Games as Conversation

Rogelio E. Cardona-Rivera and R. Michael Young
Liquid Narrative Group
Department of Computer Science
North Carolina State University
Raleigh, North Carolina 27695-8206
recardon@ncsu.edu, young@csc.ncsu.edu

Abstract
We present a metaphor through which to study games: games as conversation, which casts gameplay as a communicative exchange between player and game. We propose to view aspects of gameplay as speech acts, as defined by Austin and Searle, and present several examples that illustrate the diverse locutionary, illocutionary, and perlocutionary acts present in the design of digital games. Through our perspective, we are able to cast problems relevant to the interactive entertainment community as discourse problems, where an interactive system must determine what to “say,” in order to elicit in the minds of players a specific mental model that will allow them to perform successfully in the game. We conclude with a research agenda that proposes to leverage the artificial intelligence paradigm of discourse planning to tackle the discourse problems of interactive entertainment.

In seeking to establish a critical discourse for the study of games, the field of game design has relied on metaphors such as “games as X,” (Salen and Zimmerman 2003) or “games viewed from the lens of X” (Schell 2008), where X is typically either a discipline or a framework from which we glean useful insights for understanding or developing games. In this paper, we introduce a new perspective to this critical discourse: games as conversation. Through this perspective, we frame games as a context for communicative exchange between a game player and the game. We contend that this framing provides a fresh perspective on problems of interest within the interactive digital games community, and also provides for a research agenda to model game-related phenomena of interest. While our interest is in the computational study of interactive entertainment, we suspect that our framework will be helpful for the study of games from other, non-computational perspectives.

Importantly, we do not intend “communicative exchange” to mean the use of games to rhetorically communicate ideas (Bogost 2007). Rather, we intend to present the act of gameplay as being a dialogue between the player and the game (as represented through the designed in-game artifacts). In the sections that follow, we examine why we feel gameplay is fundamentally a communicative exchange, we discuss the implications of treating games as conversation, and we inspect relevant problems in the interactive digital games community viewed as discourse problems. We then identify how our framework is amenable to computationally-oriented inquiry, and present a research agenda for future work.

Why treat Games as Conversation?
Communication as Fundamental Human Activity

One of the insights of phenomenology (Heidegger 1962) is that activities of interpretation pervade daily life. Our activities of interpretation are so commonplace, that we are almost unaware of it. One way we can conceptualize this activity is to think of it in a more common context; for example, when we are asked to “experience” an artistic piece. We typically search for meaning within art, and this search is subject to the particular interpretation of the human who is experiencing it. Phenomenology extends this line of reasoning into the banal and mundane. In fact, to exist is to interpret the world around us.

However, this interpretation is not a passive activity, since we need a system of knowledge (a pre-understanding) against which to interpret the world. This is the view of Gadamer (2004), who argues that this pre-understanding is inherently linguistic in nature. In interpreting a particular context, we (implicitly) call upon a history of interactions in a language, with which we co-create meaning. That is, there is no objective reality, nor subjective reality, but instead it is a combination of both, because one cannot exist without the other. An example of this linguistically-grounded activity of interpretation in a more common context is the process of translation. For natively bilingual speakers, the process by which an idea in a source language is translated to a target language mandates an explicit interpretation. There is no bijective relationship between concepts in any two natural languages, so an idea of a source language in a target language must be biased, since the interpreter must co-create (with the idea to be conveyed in the source language itself) a new meaning in the target language.

If we accept that to exist is to interpret, and that interpretation is inherently linguistic, then to exist is inherently linguistic, linked to communicating and understanding.
Communication is Grounded in Social Context  To phenomenologists, there is an inevitability of a hermeneutic circle: meaning is contextual, depending on the what is to be interpreted, and the pre-understanding a person uses to make sense of the moment. However, the pre-understanding is itself a product of the person’s history of interactions in language. The act of interpretation changes the individual’s use of language, and the use of language changes through its use by individuals. However, the individual’s language is imparted through social interaction, and as such, communication is grounded in social context.

Computers as Social Actors

In human-computer interaction, computers are construed as social actors (Reeves and Nass 1996), to which we attribute a host of social attitudes and behaviors. In essence, social interactions with computers emerge due to people heuristically scanning for patterns and cues that clue them into the proper behavioral patterns for the situation at hand. When a person finds a particular set of cues, the person commits to applying a social script appropriate for navigating the particular context. However, a person can detect those cues from a computer, leading to the incorrect application of a script that results in social behavior.

