



Vol. 4.2 (2016)

ISSN 2182-8830

'Estudos Literários Digitais 2'

Manuel Portela e

António Rito Silva (orgs.)

An Interdisciplinary Model of Context Morphology

DÁNIEL LEVENTE PÁL

Eötvös Loránd University

BÁLINT DARÓCZY

*Institute of Computer Science and Control, Hungarian Academy of Sciences, and
Eötvös Loránd University*

Abstract

Using an abstract graph model we describe a hypothetical transformation of a literary entity through the active connections in the graph. First, we weakly define a set of transformations (“morphology”) over a particular entity as a series of the recent activities distinguishing them into categories according to their effect on the entity. Second, to our conjecture, the emerging paths of forms are slowly abandoning the original “birth context”, shaping a decreased, cleaned set of entities, to replace the gap with entities derived from the dynamic set of “recipient context”. **Keywords:** graph; literary entity; context morphology; morphological space-time dynamics; metatheory; interdisciplinarity.

Resumo

Usando um modelo abstrato de grafos, descreveremos aqui uma transformação hipotética de uma entidade literária através das conexões ativas no grafo. Primeiro, definiremos, em sentido fraco, um conjunto de transformações (“morfologia”) ao longo de uma entidade particular como uma série de atividades recentes distinguindo-as em categorias de acordo com os seus efeitos sobre a entidade. Depois, segundo a nossa conjectura, os caminhos emergentes das formas vão abandonando lentamente o “contexto de nascimento” original, moldando um conjunto limpo de entidades para substituírem a lacuna com entidades derivadas do conjunto dinâmico do “contexto de recepção”. **Palavras-chave:** grafo; entidade literária; morfologia contextual; dinâmica morfológica espaço-tempo; metateoria; interdisciplinaridade.

1. Introduction

As an introduction, we can articulate some general “philosophical” questions. As we know and understand, some of these questions are the most important and inspiring questions of literary theory. We also know that we can’t answer them, nor completely neither finally, nevertheless we are making an effort, and our answers may offer some new perspectives for inter- or multidisciplinary scientific thinking:

- (1) What is, or what could be the identity or identification of a literary work?
- (2) What are the (terminological) problems of inter- or multidisciplinary scientific thinking?

- (3) Is there an idea of a general (literary) theory? Is it possible to make a meta-theory of nearly all (literary) theories?
- (4) What could we gain if we resolve these three questions?

1.1. *The question of the identity of a literary artwork*

As Richard Shusterman wrote in his work *The Object of Literary Criticism*, one of the central questions for the philosophy of literary criticism is the problem of the identity of the work of literature and/or the various methods of identification of literary texts: “We have seen that it has bearing not only on the question of the literary work’s ontological status but also on the more practical critical questions of interpretation and evaluation. Unfortunately, it is as difficult as it is important, and involves a variety of puzzling questions.” (110) Nearly the same thoughts were written by René Wellek and Austin Warren in their famous book, *Theory of Literature*:

This raises an extremely difficult epistemological question, that of the ‘mode of existence’ or the ‘ontological situs’ of a literary work of art (which, for brevity’s sake, we shall call a ‘poem’ in what follows). (...) To the question what and where is a poem, or rather a literary work of art in general, several traditional answers have been given which must be criticized and eliminated before we can attempt an answer of our own. (141)

But if we get through these theories and ideas, and we can reply to this question correctly, an answer “must solve several critical problems and open a way to the proper analysis of a work of literature.” (Wellek and Warren, 1949: 141)

In this study we choose a different path, we reverse the problem. As we will see, neither the tradition of various interpretations, nor the specific individual interpreters¹ but the model itself—assuming some properties of our graph (Watts and Strogatz, 1998; Frank and Strauss, 1986; Daróczy *et al.*, 2015)—will select the necessary theoretical and other types of elements of the analytical and/or comparative interpretations, and finally it will identify a literary text, it will determine the identity, the “mode of existence”, and the “ontological situs” of a literary work of art.

¹ See, e. g.: Ingarden, 1931/1973; Wellek and Warren 1949: 139–158; Bilsky, 1953: 531–536; Bonati, 1960; Mitias, 1982: 41–52; Davies, 1996: 577–592 and 2007: 17–31.

1.2. Some problems of inter- or multidisciplinary scientific thinking

Writing about a text or a book or about any kind of literary artwork in a different field of interpretation—as a hypertext and as a possible problem of information theory, mathematics, or graph theory (etc.)—immediately produces terminological troubles and propositions in the metaphorical shadows of that other science’s different language, terminology and terminological tradition which is a cardinal and significant interdisciplinary gap among the branches of human and abstract sciences. Nevertheless, there are some cases when the scholars of different fields of science have almost the same theoretical problems, ask almost the same questions and answer them almost the same way, but in different paradigms—with totally different terminology and with a distinct and sundry way of logic which do not resemble each other.

George P. Landow wrote on it the most concise and laconic way in his recapitulative book, *Hypertext*, reissued in three successively expanded versions:

When designers of computer software examine the pages of *Glas* or *Of Grammatology*, they encounter a digitalized, hypertextual Derrida; and when literary theorists examine *Literary Machines*, they encounter a deconstructionist or poststructuralist Nelson. These shocks of recognition can occur because over the past several decades literary theory and computer hypertext, apparently unconnected areas of inquiry, have increasingly converged. Statements by theorists concerned with literature, like those by theorists concerned with computing, show a remarkable convergence. Working often, but not always, in ignorance of each other, writers in these areas offer evidence that provides us a way into the contemporary *episteme* in the midst of major changes. (1992/1997: 2; 2006: 1)²

To avoid these problematic points—and on the other hand for the reason of avoiding later terminological confusions—, we suggest that we promote and introduce a new interdisciplinary terminology as a theoretical and also a practical/technical introduction of our study: instead of using the words ‘text’, ‘book’, or ‘literary work’ (etc.), we “create” and use with a new kind of theoretical meaning the words: ‘content’ and ‘entity’.

In this creative process of overcoming the problem and the limitations of the previous scientific vocabularies and semantics, we would like to develop this kind of workaround—in a very similar way as Roland Barthes suggested in 1971:

²We are aware that this idea of Landow has been extensively criticized—and it has not been proved—nevertheless we don’t want to use it as a theoretical base of our conception, but we use it as an example to illustrate the difficulties, problems and terminological overlaps/confusions of inter- or multidisciplinary scientific thinking.

