Supporting Social Skills Rehabilitation with Virtual Storytelling

Cindy Even  
Lab-STICC UMR 7528  
Virtualys

Anne-Gwenn Bosser  
ENIB  
Lab-STICC UMR 7528

João F. Ferreira  
School of Computing  
Teesside University

Cédric Buche  
ENIB  
Lab-STICC UMR 7528

Florian Stéphan  
Laboratoire des Neurosciences de Brest  
CHU Brest

Marc Cavazza  
School of Computing  
Teesside University

Christine Lisetti  
Affective Social Computing Lab  
Florida International University

Abstract

Social skills training (SST) has recently emerged as a typical application for emotional conversational agents (ECA). While a number of prototypes have targeted the general population, fewer have been used for psychiatric patients despite the widely recognised potential of ECAs technologies in the field of mental health. Social cognition impairment is however a widely shared symptom in psychiatric patients suffering from pathologies such as schizophrenia. Going further than SST, rehabilitation programmes involving role-play, but also problem solving have been successfully used by clinicians, drastically improving the quality of life of patients suffering from such disabilities. One of the challenges of these programmes is to ensure that the patients will be able to adapt their behavior when the situation varies, rather than training them with the appropriate behavior for a set of specific situations. In this paper, we describe a novel approach for the development of a serious game supporting rehabilitation programmes for social skills, which will primarily target schizophrenia patients. We propose to use an ECA in combination with a narrative generation engine issued from interactive storytelling research to provide varied situations. This approach reflects the combination of both role-play and problem solving exercises on which remediation therapies rely, and has the potential to support patient’s progress and motivation through the rehabilitation programme.

Introduction

According to the World Health Organisation, schizophrenia affects about 29 million people worldwide. Social functioning is impacted for most patients and evidence suggests that social disability levels may predict patient outcome better than clinical symptoms.

We describe here an architecture and proof of concept application, for the construction of dialogue-based social skills exercises which may be used in the context of a social rehabilitation programme. Our proposal combines in the same application an emotional input dialogue with an Emotional Conversational Agent (ECA), driven by a generative engine to provide easily varied and coherent dialogues.

Copyright © 2016, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.


After motivating our project and reviewing related work, we will describe our proposed architecture as well as a prototype application illustrating the approach.

Social Skills and Schizophrenia

Emotion recognition and expression, memory, motivation or planification for instance are among commonly found cognitive deficits in schizophrenia patients (Green 1996). Such deficits have been identified as predictors of degraded social functioning (Prouteau et al. 2005).

While positive symptoms of schizophrenia such as hallucinatory experiences respond well to medication-based treatment, such protocols do not satisfingly address these deficits (Kucharska-Pietura and Mortimer 2013).

In order to improve the quality of life of schizophrenia patients, clinical approaches have been developed to address the social impact of this major disorder. Social skills training (Bellack 2004) attempts to address social cognition (defined by (Liberman et al. 1986) as composed of cognition, verbal and non verbal communication). Cognitive remediation techniques add a problem-solving dimension to the patient’s arsenal (Péneau and Franck 2015) and target each cognitive deficit individually. For instance, Gaïa (Gaudelus and Franck 2012) is a cognitive remediation programme targeting the ability to recognise facial information, for people with schizophrenia or related disorders.

However, evidence suggests that approaching these deficits with a combination of methods in an integrated rehabilitation programme may be more efficient (Kurtz et al. 2015; Almerie et al. 2015), and it has been proposed that it may be profitable to address social cognition as a whole in naturalistic settings: (Oker et al. 2015) argue for such an integrated approach to the evaluation and treatment of social deficits, describing the advantage of virtual characters in such an endeavour.

Related Works

Virtual Reality, Serious Games, and Mental Health

The focus of attention on an engaging distraction is thought to be a key factor promoting learning in games (Thompson et al. 2010); entertainment likely attracts and holds the player’s attention on the video game, thereby facilitating player’s exposure to behavior change procedures. As a matter of fact,
engaging a patient’s motivation is frequently necessary in health care, as patients can be required to undergo unpleasant or mundane procedures (Kato 2010). Sustained motivation is particularly difficult to achieve for schizophrenia patients, as a deficit in this respect is among commonly found negative symptoms. (Rus-Calafell, Gutiérrez-Maldonado, and Ribas-Sabaté 2014) and (Park et al. 2011) propose social skills training virtual reality applications for schizophrenia patients. The latter work reported that using such technologies in addition to regular treatment can be helpful in increasing and maintaining schizophrenia patients’ motivation. They conclude that these sustained motivation levels may on their own account for a part of the global efficiency of the solution.

