Graph Logic Model Framework for Predictive Linguistic Analysis

M. Charnine[1], I. Kobozeva[2], S Loesov[3] and I. Schagaev[4]
[1]Institute of Informatics Problems, 119333, Moscow, Vavilova St., 44/2, Russia
[2]Department of Theoretical and Applied Linguistics, Moscow State University, Leninskiye Gory, 1, Moscow, Russia, 119991
[3]Institute for Oriental and Classical Studies, Russian State University for the Humanities, Miusskaya Pl.6, korp. 1, 125993, GSP-3, Moscow, Russia
[4]IT-ACS Ltd, Shephall View, SG1 1RR, Stevenage, UK

Abstract - *In this paper we are trying to combine interests of authors and collaborators into one non-mutually exclusive concept and logic flow aiming to create a framework for searching of a migration of words from one cluster to another; this enables us to define a semantic shift well before it become obvious. We show how introduced graph-logic model can be applied for analysis of migration of the meaning of sentences indicating, quite often, a paradigm shifts. Using Artificial Intelligence as an example we illustrate development of AI from philosophy of mind to science, science fiction and technology, including games in science, technology, and further education. Several examples how proposed model with supportive searching framework applied in mentioned areas detecting evolving processes are presented*.

Keywords: A Graph Logic Model, Semantic search, Long-term trends, Google Books Ngram, Historical data, Predictive linguistic analysis

1 Introduction

During our previous research over migration of terms and areas in curriculum design [Bacon13] it was discovered that areas of research and knowledge in general are moving, changing, morphing and this "evolution" can be detected and even predicted. While practical use of this conception was already proved in number of papers [Bacon13], [Bacon14] and patent [Patent07] it is worth to investigate how it can be applied in linguistics, what are the limits and what it enables in terms of analysis and monitoring of migration of terminology and corresponding semantic shifts. At first we introduce a basic model – so called graph-logic model [Schagaev14], [Schagaev15] applying it for analysis of semantic shifts and migrations of terms. Every model that describes system initially using graph theory (GT) differentiates entities (nodes, vertices) and relations – (links, arcs, edges). GT model description is static; behavior aspect of the model is described separately by introducing rules and procedures allocated to nodes and links, such as activation, or algorithmic description of process to change state on a graph.

When rules for the graph path determination are applied algorithmically for every node then moves on the graph might be defined without contradiction.

The way of leaving/arriving to/from any node can be described with the help of logic operators (LO) from the basic {OR, AND, XOR}. LO might be chosen to apply to all and every node to define conditions of leaving/arriving. Say, if one applies "XOR" operator for the node leaving condition and chooses links along the graph, we are able to redraw the graph and actually mark activated links for each node, applying XOR rule for choosing only one link.

This scheme with some other restrictive conditions is used widely and known as Markov Process (MP). MP adds to the XOR rule (applied for every vertex) a normalization condition – either one adjacent link is activated or another, while probabilities – kind of "weights" - are used to "normalize" the chances of choosing one particular link (sum of probabilities to come out from a node equals 1, note that for incoming this condition does not stand).

To complete correct introduction of travel along the graph for XOR logic one has to introduce termination condition for this travel. Termination condition in MP graph as a whole is introduced by so-called aggregate state which must be reached and the condition that one of the nodes in a column is activated when the process is leaving i-th column (one of the states from there) and arriving to i+1-th column, as well as sum of probabilities in every column is equal to 1.

Again, MP works when "XOR" logic is applied to every node as a decision making rule to leave or to arrive the node. In system programming a using "XOR" operator can describe the mutually exclusive operations – i.e. concurrency, with separation of processes at the critical section entering. In

This work is partially supported by the RFFR, Grants 16-07-00756 and 15-07-06586

"classic" probability theory "XOR" logic applies when conditional probabilities are used. Models that use conditional probabilities require XOR operators for every node of process description by definition.

Let us consider another rule of leaving a node: AND logic. In this case semantically opposite graph model is introduced that describes transitions for every node at once, at the same time, instantly; when links exist from i-th node, say, to j-th, k-th,..., x-th nodes then all possible links are activated all together at the same time.