We might say, as a matter of fact, that interdisciplinary activity, today so highly valued in research, cannot be achieved by the simple confrontation of specialized branches of knowledge; the interdisciplinary is not a comfortable affair: it begins effectively (and not by the simple utterance of a pious hope) when the solidarity of the old disciplines breaks down—perhaps even violently, through the shocks of fashion—to the advantage of a new object, a new language, neither of which is precisely this discomfort of classification which permits diagnosing a certain mutation. (1986 [1971]: 56)

Using Barthes' logic—but with other sub-conclusions and final conclusion—, we can redefine more easily the identity and identification of a literary work which owes at least its phenomenal existence and its qualities and attributions to the former terminologies.

To clear that main confusion, we work with our own terminology and with some special terms ('content' and 'entity') which have at the same time some kind of heuristic value of usage and also an interdisciplinary resonance of meaning. Using this interdisciplinary method in this way, we can test our hypothesis on a significant sample of 20th and 21st century literary theory (including Roland Barthes, Jacques Derrida, and the various hermeneutic circle theories from Heidegger and Gadamer to the rule-based expert systems,³ to David Damrosch's and Joseph Tabbi's writings about world literature, to Franco Moretti's graph-based models of literature, to Aarseth and Eskelinen's cybertext poetics theory) to find and understand new connections, coherences, attributes, and qualities about how a (literary) content/entity works as a graph-based model, and ascertain what happened (with a philological methodology or with the theoretical experiences of reception theory), and what could or will happen with that (literary) content/entity in its own future.

³ As George P. Landow mixed the computational science and literary theory to create a possible hypertext-interpretation, Gerry Stahl in his book *Tacit and Explicit Understanding* gave a possible interdisciplinary use of Heidegger's and Gadamer's philosophical influence on the areas of design theory, Artificial Intelligence theory, and cognitive science—which is a highly interesting notion: "Heidegger and his student Gadamer revived that orientation to expound a general theory of human understanding and interpretation. Today, hermeneutics refers primarily to this philosophy of interpretation as fundamental to human existence, which Heidegger (1927) formulated and Gadamer (1960) further expounded. (...) To understand design, one must take into account the role of human interpretation. This means that a science of design—or, for instance, a theory of computer support of design—should be conceived on the model of the human sciences more than on that of the natural sciences. This is contrary to the traditional approach of AI attempts to automate design with rule-based expert systems, that look primarily to the mathematical sciences rather than the interpretive sciences for their model of scientific method. (...) In particular, the approaches of design theory, AI, and cognitive science that are important for this dissertation are philosophically close to Heidegger." (105–106)—See also the figures on his site: <http://gerrystahl.net/publications/dissertations/computer/d4.0.html>.

1.3. *How to make a meta-theory of (literary) theories*

At this point, we can ask: is there an idea of a general (literary) theory? Is it possible to make a meta-theory of (literary) theories? If we create a new terminology, one of its usages could be a new methodology for interdisciplinary approaches. But what is the benefit of that intellectual operation? What can we gain in human or literary sciences from that?

René Wellek and Austin Warren wrote in *Theory of Literature*:

While physics may see its highest triumphs in some general theory reducing to a formula electricity and heat, gravitation and light, no general law can be assumed to achieve the purpose of literary study: the more general, the more abstract and hence empty it will seem; the more the concrete object of the work of art will elude our grasp.” (6-7)

Humanist scholars—especially from the 19th century when the impact of natural sciences strengthened on the methodology of humanities—would like to make a general theory or general law of how the literature or literary text works, nevertheless “...attempts to find general laws in literature have always failed.” (Wellek and Warren, 1949: 6)

We think that using a similar logical methodology—as explained above in relation to the question of the identity of a literary artwork—to solve this “tiny” problem leads us to a kind of result that is worth considering. Ergo, we reverse the problem. We do not want to find a general literary theory above all the other literary theories, but we only replace the types of content entities in the nodes of our graph. In this way our investigation focuses not on the identity of a literary artwork (as an entity), but the “identity” of a theoretical work or text (as an entity).⁴

As we will see below, it is totally free which kind of entity is inserted in the central position of our graph-model. We can place a literary artwork, or a theoretical work, or every kind of textual content in the centre—even a newspaper article. And if our model works well, we will not get insights from the use of predetermined empirical or abstract rules—which is the common mistake of the general literary theories as we know—, but without any kind of preliminary concepts the model itself—assuming some properties of the graph—will select and situate the necessary theoretical reflections, and it will draw the importance and “social network” of that (literary) theory.

⁴ To our knowledge, our graph-based model is something new, particularly in the concept of modelling the whole graph of the literary/scientific works.

1.4. *The possible dimensions of our new terminology*

Detailing our terminology, we activate four different dimensions of our contextual or “context-morphological” analysis:⁵ (1) raw material = the (literary) text itself; (2) topics = para- and metatextual elements; (3) in-corpus entities = intertextuality, links, etc.; (4) time = “[The work of art] has something which can be called ‘life.’ It arises at a certain point of time, changes in the course of history, and may perish.” (Wellek and Warren, 1949: 156)

With this approximation, our main contribution is a model which tries to unite widely accepted and recognized theories over the birth and life of literary work into a joint abstract model with a possible explanation about why these sometimes contradictory theories have a *raison d’être*. In this paper we try to define the model as general as it can be. Meanwhile, we are acting on the border of traditional graph/network theory and the semantic analysis of raw (literary) content. While taking advantage of the recent results of both fields (Cortes and Vapnik, 1995; Watts and Strogatz, 1998; Caimo and Friel, 2011; Silkósi *et al.*, 2012; Garzó *et al.*, 2013; Cheng *et al.*, 2014; and, on the other hand, e. g.: Aarseth, Barthes, Derrida, Damrosch, etc.⁶), we can define some constraints to our graph according to traditional network, probability and graph theory (Barabási and Albert, 1999; Erdős and Rényi, 1959; Robins *et al.*, 2007; Frank and Strauss, 1986).

One of our main goals is to determine a capable set of inseparable underlying entities (such as sociological, technological, demographical, etc. elements) connected in any way to literary works (which is also an entity). Although even a carefully selected collection of entities surrounded by the birth of the literary work (“birth context”) presupposes some preliminary interpretation of the particular work, on its own it does not say much about the non-trivial, complex connections between entities.