The reasons why we exhibit social responses to computers is still an area of active research. However, several characteristics have been identified as contributing to evoking a person’s categorization of a computer as a social being (Nass and Moon 2000): words for output, interactivity (responses based on multiple prior inputs), and the filling of roles traditionally filled by humans. Games arguably must be interactive, typically employ the use of words and other signifiers for output, and quite often orchestrate (or invite players to take the role of) human or human-like artificial agents, albeit in much more dramatic settings.

Since communication is the fundamental human activity grounded in a social context, when computers are treated as social entities, the context through which we engage with them is communicative. This idea has been explored by other researchers (Winograd and Flores 1986), and we leverage it here for benefit of broadening the study of interactive digital games.

Implications of Games as Conversation

If we consider the act of gameplay as a conversation, we may ask ourselves: How do the participants, both game player and game alike, communicate? and what does the communication imply? We look at these questions in the context of games as conversation by presenting many examples that serve to illustrate the richness of our framing.

Speech Acts

Games are action-oriented environments, where players must typically exert a non-trivial work effort to navigate through and complete a game experience (Aarseth 1997). It is therefore sensical to consider game-linguistic exchange as oriented toward the performance of some action within a game environment. The seminal works on communication as linguistic action by Austin (1955), and his student Searle (1969) are therefore relevant starting points for framing games as communicative exchanges.

As noted by Perrault and Allen (1980), (prior to the work by Austin) the analysis of an utterance was typically grounded in the utterance’s truth value, allowing forms of logical arguments like (e.g.) syllogisms:

1. All humans are mortal (True)
2. Hyapatia is a human (True)
3. Therefore (; ), Hyapatia is mortal (True)

However, not all utterances can be construed in this manner. For example, the utterance “I now pronounce you a married couple” effects a change in the world, and by itself does not have an associated truth value. Austin considered these sentences to be a type of performance, or act. Austin and Searle considered all utterances to be speech acts, distinguishing three kinds:

1. The locutionary act, or the act of putting words together into a form that is legal in the language. It is the act of saying something.
2. The illocutionary act, or the intended meaning that the speaker wishes to convey. It is the act in saying something.
3. The perlocutionary act, or the effected change in the listener’s mental state and/or future actions. It is the act achieved by saying something.

In the case of gameplay, both player and game actions are locutionary acts, where the language of the conversation (which defines what is legal) is built explicitly into the domain of interaction; the game itself explicitly affords exactly that which is permissible in the communicative exchange. Part of the challenge game designers face then, is that the game provides a sufficiently rich vocabulary with which players can express themselves. Put differently, normative game design must allow players to perform the precise speech acts that they want, or are motivated to do. If we take agency2 to be the defining characteristic of games, then our perspective on normative game design is related to the idea that designers should strive to achieve a balance between what is motivated by the game, and what is afforded to the player (Young and Cardona-Rivera 2011).

When players or games perform a locutionary act (i.e. interact via the environment), it is in service of an illocutionary goal (what they intend with their action). Searle (1976) proposed a taxonomy of illocutionary goals, and we examine game-related examples for each goal in turn:

- Assertives, which commit the speaker to the truth of an expressed proposition. Tutorials for complex games present information about the game space at the game’s onset (Andersen et al. 2012), often illustrating the ground truths for how the internal mechanics of the game’s design support player actions, which in turn affect the player’s expectations for future action. The information presented is an example assertive.

---

2as defined by Murray (1997): the satisfying power to take meaningful action and see the results of our decisions and choices.
• **Directives**, which commit the hearer to take a particular action. Game environments often highlight through various discourse methods (Steiner and Vorganti 2004) certain facts of the world that the player must manipulate, or obtain to further the development of the game experience (e.g. a key that must be obtained, which opens a lock). The highlighting itself is an example directive.

• **Commissives**, which commit the speaker to some future action. When faced with a quest, players must often explicitly agree to carrying out the quest to the best of their ability (Smith et al. 2011). This explicit agreement constitutes a commissive.

• **Expressives**, which express a speaker’s attitudes and emotions towards a specific proposition. When players are faced with choices in Mass Effect 2 (BioWare 2010), they are often provided with value-judgments on the morality of specific choice options on a binary scale: Paragon (virtuous, good actions) or Renegade (hostile, evil actions). These value judgments are expressives over the available options for action.

