Poetry generation system with an emotional personality

Joanna Misztal[1] **and Bipin Indurkhya**[2]
[1] Faculty of Mathematics and Computer Science, Jagiellonian University, Krakow, Poland
[2] Computer Science Department, AGH University of Science and Technology, Krakow, Poland

Abstract

We introduce a multiagent blackboard system for poetry generation with a special focus on emotional modelling. The emotional content is extracted from text, particularly blog posts, and is used as inspiration for generating poems. Our main objective is to create a system with an empathic emotional personality that would change its mood according to the affective content of the text, and express its feelings in the form of a poem. We describe here the system structure including experts with distinct roles in the process, and explain how they cooperate within the blackboard model by presenting an illustrative example of generation process. The system is evaluated considering the final outputs and the generation process. This computational creativity tool can be extended by incorporating new experts into the blackboard model, and used as an artistic enrichment of blogs.

Introduction

> Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain. (Lister 1949)

This expresses one of the strongest requirements for AI quoted by (Turing 1950). It takes the view that the process of expressing feelings by means of artistic artifacts is a hallmark of human capability. Such requirements have created a challenging task for AI: how to design a computer program that could write a sonnet inspired by its thoughts and emotions. In recent years, various poetry-generating systems have been developed, discussed in more details below, some of which focus only on producing entertaining artifacts, while others simulate the creativity process and incorporate affective computing techniques. However, most of them do not model a sense of *self* capable of expressing its own feelings. The main goal of this project is to take up this challenge and to create a system with an emotional personality. Specifically, we plan to create an empathic system that changes its mood according to the emotions evoked by reading the given text, and expresses them in the form of a poem.

The *affective empathy* has been defined in the psychological literature as the observer's emotional response to the affective state of others (Davis 1983). Similarly, we propose a term *computational empathy* to mean recognition and interpretation of emotions of another person by the computer system. Our work introduces a system with a complex emotional model that attempts to understand affects in human artifacts, and expresses those feelings in the form of a poem. The design considers an optimism rate which is an individual feature of the system influencing its perception of the environment (the text).

This paper is organized as follows. The *Background* section presents existing approaches to sentiment analysis and emotional modeling. It also presents the blackboard idea and other poetry-generation systems. The *Overview* section explains the general idea of the system. The poetry-generation process in our approach is implemented on a blackboard model, which is described in the *System Architecture* subsection. In this approach, the poetry is composed by a group of experts - each of whom has some specific knowledge about the poetry-generation process, and all of them share a global work-space called the blackboard.

The details of the poetry-generation algorithm are presented in the *Poetry Generation Algorithm* section and explained with an illustrative example. The system takes the inspiration for its creativity from the text provided by the user. Key phrases are extracted from the text to determine the theme of the poem, and also to set its sentiment. The key phrase that is found to be the most *inspiring* by the experts is used as the title and main theme of the poem.

The experts start to perform their tasks - *words-generating experts* produce words related to the topic based on their knowledge. Some of them use lexical resources such as synonyms dictionary or word collocations. There is also one expert incorporating a model of emotional intelligence that defines the mood evoked by the given text, and generates words describing this sentiment.

The *poem-making experts* choose words from the pool and try to arrange them into phrases. Each of them uses its own Context-Free Grammar to construct phrases. Some *poem-making experts* use poetic tropes like metaphors or epithets to enrich the style.

The *evaluating experts* select the best phrases according to some constraints, considering the stylistic form.

The *control component* tries to regulate the poem composition by maximizing its diversity and choosing the experts

that were the least frequent before.

Some illustrative results are presented in the *Examples* section. *Evaluation* contains a summary of system's performance in the context of the proposed algorithm and the evaluation of the final outputs.

Current version of the system includes some basic types of experts. However, the blackboard architecture allows the system to be extended by adding new experts. Possible improvements and proposition of new experts as well as possible application of the program are mentioned in the *Conclusions* section.

Background

Sentiment analysis and affective lexical resources

The goal of text sentiment analysis is to extract the affective information or writer's attitude from the source text. Basically the sentiments may be considered within the polarity classification (*positive, negative or neutral*). However, this method does not provide us with a detailed understanding of the author's emotional state, and another approach is needed.

The computational methods for sentiment analysis are usually based either on machine learning techniques such as naive Bayes classifiers trained on labeled dataset, or use lists of words associated with the emotional value (positive-negative evaluation or sentiment score values). In our research we use ANEW database consisting of nearly 2500 words rated in terms of pleasure, arousal, and dominance (Bradley and Lang 2010) for text arousal calculation.

To extract the sentiment evaluation, we use the *Sentistrength* (Thelwall et al. 2010) sentiment analysis tool. It estimates the negative and positive sentiment values in short informal texts (rating both positive and negative scores with 1-5 scale), considering common and slang words, emoticons and idioms. The base of the algorithm is the *sentiment word-strength list* containing terms with 2-5 scale of positive or negative evaluation. The initial, manually-prepared words-sentiments list has been optimized by a training algorithm to minimize the classification error for some training texts. The system also considers a spelling correction algorithm and *booster words list* with terms that can increase or decrease other words' scores (such as *very, extremely*) as well as *negating word list* with terms which may invert emotion value (*not, never*). Additionally, the algorithm uses a list of emoticons commonly used in social web texts, and considers some other stylistic parameters such as questioning and repeated letters.

