Conversation as a Basis for Interactivity

Samuel Mann
Graham McGregor

Otago Polytechnic
Dunedin, New Zealand

smann@tekotago.ac.nz

Proceedings of the 15th Annual NACCQ, Hamilton New Zealand July, 2002 www.naccq.ac.nz

ABSTRACT

Interactivity is a widely used term. While well described in terms of mechanics, it is lacking in an underlying model. We propose conversation as a basis for considering interactivity. Key implications of this are the provision of information for inferential processing, non-sequential access to information and an approach based upon an unfolding dialogue between the user and the application. Natural language processing and artificial intelligence may not be available in the near future but this should not stop us examining further the components that make conversation work for us and attempting to apply them in our interactive applications.

1. INTRODUCTION

The purpose of this paper is to examine the potential for using conversation as the basis for considering interactivity. By using such a metaphor, it is hoped that direction can be given to those designing interactive experiences.

Despite the term 'interactive' being hard to avoid (Raskin 2000), the field of interactivity is not well defined. It is clear that animation is not interactivity, while dynamic, in that they are moving, they are not interactive: the user simply sits passively and watches them. He does not interact with them. Interactivity occurs when the user gets to do something, and when the users action affects the behaviour of the software. In its most basic form, then, interactivity implies user control and dynamic experience. Most attempts to define interactivity involve taxonomies and continuums - interactivity going from 'none' to 'maximum'.

Simple pragmatic taxonomies of interactivity focus on physical and mechanical things - what the user can do, how the program would behave in response. For example, many taxonomies of interaction identify five basic levels of interactivity: presentation; simple branching; complex branching; complex branching with feedback; and user generated content. Kristof and Satran (1995) presented a physical/mechanical taxonomy that focused more explicitly on the notion of user control - the more control, the more interactivity: pace; sequence; media; variables; transactions; objects; simulation (events).

It is not the case, however, that more interactivity is always better, indeed Murtaugh (1996) asks why anyone would want to exert control over a well-constructed story. Every project has an optimal degree of interactivity but taxonomies give little guidance to the degree or form of interactivity for a given task.

For this, we argue, we should apply a metaphor to model interaction that has better definition, and extends beyond the digital world.

The most pervasive attempt to provide a model for interactivity is to consider the narrative, indeed the two terms are often considered synonymous. The narrative refers to the story, the "crafted escalation of tension". The notion of interactive storytelling is compelling but may be a flawed vision (Aarseth 1997). Partridge (2001) argues interactivity and narrative "seem mutually exclusive rather than analogous" (p 9), Rouse (2001) examines the differences between the designer's story and the player's story, while game development conferences have entire streams struggling with the concepts of story (eg Bates 2001). Attempts to blend interactivity with storytelling often result in exponentially expensive branching for little gain and what Dove (1994, p281) calls the "fallacy of choice".

The narrative then, with its connotations of linearity is perhaps an inappropriate basis for interactivity, we suggest a change to conversation as an underlying metaphor for interactivity. Similarly, recent taxonomies of interactivity have moved beyond the physical activities the user can perform to consider the mental experiences the interactivity involves. Shedroff (1997) describes six different continua for interactivity. He claims that not just software, but all interactive experiences (e.g. talking to someone, going to the theatre, dressing a child) can be described by the six metrics in Table 1.

Shedroff then combines these six continua into three groups of similar metrics: Feedback and Control; Creativity, Productivity and Communication; and Adaptivity. These then define a three-dimensional space on which any interactive experience can be placed (Figure 1).

Shedroff places conversation at the peak of interactivity (second to Star-Trek's Holodeck- but we'll leave that for the next few decades). Conversation then, may be seen as a worthwhile metaphor for considering interactivity.

So what is conversation? Tannenbaum (1998 p286) describes three factors that moves talk to conversation: "degree to which participants in a communication have control over and can exchange roles in their mutual discourse". Two models are used in conjunction to describe conversation: a code model; and an inferential model.