• **Declarations**, which effect a change in the reality surrounding both speaker and hearer. When players obtain items related to a status in the world (e.g. when a player is able to draw the Master Sword in the Legend of Zelda: Ocarina of Time (Nintendo EAD 1998), and becomes “The Hero of Time” in the process), they are participating in declarations.

This taxonomy helps elucidate the diverse illocutionary goals game players and designers (through their designed games) may have during the communicative exchange of gameplay. Of course, game designers do not encode discourse utterances to communicate for the sake of communication; they typically have an intended experience or trajectory that they hope to elicit in the player (Nelson et al. 2006). Thus, an additional challenge game designers face is the design of discourse utterances that achieve their perlocutionary goals, typically a change in the player’s mental and/or affective state regarding the progression of the game experience.

However the production of an utterance designed to achieve an illocutionary or perlocutionary goal does not necessarily entail its comprehension on behalf of the consumer of the utterance. How do we (as consumers) come to know essentially entail its comprehension on behalf of the consumer? For this, we turn to an overview of what is intended by the producer of utterances? For this, we:

• **Conversational implicature**, mean-

Gricean Maxims of Interactive Contexts

According to the philosopher of language Grice (1957), when people engage in dialog, they cooperate on the choices of what they say and how they say it in order to facilitate an effective exchange of meaning. The Cooperative Principle is summarized by Grice as a contract between the participants of dialog, where all participants observe unstated conversational rules. Per Grice (1975), a cooperative transaction can be said to exist when participants share some common aim, the contributions of participants are mutually dependent, and there is an understanding between participants that (ceteris paribus), the communicative exchange should continue unless both parties agree it should end. The game player and the game (embodying the game designer) arguably have the common aim of the player successfully completing the game itself, and their contributions to the continuation of the game are mutually dependent because players advance the game insofar the game permits their actions to take place, and the game cannot advance without the direction of players. The idea of the cooperative transaction breaks down for the third Gricean requirement (regarding the required consent of participants in order for communication to continue), since the designed game has no agency in deciding whether or not gameplay continues. However, this breakdown does not stop people from mindlessly (Nass and Moon 2000) ascribing intentions to designed artifacts (e.g. “The game does not want me to go there.”), and we contend that while it cannot in reality be a fully cooperative transaction, it appears and is treated as such for the player engaged in gameplay. As such, the communicative context created through gameplay exhibits its own type cooperative contract: a cooperative contract of interactive entertainment.

**The Cooperative Contract of Interactive Entertainment**

Young (2002) was the first to note that in interactive entertainment contexts, players co-operate with game environments to bring about the successful completions to game experiences. Young claimed that Grice’s Cooperative Contract binds players and games just as strongly as when we interact with each other in more conventional conversational settings. In this paper, we expand upon those initial ideas, and develop them into a general framework with which to analyze games.

The Cooperative Contract is an (often tacit) agreement between participants engaged in cooperative transactions, involving the observance of four maxims. These maxims are considered rules that people engaged in cooperative transactions assume, and they help explain implicatures, meanings suggested by an utterance that are neither necessarily expressed nor entailed. If the meaning of a communicative act is inconsistent with a Gricean Maxim, and the consumer assumes the producer is being cooperative, the consumer will perform an additional search for meaning (Gerrig and Bernardo 1994) in order to make sense of the act. We present and provide examples for each maxim in turn:

1. **The Maxim of Quantity**, where the contribution is neither more nor less than what is required. Task-based game quests must provide scaffolds (Thomas and Young 2009) to ensure that the difficulty to complete the quest is neither too little (causing players to get bored), nor too great (causing players to get frustrated) (Hunicke and Chapman 2004). These scaffolds must preserve the Maxim of Utility.

---

3This is obviously not true of all games. For instance, the game **Cat Mario** (Chiku 2007) is unfairly adversarial against the player. Within our framework, the unfairness would be considered an (intentional) failure by the designer to maintain the Cooperative Contract, leading to a rather frustrating communicative exchange.
The Narrative Paradox as a Discourse Problem

As we have already mentioned, gameplay is so constrained in terms of the communicative channels available between game, and player that game designers are pressured to make utterances relevant and meaningful. Similarly, players are pressured to attempt to make their contributions (their in-game actions) meaningful in the context of the designed experience. To paraphrase Adams (2013): the game designer promises to provide a credible, coherent experience if and only if the player promises to behave in credible, coherent ways.

