

Dialogue Generation for Digital Interactive Storytelling (Dis)

Amal S. Fadak¹, Mohamed O. Khozium²

¹ Faculty of Computing and Information Technology, King Abdulaziz University,
Jeddah, Saudi Arabia.
Amal_Fadaak@hotmail.com

² Prof., Department of HITM, Faculty of Public Health and Health Informatics, ICRS
Consultant, Umm Al-Qura University, Makkah, Saudi Arabia.
mokhozium@uqu.edu.sa

Abstract

Nowadays, interactive storytelling systems have received significant attention from the research community exploring interactivity in terms of promoting engagement, enjoyment, improving storytelling content and diversity. These studies believe that IS would be used generally in different areas. The new technique uses and generates dialogue, giving the capability to produce new element of stories and enabling discourse between characters and player within the story world. In our project we endeavor to design a framework for dialogue generation within digital interactive storytelling system, taking into account novel approaches involving planning algorithms, ontology-based generative grammar and dynamic story context.

Keywords- Dynamic story context; Grammar story context; aka Friends System; Façade System; Scenejo System; Tale-spin system

I. INTRODUCTION

Entertainment is a broad field that recently became an important area of research. Researchers have carried out academic studies that focus on games and their storylines, where the stories in games play a part in raising the immersion level, build tension, and add to the interest of the player. However, one main difference between games and other media is the interactive nature of games; as they draw a player to engage with the story of the game. There are several ways for presenting a game's storyline, these include linear, branching, parallel, and threaded. Most games follow the linear approach where the events are pre-defined in a sequence to be displayed to the players. Thus, the player does not have the role of developing the flow of the plot thereby diminishing the interactivity level of a game.

In order to benefit from the story and allow the player to feel a high degree of interactivity, the concept of interactive or non-linear storytelling has to be introduced. Simply put, interactive storytelling presents the opportunity for players to create their story, to have an input in what is happening in the game world they are placed in. They are responsible for dictating how certain events may come to pass. This in turn helps players to be more engrossed in the game and enjoy them in a greater degree, since as stated by Laurel in [1], "story making is a pleasurable activity", as seen by recent studies [1,2,3,4,5,6,7].

The question arises in what ways an interactive experience can be provided to the player. The most common techniques used, and as noted [3], are branching storylines, which enable the player to select the order of the missions available. However, to achieve this goal, we can employ other techniques, such as parallel paths, threaded stories, or dynamic object oriented narrative [8]. Nevertheless, adopting dialogue generation is a new technique and offers the interaction between character-character and character-player. It leads the character to act intelligently, by interpreting the state of

their environment and applying proper behaviors or responses accordingly [9].

Dialogue generation is the ability of producing an appropriate conversation between characters in a virtual world, as in chatterbot. In case of interactive storytelling systems, dialogue generation would change the form of stories from linear to “non-linear branching dialogue-based” storytelling. This creates different storylines based on the interaction and discussion between the story characters and between the characters and the player. Developing dialogue generation system is a crucial matter to focus on, where characters can decide what to say and when to say it in order to achieve their goals.

In this paper, we introduce designs related to DIS, dialogue generation. The paper structure is organized as follows: in section 2, we briefly discuss DIS systems, the state-of-art in dialogue generation. In section 3, we present some of existing dialogue generation systems. In section 4, we discuss the negative and positive aspects of these systems. Finally, we sum-up the contribution of this paper

II. BACKGROUND

A. *Digital Interactive storytelling (DIS)*

What really make games interactive? And, what is the best way for that? These questions are some questions researchers have been trying to focus on. Digital games are way of entertainment where gaming companies have been trying to embellish games with stories. To heighten the entertainment value, the game allows a player to interact directly with the story world, which yields a story. This interaction becomes indispensable and considered as one of the main advantages of the games technology. This is known as Interactive Storytelling (IS), a new term coined by Chris Crawford and according to his definition, IS is when a story engine is capable to generate stories depending only on user's response to the story events [10]. Several different IS systems have been built for various purposes although they all contain two important elements: drama manager (DM) and characters. Drama manager is the part that coordinates the agents' behaviour and actions in the virtual world and handles the story's mechanics.