In our approach, we also use the WordNet-Affect lexical resource (Strapparava and Valitutti 2004) to build a hierarchy of words describing emotional states that are used later to generate the affective content of poems. The lexicon contains WordNet hyponyms of the emotion word, which are a subset of synsets suitable to represent affective concepts correlated with affective words. For example, for the emotional word *compassion,* we can derive a correlated set of words describing this state: *forgive, merciful, excusable, affectionate, commiserate, tender*.

Emotional modeling

As mentioned in (Cambria, Livingstone, and Hussain 2012), the research on human emotions dates back to ancient times. One of the first categorization of emotional states was made by Cicero who separated them in four categories of fear, pain, lust and pleasure. Later studies on this topic were developed by Darwin (19th century), Ekman (who defined six basic emotions as happiness, sadness, fear, anger, disgust and surprise in 1970s) and many others.

One approach towards emotional modeling that has been commonly used by scientists since 20th century is the dimensional model, where particular emotions are represented as coordinates in a multi-dimensional space. One of the first examples is the circumplex model (Figure 1) presented in (Russel 1980). In this model, the horizontal (...) dimension is the pleasure-displeasure and the vertical is arousal-sleep(Russel 1980). In the Whissel's model (Whissel 1989), the 2D spatial coordinates are evaluation (*positive-negative*) and activation (*passive-active*). The author places words from her Dictionary of Affects in Language in this space. Another example of such model is Plutchik's wheel of emotions (Plutchik 2001) consisting of 8 basic and 8 composed emotions placed in the circle, where the similarity of emotions is represented by radial dimension.

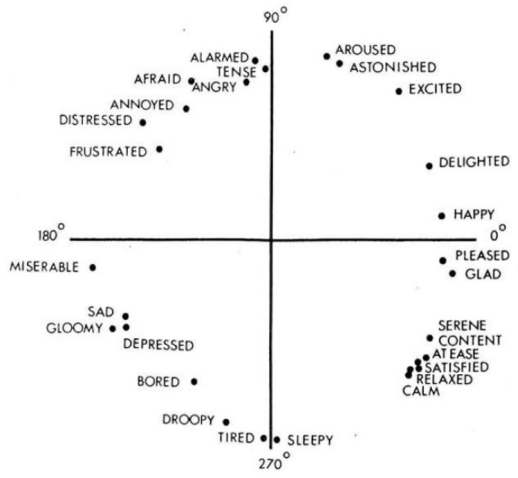

Figure 1: 2D circumplex model of emotions adatpted from (Russel 1980).

The dimensional models are a promising tool for computational modeling of emotions as they provide simple way to measure, define and compare the affective states. They are used in AI systems to simulate the emotional personality as presented in (van der Heide and Trivino 2010; Kirke and Miranda 2013). However, they have some significant limitations as they are based mostly on the verbal representation of affects. As mentioned in (Cambria, Livingstone, and Hussain 2012), they do not allow defining more complex emotions and they do not consider the situation of several emotions being experienced at the same moment.

Blackboard architecture

According to the Global Workspace Theory (Baars 1997; 2003) the brain functioning may be illustrated by a theater metaphor where:

> Consciousness (...) resembles a bright spot on the stage of immediate memory, directed there by a spotlight of attention, under executive guidance. The rest of the theater is dark and unconscious. (Baars 2003)

Thus, in the conscious part the actions are performed by a large number of autonomous specialized modules (the actors).

The blackboard architecture is a model that fulfills the assumptions of GW Theory of mind and therefore has a potential to be used in simulating cognitive processes such as creativity. The model may be visualized by another metaphor (Corkill 1991) of a group of independent experts with diverse knowledge who are sharing a common workspace (the blackboard). They work on the solution together and each of them tries to add some contribution on the blackboard until the problem is solved. The blackboard model is an appropriate solution for problems that require use of many diverse sources of knowledge, or for ill-defined, complex problems. It allows a range of different "experts" – they may be represented as diverse computational models as their internal representation is invisible at the top level.

The idea of using experts representing knowledge has been previously used to simulate cognitive tasks. For example in Word Expert Parser (Small 1979), experts cooperate to provide better understanding of text during the process of conceptual analysis of natural language.

Poetry-generation systems

Since making a system that would produce aesthetically pleasing poems based on predefined templates is not such a difficult task, there exist various poetry-generation programs working in this way. An elaborate example is Kurzweil's *Cybernetic Poet* (Kurzweil 1992), which generates a language model from a set of poems input by the user, and composes new ones in the same style. However, a really challenging task is to make a program that produces the poems in an intentional way. (Gervas 2010) notes that the simulation of human creativity may be significantly different from the original process of creativity itself. Accordingly, there exist various approaches towards computer poetry generation. The McGONAGALL system (Manurung, Ritchie, and Thompson 2012) uses evolutionary algorithms to make a poem that fulfills the constraints on *grammaticality, meaningfulness and poeticness*. ASPERA (Gervas 2001) generates poems with a forward reasoning system. (Toivanen et al. 2012) present a system that creates novelty by substituting words in existing Finnish poetry. In subsequent work, (Toivanen, Jarvisalo, and Toivonen 2013) introduce a constraint programming technique for poetry generation. There are also several projects that incorporate emotional affects in the creation process. (Colton, Goodwin, and Veale 2012) present a corpus-based poetry generator that creates poems according to *days mood* estimated from the news of the day. However, the mood is only defined as *good* or *bad*, without any further refinement of the emotional state. The *Stereotrope* system (Veale 2013) generates the emotional and witty metaphors for given topic based on the corpus analysis. Another interesting approach is MASTER (Kirke and Miranda 2013), which is a tool for computer-aided poetry generation. In this system, a *society* of agents in various emotional states influences each other's moods with their pieces of poetry. The final poem is a result of *social learning*. The poems produced by the system are not meaningful in the usual sense, but they consist of repeated words and sounds that create poeticity.