When a player decides to engage with a game, she brings with her a set of tacit (not necessarily accurate) expectations of what she will experience. As she actually experiences the designed game artifact, she is expected to perform within the rules of the game to complete a trajectory that a designer has afforded. However, there is a tension with what is available and what the player would naturally choose to do. This tension has been identified as the “narrative paradox” (Aylett 2000), or the “boundary problem” (Magerko 2005) of interactive entertainment: how to recognize and respond to attempts by the player to perform actions that deviate from the intended experiential trajectory.

Attempts to curtail the player’s available in-game actions to a specific set that is acceptable for the sake of maintaining a coherent game experience may inadvertently reduce the player’s sense of agency, diminishing the game experience as a whole.

What the player would naturally choose to do is tricky to precisely identify. Player choice can be affected by exogenous determinants, including but not limited to, wanting to play to appeal socially (Roberts and Isbell 2014), personal preference over game content (Yu and Riedl 2014), and expected in-game utility (e.g. Monti, Grant, and Osherson 2005). Player choice can also be affected by endogenous determinants, such as in-game formal (i.e. motivational) affordances (Mateas 2001), or narrative affordances (Young and Cardona-Rivera 2011), sequences of content that a player can envision as completions to their current experience due in part to the role in which the player has been cast (or has willingly adopted).

Framing the boundary problem as a discourse problem, the challenge is to design an interactive experience where a game must determine (either a priori at design time, or dynamically at run time) what it has to “say” in order to elicit in the player the correct mental model of what she has to do next (to further the conversation, i.e. the game itself). In this way, the boundary problem becomes the discourse problem of scripting the interactor in the pursuit of a role, as described by Murray (1997) in the context of the adventure game Zork (Infocom 1980):

The lesson of Zork is that the first step in making an enticing narrative world is to script the interactor. ... By using literary and gaming conventions to constrain the players behaviors to a dramatically appropriate but limited set of commands, the designer could focus their inventive powers on making the virtual world as responsive as possible to every possible combination of these commands.
We contend that as part of the process of scripting the interactor, a game must appeal to the player through endogenous influence factors, such that what is uttered to the player is contextually relevant to the moment of interaction, and to the designed experience as a whole. Of course, scripting the interactor brings with it its own set of challenges. Broadly speaking, to script the interactor involves being able to precisely codify the appropriate illocutionary and perlocutionary goals the game should act upon. However, there are many potential ways in which a game could elicit a desired player mental state, and not all ways will have the same effect on the overall interactive experience. For example, if it is desirable for the player to move to a specific location, dynamic lighting may be effective at directing visual attention to the location of interest (El-Nasr et al. 2009), but may adversely affect the player’s sense of presence, or “sense of being in the world” (Steiner and Voruganti 2004).

The preceding discussion serves to illustrate the richness of perspective afforded by framing games as conversation, as well as the potential relevance of this framing to scholars interested in games. Our particular interest is in a computationally-oriented inquiry of games: the design of systems and algorithms to procedurally generate content that achieves a game designer’s communicative goals. In the next section, we illustrate how our framework enables computationally-oriented inquiry via a decades-old approach within artificial intelligence: the paradigm of planning.

**AI Planning-based Gameplay Discourse Generation**

In general, the field of discourse generation deals with the task of developing computer systems that can put together meaningful utterances in order to meet specific communicative goals. Reiter and Dale (2000) outline three stages of discourse generation that computational systems typically cycle through:

1. **Content determination**, or identifying out what information the system needs to communicate
2. **Structuring**, or figuring out how to structure the information in a set of discourse utterances, and finally
3. **Surface realization**, or translating the discourse structure that expresses the identified content into a form that is legal in the target medium (e.g. text, film, games)

**Planning-based Models of Natural Language Generation**

As previously discussed, utterances are construed as speech acts, and the aforementioned discourse generation process essentially proposes the following problem: given a set of communicative acts that are legal in a target medium, and a goal of what the system needs to communicate, the system must compute how to structure the utterances together to achieve the communicative goal. In this light, this is exactly the problem that the field of automated planning tries to solve, albeit at varying levels of complexity, and in a diverse array of task-environments.

**Automated planning (or simply planning)** is the computational study of an explicit deliberation process that chooses and organizes actions in service of some goal, by anticipating the outcomes of those actions and the conditions necessary for their execution (Ghallab, Nau, and Traverso 2004). Although there are different forms of planning, the basic conceptual structure of the planning problem contains 1) a formal (typically logic-based) description of the initial state of some domain, 2) a formal description of a desired goal state, and 3) a state-transition system, describing how the world can change as a result of applying planning operators in states. Planning operators have preconditions, conditions of the world that must be true prior to the operator’s execution, and effects, which describe conditions of the world made true through the operator’s execution.