Another asset for serious games is their ability to adapt to the player. For instance, (Fernández-Aranda et al. 2012) describe how technologies for emotion recognition can be used in a serious game for patients with impulse-related disorders. They use biosensors together with speech and face recognition systems to track the emotional state of the player during the game. The game then automatically responds by modifying the game play difficulty which helps the patients to learn self-control and emotional regulation strategies as well as relaxation skills. Virtual human interviewers have also been used in the past to detect behaviors correlated with depression, anxiety or post-traumatic stress disorder using multimodal perception systems (DeVault et al. 2014; Yu et al. 2013).

**ECAs and Social Skills Related Applications**

Emotional Conversational Agents (Pelachaud 2005) have become a tool of choice for applications related to coaching and training, through their support of affective dialogue and multimodal interfaces. For instance, the system HapFACS is used to control a virtual health coach: 3D virtual agents engage people in spoken dialogs focused on healthy lifestyles (Lisetti et al. 2013), such as on excessive alcohol consumption. The dialogue system processes the persons’ spoken answers and facial expressions to control the flow of the conversation based on effective motivational health care interventions.

Another recent example is the TARDIS project (Anderson et al. 2013), proposing a training software for job interviews, demonstrating the potential of ECAs for social skills related applications in the general population.

The potential offered by the safe environment that ECAs provide for social skills practice has been acknowledged in the area of mental health as well. As noted by (Kim and Kim 2011), virtual avatars have an advantage over a clinician or trainer, who may have limited acting or expressing capabilities, hindering role-play exercises efficiency: virtual characters can present emotional cues in a consistent and clinically-controlled manner. In (Tanaka et al. 2015), authors describe an ECA embedded in a virtual environment for social skills training, targeting patients suffering from Autistic Spectrum Disorders.

The potential of ECA technologies in remediation therapy for schizophrenia has been recently discussed in (Oker et al. 2015) with preliminary works regarding the use of an ECA with a scripted dialogue (modeled as a state machine). This is currently the approach closest to ours, and whilst the authors report mixed-results so far in terms of patient’s engagement, they argue that improving the technological settings with more immersive features would help in this respect.

**Interactive Storytelling for Engagement and Variability**

Interactive Storytelling techniques, and especially narrative generation techniques aim to engage the audience in varying and adaptive storylines. Such techniques could then provide a natural medium to better engage the patient in the dialogue with the system: when compared to a pre-scripted dialogue tree or state machine, generative techniques allow to describe a variation of stories scaling better with the complexity of maintaining coherent dialogue unfolding. They also aim to provide a greater variation through successive use of the same system.

Narrative Generation techniques are designed to create multiple and adaptive scenarios from the same initial narrative data. Mostly based on planning techniques (Young 1999; Porteous, Cavazza, and Charles 2010; Riedl and Young 2010), recent narrative generation systems can also rely on logic based approaches (Martens et al. 2013). This approach relies on the dynamic construction of causally linked storyline structures. The trace produced by their execution may then be subject to a post-processed causality analysis, favoring the user-reflection on her experience.

Narrative Generation and Conversational Agents research have already been interleaved in the past: for instance, the system described in (Cavazza and Charles 2005) deals with the generation of dialogue and character-based storytelling in a unified way whilst the Companions project (Cavazza et al. 2010) relies on a narrative theory of influence to generate persuasive dialogues. Previous work has also addressed the modeling of stories driven by the emotions of virtual characters (Pizzi January 2011) as well as user emotions (Cavazza et al. 2009).

**Proposal**

In this paper, we propose an architecture suited to the development of programs supporting social skills rehabilitation according to the following principles:

- Our system offers an emotionally-charged dialogue between the user and a virtual agent. The software produces a trace of the interaction, which may help in reflecting on the interaction between the ECA and the patient during debriefing sessions.

- Patients facial expressions will be used to influence the virtual agent’s response, in addition to dialogue utterances, providing the user with a naturalistic feedback on her effectiveness in communicating emotions.

