

Designing Natural-Language Game Conversations

Jonathan Lessard

Concordia University

jonathan.lessard@concordia.ca

ABSTRACT

This paper reports on LabLabLab's three year experience in game-design oriented research on interactive dialogue with non-playing characters and developing natural-language conversational games. It explores the specific affordances and constraints of natural-language interaction for game conversations and offers strategies for their effective design. It also examines the general notion of conversational puzzle and proposes interface-agnostic design approaches founded on the concepts of cognitive conflict and conversational moves.

Keywords

Game design, natural language interaction, dialogue systems, narrative design, interactive storytelling, conversational games, conversational puzzle, conversational move, cognitive conflict, conversation modeling, puzzle design

INTRODUCTION

The initial impulse of the LabLabLab research-creation project (initiated in 2013) was to explore alternatives to dominant patterns in the design of interactive conversations in video games. The models for game dialogue systems in mainstream video games have remained essentially the same since late 1980s. The main ones being the familiar dialogue trees of predefined utterances and, as Bruski and Björk put it, the “‘database retrieval’ style” (2009), in which players select from a list of topics to acquire information from non-playing characters (NPCs). These are sometimes spiced up by making the available options dependant on quest flags or character attributes, but remain similar in that the player can always only choose within a short selection of predetermined inputs.

As with all game mechanics, there is of course nothing intrinsically “wrong” with a menu-driven approach to interactive conversation; but it does offer a specific set of affordances and constraints that should be acknowledged. For example, testing various dialogue systems on the same interactive drama, Sali *et al.* (2010), report that sentence selection “appears to maximize story involvement”, abstract response menu interface “maximized reasoning about the underlying game structure”, and natural language understanding “maximized a sense of presence and engagement with the overall experience”.

Though this project is related to existing research on natural language in games and interactive storytelling¹, LabLabLab's specific outlook is that of *game design*. Its main purpose is investigating how interactive conversations can be crafted as games themselves. The project's focus is conversational *gameplay* (rather than believability or drama)—that is attempting to reach a specific outcome through a series of conversational

Proceedings of 1st International Joint Conference of DiGRA and FDG

© 2016 Authors. Personal and educational classroom use of this paper is allowed, commercial use requires specific permission from the author.

“moves”. This is possible with menu-based systems, and we’ll see examples of that later, but it feels extremely limiting when compared to the experience of “real” conversation where a very wide spectrum of moves is available. Here, the ideal model (as is often the case with interactive storytelling issues) is the live or tabletop role-playing game (RPG) in which players can devise and perform their avatars’ utterances at the most fine-grained level: choosing wording, tone, accent, *etc.* Of course, it should come as no surprise that RPG conversations feel just like natural human conversation since that’s exactly what they are.

Ironically, natural-language interaction (NLI) is exactly where narrative-based computer games come from. Those are usually traced back to *Adventure* (1977) who was itself (in some respects) a computerized remediation of RPGs but also inspired by contemporary natural language interaction experiments (Lessard 2013a; Montfort 2003) such as the famous *Eliza* program (Weizenbaum 1965). In fact, NLI was a common feature of computer games until the late 1980s. Towards the end of the decade, menu-based interactions and mouse-driven graphical interfaces progressively replaced the traditional parsers and most players today have never had to type a word of text within a digital game. Although the move to GUI was perceived as “progress”, some qualities of the original experience had to be sacrificed to profit from the new interfaces’ much clearer affordances minimized input errors (Lessard 2013b). In his *Guide to Adventure Games* published in 1984, Gary McGath wrote: “[...] for telling the computer what you want to do, there is no question that words are more flexible than any joystick or trackball”.

LabLabLab’s research proposition was to revive NLI for NPC conversation and map out its affordances and constraints. The rise of natural-language agents such as Siri or Cortana confirms both the timeliness and relevance of this effort. As mainstream users become re-acquainted with NLI, we can expect a rising demand for games playing with those modalities. Games in development such as *Event [0]* and *Bot Colony* may be commercial forerunners though we must recognize the academic precedent set by *Façade* (Mateas and Stern 2005). LabLabLab draws inspiration from *Façade* for its choice of NLI but also for its research through creation approach. The intended contribution is to explore the game design potential of NLI as well as better document the design space of conversational games in abstraction of any specific form of interaction.

This paper acts as capstone report for the first three years of LabLabLab that saw the production of a series of three experimental game prototypes. It reflects a research-creation methodology (more specifically research through design) in which the design and development activities represent a key source of knowledge production that is then embodied in the actual prototypes (Godin & Zaheri 2014). The interpretations presented here are supported by references to the artifacts themselves (they are available for the reader to consult and make her own judgment), professional observations made during the process, feedback from experts as well as reactions from “real world” gaming communities (the prototypes were made available on online game portals and were the object of evaluation and direct comments).

After describing the three game prototypes that will serve as main reference, the discussion will turn to develop the concept of “conversational puzzles” and their design; and afterwards tackle the design-related issues of using natural language interaction for game conversations as opposed to alternative mechanisms of interaction.

LABLABLAB GAMES

Between 2013 and 2016, LabLabLab developed three digital games of generally similar format with varying content: *A Tough Sell* (2014), *SimProphet* (2015), *SimHamlet* (2016)². They are all single-screen games staging dialogue situations between a player character and a NPC (see Figure 1). An optional tutorial briefly informs the player of the conversation's fictional context and of its desired outcome.