A full review of planning-based approaches to discourse generation is beyond the scope of this paper (see Garoufi 2014 for a review), but the main idea behind most of these approaches is to treat human communication as a goal-oriented process, and then use planning to compute the speech operators needed to achieve communicative goals.

**Conceptualizing Gameplay Discourse Generation**

By framing games as conversation, and approaching discourse generation through a planning perspective, we have informally defined a **conversational-gameplay loop**, illustrated as a conceptual diagram in Figure 2. The loop operates given 1) a designer’s communicative goals that are specified at design time, 2) a knowledge-base of the game domain, which in our case is a planning domain model, and 3) a player model that characterizes the player’s state in some way useful to the planner.

Given the inputs, the planning system must identify what utterances it must produce to achieve the communicative goals, taking into account what the player does after the utterances have been produce. The planner can attempt to recognize the player’s mental state regarding their intended plan through the interactive experience by looking at the player’s actions, activities which are representable in terms of the knowledge-base. The planner can also attempt to further characterize the player’s mental state by looking at be-
havior beyond the knowledge-base that might be relevant, such as camera focus, or character gaze.

A Summary, and An Open Research Agenda
This paper has presented a rich framing for the study of digital games. Based on the idea that the fundamental human activity is communication, which is grounded in a social context, and which appears when we interact with computer-mediated experiences, this paper frames games as conversation. This framing allows us to look at the communicative context that emerges when the game and the player “converse” (i.e. when the player plays the game). Gameplay emerges in a communicatively constrained environment, effecting a dual pressure on both the game and the player: the game is pressured to communicate information that is relevant and meaningful for the player to advance, and the player is pressured to play in a credible, coherent way given what has been afforded by the game for action. This dual pressure causes game and player actions to be intertwined. We thus posit that gameplay is a cooperative communicative transaction, as defined by Grice (1975). The Gricean maxims provide a context for one definition of normative game design, which essentially supports the generation and interpretation of utterances that enable the game and player to operate harmoniously. We thus ask ourselves what exactly must be communicated, and to what effect, for which we leverage speech act theory as discussed by Austin (1955) and Searle (1969). Normative game design involves the appropriate specification of a game designer’s illocutionary and perlocutionary goals, with the ultimate purpose of effecting a change in the player’s mental or affective state that enables the player to understand how she fits into the progression of a trajectory that completes her game experience. Speech act theory is amenable to computational implementation via the artificial intelligence paradigm of planning. Discourse planners take as input the specification of a task environment, and compute a sequence of actions that (when followed) achieve a communicative goal of interest. We propose the use of discourse planning to dynamically (at run time) determine what the game has to “say” to elicit in players the correct mental configuration that would allow the player to successfully complete the game experience.

In contrast to other ways of eliciting the same mental configuration (e.g. through dynamic lighting), we suspect that producing contextual utterances will not adversely affect the player’s sense of agency, or sense of immersion. This hypothesis is the first part of the research agenda that our framework affords. Another key part of the research agenda is being able to precisely quantify the diverse exogenous and endogenous determinants of player choice, in order to better understand how specific utterances can support, or detract specific determinants relevant to allowing the player to envision successful completions to a game experience. A third key part of the agenda is the identification of which game locutionary acts (or combinations thereof) are effective at accomplishing illocutionary and perlocutionary goals. Certainly, the available locutionary acts (i.e. discourse utterances) will be dictated by the game type and its associated design conventions, much like the design of physical artifacts must respect conventions surrounding their use (Norman 2013). However, we suspect there exists a broad class of speech acts that elicit specific mental configurations in diverse contexts. For example, in story-driven games, communicating that there exists a trajectory in the player’s foreseeable future that the player can uniquely contribute to due to her role and her capacity for action, likely contributes to the perception of a narrative affordance (Young and Cardona-Rivera 2011) for that player.

By framing games as conversation, we can conceive of new ways to tackle problems central to the interactive entertainment community, and hopefully enable new solutions that leverage prior attempts at understanding and formalizing discourse processing.

Acknowledgements
This work was supported by the U.S. Department of Energy under grant number DE-FG02-97ER25308. Opinions, findings, and conclusions expressed here are solely those of the authors.

References


