Getting to the point: Emotion as a necessary and sufficient element of story construction*

Clark Elliott
Institute for Applied Artificial Intelligence
DePaul University
243 South Wabash Avenue
Chicago, IL 60604
and
School of Education and Social Policy
Northwestern University
Evanston, IL
email: elliott@ils.nwu.edu
web: http://www.depaul.edu/ẽlliott

Ernst Melchior
Department of Computer Science and Information Systems
DePaul University
emelchio@hawk.depaul.edu

Abstract

In this short paper we discuss aspects of storytelling using emotionally sophisticated computer agents. In particular we focus on aspects of emotional interaction as being necessary and sufficient for creating stories. We present the ideas in the framework of an extant collection of multimedia computer programs called loosely, *The Affective Reasoner* [Elliott, 1992; Elliott and Ortony, 1992; Elliott and Siegle, 1993; Elliott, 1993; Elliott, 1994b]).

Motivation

Overview

All stories need a *point*. This will be just as true of interactive fiction as it is of static fiction. One common view is that the point of a story arises largely from failed expectations (c.f. [Schank, 1986; Schank, 1990]). It is possible, however, that a computational model of stories might be mislead by failed expectations which do not yield stories (e.g., the weather was hot for four straight days in October), and by stories which actually seem enhanced when our expectations *are* met (e.g., John Wayne heroically escapes and saves the day, again). We suggest an alternative view of necessary and sufficient conditions for "storiness" (our term). In this view the explicit presence of a human emotion, or the reasonable expectation of a human emotion based on a specific configuration of emotion antecedents within a situation, is necessary, and sufficient to yield "storiness." If this view is valid, then realizing a system that has semantic knowledge about emotion, has knowledge of those aspects of personality that allow interpretations of situations in idiosyncratic ways that lead to emotion, and has knowledge about the ways in which relationships between characters lead to differing emotions, may prove to be important step in developing interactive story systems.

Engagement with stories can come in two related, but distinct, ways. Both have to do with the contract between storyteller, and listener. First, we might tend to empathetically feel what the characters are feeling, and second, we might react to what the characters are doing with an emotional response. Without at least one of these it is possible to have interesting narratives (e.g., about physics), but we suggest that it is not possible to have effective stories.

Stories need a contract between teller and listener. These can vary depending on the story-telling situation. To wit: the story of the Cyclops is acceptable if the listener wants to hear about a monster or a hero, and unacceptable if the listener wants to hear why your term paper is late. The nature of this contract is very important in developing interactive systems. It requires maintaining a delicate balance between (1) presenting novel situations to the user, (2) allowing the user freedom to guide the system, (3) gracefully constraining the user's choices about narrative action so that the computer system can follow the story, and (4) maintaining believability of the automated characters.

Each of these three issues, "storiness" as a function of emotional content, engagement of the listener, and the contract between storyteller and listener are addressed in the system we are developing,

*Preparation of this article was supported in part by Andersen Consulting through Northwestern University's Institute for the Learning Sciences. Citation appears in bib.

and will be touched upon in this short paper.

Emotion as a basis for stories – three contrasting cases

Consider the following three situations. The first is what we may call the *Hero Story*. In this story, Walter Payton, the National Football League's all-time leading running back, rushes for over 100 yards in nine straight games. (Some running backs never manage a single 100 yard game in their careers.) In the tenth game he dramatically, and as expected, sets another record by going, again, over 100 yards. In the Chicago pubs after the game, the story is told, and re-told many times, with relish.

If we say, "The boy sat in his chair," there is no story. On the other hand, if we say, "The boy sat in his chair, but he did not want to," "The boy sat in his chair, even though Jane believes he should not do so," or, in particular, "The boy sat in his chair because he wanted to, but knew that he was wrong to do so," we have added an essential element to the narrative that at least allows us to argue for "storiness." (C.f., Sesame Street's send-up of day-time soap opera, "Love of Chair.")

In the Greek tragedy, Oedipus Rex, we are certain that Oedipus will discover that he has slept with his mother (that is, we *expect* that this will come about), but it does not seem to diminish the effect of watching the play, even if we experience it repeatedly. The same can often be said of humorous situations as well, where we, as omniscient observers, are just waiting for our expectation to be *fulfilled* when the character discovers he has his zipper undone, or has lentil beans stuck in his teeth as he addresses the national press club.

While we could argue that in each case an expectation has been violated (e.g., that Payton will fail, that Oedipus is not expecting the outcome), these are not *our* expectations, and it is for us that "storiness" is critical, and present. In addition, many stories can be enjoyed, like music, many times over, even though we know what is going to happen. In fact, if our expectations were not met, it would likely most often be a great disappointment (imagine missing the *precise* notes and orchestration of the recapitulation in a beloved Beethoven symphony).