Among the above-mentioned systems, we can distinguish two different approaches towards modeling the system's *personality*. In the first approach, the system's behavior is determined by some predefined parameters (e.g. in MASTER - agents have initial moods and words). Another alternative is to adapt the emotional state to some environmental factors. This approach is taken by (Colton, Goodwin, and Veale 2012), where the mood of the day is calculated from the sentiment value in daily news. The *Cybernetic Poet* also builds a data-driven model, but it does not exhibit any creative nor emotional behaviors – the system can only replicate the style of the existing poetry.

In our system, we combine both approaches - the emotional state is acquired based on the affective information extracted from the blog text, but it is also dependent on the individual features of the system – the model of emotions and its optimism rate that give the system an individual personality. Hence the external factors are used only as an inspiration for the theme and stimulus for the affective state.

Our approach may be also compared to MASTER, which is also a multi-agent model for poetry generation with emotions. In MASTER (Kirke and Miranda 2013), the agents interact by reciting their own pieces of poetry to each other. Thus, in contrast to our model, they do not share any global knowledge. The mood-defining factor for MASTERs agents is the poetry produced by the societal agents themselves. Hence the method for calculating the emotional state differs from ours, where we extract sentiments from web text. Moreover, all of the agents in (Kirke and Miranda 2013) have the same structure, while in the blackboard model they represent diverse computational units with distinct knowledge sources and roles.

Our approach may be considered as similar to the idea of using specialized families of experts that cooperate during the poetry-generation process incorporated in the later version of WASP (Gervas 2010). Groups of experts work there as a *cooperative society of readers/critics/editors/writers*. However, WASP does not incorporate the blackboard model directly.

Evaluation approaches

The evaluation of any creative system is a nontrivial problem. As the task is not only to generate a satisfying output but also to imitate the creation process, the evaluation needs to consider both the aspects. The most obvious way to evaluate the output is to make a kind of Turing test (Turing 1950) for poetry as in (Kurzweil 1992). In such a test, some computer-generated poems mixed with the human-authored

poetry are presented to the human judges. The score is based on how many poems composed by the system were classified by judges as human-authored. However, the domain-specific Turing test does not consider the evaluation of the creation process. Another approach, taken in FACE descriptive model (Colton, Charnley, and Pease 2011), is based on evaluating the generative act performed by the system and its impact. FACE introduces a set of parameters evaluating the creativity of the program, and considers not only the artifacts produced by the system but also the process of generation, which is essential for creativity evaluation. A *creative act* that satisfies all FACE criteria is denoted by a tuple $< F^g, A^g, C^g, E^g >$, where the *C – concept* means the system taking input and producing outputs denoted by *E – expressions*, the *A – aesthetic measure* is the fitness function evaluating the (concept, expression) pairs with real-number values and the *F – framing information* is the linguistic comment explaining the context or motivation of the outputs.

Overview

The system structure is based on the blackboard model. It consists of a group of experts that represent diverse sources of knowledge, the common blackboard workspace and the control component that regulates the process by choosing one of competing experts that will contribute to the final solution. The modules are described in the *System architecture* subsection.

At the beginning of the poetry-generation process, the input text is set on the blackboard and the agents start to work on it. Each agent has a special role and knowledge and it waits until it finds something on the common workspace that it can use for performing its task. When something interesting appears, the agent processes the information using its individual knowledge and adds new partial solution to the blackboard. The control module decides which agent's contribution should be used for the final poem. The algorithm is explained in more details in *Poetry generation algorithm* subsection along with an illustrative example of the generation process.

System architecture

The system architecture is presented in Figure 2. The main modules of the system are described below.

Blackboard is a common workspace with partial solutions and other information about the problem, shared by the experts. In our system, it consists of:

Text – The input text which is used as an inspiration for the poem. The experts analyze it to define the main theme and sentimental content for the poem.

Constraints – The initial constraints and information about the poem. In the example, we use constraints on the number of lines, the number of syllables in each line and the grammar constraints on tense and person to ensure grammatical consistence of the poem. These constraints are set manually at the beginning of the process or chosen randomly by the system.

Key phrases – Most frequent noun phrases retrieved from the text by one of the experts. Each phrase has its *inspiration* value defined by $W * Cat$, where *W* is number of words that the experts can generate from this phrase and *Cat* is number of non-empty categories of words (categories are explained in *Pool of ideas*).

