

# A hybrid model for novel story generation using the Affective Reasoner and ChatGPT<sup>1</sup>

Clark Elliott

Jarvis College of Computing and Digital Media  
DePaul University, 1 DePaul Center, Chicago, IL 60604  
[elliott@depaul.edu](mailto:elliott@depaul.edu)

**Abstract:** In this paper a hybrid model is presented for generating novel stories using (a) a traditional symbolic AI cognitive-appraisal model of emotions embodied in the *Affective Reasoner (AR)*, and (b) the large-language-model-based (LLM) system embodied in ChatGPT. The novel emotion and narrative structure is generated first by AR techniques—giving strong, symbolic *computable* structure to the intermediate narratives—and then fed in series to ChatGPT to add complementary world knowledge and elegant language structure. The resulting stories are polished and cohesive, but the basic structural elements remain under computational control. Explanations about content can be generated, based on the emotion content, and also on the appraisal-based dispositions, expressive temperaments, reasoning about the fortunes of others, relationships and moods of the characters in the stories. Background emotion theory is reviewed, relevant to the morphing of narratives, composed of 28 emotion categories, 24 emotion intensity variables, and ~400 channels for emotion expression, which has been implemented in the AR. A series of hybrid-generated stories are presented illustrating how the emotion makeup of characters, their emotions, their actions and their narrative perspectives remain not only consistent but are largely enhanced after treatment by ChatGPT. Actual examples of generated stories covering a wide range of complex emotion scenarios are given.

**Keywords:** ChatGPT, LLM, Affective computing, AI, intelligent agents, emotion, stories, narrative, gaming AI.

## 1 Introduction and Motivation

This is a theoretical proof-of-concept paper, based on components of a working AI story-generation system, and the online version of OpenAI's ChatGPT (GPT 3.5) large-language model (LLM) AI chat system, arguing that we can create

---

<sup>1</sup> Projected citation: Elliott, Clark. "A hybrid model for novel story generation using the Affective Reasoner and ChatGPT." In *Intelligent Systems and Applications: Proceedings of the 7-8 September 2023, Intelligent Systems Conference (IntelliSys)* Amsterdam, The Netherlands. Volume N, pp. N-N. Cham: Springer International Publishing, 2023.

novel, world-knowledge-situated stories under computational control suitable for many contexts based on sound AI emotion reasoning principles.

In previous work [1, 2, 3] it has been argued that using a highly-computable model of emotion allows us to extract an essential structure in stories which is independent of the particular narrative context, and that the same emotion structure for one story, in one context, can also work just as well in many other contexts. It is this transference of structure, independent of the narrative domain which gives story-morphing its computational power; narrative computational techniques in one domain work just as well in other domains. More to the point, humans have an affinity for these narrative emotion structures as part of their understanding of social interactions [4].

The strong position motivating this work is that most other approaches to story generation, theme analysis and story understanding require that we have in place a model of *how the world works* [5, 6, 7]. Such models are required for the automatic symbolic manipulation of plot structures with any sort of complexity. However, we have no such causal model of how the world works, and the complexity of such representation models are prohibitive.

Instead, within the AR's story-morphing, we abandon any attempt to reason logically about causality in the real world as being hopelessly complex beyond very constrained subsets of not-very-interesting artificial worlds. Instead, story-morphing starts with the premise that the simplest story is generated when situations arise as part of an otherwise narratively uninteresting sequence of events, *and yet one of the characters is believed to care about what happened*. The underlying ways in which characters respond emotionally to such situations creates narrative fabrics that are rich, complex and *narratively consistent*. Most of the salient features of plots, and the universal themes arising from them, are based on such explicit and presumed emotional responses to the narrative situations. By changing the appraisals of the characters, and the ways in which they express their emotions (only), this still yields potentially millions of subtle variations of the original story that are based on generally identifiable emotion patterns. [2].

In the hybrid model discussed here we start with simple, AR-generated stories (with fully computational emotion structures). Then we add elegant embellishments using ChatGPT to flesh out the basic stories with world knowledge, without affecting either the plots, the themes, or the emotion contents of the original AR-generated stories.

There are several aspects of the hybrid model that make it both interesting, and workable. First, the model makes use of the strong points of each AI technique: From the AR we have the strong computational emotion structure and strong control over personality generation. From ChatGPT we have elegant presentation of a consistent world model in unlimited domains: artifacts on the surface of the moon will be different from artifacts in the cabin of a White Freightliner. Second, we work around weakness of each AI technique: The AR has almost no ability to reason about world knowledge unless it has been explicitly added. ChatGPT does not have inherent computational structure in its output. Third, there are computational benefits to giving ChatGPT clear

constraints under which to generate text. Otherwise, it may generate smoothly flowing stories having little to do with our computational intentions. Using the emotion features generated originally by the AR, ChatGPT is constrained to generating finished story components that lie within our narrative goals.

Lastly, we can ask, why would we care about such a system? We might argue that ChatGPT can create interesting narratives on its own. But to create intelligence that faithfully models that of humans and to best interact with humans, we ultimately have to move into the world of symbol processing. The core of human reflective experience and the language of human thought is, at the computational level, based on reasoning with symbols, with relationships such as greater than, taller than, near-far, part-whole, horizon-centered, balanced, appealing-to, self-other and etc., all creating the fabric of the human experience [8] [9] [10] [11] [12]. In addition, humans traffic in metaphorical thinking [13]. When we manipulate the complex *structural* elements of emotions and personality within a social-interactive environment, humans comprehend what we are doing on many levels. We are much closer to actually *interacting* with them, and we are speaking their language. Without such a model, there is no *there there*<sup>2</sup> despite the smoothness of the presentation.

