



Cognitive Architecture of Collective Intelligence based on Social Evidence

Anton Kolonin¹, Evgenii Vityaev² and Yuriy Orlov¹

¹*Institute of Cytology and Genetics, Novosibirsk, Russian Federation*

²*Institute of Mathematics, Novosibirsk, Russian Federation*

akolonin@gmail.com, vityaev@math.nsc.ru, orlov@bionet.nsc.ru

Abstract

Development of collective intelligence and consciousness starts playing key evolutionary role given spread of social networks and massive Internet communications. Paper considers cognitive model and architecture capable to capture cognitive phenomena in social environments, which could be used for modeling of collective cognition as well as developing practical applications for individual interaction with collective intelligence. The model is evaluated by means of qualitative and quantitative experiments.

Keywords: cognitive architecture, collective intelligence, social evidence, knowledge, consciousness

1 Introduction

As it has been pointed out by Valentin Turchin [1], the evolution is going over so-called “meta-system” transitions from simple systems to complex ones. The same is true for intelligence, “as ability to reach complex goals in complex environments, using limited resources” (accordingly to Ben Goertzel [2]). The intelligence, from cybernetic perspective, is grounded on capacity to store information. In this regard, one of the “meta-system” transitions has been experienced with appearance of mammals with information capacity of their brains (supplied with number of neural connections in the cortex) exceeding capacity of mammalian genome (supplied with number of possible gene interactions). The other transition is being experienced right now while the tight connectivity of humans in world-wide social networks is getting close to the one of human brain. Importance of the latter transition for evolution and humanity could not be underestimated because of amount of nodes and links in modern computer networks is exponentially growing with addition of artificial agents being involved in “hybrid” human-computer networks. Each of such agents, from cloud-based services to in-pocket “personal assistants”, may carry simple-to-complex forms of artificially intelligent behavior possessing increasingly growing memory capacity and interaction speed beyond human capabilities.

The latter fact makes it critically important to understand phenomena of social interactions “en masse”, so that effects of social behavior based on collective consciousness could be well understood, predictable and manageable – from perspective of humans exposed to modern information networks.

Pioneer work on mathematical modeling of social behavior has been presented by V. Lefebvre [3], who has shown the ability of a formal model to predict development of social interactions performed by individual agents. From phenomenological perspective, the most representative analysis of possible social behavior phenomena has been done by R. Cialdini [4], where multiple cases of society-bound effects of individual behavior are considered as having either positive or negative impact on survival and benefits of single individual or entire society.

In the further work, we consider approach for building computational model encompassing knowledge acquired by entire society by means of evidence supplied by each of its members. In such a model, it is anticipated the reasoning within the model can be performed in terms of probabilistic inference or fuzzy logic such as non-axiomatic reasoning systems by P. Wang and E. Vityaev [5,6]. The environmental constraints for the computational model are imposed by intelligence definition made by B. Goertzel [2], so that intelligent behavior is constrained by available physical resources, available to either alive, artificial or “hybrid” being. This is assumed to be done on the basis of “evidence-based knowledge representation model” with resource constraints as suggested earlier by A. Kolonin [7,8]. The latter would make it possible to use this model to predict and manage behavior of animal swarms, human societies or multi-agent computer systems. Finally, we suggest to generalize the theory of functional systems (TFS) developed by P. K. Anokhin and may others to extend over representing single intelligence system [9] to describe behavioral aspects of society of such systems.

2 Approach for Cognitive Architecture

We are developing the following cognitive architecture, capable to model collective consciousness as well as individual one (as extreme case with only one member in society) based on “social evidence, constrained by resources”, as originally suggested earlier [7,8]. Primarily, this can be qualified saying that meaning of a fact or relationship for subject of cognition (individual or society) is determined as cumulative function of the meanings of all of the fact or relationship appearances for all social referees of the object, with account to extent of social connection between the subject and a referee. Secondary, this is constrained stating that amount of appearances of a fact or relationship as well as amount of social referees and strengths of social connections to them is defined by cognitive capabilities of the subject of cognition. The subject of cognition could be either individual alive being, or society or artificial system or society of such systems etc. – we will denoting it as “agent” further.

Within this approach, there are requirements to support different kinds of data processed within the cognitive architecture possessed by agent, from three different perspectives (Fig.1).