To overcome this, we try to model the “birth context” of the previously known literary work with an undirected graph over a finite set of entities as a union of the just defined “birth context” and another, more various set of entities connected to the recipient environment (“recipient context”). It is worth mentioning that these sets are not necessarily disjoint (they share some entities). An example for shared entity can be the author. It can appear as the author of a particular work and as a reader of another or even his own work. Although the role of the author—even his/her existence as an individual

⁵ At this point there is a possibility of theoretical continuation: reading together our categories (which have an origin or etymology of information theory) with Gérard Genette’s theory of textual transcendence, also known as transtextuality—which is “all that sets the text in relationship, whether obvious or concealed with other texts” (1992: 83–84)—, and with the five subdivided categories of transtextuality (intertextuality, paratextuality, metatextuality, hypertextuality, and architextuality). See also: Genette, 1992; 1997a; 1997b.

⁶ See further references at the end of the article.

entity—is a distinguishing and therefore a key element of the above-mentioned theorems, our model does not coerce any entity prior time or space evaluation of our graph.

Additionally, using the resulting underlying abstract graph model we are describing a hypothetical transformation of a literary entity through the active connections in the graph. First, we weakly define a set of transformations (we refer to it as “morphology”) over a particular entity as a series of recent activities distinguishing them into categories according to their effect on the entity. Second, to our conjecture, the emerging paths of forms (descendants of the first occurrence of the original literary work) are slowly abandoning the original “birth context”, shaping a decreased and cleaned set of entities to carry on and expanding or replacing the gap with entities derived from the dynamic set of “recipient context”.

In the next part, our main contributions are:

1. Defining the fundamentals of a hypothetical graph model to capture the birth of a literary work.
2. Defining five basic transformations on the graph, which can alter both the set of nodes and/or the set of edges.
3. Defining an undirected graph as an assemble of context sub-graph, analysis of its properties and a possibility to use it for predicate properties of the literary work.

2. *The context sub-graph*

As we mentioned briefly, our main objective is to define an undirected graph with a set of vertices (nodes) and a set of edges between them. In this section, we only examine the birth and the basis of the evaluation of the contextual sub-graph. We think of a contextual sub-graph as a union of contextual elements and their relation to a set with any tangible or intangible concepts contain at least one of the representations of a particular work. We do not limit the number of elements in either sets, therefore our graph can be infinite or finite. First, we define an important, unique, but not necessary key node in the graph, the *hidden unit*.

It is difficult to define what the *hidden unit* is. The *hidden unit* is not a computable object, it is the existence of the (literary) content/entity which never existed in this form. The *hidden unit* is a possible imagined form of being, the complex of all the possibilities of a (literary) work of art. It could be similar to the meaning of the term Text (“The Text is a methodological field.” [Barthes, 1986: 57]⁷), but it is not exactly the same. We think we only

⁷ A possible way of continuation is Jerome J. McGann’s “quantum field” metaphor—an excellent example of interdisciplinary metaphorical borrowing—which says that certain text-types are quantum fields: “...poems and other imaginative kinds of social texts are quantum fields”. (B7)

know some properties of it, for example: The *hidden unit* “can exist outside its printed version and that the printed artifact contains many elements which we all must consider as not included in the genuine poem...”, or: “...while the mere destruction of the copy of a book or even of all its copies may not touch the work of art at all” (Wellek and Warren, 1949: 142)—where the terms “poem” and “work of art” means the *hidden unit* itself—in our terminology.⁸

Our reasoning has multiple purposes. One of them is that we want to highlight a point of connection among any contextual elements. Therefore, any node in the context sub-graph will be connected to the *hidden unit*. This results in a strong connection between the elements since the maximal distance between two elements will be at most two (either connected or there is a path between them through the *hidden unit*). We define three additional sets of nodes, which are not necessarily disjoint but they do not include the *hidden unit*:

1. “Birth context”: Any tangible or intangible knowledge is connected in any way to the *hidden unit*. There are parts of the “birth context” which never change, but it can be expanded with previously non-included elements. There are also “false” elements—based on a kind of false biographical information. These “false” elements can be removed from the graph in comparison to the fixed part. For the sake of completeness it should be mentioned that the “birth context” can include elements, which are never found or may have been lost over time—e. g.: bibliographical facts, items researched by literary historians, etc. These kinds of elements differ in their type of connection to the “recipient context”. The former elements never had any connection to any element of the “recipient context”, while in the latter case the connection deactivates (lost) over time—or on the other hand it reactivates (found) after a long time.⁹

⁸ It is also important to note that a hidden unit has something which can be called “life”—in time and space—and the other three dimensions of the “context-morphological” analysis (raw material, topics, in-corpus entities); see 1.4.

⁹ The history of the Archimedes Manuscript can be a precise example for this type of rare but theoretically important possible type of elements: “LOST for centuries. FOUND by the Walters Art Museum. (...) In Jerusalem in 1229 AD the greatest works of the Greek mathematician Archimedes were erased and overwritten with a prayer book by a priest called Johannes Myronas. In the year 2000 a project was begun by a team of experts at the Walters Art Museum to read these erased texts. By the time they had finished, the team that worked on the book had recovered Archimedes' secrets, rewritten the history of mathematics and discovered entirely new texts from the ancient world. This exhibition will tell that famous story. It will recount the history of the book, detail the patient conservation, explain the cutting-edge imaging and highlight the discoveries of the dogged and determined scholars who finally read what had been obliterated. (...) The manuscript sold at auction to a private collector on the 29th October 1998. The owner deposited the manuscript at The Walters Art Museum in Baltimore, Maryland, a few months later. Since that date the manuscript

We will refer to an element of this set as s_{ij} where the index “i” corresponds to the i-th *hidden unit* (H_i) and “j” is the index of the j-th element of the set $S_i = \{s_{i1}, \dots, s_{ik}\}$.

The possible nodes of the “birth context” are, inter alia: biographical facts of the author itself (e. g.: age, sex, gender, profession, social network, language and/or other skills, religion, political position, physical and mental state, etc.); information about the various forms of publishing (e. g.: the date and place of publishing, the publishing house, the name of the editor, infos about pre-orders, etc.); the author personal feelings and opinion about the work (related texts from diaries, letters, etc.), and any kind of influences on the author (books, poems, plays in a theatre, films, political thoughts, philosophies etc.).—See: literary history, biographical and historical researches.¹⁰

2. “Representations’ context”: These are the accessible points to a particular *hidden unit*. Here belong the printed copies, handwritten notes, electronic versions or even any part of any previously known representations including the various oral presentations too. Similarly to the “birth context”, we denote the elements in the graph with r_{ij} where i is the index of the *hidden unit* and j refers to the substantive representation from the set $R_i = \{r_{i1}, \dots, r_{i1}\}$.

