

## Poetry beyond Poetry: Applying GraphPoem Outcomes in DH, NLP, and Performance

Chris Tanasescu (MARGENTO), Université Catholique Louvain (Belgium)

Diana Inkpen, University of Ottawa (Canada)

The *Graph Poem* project has been awarded an Insight Development Grant in 2014 by the Social Sciences and Humanities Research Council of Canada to digitally define, compute, represent, expand, and evaluate graphs—networks of nodes and edges, where the nodes are the poems, and the edges, links between the poems defined according to genre-based criteria—as computational tools to (re)organize North American (and in the longer run, world) poetry, and to discover new relevant commonalities and paradigms among various poems and poetry corpora.

In these years we have developed a number of algorithms and natural language processing (NLP) tools for poetry analysis, including a topic classifier, a rhyme classifier, a meter parser and classifier, a diction analysis tool, statistical and deep-learning metaphor detection tools, etc. Although further development is needed towards holistically analyzing poetry within our project, as far as these classifiers are concerned, our research is in its final stage.

All of these results and implements have been used for quantifying commonalities between poems and representing corpora as network graphs. By analyzing the specific features of those graphs we are able to draw pertinent conclusions regarding the corpora they represent and single out poems of remarkable status or consequential positioning. Identifying cut-nodes in the network graph, for instance, is synonymous to identifying poems that, if removed, would render the network—and consequently the corpus the latter represents—disconnected.

One of our main contentions in this paper is that some of the results obtained in developing poetry tools prove not only applicable in more general contexts, but are also better than the tools trained on non-poetry data and perform better than the best general-use affordances for both poetry and non-poetry tasks.

We trained our own word embeddings on the Poetry Foundation (PoFo) poems and named the resulting model the Graph-Poem model. The options in terms of available generally used embeddings were GloVe (Pennington et al 2014) and word2vec (Mikolov et al 2013) vectors, yet we chose the former as they had been shown to work better for many lexical-semantic tasks. Still, these word embeddings of ours turned out to be better than the GloVe ones. They are moreover working better on any other type of data, which helps us make a point about the wider relevance of developing poetry-based NLP tools, and the further applicability of the latter in text analysis in general. Also, as detailed in (Kesarwani et al 2017) and in (Tanasescu et al 2018), our metaphor classifier turned out to work better on non-poetry data even if trained on poems.

One of our recent projects draws on the methods used in putting together a graph-poem anthology and deploys that in a wider text analysis/NLP, DH, and performance contexts (MARGENTO 2018). In the book we deployed Python in-built algorithms for representing the corpus as a graph, yet for the algorithmic expansion of the corpus we used our own diction

classifier (Kesarwani 2018), developed as part of the *Graph Poem* project. Namely, while analyzing the graph for centralities, we noticed that a number of poem-nodes ranked very low in closeness centrality but were better than most for betweenness centrality. In other words, while being very marginal in terms of distance to the vast majority of the network, these poems had a special status in helping other poems connect to each other within the corpus. The automated expansion of the corpus involved our diction classifier that helped track down potential new candidates for inclusion that, in their turn, after joining the network proved also to be such marginal connectors.

In a live computational performance event we will be presenting this year—“E-poetry event for DHSI 2019: *GraphPoem*” (June 2019, Victoria, Canada)—we deploy similar methods and algorithms only applied to texts of unrestricted genre in an interactive context in which the participants in the most important digital humanities summer institute in the world will have the opportunity to feed themselves texts—by the hundreds or more—to the real-time expanding network. The event and the applications will take the poetry-modeled methods and tools to a more generally relevant level while laying the foundation for a communal digital-humanities-based concurrence of transdisciplinary interests and topics.

While detailing the research and creative experiments that led us to currently extrapolating computational poetry analysis and generation to examining networked textualities in the digital medium, our paper investigates nonetheless the relevance of poetry to such developments and the qualities of the genre that have invited the abovementioned extrapolations.

There are quite a number of features that make the genre especially conducive—in a sort of epitomizing capacity—to peremptory operational instantiations and manifestations of the very architectures and processes typical of digital space and the digital medium. As argued elsewhere (Tanasescu 2016), poetry is in certain ways *the* digital space genre through its inherent performativity—a feature embedded in the form and not in its occasional public exposition only [a point for which we drew on established authors in cognitive poetics such as Reuven Tsur (Tsur 2012)]—and its consonance with digital space’s sonic economy [an argument based on a Stephen Kennedy’s “chaos media” and his appropriation of Gaston Bachelard’s notion of echoic poetry in describing digital space as the universe of “echostates” (Kennedy 2015)].

We are furthering that argument in this paper by looking closer into the opportunities offered by applying graph theory in poetry, and thus advancing an *asymptotically holistic* approach to the genre that turns out significantly beneficial to much wider concerns as well. Since our research undertakes to develop computational tools for analyzing (asymptotically) all poetic features, the growing connectivity of the resulting networks will (also asymptotically) exhaust the textual possibilities of the poem-nodes.

These networks both enact and build on the performativity of the digital medium through the said poem-nodes’ own in-built (genre-relevant) performative drive and foundational memetic expression. This potential completeness and asymptotic completion marks poetry in computational approaches and networked representations or assemblages as one of the virtually most valid models for inhabiting, charting, and developing the (post)digital.

## Bibliography:

- Ding, Cangfeng, and Li, Kan. 2018. "Centrality Ranking in Multiplex Networks Using Topologically Biased Random Walks." In *Neurocomputing* 312: 263–275.
- Kennedy, Stephen. 2015. *Chaos Media: A Sonic Economy of the Digital Space*. New York: Bloomsbury Academic.
- Kesarwani, Vaibhav. 2018. *Automatic Poetry Classification Using Natural Language Processing*. Ottawa: University of Ottawa. Web: <https://bit.ly/2ImNFUM>. Accessed October 01, 2018.
- Kesarwani, Vaibhav; Inkpen, Diana; Szpakowicz, Stan; Tanasescu, Chris. 2017. "Metaphor Detection in a Poetry Corpus." *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*: 1-9. Web: <http://bit.ly/2SVEnIE>. Accessed August 1, 2019.
- MARGENTO, ed. 2018. *US Poets, Foreign Poets*. Bucharest: Fractalia Press.
- Mikolov, Tomas; Yih, Wen-tau; Zweig, Geoffrey. 2013. "Linguistic Regularities in Continuous Space Word Representations". *HLT-NAACL*: 746–751.
- Pennington, Jeffrey; Socher, Richard; Manning, Christopher D. 2014. "GloVe: Global Vectors for Word Representation." Web: <https://nlp.stanford.edu/projects/glove/>. Accessed October 2, 2018.
- Tanasescu (MARGENTO), Chris. 2016. "Community as Commoning, (Dis)Placing, and (Trans)Versing. From Participatory and 'Strike Art' to the Postdigital." In *Dacoromania litteraria 3*: 10-44. Web: <http://bit.ly/2BQr6Eq> Accessed October 2, 2018.
- Tanasescu, Chris; Kesarwani, Vaibhav; Inkpen, Diana. 2018. "Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities." *AAAI Publications, The Thirty-First International Flairs Conference*. Web: <http://bit.ly/2sRi9V7>. Accessed Aug 1, 2019.
- Tsur, Reuven. 2012. *Poetic Rhythm Structure and Performance: An Empirical Study in Cognitive Poetics*. Revised and expanded edition. Brighton and Portland: Sussex Academic Press.