The player is invited to type the desired character utterance in a text window before validating. The line is then displayed on-screen, triggering an answer from the NPC. The speech of each character is conveyed by cartoonish bubbles. This textual exchange constitutes the only form of input and the main form of feedback, though additional visual signs are embedded in the screen to inform on progress and current game state. In essence, these interfaces could be described as fancy chat rooms. The conversation history log is not immediately visible but can be summoned via a button.



Figure 1: Screenshots from *AnonymousLab* games (from left to right): *A Tough Sell* (2014), *SimProphet* (2015), and *SimHamlet* (2016).

These games are built on two main technologies. The natural language processing component is the open-source [ChatScript](#) chatbot engine by Bruce Wilcox which functions as a server. This technology was chosen because it was mature (having won multiple Loebner prizes between 2010 and 2015), open source, well documented, and, most importantly, featuring a very legible script language accessible to non-specialist content authors (Wilcox 2011).

The client applications were developed with the [Unity](#) engine and initially distributed as Unity Web Player then WebGL content. Players can interact with the games online from within a web browser without having to download and install anything. The actual conversational code of the characters is hosted on a single server which keeps records of all player logs (independent of where the client application is hosted) and can be updated without re-publishing all clients. The three games were published on the following free online gaming sites in order to reach actual gaming communities: [Newgrounds](#), [Kongregate](#) and [Gamejolt](#).

A Tough Sell

In LabLabLab's first prototype, *A Tough Sell* (2014), the player is cast as the Evil Queen of the *Snow White* fairy tale. The action begins *in media res*, at the specific moment when the queen knocks on Snow White's door (actually, the seven dwarves' door) disguised as an innocent old woman. Her intent is to have Snow White eat the poisonous apple she's prepared. At this point of the narrative, Snow White is aware that her stepmother is trying to kill her and is quite wary of this stranger offering her an unsolicited apple.

This prototype explores “persuasion” as its core conversational objective. It revolves on an economy of “trust” which is internally represented as a simple integer variable and exposed as a progress bar. This trust bar will fill or empty according to Snow White’s reaction to player inputs. When it is filled, Snow White will accept to eat the apple, having now full confidence in the old woman. The game also displays a “patience” bar which progressively empties until Snow White has enough and closes the door.

Players are challenged to build an understanding of Snow White’s personality on the basis of her answers to devise and test trust-building approaches. A variety of dialogue moves will work towards that end: offering a good justification for a stranger offering an apple, inquiring about and empathizing with her current situation, delivering false news about the Evil Queen’s death, *etc.* Other moves like threats and contradicting statements will go the opposite way (see Transcript 1 for an excerpt).

The game was played around 13 000 times at the time of writing and received ratings averaging 3.4/5. Players report enjoying coming up with ways to deceive the NPC: “This game is so fun! I really enjoyed both making her mistrust me and leading her to trust me. It is very interactive with a very smart A.I. It was great!” (bunnys2 on *Newgrounds* 2014). Some also confess to the guilty pleasure of successfully realizing an evil character’s plans: “5/5, would kill an innocent girl to satisfy my vanity again” (LordFell on *Gamejolt* 2015). Many comments include interesting or funny excerpts from the conversation. Common complaints concern the game being too easy and Snow White misunderstanding inputs. The latter was especially an issue when miscommunication led to an irrelevant gain of trust. Unsurprisingly, players enjoy winning as a result of their clever inputs and not of system bugs.

```
# Excuse me, but I don't understand who you are and what you're doing here.
> I just wanted to see if you wanted this fresh, juicy apple.
# It does look fresh. (trust+1)
> And juicy too! I've heard your stepmother is around here and wanted to warn you
# Thanks but I know about my stepmother's plan to assassinate me. (trust+1)
> Try to kill you, such a sweet girl!? What a hag!
# You are a so understanding. (trust+1)
> I would like to punch your stepmother in the head
# You would really do that?! Thanks! (trust+1)
```

Transcript 1: A player convincing Snow White to eat the poisonous apple in *A Tough Sell* (2013)

SimProphet

SimProphet (2015) explores how NLI can allow players to contribute new elements to the conversation. The player character (invisible and undefined) has allegedly met some form of transcendent being and is on a mission to spread its message. The game begins as the player character encounters his first potential convert. The NPC is Ambar, a basic Sumerian shepherd surrounded by his sheep. Ambar is very curious about the details of this newfangled religion and leads the conversation by asking questions related to various issues such as cosmogony, worship practices, priesthood, sacred buildings, taboos, *etc.* The game itself has little specific narrative content but challenges the player to invent answers to these questions, progressively building the outline of a fictitious faith. Ambar will remember some of the player’s propositions and use them as material for comebacks or follow-up questions (see Transcript 2). The game also displays two meters representing the level of conversion for both Ambar and his sheep. Although the explicit

goal is to convert both, this objective acts more as a pretext to frame an often comedic exchange between the player and Ambar.

```
# Anything special happens to my body once I'm dead, as a Birdena?  
> It will be eaten by birds  
# A classy god wouldn't let birds eat my body.  
---  
# What will Pepperoni Titan do to those who refuse to adore it?  
> Kiss them on the lips  
# I'd hate to be kissed by a God.
```

Transcript 2: *SimProphet's* NPC is not convinced.