The possible nodes of the “representations’ context” are, inter alia: each printed copy; all editions; all the philologically or textologically different versions of a (literary) artwork; all electronic versions; each oral presentation, lecture and live performance; and also: all the translations into all languages.—See: philology, textology, textual criticism.

3. “Recipients’ context”: Usually, we think of them as individual entities, but as we will see, they can act together to achieve changes in the graph. Their main role is to imply connections and develop new elements into any of the three contextual sets. The elements are referred similarly as $B_i = \{b_{i1}, \dots, b_{im}\}$.

The possible nodes of the “recipients’ context” are, inter alia: every piece of the activity of the scholars and critics (from monographs and published papers to blogs or other types of

has been the subject of conservation, imaging and scholarship, in order to better read the texts. The Archimedes Palimpsest project, as it is called, has shed new light on Archimedes and revealed new texts from the ancient world. It has also generated a great deal of public curiosity, as well as the interest of scholars throughout the world. On 29th October 2008, we celebrated the ten year anniversary of the project. What was erased text, in terrible condition, impossible to access, and yet foundational to the history and science of the West, is now legible, and instantly available for free by clicking [HERE](#).” (See: <http://thewalters.org/exhibitions/archimedes/> and <http://archimedespalimpsest.org/>) All in all this type of element is very rare, as W. Robert Connor, the president of the Teagle Foundation said about the Archimedes Manuscript to *The New York Times* in 2006: “The number of times you get a new text is very small. It’s like hearing an old violin played at a superb level.” (Lee, 2006)

¹⁰ It can be a conceptual problem that in the description of the “birth context” we don’t want to neglect the results and outcomes of the positivist researches. There can be some kind of difficulties in upholding the distinction between “birth context” and “recipient context”.

posts/comments in the social media), all the referring intertextualities and (scientific, philosophical) replies or counter-arguments.—See: reception theory, hermeneutics, etc.

2.1 Birth of the “context sub-graph”

Before we go into the details about the possible transformations we define a three-step procedure (as seen in Figure 1) to give birth to a context sub-graph.

- **0th iteration:** We think of the starting graph with only a *hidden unit* and the fixed part of the birth context. At this point, both the representations and recipients are either empty or inactive, or in other words there is no direct and/or active connection from the graph.
- **1st iteration:** At least one representation appears in the graph and is connected to the *hidden unit* and to a subset of the birth context. In the example two new representations appear, $R_i = \{r_{i1}, r_{i2}\}$.
- **2nd iteration:** Some set of the recipients encounter with one of the representations. All of them are connected to H_i and at least to one of the representations. In the example three recipients appear and they are connected to only one representation, r_{i2} .

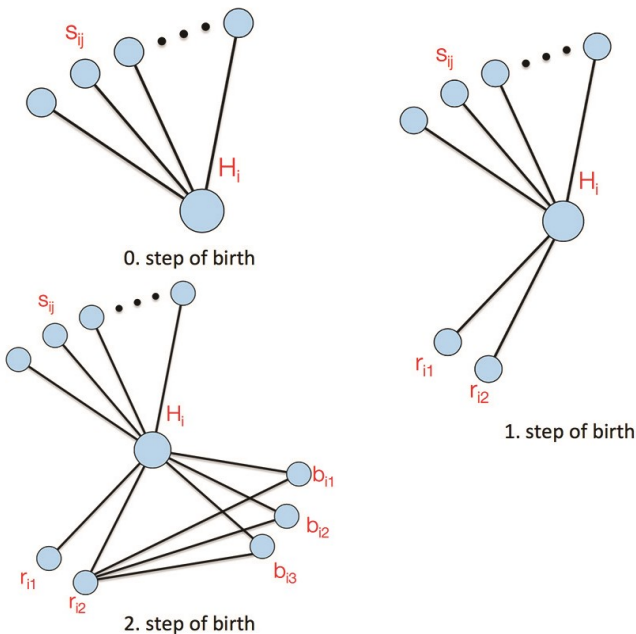


Figure 1. The three steps of the birth of the context sub-graph.

After the 2nd iteration our graph will be self-aware (or developed), any kind of further change in the graph will only be implied by a subset of recipients, B_i^* . This is true even for any expansion of the recipients. It may seem at first sight as a trade-off. To prove otherwise, we define five basic transformations and discuss the coverage they intend. We differentiate two types of nodes and therefore two types of edges at a particular state of the graph. There are active nodes and edges to imply further changes and deactivated or inactive elements, which at a state cannot be part of any transformation. In the figures we will mark the inactive edges with grey and the active edges with black.

2.2 Transformations

As we mentioned previously, all the transformations assume an already developed context sub-graph with at least one element in all the three contextual sets. We define five basic transformations, which can alter both the set of nodes and the set of edges. In the figures we mark the corresponding edges and nodes with red:

1. New representation: As an impact of an active subset of the recipients (B_i^*), a new representation appears. It will be connected to the *hidden unit* H_i , to the elements of B_i^* , and at least to one of the already existing representations which were connected to B_i^* . In Figure 2 we can see an example.

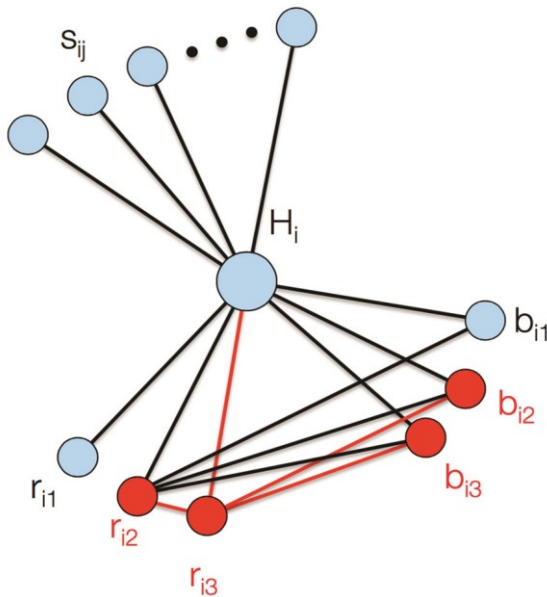


Figure 2. Due to the activity of $B_i^* = \{b_{i2}, b_{i3}\}$ a new representation, r_{i3} , joins the graph. The r_{i2} is the origin of the new representation.