We suggest that an alternative set of necessary and sufficient conditions can address "storiness," in some cases avoiding problems raised by the failure-of-expectation approach: In each of the above situations, the agent involved, with whom we are engaged, experiences, or we as observers, can reasonably be expected to experience, emotions. In looking at the above three examples, consider first that in the Walter Payton Hero Story, the telling of the story engenders *admiration* for Payton, or *pride* as Chicagoans represented by "Walter." There may additionally be *gloating* over a hated rival such as the Green Bay team.

In the second example, although the *Boy*'s emotions are not explicitly stated, our underlying theory of emotion suggests that in the first case we have no basis for emotion to arise, but in the latter three we do: In the first case, that the boy is sitting in his chair, does not suggest anything that might be of concern to the boy by embracing his desires, his principles, or his preferences – nor does it suggest anything that embraces ours. In the second case, that the boy is sitting in the chair but does not want to, reasonably suggests that he might, according to our theory, be *unhappy* about it. In the third case, that the boy is sitting in the chair, but should not be, suggests that he is, in Jane's eyes, blameworthy in doing so, and so might reasonably be expected to cause her to feel *reproach*. In the last case the boy desires to be doing something which he himself feels is blameworthy. This can reasonably be seen as leading to conflicting emotions, such as *remorse* and *gratification*, *satisfaction* and *guilt*, and so forth.

In the final example, in classic tragedy such as that of Oedipus, and in fact with classic comedy as well, it is often most effective if we anticipate, and expect, the strong emotions of the characters. The more clear our belief in the impending emotions, the stronger our engagement with the characters, and the more effective the manifestation of the respective theme. That we, as the audience, anticipate Oedipus's downfall engages us all the more in his extreme remorse and shame.

This latter situation embraces a construct we feel to be especially important to story themes: that of simultaneous, but contrasting – and especially conflicting – emotions caused by the same eliciting situation. These can take place between characters, and, even more strikingly, within a single character. In our emotion theory, and the social interaction it supports, consistency of goals between agents, and within a single agent, is allowed, but not implied. Many proposals for multi-agent systems appear to support only a cooperative model. In our view this is needlessly limited, and not suitable for the generation of stories. Since our representation does not require consistency in the ways individuals' appraisals give meaning to situations that unfold, conflicts between agents, and within the same agent, may arise. Particularly germane to the generation of deeper story themes is one particular group of conflicts that arise in our agents when situations are interpreted one way with respect to the goals (i.e., desires) of some agent, but in another way when interpreted with respect to their principles (i.e., appraisal of right and wrong). Consider, for example, the pain we feel for the "ill-made knight" Lancelot, when he struggles with his love for Guinevere, and his duty to Arthur; or for Hamlet in his sworn duty to avenge his ghostly father — but which duty is at odds with his love for his mother and his king (uncle). In the current project we use a rich representation for these conflicts since our theory explicitly differentiates between those emotions, and antecedents to emotion, which involve an agents principles, and those which involve an agent's goals (c.f. [Reeves, 1991]).

In the computer model we use, our agents respond emotionally to those situations that are of concern to them. They do this in idiosyncratic ways according to their dispositional, or mood-effected, ways of seeing their world. In this way the theme of the story, as it unfolds through the experience of the characters, is independent of what actually takes place externally, being dependent rather upon the agents' emotional response to the situations that unfold in their world.

In one exploratory application we created a series of simulation events representing the plot events in O. Henry's story, "The Gift of the Magi." In response to the same set of plot tokens, represented as the simulation events, we developed differing themes, and stories, by altering only the agents' goals, principles, and preferences. In this way were able to, for example, create a love story, as originally intended by O. Henry. In this scenario Jim and Della are happy and feel blessed that they care for each other so much that they inadvertently diminish each other's Christmas gift by sacrificing a precious possession out of love. Alternatively, and without changing either the dialog, or the external events that took place, we were able to create a drama wherein the one who "suffers most wins" representing a battle of wills, and each, unaltered, event is taken in almost the opposite way from the original interpretation. In this later scenario Della is furious that Jim has thwarted her martyrdom in selling her hair to buy him a watch chain, but then realizes she can have enhanced suffering because now her coveted combs come, tragically, too late for her to use them.

That our agents can respond to any situations within their realm of appraisal, including those that arise in real time, as antecedents to emotion unfold in an interactive environment, leads us to believe that we have a least one "hook" into creating interesting, real-time stories, that have a point.