Cognitive perspective

First, from primary cognitive perspective, there would be three different segments of the graph representation of the entire set of knowledge possessed by the system. The top layer is called “foundation graph” of basic knowledge, which is necessary for any social system to be shared by all of its members in order to communicate. The middle is called “imagination graph” keeping “inferred” knowledge, based on foundation terms and relationships residing in the “foundation graph” and supported by “evidence” coming from the bottom layer. The bottom layer is called “evidence graph” which contains everyday life-time experiences possessed by the agent (left side of Fig.1).

The “foundation” layer can be called “belief system” encompassing unconditional or “commonly accepted” information regarding the surrounding environment and agent itself. It can be considered as cornerstone cognitive base, or set of “absolute truths” about the surrounding world, storing basic

“belief system” of the agent. Without having that shared, the two agents speaking the same language syntactically, would not understand each other. Reasoning on this part of knowledge might be called orthodox, stereotypic or closed-minded thinking. It is anticipated for cognitive architecture to have a mechanism for either accepting the knowledge coming to this layer from the outer world via “imagination graph” (if it is compatible with the belief system), or rejecting it (in the opposite case). It is assumed that, under normal circumstances, foundation graph does not need any probabilistic or fuzzy inference applied to it, so the cognitive functions and respective decision making operations executed efficiently (in terms of resource consumption) by means of “binary logic” inference. The architecture should ensure that portions of “imagination graph” exceeding given threshold of relative amount of evidence can be “hardwired” to the foundation graph if it is considered beneficial for the system from the perspective of resource saving.

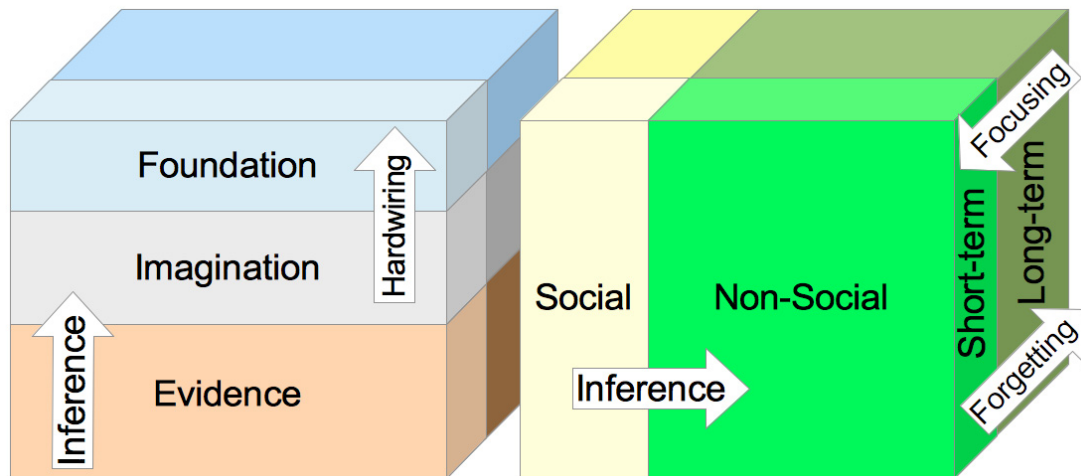


Figure 1: Segmentation of data from three different perspectives.

The “imagination graph” can be thought as a dynamic pool of novel evidence-based knowledge coming to an agent via communication channels over time. This part of the knowledge graph can be considered as dynamic, non-stereotypic or open-minded core. It is inferred given the scope of “evidence data” currently residing in the attention focus and the trust levels specific to particular social referees providing the inputs or expressing their attitude to them, such as sign of evidence supplied for relationships in the “evidence graph”. The inference engine is expected to maintain “image of the world” in the “imagination graph” constantly, relying on the scope of evidence data within actual time frame and social context (identified by set of relevant social referees). The cognitive architecture should provide capability to collect cumulative evidence and draw inferred trust values for respective relationships to communicate them back to the outer world later or “hardwire” to the foundation graph eventually. Obviously, maintenance of the dynamic “truth values” by means of probabilistic inference might be more time consuming and allocate more resources than operations in “foundation graph”.

The “evidence graph” records temporal events or facts of evidence to draw cumulative probabilistic assertions in the “imagination graph” on that basis, with account to social evaluation. This pool of socially relevant temporal facts serves as an evidence base for the inference engine calculating the truth values with account to subjective grounds and temporal context. Each piece of information here is timestamped and labeled by a social referee communicating it or expressing their opinion in its regard. Data stored here can be also subject of “evidence compression” with either clustering of fractional time slices into larger time intervals or aggregating evidences from individual peers into larger groups of peers. Further, in the course of “evidence consolidation”, it can be removed

from this pool with transition of knowledge (derived from the evidence) from the imagination graph to the foundation graph – if the cumulative evidence gets high enough. Also, the “evidence forgetting” can effect in complete removal of evidence from the graph if no extra supporting evidence is experienced for long time. The processes of “compressing”, “consolidating” and “forgetting” evidence are driven by physical resource constrains, so the system assures the amounts of all data fit the existing memory and allocates less resources to store and process the knowledge. The basic goal of the agent is – maintain the most reliable knowledge fitting the system’s internal belief to a greater extent, spending as much less resources on it as possible.

