Incremental Alignment of Metaphoric Language Model for Poetry Composition

Marilena Oita

The Swiss AI Lab IDSIA, USI, SUPSI
marilena@idsia.ch

Abstract. The ability to automatically generate meaningful text with respect to a topic is an important AI mission. In particular, the automatic generation of content which is deemed creative is a great challenge. In this paper, poetry generation is approached through the lenses of a new architectural design for creativity, that leverages semantics for the creation of variance and the preservation of the content coherency throughout the generation process. The full implemented system is made available on github [1].

Keywords: creativity, automatic poetry generation, metaphor identification, neural language model, semantic alignment

1 Introduction

Understood as the generation of novelty that is goal-appropiate, creativity is the most valued of human abilities [1]. Modelled in artificial agents, creativity or *useful imagination* [1] is expected to bridge the gap to generic AI [31], with considerable implications towards speeding up the creation of value.

The composition of poetry is a unique artifact of human language, in itself a compelling manifestation of creativity.

Similar to science, poetry decodes the reality, but in a more unconventional, subjective way.

One demands two things of a poem. Firstly, it must be a well-made verbal object that does honor to the language in which it is written. Secondly, it must say something significant about a reality common to us all, but perceived from a unique perspective. —W. H. Auden

Expressing the perceptions of an unexpressed (encoded) internalized reality, poetry decodes it by acting "as a shortcut to the truth" [2]. By expressing in new ways the notions of reality [8,32], ie. making explicit a novel relation between observations of reality [32], poetry generation is a core creative process that can be reverse-engineered using the INNGenuity framework [21].

[1] https://github.com/marilenaoita/poetry-composition-resources
[2] Sarah Howe, https://www.newscientist.com/article/2073697-verse-in-the-universe-the-scientific-power-of-poetry/

Building autonomous creative systems that automatically generate poetry which is both meaningful and original (genuinely creative) is a huge challenge in AI. Besides the practical interest in augmenting the human expertise in poetry composition, by speeding up the process and widening its reach, a system which generates meaningful and original poetry is however most interesting for the creative breakthrough artificial agents would demonstrate.

Machine learning models, by nature, are only retaining, by heart, the patterns of the input corpus. Since creativity needs understanding [7], pure learning of sequences does not produce something truly new, i.e., *different*.

Current state-of-the-art deep learning models for natural language generation have perfected the mechanisms of generating next-best-sequence starting from a seed text. Despite that, the results are far from being recognized by humans as coherent, and even less as creative. Despite mimicking the patterns of expression of the emulated author or style, the generated text either does not have coherency and meaning, either obviously reproduces too much of the original works. More than putting a word or character after another, poems develop ideas that unfold incrementally as the discourse evolves.

An originality of this work is the use of neural networks in alignment with semantics. Semantics encourages the creation of what we perceive as imaginative by activating the creativity conditions [27]: 1. the creation of variation (novelty), and 2. the assessment of meaning (coherency).

The contributions of this paper are:

1. identifying by using NLP techniques metaphoric expressions and exploiting them with the goal of poetic imagination modelling;
2. training a figurative language model to predict the next best metaphor given a seed text sequence;
3. employing Universal Sentence Encoder (uSE) [3] in an incremental way on the output of the language model to semantically align 'imagined' sequence to verse; to this end, an index with more than 3 million uSE embeddings of the Gutenberg poetry corpus [4] has been built.

This work continues as follows: Section 2 outlines related work to automatic poetry generation. Further in Section 3, the collaboration between computational linguistic, semantics and deep learning is described. The system's architecture and pipeline output results are illustrated in Section 4. Section 5 wraps up with discussion, conclusions and future work.

2 Related Work

Automaticaly identifying and generating metaphoric constructs is an important step in the process of poetry composition. Metaphor models have been explored in [26], but from a pure computational linguistics perspective. A semantic approach to metaphor interpretation and generation is described in [30]. In this

[3] Google module, tensorflow_hub, 2018
[4] https://github.com/aparrish/gutenberg-poetry-corpus

paper, the metaphor model is in contrast a figurative language model trained with a deep neural network.

By using Gutenberg poetry corpus, the work presented here classifies as corpus-based poetry-generation. A template method for this category is described in [5].

NLP techniques and semantics have been already used for poetry generation, but not in collaboration with neural networks.