Engagement of the listener

Stories, interactive or otherwise, are most effective when they engage the listener. We see this as taking place in two different ways. First the user may empathize with one or more of the characters in the story, and through identification with that character become engaged in the character's outlook within the context of the story (that is, "feel what the characters feel"). Second, the user may respond emotionally to the characters themselves, and the situations those characters find themselves in (e.g., with pity, or disdain). In neither case is it important that the characters be intellectually sophisticated, but it is essential that they be emotionally believable. For example, it is possible to identify with, and to respond emotionally to, animals (e.g., as in Jack London's *White Fang*), and simple-minded humans (e.g., as in Steinbeck's *Of Mice and Men*). On the other hand, the characters must be emotionally believable, and to some extent, sophisticated: even simple characters can feel real pain and triumph – and thus we can too, either through them, or on their behalf.

In designing applications of our automated "emotionally sophisticated" agents, both believability and engagement have been considered extremely important. Without believability the agents tend to be experienced as either simplistic, or confusing. If they do not engender engagement they are boring, because the user does not care about them.

With respect to believability we have relied on a rich, comprehensive, theory of emotion to create interesting emotional lives for our agents which are nonetheless internally consistent (see above, and [Elliott, 1994a]). It is our hope that using a theory-based approach to emotion will afford us a better chance of creating agents which are unpredictable, but still make sense.

We have further addressed believability in the way we use multimedia techniques. Through empirical exploration with the running system we have found that it is often true that, "less is more" when it comes to projecting believable automated entities. An important class of problems can arise when users' expectations about the sophistication levels of agents are raised in ways that were not intended, and cannot be fulfilled. When this happens, frustration occurs, the contract with the user is broken, and believability is sacrificed.

These problems can arise in a number of ways. We have found that in some cases sophisticated 3-D graphic facial images can actually be *less* effective than cartoon-like schematic faces in trying to express emotional responses to situations that have arisen. This seems to be because the greater resolution of the faces allows the user less freedom for adjusting the representation to meet the situation. A generically *angry* face can be adapted, in the user's imagination, to be appropriately angry for a specific situation (e.g., equally well for pure annoyance as for annoyance tinged with confusion). On the other hand, a 3-D face representation, which has much greater resolution of detail, must bear the burden of much great faithfulness to detail, which, unless it is explicitly tuned for each particular situation, will fail to meet a number of expectations. By contrast a schematic face does not actually have to look at another character to be seen as communicating with it, and will never fail to meet expectations about, say, the small muscles around the corner of the eyes – simply because it does not explictly represent them. Thus the computational burden of using 3-D is actually much greater for this type of application than merely supporting the enhanced graphics: the representation itself must also necessarily contain a great deal more semantic information about the scenarios the face-display is to be presented in. If this information is not correct, and correctly expressed, believability suffers.

We have also been careful in the use of our acoustic speech interface as well. We currently use a continuous speech recognition package (Verbex *Listen for Windows*) that is quite decent at detecting a limited subset of normal conversational speech. In addition we have great deal of generative capability for our text-to-speech system, since the text strings are created dynamically in LISP. However, we have found here as well that making full use of the technological power can, in some cases, actually lead to more problems than it solves. Since we have the ability to parse (small portions of), and spit out, relatively complex utterances, it is easy to mislead the user into believing that our agents actually understand the utterances (while in fact they only understand small portions of the relatively explicit emotional content). In addition, it seems to be true that aural interfaces are particularly prone to raising unfounded expectations regarding the natural language understanding abilities of computerized agents. Again, once the user forms mistaken beliefs, even temporarily, the fragile contract with the user is easily broken, and believability suffers. But we *do* want to highlight what can be, at times, a sophisticated emotional awareness on the part of agents. On the one hand, when the computerized agent says something like, "Gee, I am really really sorry for you," this can tend to raise conversational expectations, whereas if the agent says, "I am just a stupid computer program, but in my own simple way I consider you to be my friend, and I am sorry that this bad thing happened to you," it is less likely to do so.

To some extent the "less is more" issue is mirrored in the inflectional capabilities of the text-to-speech packages as well. In contrasting *Monolog for Windows*, a relatively rudimentary package, with *Text Assist for Windows* (DecTalk), a much more sophisticated package, we find that even though the latter allows us much more control over the voices, and the representation of characters, we can not always as believably represent different emotions. In using the more sophisticated package, because of the more distinct features of the different voices the simple techniques we use to represent emotional speech (i.e., global changes to speed, pitch, and volume) can be significantly less effective in communicating affect: a "little boy" character speaking slowly does not sound sad, he just sounds like he is speaking slowly; a "middle-aged woman" character speaking slowly and in a high-pictched voice does not sound afraid, she just sounds like she is speaking slowly, and in a high voice.