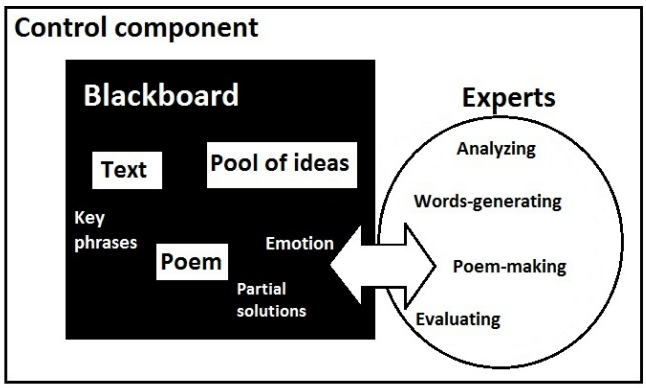

Figure 2: Blackboard architecture used in the system. A group of experts that represent diverse sources of knowledge works on the common blackboard workspace. The control component regulates the process by choosing one of competing experts that will contribute to the final solution.

Topic – The main theme for the poem selected from the key phrases as the phrase with highest *inspiration* score. If there are more phrases with the same value, one is selected at random. Once the topic is set, the experts start to produce their artifacts associated with it.

Emotion – The emotional state for the poem defined by one of the experts by analyzing sentiments in sentences from the text containing the *topic* phrase.

Pool of ideas – A part of blackboard that is used as a workspace for experts. It contains all words and partial solutions produced by the experts. It is also a source of inspiration, as some of them use artifacts generated by others to produce new ones. The expressions in the pool are divided into categories based on their grammatical form and meaning. The main categories are:

Nouns – list of nouns from the topic phrase and their synonyms.
Adjectives – list of adjectives from the topic phrase and their synonyms.
Epithets – lists of adjectives that are most frequently preceding the noun for each noun from the topic phrase.
Verbs – lists of verbs that are most frequently following the noun for each noun from the topic phrase.
Comparisons – lists of nouns that are most frequently following the adjective for each adjective from the topic phrase.
Hypernyms – lists of hypernyms of the noun for each noun from the topic phrase.
Antonyms – lists of antonyms of the words for each noun and adjective from the topic phrase.
Emotional words – words describing the emotional state defined for the poem.
Phrases – list of expressions generated by experts, candidates for the new line in poem.

Poem draft – Current version of the poem consisting of lines. Each line is selected from *phrases* candidates by the *evaluation experts*.

Model of emotions – A 2-dimensional model, where each emotional state is represented by coordinates in *(valence, arousal)* space. The emotions used in the model are WordNet hyponyms of the word *emotion* used in WordNet-Affect lexicon in the hierarchy of emotional categories. The *(valence, arousal)* coordinates for emotional labels in the model have been retrieved from the ANEW database. The choice of emotional categories is based on the lexical resources that we use. It is possible to improve the model by rearranging the categories or their spacial coordinates or to use other more complex models of emotions as mentioned in the *Background* section.

Experts – Independent modules that have access to the common blackboard. They are triggered by events on the blackboard – when they find something that they can use, they try to add new information to the blackboard. Each of them has an individual knowledge and they have diverse roles in the system.

Analyzing experts – Experts that retrieve information from the initial text and add their data to the blackboard.

Keywords expert – Extracts the most frequent noun phrases from the text and adds them to the *key phrases* section on the blackboard.

Emotion expert – Defines the emotional state for the poem and sets the *emotion* on the blackboard. As the whole text may be long, and the emotional attitude may vary within it, the sentiments are considered only for the sentences containing the topic of the poem. Sentiments are calculated in terms of *valence* (positive/negative evaluation of pleasure scaled to -5 to 5) and *arousal* (passive/active scaled to -5 to 5) levels. The valence of the text is calculated by using SentiStrength tool, which estimates the negative and positive sentiment strength in sentences based on the *Emotion Lookup Table*. However, as we want our system to represent an independent emotional intelligence, it should perceive the affects of the text in a more subjective way. Therefore, we introduced the *optimism rate* which is a parameter set at the beginning of the algorithm (or chosen randomly) that biases the valence result so that the perception of the text may be more optimistic or pessimistic. Thus, the final valence estimated by the program is given by:

$$V = \alpha_{opt} \cdot \sum_{s \in Text} Sent_{pos} + (2 - \alpha_{opt}) \cdot \sum_{s \in Text} Sent_{neg} \quad (1)$$

where α_{opt} is the optimism rate of the system (between 0,7 and 1,3), $\sum_{s \in Text} Sent_{pos}$ and $\sum_{s \in Text} Sent_{neg}$ is the sum of positive and negative sentiments respectively for all sentences in the text.

The arousal value has been calculated with use of ANEW. The algorithm combines the average ANEW arousal value for the words in text. The basic formula for arousal calculation:

$$A = (\sum_{w \in Text} A_{ANEW}(w))/length(Text) \quad (2)$$

where $A_{ANEW}(w)$ is the arousal value of word w retrieved from ANEW database. However, the sentiment in the text may be expressed not only within words but by other features of the text, similarly to expressing emotions with voice intonation in a spoken message. For example, the text *"That's great..."* can be perceived as less arousing than the same words written in a different way: *"That's GREAT!!!"*. Hence, the arousal calculation uses a punctuation-sensitive algorithm, i.e. some punctuation marks in the text increase the arousal value, while others decrease it. The calculated arousal score may be modified according to the rules:

$$f(A) = \begin{cases} A - 1 & \text{if "..." in text} \\ A + 1 & \text{if "!" in text or word in capitals in text} \\ A + 2 & \text{if "!!!" in text} \end{cases} \quad (3)$$

where A is the text arousal.