The information processing paradigm holds that communication is seen as a potentially perfect encoding-decoding process involving the attempts of interlocutors to transmit and reconstruct messages. This is a rationalistic model and was developed at least in part from electrical engineering (Shannon and Weaver 1949). Language forms a conduit for ideas that are encoded and decoded (hence the 'code' model). Later researchers added detail to this model: feedback loops; fields of experience; role exchangeability; recognition that communication is a process (rather than a series of discrete events); fidelity, the effectiveness of the communication in achieving the purpose of the sender (all reviewed in Tannenbaum 1998). The language is not just spoken but includes kinesics (gesturing), haptics (touch), orientation and proximity. These factors together

Feedback	the user receives a response
Control	the user can control the behaviour of the interaction
Creativity	the user can contribute to or modify the content of the experience
Productivity	the user is able to produce some novel entity
Communication (the computer or another user)	the user can communicate with another entity
Adaption	the experience changes based on the behaviour of the user

Table 1: Six Dimensions of Shedroff's interactivity

provide a complex symbolisation system that must be agreed upon.

McGregor (1994 p24) argued that the code model provides an explanation of how communication is possible at all. It requires a sender 'speaker' and a receiver 'hearer'. Where it fails is in recognition that listening is an active process, that it is not a perfect process distorted by 'noise'. It also fails to recognise the "complexity of interactional processes, and the principles of contextual appropriateness, that are inherent in successful communication" (McGregor 1994 p25). Sperber and Wilson (1986) proposed an inferential model whereby communication is produced by interpreting evidence.

McGregor (1986) argued "everyday talk is a joint production and involves individuals trying to make communicational sense of others, in order to know to say next" (p154). "It takes the cooperative effort of the individuals concerned to arrange a conversation that makes sense". McGregor discusses how a conversation might iterate towards a meaning and how "any utterance can be understood in numerous different ways, and that people make decisions about how to interpret any given utterance or gesture based on their definition of what is happening at the time of interactions" (p158).

A consideration of interface in terms of conversation also needs to take into account the fact that most communicators are "sometimes sceptical, crafty and less than veracious" (Coupland *et al.* 1991 p2.), a conversation has strategic elements and encounters are based on wider ongoing relationships.

Conversation then, is context bound, often good enough rather than perfect and miscommunication is the norm (Coupland *et al.* 1991). Information is accessed non-sequentially as people infer links and