2. Representations connect to the birth context: As an impact of an active subset of the recipients (B_i^*), a representation and an element of the birth context will be connected, or an already existing connection disappears between a representation and an element of the birth context. It is important to note that only the representation connected to the B_i^* can be affected. It is also worth to mention that the recipient does not necessarily need to be connected to the birth context before the transformation, but afterwards they are. In Figure 3 we can see an example.

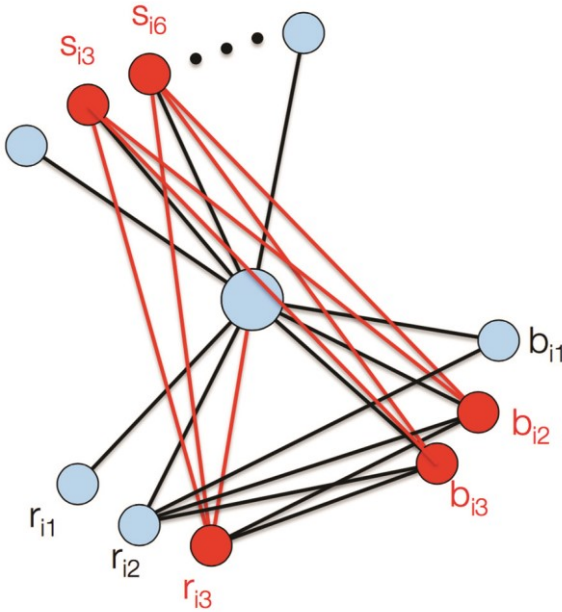


Figure 3. Due to the activity of $B_i^* = \{b_{i2}, b_{i3}\}$, a representation, r_{i3} , will be connected to two elements of the birth context ($S_i^* = \{s_{i3}, s_{i6}\}$).

3. False contextualization: As an impact of an active subset of the recipients (B_i^*), a previously non-existent contextual element gets connected to one of the representations and B_i^* . It should be pointed out that, by definition, this element was not part of the constant set of the birth context, which can imply that it is compulsorily a false element, while it is not obvious to distinguish them from the latter since we defined the fixed set as not necessarily tangible. In Figure 4 we can see an example.

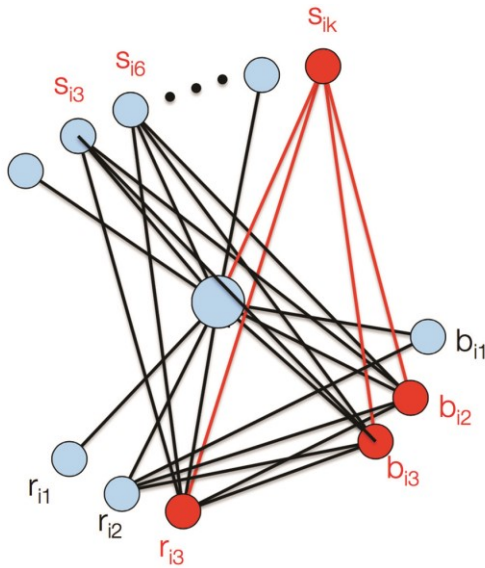


Figure 4. Due to the activity of $B_i^* = \{b_{i2}, b_{i3}\}$, a representation, r_{i3} , will be connected to a false element of the birth context, s_{ik} .

4. Recipient enrichment: As an impact of an active subset of the recipients (B_i^*), a new recipient element appears and gets connected to one of the representations and B_i^* . Through the procedure, the new recipient will be “copying” some of the edges of B_i^* . In Figure 5 we can see an example.

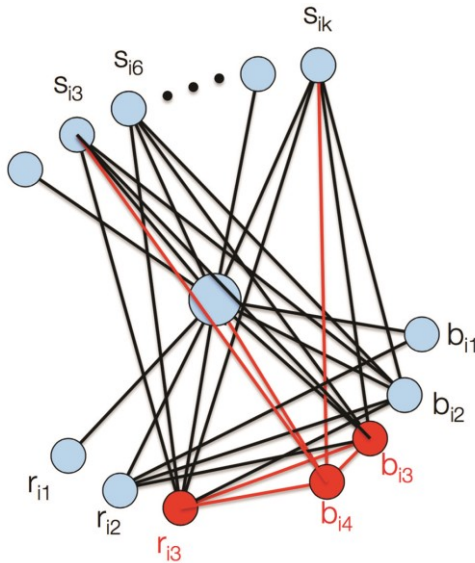


Figure 5. Due to the activity of $B_i^* = \{b_{i3}\}$, a new recipient b_{i4} , will join the graph and “copies” some of the connections of b_{i3} : r_{i3}, s_{i3}, s_{ik} . It is worth noting that the false contextual element is already spreading.

4. Deactivation/Activation: An active subset of the recipients (B_i^*) will change the state of entities from active to inactive or from inactive to active. As a consequence, all their connections will change their state too.

Since any change in the graph can only be triggered by an active subset of the recipients, we can safely play with the idea of activity reduction over time. So it leads us to a strange phenomenon: some of the elements of the fixed part of “birth context” will disappear (or in other words they will lose their importance) in time. Formally, with a simple assumption about the likelihood of the disappearance, even with a very conservative rate, we rapidly get a significant loss. In Figure 6, we can see an example of active and inactive edges where for example r_{i1} is no longer available for active recipients or, in other words, it is “lost forever”.

