would implement a specific serious game mechanism combining user knowledge acquisition and natural, pleasant interaction. In general, serious games are not interactive dramas, such as the well known Façade [5], where dialogue is the key element. In serious games, the user text/speech interaction is limited.

Reading text in games – whose entertainment level is anyway a reference for successful serious games as well – is generally perceived as boring - and, as a matter of fact, it is avoided, if not perceived as strictly necessary - since it slows down the pace of the gameplay. Players are drawn by the plot and/or the desire to achieve the targets, and they usually do not like reading long texts, also because of possible time constraints. The case of dialogs is different, since they are interactive and they may be directly useful to advance in the game. Anyway, it is important to keep into consideration that the target is short dialogues finalized to the extraction of specific knowledge from the NPC.

We have developed a CA architecture and are testing it in the context of a serious game (SG) for promoting safe maritime behavior in coastal areas (SeaGame [6]) and one for cultural heritage promotion (TiE, [7]). In this SG context, the NPC's target is both to provide useful information on the basis of the user request and ask player questions in order to stimulate reasoning (e.g., about facts and context). Our current implementation is text-based. However, a speech-to-text module could be added in order to further increase natural interaction.

The remainder of this paper is organized as it follows. Section 2 shows the background of the study and related work. Section 3 presents the system, and a use case is shown in section 4, while section 5 draws the final conclusions.

2. BACKGOUND AND RELATED WORKS

Natural language speech interaction represents a key opportunity for simulation and gaming. Façade, Scribblenauts and Bot Colony are major samples of games that have deeply exploited Natural Language Processing (NLP), while the diffusion of full natural language games looks to be "only a matter of time" [7]Several techniques and tools have been used for NLP in gaming and human-behavior simulation. The Artificial Intelligence Mark-up Language (AIML) is an XML based authoring tool that allows developers to specify rules (they are actually called categories) where patterns of words (input) correspond to a template (output clause) that is executed when the player's input matches the corresponding pattern. Templates may implement recursive substitution. AIML is very low level, and its use of wilde-cards and recursion implies self-modifying input, which is very problematic, in particular if we consider the maintenance of the rules. A. L. I. C. E. [8], a high-performance chatbot programmed in AIML, involves 1120,000 rules, which makes the writing of new ones awkward. Also, the use of the simple wildcard is prone to false positives (unwanted matches).

Façade is an interactive drama based on 800 template rules. Façade does not aim at fully understanding the player's input, but to map sentences into discourse acts, such as agree, disagree, criticize, flirt. As an intermediate step, in order to use the underlying Jesse rule system, sentence's words are mapped into intermediate facts. In order not to break the flow, the story moves on some way, even when the system is not able to interpret the input. Façade introduces the OR operator to match sets of words and the NOT to exclude the presence of words. It also uses WordNet as a set of synonyms, but not as an ontology.

Chatscript is a recent language that is the basis of Suzette, the chatbot that won the 2010 Loebner Competition (Turing test) [7]. ChatScript has been designed to favor the extraction of meaning from sentences and provide a syntax that can be easily implemented, thus enhancing maintainability of code. ChatScrpt introduces new powerful rule types for language pattern matching, considers punctuation in text analysis and has a clear syntax for nested dialogues. Chatscript supports the matching of canonical and original forms of verbs and the lumping of words into "concepts", also exploiting WordNet ontologies. This allows approximating patterns of meaning. Words from the player can be

captured and used in the bot's response. Beside the AND, OR and NOT pattern operators it includes also the possibility of specifying if one or more words are at the beginning or end of a sentence, or if they can be present in any order. Ranged wildcards give the possibility of matching an exact number of words. It is possible to define variables and assign values to system variables (e.g., about the time).ChatScript does reflection, tracking its decisions. The control script strategy allows also to track the player's profile and to insert transitional sentences to change topics. Rules are subdivided into topics, accessible via keywords, thus making large collections of rules efficient to search and simplifying the authoring of orthogonal content.

The above represent the state of the art for gaming and human-behavior simulation, and looks particularly accurate and valuable in the perspective of building a realistic general-purpose dialoguing system. However, in our case of developing a system able to provide information to a player from a specific knowledge basis, it seemed to us appropriate to explore other techniques from the information search and retrieval research domain. In particular, we could state the problem as it follows: we have a list of possible answers of the bot and we need an algorithm that, at each player's question (query), is able to rank each possible answer (document) so that the most appropriate one can be selected.