The game was played approximately 12 500 times and earned reviews averaging 3.6/5. It was featured on the front page of *Gamejolt* which led to an important traffic and many comments. This specific community enjoyed the game particularly with an average rating of 4.4/5 and numerous very positive comments such as: “Just 3 words: Best game ever” (knightrunner 2015), “Amazing game, I love it.” (EchoDJ 2015), “this is HILARIOUS” (hexiel 2015), etc. The game’s unique form of AI-collaborative comedy routine is the most noted aspect though some commentators actually adhered to its simulative pretenses: “It provides an excellent training for evangelists like me. The questions are realistic and I get to use whatever style I want to answer them. This is truly a unique idea” (AlexMario_Media 2015).

SimHamlet

LabLabLab’s third prototype focuses on “interrogation” as main conversational objective. Game dialogue is often a means for players to acquire information on the fictional setting though it is most often a simple matter of systematically going through all available topics and rarely a challenge in itself. *SimHamlet* explores the potential gameplay of retrieving information from a reluctant or non-cooperative character. The game begins in the aftermath of Shakespeare’s play. The player is cast as a government official with the mission of clarifying the recent events to write an official report. A gravedigger must be interrogated in order to establish the “how”, “when”, “where”, “why” and “by whom” of each murder. He apparently knows everything there is to know however the process is complicated by the NPC’s idiosyncratic perception of the events. As the player progresses, the epitaphs on the victims’ tombstones are increasingly completed.

```
> How did Ophelia die?  
# She stopped breathing, was a very bad idea.  
> Why did she stop breathing?  
# How am I supposed to know why people do things? I wasn't there when it happened!  
> Where did she die?  
# Well I imagine it was in the cold water she was immersed in.  
> What water?  
# Well, sure, she was immersed in the river. Can't be immersed in a cup of water!
```

Transcript 3: A player-driven interrogation in *SimHamlet* (2016)

DESIGNING CONVERSATIONAL PUZZLES

Developing the LabLabLab games did not only raise issues concerning natural-language interaction (which will be covered in next section) but also the more general problem of

designing challenging interactive conversations with fictional characters. Concepts and approaches that were developed to assist this design process will be presented in this section. These were informed not only by experience, but also by the analysis of existing games and borrowing from linguistic and computational modeling of conversation.

Conversational puzzles

Goal oriented challenging conversations with NPCs are not uncommon in video games. They are usually found in narrative driven games such as adventure or computer role-playing games. Amongst the most famous examples is the insult sword-fighting of *The Secret of Monkey Island* (Lucasfilm 1990) which consists in learning a number of pirate insults and their appropriate comebacks in order to win insult duels. More modern examples include choosing the adequate lines to seduce NPCs in *Dragon Age: Origins* (Bioware 2009) series, selecting the right attitudes to expose suspects in *L.A. Noire* interrogations (Team Bondi 2011) or to persuade characters to act the way you want in *Deus Ex: Human Revolution* (Eidos Montreal 2011).

These types of in-game challenges are often considered puzzles as they are problems with a finite number of pre-determined solutions. They can be considered as a subset of the general category of “fiction puzzles” (expression by Karhulahti 2014). Comparing with a jigsaw puzzle, the fiction puzzle requires a player to piece together story rather than picture fragments; in order to reconstruct one of the predefined valid narratives in lieu of a reference image. This is the staple of story-based progression games: in order to generate the valid story of King Graham in *King’s Quest V* (Sierra On-Line 1990), one must have, amongst other things, have found a silver coin, bought a pie, to eventually through the pie in the face of a yeti, so that said yeti may fall down a cliff. Conversational (or dialogue) puzzles proceed from a similar logic except that their “pieces” are utterances between characters—or “conversational moves”.

Conversational Moves

I use here the expression “conversational move” borrowed from the linguistic theories of conversational (or dialogue) games (Schiffrin 2005, p. 120) to abstract the notion from any specific game mechanic such as dialogue trees, topic or attitude selection and, of course, natural language interaction. Conversational moves represent a single or a series of utterances intending to change the state of the conversation, that is: to make a point. They are not attached to specific wordings and the same move can be performed in multiple ways. For example, the conversational move “sympathize with Snow White” in *A Tough Sell* could be worded as “Oh, poor thing”, “This is so unfair”, *etc.* The purpose here is not to use rigorous linguistic terminology, but rather find a convenient notion with an appropriate level of abstraction for the design of NPC conversations. Conversational puzzle-solving would then consist in playing the right moves at the right time in order to reconstruct a conversation leading to a desired outcome.

Different dialogue systems complicate this process in different ways. Dialogue trees are challenging because they require players to discover more or less strict sequences of moves. A blog post on “bad” adventure game puzzles states: “In *Countdown*, only trial and error can yield the correct chain of dialogue to trigger the correct response from the informant” (Luoranen 2009). In some systems, making a particular move available is the obstacle. For example, an event in the game world must have been triggered in order for the player character to have a specific line to say to a NPC. Many *Dragon Age* (Bioware 2009) seduction conversational puzzles require the player to perform quests and find specific items to “unlock” romance-enhancing moves. In open NLI games like

LabLabLab’s, players are confronted with the difficulty of coming up with relevant moves and formulating them in such a way that the game understands rather than selecting them from a list.