In terms of characters, the virtual agents who inhabit the virtual environment have a plan to follow and goals to accomplish, and they communicate to do so. As result of communication, the sequence of actions will be generated in such a way that they are organized and coordinated in order to constitute a story. A planner technique is applied as it is important to derive the sequence of actions forming the story and controlling the unfold story. Regarding the dialogue generation, Cavazza & Charles in [11] remark how this occurs through a number of steps. First, it is determined how the utterance would be expressed. Second, the semantic structure will be formed based on the dialogue. Finally, lexicalisation, that is, the generation of lexical form from the semantic content, will occur.

Story production in games is a significant feature for ensuring players feel immersed and involved. It consists of events that change according to character's communication, representing them in different ways, namely, linear, branching, parallel, and threaded. Linear is when the sequence of events is fixed and organized in a logical and rational sequence. Branching is where players are able to choose the order of events from several available options. Parallel storylines have multiple protagonists, each of whom has their own stories, which intersect at certain points. Threaded storylines occur when the story is divided into small parts. The player can choose different paths, which will lead to different storylines in terms of beginning, middle, and ending [12].

Some games have attempted at creating interactive stories such as Heavy Rain, Mass Effect 3. In the first, the story is about four playable characters involved with the mystery of a serial killer. The player has to perform some actions by pressing keyboard buttons or using motions of the joystick controller. The player's actions and decisions might lead to certain actions that may lead to different endings. The second, is 3rd person action Role-Player Game (RPG), the player takes a role as Commander

Shepard whose work is saving the earth from alien invasion. The game takes into consideration all decisions made in the previous two parts of Mass Effect 1 & 2 then feeds them into the narrative system. These decisions along with any upcoming decisions are made through the game while trying to identify what the ending could be, or whether or not the alien invasion could be stopped.

B. Dialogue Generation

Dialogue generation is a producer of discourse acts involving characters within a context generated by the storytelling system. This type of system allows players to direct the story according to their preferences. Being such a system, it should be understood and generated. First, understanding what is being told by the user and identifying the concept of the utterances. Second, merges phrases for generating the appropriate responses according to the current discourse context.

As regards the dialogue systems functionalities, they are many at Smith states [13]:

1. It has to accomplish the user's goal, where any requested information by the user must be provided,
2. It has to manage and deal with any task that comes up during the conversation.
3. The generated texts have to be well-structured and that depends mainly on the representation of the conversation data.
4. It supports different degrees of initiative for the players. If the dialogue guided by the system then it is system-initiative while if the user have the initiative and guide the dialogue to the goal then it is user-initiative.
5. Predicting the player's answers and responses.

As mentioned earlier, understanding and generation should be met by the systems and these two features are Natural Language Processing (NLP) types: Natural Language Understanding (NLU) and Natural Language Generation (NLG) and that relevant to dialogue generation. As Reiter states, the generation process divided into three phases:

1. *Text planning*: deciding the utterances' goals are to achieve, what information should be included in the text and what order they take within sentence structure.
2. *Sentence planning*: how the goals might achieve including grouping the information into sentences and deciding which words and phrases should be selected to provide readable texts.
3. *Realization*: applying the grammar rules to produce texts syntactically, morphologically and orthographically correct.

C. Artificial Intelligent Planning

Planning techniques have been applied successfully in developing interactive storytelling. As long as characters have sequence of actions to perform and make up a story, they need to be controlled and organized to get the goals accomplished. In regards to this, planning technique, a type of AI, has been

integrated for unfolding individual characters' roles. Using planning techniques within interactive storytelling would increase diversity and cohesion of the story and thus increase the user's experience [14], [15]. Ghallab et al., in [16] defined the planning as "subarea of AI that studies the abstract and explicit deliberation process of choosing and organizing actions, by anticipating their expected outcomes, in order to achieve objectives."

There are many planning techniques involving Hierarchical Task Network planning (HTN) and Heuristic Search Planning (HSP), Graphplan, Fast Forward (FF), Partial Order Planning (POP) [17], [18] The importance of adopting planning approach is in its ability to generate narrative events and dramatic interest based on its own procedure. Within HTN, each task is decomposed into sub-tasks until they are described in terms of primitive actions, the actions that the story consists of. The top

node/task is the main aims to accomplish and it is divided into sub-nodes, the first layer nodes describe the scenes and in turn each of such nodes divided into sub-nodes as well (second layer nodes) and describe the steps required for characters to achieve their objective [19].