With respect to engagement, multimedia computers have a powerful tool in intelligently applying music to communicate affect (c.f., [Brown, 1981; Gabrielsson, 1989; Pignatiello *et al.*, 1986; Kenealy, 1988]). Music can engage users in two ways. First, it can be used to give additional cues to the user about the emotional states of the characters. "Upbeat" music can, for example, help the user to interpret a character's actions as those of one who is happy, or relieved, whereas "serious" music can help the user to interpret other expressions (e.g., facial cues, voice inflection), as representing anger or remorse. More importantly, we see music as a powerful arousal, and mood-induction, tool which can greatly increase a user's visceral sense of identification with, or emotional response to, a story character. Though it is difficult to capture in text, we find that the extensive MIDI capability of the multimedia PC, when combined with high resolution sound generation through professional synthesizers, allows us to make this a feature of the real-time system that can be, at times, striking.

In our current project we index a large MIDI music database according to the general themes present in the emotion theory.[1] Current technology allows us to retrieve hundreds of hours of music so that it can be indexed down to the millisecond, and loaded and played with subsecond response time. We use *Proteus* sythesizers so that playback quality is relatively high. In storytelling, we use the music to help represent individual emotions (in short fragments), and long, underlying themes (in long fragments).

That music is so emotionally arousing, but simultaneously non-specific in what it represents, works largely to our advantage in that we have a great deal of freedom operate without violating believability.

[1] This database is currently being re-recorded using exclusively live (but re-mixed) performances.

References

[Brown, 1981] Roger W. Brown. Music and language. In *Documentary Report of the Ann Arbor Symposium.* Music Educators National Conference, 1981.

[Elliott and Melchior, 1995] Clark Elliott and Ernst Melchior. Getting to the point: Emotion as a necessary and sufficient element of story construction. In *AAAI Technical Report for the Spring Symposium on Interactive Story Systems*, Stanford University, March 1995. AAAI, American Association for Artificial Intelligence. THIS PAPER.

[Elliott and Ortony, 1992] Clark Elliott and Andrew Ortony. Point of view: Reasoning about the concerns of others. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pages 809–814, Bloomington, IN, August 1992. Cognitive Science Society.

[Elliott and Siegle, 1993] Clark Elliott and Greg Siegle. Variables influencing the intensity of simulated affective states. In *AAAI technical report for the Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, pages 58–67. American Association for Artificial Intelligence, 1993. Stanford University, March 23-25, Palo Alto, CA.

[Elliott, 1992] Clark Elliott. *The Affective Reasoner: A Process Model of Emotions in a Multiagent System.* PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.

[Elliott, 1993] Clark Elliott. Using the affective reasoner to support social simulations. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 194–200, Chambery, France, August 1993. Morgan Kaufmann.

[Elliott, 1994a] Clark Elliott. Components of two-way emotion communciation between humans and computers using a broad, rudimentary, model of affect and personality. *COGNITIVE STUDIES: Bulletin of the Japanese Cognitive Science Society*, 1(2):16–30, 1994. In Japanese with English abstract.

[Elliott, 1994b] Clark Elliott. Research problems in the use of a shallow artificial intelligence model of personality and emotion. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 9–15, Seattle, WA, August 1994. AAAI, American Association for Artificial Intelligence.

[Gabrielsson, 1989] Alf Gabrielsson. Intense emotional experiences in music. In *Proceedings of the First International Conference on Music Perception and Cognition*, pages 371–376, Kyoto, Japan, October 1989.

[Kenealy, 1988] Pamela Kenealy. Validation of music mood induction procedure: Some preliminary findings. *Cognition & Emotion*, 2(1):41–48, 1988.

[Pignatiello *et al.*, 1986] M. F. Pignatiello, C. J. Camp, and L. A. Rasar. Musical mood induction: An alternative to the velten technique. *Journal of Abnormal Psychology*, 95:295–297, 1986.

[Reeves, 1991] John F. Reeves. Computational morality: A process model of belief conflict and resolution for story understanding. Technical Report UCLA-AI-91-05, UCLA Artificial Intelligence Laboratory, 1991.

[Schank, 1986] R.C. Schank. *Explanation Patterns: Understanding Mechanically and Creatively*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.

[Schank, 1990] Roger C. Schank. *Tell Me a Story*. Charles Scribner's Sons, New York, 1990.