Thus, we implement a question answering (QA) module, which addresses the problem of finding answers to natural language questions. The common protocol of a QA system includes answer extracting & meriting and confidence calculation & sorting. An example of QA dialogue manager system is Malin, which is used in BirdQuest [9]. Differently from that work, our application context is not general web info searches, but knowledge transmission in a SG.

The problem of implementing a dialogue system may be tackled also in terms of text retrieval. [10] provides an excellent overview of Text Retrieval (TR) and NLP techniques. The classic approach consists in representing the query and the documents as vectors over the collection vocabulary and rank documents on the basis of vector similarity. Supposing no inter-word dependences, the algorithm computes key statistics, among which Term Frequency (TF) - a measure of term salience estimated as the relative frequency of that term in a document - and Inverse Document Frequence (IDF) – the inverse of the frequency of the term among all the documents in the collection. The higher the value, the more important should be that term for identifying the correct document. Beside the vector-similarity approach, there is the probabilistic approach to TR. For instance, the BM25 probability model adds to the TF-IDF basic statistics several parameters for tuning on development data [11].Another approach, similar to TF-IDF, relies on language modeling (LM) [12], whose advantage is expected from the possibility of exploiting theoretical foundations and estimation techniques on speech recognition.

Enhancing TR with NLP techniques has long expected to provide benefits, even if results until now have not proven to be significant [13]. A key, and obvious, step in this direction is given by the syntactical analysis, since syntax is generally considered important for interpretation. As a matter of fact, parsing, which is the identification of a statement's syntax, has been useful in various practical cases [13]. The Penn Treebank (PTB), based on two million manually annotated texts and conversations, is a major reference for current statistical parsing [14]. Another NLP-related aspect that may enhance TR concerns the lexical relations, which can exploit manual or statistically-based work [15]. The most important factors that generate ambiguities in a text, in fact, are lexical relations such as synonymy, polysemy, anaphore, metonymy, etc [16]. Hyperonymic/hyponimic relations place the lexicon in a vertical disposition, which allows creating a connection between, say, museum and art gallery, where art gallery is a hyponym of the hyperonym museum. Similarly, synonym relations define horizontal connections among words and/or terms. A lexical analysis allows identifying semantic concepts, which are key for a comprehension of the text.

For such analysis, our system uses WordNet, one of the most known lexical databases for the English language. It collects words into synset, sets of synonyms. WordNet provide also definitions and records some of the most common semantic relations between these synsets (hyperonym/hyponym and holonym/meronym for nouns; hypernym, troponym and entailment for verbs; related nouns, similar to and participate of verb for adjectives; root adjectives for adverbs). WordNet also provides the polysemy count and distinguishes grammar categories.

Beside lexical analysis, our approach intends to perform a simple syntactical analysis, that captures the most relevant structure of the target dialogues, and can be easily managed by authors.

3. THE SYSTEM

As stated in the introduction, the design of the system was driven by three major requirements about the NPC.

- Every NPC embodies certain well defined units of knowledge
- NPCs are finalized at specific knowledge acquisition by the player
- The context as a key factor
- Each NPC's knowledge and capabilities should be easy to edit and maintain

This means that the system is not general purpose but application-specific. This restriction of the field is determined by the specificity of the target application (i.e., serious games) and the need to keep the content authoring framework as simple and usable as possible by requiring only the strictly necessary contents. However, our approach introduces some features that could be exploited also to enhance the CA and NLP technologies in general.

Following an established trend of the robotic system theory, the designed CA system involves a strategic level and a tactical level. The first one – implemented by the Strategy Manager (SM) module - is responsible for the overall management of the dialogue, also considering game objectives; the second one – implemented by the Tactics Manager (TM) module - reacts to the player's text input.

3.1 The Strategy Manager

The SM component involves a number of abilities, such as the following:

- Manage the beginning and the end of the conversation. For instance, the SM could welcome the player and invite him to

speak, introduce itself, answer to a welcome, etc. An important first step could be the presentation of an overview of the knowledge set available by the NPC

- Change the knowledge set of the tactical level (e.g. considering the level of player or the emotional states)
- Intervene during the conversation, providing help/indications if the conversation is lagging or should terminate as all (or a sufficient level of) the NPC's knowledge has been told to the player.
- Change the language register of the conversation, based on the player's interactions.

From the point of view of the implementation, the SM module consists of the following sub-modules.

The Position Manager, that usually makes the NPC get closer to the player's avatar when it enters a given triggering area. An indifference behavior may be employed, for higher difficulty game levels, so that the player has to "find the right NPC".