Whereas HSP extracted heuristic from STRIPS formulation (STanford Research Institute Problem Solver) and estimating the goal distance from the initial state by using both a heuristic search algorithm [20]. It works by measuring the path from any generated state to the goal state [17].

Yet Graphplan is a graph-based plan, and it consists of several levels of both literals and actions. Many steps are created to get a graph. Level (0) represents only the true-value literals. Then, the following level represents actions for which the preconditions hold in the first level, (0), are added. Repeating the process until getting the graph, in case all literals have no mutex relation between them or until a new action level is added [21].

FF, which is short for Fast-Forward, is one of the planning techniques which can be used for the extraction stage in different approaches e.g. HSP, Graphplan. It is a forward state space search technique, which is characterized by using heuristic functions in order to estimate the distance to the goal and avoid mutually exclusive actions. Furthermore, the heuristic derived utilizing a relaxation process known as an enforced form of hill-climbing, which combines both local and systematic search [22].

III. DIALOGUE SYSTEMS EXAMPLS

Different areas have used dialogue generation systems such as education, training, and medical areas, the following will introduce some of them.

A. *I-Storytelling (aka Friends)*

The systems using agents having goals to follow based on and HTN to generate plan for achieving these goals [14]. I- storytelling system is character-based where the story is described through the individual roles of each actor [23]. The Friends-based system has character Ross wants to invite character Rachel on a date, for that HTN plan is used. It breaks Ross's goal into different sub-goals and each sub- division corresponds to actions and has a specific narrative significance [3]. For example, Ross's goal is to attract Rachel, one of the sub-goals could be finding information about her, or talking to her in some way, inviting her out, offering her gifts. Each of these sub-tasks can be considered a goal to achieve as well in its first level until reaching the terminal nodes, which are primitive actions that represent the actual action for the character.

Cavazza and Charles state in [2] that the sentences structure is affinity-based generated then analyzed semantically using Tree adjoining Grammar (TAG). According to the information being used, the semantic structure is identified as: informative, interrogative, questions, or requests. Then, the surfaces of the texts are created by checking the correctness and selecting the suitable elements that express the character's goals.

B. *Façade*

First interactive system been developed and offers a high level of interactivity. The story Mateas and Stern [24] created is about a married couple (Grace and Trip) and a player who walks into their apartment and later on, a tense position arises when they start arguing more and more in front of the player who finds himself entangled in it. The player acts and the characters decide the reactions accordingly. The story ending is based on the player's reaction and there a drama manager who monitors the sequence of the events. The manager acts as a planner for selecting, ordering and executing beats which are responsible for unfolding the story's events: guides characters, describes their actions.

The characters-player interaction is text-based; each character has a set of behaviours enabling them to perform several intelligent activities, which are written with A Behaviour Language (ABL). The activity to be considered as a goal and to be accomplished and should be supported with one or more behaviours, where these behaviours consist of sequential or parallel steps. If all the steps have been completed then the behaviour is achieved, it is indicated as a failure and eventually the goal (activity) would start looking for another behaviour instead for satisfying the goal.

Furthermore, NLP system is involved and it is responsible for beats selection based on the player interaction. Through Natural language processing the beats translated into “discourse act”, determines what action’s type is: a compliment, a criticism, or even flirtation in a manner appropriate to the current total context. This comes later after a potential reaction have chosen based on the beats, “what effect” and what reactions have been adapted for being compatible with the current situation [11,25].

C. *Virtual Storyteller*

It is multi-agent framework its architecture consists of three agents: director, character, and plot. The reason behind building such a system is creating short stories, fibula, in fairytale domain [26]. The generated stories is character-based interaction on the other hand is supervised by the plot agent, therefore the characters known semi-autonomous.

In the first phase for creating the story, a document plan is produced from the fabula, which means a formal story representation, takes the form of a binary tree. Its node is the story elements and edge is relations between them and annotated with some terms for more information. As long as there are elements that are not used are eliminated. Next, a syntax tree is replaced with the document plan whose nodes and edges indicate lexical units and relationships between them, respectively [27]. More sentences produce more trees, which are combined and the way they combine is determined by the edges. Finally, the natural language is generated coherently making sure the words in trees are in correct order.