- From one dialogue description written for the system by the designer, we potentially obtain varying dialogue unfoldings, which may prevent the patient from only learning a rehearsed behavior. This may also help sustain patient’s motivation.
We also describe a prototype developed in collaboration with a clinical care team as a proof of concept for a serious-game supporting social rehabilitation therapies for schizophrenia patients.

**System Overview**

![System Overview Diagram](image)

The system uses the components described on figure 1:
- **ECA**: HapFacs (Amini and Lisetti 2013) is an open-source framework providing virtual characters, animated in real-time, and providing verbal as well as non-verbal (emotional) communication;
- The TeLLer open-source suite provides a Linear Logic based narrative engine, allowing varying interactive scenarios, from a declarative description;
- Emotional Input: an off-the-shelf software from Intel® Perceptual Computing SDK\(^2\) provides emotion recognition from the patient’s facial features.

**ECA**

HapFACS allows to design a set of facial expressions on virtual characters, based on Emotional Facial Action Coding Systems (FACS) (Friesen and Ekman 1983). It uses the Haptek Player\(^3\) and characters created in the commercial software PeoplePutty\(^4\) from the Haptek company. The Haptek software is an avatar system widely used in research projects over the years (Smith et al. 2011). The HapFACS software provides an accessible programming interface to control facial expressions in real-time based, which has already proven useful in health coaching applications such as (Lisetti et al. 2013). In order to animate the ECA, we need to specify the sentence of the dialogue, with the emotion associated (anger, happiness...) and its intensity.

**Dialogue Generation**

We use the narrative generation engine provided with TeLLeR, an open-source\(^5\) suite of tools comprising an interactive linear logic proof explorer that is focused on story causality and the generation of variants. Designed to unfold causally linked sequences of narrative actions according to constraints and with an aim in mind, this engine can be used for manipulating speech acts as atoms of an unfolding conversation. Generation of consistent dialogue in Linear Logic has actually been proposed in the past (Dixon, Smaill, and Tsang 2009).

This approach is suitable for the verbal communication in our system: verbal communication from the user of the system will be given in the form of possible choices in a simple turn-by-turn mode of interaction, and the system will use such speech acts as the atomic building block in the verbal conversation as well.

Linear Logic is a logic of resources: when compared to classical logic, inference rules represent the consumption and production of formulas which may be used to represent complex states. This paradigm is naturally suited to succinctly specify dynamic systems which may traditionally be described by state transition systems such as petri nets for instance, and why it has also been proposed for modeling narrative applications (Bossier, Cavazza, and Champagnat 2010; Bossier et al.). Using this paradigm implies to treat uniformly elements of the domain modeled as resources, rules governing over their evolution. Resources present in the environment at any given time define the state of the narrative.

One particularity of generation engines based on linear logic is thus that they represent the narrative environment in terms of resources, which may be consumed or produced through narrative actions unfolding. In our setting, the most significant resource we manipulate are the emotions expressed by the user (recognised by the system) and by the ECA during the conversation. Speech acts are therefore modeled as actions transforming social emotions: for instance, insulting the ECA may result in transforming a neutral expression to an angry one.

At its core, TeLLeR is a forward chaining prover that implements a fixed-point strategy which offers interaction at certain choice points (e.g. when there are different possibilities for the proof to progress). Some of TeLLeR’s features are well-suited for our goals.

TeLLeR supports the introduction and removal of resources at any point in time which is how it can be interfaced easily with other components of the systems such as user input. This permits the perceived emotion of the patient to be updated regularly in the sub-system runtime environment, and to drive the ECA sub-system answer consequently.

TeLLeR allows us to change aspects of the interaction like the granularity of the forward chaining search; we use

---
\(^{3}\)http://www.haptek.com
\(^{4}\)http://www.haptek.com/products/peopleputty/
\(^{5}\)http://github.com/jff/TeLLeR
this feature to control how many causal links the ECA follows between each interaction point with other sub-systems. A finer granularity will also allow more interaction points where new resources and actions can be added: it provides a more flexible way to deal with interactions.

We implemented a new feature that allows the possibility to backtrack and try different actions. At the application level, the rationale is that patients can try new behaviors and see what would happen if they had chosen different actions. It also allows them to replay ECA’s answers that may not be understood the first time.

**Emotional Input**

Patients suffering from schizophrenia frequently have a difficulty expressing the appropriate emotion. In order to provide a system addressing social cognition as a whole, it has therefore seemed important to the clinical team to embed facial emotion recognition software as an interface to the system. In order to respect the patient’s medical confidentiality we chose an off-line system.