Once the valence and arousal of the text are calculated, the emotional state is defined as follows:

$$emotion = \arg\min_{x \in S} d((v_t, a_t), (v_x, a_x)), \quad (4)$$

where *emotion* is the current emotional state, *S* is the set of all emotional states from the *model of emotions*, v_t and a_t are the valence and arousal of the text, v_x and a_x are valence and arousal of the emotional state and $d(x_1, x_2)$ is Euclidean distance.

Words-generating experts – Experts that have some lexical knowledge. They generate words associated with the topic and add them to the *pool of ideas* sections.

WordNet expert – generates synonyms, hypernyms and antonyms for nouns and adjectives based on the WordNet lexical resource (Miller 1995). Adds to *nouns*, *adjectives*, *hypernyms*, *antonyms* sections of the pool.

Collocation expert – generates words that are frequently used together with given nouns and adjectives. Retrieves information from 2gram model of texts from Brown Corpus. Adds adjectives that describe nouns to the *epithets* section, verbs that follow nouns to the *verbs* section and nouns that follow adjectives to the *comparisons* section of the pool.

Emotional-Words expert – generates words that describe the emotional state defined for the poem. The affective words are derived from WordNet Affect as the hyponyms of given category name. For instance, if the emotional state was defined as *calmness*, the generated set of words would contain *peace, calm, tranquilly, easiness, cool, still*.

Poem-making experts – Experts that compete to produce new lines for the poem. They use partial solutions generated by other experts in the *pool of ideas* to produce new phrases. Their outputs are added to the *phrases* section of the *pool* and are evaluated by the *selection experts*. These phrases may be also extended by others. These experts are triggered when they find something on the blackboard that they could use for their phrases. They can generate a number of phrases proportional to their *importance factors* that are set manually at the beginning of the algorithm. Some of these experts compose stylistic forms typical for poetry.

Grammar experts – Experts that use Context-Free Grammar rules to produce phrases.

Apostrophe expert – Generates apostrophes with the noun,

its description and hypernym. For example: *O life the heavenly being*
Comparison expert – Generates comparisons for adjectives using nouns that are most frequently described by them. For example: *As deep as a transformation*
Epithet expert – Generates expressions with a noun and its epithets or emotional adjectives. For example: *marvelous sophisticated fashion*
Metaphor expert – Generates metaphors by comparing the person to an object. For example: *You were like the downtalking style*
Oxymoron expert – Composes phrases with antonym words. For example: *good and bad*
Rhetorical expert – Composes rhetorical questions about noun, or noun and its epithets. For example: *why was the style so peculiar ?*
Sentence expert – Generates sentences according to its grammar rules. Uses all the words categories, and also the emotions describing words. For example: *She loved the peaceable new york*

Recycling experts – Experts that generate new phrases by transforming phrases generated by other experts.
Exclamation expert – Generates a new phrase by adding "!" exclamation mark to the phrase from the pool.
Overflow expert – Generates a new phrase by breaking phrases from the pool into two lines.
Repetition expert – Generates a new phrase by repeating a phrase from the pool .

Selection experts – Experts that select the best solutions according to given constraints and heuristics.
Inspiration expert – Selects the topic for the poem from the set of key phrases according to formula:

$$Topic = \underset{x \in Keyphrases}{\arg\max} \; W_x \cdot Cat_x, \tag{5}$$

where W_x is the number of words that the experts can generate from this phrase, and Cat_x is the number of non-empty categories to which these words belong.
Syllables expert – Selects phrases that have the number of syllables closest to the target number of syllables for the current line in poem.

$$Lines = \underset{x \in phrases}{\arg\min} \; |S_x - S_t[i]|, \tag{6}$$

where i is current line number, S_x is number of syllables in phrase x, $S_t[i]$ is number of target syllables for line i The syllables are counted using the CMU Pronouncing Dictionary combined with the syllables-estimating algorithm used for words that are not included in the dictionary.

Control component – the unit responsible for setting initial constraints for the poem, setting experts' probabilities and evaluation expert whose contribution should be used for the current line of poem. In the current version of the system, the constraints are set for the number of lines and the numbers of syllables in each line, grammar form and tense. The stylistic constraints are selected at random from a set of templates. The experts' importance factors are chosen manually, and are used during the generation process when an expert produces a number of phrases proportional to its importance factor. The control module also tries to maximize the diversity of the poem by giving preference to the artifacts generated by those experts that contributed less frequently before. For instance, if the poem consists of two lines generated by the *grammar expert* and one by *apostrophe expert*, and for the fourth line the *grammar expert* is competing with the *oxymoron expert*, the control component will give preference to the *oxymoron expert.*

Poetry generation algorithm

We present below the generation process along with an illustrative example. The algorithm can be divided into following phases:

Modules initialization Blackboard is initialized with the text input by the user. The form of the poem is selected from a set of templates, and grammar constraints are defined for stylistic consistency.