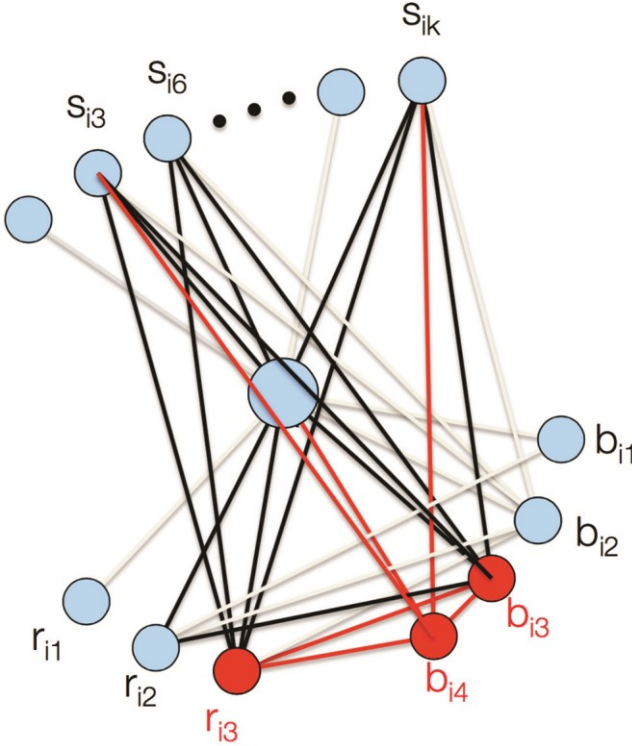


Figure 6. Inactive edges (gray) and the disappearance of entities.

It should be mentioned that although the reason for the disappearance is not immediately relevant, one can simply assume that the abandonment of any element related to a permanently inactive entity will be more likely in time. Here we suggest, without any actual dataset, a simple model to address the alteration. Now let be $|C_{ik}(t)|$ the number of edges related to the representation r_{ik} , $|B_{ik}(t)|$ the number of recipients and $|S_{ik,false}(t)|$ the number of

false contextual elements connected to r_{ik} after t iterations (aka transformations). Our hypothesis is an additive formula,

$$|C_{ik}(t)| = \alpha |C_{ik}(t-1)| + |C_{ik}^*| + |S_{ik,false}(t)| + |B_{ik}(t)|$$

where $0 < \alpha \leq 1$ and C_{ik}^* is a subset of C_{ik} which likely never vanish only the remaining part C_{ik} . Even if we only assume that $\alpha = 0.99$, after 100 iterations the representation will lose more than 50% of its original, fixed “birth context” (C_{ik}^* and C_{ik}), as we can see in Figure 7. This slow procedure shapes a decreased set of cleaned entities, though if the activity is high enough, the gap will be replaced with entities derived from the dynamic set of “recipient context” or with “false” contextual elements ($S_{ik,false}$).

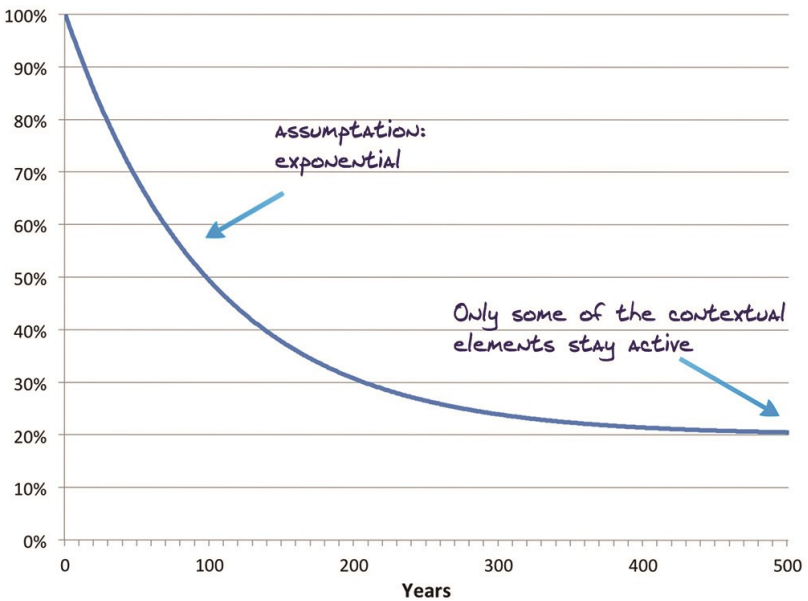


Figure 7. Proportion of active elements of the original, fixed “birth context” in time. (There is a suspicion that if turn *time* into *space*—where the space means the space of interpretation—we will have almost the same function...).

As a possible way of continuation we can say that these were until now “small networks, in which intuition can still play a role, they are like the childhood of network theory for literature; a brief happiness, before the stern of adulthood of statistics.” (Moretti, 2013: 197) In these previous cases the intuition was a good thing, but concepts are better. So we would like to continue our theoretical concept in a way similar to Franco Moretti’s work at the Stanford Literary Lab: “... a much larger study of drama and network theory is in progress at the Literary Lab: a collective project, on hundreds of plays from a variety cultures and historical periods.” (195) Moretti’s projects apply statistical models to textual strings in large corpora of texts. He is interested

in extrapolating cultural and ideological analyses (such as the emergence and features of genres) based on title analysis, or character’s name occurrences, specific linguistic collocations, etc. Our model includes, among others, textual string analysis as an example for the second transformation (“representations connect to the birth context”) in which the recipients are the researchers, the representations are the textual strings and the results are new connections between the representations/recipients and elements in the birth context. After all, the main similarity of our proposal to Moretti’s project is the timing of “do it bigger”—if we are right, this point of our argumentation could be the best to change the perspective, and examine our arguments “globally”—this is the point when an entity as a graph turns to an entity as a sub-graph of a much larger graph (which is after all the world literature itself).

We can see on Figure 8 that our sub-graph is just a small, tiny something in the universe of literary network, which suggests that if we can build a map from all the contents/entities we may have the possibility to understand in an abstract way how a literary artwork becomes a small part of the (world) literature or how a scientific/theoretical thought becomes a part of the “Zeitgeist”.

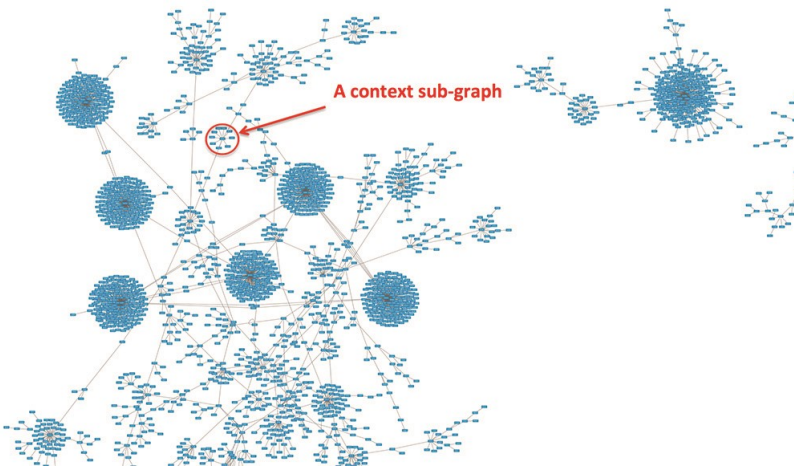


Figure 8. A piece of the whole context graph, an ensemble of all the sub-graphs. The sub-graph in the red circle is a context sub-graph. Note that the graph is not necessarily fully connected.