This logical condition assumes a parallel movement along the graph from any node to all connected nodes. We name this rule applied to a node as "AND" logic. Examples of the systems that are using AND logic for every node are:

- Broadcasting networking
- Salesman problem analysis,
- Diffusion processes in physics,
- Quantum effects model,
- Parallel calculations when hardware resources are unlimited, etc., etc.

Finally, "OR" logic might be applied for each and every node on the graph, assuming that only one or some links are selected, therefore flexible parallelism might be described – when it is not necessary to start everything at once.

Three natural questions arise here:

A. Can one apply various logic operators for various nodes?

B. Do conditions at output vertices dictate input conditions to other nodes?

C. How to separate logic applied? Do they affiliate to the vertex? Or Link?

It is clear that when every node has one input and one output it does not matter. On the contrary, if a graph has several links to or from its nodes it does...

Instinctive reply to question A is yes: what one does in parallel might be exclusive at the other end of the link; therefore an answer to question B is no. It is worth mentioning that conditions to leave a node (vertex) and arrive to another vertex should be attributed to edge not vertex, Fig.1

Figure 1. Logic of the edge

Their (logic operators) combination might be even more interesting: say leaving condition is "OR" but arriving condition is "AND" for each input - and we have Petri net described. To illustrate descriptive "power" of the combination of graph and logic models for behavior of graph let us draw a graph - Figure 2 - applying various logic operators for different nodes.

Let us describe Figure 2 in some details. Nodes (vertices) have output and input links. Those links that might be activated either one or another or both (node a links to node b and d are described in callout **ORo(b,d)** - that means that no order or imperative timing is required to move from vertex **a** (OR logic).

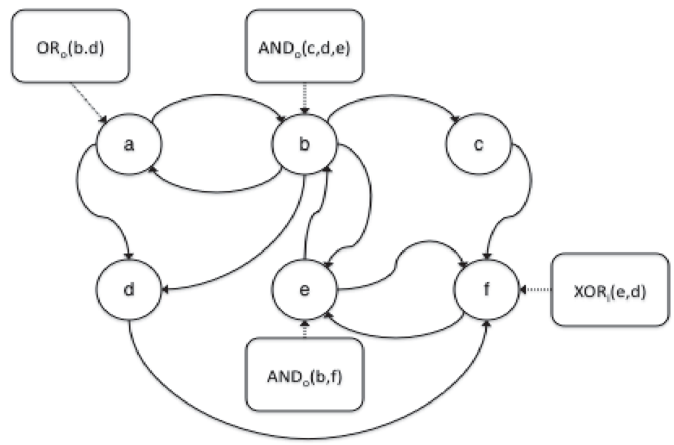

Figure 2. Various operators applied to leave and arrive

In turn, node **b** assumes parallel activation of links to nodes **d**, **e**, and **c** as it is described in special callout with operator **ANDo(c,d,e).** Note that link from node **b** to node **a** is not included in this list. Finally, input links that are required to be mutually exclusive at the node f are described by special callout as (XORi (e,d)). Again, note that incoming link from node c is not included as an input XOR operator of node f and thus might be analyzed separately.

The model proposed here in general is quite simple; it defines a graph behavior with various assumptions of leaving and arriving conditions for all vertices. This approach allows an analysis of large scale graph behavior in much more details and greater variety that the "standard" graph theory. Every node **x** of the graph such as Figure 5 thus might be described as a string:

x: AND-(j,j,k) AND+(l,m,n) OR-(p,q) OR+(r,s,t) XOR-(u,f) XOR+(g,h)

Minus "-" or "o" stand for every logic operator for output link, Plus "+" or "i" for every input link. Logic operators that have to be applied to various combinations of output links are explicitly presented in this notation.

Interestingly, leaving conditions do not obligatory match arriving ones: leaving one place together with all the rest such as parallel (AND-) calculations might be mutually exclusive at the arriving XOR – concurrency.

The addition of weights or normalized weights on the edges of GLM defines Bellman optimization model. For this we have to rename weights as rewards and penalties affiliated to each edge and input and output operators all must be “XOR”.