Text:

When someone leaves you, apart from missing them, apart from the fact that the whole little world you've created together collapses, and that everything you see or do reminds you of them, the worst is the thought that they tried you out and, in the end, the whole sum of parts adds up to you got stamped REJECT by the one you love. How can you not be left with the personal confidence of a passed over British Rail sandwich? [1]

Constraints:

Number of syllables in lines: (line 1: 8; line 2: 8; line 3: 8; line 4: 8)
Grammar form:
Person: *she*; Tense: *present*;

Poem-making experts are initialized with individual *importance factors* varying from 1 to 5, determining how many phrases they can generate in each turn. The default values presented below may be modified manually.

Poem making experts importance factors: *Apostrophe expert: 2, Comparison expert: 3, Epithet expert: 5, Metaphor expert: 2, Oxymoron expert: 2, Rhetorical expert: 3, Sentence expert: 5, Exclamation expert: 1, Overflow expert: 1, Repetition expert: 1.*

Emotional expert is initialized with a random *optimism factor* between 0,7 and 1,3. A higher value means a more optimistic attitude.

Optimism factor: 0,84

Topic selection The topic is chosen as the most *inspiring* key phrase from the text. To define it, first all key phrases are retrieved and evaluated with the *inspiration* score.
Keywords expert extracts key phrases as the most frequent phrases consisting of a noun and descriptive adjectives.
Key phrases:

[someone, end, whole little world, whole sum, british rail sandwich, parts, personal confidence, fact]

[1] `http://www.jaceandjenelle.com/my-personal-blog.php`

Words-generating experts estimate how many words they can produce from each key phrase. The *inspiration* for each phrase is calculated according to formula (5). The *inspiration expert* selects the *most inspiring* phrase for the topic.

Inspirations: *whole little world: 6920, personal confidence: 3920, whole sum: 3880, someone: 2324, parts: 1918, fact: 1512, end: 910.*

Poem topic: *Whole little world*
Emotional expert defines the emotional state for the poem. The sentiments are retrieved from sentences containing the topic phrase. The expert calculates *valence* and *arousal* according to (1), (2) and (3). Then the emotional state is defined as in (4).

Sentences containing topic phrase :

When someone leaves you, apart from missing them, apart from the fact that the whole little world *you've created together collapses(...).*

Valence: -0.94;*Arousal:* 2.0; *Emotional state: despair*.

Words generation Once the topic and emotional state for the poem are defined, the *words-generating experts* start to produce their ideas. They store their artifacts under appropriate categories in the *pool of ideas* section of the blackboard.

Pool of ideas:

Nouns – *[macrocosm, existence, universe, cosmos, world, creation]*
Adjectives – *[whole, little, small]*
Verbs – *existence: [loses, reflects, becomes, fails, is, belongs], world: [centered, admired], universe: [is, had, are, was], creation: [is, does, prevents]*
Epithets – *world: [little, contemporary, real, previous] , existence: [happy, celestial, historical], universe: [interdependent, entire], creation: [own, inventive, artistic*
Comparisons – *whole: [lines, block, incident, country]*
Hypernyms – *existence: [state], world: [natural object], creation: [activity]*
Antonyms – *whole: [fractional], little:[big]*
Emotional words – *[pessimistic, cynical, resignation, discourage, hopeless]*

Phrases generation As the words start appearing in the *pool of ideas*, the *poem-making experts* start to produce phrases for new lines according to grammar constraints. They add their artifacts to the *phrases* section.

Phrases:

Epithet Expert: *corporate existence, great world*
Apostrophe Expert: *oh world the little natural object*
Sentence Expert: *the creation prevents abjectly, she likes the hopeless, she loves the pessimistic cosmos*
Comparison Expert: *as whole as a story, whole like a convocation*
Metaphor Expert: *she is like the human existence*
Exclamation Expert: *as whole as a story!*
Rhetorical Expert: *why is the existence so nonfunctional?*
Oxymoron Expert: *whole but fractional*

Line phrase selection When all experts finish their generation, the phrases that fulfill the line constraints best are selected by *selection experts*. Then the *control module* makes the final selection judging by the experts' frequencies in former lines. The same algorithm is repeated for each line of the poem.

Generating line 4. Target syllables number: 8 *Poem:*

line 1: *what is the jewish cosmos?* (Rhetorical Expert)
line 2: *o existence the daily state* (Apostrophe Expert)
line 3: *perceptual physical world* (Epithet Expert)

Syllables expert – best phrases candidates:
happy corporate existence (Epithet Expert) : 8,
she sees the pessimistic world (Sentence Expert): 8
Control module – selecting less active experts in former lines generation:
Epithet Expert: 1 line,
Sentence Expert: 0 lines
Line phrase selection: *she sees the pessimistic world* (Sentence Expert)

Examples

Below we present some example outputs of the system inspired by three input texts. We include some remarks on the interpretation of the produced poems, which are further analyzed in *Evaluation* section.

Compassionate poem about the life

Inspired by the text:

With the holiday craziness yesterday, and having to work, i didn't get to finish posting all of my thankfulness pictures. So you might see them pop up over the next few days.this morning i am thankful for the adult men in my life. My dad and mr P. i am fortunate to have both of them in my life to encourage me, support me, take care of me, and love the kids with all of their hearts.[2]

Topic: *Life*, Emotion: *compassion*
Poem:

O life the personal beingness
You are like the simple life!
Musical sacrificial life
You are like the general life
You see the excusable life
Emotional musical life
O life the heavenly being

Remarks:

The topic *Life* provided a wide range of epithets associated with the main phrase. Produced output presents a big lexical diversity of adjectives describing *life* what creates the poetical stylistics. The apostrophes are used in the first and

[2] http://storyofmylifetheblog.blogspot.com.es/

last lines of the poem, giving it a closed form. This effect was accidental, however it could be an interesting improvement to order experts in this way. The emotional state is expressed only by the adjective *excusable* as numerous adjectives dominated the emotional words.