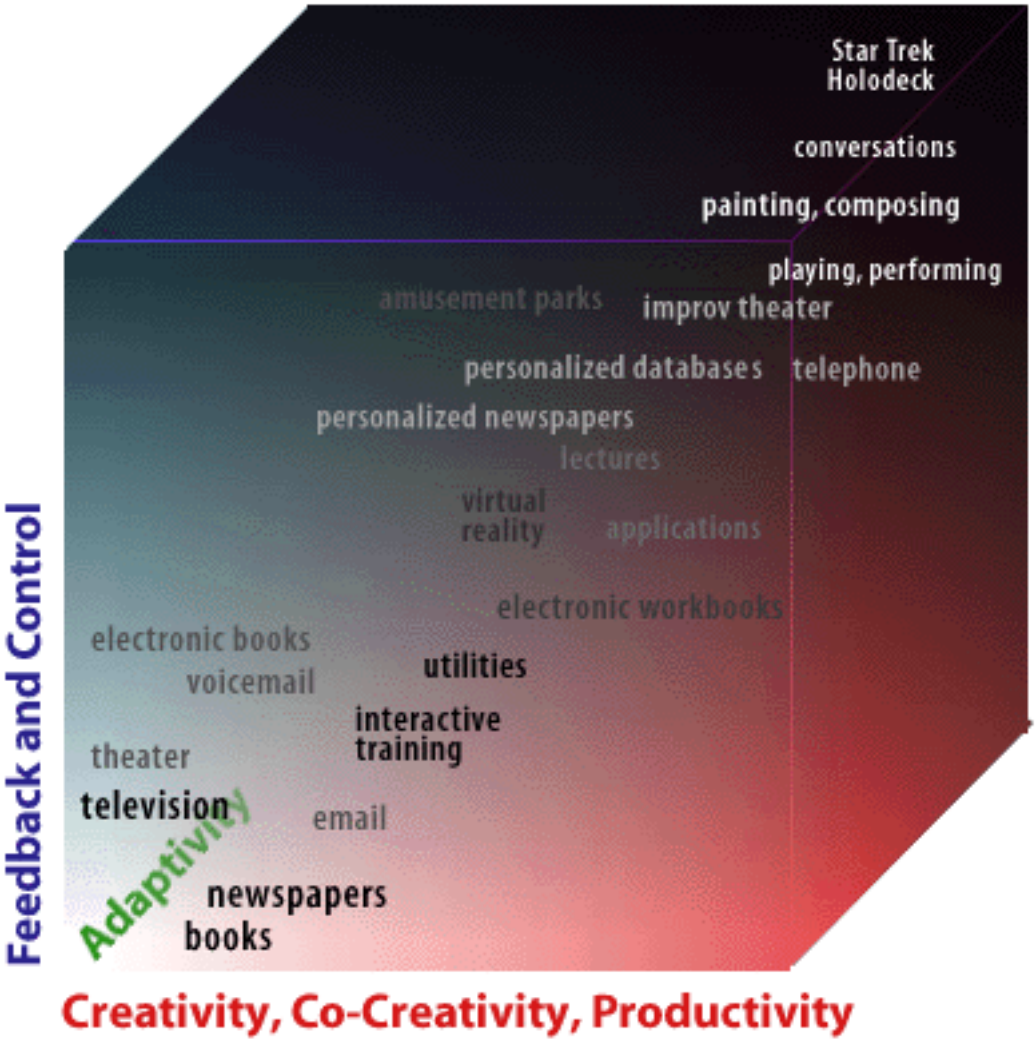

Figure 1: Shedroff's interactivity cube

hidden meanings and are inspired on different tangents by the conversation. Communicators may be fickle and deceitful and undertake strategies to help the conversation towards their goal. Conversations are usually immediate and heavily reliant on feedback and adaptability (Tannenbaum 1998 p293). Furthermore, the processes by which we infer information are immensely rich and poorly understood.

Why then should we try to apply conversation to help us understand interactivity? The answer is that conversation is immensely richer the dry and mechanical rules we usually apply to computer design. Taking even user centred guidelines such as Shneiderman's eight golden rules of dialog design (1998), it can be readily seen that the two are much separated. Consistency, yielding closure and maintaining locus of control are not aspects one would normally use to describe a good conversation. Tannenbaum (1998 p292) argues that even the most advanced multimedia "cannot perfectly replicate this kind of interactivity, or capture the subtle complexities of conversational management".

2. ASPECTS OF CONVERSATION APPLIED TO INTERACTIVITY

Whether we adopt the coding model or the inferential model, there are significant implications for conversational interactivity. Tannenbaum (1998 p292) contends "the user should be intimately involved in an unfolding dialogue between him or herself and the (multimedia) production".

2.1 CONTROL

Tannenbaum (1998) argues that the optimal interactivity occurs when there is a balance of control between the human and the computer. This differs from the usual approach where it is assumed that the human should always be in control (Nielson 1999). Giving the computer back some control is one of the advantages of the interactive TV approach. Instead of a web-based drilldown encyclopaedia approach that is heavily user control and pull focus, the conversational interactivity mixes push and pull interactivity in a way that emulates a conversation. In an advanced case, Porter (2001) describes her approach as "a strangely authored awareness of distance between player and avatar, characters acknowledge the player's authoritative role".