Neural networks trained at word and/or character level to learn style and rhythm [33], model the task of poetry composition as text generation using sequence models [17]. But poetry is anything but plain text generation [19]. More recent works on natural language generation do optimize their architecture by integrating semantics [28], although in a different manner than in this paper.

Alternatives to neural networks are represented by either evolutionary algorithms [15] or Hidden Markov Models, employing a list hard-coded rules modelled as a set of constraints on meter, word similarity and rhythm [9]. Meaningfulness, grammaticality, and 'poeticness' are properties of a goal state, as poetry generation is reduced to a state space search problem in [16].

Not surprisingly, semantics has been pervasively employed when the focus of the content generation task was less on rhyme, and more on poem coherency, which is declared to enhance "feeling, insight and wit' [29]. Semantic features of poems, like type token ratio, concrete or abstract object tokens, etc. have been emphasized as characteristic of good poems [11]. Semantics integrated from Wordnet [2] can be useful by exploiting the relations between concepts present in the seed text. A work which integrates semantics as central in their technique is [22]. PoeTryMe [5] builds verses with spans of text that contain the seed tokens involved in a semantic relation. The method presented in this paper uses the semantics once in an indirect way through the uSE embedding scheme, and another through the use of Wordnet synsets for the metaphoric expressions identification.

Other related works on text-to-poem, e.g., plot-to-poem [6], or on meaning transfer using alignment [7], consider solely the direct correspondence of the initial text to the target (poetic verse) by means of concepts *similarity*, but in this ”translation” no creativity is involved. The reason is first, there is no creation of variation: there will always be one single correspondence from the first text structure to the target one, and second: in this linear transformation no overall creative goal is enforced (global coherency of the generated content). These two shortcomings have been addressed in this paper. Preventing "mindlessness" [13] by creating (in our case, *linguistic*) variation and enforcing semantic coherency have been achieved with subtle architectural changes, but which lead to a non-linear impact on results.

[5] http://poetryme.dei.uc.pt/
[6] http://static.decontextualize.com/plot-to-poem.html
[7] http://yknzhu.wixsite.com/mbweb

3 Methodology

The present work follows INNGenuity architecture [21] for creativity by modelling in the context of poetry composition the following hubs: imagination, executive and salience (attention to goal). Having as input a seed text (e.g. phrase, story, poem, etc.), the goal is to create a corresponding poetic structure that preserves the intention communicated by the original seed text, but expresses it in new ways.

3.1 Main Idea

Figure 1 outlines an end-to-end method for poetry composition that is fully unsupervised.

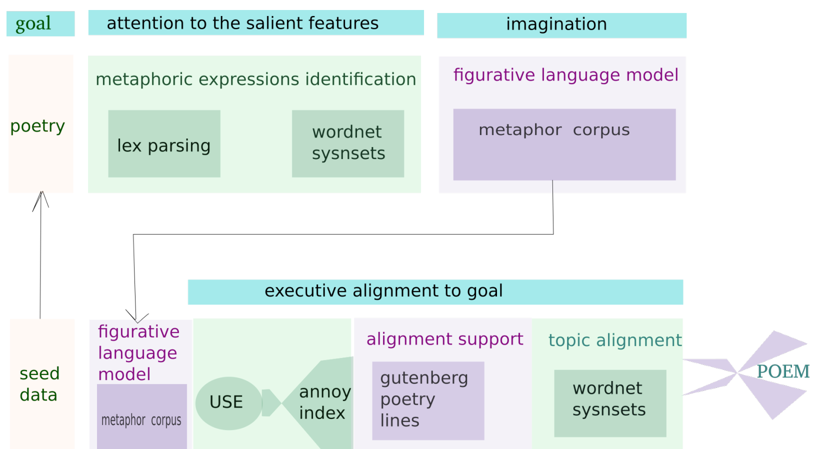

Fig. 1: The proposed pipeline for poetry composition

Typical deep learning approaches to poetry generation train a language model on the poetry corpus itself [8], which results in a close-to-zero-creativity text generator.

The strategy implemented in this paper is built on the following reasoning: the poet uses *imagination* for encoding the message of the poem, in the form of metaphors: succint, semantically-rich, unexpected associations of concepts. Therefore, what we need is to model is 'imagination' as a *metaphor generator*, implemented using a generative language model. The imagined output is next put in context by using semantic alignment, which also takes care of the verse

[8] http://karpathy.github.io/2015/05/21/rnn-effectiveness/

structure. Similar to the human process of creation, this latter step models the executive part of our brain that puts an idea in its most rightful form given the abilities (here, the expressivity of the semantic model) and the resources (the reach of available expression means).