Social perspective

Within the dynamic social evidence-based knowledge representation model [7,8], truth value of any piece of information for given period of time is calculated as sum of truth values of evidential facts supporting it in the given time frame and communicated by peer agents or social referees or just evaluated by them, multiplied by the trust levels given to each of these referees. Each fact or relationship or event of elementary evidence is assumed specific to particular source in social environment and time. To deal with this assumption, segmentation of the entire agent's “knowledge graph” can be further split in “social” and “non-social” segment (right side of Fig.1).

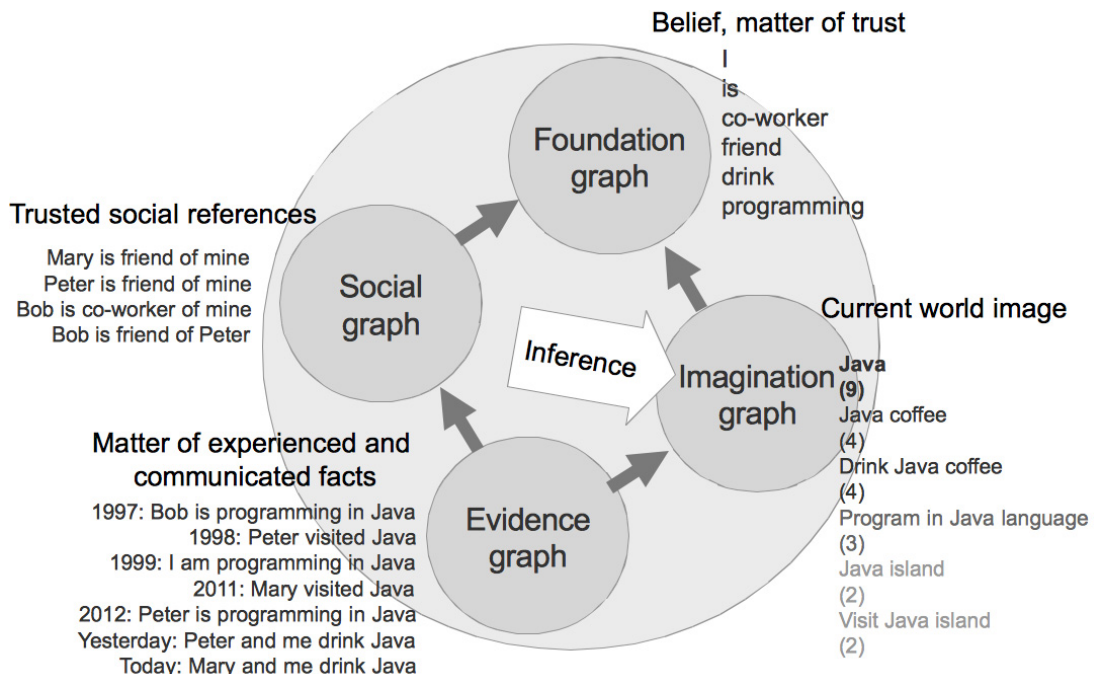


Figure 2: Inference on socially and temporally constrained truth values.

The “social graph” describes social interaction channels possessed by the agent and provides the basis for account of subjectivity, so that each fact in the “imagination graph” is supplied by trust given to a particular social referee at a time. This is effectively the social core, or personal social network of the agent, maintaining trust levels for each of peer agents – how much confidence can be given to incoming information communicated by the peer.

Based on this, in the course of probabilistic inference on any sorts of knowledge of question, the scope of supporting evidence is not only restricted with time frame, but is also accounts for judgments

regarding reliability of different evidences, which can be done based on the amount of evidence associated with these facts, evaluated in terms of trust towards the evidence source – social referee supporting the evidence by means of communicating it or expressing their attitude in its regard (Fig.2).