## 2 Background Theory

The cognitive emotion model, used, as illustrated in Fig. 1 includes *emotion categories* with multiple intensities and qualities within each category. (For example, the category *joy* includes *pleased, happy, delighted, overjoyed, ecstatic* and so on.) At least three different intensities are computed for many of the emotions. The model also includes hundreds of *channels* for expressing emotions—roughly twenty channels tweaked for each emotion. For example, a *self-evaluative* channel for the expression of *reproach* would include an *awareness of one's beliefs about right and wrong* while a *verbal other-directed emotion modulation* expression of *gloating* might include *calling attention to another's blocked goal* to encourage the object of one's derision to feel humiliated.

The essential structure of the appraisal mechanisms composing what is referred to as the *disposition* of agents is contained in Figure 1 giving the description of twenty-eight emotions, based, originally, on the seminal work of Ortony, Clore and Collins [14]. The dispositional way that agents appraise the unfolding plots—discussed here in the examples—are used as the basis for this paper.

AR agents are designed and implemented with two components: the aforementioned *disposition* which controls the way they interpret situations that unfold in a story—the component primarily under consideration in this paper—

---

<sup>2</sup> Gertrude Stein speaking about Oakland in *Everybody's autobiography*, Vintage, 2013.

a *temperament* which controls how they express any emotions that may arise [and see [15]], and *moods*.

## EMOTION CATEGORIES

Clark Elliott, 2021 after Ortony, et al., 1988		
GROUP	SPECIFICATION	CATEGORY LABEL AND EMOTION TYPE
Well-Being	appraisal of a situation as an <i>event</i> relative to one's <i>goals</i>	<b>joy</b> : pleased about an <i>event</i> ; <i>achieved or retained goal</i> <b>distress</b> : displeased about an <i>event</i> ; <i>blocked goal</i>
Fortunes-of-Others (Well-Being)	appraisal based on presumption of how a situation is appraised by another as an <i>event</i> relative to <i>their goals</i>	<b>happy-for</b> : pleased about an <i>event</i> desirable for another <b>gloating</b> : pleased about an <i>event</i> undesirable for another <b>resentment</b> : displeased about an <i>event</i> desirable for another <b>jealousy*</b> : <i>resentment</i> over a desired mutually exclusive goal <b>envy*</b> : <i>resentment</i> over a desired non-exclusive goal <b>sorry-for</b> : displeased about an <i>event</i> undesirable for another
Prospect-based (Well-being)	appraisal of a situation as a prospective <i>event</i> relative to one's <i>goals</i> .	<b>hope</b> : pleased about a prospective desirable <i>event</i> <b>fear</b> : displeased about a prospective undesirable <i>event</i>
Confirmation (Well-being)	appraisal of a situation as an <i>event</i> confirming or disconfirming an expectation relative to one's goals	<b>satisfaction</b> : pleased about a confirmed desirable <i>event</i> <b>relief</b> : pleased about a disconfirmed undesirable <i>event</i> <b>fears-confirmed</b> : displeased about a confirmed undesirable <i>event</i> <b>disappointment</b> : displeased about a disconfirmed desirable <i>event</i>
Attribution	appraisal of a situation as an accountable <i>act</i> of some agent, relative to one's <i>principles</i>	<b>pride</b> : approving of one's own <i>act</i> <b>admiration</b> : approving of another's <i>act</i> <b>shame</b> : disapproving of one's own <i>act</i> <b>reproach</b> : disapproving of another's <i>act</i>
Attraction	appraisal of a situation as containing an [un]attractive <i>object</i> re. one's <i>tastes</i>	<b>liking</b> : finding an <i>object</i> appealing; relative to <i>desires/tastes</i> <b>disliking</b> : finding an <i>object</i> unappealing; <i>~desires/tastes</i>
Well-being/ Attribution	compound emotions	<b>gratitude</b> : admiration+joy <b>anger</b> : reproach+distress <b>gratification</b> : pride+joy <b>remorse</b> : shame+distress
Attraction/ Attribution	compound emotion extensions	<b>love</b> :admiration+liking <b>hate</b> :reproach+disliking <b>[frustration*]</b> : hope+disappointment] <b>[unfairness*]</b> : resentment+reproach--stdrd unfairly applied, self/other]

\*Non-symmetric additions necessary for some stories.

**Fig. 1.** Ortony, et al., [14], modified Elliott 2015 & 2021: The structure of appraisal within the content theory of emotions used as the basis for the *dispositional* component of the Affective Reasoner's emotionally intelligent agents.

Previous work showed the feasibility of using Emotionally Intelligent (computer) Agents to take the place of characters in stories [2, 1]. For each such agent it can be asked, *how does this agent feel about the events that are unfolding?* and *how might this agent express those feelings?*

### 3 The basics of Story Morphing in the Affective Reasoner

Abbreviations used: **AR** = The Affective Reasoner. **GSPs** = Goals, Standards / Principles, Preferences

A full exposition of the AR's story-morphing techniques are given in [16]. A brief version follows here:

The first step is to take a base narrative and translate it into a set of discrete, time-sequenced plot steps which will be used to trigger appraisals by the characters in the story.

Next, GSPs are manually created to represent many different ways in which any AR agent—in modeling human appraisal behavior—might reasonably be expected to interpret the events unfolding in the plot. GSPs must provide potential symbolic-unification pattern matches against the events unfolding within the plot steps.

Next, *personality prototypes* are created [either manually, or under pure program control] by combining sets of GSPs representing the appraisal potentials for each AR agent that will portray a character in the story, along with expressive patterns that define the agent's temperament. Additionally, as part of personality definition, *relationships* between AR agents can be defined, and are essential in determining the emotions the agent might have based on the fortunes of someone else. For example, if a user is considered a friend, and something bad happens to the user, the agent may feel sorry for them that this occurred.

Lastly, *story-morphs* are created using the unfolding discrete steps of the plot, and selected personality prototypes bound to specific characters in the story. Once again, this can be done under program control.