This paper proposes therefore modelling 'imagination' as a language model trained on (possibly) multi-token expressions identified from a metaphor corpus, by leveraging lexical and semantic features. Imagined outputs will be fuzzy and imperfect, typical to what a language model can currently achieve, but inline with the nature of what imagination typically generates.

Culture or 'experience', inherently semantic, comes to complement and correct the imagination. Its presence is needed as a way to aprehend the universe to which poetry refers to, and distills "truth" from, in the form of poetic constructs. In this paper, knowledge is modelled inherently using the Universal Sentence Embedding (uSE) model [4], which has been trained by Google on very large generic datasets like Wikipedia, web news, web question-answer pages and discussion forums. Inherent semantics helps achieving, through alignment, the best mapping of an idea to a poetic 'universe' denoting the existing manners of expression (here, the Gutenberg lines of poetry).

3.2 Metaphor Expressions Parsing

Language at its most distilled, *metaphors* are essential to poetry [10]. For their identification in text, the current approach considers parsing the text in a specific way.

Word associations are an important element of linguistic creativity, because of their direct link to metaphoric constructs [12]. In particular, lexical associations have been found to play an important role in poetic text [20].

Multi-token structures, e.g., *tender_light*, *stormy_flames*, are particularly important for the meaning preservation, an important goal in this paper.

In contrast with other poem generating techniques that simply parse the text token by token, here the text is parsed using a lexical parser first. Candidate constructs to metaphoric expressions are obtained by serializing the tokens linked by the `compound` or `relcl` relations, as identified by a dependency parser (with additional rules , aka POS tag). Using simple rules on the part-of-speech tag of tokens composing the candidates, non-metaphoric structures like *one_shade* or *these_lilies* can be pruned.

The advantages of segmenting the text at (possibly) multi-token are in this context the following: I) it allows the training of a more appropiate model, since our goal is to generate metaphors which are multi-token expressions; II) it prevents the loss of meaning: empyrically, a language model loses sight of meaning very quickly when trained at smaller granularity, i.e. word or character level); III) it optimizes the training since the sequence window has to be much smaller for the model to learn useful patterns, while capturing more useful dependencies.

This specific parsing also allows the assessment whether a text sequence may contain metaphors or not at all. This can be important in the final step, where more metaphoric constructs can be given more attention, and verses aligned semantically but containing no metaphors could be ignored. The metaphoric potential of a sequence text is computed based on the existence and the quality of metaphoric expressions. The parameters for quality are whether the expression contains highly ambiguous and semantically-rich token components. Ambiguity and rich semantic properties [11] (e.g., number of attributes, relations to other concepts in an ontology) are both qualities derived using Wordnet [18]. Ambiguity is measured based on the number of synsets a token has in Wordnet. For the semantic interpretation, we need to fix the joint meaning of a candidate metaphor. For this, the associated Wordnet synset definition sets to each component token are extracted. Each (typically,) pair of synset definitions is next considered. Pair definitions are embedded using the uSE model, and best joint interpretation is chosen to be the pair closest in terms of semantic coherency.

3.3 A Figurative Language Model

The process of creation of metaphors has the merit of engaging both cognition and imagination [24]. A metaphor corpus [9] containing verses from Poetry Foundation, known to contain metaphors, has been selected for building the figurative language model. Other sources [10] of metaphors could be added to increase the expressivity and robustness of the model.

The training is done on metaphoric expressions parsed as described above. A neural network consisting of 4 stacked bidirectional LSTMs, with a dropout of 0.4 at each layer, and the hidden layers of 1024 dimensions has been considered. A sequence length of 8 is chosen, with a learning rate of 0.0001, and all is trained over 60 epochs. The considered vocabulary size of metaphor lexical expressions is 1624. This is not much, but since the training is done at expression level and not at word or character level, it is sufficient to create variation.

The general shortcoming of training at expression level is that the generated expressions will not be linked by typical stop words, etc. Here, the priority is given to the preservation of meaning, rather than the form. Additionally, the generated sequences will be easily 'corrected' through alignment in a next step.