Angry poem about the end

Inspired by the text:

> *I remember being endlessly entertained by the adventures of my toys! Some days they died repeated, violent deaths, other days they traveled to space or discussed my swim lessons and how I absolutely should be allowed in the deep end of the pool, especially since I was such a talented doggy-paddler.* [3]

Topic: *Deep end*, Emotion: *anger*
Poem:

I knew the undisrupted end
I was like the various end
As deep as a transformation
O end the left extremity
Objective undisrupted end
I hated the choleric end
O end the dead extremity

Remarks: The emotional state for the poem is *anger*, which may correspond to some negative expressions in the text (*died, violent deaths, deep end*). The mood is expressed in the poem by words *choleric* and *hated*.

Fearful poem about the way

Inspired by the text:

> *Lately everyone has been wondering "Is Jenelle and Gary going to get back together?!". NO! He is living his life and I'm living mine. We are both happy with our lives the way they are at the moment, I know for me at least I'm EXTREMELY happy. Gary might of been tweeting things because he might of been jealous in a way that I was dating Courtland but he agrees to stop today.* [1]

Topic: *Way*, Emotion: *fear*
Poem:

O mode the symbolic property
Quickest moderate way
She was like the mode
She seemed hysterical because the way left

Remarks: We can observe here that the system does not do well with ambiguous words. The *way* is once interpreted as *property* or *mode* but the algorithm does not consider what was the phrase context in the text. However, the poetry may allow some less strict interpretations of meaning as the ambiguity can be used as an intentional poetical operation.

[3] `http://hyperboleandahalf.blogspot.com`

Evaluation

The evaluation of a creative system is a difficult and ill-defined problem. As the goal is not only to generate a satisfying output but also to imitate the creation process, the evaluation needs to consider both the aspects.

Output evaluation

As the human interpretation of poetical artifacts is a subjective process, we claim that the *Turing tests* are not reliable ways to evaluate poetry. However, the system requires some kind of evaluation for its outputs. Hence, according to (Manurung, Ritchie, and Thompson 2012) we assume that generated texts need to meet the constraints of *grammaticality, meaningfulness and poeticness* to be considered as valuable poetic artifacts. Below we evaluate our outputs along these dimensions.

Grammar The consistency of grammatical form is controlled by the constraints on person and tense. Use of *Context-Free Grammars* as the knowledge for poem-making experts provides the poem with a proper grammatical structure. As we can observe in *Examples* section, the outputs generally represent proper grammar. Some minor mistakes are caused by mis-classification of ambiguous words. This problem could be solved by improving the text-analyzing phase so that the key phrases are analyzed considering the context in which they are used.

Meaning The meaning of the poem is derived from the lexical (WordNet) and statistical (Brown Corpus analysis) associations of words in the topic phrase. Poems contain synonyms, hypernyms and antonyms as well as words that are most commonly used together with the main phrase. This combination results in a higher diversification of produced poems. The choice of the topic as the most *inspiring* phrase causes more possibilities to produce varied and meaningful poems. Also, the use of phrases describing the emotional state gives the impression of intentionality in produced compositions.

However, as observed in the last example in *Examples* section, the algorithm lacks handling of ambiguous phrases. Thus. the interpretation may differ from the meaning of the phrase in the initial text, and may not be consistent throughout the poem. This problem could be resolved by analyzing the context of words in the text but, as mentioned above, for poetry the ambiguity may sometimes be perceived as an intentional operation.

Poeticness The poetic form of generated poems is created by two main factors – the experts using poetical forms for their phrases and the stylistic constraints for lines. As can be observed in presented outputs, the poetical forms used by experts, such as epithets and apostrophes, make an important contribution to the overall perception of poetical composition.

The stylistic constraints in the current version consider only the number of syllables for each line, and are used for selecting best candidates for lines. This approach does not allow more elaborated poetical operations, such as the use of rhymes or rhythm. However, this could be easily improved

by adding new *selection experts* to the blackboard architecture. Each expert should use some heuristics to evaluate the competing phrases and the final selection should respect all criteria.

Another important aspect influencing the poetical character of outputs is the use of emotionally rich words that evoke imagery and are typical for poetical expressions.

Output evaluation summary

As presented above, the products of the system meet the triple constraints on *grammar, meaning and poeticness* to some extent. Further improvements of these factors in the system should include context-based analysis of words and introducing more stylistic constraints for the poetical form.

Model evaluation

As the main focus of computational creativity systems is to produce their outputs in an intentional way, the generation process should consider this as an important concern for evaluation. We propose evaluation of our system using the FACE model (Colton, Charnley, and Pease 2011) which is aimed at evaluating *creative acts* performed by a computer. The details of the model are presented in the *Background* section. We present below how our system architecture corresponds to these criteria.