2 GLM use

The model proposed here enables to define:

- Mutual dependence
- Structural unconditional parallelism
- Concurrency
- Transitive dependence
- Loops, including Hamilton ones
- Search of special sub-graphs that match a chosen of pattern

Graph analysis model is complete when termination condition for travelling along the graph is introduced. It is clear, that when AND, OR or XOR logic is becoming part of the model then the termination conditions might be absolutely different from the ones of the known models. Examples of termination condition might be:

A. Visiting (numeration) of all nodes, or
B. Finding the structures with:
i. Selected condition such as existence of Hamilton loop
ii. Searching of all sub graphs in the graph
iii. Searching of particular weighted sub trees
iv. Searching of selected sub-graphs to match required searching pattern
C. Quantitative termination;
D. Formation of the table about all distances for every node to every other;
E. Finding heaviest and shortest paths for each pair of nodes when they are transitively connected;
F. Defining strategies which node to choose to maximize gain along the travel from the selected node to the terminal node (horizon);
G. Development of balanced schemes (sub graphs) for each node that will define critical path along the graph travel.

Using the search of dependency in the graph of the model within complex system (like results of the search for new trends over the web, see [Bacon13],[Charnin15] might exceed dimension of the matrix 105 by 105.

When GLM applies for analysis and simulation of impacts propagation of event, say for aircraft, the model of aircraft can be described as <GT, AND,P>, where AND stands to the logic operator used and P is normalized weight of every node – say, probability of use [Patent07].

Note that there is no Markov condition to leave the node (XOR and sum of leaving probabilities is 1; Markov attributes vertex, while GLM model attributes an edge; the start and the end points of the edge might be completely different.

Instead of visiting all nodes and forming all possible outcomes a probabilistic weight of the links (edges) might be extremely useful to converge algorithm when ξ threshold is applied: $Pi \leq \xi i$, where Pi is probability of i-th link activation, ξ - empirically assigned value.

Thus for every vertex of potential dependencies of elements it is possible to form hierarchy of lists with various termination weights $\{\xi 1, \xi 2, ..., \xi x\}$. The model briefly described here already has become a logic core of a “Method of active system safety” recently patented [Patent07]. Regarding monitoring and analysis of semantic change a model can be applied as follows.

2.1 GLM use in linguistic domain

An application of GLM enables us to analyze migration of the meanings, semantic shifts and observe behavior of a knowledge domain. We think that popular term of Artificial Intelligence might be interesting to analyze along its evolution.

At first logician Immanuel Kant explicitly introduced the way of thinking and defining one’s own intelligence in the end of 18th century. It was done using categories that described elements’ dependencies in terms of schemes “one to many”, “one to one”, “many to one”. These categories we extend by GLM that makes logic operators map to natural language elements as Table 1 illustrates. Using GLM it is easy to introduce and analyze the leaving and arriving conditions in word dependencies [Bacon13]:

Table 1. Language statements and Logic Operators

Statement in language	**Logic operator**
One of	XOR
Maybe, some	OR
Always	AND
Never	NOP
Some	Combination of XOR, OR, AND

The meaning of the word can be modified in syntactic context: the word “bin” might be a noun (*a bin*) or a verb (*to bin*), as well as *Google* or *Hoover* – *googling* and *hoovering* indicate the action of using the companies made systems, thus noun became a verb in proper contexts while staying a noun in other contexts. A word can change its positive connotations into negative and then again to positive as Russian *mecenat* ‘Maecenas’. Subtle historical semantic shifts of the key words of different cultures are well represented in ethno-linguistic studies [Wierzbicka97].

Then defining terms and observing their shift we can:

- In the example above describe a shift from a noun to a verb;
- Change the imperativeness of use making another table hierarchy of will;
- Create new dependencies, reflect shifts of meanings or actions (from guessing and pure research to industry and detect unavoidable actions instead of possible ones);
- Do fine-tuning of curriculum - in other words we can check efficiency of advertising, detect technological revolution before it becomes visible even for authors;
- Detect direction of change where and when nobody actually sees that.

To summarize these bullet points: we will be able to see the changes of the world BEFORE the world itself realizes it.

The concept of language transformation and change of lexical meaning as a reflection of national, economic, territorial and ethnic changes as well as the change of the form of words is an established fact in historical linguistics.

See below an example of such changes in Babylonian branch of Akkadian languages – Table 2.