3.4 Semantic Alignment of Imagined Text to Verse

In contrast with other works, the semantic similarity considered here does not work at token level, but at phrase level instead. Universal Sentence Encoder model, which has been trained with a deep averaging network, encodes given text into a 512-dimension vector. In addition to carrying semantics from external generic sources, uSE allows the transformation of multiple verses into a single "resuming" sentence, an appealing quality which is next leveraged in the incremental alignment.

---

[9] https://github.com/marilenaoita/nips-poetry-composition-resources/blob/master/metaphor_annotated.txt

[10] http://ota.ahds.ac.uk/headers/2541.xml

The alignment proceeds between the 'imagined' sequence generated at each step by the figurative language model and the Gutenberg poetry corpus. Note that the corpus on which the language model is trained is not the same as the corpus of alignment: this ensures the creation of meaningful variation: each time we run the process, we may obtain different results given a appropiate *temperature* parameter value to the generative model. For efficiency reasons, an annoy [11] index of dimension 512, and built with 20 trees, is storing the 3,085.117 uSE embeddings corresponding to the Gutenberg poetry lines.

The transformation from the intial seed text to a poem is done *incrementally* as follows. The seed text is segmented into sentences. A first seed sentence (or the title) is given to the metaphoric language model. The predicted next sequence of the model is aligned to the closest verse from the gutenberg corpus by means of vector cosine similarity. Starting with the next seed sentence, at each step, the topk alignment candidates to the imagination output will be ranked based on semantic coherency with the poem lines generated so far, and best chosen. That means that when searching for the best alignment there exists a retroactive check for balancing the variation (given by the imagination output), and the content consistency with respect to what already exists as poem draft. This contributes at ensuring, at the best of the uSE model capabilities, that a narative flow is present.

Finally, structural constraints can be integrated by means of mimicking the style of expression of the initial poem, using explicit prononciation hints and the syllables number in verses from cmudict [12]. There exists two options: either when extracting the alignment candidates we select the one which mostly respects these constraints, either we replace the ending tokens of the verses such that the rhythm is present (using an index of semantically similar tokens that exhibit the desired structure for instance). Since the primary concern in this paper is not the rhyme, this more straightforward part has not been currently addressed.

4 Experiments and Results

For illustrating the resulting creations of this pipeline, the seed text has been chosen to be either short stories (e.g. about the universe [13]), other poems, or the interpretation of the poems directly [14].

4.1 "The Big Bang"

short story about the beginning of the universe [15]

[11] https://github.com/spotify/annoy
[12] http://www.speech.cs.cmu.edu/cgi-bin/cmudict
[13] https://www.esa.int/esaKIDSen/StoryoftheUniverse.html
[14] http://classicalpoets.org/10-greatest-poems-ever-written/
[15] https://www.esa.int/esaKIDSen/SEMSZ5WJD1E_OurUniverse_0.html

Aligned to Gutenberg (current pipeline) :
"Slow-syllabled of weed and bloom.
But time and earth case-harden us to live;
The Serpent sleeping, in whose mazie foulds
The clearest stream through Phrygias land which flows.
To print our poems, the propulsive cause,
Where sunset spreads serenest."

We can observe here creative transformation of the original, rather formal content, to a poem which has a flow and conveys the "big bang" idea and semantic threads about the genesis of the universe.

Aligned to the metaphor corpus itself (experiment) [16].

Unfathomable sea ! whose waves are years,
while what lurks below the surface is another story,
the ocean was salt before we crawled to tears.
the western wave was all a-flame
earth is a door i cannot even face
the dead man is the flywheel of the spinning planet
the breath of the moist earth is light
over why why. causation is sequence
hath guest fire-fledgd as thine.. whose lord is love ?

4.2 "Ozymandia"

by Percy Bysshe Shelley.