The Welcome/Farewell Manager, that manages a simple welcome dialogue. In the welcome dialogue, the NPC may provide (as a sort of help) an overview of the knowledge available (questions it can answer). The Welcome Manager (WM) exploits a set of questions predefined by the author, and outputs the most likely answer according to a dialogue processing technique based on lemmatization and a subsequent ITF-IDF procedure. The ITF-IDF aims at identifying, in the set of the predefined possible user questions, the closest one (in the author-defined set of questions) according to vector similarity. Then, the corresponding author-defined answer is output. The dialogue may be initiated by the NPC as well, which presents itself and invites the player to make questions.

The Conversation Monitor (CM) is responsible for continuously observing and assessing the conversation that is being carried out at tactical level. If the conversation is difficult – for instance, if a certain number of (subsequent) user input sentences have not been recognized, or if the user is slow in inputting text, or there are still some topics not covered in the conversation -, the CM intervenes with some author-defined help/motivation texts. The CM is implemented as a rule-based expert system. Rules are simple to implement, in order to allow composition by authors themselves. Thus, they are implemented as boolean conditions where parameters are given by number of unrecognized texts/covered topics, number of correct interactions, time of interaction. The CM is also aware of the advancement of the player with respect to the knowledge plan (topics to be covered). Thus beside providing hints to possible missing questions on the players parts (see above), it can also trigger the Welcome/Farewell Manager for letting the player know that it has said everything it should have.

The User language Register Detector performs a simple word frequency analysis in order to estimate the language register of the user. If a specific register is detected, then the CM either outputs some comments (e.g., in case of bad words), and/or adapts its language (e.g., by simply adding gentle starting or closure phrases). Detection is performed through semantic analysis, in particular exploiting WordNet concepts.

The simple dialogue management procedure presented in the Welcome Manager description (i.e., lemmatization + ITF-IDF analysis of a user input with a document corpus) is the basic processing unit that we use, also in the following, with different inputs and targets. We call it Fundamental Statistical Text Procedure (FSTP).

The procedure is parametric, in the sense that the author can define a threshold for recognition. If the procedure cannot associate the response with a confidence level above the threshold, then the system outputs a statement saying that the phrase has not been recognized. At application level, the CM is the responsible for managing these situations, intervening proactively. For instance, after two times the system has informed the player that his text was not recognized, the CM may give help to the player about its knowledge topics, or invite player to go on with the game, giving him a clue even if it was not deserved.

3.2 The tactics level

The TM is responsible for directly reacting to the user input. Here we use a statistical approach, based on the above defined FSTP procedure. However, not the whole corpus of possible answer documents is analysed in order to match the player's question. This is the major innovation proposed in this work and comes from our analysis of previous works on CAs in SGs [6]. The objective is both to reduce the response latency, which is a key requirement in a real-time system such as a videogame, and, overall, to enhance the search by exploiting syntactic cues. The syntactic analysis simply consists in finding the type of question asked. Consequently, the corpus of knowledge of each NPC is structured in answers to different types of questions. The type of a question is individuated by its introducing interrogative adjective/pronoun (IAP). Here are the possible types of questions: Where; When (also including "at what time"); Who; Why; What and which (pronoun); What and which (adjective); How; How much; How many; Interrogative Adjective/Pronoun Free (IAPF) questions.

We expect that this question-based schema should be appropriate for presenting information for a wide coverage of a topic.

Considering the authoring point of view, authors have to write the answers, that have to be tagged only by the corresponding IAP, in order to speed-up the authoring process. An answer may be tagged with more than one tag (e.g., see interaction b in example 2), which allows sparing authoring time.

At runtime, when the user inputs a text, the actual question statement is extracted. If a question cannot be detected (no question mark, nor IAP), the CM invites the player to ask questions. Otherwise, the IAP is singled out and the question input is first lemmatized, and then processed with ITF-IDF with respect to the corpus represented by all the possible answers tagged with the same pronoun. The words inside every document (answer) are not weighted the same. The first sentence in the document is weighted more, as it is expected to correspond to the question. While the rest of the document is expected to provide more information on the topic introduced in the first sentence. Authors have to be aware of this when they do write the documents, and also that long texts are rarely read by players.

If no answer matches the input with a sufficient level of precision, the system outputs an author-defined expression and the CM may intervene for managing the situation.