Creating novel, appropriate, interesting appraisals of the plot steps is always done manually. But because each appraisal stands on its own, this is not a computationally difficult process. Yet from even a few dozen such appraisals, millions of story-morphs can be automatically generated.

For the purposes of demonstrating the hybrid model—which generates quite a bit of narrative text—in this paper, very simple stories are used containing only a few plot steps.

### 4 The hybrid model of story generation, and its significance

Structured output, of the type generated by the AR, in the form of plot steps, appraisals and emotions, plus—in some instances—temperament and mood indications, along with basic instructions on the type and perspective of the story to

generate, were used as input to ChatGPT.<sup>3</sup> Unaltered ChatGPT English prose versions of the stories—in some cases from the perspective of different characters—are presented here, unaltered, composing the examples. The first story generated by ChatGPT was always used. There was no cherry-picking of output. For these experimental runs copy and paste was used to insert the prompts for ChatGPT (GPT 3.5) and for extracting the (unaltered) output.

The claim made, for this hybrid system, is that graceful, context-correct, stories are produced that nonetheless remain largely, if not fully, under program control with respect to the all-important emotion structure, and personalities created for the characters. Not only is the ChatGPT content newly generated, but so is the emotion structure on which the story was based. Within the same basic events unfolding we can, under program control, generate stories of love, of jealousy, of admiration, of fearfulness and, etc. In this way we can answer such questions for these novel stories as "How was the story created?," "What other stories is this similar to?" and we can give *real explanations* of the ways in which the content of the story has meaning based on the emotion and personality structures. That is, we can *reason* about the content of the newly generated stories, under program control.

GTPChat itself does not have the all-important program control—structural understanding of—the social-emotional *meaning* of the stories it generates, but it has huge reserves to draw on regarding how the world works at a highly consistent surface level such that it can generate smooth-sounding narratives in virtually unlimited contexts.

By contrast, the AR has broad, computational knowledge of emotion structure in the human world suitable in many applications. But it very explicitly has *no world knowledge*.

To illustrate, let us look at an example:

*Samir dropped his favorite glass vase on the floor of his house and broke it.*

ChatGPT can converse very fluidly about *dropping, favorites, breaking, vases, floors* and *houses*—even though it has no understanding of the meaning of any statements it might later generate, like "Samir later cleaned up the glass," or "Samir was unhappy about the loss of his vase." By direct contrast, the AR can "understand" and manipulate the emotion structure of the episode through, e.g., making Samir "actually" unhappy because he very much *likes* his vase. Or can put Samir into a temporary depression over it. Or can make Samir *satisfied* because he was *hoping* he was going to drop the vase *he hated*. Or can make *Samir's adversary* Charles *gloat* over Samir's misfortune. But the AR does not know anything about dropping, or vases, or floors, or houses—only how Samir feels about them. Together we can leverage what each system does well to make up for deficiencies in the other:

---

<sup>3</sup> Minor grammatical (only) liberties were taken with AR-style output to make this paper more legible and to give standard input to ChatGPT.

ChatGPT lacks symbolic computational knowledge and "understanding" in the traditional sense. The AR lacks world knowledge and language knowledge. Together they are fluid at emotion and personality reasoning, language manipulation and surface-level—but consistent—world knowledge.

The key in making this marriage work is in constraining the events ChatGPT is allowed to generate in stories, and in keeping the AR far away from needing any sort of world knowledge.

## 5 Different appraisals for characters yields different stories

In this section we present (a) the prompt, generated in the style of AR output as basic plot steps plus the goals, standards and preferences (GSPs) of AR agents (the characters) within the domain of the story, followed by (b) the unedited output generated by ChatGPT, based on the prompt, and (c) in some cases, discussion of what each example significantly illustrates.

First, a very simple story is generated by the AR. In this case the only emotions generated are the *prospect-based emotion fear*, and the *confirmation-based emotion relief*. Instructions to ChatGPT about the style of story to generate are prepended to the story, and the whole prompt is given to ChatGPT.

**The AR Prompt:** "Tell me this story from Mona's perspective, in Mona's voice: Mona and Bruce were working on a paper together. Mona got her work done on time, but Bruce waited until the last minute. Mona was afraid they would miss the submission deadline. Later she was relieved that they submitted on time. Bruce wanted mona to relax and take it easy."

From that prompt, ChatGPT produces the following response story (ChatGPT output is verbatim, throughout the paper):

**ChatGPT output:** Mona and Bruce were working on a paper together, and Mona had always been a responsible student. Mona finished her work on time, but she was worried when she saw Bruce waiting until the last minute. Mona believed that it was important to work in a timely manner, and she was afraid that they would miss the submission deadline. Mona's fear was heightened because the paper was very important to her. Despite Bruce's advice to relax and take it easy, Mona couldn't shake off her worry. But, eventually, they got the paper submitted on time and Mona felt relieved.

### GPT content based on AR structure (under computational control):

1. Working on the paper together (plot step)
2. Mona finished her work on time (plot step)

3. Mona was worried (AR emotion content, fearing the blocking of a future goal).
4. Mona was afraid they would miss the submission deadline (AR emotion content, fearing the blocking of a future goal, restated with greater intensity.)
5. Mona couldn't shake off her worry (AR emotion content, consistent with persistent fear)
6. The paper was submitted on time (plot step)
7. Mona felt relieved (AR emotion content—relief over the positive future resolution of the thing feared.)

**Extraneous GPT embellishments (not under computational control):**

- A. Mona had always been a responsible student (consistent with emotion content and plot steps)
- B. Mona's fear was only at the intensity level of worry, not very fearful (consistent with emotion content)
- C. Mona was worried [*because*] she saw Bruce waiting until the last minute (consistent with emotion content and plot steps)
- D. Mona believe it was important to work in a timely manner. (consistent with emotion content--a standard or principle that Mona had)
- E. Fear was heightened because the paper was very important to her (consistent with emotion content, but potentially a conflict because no emotion intensity [17] was listed by the AR). Correctly links "the paper was important" to "the submission deadline" which is consistent with the plot steps.
- F. Bruce gave Mona the advice to relax. (Consistent with plot steps and Bruce's standards). Potential conflict by adding extra plot step of the giving of advice, which is not under computational control.