Meaning of the poem : the poem's interpretation as extracted from [17]. "In this winding story within a story within a poem, Shelley paints for us the image of the ruins of a statue of ancient Egyptian king Ozymandias, who is today commonly known as Ramesses II. This king is still regarded as the greatest and most powerful Egyptian pharaoh. Yet, all that's left of the statue are his legs, which tell us it was huge and impressive; the shattered head and snarling face, which tell us how tyrannical he was; and his inscribed quote hailing the magnificent structures that he built and that have been reduced to dust, which tells us they might not have been quite as magnificent as Ozymandias imagined. The image of a dictator-like king whose kingdom is no more creates a palpable irony. But, beyond that there is a perennial lesson about the inescapable and destructive forces of time, history, and nature. Success, fame, power, money,

[16] https://github.com/marilenaoita/nips-poetry-composition-resources/blob/master/metaphor_annotated.txt
[17] http://classicalpoets.org/10-greatest-poems-ever-written/

health, and prosperity can only last so long before fading into "lone and level sands." "

Generated poem from the interpretation of the original poem directly:
"To the green doublet; bitter is the wind, as though it blew
The steepy rock, and frantic tide,
Winding and vague was the family road–
The fires are dead, the gold is stone."

4.3 Evaluation of Generated Poetry

The task of poetry generation is challenging, but even more is evaluating its results in terms of the creativity involved [18]. Current means to properly evaluate automatically generated content are very limited [23].

An aesthetic measure introduced by [3] formalizes a poem's beauty solely based on phonemic features, therefore this measure cannot be used for semantic coherency since it fails to quantify the meaning of poetic texts.

Common in language generation tasks, the perplexity [34] measure is used to evaluate language models. Capturing the likelihood of correctly predicting the next character/word/expression in a sequence, perplexity can however only measure the learning capabilities of a model. Indeed, the pipeline proposed in this paper does contain multiple modules, not only a language model, therefore the usage of perplexity to evaluate the generated poems is not suitable. To study the different results that are obtained using a typical language model, a language model has been trained on the Gutenberg corpus directly, with the Adam optimizer, and the following parameters: hidden_units of 128, batch_size of 512, num_epochs of 300, using an Embedding layer, two stacked bidirectional LSTMs, and dropouts in between. The training of this alternative for poetry generation, takes several days to train on the entire Gutenberg corpus. Nevertheless, even with this state-of-the-art model, the results mostly reproduce the context of the original poem. That is, the model correctly predicts the next possible sequence of words, given the seed, but without much variation from the original poem's content. This is not surprising, since a language model trained with deep networks is the result of a learning process which is not creative per-se.

More semantic-oriented measures would be needed to capture the subtlety of the imagination or richness of meaning. For this, empyrical tests involving human interpretation are still the most trusted [6].

Besides the illustrated poem generated by the method, chosen short for practical reasons, more examples are given at [19]. The code and other ressources (alignment index and models) are available on github for further, direct experiments.

---

[18] http://www.thepaintingfool.com/papers/pease_aisb11.pdf
[19] https://github.com/marilenaoita/nips-poetry-composition-resources

5 Discussion and Further Work

This work demonstrates the potential of the interaction between a figurative language model and semantics, in order to enable meaningful poetry composition. Towards this goal, all main dimensions of text: linguistic, semantic and statistical have been leveraged in collaboration, in a fully unsupervised manner.

Although the automatic evaluation techniques are still in development for assessing the quality of generated content (be it video, image, or text), empyrical tests attest that poems generated with this pipeline show a preservation the global meaning, while developping threads of the ideas of the original poem. The resulting verses are also being expressed creatively, and not mimicking at all the original content, as it is the case for typical language models.

Further work includes the usage of a joint embedding with visual features on the metaphor structure representation. A multimodal [14] representation can enrich the semantic properties of the concepts involved in metaphors, and further improve the quality of the poetic composition. Illustrating the visual images the poems are creating by means of their metaphoric language is another promising direction that can be pursued by adapting the ideas presented in ChatPainter [25], in order to create *PoemPainter*. Increasing the quality of generated metaphoric expressions by incorporating semantic relations in the learning process is another line of improvement.

Creative text generation needs an architecture that supports the integration of semantics to neural networks in a dynamic and flexible way [21]. The technical means towards creativity are being forged, and there exists a global effort towards the integration of meaning, as a way towards automatic machine *understanding*, a more flexible and interpretable 'generator' as opposed to the shallower and more rigid machine *learning* 'translator'.