Table 2 Evolution of meanings in Babylonian languages

English	Old Babylonian	Middle Babylonian
so, thus	Kīam	akanna
All	Kalûm	gabbu
good, beautiful	Damqum	Banû
one another	aḫum aḫam	aḫāmiš
Urgently	arḫiš	ḫamutta
Work	Šiprum	dullum
towards (a person)	ana ṣēr	ana muhhi
chez (French)	itti, ina mah(a)r-	itu-
there is no	ul ibašši	Yānu
dispatch	ṭarādum	šūṣû
when (conjunction)	Inūma	kī
sunrise	ṣīt šamšim	napāḫ šamši
return	Târum	naḫāsum
Neglect	egûm,aham nadûm	mekû
suit, fit	redûm, naṭûm	alākum

Bodies of languages compared here are not equal Old Babylonia is extremely large, Middle Babylonian is much smaller, A comparison was made on a basis of what is possible in the 1800-1700 BC for Old Babylonian and 1350-1200 for Middle Babylonian [Loesov14].

The handling of textual data and extraction of required properties can be done nowadays using a GLM-based framework in combination with Google Ngram tool.

That help to see a growth of “neighbors” for keywords and therefore predict further shifts of their semantics and changes of knowledge in the corresponding knowledge domain.

2.2 GLM use in “macro linguistic”

Another approach of using GLM stretches up to “migration of ideas” [Charnin15], [Jacob13]. In this domain we can divide knowledge areas into almost non-overlapping segments allocating to them macro-nodes of GLM.

Then links that represent dependencies might be considered as weights, normalized weights or probabilities of changes/migration.

This approach enables us to trace a trend. E. g. the manifestation of new concepts in the Internet can cause their reflection in scientific papers and later in products, industrial systems or other developments.

This is also useful as an instrument to evaluate probability and delays between the appearance of an idea and its reappearance translated into some other language. Say, an appearance of “artificial intelligence” as a term in English required several years to reappear in Russian scientific works.

Probabilities and delays in this process are traceable and therefore define handicap or advances of various scientific groups.

“An idea” can be represented as a group of semantically close phrases and terms that are used similar to such methods as LDA [Blei2003]. Ideas can be organized or be part of hierarchy, elements of which can be again vertices or nodes of GLM.

Figure 3 illustrates main levels of hierarchy. Analyzing link’s strength, frequency, occurrence area (academic or industrial, political) might help to evaluate where the world goes, and to do it automatically; when the program itself searches terms and creates nodes and links drawing a map that is a semantic map of research evolution.

As for the immediate application or implementation aspect, one can find dependencies between terms using Google Books Ngram [Roth13];

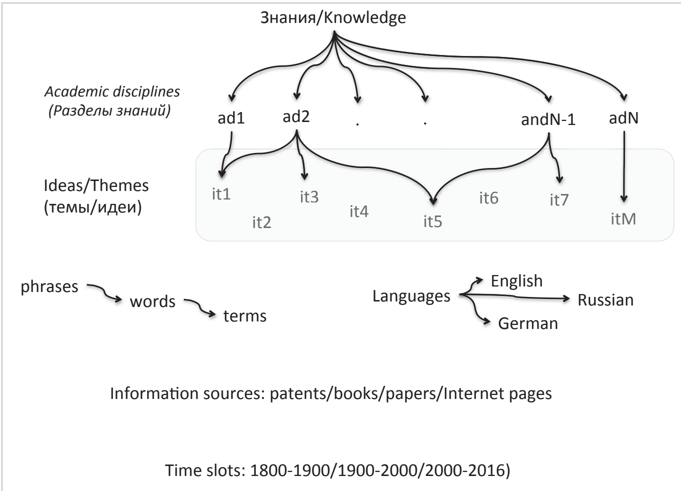

Figure 3. Hierarchy of knowledge

Figure 4 presents an example of dependency between appearances of “can machine think” in English after Turing’s book [Turing1950] and the question “can machine think” in Russian scientific literature Russian that appeared after 1960.

3 Migration of term *Artificial Intelligence*: an example

Migration of terms from English books into Russian took several years in 1940-1960.

For example, the term “Artificial Intelligence” became used in English language books since 1957, and word-by-word translated as “искуственный интеллект” it migrated into Russian scientific literature in 1965, as Books NGram indicates, Figure 5.