In order to capture relations among words, the concept of "concept" is implemented, as a set of words that may be used interchangeably in that question. Our implementation exploits the WordNet ontology and allows the author to define ad-hoc word sets as well for representing concepts. For instance, the concept vegetable could involve the terms: vegetable, carrots, artichokes, cabbages, etc. Exploitation of a whole ontology, as in the case with WordNet, is important, in particular to allow the use of hyperonyms in questions. For instance, "The car is red" may correspond to "What color is the car?".

Pronouns inside the player's question are tackled statistically, exploiting the concept of context. The context is a set of words. It is initially represented by few nouns representing the location where the player's avatar meets the NPC. This allows, for instance, to resolve questions like "What is this?", when the NPC is in front of a monument. The context nouns (e.g., "monument") are added to the player's direct input, so that they can be processed by the FSTP. The context evolves as the conversation goes on. In particular, it is enriched with the nouns said by the NPC. At each NPC's answer, the weight of the previous context is strongly lowered and the new nouns are added, with an exponential weighting function, that gives more value to the nouns in the end of the sentence. Similarly, previous nouns in the player's question are considered as well with a higher exponential weight (e.g., "I like that palace! Who has built it?"). This example also shows that if the player inserts two or more sentences, only the last one recognized as a question is used as input in the FSTP. If the CA is not able to recognize the question (i.e., all the candidate answers score below the recognition threshold), the FSTP analysis is done again, with an extended set of words, including also the previous sentence.

Another aspect is related to possessive adjectives/pronouns that are inverted in the original question (i.e., you becomes me, your my, and vice-versa) in order to allow the matching among questions and answers.

3.3 Additional CA behavior

We have also added two features in order to improve effectiveness of the dialogue in a game context, and they are both related to the provision of more in-depth information.

The first one is a simple implementation and consist in dividing a CA's answer into sequential pieces. At the end of each piece, the CA asks the player if he wants to continue listening on the topic or prefers making another question. We believe that such a subdivision of the provision of information is important in order to meet the player's expectations. Of course, this text (answer)-authoring approach enables the fact that the level of depth of information provision may also be adaptively defined through user profiling.

The second feature consists in creating a hierarchy of questions. In particular, the author can define a tree of questions, so that one question may enable the player to ask further questions, typically related to enabling one. This again allows the player to get more in-depth information, but also introducing a compelling game mechanics - the need to identify "opening" questions -, that represents a very simple model of the status of the conversation.

3.4 Authoring

A major driver of the proposed system is given by the need to allow authors to write dialogue text in a simple, efficient and effective way, and that the resulting documents are easy to maintain. In this subsection we summarize the characteristic of the authoring system, that is being implemented as a user-friendly visual tool for authoring.

In particular, the author has to prepare:

- The set of questions the Welcome Manager is able to recognize and, for each one of them, the relevant answer.
- The help texts and the trigger conditions through which the Conversation Monitor proactively intervenes to support the player
- The question recognition threshold for the possible answers
- The texts of all the possible answers, that represent the specific knowledge of the CA. Every answer has to be tagged only by the corresponding IAP
- Ad-hoc word sets representing concepts.

The author has to be aware of the fact that in the matching search algorithm, the first sentence in the document (answer) is weighted more, as it is expected to directly correspond to the question.

The author can also set dependencies among questions, defining a tree, so that one question may enable the player to ask further questions.

Finally, we have implemented also the Parametric Question/Answers (PQAs) feature, through which the author can define a template for a set of homogeneous questions and answers. For instance, we can imagine that the CA represents an art expert that can be queried by the player about the place of birth of several different painters. In order to spare time, the author can simply specify a "Where" question, and one single answer where the name of the painter and the place of birth are parametric and depend on the name inserted by the player in the question. The integration of this approach with a database table is a subsequent straightforward step.

4. USE CASES

In this section we present some application examples of the developed algorithm. We consider two use cases: Travel in Europe, a serious game for promoting the cultural heritage, and SeaGame, a serious game to promote safe behavior in coastal areas.

4.1 Cultural heritage context

Our game aims at promoting the cultural heritage of European cities of art. The user explores faithfully reconstructed areas of such cities, while accomplishing missions in a sort of treasure hunts with artistic topics. During his exploration the player may encounter some Non Player Characters (NPCs) with which he can interact in order to get information about the visited areas/buildings.

In the first situation, the player has to visit Palazzo Bianco – an artistic Renaissance villa in the Genoa city center - and collect information about the building and its museum.