Defining an objective

Conversations are not necessarily goal-oriented. A conversational puzzle, however, implies a desirable outcome that is not trivial to achieve. The first step in designing one is thus to establish what conversational state needs to be reached by the player and the reasons why this is a problem. If the objective is to convince a NPC to give the key to a door and this can be done by simply asking, you have the conversational equivalent to a one-piece jigsaw puzzle. Although there might be other ways to change someone’s mind (physical violence, material bribes, suggestive body language, *etc.*), we’re looking here for problems that can be solved solely through talking—which puts us in the realm of *argumentation*.

In his work on modeling argumentation in everyday conversation, Jean-Louis Dessalles notes that it is not a routine activity:

Conversational argumentation can potentially deal with any issue. Contrary to many verbal tasks of daily life, like ordering a taxi, there is no pre-definite script for such interactions. Arguments cannot be retrieved from previous mastery of dialogue games [...] and must be computed anew (Dessalles 2008).

This partly explains the difficulty in designing more systemic or procedural approaches to conversations with fictional characters. From our perspective, this also means that any specific argument situation is an opportunity for an original conversational puzzle. To better define those situations in a way that will help us break them down in “puzzle pieces” we can further follow Dessalles in observing that: “aspects of argumentation have to do with incompatible beliefs and desires and with belief revision” (2008). These beliefs and desires can be found in various “strengths” (positive or negative) and a “cognitive conflict” occurs when, in a conversation, two people realize they attribute opposite strengths to a same proposition. In order to resolve this conflict, conversational moves can be played from all parties in order to revise those beliefs until they are of equivalent strength.

Cutting out the pieces

In *A Tough Sell*, we can define the main conversational problem as: (1) the Evil Queen desires that Snow White eats the poisoned apple and (2) Snow White believes it is dangerous for her to accept food from a stranger³. This conflict is made immediately manifest when the player offers the apple to Snow White and discovers that she will not touch it. It then becomes clear that conversational work will have to be done to bring her to stop considering this apple as a threat.

What is less clear is what kind of work the player will have to do to reach the desired conversational goal. A useful approach here is to unpack the main conflict into more granular constituents. Again, in the case of Snow White, we can identify a series of sub-beliefs that inform her general attitude towards the apple. Having this in hand, we’re well off in defining resolution and aggravation moves for this puzzle:

Subconflict #1	Snow White believes the old woman is a stranger.
Resolution moves	a) Pretend to be an itinerant apple peddler

	b) Pretend to be lost c) Pretend to be amnesiac
Aggravating moves	a) Pretend to be a neighbor b) Pretend to be a family member c) Tell true identity (stepmother)
Subconflict #2	Snow White believes her stepmother is trying to kill her
Resolution moves	a) Pretend her stepmother is dead b) Pretend her stepmother wants to make amends
Subconflict #3	Snow White believes the dwarves are well-intentioned when they say she shouldn't talk to strangers.
Resolution moves	a) Praise the dwarves. b) Suggest the dwarves are retaining Snow White as a domestic slave.
Aggravation moves	a) Insult the dwarves.

Table 1: Part of the cognitive conflicts to be solved in *A Tough Sell* (2014) and some of their associated conversational moves.

In order to come up with such a list of moves, it is also very useful to map out the initial conversational state from the perspective of the participating characters. This includes the “initial assumed common ground” (Shiffrin 2005, p. 203). In *A Tough Sell*, this common ground is quite thin as the two characters are supposed to be complete strangers to each other. In *SimHamlet*, the gravedigger is aware of the player character’s role and purpose so that communicating this knowledge needs not be the object of conversational moves.

To this “public” common ground, we can further detail respective relevant private knowledge. In *A Tough Sell*, the knowledge distribution is initially very unbalanced as Snow White doesn’t know her interlocutor is the Evil Queen and that the apple is poisoned. On the other side, the Evil Queen doesn’t necessarily know that Snow White is living with seven dwarves that have forbidden her to talk to anyone. Acquiring this information in order to exploit it can be the object of interesting conversation moves.

This leads us to recognize that the conversational state also include the “public utterances so far” (Shiffrin 2005, p. 203). The Evil Queen cannot be expected to say something about the dwarves unless their existence has previously been established during the conversation. Though a complete simulationist model of conversation would need to rigorously keep track of all the changes to the state, defining a few rules of entailment to conversational moves can suffice to afford chained argumentation and help give a sense of a progressive shift in the NPC’s mental state. In *A Tough Sell*, for example, the move “claim good intentions” will only come through if some trust points have already been established, thus reinforcing an established favorable impression.

A Tough Sell offers the example of a clear-cut conflictual situation. The notion of cognitive conflict can also be understood in a broader sense, encompassing such things as misunderstanding and doubt. For example, unpacking the romance option between *Dragon Age: Origins* (Bioware 2008)’s player character and NPC Morrigan reveals cognitive conflicts that mostly amount to rectifying preconceptions about the other. For one thing, Morrigan seems to presume that the Warden (the player character) does not appreciate her as she is. Amongst the moves that will change that belief are: stating that shapeshifters (like her) are useful, and praising her for being daring in her youthful explorations. Morrigan also seems to desire the Warden to approve of her witch mother, which can be done by recognizing the value of her seemingly harsh parenting methods.

Once again, knowing the gap between two characters' beliefs and desires highlights the possible steps that can be taken to resolve the cognitive conflict.