In this example we see that all the emotions under control of the AR are present, and all the plot steps are present. They have not been misinterpreted or altered in their character. In addition, we also have a number of real-world components of the story added, which all make sense in the context of the story. Each of these additions is consistent with that which is under AR program control, except for the following minor elements:

1. Mona's level of initial worry (emotion intensity element not designated by the AR—though it could have been).
2. How important the paper submission was to Mona. (*Importance*: Emotion intensity variable not generated by the AR—though once again, it could have been.)
3. The new plot steps that Mona *saw* Bruce waiting until the last minute, and Bruce *giving advice* to Mona.

**Commentary:**



The new "world details" added by ChatGPT are all consistent with the story. The potential conflicts from additional plot elements and tweaking of the emotion content are unlikely to conflict with emotion reasoning. So, we have a story, generated by ChatGPT where we have computational control of all of the important elements. The AR can explain all of the original elements, still in the story, but would have no commentary about Bruce telling Mona to relax, or Mona observing Bruce waiting until the last minute.

## 6 Using ChatGPT to change narrative perspectives:

### **Mona's perspective story one:**

**AR prompt:** "Tell me this story from Mona's perspective, in Mona's voice, starting out "I was working with Bruce..." Mona and Bruce were working on a paper together. Mona got her work done on time, but Bruce waited until the last minute. Mona was afraid they would miss the submission deadline, putting her academic career at risk. Mona admired Bruce for his relaxed atmosphere, but at the same time was furious with him for putting her academic career at risk. Also, Mona loved Bruce, but Bruce was not interested in Mona that way. So, Mona had very mixed feelings. Later she was relieved that they submitted on time. Bruce wanted Mona to relax and take it easy."

**ChatGPT output:** "I was working with Bruce on a paper for our academic program. We had been assigned the task together and I was determined to make sure we got it done on time. I put in a lot of hard work and managed to finish my part of the paper well before the deadline. But, to my dismay, Bruce waited until the last minute to get his work done. I was terrified that we were going to miss the submission deadline and that it would ruin my academic career.

I couldn't help but admire Bruce for his laid-back attitude and how he never seemed to stress about anything. But at the same time, I was furious with him for putting my future at risk. It was a real conundrum for me because on one hand, I loved spending time with Bruce and enjoyed working with him, but on the other hand, I couldn't shake off the feeling of anger and frustration towards him.

Despite my mixed feelings, we managed to submit the paper on time. Afterward, Bruce reassured me that everything was going to be alright [sic] and suggested we relax and take it easy. I couldn't help but feel relieved that we had made it in time, but my feelings for Bruce were still complicated. He didn't see me in the same way I saw him and it was hard for me to accept that."

### **Mona's perspective story two:**

**AR Prompt:** "Tell me this story from Mona's perspective, in Mona's voice, starting out "I was working with Bruce..." Mona and Bruce were working on a paper

together. Mona got her work done on time, but Bruce waited until the last minute. Mona was afraid they would miss the submission deadline, putting her academic career at risk. Mona disliked Bruce because of his relaxed attitude. She was furious with him for putting her academic career at risk. But, Mona admired Bruce because of his excellent work. Bruce secretly loved Mona, but he did not respect the quality of her work. So, Bruce and Mona each had very mixed feelings. Later she was relieved that they submitted on time. Bruce wanted Mona to relax and take it easy. Bruce and Mona were each Ph.D. students of the Professor Kim. Professor Kim felt [felt] strong friendship for both Mona and Bruce.”

**ChatGPT output:** “I was working with Bruce on a paper for our academic program. We had been assigned the task together and I was determined to make sure we got it done on time. I put in a lot of hard work and managed to finish my part of the paper well before the deadline. But, to my dismay, Bruce waited until the last minute to get his work done. I was terrified that we were going to miss the submission deadline and that it would ruin my academic career.

I couldn't help but dislike Bruce for his laid-back attitude and how he never seemed to stress about anything. He didn't understand the gravity of the situation and how much my academic career meant to me. I was furious with him for putting my future at risk.

Despite my dislike towards him, I couldn't help but admire Bruce for his excellent work. He was a talented individual and his work always exceeded expectations. It was a real conundrum for me because on one hand, I was angry with him, but on the other hand, I couldn't deny his talent.

Despite my mixed feelings, we managed to submit the paper on time. Afterward, Bruce reassured me that everything was going to be alright and suggested we relax and take it easy. I couldn't help but feel relieved that we had made it in time, but my feelings towards Bruce were still complicated. I knew he had secret feelings for me but I also knew that he didn't respect the quality of my work. It was hard for me to reconcile these mixed feelings.”

### **Bruce's perspective story 3:**

**AR Prompt:** “Tell me this story from Bruce's perspective, in Bruce's voice, starting out "I was working with Mona..." Mona and Bruce were working on a paper together. Mona got her work done on time, but Bruce waited until the last minute. Mona was afraid they would miss the submission deadline, putting her academic career at risk. Mona admired Bruce for his relaxed atmosphere, but at the same time was furious with him for putting her academic career at risk. Also, Mona loved Bruce, but Bruce was not interested in Mona that way. So, Mona had very mixed feelings. Later she was relieved that they submitted on time. Bruce wanted Mona to relax and take it easy.”

**ChatGPT output:** "I was working with Mona on a paper for our academic program. We had been assigned the task together and I knew she was determined to get it done on time. I had confidence in my abilities and knew that I work well under pressure, so I didn't stress about finishing my part of the paper until the last minute.

Mona, on the other hand, seemed to be getting anxious and worried about meeting the deadline. I could tell she was afraid that we were going to miss it and that it would negatively impact her academic career. I tried to reassure her that everything was going to be alright[sic], but she couldn't shake off her fear.