P: Is there a museum inside? NPC: Yes, there is a museum inside	(a)
P: At what time does it open? NPC: The museum opens at 3 pm.	(b)
P: Why is it named Palazzo Rosso? NPC: It is named Palazzo Rosso because of its red color	(c)
P: When was it built? NPC: Palazzo Rosso was built in the 17th Century.	(d)
P: Who designed Palazzo Rosso? NPC: Architect Pietro Antonio Corradi designed Palazzo Rosso	(e)
P: What can I see inside? NPC: You can see a lot of pictures inside	(f)
Enabled additional questions	
P: Who are the painters of the pictures exhibited? NPC: The painters are Rembrandt and Van Dick.	(g)
P: Which are the most important pictures of Rembrandt? NPC: The pictures by Rembrandt are the following…	(h)
Level 1	
P: How much does the ticket cost? NPC: The ticket costs 10 €	(i)

Table 1 Example from cultural heritage context

Question (a) is a IAPF question. (b) is detected as a time ("when") question. Pronoun disambiguation is obtained by processing also context information (the museum term comes from the dialogue history, as it has been used in the previous interaction). The same applies to (d) for the term "Palazzo Rosso". In (d), the term build is part of a concept including also the term construct. Similarly, in (e), the term design is in a concept including also construct and build. Question (f) enables other possible questions about the paintings inside. (g) and (h) may be below the recognition threshold, in which case the CM may intervene. In (e) the "how much" tag is recognized. In Example 2, the player has entered Palazzo Rosso and is visiting the Museum and asking for information about the pictures on exhibit.

P: Hi! Can I play volleyball ball here? NPC: No you cannot play volleyball here.	(a)
P: Where can I play volleyball? NPC: You can play volleyball just outside the beach.	(b)
P: Can I use it (a ball that the player is holding)? NPC: Yes, you can use the ball.	(c)
P: Can I listen to the music? NPC: You can listen to the music, but not aloud.	(d)
P: Can I take my dog with me? NPC: You cannot come with the dog.	(e)

Table 2 Example from SeaGame context

Question (a) shows the use of the tag "how many", while the NPC answer in (b) may be tagged both as a "which" and "where" question (Where can I see Van Dick pictures?). In (b) there is a word concept that includes the verbs be, see and watch. Pronoun disambiguation (through history context) is applied in (c). In (d), there is a recognition failure, because the player sentence refers to a previous question of his. The CM intervenes by asking more detail (since there are too few words). The word concept in questions (e, f) includes the terms "there" and "on exhibit". The answer in (g) is not perfectly appropriate. This is an example showing that consequential interactions are problematic ton address.

4.2 Seagame context

The second context concerns Seagame [17]. The player arrives to the beach and asks information about the behavior rules at the seaside.

P: How many rooms are there in the museum? NPC: There are 3 rooms in the museum	(a)
P: In which room can I watch pictures by Van Dick? NPC: You can see Van Dick pictures in room two.	(b)
P: How many rooms are dedicated to him? NPC: Two rooms are dedicated to Van Dick.	(c)
P: And to Rembrandt? NPC: Not understand – The CM intervenes	(d)
P: How many paintings are on exhibit? NPC: There are over forty paintings.	(e)
P: Which kind of paintings are on exhibit? NPC: There are two kinds of paintings: portraits and landscapes.	(f)
P: Are they painted by Rembrandt? NPC: Rembrant has painted a number of interesting portraits	(g)

Table 3 Example after entering in a museum

Question (c) resorts to context information for the pronoun resolution. In (e) the word concept includes take, come with and bring.

In a second situation, the player meets a sailorman and asks him information about the history of the site.

P: Hi! I have just arrived here! NPC: Welcome! My name is Giobatta, I am a sailorman…	(a)
P: What is the name of this bay? NPC: The name of this bay is San Fruttuoso.	(b)
P: What can I see here? NPC: You can see the abbey with Doria's graves, here.	(c)
Level 2	
P: Who are the Doria? NPC: Doria are a famous doge family	(d)
P: What is a doge? NPC: A doge was the commander of the city of Genoa in the middle ages.	(e)
Level 1	
P: How can I go to the abbey? NPC: You can go the abbey by following the track just below the mountain.	(f)
P: When was it built? NPC: The abbey dates from the 10th Century.	(g)

Table 4 The player meets a sailorman

Answer (e) is tagged both as a "what" and a "who". In (h), the word concept of "dates from" includes also "build in" and "constructed in".

5. CONCLUSIONS AND FUTURE WORKS

Serious Games represent a promising tool for improving instruction, given their high realism and ability to attract a wide basis of users thanks to the appeal of technology and entertainment, and capture the player attention for long time spans.