Puzzle-solving

Having a conversational objective as well as the conversational moves that could get a player there, we can now attend to the interactional aspects of solving this puzzle. Here, the specifics of the chosen conversational system will affect greatly the structure and pacing of the actual puzzle solving (and we'll see in the next section the particular affordances of natural language interaction in this context). However, if we accept this particular form of interactive conversation to be a sort of puzzle, we can assume that general puzzle design guidelines will apply. Let's reproduce here designer Jesse Schell's tips on the topic:

1. Make the goal easily understood
2. Make it easy to get started
3. Give a sense of progress
4. Give a sense of solvability
5. Increase difficulty gradually
6. Parallelism lets the player rest (Schell 2007).

Adapting this advice to our current approach, we could start by suggesting we make the main cognitive conflict clear as soon as possible. This helps players understand their character's as well as the NPC's mental states and motives, and appeals to the common urge to engage in arguments. Considering subconflicts as milestones in the process, we should try to expose a basic one in the early stages. For example, the moment she's offered the apple (which is usually amongst players' first move), Snow White wants to know who the player is to be offering an apple to stranger, thus revealing a key subconflict. This direct question helps players get started by trying to find a plausible answer such as: "I'm just an old woman walking in the woods". Early successes help give a "sense of solvability".

Having milestone objectives is also useful to give players "a sense of progress". In *A Tough Sell* and *SimProphet*, this takes the form of simple progress bars which fill up when points are made towards the objective. This might seem a bit crude but it was found that the LabLabLab games were unfamiliar enough in their form that this simple, explicit feedback helped players stay in tune with the games' proposition. In *SimHamlet*, the sub-objectives are more explicitly singled out and the player can gauge progress made in each of them individually. All the LabLabLab games allow for nonlinear puzzle solving ("parallelism"), letting players tackle subconflicts in no specific order, jumping back and forth as potential conversational moves occur to them. This naturally establishes a progression in difficulty as players resolve the subconflicts that seem the most obvious to them, leaving the more challenging ones for later.

Depending on the chosen interactive dialogue system, nonlinear argumentation is not always possible. Also, as we've seen earlier, some conversational moves may depend on aspects of the conversational state having been established earlier. Enforcing an ordered chain of conversational moves can be a way to increase difficulty or to ensure a more dramatic progression in the shift of one psychological state to the other.

In *SimHamlet*, the NPC will initially repeat the official story according to which the old king of Denmark (Hamlet's father) has died from a snake bite. The player needs to

question this assertion or make the observation that the NPC seems nervous to learn that the latter has received threats concerning this information and fears for his life. The player can then relieve the NPC by reminding him that everyone is dead in this story to obtain the full confession.

Designing a fun conversational puzzle for players to solve is a difficult task that will greatly vary depending on specific fictional contexts. Though it is by no means the only way to tackle this problem, the cognitive conflict and conversational moves approach adopted by LabLabLab in the development of its three games has proven useful to break down the main objectives into a series of manageable, smaller-scale challenges. We believe it could be of use as a narrative design methodology for most forms of interactive conversations. However, an important part of the LabLabLab project contribution revolves specifically around the use of natural language interaction.

NLI FOR GAME CONVERSATIONS

Affordances of Natural Language Interaction

We've outlined in introduction LabLabLab working hypothesis that natural language interaction (NLI) might be a means towards more interesting (or at least different kinds of) conversation with NPCs. In a previous publication (Lessard 2015) were outlined some unique affordances of NLI for game conversations which we'll summarize here.

Creative Conversational Play

Menu-driven dialogue systems allow players to select between a few predetermined utterances, leading to another such menu, and so on. The options being explicit, any challenge can only be derived from finding the right *path* through that node-based graph. NLI, on the opposite, allows players to formulate (literally and metaphorically) their own “conversational moves”, devising rhetorical tactics informed by their understanding of the interlocutor's personality and the state of the discussion. NLI offers a shift of initiative, putting players in a situation to *act* upon the conversation rather than always *react* to a proposition. This shift allows players to feel they've generated the solution (and feel ownership towards it) rather than having simply *found* it.

Role-Playing

Some games (often computer role-playing games) afford a deep level of customization for player characters. Players can become very invested in their avatar, projecting personality traits over the attributes they've largely contributed to define. However, when comes the time to engage a conversation with a NPC, the player is typically given a handful of possible lines that might not reflect at all how one would imagine that character to talk. Some systems will partially acknowledge character traits in the selection of proposed dialogue lines, but this can only go so far as all this content needs to be handcrafted in advance. NLI, on the opposite, leaves complete room for players to converse “in character”, fleshing out their avatar through personality-laden discourse. This does not mean that the game will necessarily acknowledge all aspects of the characterization but players can at least have the satisfaction of maintaining their avatar's coherence at the discourse level.

Contributing Fictional Content

Menu-driven conversations leave no opportunities for players to introduce any element that hasn't been pre-determined. NLI opens room for players to actually provide new content that can be (to an extent) acknowledged by the game. This aspect was the main

focus of *SimProphet*. The player logs reproduced in Transcript 2 show conversational exchanges that would have been impossible without natural language interaction. Here, players have not only defined the names of the deities and of their followers, but also such specific notions as birds eating a follower's body after death, or the kiss of a god as a punishment. None of these things were predetermined, only the notions of *something* happening to a body after death and divine punishment.