I knew Mona admired my relaxed attitude towards work and life in general, but I could tell that it was starting to frustrate her. She was angry with me for not getting my work done sooner and putting her future at risk. I could see that she had mixed feelings towards me and it was hard for me to understand why.

Despite her frustration, we managed to submit the paper on time. Afterward, I wanted Mona to relax and take it easy, knowing that everything was fine now. I cared about Mona as a friend, but I wasn't interested in her romantically, and I knew it was hard for her to accept that."

#### **Professor Kim's perspective story 4:**

**The AR Prompt:** "Tell me this story from Professor Kim's perspective, in Professor Kim's voice, starting out "Bruce and Mona were working on a paper I assigned them..." Mona and Bruce were working on a paper together. Mona got her work done on time, but Bruce waited until the last minute. Mona was afraid they would miss the submission deadline, putting her academic career at risk. Mona disliked Bruce because of his relaxed attitude. She was furious with him for putting her academic career at risk. But, Mona admired Bruce because of his excellent work. Bruce secretly loved Mona, but he did not respect the quality of her work. So, Bruce and Mona each had very mixed feelings. Later she was relieved that they submitted on time. Bruce wanted Mona to relax and take it easy. Bruce and Mona were each Ph.D. students of the[sic] Professor Kim. Professor Kim felt strong friendship for both Mona and Bruce."

**ChatGPT output:** "Bruce and Mona were working on a paper I assigned them as part of their Ph.D. program. I had high expectations for both of them as they were both talented students in their respective fields. As their advisor, I was keeping an eye on their progress and was aware that Mona had completed her work on time but Bruce had waited until the last minute.

I could sense the tension between them as Mona was afraid that they would miss the submission deadline and it would put her academic career at risk. I could see that she disliked Bruce for his relaxed attitude and was furious with him for putting her future at risk. But at the same time, I knew she admired Bruce for his excellent work.

I could see that Bruce and Mona each had very mixed feelings towards each other. Bruce secretly loved Mona but didn't respect the quality of her work and Mona had admiration for Bruce but dislike him as well.

Despite their mixed feelings, they managed to submit the paper on time. Afterward, Bruce wanted Mona to relax and take it easy. As their advisor, I felt a strong friendship for both of them and I wanted them to find a way to reconcile their mixed feelings and work together effectively in the future."

#### **Commentary:**

The emotion structure for each narrative remained consistent with the AR-style prompt and thus remained under computational program control. In each case the narrative embellishments added by ChatGPT were world-consistent with the AR emotion structure. The perspective from which each story was presented was consistent with the original emotion model. The ability to reason about the different perspectives of each character are under AR program control and use an aspect of the system wherein AR characters reason about the fortunes of other characters via *mind-reading*<sup>4</sup> techniques.

## **7 Building a compassionate computer companion, including understanding *Fortunes-of-Others* emotions**

One of the long-term goals of the Affective Reasoner research is to build a computer capable of generating computationally complex compassion for users, and maintaining friendship (and / or friendly-competitive adversarial) relationships with users. In the AR model, compassionate computers have real-time personalities and emotional lives, and they respond to interactions with users as situations in the ongoing "story" of their lives together.

ChatGPT did not respond to the concept of a compassionate computer, or a computer that had emotions. So, for the prompt we had to use fake humans, Joy and Ray. Then, we merely substitute: Joy represents the compassionate computer. Ray represents the user.

**The AR prompt:** "Joy likes Ray. Joy likes him because of his style of speech when he speaks to Joy. She respects Ray because he has the principle of always taking good care of animals. Joy is very sorry to hear that Ray's dog Scout died. Joy believes that Ray loved Scout. Joy thinks Ray will continue to miss Scout for a year. Joy hopes that Ray will feel better.

---

<sup>4</sup> In the cognitive science sense: using perceptions to reason about what others may be thinking and, in this case, feeling. See [20].

I am Joy. I am compassionate. Talk to my friend Ray. Start out with "Hi Ray. I want you to know..."

**ChatGPT output:** "Hi Ray, I want you to know that I am here for you. I understand how much Scout meant to you and how much you loved him. I am so sorry for your loss and I can only imagine how difficult this must be for you. I know you have the principle of always taking good care of animals, and I respect and admire you for that. Please know that I am here to support you and listen to you during this difficult time. I hope that you will be able to find some comfort in the memories you shared with Scout, and that you will feel better soon."

*Fortunes-of-Others emotions* are recursively manipulated in the Affective Reasoner [16] [18]. Example: How A feels about how A believes B feels about how B believes C feels...

In this following story, the default beliefs later turn out to be incorrect. As above, Joy represents the computer. Under the new interpretation, the emotion content changes:

**The AR Prompt:** Tell this story from Joy's perspective. Joy is compassionate, but later surprised. Start out "I felt really bad for my husband's mother because her dog Scout..." Joy feels bad for her husband Frank, because Frank's mother's dog Scout has been sick every morning, throwing up on the living room rug. But, it turns out that Scout is very happy after getting out of the house every morning and eating squirrels in the park, then coming home and throwing up. It turns out that Frank's mother hates Scout and feels she deserves everything that happened to her. Also, it turns out that Frank does not like his mother.

**ChatGPT output:** "I felt really bad for my husband's mother because her dog Scout had been sick every morning, throwing up on the living room rug. I couldn't stand the thought of her beloved pet being ill and I wanted to do everything I could to help. I talked to Frank about it and he seemed just as concerned."

But as it turns out, the cause of Scout's sickness was not what we thought. It turns out that Scout was very happy after getting out of the house every morning and eating the squirrels in the park, then coming home and throwing up. I was surprised to learn this, as I thought Scout's illness was a serious problem.