Natural language interaction techniques can provide an important added value in this context, increasing realism and ease and pleasantness of user interaction with information.

We are implementing a natural language interaction system with a clear focus on instructional dialogues aimed at favoring a player acquisition of knowledge on specific topics through interaction with specialized NPC. Another key requirement for the system is easy and efficient writing and maintenance of the texts, also by author with no specific expertise on computational language technology.

The system we propose includes a strategy and a tactical level. The former is responsible for managing the high-level aspects of the conversation, while the latter responds to the player's queries by relying on an original combination of an essential syntactical analysis and an ITF-IDF-based procedure.

The technique specifically targets a (educational) gaming application domain – for instance, we could consider that the difficulty of getting information from an NPC may be seen as a challenge of the game. However, we believe that the technique could be tried and extended to meet the requirements of other application domains as well.

We are still finalizing the implementation of the system. So, the next steps will concern an extensive testing phase, that will help to assess the system in authentic contexts of use and give indications on enhancements.

We believe that the system performance could be improved in particular in the pronoun resolution module and in better specifying the notion of status of the conversation. Furthermore, we expect to get precious information from the tests on how to enhance the syntactical and statistical sentence analysis in order to better deal with less regular and common text entry patterns by the players.

6. REFERENCES

[1] Bellotti F., Berta R., De Gloria A., Panizza G., Pellegrino M., Primavera L., Designing Serious Games for Cultural Heritage Purposes, in S. de Freitas, P. Maharg (Ed.), Digital Games and Learning, Continuum, 2011

[2] Zielke M. A., Evans M. J., Dufour F., Christopher T. V., Donahue J. K., Johnson P., Jennings E. B., Friedman B. S., Ounekeo P. L., Flores R., "Serious Games for Immersive Cultural Training: Creating a Living World," IEEE Computer Graphics and Applications, vol. 29, no. 2, pp. 49-60, Mar./Apr. 2009.

[3] Losh, E. The palace of memory: virtual tourism and tours of duty in Tactical Iraqi and Virtual Iraq. In Proceedings of the 2006 international Conference on Game Research and Development (Perth, Australia, December 04 - 06, 2006.

[4] Johnson W. L., Developing intercultural competence through videogames, International Workshop on Intercultural Collaboration, Palo Alto, Ca, 2009.

[5] Mateas M. and Stern A., *Natural Language Understanding in Façade: Surface-text Processing* Best Paper Award, Technologies for Interactive Digital Storytelling and Entertainment (TIDSE), Darmstadt, Germany, June; 2004.

[6] Bellotti F., Berta R., Gloria A. and Primavera L.. "Enhancing the Educational Value of Video Games", ACM Computers in Entertainment (CIE), Vol. 7, June 2009, 2009.

[7] Wilcox, 2010 – Wilcox B., Beyond Façade: Pattern Matching for Natural Language Applications, available online at: http://www.gamasutra.com/view/authors/928/Bruce_Wilcox.php.

[8] Wallace R, http://www.alicebot.org /TR/2005 /WD-aiml/, 2005.

[9] Flycht-Eriksson A. & Jönsson A., Dialogue and Domain Knowledge Management in Dialogue Systems Proceedings of the *1st SIGdial Workshop on Discourse and Dialogue*, Hong Kong; 2000.

[10] Lease M., Natural language processing for information retrieval: the time is ripe (again). In Proceedings of the ACM first Ph.D. workshop in CIKM (PIKM '07). ACM, New York, NY, USA; 2007.

[11] Sparck Jones K. et Al., A probabilistic model of information retrieval: development and comparative experiments (parts I and ii). Information Processing and Management, 36:779–840, 2000.

[12] Ponte J. M. and Croft W. B., A language modeling approach to information retrieval. In Proceedings of the 21st annual international ACM SIGIR conference, pages 275–281, 1998.

[13] Lease M. et Al., A look at parsing and its applications. In Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06), 16–20 July 2006.

[14] Marcus M. et Al., Building a large annotated corpus of English: The Penn Treebank. Comp. Linguistics, 19(2):313–330, 1993.

[15] Etzioni O et Al., Machine reading. In Proceedings of the 21th National Conference on Artificial Intelligence (AAAI), 2006.

[16] Chiari I., Introduzione alla linguistica computazionale. Ed. Laterza, Bari, 2007.

[17] Bellotti F., Berta R., De Gloria A., Primavera L., "Adaptive Experience Engine for Serious Games", IEEE Transactions on Computational Intelligence and AI in Games, Vol 1, No. 4, 2009, pag. 264-280