Concept and concept expression In our case, the *concept* is the blackboard architecture with the set of experts cooperating to compose the poem. The motivation to use the blackboard architecture as presented in *Background* section is the Global Workspace theory which compares the brain functioning to a group of independent modules sharing a public workspace. The program takes a text as an input and produces the *concept expressions* in the form of poems. The outputs are evaluated in the *Output evaluation* subsection. In this approach we could also consider each expert as an independent *concept* producing its own *expressions* as partial solutions for the problem.

Aesthetic measure The *aesthetic measure* in the system may be considered as the heuristic functions evaluating candidates for new lines in poem. Each pair expert (*concept*) – phrase (*expression*) is evaluated respecting the stylistic constraints (6) and the expert's frequency before. The result is a real number. Another measure is used for topic selection — each key phrase is evaluated according to its *inspiration* value as in (5).

Framing information The *framing information* in system might be found only in the name of the emotional state defined according to the model of emotions (4). This output provides some information about the context of the poem.

FACE evaluation summary

As presented above, the generation process performs the *generative acts* of the form $< A^g, C^g, E^g >$. The F^g is provided by description of the emotional state only, but it may not be sufficient to satisfy the *framing information* criterion.

Conclusions

We proposed a system that is capable of expressing its own feelings in the form of a poem. The emotional state is generated by empathic perception of the text, and the mood is modulated by the optimism rate factor given to the character.

The blackboard architecture used in the system provides an effective way to model creativity: it is easily extensible with new linguistic resources and stylistic constraint. It could even incorporate experts representing other existing poetry generation systems such as *Stereotrope* for generating metaphors. Moreover, the blackboard model is a computational representation of Global Workspace theory of mind, which makes it a promising tool for simulating cognitive processes.

The poems produced by the system generally satisfy the triple constraints of *grammar, meaningfulness and poeticness*. However, in the future work, more attention should be paid to the context of analyzed words. According to the FACE evaluation, our system performs the creative acts of the form $< A^g, C^g, E^g >$. The *aesthetics measure* could be improved by defining more stylistic constraints for the poem.

The approach presented here can also be applied for generating poetry based on blogs.

References

Baars, B. J. 1997. In the theater of consciousness. *Journal of Consciousness Studies* 4.

Baars, B. J. 2003. The global brainweb: An update on global workspace theory. *Science and Consciousness Review*.

Bradley, M. M., and Lang, P. J. 2010. Affective norms for english words(anew): Affective ratings of words and instruction manual. Technical report, NIMH Center for the Study of Emotion and Attention University of Florida.

Cambria, E.; Livingstone, A.; and Hussain, A. 2012. *The Hourglass of Emotions*. Springer.

Colton, S.; Charnley, J.; and Pease, A. 2011. Computational creativity theory: The face and idea descriptive models. In *Proceedings of the International Conference on Computational Creativity*.

Colton, S.; Goodwin, J.; and Veale, T. 2012. Full-face poetry generation. In *Proceedings of the International Conference on Computational Creativity*.

Corkill, D. D. 1991. Blackboard systems. *AI Expert* 6.

Davis, M. 1983. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology* 44(1):113–126.

Gervas, P. 2001. An expert system for the composition of formal spanish poetry. *Knowledge-Based Systems*.

Gervas, P. 2010. Engineering linguistic creativity: Bird flight and jet planes. In *Proceedings of Workshop on Computational Approachesto Linguistic Creativity*.

Kirke, A., and Miranda, E. 2013. Emotional and multi-agent systems in computer-aided writing and poetry. In *Proceedings of the Artificial Intelligence and Poetry Symposium*.

Kurzweil, R. 1992. *A (Kind of) Turing Test*. The MIT Press.

Manurung, R.; Ritchie, G.; and Thompson, H. 2012. Using genetic algorithms to create meaningful poetic text. *Journal of Experimental and Theoretical Artificial Intelligence*.

Miller, G. A. 1995. Wordnet: A lexical database for english. *Communications of the ACM*.

Plutchik, R. 2001. The nature of emotions. *American Scientist* 89(4):344–350.

Russel, J. A. 1980. A circumplex model of affect. *Journal of personality and social psychology*.

Small, S. 1979. Word expert parsing. In *Proceedings of the 17th annual meeting on Association for Computational Linguistics*.

Strapparava, C., and Valitutti, A. 2004. Wordnet-affect: an affective extension of wordnet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation.*

Thelwall, M.; Buckley, K.; Paltoglou, G.; and Cai, D. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*.

Toivanen, J. M.; Toivonen, H.; Valitutti, A.; and Gross, O. 2012. Corpus-based generation of content and form in poetry. In *Proceedings of the International Conference on Computational Creativity*.

Toivanen, J. M.; Jarvisalo, M.; and Toivonen, H. 2013. Harnessing constraint programming for poetry composition. In *Proceedings of the International Conference on Computational Creativity*.

Turing, A. 1950. Computing machinery and intelligence. *Mind*.

van der Heide, A., and Trivino, G. 2010. Simulating emotional personality in human computer interfaces. In *Proceedings of IEEE International Conference on the Fuzzy Systems (FUZZ).*

Veale, T. 2013. Less rhyme, more reason: Knowledge-based poetry generation with feeling, insight and wit. In *Proceedings of the International Conference on Computational Creativity*.

Whissel, C. 1989. The dictionary of affect in language. *Emotion: Theory, Research, and Experience*.