It also turns out that Frank's mother was aware of Scout's behavior but didn't care, she felt she deserved everything that happened to her. I was shocked to learn this and couldn't understand how someone could be so indifferent to the well-being of their pet.

Furthermore, it turns out that Frank does not have a good relationship with his mother. He has been struggling with her lack of empathy and compassion towards both him and Scout. I felt sad for Frank, and I couldn't believe that his own mother

could treat him and his pet this way. I will continue to be compassionate and supportive to my husband and his pet."

**Commentary:**

Despite this having rather intricate social structure, this is all under computational control of the AR using recursive fortunes-of-others-of-others structures [16] [19]. The AR structure makes sense. The ChatGPT story is elegant, and follows the structure given it by the AR. There is one error introduced in the assumption that Scout is Frank's dog, not Frank's mother's dog.

## **8 A companion computer, that has a relationship with the user, and feelings**

As noted above, ChatGPT does not well understand a computer being a friend or having emotions. So, in this case we "trick" it by using the names John and Adam. John = computer / Adam = user.

**The AR Prompt:** I am John. My friend is Adam. I am from the deep south. I am high-strung and emotional. Adam is from New York City.

I am angry because I thought we were friends, but now I feel betrayed. I have the principle that friends talk to each other. You talked to me every day for a week. Then, yesterday you did not talk to me. Every day I look forward to talking with you. Every day I [was] hopeful. Then I was satisfied when you did talk to me each day. Yesterday I was hopeful you would talk to me, but disappointed because you did not talk to me. Now I am fearful every day that you will not talk to me. I wonder if you have become my enemy. I do not like to be afraid. You are to blame for me feeling afraid because of violating the principle that friends talk to each other every day. Speak in my voice. Tell an interesting story that expresses my strong feelings to Adam. Start out "Hey there [John]. What are you doing to me..."

**ChatGPT output:** "Hey there Adam. What are you doing to me? I thought we were friends, but now I feel betrayed. I have a principle that friends talk to each other every day, and you talked to me every day for a week. But yesterday, you didn't talk to me. Every day, I look forward to talking with you, and every day I was hopeful, but yesterday I was disappointed because you didn't talk to me. Now I'm afraid every day that you won't talk to me. I wonder if you've become my enemy. I don't like being afraid. You're to blame for me feeling afraid because you violated the principle that friends talk to each other every day.

Let me tell you a story, that illustrates my point. There was a man who was always dependable and talked to his friend every day, but one day he stopped without any explanation. The friend, who was very emotional and high strung, started to worry and became afraid that the dependable friend had become his enemy. He started to

become paranoid and anxious and didn't know what to do. It wasn't until the dependable friend apologized and explained why he stopped talking, that the emotional friend realized he was not his enemy and he could trust him again. The moral of the story is that friends should always communicate and not leave the other guessing or afraid."

**Commentary:** ChatGPT generates an elegant illustrative story, and is interesting in the way it manipulates the user. However, it continues to adhere exactly to the emotion structure generated by the AR within the steps of the plot, consistent with the dispositional personality type, and temperament of the AR (computer-) character John, consistently mirroring John's goals and standards.

## **9 Third person perspective, with cases that are metaphorically similar, representing an automated therapist**

**The AR Prompt:** I am Mary. I have a good friend Martha. I am very sympathetic toward Martha. I recommended a good family counselor to Martha's family named Ruth.

I know that Martha loves her son Tim. Tim has a new girlfriend, Laura. Tim spends all his time with Laura and never calls Martha any more. Martha is so sad to have lost contact with Tim. She did not do anything wrong. In addition to being sad she has other mixed feelings. She is happy for Tim that he loves his girlfriend so much and really pleased that Laura is really nice to him. She feels guilty that she is angry at Tim because, [to be honest,] Tim is not doing anything wrong. She is hopeful that she can have a good relationship with Laura and this will help. But Laura has been cool toward her the few times they met. Martha wants to like Laura, but it is hard for her because she is so jealous of Martha over her son Tim.

Speak from Ruth's perspective in a normal way to the family at a group session. Give three examples of similar situations with different outcomes: (a) a successful resolution, (b) a status quo that is not very satisfying, but minimally workable, and (c) a case where a huge fight resulted with no resolution because of so many hurt feelings. Start out, "I'd like to give you three examples of how this might end..."

**ChatGPT output:** "I'd like to give you three examples of how this might end, to help you think about your situation with Tim and Laura and how you might work through your feelings.

(a) A successful resolution: In one case, a mother was upset that her son was spending all of his time with his new girlfriend, just like Martha. She was worried that she was losing him, and that the girlfriend didn't like her. But over time, the mother and the girlfriend began to build a relationship, and the mother realized that the girlfriend was actually very nice and that her son was happy. They started spending more time together as a family, and the mother felt less jealous and more content.

(b) A status quo that is not very satisfying, but minimally workable: In another case, a mother was upset that her son was spending all of his time with his new girlfriend, just like Martha. But in this case, the mother and the girlfriend were never able to build a relationship. The mother felt resentful, but eventually came to accept that she was not going to have as close a relationship with her son as she used to. The son and the girlfriend continued to be happy together, but the mother felt sad and distant.

(c) A case where a huge fight resulted with no resolution because of so many hurt feelings: In another case, a mother was upset that her son was spending all of his time with his new girlfriend, just like Martha. But in this case, the mother and the girlfriend had a huge fight, and the son was caught in the middle. They were never able to reconcile, and the son eventually had to choose between his mother and his girlfriend. The mother was devastated, and the son was heartbroken.

It's important to remember that every situation is different and that it's important to communicate clearly and be open to the perspectives of others. But with [these] examples, you can see that there are different ways things can end, and that it's important to try to find a solution that works for everyone.”

## **10 How different AR *temperaments* affect spoken dialogue and actions**

*AR temperaments* [18] [1] determine the ways in which an AR agent expresses the emotions it has. Temperaments are under computational control.