Similar delays were observed in migration of a word combination “machines think” (from 1944 up to 1955) or a complex term “Turing test” (1960 up to 1966), Figures 6,7.

It is highly likely that initially the term *artificial intelligence* was not defined explicitly, denoting a kind of new phenomena related to computer science and computer technology.

It is worth to investigate further some other combinations of words closely related to the same area which did not become widely accepted terms *machine intelligence*, *intelligent machines*, *intelligent machinery* and Russian equivalents: *думающие машины*, *машинный интеллект*.

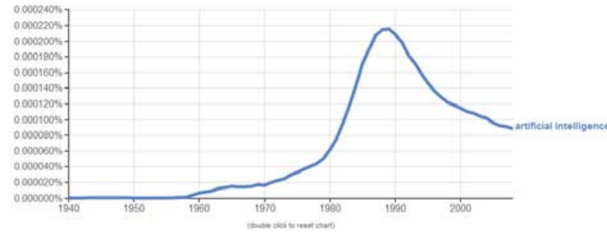

Figure 4. Trend of the term “Artificial Intelligence”

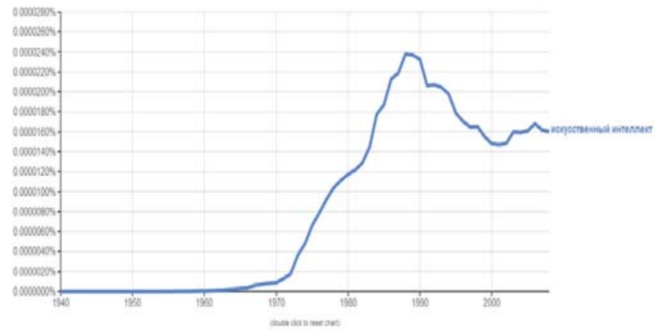

Figure 5. Trend of the term “искусственный интеллект”

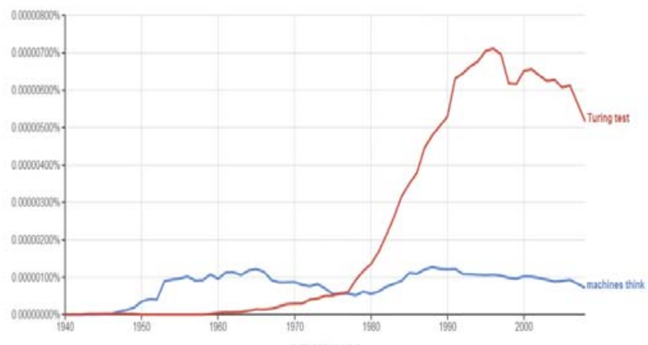

Figure 6. Trends of “Machine thinks” and “Turing test”

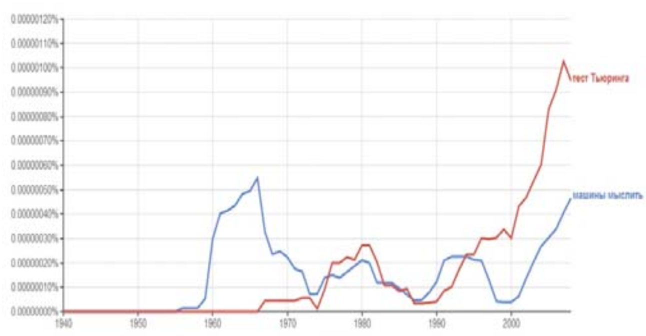

Figure 7 Use of words “машины мыслить” , “тест Тьюринга”

The biggest impact on frequency of use for a combination of words “machine thinks” was no doubt caused by Alan

Turing's paper "Can the Machine Think?" [Turing1956] initially published under the title "Computing Machinery and Intelligence" in 1950 [Turing1950]. This paper was translated into Russian in 1960 and the growth of frequency of Russian translational equivalent of "machine thinks" indicates this. John McCarthy who coined the term Artificial Intelligence (AI) in 1955, defined it as "the science and engineering of making intelligent machines". There is no strong evidence though that he produced a reasonable definition of an "intelligent machine" that is required for the explicit definition of artificial intelligence.