3. The global context graph and evolution

We have already noted that the context graph is only a sub-graph. Let us define an undirected graph as an assembly of context sub-graphs. For any immediate neighbourhood of a *hidden unit* type node has properties of the context sub-graph. This includes the transformations too. By definition, a

node can be in multiple context sub-graphs, acting differently. This variability is the key to the underlying structure. For a particular context sub-graph the three sets (birth, recipient, representation) are disjoint. Although any node could appear as an element of the birth context in other context sub-graphs, there are necessary restrictions. First, the recipients and the representations are not interchangeable. Second, a recipient, unlike the representations, can also be a recipient in other sub-graphs. Third, the fixed part of the birth context can contain only recipients which have already existed or representations from another context sub-graph.

Since the graph can only change through the transformations, we can offer a brief description of the properties of the graph. Let us start with a very interesting consequence of the *hidden units* and the birth context. As we mentioned, the *hidden unit* is connected to any related nodes in a context sub-graph with a maximal distance of two. Since the definition of the constant part of the birth context suggests connections between the elements of the birth context and popular events, topics, or representations (overall popular nodes), we can assume that this set causes a direct connection to sub-graphs with low average distance among the entities (inside the connected parts). We call these sub-graphs dense. This distance is even smaller if we only examine the active nodes. Without an exact dataset we cannot provide any particular experiment about the degree distribution or other properties of the graph, but as a hypothesis, we think we do not make a big mistake in suggesting that the graph not only follows log-norm or power law as a degree distribution, but it is highly clustered¹¹ and the small-world property¹² is also valid.

Another feasible utilization of this undirected global context graph is the possibility to measure a similarity between two *hidden units* according to their position and their context sub-graph. Here we think of the nodes of the graph as random variables and the edges indicate the neighbourhood of the nodes. If we assume local and global Markov property our graph can be interpreted as a Markov Random Field, which could be used to form a kernel value, some kind of similarity (Daróczy *et al.*, 2015). Furthermore, this similarity can be used to assign or predicate on properties to a literary work at a state through the *hidden unit* (the most unique element for a literary work,

¹¹ “A clear deviation from the behavior of the random graph can be seen in the property of network transitivity, sometimes also called clustering, although the latter term also has another meaning in the study of networks ... and so can be confusing. In many networks it is found that if vertex A is connected to vertex B and vertex B to vertex C, then there is a heightened probability that vertex A will also be connected to vertex C. In the language of social networks, the friend of your friend is likely also to be your friend.” (Newman, 2003: 183)

¹² “The neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors are shown to be small-world networks.” (Watts and Strogatz, 1998: 440)—which means “philosophically” or “ontologically” that our graph-model is a small-world network, and in this way it is similar to other graph-models of natural and human phenomena.

which is also a conceptual construction without a limited and known set of material instantiations).

4. Conclusions

One of the interesting impressions of this “morphology” is an unbalanced context suggesting that the importance of “birth context” is smaller than the importance of the “recipient context”, although this alteration is substantially and unwillingly depending on the cardinality of the possible changes in the “recipient context”. Additionally, in the study of world literature the impact may distinguish clearly national literature from newly born world literature (see the set of criteria for becoming “world literature”—Damrosch, 2003: 281; Tabbi, 2010: 23).

Another feasible utilization of our probabilistic graph is a novel and simple (literary or meta-theoretical) comparatistic method, assuming some properties of our graph: not the interpreter (subjectively) but the “algorithms” as properties or characteristic features of our graph-based model will select (objectively, from the dataset) the necessary theoretical elements of the analytical interpretation and finally they will determine the identity and ontological status of a literary entity.

We feel it is important to remark that we are not arguing against the already mentioned quote: “the more general, the more abstract and hence empty it will seem; the more the concrete object of the work of art will elude our grasp.” (Wellek and Warren, 1949: 6-7). The graph model we presented acts like a set of movements instead of fixing hard rules. Therefore we think of our model as a tool, which can lead to a wide range of possible outcomes. In future work we would like to examine and compare well-known literary theories, through identifying the driving, frequent patterns in the morphology.

References

- AARSETH, Espen J. (1997). *Cybertext: Perspectives of Ergodic Literature*. Baltimore and London: The Johns Hopkins University Press.
- BARABÁSI, Albert-László (2002). *Linked*. New York: Perseus Books Group.
- , Réka Albert (1999). “Emergence of Scaling in Random Networks.” *Science* 286.5439: 509-512. 1 Jun. 2016.
[http://www.nd.edu/~networks/Publication%20Categories/03%20Journal%20Articles/Physics/EmergenceRandom_Science%20286,%20509-512%20\(1999\).pdf](http://www.nd.edu/~networks/Publication%20Categories/03%20Journal%20Articles/Physics/EmergenceRandom_Science%20286,%20509-512%20(1999).pdf).
- BARTHES, Roland (1970). *S/Z*. Paris: Seuil.
- (1986). “From Work to Text.” *The Rustle of Language*. Roland Barthes. Trans. by Richard Howard. New York: Hill and Wang. 56-64.