2.2 NATURAL LANGUAGE AND AI

Conversation has long been a goal of artificial intelligence (AI) researchers, indeed the Turing test for intelligence is rooted in conversation. Perhaps the most enduring AI approach is that of ELIZA (Weizenbaum 1976), a computer program that emulated a psychotherapist. ELIZA had almost no intelligence whatsoever, only tricks like string substitution and canned responses based on keywords. Yet when the original ELIZA first appeared in the 60's, some people actually mistook her for human.

Tannenbaum (1998 p260) argues that the human symbolisation system is so complex that computers cannot be made to completely understand the natural language. Shneiderman (1998 p166) describes natural language processing (NLP) as a "wonderful fantasy" but the subtleties, special cases, abstracted meanings and context mean it will probably remain a fantasy for some time. Shneiderman goes on to describe NLP and AI as "mind limiting distractions that inhibit designers from creating powerful tools" (p167). Conversation based interactivity is not necessarily based on simulating the actual spoken conversation.

Porter (2001) had much success with cartoon speech and thought bubbles to represent conversation amongst a group of characters. Using these conventions helps players see the distinction between what a user thinks and actually says, which gives the player the ability to choose to 'lie' in a conversation.

Probably more useful than NLP is an AI system that manages the path of the conversation. Tannenbaum (1998 p289) puts this as the fundamental question: how conversations occur at all without bumping into each other. Much work has been performed in the area of game AI, Rouse (2001), for example describes how the player affects story through responses of the AI characters and environment. This does not necessarily need advanced programming. Partridge (2001 p13) describes a narrative management system that "imbues the environment with simulated physical, spiritual and psychological properties and it lends intelligence to the simulated characters that inhabit those environments". This ability to 'create stories dynamically' is closely linked to the 'negotiated meaning of conversation'.

2.3 PROBLEM OF THE SOURCE

Tannenbaum (1998 p119) discusses the "problem of the source". How can one have an interactive conversation when one of the parties is inanimate? This is akin to Brody's (1996) example of the limitation of the technical interface disturbing the primary experience: a coffee cup that requires instructions. How can you talk to something that can't answer back? Despite this, there are many reports of users humanising their computer - talking to it, verbally arguing with it. It does not matter that the computer does not answer back verbally - it is providing evidence for inference through actions and the conversation has two parties.

Kress and van Leeuwen (1996) present a model of visual language whereby a single picture can operate on several different levels, permitting an almost interactive experience as the viewer examines the different levels. They include discussion of interaction; in their case they discuss the interaction between various participants in an image (including the viewer) approached as a conversation through the otherwise static image.

An important area in linguistic research is that of deterministic conversations. These are conversations that are, to some extent, and by one party at least, staged. Conversations between a pharmacist and customer, air-traffic control and legal discourse fall into these areas. The processes used in such discussions, while not leading to riveting conversations, may prove useful in developing the computer end of conversational interactivity.

2.4 STORYTELLING AS PART OF CONVERSATION

Smith (2002) argued that computer interactive fiction is at a dead end "teeth-grindingly linear". Telling stories however is a critical component of conversation. Bates' (2001) solution is to "create areas in which the player has freedom, and then to string these areas together in a linear series". This though is hardly conversational.

The significant change suggested by conversation is to move from seeing the "user as a problem" in terms of interactivity to the "user as a resource" (Smith 2002). This is a deistic approach (God set up the framework and left the humans to it): the crux of the application is the framework, the content is the area of the user.

Seely Brown (1999) gives an example of an application based on conversation. He describes tech-reps fixing photocopiers where most of their learning came from "just together weaving together a narrative". They created a distributed network to encourage these conversations then because fragments were lost to the ether created a web-based system and a process of transforming opinions and stories into warranted beliefs.

2.5 EFFICIENCY OF COMMUNICATION

The cooperative principle developed by Graice (1975 - in McGregor ,1994, p28) consists of four basic maxims: quantity, quality, relation and manner. These specify what people have to do in order to converse in a maximally efficient, rational and co-operative way; they should speak relevantly and clearly, while providing sufficient information. When we actually converse it is usual to deliberately flout at least one of these maxims, and doing so contains implied information ('implicature') that requires work by the listener.