In 1955 the term AI appeared, that was used before as a free combination of words, associated with variety of meanings. Before this *artificial intelligence* was used in books in the context of philosophy, occurring in philosophical arguments about the nature of mind and ethics of creating artificial beings. Then it was not clear whether it was possible to create a complete working AI. It must have been highly unlikely as Turing test was not known then and therefore criteria of artificial intellect were vague. Surprisingly, there is no drastic change in this direction – while much more developed programs are appearing and passing various tests including the Turing test, there is still no clear understanding of what is possible and what is not in the domain of AI and to what extent AI can match human intelligence.

Since I. Kant [Kant34], followed by philosophers/novelists [Capek21], [Asimov50], [Asimov53], [Lukas65], [Minskly55],[Minsky58], [Chomsky12] a discussion of understanding, of how brain becomes mind and how we can make a schema of this process was primarily conceptual, virtual.

Philosophical arguments about the nature of mind and the ethics of creating artificial beings were pretty intensive. Initially, the term AI was used together with such words as *brain*(7), *consciousness*(5), *thinking machine*(3), *human mind*(2), *philosophical*(1), *philosophers*(1). Here in brackets we count number of uses of these words in Turing's paper "Can the Machine Think?" [Turing1956]. All these words belong to philosophic context and appeared well before 1800 – as Google Ngram demonstrates (Figure 8).

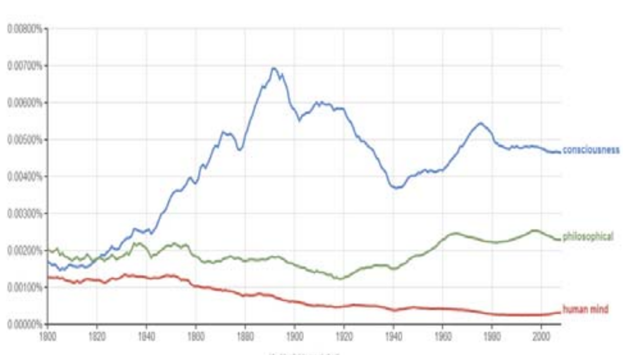

Figure 8. Trends of "consciousness", "human mind" and philosophical"

To some extent human ability of creation of an artificial mind that is equal to human one is far from clear, in spite of recent €2bn projects funded by EC and similar one in USA. This might be explained by the absence of adequate technology and math methods to attempt any implementation.

With appearance of automata theory and later transistors Information Computer Technology (ICT) emerged and the term AI migrated to the area of engineering and became associated with development of simulation systems as well as play of Chess and Go.

Other terms, which are now used most intensively in the context with AI, such as *robotics* (1970), *computer vision* (1974), *natural language processing* (1964) and speech recognition (1957) appeared after 1957. Here brackets show years of appearance of these terms – according to Google Books Ngram.

The strength of associative relation and relative frequency of AI in context with other terms (according to data of Google Search Engine) can be ordered as Table 3 shows.

It is clear that nowadays AI is considered as a technical term and is used in technical environment with terms such as "speech recognition", "robotics", "computer vision", "natural language processing").

Philosophical context is still present but appears less often: ("brain", "human mind", "philosophers", "philosophical", "consciousness").

Therefore paradigm shift from science to implementation is taking place, while philosophical issues concerning this concept are not completely resolved.

Table 3. AI and related terms

Terms used wih AI	Number of appearances	Freq-cy
"AI" "speech recognition"	1,670,000 / 6,020,000	= 0.277
"AI" "robotics"	16,500,000 / 65,700,000	= 0.251
"AI" "computer vision"	1,470,000 / 8,580,000	= 0.171
"AI" "natural language processing"	543,000/4,600,000	= 0.118
"AI" "brain"	19,300,000/532,000,000	0.0363
"AI" "human mind"	254,000 / 8,140,000	0.0312
"AI" "philosophers"	356,000 / 21,700,000	0.0164
"AI" "philosophical"	516,000 / 55,700,000	0.00924
"AI" "consciousness"	591,000 / 84,600,000	0.00698