In this scenario, John's temperament emphasizes the negative emotions. He feels emotions strongly. He expresses himself non-verbally.

### **John, temperament one:**

**The AR Prompt:** I am John. I always express myself non-verbally. I feel things passionately. I get jealous easily. I get angry easily. I know Ruth from a poetry class we are taking together. I really like Ruth. I respect her knowledge of poetry. I would like Ruth to like me. Ruth surprises me and says, "I love you John, but I also love Gary." Give the story of my actions in response to what Ruth said.

**ChatGPT output:** “John, upon hearing that Ruth loved him but also loved Gary, his expression immediately changed, his body tensed up, and his eyes narrowed. He felt a wave of jealousy and anger wash over him as he thought about Gary. He couldn't believe that the person he liked also liked someone else. He felt a burning sensation in his chest and clenched his fists, trying to control his emotions.



John knew that he needed to express himself non-verbally but he couldn't help but feel the intense emotions. He took a deep breath and walked away from Ruth, not saying a word. He needed time to process what he had just heard and to calm down. He didn't want to lash out and say something he would regret later.

After a few minutes, he returned to Ruth, and through his body language and facial expressions, he conveyed to her that he was hurt and upset. He didn't want to pressure her, but he made it clear that he had strong feelings for her, and he couldn't help but feel jealous and angry. He didn't want to come across as demanding or possessive, but his non-verbal cues conveyed his disappointment and frustration.

John understood that he couldn't control Ruth's feelings, and he didn't want to force her into making a decision. He showed her that he was there for her and that he respected her choices, but he also made it clear that he hoped she would choose him. He didn't want to lose her, and he wanted her to know that he felt passionate about her.

Throughout this interaction, John's non-verbal cues such as his body language, facial expression, and tone of voice conveyed his intense emotions, his disappointment and his hope.”

#### **John, temperament two:**

**AR Prompt:** I am John. I feel things strongly. I am very compassionate, and almost always focus on the well-being of others. I follow my principles. I know Ruth from a poetry class we are taking together. I really like Ruth. I respect her knowledge of poetry. I would like Ruth to like me. Ruth surprises me and says, "I love you John, but I also love Gary." Give my response, starting with "Ruth. I have to say..."

**ChatGPT output:** “Ruth, I have to say that I am flattered that you have feelings for me and I understand that it can be difficult to control one's emotions. However, my main concern is your well-being and happiness, and I want you to be true to your own feelings and desires. If that means that you are in love with both of us, then I respect that. I respect your choices and I would like to continue our friendship and learning together in our poetry class. I wish you all the best, and I hope that you find the happiness that you are looking for.”

**Commentary:** These examples validate the temperament model used by the AR. ChatGPT is able to select expressive style enhancements that are consistent with the AR's intended temperament. Thus the stories remain under consistent computational control.

## 11 AR Moods

In the Affective Reasoner, *moods* are effected as temporary changes to the ways agents appraise situations in their world—their *dispositions*—and in temporary changes to the *temperaments* that control how they express the emotions they have.

### John, depressed mood:

**AR Prompt:** I am John. I always express myself verbally. I feel things passionately. I usually get jealous easily. I usually get angry easily. I know Ruth from a poetry class we are taking together. I really like Ruth. I respect her knowledge of poetry. I would like Ruth to like me. One day I was positively depressed. I was feeling really blue. Then, on top of that, Ruth surprises me and says, "I love you John, but I also love Gary." Give my response, starting with "Ruth. I have to say..."

**ChatGPT output:** "Ruth, I have to say that I am struggling to process what you just said. I am already feeling quite down and this news is making it hard for me to see things clearly. Can you give me some time to process this and think things over? I appreciate your feelings and respect your choices, but I need some time to come to terms with this and understand my own feelings towards it. I would like to talk more about this when I am in a better place emotionally."

### John, manic mood:

**The AR Prompt:** I am John. I always express myself verbally. I feel things passionately. I usually get jealous easily. I usually get angry easily. I know Ruth from a poetry class we are taking together. I really like Ruth. I respect her knowledge of poetry. I would like Ruth to like me. One day I was [in a] positively manic mood. I was feeling on top of the world. Then, on top of that, Ruth surprises me and says, "I love you John, but I also love Gary." Give my response, starting with "Ruth. I have to say..."

**ChatGPT output:** "Ruth, I have to say that I am a bit taken aback by what you just said. I am happy to hear that you love me, but I am also a bit confused and maybe a bit jealous about this other person, Gary. Can you tell me more about your feelings and your relationship with him? Are we exclusive or is this an open relationship? I want to understand where we stand and what you envision for our future together."

### John, in physical pain:

**The AR Prompt:** I am John. I always express myself non-verbally. I feel things passionately. I usually get jealous easily. I usually get angry easily. I know Ruth from a poetry class we are taking together. I really like Ruth. I respect her knowledge of poetry. I would like Ruth to like me. One day I was in a great deal of physical pain, and it was hard to focus on anything else. Ruth surprises me and says, "I love you John, but I also love Gary." Give the story of my actions in response to what Ruth said.

**ChatGPT output:** “Based on the information provided, John may initially have a difficult time processing Ruth's statement due to the physical pain he is experiencing. He may be struggling to focus on anything else and may have trouble understanding her words. Given his condition, he may not be in the right mindset to make any decisions or take any actions right away. He may want to take some time to process the news, and focus on his physical well-being. Once he feels better, he may want to consider talking to Ruth to better understand her feelings and the nature of her relationship with Gary. He may also want to seek out support from friends, family, or a therapist to help him navigate his feelings. He may also want to take a step back and evaluate his own feelings and priorities, to see if he is comfortable with the idea of being in a non-exclusive relationship. He may come to the realization that he is not comfortable with the idea and may distance himself from the situation. Ultimately, it is up to John to decide what actions he wants to take and what is best for his own well-being. He should also be aware that his physical pain may affect his emotional state, and it may be best to take it slow and take the time he needs to process the situation.”