In the coding model, extra material may be considered "noise", in the inferential model it is valuable information, but for both is an important part of conversation: people ask for clarification, provide redundancies and enjoy language systems that are imprecise.

In interface design there is some work on the information efficiency (Raskin 2000) but it is always assumed that we should strive for maximum efficiency. A conversational interaction might not be the most efficient, but it may be more effective.

2.6 MASS/INTERPERSONAL COMMUNICATION

Traditional consideration of conversations was face to face. This provides an interaction that is intimate, private, peculiar, plastic and malleable and highly interactive (Tannebaum 1998 p273). It is based upon two communicators intentionally orienting towards each other as both subject and object. This may be contrasted with both impersonal and mass communication. Impersonal communication increases the number of communicators, is limited in the channels available, knowledge of partner and usually has a formal relationship. Mass communication is heterogeneous, between strangers, spatially separated and has a low level of interaction.

These distinctions are “now blurred on all counts” (Tannenbaum 1998 p284) as we begin to deliver conversations via new technologies.

2.7 POINT OF VIEW

One of the conversational characteristics described by Tannebaum was that of the ability to exchange roles (to empathise). In terms of interactivity this can be seen as the point of view, both graphically and conceptually. Freidman (2001) discusses the difference between maps and tours: Maps are abstracted accounts of spatial relations (‘the girl’s room is next to the kitchen’), whereas tours are told from the point of view of the traveller/narrator (‘You turn right and come into the living room’). Maps document places; tours describe movements through spaces.

Tours, in other words, are the subjective, personalized experiences of the spaces described abstractly in maps. You start your journey with a map. Then, as you navigate the geography, that abstract knowledge becomes the embodied first-hand experience of a tour. This ability to change the presentation and perspective of the information, perhaps as the user infers meaning, would be a great step towards conversational interactivity.

2.8 FUTURE CONVERSATIONAL INTERFACES

We believe augmented reality and ubiquitous computing will require conversational interactivity. Feiner (2002) argues that within ten years “the user’s view of the world and the computer interface literally become one”. For computers to become ubiquitous, inextricably and transparently incorporated into our daily lives, new approaches to interface design will be needed. “Always on” and “always aware”, means we are going to have ongoing conversations with our computers (Mann, 1998), they, like humans, are going to have to know how manage these conversations or risk becoming very annoying. Electrophysiologically interactive computer systems enable affective computing (how users feel) and raise the possibility of the computer being able to communicate over multiple channels in conversation (Allanson, 2002).

3. CONCLUSION

While researchers persist with mechanically based taxonomies of interactivity, it lacks a coherent underlying model. Conversation is hugely complex but, perhaps because of this complexity, rather than despite it, provides a potential metaphor for interactivity. This paper has considered some of the aspects of conversation and explored how consideration of interactivity would benefit from the application of these aspects. The next step in formalising this model would be to assess the model by examining systems. The environmental accreditation system described by Mann and Brown (1998) may be considered to have a high degree of conversational interactivity; the unfolding dialogue was indeed designed to emulate a series of conversations between farmers and advisors. This model seems promising and is worthy of further development.

REFERENCES

Aarseth, E. (1997). Cybertext - perspectives on ergodic literature. London, John Hopkins

Allanson, J. (2002). “Electrophysiologically interactive computer systems.” Computer (March). 65-60

Bates, B. (2001). Story: Writing Skills for Game Developers. Game Developers Conference, San Jose http://www.gdconf.com/archives/proceedings/2001/game_papers.html

Braben, D. (2002). Another five years from now: future technologies. Game Developers Conference http://www.gdconf.com/archives/proceedings/2002/

Coupland, N., J. M. Wiemann, *et al.* (1991). Talk as ‘problem’ and communication as “miscommunication”: an integrative analysis. in Coupland, H. Giles and J. M. Wiemann.(eds) “Miscommunication” and problematic talk. London, Sage: 1-17.