4 Conclusion and next steps:

- Google searching machine is a tool that provides a raw data. Requirements and ways to interpret them, creation of the next iteration of search has to be formalized.
- Regretfully our subjectivism limits rigorous monitoring of trends in the languages and knowledge areas (the concept of an idea might be not really strong). Introduction of Graph Logic Model as a framework for autonomous monitoring of the knowledge provides the most powerful framework for analysis of objects in their dynamic.
- This paper illustrates it using linguistic knowledge. The growth of the amount of neighborhood terms means that the subject evolves. And vice versa. Applied in combination with Google tool Graph Logic enables a self-tuning framework up to automatic generation of scenarios for next search.
- This "self-adjustments" might help to modify, say, university curriculum, shifting some elements to skills and some others to advanced studies, or timely delete modules or disciplines that became obsolete.
- We propose an approach of using a graph logic model for linguistic methods of analysis showing "where the meaning goes".
- Shown that evolution of semantics in various domains, interaction of terms and change of their internal properties might be done using automatic searching tool "supervised" by proposed graph logic model.
- As immediate application of proposed approach we see:
  - Evolving scheme of application of GLM as a part of searching framework with reflections of knowledge in the chosen domain
  - Visualization support for terms migration and domain change in the way of how it works with the subject.
- Using term "Artificial Intelligence" evolution we demonstrated that our approach works.

5 References

[Bacon13] Bacon E, Schagaev I, Charnine M, Foggie R, Kravtsov G, Kirk B., "Web-enhanced design of university curricular", http://worldcomp-proceedings.com/proc/p2013/FEC4025.pdf

[Blei2003] D. M. Blei, A. Y. Ng, and M. I. Jordan. "Latent Dirichlet allocation", Journal of Machine Learning Research, 3:993–1022, Mar. 2003.

[Bacon14] "Essential Knowledge Aggregation, Delivery, and Assessment", E-learn Volume 2014 Issue 5, 05-01-2014 Article No. 1

[Capek21] Capek K., "R.U.R" Play, 1921

[Chomsky12] Chomsky N "Where AI went wrong", The Atlantic, Nov 1, 2012, http://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/

[Charnin15] M.Charnine, S.Klimenko, "Measuring of "Idea-based" Influence of Scientific Papers", Proceedings of the 2015 International Conference on Information Science and Security (ICISS 2015), December 14-16, 2015, Seoul, South Korea, pp.160-164.

[Jacob13] Jacob S., Jerry A. 2013. "Receptivity Curves: Educational Research and the Flow of Ideas: Expanded Version", Population Studies Center, University of Pennsylvania, PSC Working Paper Series, PSC 13-10, http://repository.upenn.edu/psc_working_papers/50

[Kant34] Kant I. Critique of Pure Reason, Everyman, 1934, ISBN 0-460-87358-X

[Loesov14] Aramaic languages http://ivka.rsuh.ru/binary/78809_46.1236189872.35672.pdf

[Lukas65] Lucas, J. R. 1965. "Minds, Machines, and Gödel." Philosophy 36: 112-127.

[Minsly55] McCarthy, J., M. L. Minsky, N. Rochester, C. E. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence."

[Minsky68] Minsky, M. 1968, "Semantic Information Processing", Cambridge, MA: MIT Press.

[Patent07] Patent GB 0707057.6 http://www.it-acs.co.uk/files/Grant_for_a_patent.PDF

[Roth13] Roth, Steffen, "The fairly good economy. Testing the economization of society hypothesis against a Google ngram view of trends in functional differentiation (1800-2000)", Journal of Applied Business Research, 2013, Vol. 29 No. 5, pp. 1495-1500

[Schagaev14] Schagaev I, "Control Operators Vs Graph Logic Model", Proc of Int-l Conference on Foundation of Computer Science, FCS14, pp180-182,

[Schagaev15] Schagaev I, Kaegi-Traschel T Software Design for Resilient Computers Springer 2015 ISBN 978-3-319-29463-6

[Turing1950] A. M. Turing "Computing Machinery and Intelligence", Mind, New Series, Vol. 59, No. 236 (Oct., 1950), pp. 433-460.

[Turing1956] A. M. Turing, "Can the machine think?", The World of Mathematics, 1956, vol. 4, Ed. J R Newman (New York: Simon and Schuster), pp. 2099-2123.

[Wierzbicka1997] A. Wierzbicka, " Understanding Cultures Through Their Key Words: English, Russian, Polish, German, and Japanese", 1997, N.Y., Oxford Univ. Press, pp. 1-317.