Consciousness as a brain complex reflection of the outer world causal relationships [1]

Evgenii Vityaev[1,2][0000-0002-5651-6781]

[1]Sobolev institute of mathematics, Koptuga 4, Novosibirsk, Russia
[2]Novosibirsk State University, Pirogova 2, Novosibirsk, Russia
vityaev@math.nsc.ru

Abstract. In the previous works we analyzed and solved such problem of causal reflection of the outer world as a statistical ambiguity. We defined maximally specific causal relationships that have a property of an unambiguous inference: from consistent premises we infer consistent conclusions. We suppose that brain makes all possible inferences from causal relationships that produce a consistent model of the perceived world that shows up as consciousness. To discover maximally specific causal relationships by the brain, a formal model of neuron that is in line with Hebb rule was suggested. Causal relationships may create fixed points of cyclic inter-predictable attributes. We argue that, if we consider attributes of the outer world objects regardless of how we perceive them, a variety of fixed points of the objects' attributes forms a "natural" classification of the outer world objects. And, if we consider fixed points of causal relationships between the stimuli of the objects we perceive, they form "natural" concepts described in cognitive sciences. And, if we consider the information processes of the brain when the system of causal relationships between object stimuli produces maximum integrated information, then this system may be considered as a fixed point which has a maximum consistency in the same sense as the entropic measure of integrated information. It was shown in other works that this model of consciousness explains purposeful behavior and perception.

Keywords: clustering, categorization, natural classification, natural concepts, integrated information, concepts.

1 Introduction

Causality is a result of physical determinism: «for every isolated physical system some fixed state of a system determines all the subsequent states» [1]. But let us consider an automobile accident [1]. What is the reason for it? It might be a road surface condition or humidity, position of the sun with respect to the drivers' looks, reckless driving, psychological state of driver, functionality of brakes, etc. It is clear that there is no any certain cause in this case.

[1] The work is supported by the Russian Science Foundation grant №17-11-01176.

In the philosophy of science the causality is reduced to forecasting and explanation. «Causal relation means predictability … in that if the entire previous situation is known, an event may be predicted …, if all the facts and laws of nature, related to the event, are given» [1]. It is clear that nobody knows all the facts, which number in the case of an accident is potentially infinite, and all the laws. In the case of a human being and animals, the laws are obtained by training (inductive reasoning). Therefore, causality is reduced to predicting by inductive-statistical (I-S) reasoning, when predictions logically inferred from facts and statistical laws with some probabilistic assessment.

Causal relationships and laws, discovered on real data or by training, face with the problem of statistical ambiguity – contradictions (contradictory predictions) may be inferred from discovered causal relationships [2-3]. To avoid inconsistences, Hempel [4-5] introduced a requirement of maximal specificity, which implies that a statistical law should incorporate maximum of information, related to the predicted property. Following Hempel, we defined maximally specific rules for which we proved a consistency of (I-S) inference, which use only these maximum specific rules [2-3]. A special semantic probabilistic inference was developed that discover maximum specific rules, which might be considered as the most precise causal relationships (having maximum conditional probability and using maximum available information). Causal relationships may create fixed points of cyclic inter-predictable attributes.

The structure of the outer world objects was analyzed in the form of «natural» classifications. It was noted by naturalists that «natural» classes of animals or plants have a potentially infinite number of different properties [6]. Naturalists, who were building «natural» classifications, noted that construction of a «natural» classification was just an indication: from an infinite number of attributes you need to pass to the limited number of them, which would replace all other attributes [7]. This means that in «natural» classes these attributes are strongly correlated, for example, if there are 128 classes and the attributes are binary, then the independent «indicator» attributes among them will be about 7 attributes as $2^7 = 128$, and others can be predicted based on these 7 attributes. We can select different 7-15 attributes as «indicators» and then others potentially infinite attributes can be predicted based on these selected attributes. So, there is a great number of causal relationships between the «natural» class attributes.

We formalize the «natural» classification based on the generalization of the Formal Concept Analysis (FCA) [8]. Formal concepts emerging in the FCA may be specified as fixed points of deterministic implications (with no exceptions) [8]. We generalize formal concepts for probabilistic case by introducing probabilistic maximum specific rules instead of deterministic implications and defining fixed points for probabilistic implications [9-10]. We argue in [11] that, if we apply this generalization to some sample from a general population, then we receive a «natural» classification of that sample, which meets all the requirements that naturalists made for «natural» classification.

In the work [12] we demonstrate that semantic probabilistic inference might be considered as a formal model of neuron that satisfy the Hebb rule, in which this inference produce all the most precise causal relationships. Causal relationships discovered

by neurons may form fixed points of cyclic inter-predictable properties that produce a certain «resonance» of mutual predictions (excitations) of neurons.