Nonlinear Conversations

In menu-driven systems, available moves at any given moment are very dependent on the exact state of the conversation. In order to say something that was previously accessible, one needs to find the sequence of choices that will lead back to a previous menu offering. Some crucial dialogue line can sometimes be buried deeply in a tree and then become unavailable if the opportunity to say it was missed. Actual human conversations don't work like this. Almost any topic previously mentioned can always be immediately reactivated (as long as both parties remember about it) and it is common for speakers to jump from one thread to the other and back. NLI allows a nearly stateless structure in which most content is always available.

Issues of Natural Language Interaction

Considering methods for implementing NLI have been known since the 1960s, there are certainly reasons why it isn't the prevalent mode of interaction in games. Here are amongst the main issues that narrative designers will have to deal with when working with NLI.

The "Eliza Effect"

The beauty of NLI is that it gives players the opportunity to "talk" to the system as if it were a person. The counterpart to that affordance is that it creates the expectation that the system will understand and respond like a person. And indeed, NLI often attempts to fulfill that expectation by answering in a human-like manner. However no NLI systems yet have managed to maintain that illusion very long and players will sooner or later (usually sooner) witness the NPC saying something utterly irrelevant (see Transcript 4). As Wardrip-Fruin summarizes: "The Eliza effect creates a surface illusion of system complexity—which play (if allowed) dispels" (2012, p. 146). Menu-driven conversations avoid this issue by: "never building up the Eliza illusion and instead clearly representing the operations of a simple system on the work's surface" (p. 15).

> Have your stepmother ever tried to kill a dwarf? # I don't have any stepmothers > Pardon me? # I hate to be rude but there's an awful lot to do in this house > Who's trying to kill you? # My stepmother wants to kill me so that she can claim to be the most beautiful living woman. > But you just told me you don't have a stepmother! # How do I know that? > Did you take your medication today my child? # I guess that makes sense.

Transcript 4: Miscommunication in *A Tough Sell*

Leaky Fictional Coherence

By conserving strict control over everything that can be said and answered, menu-driven conversational systems can easily maintain fictional coherence. With current NLI systems, it is to be assumed that miscommunication will happen, which might be damaging to fictional coherence and immersion. In Transcript 4, for example, Snow White claims she has no stepmother, voiding temporarily the key issue of the plot.

Amnesia

This is not so much linked to NLI itself but LabLabLab's choice to use technology designed for chatbots. Chatbots are concerned with responding something relevant to an input but often do not have a strong model of the conversational state and its evolution. This affords the nonlinear conversations and puzzle-solving discussed earlier as most of the scripted answers are available at all times. The counterpart is that chatbots are not at their best in sequences of exchanges. Important points made in the conversation can be tracked manually but it often occurs that the NPC will say something that makes it seem like it has forgotten things that were already said. It is obvious in Transcript 4 that Snow White has no recollection of having said she did not have a stepmother.

Unrestricted Input

An interesting affordance of NLI is that it allows players to say anything. A major problem of NLI is that it allows players to say anything. In other words, no amount of scripting will ensure that the NPC has a relevant answer for everything the player might come up with. In Transcript 4, the player asks whether Snow White has taken her medication and the system doesn't have (yet) an appropriate answer.

This is not only an issue for the designer; this is also a user-experience issue. Free text input brings back command line interaction difficulties such as what Donald Norman called "the tyranny of the blank screen" (2002). With no explicit options to choose from, the user can easily be at a loss as to what to do next.

Designing with NLI

This mixed account of NLI underlines why it is not to be considered as an objectively superior replacement for other systems but rather as an interesting alternative offering unique possibilities. In the following subsections will be presented some approaches developed by LabLabLab to make the best of NLI's affordances and constraints.

Scripting the Interactor

A common issue with chatbots is that they claim no other purpose than to pass as human conversationalists. Users are often at a loss as to what they could be talking about with them and often take this encounter as a challenge to expose the non-humanness of the agent. A common behavior is thus attempting to "break" or expose the bot as machine. Indeed this can prove to be quite fun though usually also quite easy.

In *Hamlet on the Holodeck*, Janet Murray wrote that:

The lesson of *Zork* is that the first step in making an enticing narrative world is to script the interactor. The Dungeons and Dragons adventure format provided an appropriate repertoire of actions that players could be expected to know before they entered the program (Murray 1997, p. 78).

A fruitful approach to NLI game conversations is to seduce the player into *playing* along rather than playing *against*. This can be done by providing an understandable fictional situation as well as an objective that offers an interesting challenge. If players buy into the fiction, they will have a good reason to explore the designed conversational space for what it can offer rather than finding rewards mostly in exposing its limits. In this context, the glitches of NLI will be interpreted as the unavoidable boundaries of any storytelling machine—just as the lack of choices in menu-based systems can be seen as shortcomings. Of course most players will enjoy ridiculing the NPC at times but if they care enough, they will come back on track to pursue with the fiction. Film viewers making jokes of a supporting role’s acting might pop out momentarily of fictional immersion but they will as easily tune back in if they are committed to the fiction’s stakes.

Circumscribing the Conversational Domain

The fact that general chatbots pose as universal conversationalists, inviting discussions on any topic from politics to movies, passing by weather and philosophy, it them that much easier to break. Part of “scripting the interactor” is circumscribing the relevant conversational domain, setting up expectations as to what it is that we can talk about in the context. Once again, this form of tacit convention is very common. Readers and film viewers accept, for example, that the narrative will elude a large part of the characters’ lives in order to focus on the salient events. Even the richest transmedial worlds do not go into great details as to how, for example, one does the laundry on a spaceship or the specifics of hobbits’ dental hygiene.