Functional Perspective

In our early work we build intelligent decision-making functional system FS combining a sequence of functional systems in case they implement some standard sequence of actions that is implemented always the same way [9]. Suppose any elementary system has a set of sensors S_1, \dots, S_n , which characterize both the state of the system itself and of external environment perceived by it using the sensors. Notably, in context of the current discussion, the environment includes physical objects and effects measured by the sensors as well as social events and impacts experienced directly or indirectly. Each sensor has a set of possible indications V_{S_i} . The system also has a set of available actions in the environment. Any action that system performs at a moment t_i may result at a moment t_{i+1} in some changes in the surrounding environment, and consequently, in his sensors indications.

The sequence of functional systems $\langle FS_1^1, \dots, FS_n^1 \rangle$,

$$FS_1 = s_0 \left\{ \begin{array}{c} \xrightarrow{FS_1^1, \dots, FS_n^1} \\ \rightarrow s_1^1 \rightarrow s_2^1 \rightarrow \dots \rightarrow s_{goal}^1 \quad P_{FS_1^1} \end{array} \right\} s_1, \dots, FS_n = s_{n-1} \left\{ \begin{array}{c} \xrightarrow{FS_n^n, \dots, FS_1^n} \\ \rightarrow s_1^n \rightarrow s_2^n \rightarrow \dots \rightarrow s_{goal}^n \quad P_{FS_n^n} \end{array} \right\} s_n,$$

automatically combines into a functional system $FS = (s_{Goal}, \{\langle FS_1, \dots, FS_n \rangle\}, P_{FS})$ with one element in the set $\{\langle FS_1, \dots, FS_n \rangle\}$ of the form:

$$FS = s_0 \left\{ \begin{array}{c} \xrightarrow{FS_1, \dots, FS_n} \\ \rightarrow s_1 \rightarrow s_2 \rightarrow \dots \rightarrow s_{goal} \quad P_{FS} = P_{FS_1} \cdot \dots \cdot P_{FS_n} \end{array} \right\} s_{goal}.$$

If the sequence of actions is standard, and not switched in the middle to the execution of a different sequence, then the probability P_{FS} of functional system will be equal to the product $P_{FS_1} \cdot \dots \cdot P_{FS_n}$ of the probabilities of its constituent functional subsystems.

Moreover, automatic unification of functional systems occurs for the same reason as the elaboration of the conditioned relations for rules: the elaboration by the inner loop the conditional connection between the execution of the first functional system FS_1 and the result of the entire sequence of actions if it always obtains a sequence of results $s_1 \rightarrow s_2 \rightarrow \dots \rightarrow s_n$. Given the elementary system represented in such recursive manner, the higher-level social system may be built as superposition of its lower-layer elementary members.

3 Evaluation of the approach

Relying on the architecture described above and partially implemented accordingly to our earlier work [7,8] and probabilistic semantic inference engine [6], the following results of qualitative modeling of social behavior and quantitative study of social patterns have been obtained.

Quantitive research on determination of the “social clusters” have been performed using real-time social network data (the sources and content can’t be disclosed because of privacy reasons). Using profile information and content analysis in social network data, with use of probabilistic semantic inference [6], it has been possible to determine development of socially significant behavioral patterns confirmed in the field. The amount of 2784 social network users has been used, with 36 individual characteristics inferred and involved in the analysis and 21 behavioral types inferred in the end. One of them could be described as “married human female, living in the city of her childhood, focused on her family and children, valuing kindness and honesty on behalf of other people, treating smoking and alcohol negatively”. The other one can be expressed as “unmarried human male, likely teenager,

focused on self-development, valuing kindness and honesty on behalf of other people, tolerant to smoking and alcohol”.

Qualitative model for long term social interaction within randomly distributed and sized human societies has shown the following patterns. Within initially random distribution of individuals with semi-equal “belief system” in an environment with sufficient resources and low stress from the environment, there is a trend with eventual formation of clusters with common “belief systems” within each of the clusters, different from one of the other cluster. Over the time, the speed of “belief system” divergence is increased, due to the fact that social evidence of the information within the cluster is higher than the outer one and so the more and more information “friendly” the cluster inhabitants is uploaded their to “belief system” making it less and less open for “alien” information with evidence from outside of the cluster.