D. *Tale-spin*

It is a system able to generate simple stories through interaction of simulation characters and the user allows initializing both the story and the goals. The system is plan- based algorithm, the story generated according to the plan and afterwards it represents as natural language [on craft of interactive storytelling]. Although the system lets the user influence the story, nevertheless, it has no prior knowledge about the story in addition it lacks a good structure and poor linguistics and thus creating difficulties in reading stories.

Tale-spin operates in bottom up approach, a world is constructed and characters are given goals to pursue. As a start, the story describes where the characters live while the knowledge each agent should have is set up by the user. Finally, it comes to the user responsibility to decide what goal should attain to generate the story according to.

Tale-spin uses a Conceptual Decency model (CD) to represent the knowledge acquired, including the following representation tokens:

- Real world objects
- Real world actions
- Times
- Locations

Thus, the generation process is directly from the CD to English texts, match every possible token with its correct position in the sentence. Then, from the conceptual dependency a semantic representation is built up e.g. a lexical entry is given for each verb includes the CD and verb tense.

Although the system is limited to generate subject-verb pair, four parameters have been used to identify the structure of the sentence in terms of the needed subject, the tense, sentence's form type whether negative or interrogative, and meaning of actions. Once they are created, the story represented in its final results as a natural language [28].

E. Scenejo

This is a conversational system which enables different A.L.I.C.E agents and users to have a multi-agent text conversation with each other [Learning with interactive stories]. The created dialogue's contents are influenced by user's inputs. To achieve the result, a database attempts to match these inputs rendering them into talks embraced by the animated characters however the central part of the system's design is an authoring tool whose role is controlling and managing the expected outputs.

Regarding the generation principle, the system has two ways to fulfill either to re-arrange the story into several pre-defined plot lines or the story derived according to stated changes. Many chatbots have been used in a turn-taking loop and each has its own AIML knowledge base. Each AMIL includes a "pattern-template". Scenejo adapts "question-answers" principle for creating dialogues. Utterances are conveyed into intermediate structure called abstract acts, which they can be described in multiple discourses. Once this done, it is the time to get a response depending on a "question". Therefore, Scenejo allows the definition of short dialogue sequences, narrowing down the context for the pattern recognition of possible following acts. Changing states play the main role with regards to the way the system responds. The generated result is affected according to pre-conditions of the agents so, a predictive value should be examined first for being capable of generating a proper answers.

I. COMPARISION

Since Tale-spin system was one of the initial systems that addressed generation idea according to the applied planning approach. Thus, it fulfilled the organizing and controlling of the system over the generating texts. On the other hand, most of the stories did not do so well and were either too short or uninteresting despite the characters performed coherently.

The nature of interactivity of Tale-spin is to write standard linear stories, audience interaction is absent where the stories rely on pre-written storylines.

Another important characteristic, the range of sentence generated is limited because the system has an inflexible- structure adopted, subject-verb pair. Sentences are built up straight from the CD's semantic representation due to the absences of syntactic processing. Therefore, generated structure may be inadequate or fails to express the variety of discourse phenomena found in natural storyline.

While I-storytelling system is similar to Tale-spin in using planning paradigm that plans and controls over agents oriented goals and produces a high quality narratives. The importance of using it lies in the generating of stories and how they emerge simply from agents actions. Friends system characterized with 3D environment based on *Friends*-sitcom, generating several storylines based on both agents' interaction and user intervention by moving objects or speech recognition. The dialogue is shaped based on characters' relations and affinities. Because of different analysis levels for characters' interaction, it identifies and displayed correctly as well as triggered the right reactions and responses. Exploitation of grammar rules within the I-storytelling system is a beneficial matter, that way the user interaction is more efficient and becomes a spectator instead of a director.

In Façade, the communication is text-based, by presenting a dramatic situation to the player whose role is dealing with the animated characters according to their reactions. Regarding the characters, ABL is a language which allows behaviors so the character becomes more interesting. An important feature is beats concept, which is responsible for changing the story state, managing goals situation

whether successful or not. Each of these beats has pre-conditions that modify the state of the story.