A Tough Sell probably has the most restricted domain among LabLabLab’s prototypes. The Evil Queen is posing as a stranger which means she and Snow White have very little common ground they could talk about. This is further restrained by the “doorstep” nature of the conversation which requires being brief and to the point—unless the visitor is allowed in, which won’t be the case. In this context, the player understands that Snow White will not be receptive to small talk unrelated to the stranger’s identity and purpose. *SimHamlet* is also quite constrained and makes clear that any input not relevant to the tragic murders will likely be ignored.

All aspects of the fiction can contribute to limiting the scope: the characters, the context, the conversational objective, as well as the duration of the encounter. Smaller domains not only help managing expectations but also allow developers to focus on a limited set of possible moves and make that conversational space that much richer.

Funneling

In the general design section of this paper we’ve suggested breaking up the solving of conversational problems as series of relevant conversational moves. The problem is how to reconcile this finite repertoire with the potentially infinite number of player inputs. LabLabLab’s approach is to “funnel” wide portions of varying natural language formulations towards a limited number of relevant moves. Although those funneled utterances are not exactly equivalent in meaning, they are considered to be similar enough in intent for the scripted answer to feel acceptably relevant. This structure also helps a lot in making the script readable and scalable.

Error Handling

When choosing NLI, the designer must acknowledge that errors *will* happen. Choosing how to deal with those errors is an important part of crafting the experience. We recognize three types of errors: true negatives (TN) mean the system is right in thinking it

has no answer to the current input; false negatives (FN) occur when the system is wrong in considering there is no valid output for the current input; and false positives refer to the system being wrong in considering it has a valid answer.

False positives are the worse as the NPC is unaware that an error is happening and delivers an often incoherent response. False negatives are lost opportunities since an appropriate answer does exist for the player input except that the specific wording is not recognized; however they trigger some form of error handling and as such are not as damaging as FPs. Both can only be eliminated through offline testing and iterating as they are not recognized at runtime.

True Negatives, on the other hand, require designed answers that somehow address the miscommunication. Different approaches were tested with the LabLabLab prototypes. *A Tough Sell* uses TN errors as opportunities to steer the player back to the relevant conversational domain and also constitutes a form of hint giving. When Snow White doesn't know what to say, she cycles through general statements (related to current active topic) revealing her point of view on the situation and giving the player leads. She will say, for example: "It will be hard for me to trust anyone when I know my stepmother could be disguised to kill me". For *SimProphet's* Ambar, a TN error is simply a trigger to ask another question, bringing back the player immediately on track by prompting an answer. In *SimHamlet*, TN errors are made explicit by having the character shrug and display an interrogation mark, unambiguously inviting players to rephrase or try another approach.

Crowdsourcing

A NLI-driven dialogue system makes it possible to easily benefit from the creativity of testers throughout the development process. Instead of only informing developers of players' chosen conversational moves, play logs continuously reveal new, unplanned moves which can be used to augment the conversational puzzles. LabLabLab's experience is that testing should begin as early as possible, even with a minimal interactive framework, in order to get a good sense of the range of user inputs in the given conversational puzzle and assess the relative challenge of its components.

Odd Characters

NLI-driven characters will inevitably sound a bit odd at times. A way to circumvent this problem is to justify the oddity diegetically by casting an odd character. An early example is *Eliza's* psychoanalyst who justifiably returns many statements as questions. The most common solution is to make the NPC no more than what it actually is: a robot. Although there are relatively few NLI conversational games, most of them, *Façade* excepted, feature robots as main NPCs: *A Small Talk at the Back of Beyond* (Scriptwelder 2013), *Event [0]* (Ocelot Society 2014), *Bot Colony* (North Side 2014). This works very well, of course, but it does greatly restrict the scope of potential characters. The LabLabLab's NPCs all have their own excuses to be sometimes off: Snow White is young, naïve and stressed out while Ambar and the Gravedigger are simply dumb.

Humor

Besides using robots, another effective ploy to deal with NLI's shortcoming is to set a comedic tone. Gross miscommunications are inherently funny and it is worth considering embracing that tone rather than fighting it. In the best scenario, the game is functioning as comedy both when it's working as intended but also when it's failing. Funny excerpts shared by commentators of *SimProphet* feature almost as many unintended exchanges as

designed ones. Of course, this is also a very restrictive solution in terms of fictional scope.

CONCLUSION

Natural language interaction is, unsurprisingly, a very natural thing for players. Even though the LabLabLab prototypes are very different from the main genres of video games, all players (including very casual ones) could get started playing immediately with very few instructions. As new designs and technical solutions are found, conversation driven games (or aspects of) can be expected to grow in importance. LabLabLab's first series of prototypes represented a step in understanding the mechanics of conversational games and their design, as well as the specific affordances and constraints of NLI.

An outstanding limitation of current dialogue systems (including NLI) is their scripted nature which bounds them to the domain of puzzles with very little room for emergence. The next step for LabLabLab is to research ways to connect natural language conversation to more dynamic, procedural systems—a far from trivial step that would require computational modeling of NPCs and their perception of the game world coupled with methods of natural-language generation. Emerging research in those areas (such as Ryan *et al.* 2015a; 2015b, for example) are opening interesting opportunities.