- BILSKY, Manuel (1953). "The Significance of Locating the Art Object." *Philosophy and Phenomenological Research* 13.4: 531-536.
- BONATI, Félix Martínez (1960). *La estructura de la obra literaria*. Santiago de Chile: Universidad de Chile.
- CAIMO, Alberto, and Nial Friel (2011). "Bayesian Inference for Exponential Random Graph Models." *Social Networks* 33: 41-55.
- CHENG, Justin *et al.* (2014). "Can Cascades Be Predicted?" *WWW 2014 – Proceedings of the 23rd International Conference on World Wide Web*. New York: Association for Computing Machinery. 925-936.
- CORTES, Corinna, Vladimir Vapnik (1995). "Support-vector Networks." *Machine Learning* 20: 273-297.
- DAMROSCH, David (2003). *What is World Literature?* Princeton: Princeton University Press.
- DARÓCZY, Bálint *et al.* (2013). "Fisher Kernels for Image Descriptors: A Theoretical Overview and Experimental Results." *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae – Sectio Computatorica* 40: 201-214.
- (2015). "Text Classification Kernels for Quality Prediction over the C3 Data Set." *WWW 2015 – Proceedings of the 24th International Conference on World Wide Web*. Geneva: International World Wide Web Conferences Steering Committee. 1441-1446.
- DAVIES, David (1996). "Interpretive Pluralism and the Ontology of Art." *Revue Internationale de Philosophie* 50.198: 577-592.
- (2007). *Aesthetics and Literature*. London: Continuum International Publishing Group.
- DERRIDA, Jacques (1998). *Of Grammatology*. Trans. Gayatri Chakravortti Spivak. Baltimore and London: The Johns Hopkins University Press.
- ERDŐS, Paul, and Alfréd Rényi (1959). "On Random Graphs." *Publicationes Mathematicae* 6: 290-297.
- ESKELINEN, Markku (2012). *Cybertext Poetics: The Critical Landscape of New Media Literary Theory*. London: Continuum.
- FOUCAULT, Michel (1969). *L'archéologie du savoir* [The Archeology of Knowledge]. Paris: Éditions Gallimard.
- FRANK, Ove, and David Strauss (1986). "Markov Graphs." *Journal of the American Statistical Association* 81: 832-842.
- GADAMER, Hans-Georg (2004). *Igazság és módszer* [Truth and Method]. Trans. Gábor Bonyhai. Budapest: Osiris.
- GARZÓ, András *et al.* (2013). "Cross-lingual Web Spam Classification." *WWW 2013 Companion – Proceedings of the 22nd International Conference on World Wide Web*. Republic and Canton of Geneva (Switzerland): International World Wide Web Conferences Steering Committee. 1149-1156.
- GENETTE, Gérard (1992). *The Architect: An Introduction*. Trans. by Jane E. Lewin. Berkeley CA: University of California Press.

- (1997a). *Palimpsests: Literature in the Second Degree*. Trans. by Channa Newman and Claude Doubinsky. Lincoln, NB: University of Nebraska Press.
- (1997b). *Paratexts: Thresholds of Interpretation*. Trans. by Jane E. Lewin. Lincoln NE and London: University of Nebraska Press.
- INGARDEN, Roman (1931). *Das literarische Kunstwerk: Eine Untersuchung aus dem Grenzgebiet der Ontologie, Logik und Literaturwissenschaft*. Halle: Max Niemeyer.
- (1973). *The Literary Work of Art: An Investigation on the Borderlines of Ontology, Logic, and Theory of Literature*. Trans. George G. Grabowicz. Northwestern University Press.
- JAUSS, Hans Robert (1982). *Aesthetic Experience and Literary Hermeneutics*. Trans. Michael Shaw. Minneapolis: University of Minnesota Press.
- (1982). *Toward an Aesthetic of Reception*. Trans. Timothy Bahti. Minneapolis: University of Minnesota Press.
- LANDOW, George P. (1992). *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: The Johns Hopkins University Press.
- (1997). *Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: The Johns Hopkins University Press.
- (2006). *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization*. Baltimore: The Johns Hopkins University Press.
- LEE, Felicia R. (2006). “A Layered Look Reveals Ancient Greek Texts.” *The New York Times*, Nov. 27. May 8 2016.
http://www.nytimes.com/2006/11/27/arts/27greek.html?_r=0.
- MCGANN, Jerome J. (2002). “Literary Scholarship in the Digital Future.” *The Chronicle of Higher Education: The Chronicle Review* 49.16: B7–B9.
- MITIAS, Michael H. (1982). “The Ontological Status of the Literary Work of Art.” *Journal of Aesthetic Education* 16.4: 41–52.
- MORETTI, Franco (2005). *Graphs, Maps, Trees*. London: Verso.
- (2013). *Distant Reading: The Formation of an Unorthodox Literary Critic*. London: Verso.
- NEWMAN, Mark E. J. (2003). “The Structure and Function of Complex Networks.” *SLAM Review* 45.2: 167–256.
- PÁL, Dániel Levente (2010). “Közelítés a kiberfilológiához” [Towards Cyberphilology]. *Prae* 4: 23–33.
- (2012). “Szövegparadigmák – vektorok, hálózatok, önhasonló alakzatok” [Paradigms of Textuality – Vectors, Graphs, Fractals]. *A szótól a szövegig* [From Word to Text]. Ed. Vilmos Bárdosi. Tinta, Budapest. 193–199.
- (2014). “Visualizations (Maps) & Visions (Metaphors) of the Internet.” *5th Visual Learning Conference: Pictures – Parables – Paradoxes*. Budapest University of Technology and Economics, Department of Technical Education. Budapest, November 15, 2014. (in print).

- RAMSAY, Stephen (2011). *Reading Machines: Toward an Algorithmic Criticism*. Chicago, IL: University of Illinois.
- ROBINS, Garry *et. al.* (2007). "Recent Developments in Exponential Random Graph (p^*) Models for Social Networks." *Social Networks* 29: 192-215.
- SHUSTERMAN, Richard (1984). *The Object of Literary Criticism*. Atlantic Highlands, N.J.: Humanities Press.
- SIKLÓSI, Dávid *et. al.* (2012). "Content-Based Trust and Bias Classification via Biclustering." *Proceedings of the 2nd Joint WICOW/AIRWeb Workshop on Web Quality WebQuality'12*. New York: Association for Computing Machinery. 41-47.
- STAHL, Gerry (2010). *Tacit and Explicit Understanding*. Published by Gerry Stahl at Lulu.com, printed in the USA.
- TABBI, Joseph (2010). "Electronic Literature as World Literature; or, The Universality of Writing under Constraint." *Poetics Today* 31(1): 17-50.
- WATTS, Duncan J., Steven H. Strogatz (1998). "Collective Dynamics of 'Small-world' Networks." *Nature* 393.6684: 440-442.
- WELLEK, René, Austin Warren (1949). *Theory of Literature*. New York: Harcourt, Brace & World.

© 2016 Dániel Levente Pál and Bálint Daróczy.

Licensed under the [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 International \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).