Dove, T. (1994). “Theater without actors — immersion and response in installation.” Leonardo 27(4): 281.

Egenfeldt-Nielsen, S. and J. H. Smith (2002). Computer games, media and interactivity. http://www.game-research.com/art_games_media.asp

Falstein, N. (1999). The five stages of interactive writing http://www.theinspiracy.com/Ar5Stage.htm

Feiner, S. K. (2002). Augmented reality: a new way of seeing. Scientific American. http://

www.sciam.com/techbiz/0402feiner.html

Kaufman, H. and C. Kline (2001). Designing Immersive Public Entertainment Spaces. Game Developers Conference, San Jose http://www.gdconf.com/archives/proceedings/2001/game_papers.html

Kress, G. and T. van Leeuwen (1996). Reading images: the grammar of visual design. London, Routledge.288p

Kristof, R. and A. Satran (1995). Interactivity by design: creating and communicating with new media. Mountain View, CA, Adobe Press.131p

Mann, S. and Brown, I. (1998). Decision support for Environmental Accreditation, New Zealand Journal of Applied Computing and Information Technology <http://naccq.cit.ac.nz/jacit.htm>, 2(2):39-53

Mann, S. (1998). "Humanistic Intelligence: 'WearComp' as a new framework and application for intelligent signal processing." Proceedings of the IEEE 86(11): 2123-2151. http://wearcam.org/hi/index.html

McGregor, G., Ed. (1986). Language for Hearers, Permagon Books

McGregor, G. (1994). Third person interpretation and the sociolinguistics of verbal communication, Unpub PhD thesis, University of Newcastle upon Tyne.

Mortensen, C. D. (1994). Problematic communication: the construction of invisible walls. Westport CT., Praeger

Murtaugh, M. (1996). The Automatist Storytelling System: Putting the Editor's Knowledge in Software. MIT Media Lab, Masters Thesis,. http://ic.media.mit.edu/Publications/Thesis/murtaughMS/HTML/

Nielson, J. (1999). Designing Web Usability: Designing for Simplicity. Indianapolis, New Riders.419p

Partridge, A. (2001). Real-time interactive 3D games. SAMS, Indianopolis,.604

Porter, E. (2001). Cocktails and Conquests: an experimental narrative for the video game format. Center for New Media, Dept Information Design and Technology, Georgia Institute of Technology. Unpubl Masters Thesis, http://www.mindspring.com/~ekp23/thesis/thesis_desc.htm

Raskin, J. (2000). The humane interface: new directions for designing interactive systems. Reading, Mass, Addison-Wesley.233p

Rouse, R. (2001). Game design: Theory and Practice. Plano, Texas, Wordware.584p

Schell, J. and J. Shochet (2001). Designing Interactive Theme Park Rides Lessons Learned Creating Disney's Pirates of the Caribbean - Battle for the Buccaneer Gold. Game Developers Conference, San Josehttp://www.gdconf.com/archives/proceedings/2001/game_papers.html

Seely Brown, J. (1999). Learning, Working and Playing in the Digital Age. Conference on Higher Education, http://serendip.brynmawr.edu/sci_edu/seelybrown/seelybrownorig.htmlhttp://serendip.brynmawr.edu/sci_edu/seelybrown/seelybrownorig.html

Shedroff, N. (1994). "Information Interaction design: A unified field theory of design." . http://www.vivid..com/form/unified/unified.html

Shneiderman, B. (1998). Designing the User Interface: Strategies for effective human-computer interaction. Reading , Mass, Addison-Wesley.639

Smith, J. H. (2002). The dragon in the attic - on the limits of interactive fiction. http://www.game-research.com/art_dragon_in_the_attic.asp

Sperber, D. and D. Wilson (1986). Relevance: communication and cognition. Oxford, Blackwell

Tannenbaum, R. S. (1998). Theoretical Foundations of Multimedia. NY, Freeman

Weizenbaum, J. (1976). Computer power and human reason. San Francisco, CA: W.H. Freeman.