On the other side, Façade conveys the sentences into discourse acts form: agree, disagree, criticize, anger ignoring the fully- understanding for the entered text but in such a situation the system could not recognize the meaning of such a word, the system exploits OR operator to find other equivalent set of words to use instead and NOT to eliminate the presence of words.

The concept behind Scenejo's game is for educational purposes, how the dialogue affected the inner states of the bots and their reactions. Using A.L.I.C.E chatterbots as talking heads limits the interaction to verbal interaction. The conversation is based on AIML pattern matching that allows for representing the adjacency pairs. Although it fails to respond correctly if a user makes some spelling mistakes or a different syntax than what is expected. In the context of the utterances, being expressed as dialogue acts plays a vital role in identification of the discourse structure.

Yet in Virtual Storyteller, the story generation is emergent- narrative based, which allows a flexible structure of the plot depending on characters-player interaction. Consequently, increases the authorship and diversity of user interaction,

while their characters are capable of justifying their goals and build the plans. That occurs by choosing the suitable events for the story to pursue these goals. On one hand, the system permits the character to experience his own story which allows for more than one story to be created but on the other hand, it still capable of generating very simple story texts. Furthermore, the director agent has limitation in terms of story rules, guiding the characters, and disallowing character's intended action. These characters require more characteristics and emotions and as a result, the story becomes boring and tedious.

II. CONTRIBUTION

An important objective of IS systems is to be compatible for natural modality, natural language based interaction is a significant endeavor within AI area. We address creating a framework for dialogue generation that ultimately works within the interactive systems and allow agents to obtain knowledge from the surrounding environment through dialogue and acting according to that knowledge. The research has investigated this technique and evaluated its suitability within interactive systems. The first contribution is producing utterances based on autonomous agents' interaction and the second is assuring these dialogues perceived by the user as natural, relevant, consistent, and coherent within the context of the dialogue. Of course, it is important to consider grammatical side of dialogue which offers theoretically attractive results. Clearly, that approach might be sufficient for wide-coverage in verbal interaction rather than being fixed and restricted for specific domains.

III. CONCLUSION

This paper discussed several existing areas that are considered important and widely used within research community. Idea is presented to combine dialogue generation and planning techniques within interactive storytelling area. That way, it would lead us to explore the development of a new framework for multi-agent system which is capable of using dialogue generation for DIS. By incorporating natural language techniques, it contributes to modifying and changing human-computer interaction (HCI) from ad-hoc to being generated on the fly by planner.