When such a social cluster is exposed to pressure from outer environment or experiencing lack of resources, the clusters are forced for more tight interactions. Even if no external pressure to all of the clusters is imposed, given limited resources, existing clusters still may be interacting one with another so these interactions would act as external impact for each of the interacting ones. And then, for a social cluster, the external impact may have different effects on the internal cluster structure. For one case, if say there is negative impact which may be better handled having members of the cluster co-operated, joining their “belief systems” towards co-operated mitigation of the negative impact, the cluster may have its internal structure consolidated (proven by ability of nations to stand against alien interventions based on common religion or trust into authorities). For another case, if a negative impact may be better handled by some fraction of society having it isolated from the other fraction, the internal structure of the cluster may experience further granulation (proven by the record of breaking the nations broken into separate principalities). It may also have different effects given the positive impacts applied. Positive impact disjoining the community could be limited amount of external goods supplied to it so the ability of ones to consume more goods at the expense of the others may increase inequality and hence increase granularity. In turn, when there is an external good non-consumable by smaller social group but consumable by larger community, it could be unifying social clusters together.

4 Interpretation of the results

The process of splitting society into clusters is generally limited by the fact that any separate group incapsulated in social environment immediately becomes subject of external impacts from surrounding clusters which typically prevents its further granulation.

That is, variability of collective consciousness may be described as continuous process of splitting larger societies into smaller ones and them having smaller societies merged back – possibly, in other superpositions. Obviously, speed of such variations, as well as upper and lower limits on cluster size depend on cognitive capabilities and properties of individuals within these societies, degree of connectivity within them, speed of information spread and nature and strength of external impacts.

For one instance, low level of intelligence and slow information spread would decrease the speed of the processes and make the limits closer, so that communities stay intact longer with semi-constant size (such as size of swarm, family or tribe). On the opposite side, high level of intelligence coupled with high connectivity and speed of information exchange boost the speed of changes and make their scale relaxed, like in modern world we know certain communities joined around a novel “belief system” may be formed world-wide just in years or even months. At the same time, when the social structure of a society is grounded onto environment of its habitat, limitation of resources in the environment would trigger internal conflicts with intensity proportional to degree of granulation within the society.

That means, in the physical world with limited resources, variability of collective consciousness could potentially lead to conflicts between the clusters within the society due to growth of self-

oscillation processes driven by growth of connectivity and information transfer speed as well as because of external powers applied to society.

5 Conclusion

The discoveries of the patterns and trends like discussed above, with possibility to track their changes over time correlated with known changes in environment, effectively enables to capture collective consciousness for target social groups or communities – human or animal and predict their behavior. It makes it possible to create practical applications for corporate, governmental or personal use, such as “Artificial Intelligent Internet Agents” (Aigents at <http://aigents.com>) intended for personal use by people with everyday exposure to social interaction or information on the Internet [10]. In the future, we plan to implement entire cognitive architecture described above coupled with probabilistic inference engine [6] within the personal prediction and recommendation system based on Aigents platform.

6 Acknowledgment

The work was supported by Russian Science Foundation (grant 14-14-00269).

References

1. Turchin, V. (1977). *The Phenomenon of Science: A Cybernetic Approach to Human Evolution*. Columbia University Press, New York
2. Goertzel, B. (2012). *CogPrime: An Integrative Architecture for Embodied Artificial General Intelligence*. Open Cog
3. Lefebvre, V. (2001). *Algebra of conscience*. Springer-Verlag
4. Cialdini, R. (1984). *Influence: The Psychology of Persuasion*. ISBN 0-688-12816-5
5. Wang, P. (2013). *Non-Axiomatic Logic: A Model of Intelligent Reasoning*. World Scientific Publishing Co.
6. Vityaev, E. (2015). *Unified formalization of «natural» classification, «natural» concepts, and consciousness as integrated information by Giulio Tononi*. The Sixth international conference on Biologically Inspired Cognitive Architectures (BICA 2015, November 6-8, Lyon, France), *Procedia Computer Science*, v.71, Elsevier, 2015. pp 169-177.
7. Kolonin, A. (2015). *Distributed Knowledge Engineering and Evidence-Based Knowledge Representation in Multi-agent Systems*. Knowledge Engineering and Semantic Web Volume 518 of the series Communications in Computer and Information Science, Springer, pp 291-300.
8. Kolonin, A. (2015). *Computable cognitive model based on social evidence and restricted by resources: Applications for personalized search and social media in multi-agent environments*. International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON), 2015, Novosibirsk, Russia
9. Vityaev, E. (2015). *Purposefulness as a Principle of Brain Activity // Anticipation: Learning from the Past*. (ed.) M. Nadin. Cognitive Systems Monographs, V.25, Chapter No.: 13. Springer, 2015, pp. 231-254
10. Kolonin, A. (2015). *Personal Agents for Social Intelligence*. Knowledge Engineering and Semantic Web, 6th International Conference Proceedings, ISSN 1865-0929, Moscow, Russia, 2015, pp 283-290