ACKNOWLEDGMENTS

This research was funded by the Fonds de Recherche du Québec – Société et Culture.

ENDNOTES

1 Three “Games and NLP” workshops have been conducted in various AI-related conferences between 2012 and 2014.

2 Documentation and links to the playable prototypes can be found on lablablab.net.

3 Dessalles recognizes important differences between beliefs and desires but argues that they can effectively be treated equally in a simplified argumentative model (2008).

BIBLIOGRAPHY

Bioware (2009), *Dragon Age:Origins* [Video Game, Multiple Platforms], Electronic Arts.
Brusk, Jenny, et Staffan Björk. 2009. « Gameplay Design Patterns for Game Dialogues ». In *DiGRA '09 - Proceedings of the 2009 DiGRA International Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory*. Vol. 5. Brunel University.

Dessalles, Jean-Louis. 2008. « A Computational Model of Argumentation in Everyday Conversation: A Problem-Centred Approach ». In *Computational Models of Argument - Proceedings of COMMA 2008*, directed by Philippe Besnard, Sylvie Doutre, et Anthony Hunter. Amsterdam: IOS Press.

Eidos Montreal (2011), *Deux Ex: Human Revolution* [Video Game, Multiple Platforms], Square Enix.

Godin, Danny, et Mithra Zahedi. 2014. « Aspects of Research through Design ». In *Proceedings of DRS 2014: Design's Big Debates*. Umeå, Sweden: The Design Research Society.

Karhulahti, Veli-Matti. 2014. « Fiction Puzzle: Storable Challenge in Pragmatist Videogame Aesthetics ». *Philosophy and Technology* 27 (2): 201-20.

LabLabLab (2014). *A Tough Sell* [Browser-based Game].

LabLabLab (2015). *SimProphet* [Browser-based Game].

- LabLabLab (2016). *SimHamlet* [Browser-based Game].
- Lessard, Jonathan. 2013a. « Adventure Before Adventure Games A New Look at Crowther and Woods's Seminal Program ». *Games and Culture* 8 (3): 119-35.
- Lessard, Jonathan. 2013b. « Histoire formelle du jeu d'aventure sur ordinateur (le cas de l'Amérique du nord de 1976-1999) ». Ph. D. Cinema Studies, Montréal: Université de Montréal.
- Lessard, Jonathan. 2015. « Design Rationale for Natural-Language Based Game Conversations ». In *Proceedings of the 10th International Conference on the Foundations of Digital Games*. Pacific Grove: CA.
- Lucasfilm Games (1990), *The Secret of Monkey Island* [Computer Game], LucasArts.
- Luoranen, Adam. 2009. « Adventure game puzzles we have known and hated ». *Adventure Classic Gaming*.
<http://www.adventureclassicgaming.com/index.php/site/features/451/>.
- Mateas, Michael; Stern, Andrew (2005), *Façade* [Computer Interactive Fiction].
- McGath, Gary. 1984. *Compute's Guide to Adventure Games*. Radnor: Compute! Books.
- Montfort, Nick. 2003. *Twisty Little Passages: An Approach to Interactive Fiction*. Cambridge, MA: MIT Press.
- Murray, Janet H. 1997. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New York: The Free Press.
- Norman, Donald A. 2002. *The Design of Everyday Things*. New York: Basic Books.
- North Side. (2014). *Bot Colony* [Early Access Computer Game].
- Ocelot Society (2014). *Event [0]* [Computer Game in development].
- Ryan, James Owen, Andrew Max Fisher, Taylor Owen-Milner, Michael Mateas, et Noah Wardrip-Fruin. 2015b. « Toward Natural Language Generation by Humans ». In *8th Workshop on Intelligent Narrative Technologies and 4th Workshop on Social Believability in Games*. AAAI Press.
- Ryan, James Owen, Adam Summerville, Michael Mateas, et Noah Wardrip-Fruin. 2015c. « Toward Characters Who Observe, Tell, Misremember, and Lie ». In *2nd Workshop on Experimental AI in Games*. AAAI Press.
- Sali, Serdar., Noah Wardrip-Fruin, Steven Dow, Michael Mateas, Sri Kurniawan, Aaron A. Reed, et Ronald Liu. 2010. « Playing with Words: From Intuition to Evaluation of Game Dialogue Interfaces ». In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, 179–186.. New York, NY, USA: ACM.
- Schell, Jesse. 2008. *The Art of Game Design a Book of Lenses*. Amsterdam; Boston: Elsevier/Morgan Kaufmann.
- Scriptwelder (2013). *A Small Talk at the Back of Beyond* [Browser-based game].
- Schiffrin, Amanda. 2005. « Modelling Speech Acts in Conversational Discourse ». PhD Thesis, University of Leeds.
- Sierra On-Line (1990), *King's Quest V: Absence Makes the Hear go Yonder!* [Computer Game], Sierra On-Line.
- Team Bondi (2011), *LA Noire* [Video Game, Multiple Platforms], Rockstar Game.
- Wardrip-Fruin, Noah. 2012. *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. Cambridge, MA : The MIT Press.
- Weizenbaum, J. (1965). *Eliza: Doctor* [Mainframe Computer Program].
- Wilcox, Bruce. 2011. « Beyond Façade : Pattern Matching for Natural Language Applications ». Gamasutra.
http://www.gamasutra.com/view/feature/134675/beyond_fa%C3%A7ade_patte rn_matching_php?page=1.