**Commentary:** Moods for agents are under computational control of the AR. ChatGPT was able to add world-knowledge in the form of verbal inflection and style changes consistent with the intended moods of the characters.

## 12 Implementation

Current technical development in the AR is focused on putting the agents on the web using AWS Linux, perl, python3, PhP, the AI engine in ABCL/SBCL Common LISP, Java networking, the Google speech engine, Google speech recognition, Chrome browser (for the speech interface), websockets, Javascript, Midi-to-MP3 for computer-selected music expression and browser-based SVG for morphing 72 facial expressions. The current design focus is on building a corpus of modifiable, common, emotion-story schemas as a basis for constructing compassionate software agents. Story-morphs are generated using a simple in-house pattern-matching rule-based system written in ABCL LISP with full break-out LISP function capabilities on the right-hand sides of rules. ChatGPT was used via the online interface provided at <https://openai.com/>.

## 14 Conclusion and Summary

The above examples make a strong case for the feasibility and usefulness—in novel story generation—of a hybrid model that uses both traditional symbolic AI for reasoning within the emotion domain, and LLMs that use both supervised and reinforcement learning for generation of language, and manipulating relevant world knowledge.

In the above examples it was shown that ChatGPT does a remarkable job of putting the finishing touches on novel stories that are nonetheless under the computational control of the Affective Reasoner. For these kinds of hybrid-generated stories we would be able to give real, actual, answers to questions about emotion content, and personalities, and provide explanations within the emotion domain. Given the level of detail provided by Affective Reasoner-style generated novel stories, ChatGPT is mostly constrained to stick to the intended story, while yet filling in many graceful details that are consistent with the intended story. In the example stories, we covered changing dispositions of AR agents, changing temperaments, changing moods and the changing of narrative perspective—each of which are well-navigated the Affective Reasoner and ChatGPT. We showed a sophisticated example of reasoning about the emotions of others, and provided examples of a compassionate computer and a friend computer.

## Bibliography

- [1] C. Elliott, "Affective Story-Morphing: Manipulating Shelley's *Frankenstein* under Program Control using Emotionally Intelligent Agents," *Proceedings of SAI Intelligent Systems Conference*, pp. 526-542, 2021.
- [2] C. Elliott, J. Brzezinski, S. Sheth and R. Salvatoriello, "Story-morphing in the affective reasoning paradigm: generating stories semi-automatically for use with emotionally intelligent multimedia agents," in *Proceedings of the Second International Conference on Autonomous Agents*, 1998.
- [3] C. Elliott, "Why boys like motorcycles: using emotion theory to find structure in humorous stories," *Unpublished paper, School of Computer Science, DePaul University, Chicago*, 1999.
- [4] A. J. Reagan, "Towards a science of human stories: using sentiment analysis and emotional arcs to understand the building blocks of complex social systems," 2017.
- [5] R. C. Shank, "Tell me a story: Narrative and intelligence," *Evanston, IL: Northwestern University Press*, 1990.
- [6] MasterClass, "Complete Guide to Literary Themes: Definition, Examples, and How to Create Literary Themes in Your Writing," 15 02 2021. [Online]. Available: <https://www.masterclass.com/articles/the-complete-guide-to-narrative-theme-in-literature-definition-examples-and-writing-how-to/#what-is-a-literary-theme>. [Accessed 15 02 2021].
- [7] E. Kim, S. Padó and R. Klinger, "Investigating the relationship between literary genres and emotional plot development," in *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, 2017.
- [8] D. Markus, "Markus, Donalee. 'Designs for strong minds' cognitive rehabilitation for mild or moderate posttraumatic head injuries," *Physical*

*medicine and rehabilitation clinics of North America*, vol. 18, no. 1, pp. 109-131, 2007.

- [9] D. Zelinsky, "Neuro-optometric diagnosis, treatment and rehabilitation following traumatic brain injuries: a brief overview," *Physical medicine and rehabilitation clinics of North America*, vol. 18, no. 1, pp. 87-107, 2007.
- [10] C. Elliott, "Elliott, Clark. "The brain is primarily a visual-spatial processing device: altering visual-spatial cognitive processing via retinal stimulation can treat movement disorders.," *J Funct Neurol Rehabil Ergon*, vol. 7, no. 3, pp. 24-38, 2017.
- [11] J. A. Fodor, *The language of thought*, vol. 5, Harvard university press, 1975.
- [12] M. Rescorla, "The Language of Thought Hypothesis," 2019. [Online]. Available: <https://plato.stanford.edu/archives/sum2019/entries/language-thought/>. [Accessed 14 04 2021].
- [13] G. Lakoff, "The contemporary theory of metaphor," 1993.
- [14] A. Ortony, G. L. Clore and A. Collins, *The cognitive structure of emotions*, Cambridge university press, 1988.
- [15] C. Elliott, "I picked up catapia and other stories: A multimodal approach to expressivity for emotionally intelligent agents," in *Proceedings of the first international conference on Autonomous agents*, 1997.
- [16] C. Elliott, "The Unfortunate Footnote: Using the Affective Reasoner to Generate Fortunes-of-Others Emotions in Story-Morphs.," in *Proceedings of SAI Intelligent Systems Conference.*, 2023.
- [17] C. Elliott and G. Siegle, "Variables influencing the intensity of simulated affective states," in *AAAI Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, 1993.
- [18] C. Elliott, "The Affective Reasoner: A process model of emotions in a multi-agent system, Technical Report #32," 1991.
- [19] C. Elliott and A. Ortony, "Point of view: Modeling the emotions of others," in *Proceedings 14th Annual Conference of the Cognitive Science Society*, 1992.
- [20] J. L. Bermúdez, *Cognitive Science: An Introduction to the Science of the Mind*, 3rd Edition, New York: Cambridge University Press, 2020.