The clinical team proposed to restrict the list of emotions to include dialogues between a patient and his boss, which is not relevant for the patients that this clinical team cares for. The theme is directly inspired by a role-play session impossible to meet.

**Prototype Dialogue**

The prototype we have developed comprises three modules. Modules 1 and 2 train the patient to recognize facial expressions, in or out of context, and are not described in this article. In the third module, which relies on our proposed architecture, the patient can maintain a conversation with the ECA. A list of answers is suggested to the patient and dialogue progresses depending on his choice and his facial expressions.

**Dialogue Construction**

Figure 2 shows the linear dialogue constructed with the clinical team. The theme is directly inspired by a role-play session run at the psychiatric care unit. Such exercises can teach the patients how to introduce themselves, start, maintain or interrupt a discussion, request or refuse something, receive or give a compliment, or criticism. In order for the exercises to be effective, themes relate to the patient’s experience. The game RC2S (Peyroux and Franck 2014) for example includes dialogues between a patient and his boss, which is not relevant for the patients that this clinical team cares for because most of them are unemployed.

The dialogue and its variations are implemented in a declarative manner following the narrative engine language: for each stage of conversation topic, declaration of transitions describe possible sentences depending on the resources currently in the environment (including user emotions). The following code extract implements possible answers to greetings by the ECA.

```plaintext
B * (pHappy + pNeutral) -@ subjectDone
  "agent/I am good thank you./aHappy"
B * (pSad + pAngry) -@ E
  "agent/Are you sure/?aNeutral"
B * pSad -@ G
  "agent/Me neither./aSad"
```

B denotes here a particular stage in the dialog. Various actions are defined using the linear implication -@, each associated to a dialogue utterance (here by the agent). pHappy, pNeutral, pSad and pAngry represent emotions that the emotional input sub-system will dynamically feed into the environment. If the emotional input sub-system interprets the user to be sad, either the second or third action may be selected. (+) and (*) are respectively choice and conjunction operators. Annotations associated to the declaration are then processed and sent to the ECA sub-system.

**Running Example**

Figure 3 represents the ECA asking for money. In the modeled situation, the patient should refuse to help in an appropriate way. In the list of answers suggested, each one reflects a different tendency: for instance, the first one (“Ok I will give you that”) reflects the impulse to always say “yes” when the third one (“Yes but you have to give me the money back tomorrow.”) reflects a tendency to agree but with conditions impossible to meet.

N-@Yes "patient/Okay I will give you this"
N-@NoPart "patient/Yes but you have to give it back tomorrow!"

---

Figure 2: Example of dialogue module
The reaction of the ECA and ensuing dialogue will be different for each answer selected and facial expression of the patient, providing a naturalistic feedback to the patient, who may safely experiment and reflect on the expect of his expressions and actions by going back to a previous point in the dialogue.

Conclusion and Future Work

We have presented an architecture to support a specific therapeutic approach to improve schizophrenia patients everyday-life functioning, as well as a proof of concept application. With this work, we have paved the way for the construction of a more sizeable development. At the moment, we are working on more example dialogues, which would allow to test the acceptance of the application in the clinic.

The role of debriefing is considered as very important in order to ensure the transfer of new skills from the application to everyday-life in our settings (Thompson et al. 2010; Peyroux and Franck 2014). In order to improve this aspect in our application, we plan to investigate the potential offered by a linear logic based generation engine: they allow to present the events in the form of a causal diagram structure (Martens et al. 2013), which may help patients reflect further about the consequences of their actions. After a technical evaluation of the variability level of the system, we will evaluate its usability and performance in clinical settings to see if the potential of this approach is verified.

As also noted by (Tanaka et al. 2015), deficit in social skills is aspecific, and the approach we propose here may be usable for other psychiatric disorders such as Autistic Spectrum Disorders for instance.

Acknowledgments

We would like to thank Dr Lucie Debarre for her help in reviewing the state of the art, and the clinical team at the Ponchelet day-care center in Brest.

Part of this work has been supported by the ENI Brest mobility fund.

Part of the research described in this article was developed with funding from the National Science Foundation grant number IIS-1423260.

References


Green, M. F. 1996. What are the functional consequences of neurocognitive deficits in schizophrenia? *The American Journal of psychiatry*.