IV. REFERENCE

- [1] Laurel, B. “Narrative construction as play”. *Interactions* 11(5), pp.73– 74,2004
- [2] Rollings, A., and Adams E. . Andrew Rollings and Ernest Adams on Game Design. USA : New Riders Publishing, 2003, pp. 446
- [3] Rouse III, R. *Game design, theory and practice*, 2nd ed., Plano, Texas: Wordware Publishing, USA, 2005, pp. 76-77.
- [4] Rollings, A.; Morris, D. *Game architecture and Design*. The Coriolis Group, 2000
- [5] Salovaara, A., Johnson, M., Toiskallio, K., Tiitta, S., Turpeinen, M. “Playmakers in multiplayer game communities: their importance and motivations for participation”. *Proceedings of the 2005 ACM SIGCHI international Conference on Advances in Computer Entertainment Technology*, 2005, pp. 26.
- [6] Champagnat, R.; Estrailier, P; Prigent, A. “Adaptative execution of game: unfolding a correct story”. *Proceedings of 2006 ACM SIGCHI international conference on Advances in Computer Entertainment Technology* 2006, pp. 63.
- [7] Popova, Y and Cuffari, E. (2018), “Teporality of sense-making in narrative interactions.”. *Cognitive Semiotics*. 5-2018
- [8] “Foundations of Interactive Storytelling”. [online] Available at: <http://igba.org/> [Accessed: 2 December 2011]
- [9] Yu, M.-H.; Li, J.; Liu, D.; Zhao, D.; Yan, R.; Tang, B.; and Zhang, H. 2020. Draft and edit: Automatic storytelling through multi-pass hierarchical conditional variational autoencoder. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 1741–1748
- [10] Crawford, C. *Chris Crawford on Interactive Storytelling*. New Riders Games. 2004, pp. 223
- [11] Cavazza, M. and Charles, F. “Dialogue Generation in Character Based Interactive Storytelling”. *Proceeding of AAAI First Annual Artificial Intelligence and Interactive Digital Entertainment Conference*. Marina del Rey, California, USA, 2005, pp. 21-26.
- [12] Merabti, M., El Rhalibi, A., Shen, Y., Daniel, J., Melendez, A. and Price, M. “Interactive Storytelling: Approaches and Techniques to Achieve Dynamic Stories”. *International Journal of Transactions on Edutainment I*, Vol. 5080, pp. 118-134, 2008
- [13] Smith, R.W., Hipp, D.R., and Biermann, A.W. “An architecture for voice ds based on prolog-style theorem proving.” *Computational Linguistics*, Vol. 21 No. 3, pp.281-320, 1995.
- [14] Guilherme da Silva, Fabio, Angelo Ciarlini, and Sean Siqueira. "Nondeterministic planning for generating interactive plots." *Advances in Artificial Intelligence–IBERAMIA*, 2010, pp.133-143.
- [15] P. Ammanabrolu, W. Cheung, D. Tu, W. Broniec, and M. O. Riedl. Bringing stories alive: Generating interactive fiction worlds. In *1st Joint Workshop on Narrative Understanding, Storylines, and Events (NUSE) at ACL*, 2020. URL <https://arxiv.org/abs/2001.10161>
- [16] Ghallab, M., Dana, N., and Paolo, T. *Automated Planning: Theory & Practice*. Morgan Kaufmann, 2004, pp. 98.
- [17] Sali, S. and Mateas, M. “Using Information Visualization to Understand Interactive Narrative : A Case Study on Façade”. *Interactive Storytelling 4th International Conference on International Digital Storytelling*, 2011, pp. 284-289.
- [18] Mateas, M. and Stern, A. (). “Natural Language Understanding in Façade: Surface Text Processing”. *Proceedings of Technologies for Interactive Digital Storytelling and Entertainment*, Darmstadt, Germany, 2004, pp. 5-6.
- [19] Cavazza, Marc, Fred Charles, and Steven J. Mead. (). "Character-based interactive storytelling." *Intelligent Systems, IEEE* , Vol. 17 No. 4, , pp.17-24, 2002.
- [20] Porteous, J., Cavazza, M., and Charles, F. “Applying Planning to Interactive Storytelling: Narrative Control using State Constraints”. *ACM Transactions on Intelligent Systems and*

- Technology (ACM TIST)*, Vol. 1 No. 2, 2010, pp. 1–21.
- [21] Planning Graph as a (Dynamic) CSP: Exploiting EBL, DDB and other CSP Search Techniques in Graphplan. (n.d.)
- [22] Charles, F., Lozano, M., Mead, S. J., Fornes Bisquerra, A. and Cavazza, M. "Planning Formalisms and Authoring in Interactive Storytelling". *Proceedings of 1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment. Darmstadt, Germany, March 2003*, pp. 24–26,.
- [23] Charles, F., and Marc, C. "Exploring the scalability of character-based storytelling." *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2*. IEEE Computer Society, 2004, pp. 9-11.
- [24] Cavazza, M., Charles, F. and Mead, S. J. "Under The Influence: Using Natural Language in Interactive Storytelling". *1st International Workshop on Entertainment Computing: IFIP Conference Proceedings*, 240. Netherlands: Kluwer, 2002, pp. 3-11.
- [25] Attia, S., Lioure, R. and Declaude, Q. (2020), "Future trends and main concepts of adaptive facade systems", in *Energy Science and Engineering*, vol. 08, issue 9, pp. 3255-3272. [Online] Available at: doi.org/10.1002/ese3.725.
- [26] Wijnia, S. "Towards an interactive Virtual Storyteller." EEMCS, University of Twente, 2005.
- [27] Langohr, L. "Virtual Story Generation: from TALE-SPIN to the Virtual Storyteller." *Seminar on Computational Creativity, Dept. of Computer Science*, University of Helsinki, September 22, 2011.
- [28] Lee, M. "A model of story generation." *University of Manchester, Manchester*, 1994.