References

1. R. M. A. and J. G. J. The standard definition of creativity. *Creativity Research Journal*, 2012.
2. M. Agirrezabal, B. Arrieta, A. Astigarraga, and M. Hulden. Pos-tag based poetry generation with wordnet. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 162–166, 2013.
3. G. D. Birkhoff. *Aesthetic measure*, volume 38. Harvard University Press Cambridge, 1933.
4. D. Cer, Y. Yang, S. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, Y. Sung, B. Strope, and R. Kurzweil. Universal sentence encoder. *CoRR*, abs/1803.11175, 2018.
5. S. Colton, J. Goodwin, and T. Veale. Full-face poetry generation. In *ICCC*, pages 95–102, 2012.
6. S. Colton, A. Pease, J. Corneli, M. Cook, and T. Llano. Assessing progress in building autonomously creative systems. In *ICCC*, pages 137–145, 2014.
7. T. Dartnall. *Artificial intelligence and creativity: An interdisciplinary approach*, volume 17. Springer Science & Business Media, 2013.
8. R. W. Gibbs et al. The poetics of mind, 1994.

9. J. Hopkins and D. Kiela. Automatically generating rhythmic verse with neural networks. In *Association for Computational Linguistics*. Association for Computational Linguistics, 2017.
10. B. Indurkhya. Creativity in interpreting poetic metaphors. *New directions in metaphor research*, pages 483–501, 2007.
11. J. Kao and D. Jurafsky. A computational analysis of style, affect, and imagery in contemporary poetry. In *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, pages 8–17, 2012.
12. S. Krishnakumaran and X. Zhu. Hunting elusive metaphors using lexical resources. In *Proceedings of the Workshop on Computational Approaches to Figurative Language*, FigLanguages '07. Association for Computational Linguistics, 2007.
13. E. J. Langer and A. I. Piper. The prevention of mindlessness. *Journal of Personality and Social Psychology*, 53(2):280, 1987.
14. D. Liu, Q. Guo, W. Li, and J. Lv. A multi-modal chinese poetry generation model. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.
15. H. Manurung. An evolutionary algorithm approach to poetry generation. 2004.
16. H. Manurung. An evolutionary algorithm approach to poetry generation. 2004.
17. T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, and S. Khudanpur. Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*, 2010.
18. G. A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
19. J. Moss. *What is imitative poetry and why is it bad?* na, 2007.
20. Y. Netzer, D. Gabay, Y. Goldberg, and M. Elhadad. Gaiku: Generating haiku with word associations norms. In *Proceedings of the Workshop on Computational Approaches to Linguistic Creativity*, CALC '09. Association for Computational Linguistics, 2009.
21. M. Oita. Reverse engineering creativity into interpretable neural networks. In *Proceedings of the Future of Information and Communication Conference*. FICC, 2019.
22. H. G. Oliveira. Poetryme: a versatile platform for poetry generation. *Computational Creativity, Concept Invention, and General Intelligence*, 1:21, 2012.
23. H. G. Oliveira. A survey on intelligent poetry generation: Languages, features, techniques, reutilisation and evaluation. In *Proceedings of the 10th International Conference on Natural Language Generation*, pages 11–20, 2017.
24. P. Ricoeur. The metaphorical process as cognition, imagination, and feeling. *Critical inquiry*, 5(1):143–159, 1978.
25. S. Sharma, D. Suhubdy, V. Michalski, S. E. Kahou, and Y. Bengio. Chatpainter: Improving text to image generation using dialogue. *CoRR*, abs/1802.08216, 2018.
26. E. Shutova. Models of metaphor in nlp. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10. Association for Computational Linguistics, 2010.
27. C. D. T. Blind variation and selective retention in creative thought as in other knowledge processes. *Psychol. Rev. 67*, 1960.
28. V. Tran and L. Nguyen. Neural-based natural language generation in dialogue using RNN encoder-decoder with semantic aggregation. *CoRR*, abs/1706.06714, 2017.
29. T. Veale. Less rhyme, more reason: Knowledge-based poetry generation with feeling, insight and wit. In *ICCC*, pages 152–159, 2013.

30. T. Veale and Y. Hao. A fluid knowledge representation for understanding and generating creative metaphors. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 945–952. Association for Computational Linguistics, 2008.
31. R. W. Weisberg. The creative mind versus the creative computer. *Behavioral and Brain Sciences*, 17(3):555–557, 1994.
32. P. Willis. Don't call it poetry. *Indo-Pacific Journal of Phenomenology*, 2(1), 2002.
33. S. C. Xie. Deep poetry : Word-level and character-level language models for shakespearean sonnet generation. 2017.
34. R. Yan. i, poet: Automatic poetry composition through recurrent neural networks with iterative polishing schema. In *IJCAI*, pages 2238–2244, 2016.