High correlation of attributes for «natural» classes was also confirmed in cognitive science. Eleanor Rosch formulated the principles of categorization, one of which is the following: « the perceived world is not an unstructured set of properties found with equal probability, on the contrary, the objects of perceived world have highly correlated structure» [13-14]. Therefore, directly perceived objects (so called basic objects) are rich with information ligaments of observed and functional properties, which form a natural discontinuity, creating categorization. Later, Bob Rehder suggested a theory of causal models, in which the relation of an object to a category is not based on a set of attributes but on the proximity of generating causal mechanisms: «the object is classified as a member of a certain category to the extent that its properties are likely to have been generated by this category of causal laws» [15]. Thus, the structure of causal relationships between the attributes of objects is taken as a basis of categorization. Therefore, brain perceives a «natural» object not as a set of attributes, but as a «resonant» system of causal relationships, closing upon itself through simultaneous inference of the total aggregate of the «natural» concept features. At the same time, «resonance» occurs, if and only if these causal relationships reflect some integrity of some «natural» class, in which a potentially infinite number of attributes mutually presuppose each other. To formalize causal models, Bob Rehder proposed to use causal graphical models (CGMs) [16]. However, these models are based on «deployment» of Bayesian networks, which do not allow cycles and cannot formalize cyclic causal relationships. Instead of it, our model is directly formalizing the cyclic causal relationships by the fixed points of predictions on causations [8-12].

Neuron transmits its excitation to the other neurons through multiple both excitatory and inhibitory synapses. Inhibitory synapses may slow down other neurons and stop their activity. It is important for «inhibiting» alternate perception images, attributes, properties, etc. Within our formal model it is accomplished by discovering «inhibitory» causal relationships, which predicts absence of an attribute/property of the object (the perceived object shall not have the respective attribute/property) as compared to the other objects, where this characteristic is found. A formal model specifies it by predicates' negations for corresponding attribute/property. Absence of inconsistencies at the fixed points, that is proved in [10-11] for the most precise causal relationships, means that there are no two most precise causal relationships simultaneously predicting both availability of some attribute/property with an object, and its absence.

It should be specially noted that the «resonance» of mutual predictions of the perceived properties of the objects is carried out continuously in time and therefore the predicted properties must coincide with the ones that have just been perceived from that objects. The absence of contradictions in the predictions is also the absence of contradictions between the predicted stimulus and the really received stimulus.

2 Comparison with the integrated information theory

If the «natural» classification describes objects of the external world, and «natural» concepts are the perception of these objects, then the integrated information theory describes the information processes of the brain when these objects are perceived.

G. Tononi defines consciousness as a primary concept, which has the following phenomenological characteristics: composition, information, integration, exclusion [17-18]. For a more accurate determination of these properties G. Tononi introduces the concept of integrated information: «integrated information characterizing the reduction of uncertainty is the information, generated by the system that comes in a certain state after the causal interaction between its parts, which is superior information generated independently by its parts themselves» [18].

The process of reflection of causal relationships of the outer world (Fig. 1) shall be further considered. It includes:

1. The objects of the outer world (car, boat) which relate to certain «natural» classes;
2. The process of brain reflection of objects by causal relations marked by blue lines;
3. Formation of the systems of interconnected causal relationships, indicated by green ovals.

In G. Tononi's theory only the third point of reflection is considered. Integrated information is also considered as a system of cyclic causality. Using integrated information the brain is adjusted to perceiving «natural» objects of the outer world.

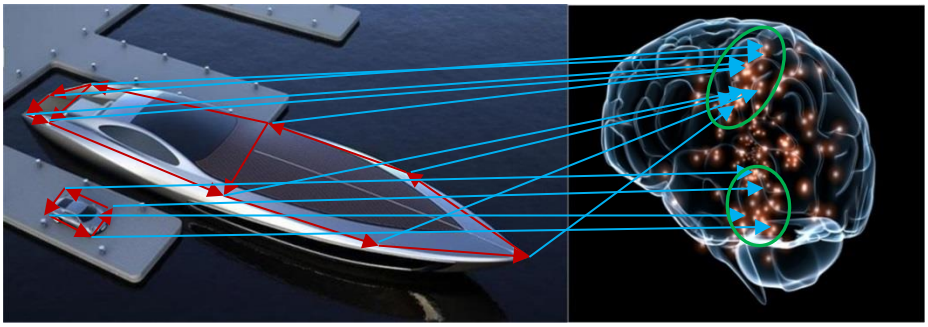

Fig. 1. Brain reflection of causal relationships between objects attributes

In terms of integrated information, phenomenological properties are formulated as follows. In brackets an interpretation of these properties from the point of view of «natural» classification is given.

4. Composition – elementary mechanisms (causal relationships) can be combined into the higher-order ones («natural» classes in the form of causal fixed points produce a hierarchy of «natural» classes);

5. Information – only mechanisms that specify «differences that make a difference» within a system shall be taken into account (only a system of «resonating» causal relationships (differences), forming a class (difference) is important;
6. Integration – only information irreducible to non-interdependent components shall be taken into account (only system of «resonating» causal relations, indicating an excess of information and perception of highly correlated structures of «natural» object is accounted for);
7. Exclusion – only maximum of integrated information counts (only values of attributes that are «resonating» at the fix-point and, thus, mostly interrelated by causal relationships, form a «natural» class or «natural» concept).

These phenomenological properties are defined as the intrinsic properties of the system. We consider these properties as the ability of the system to reflect the complexes of external objects' causal relations, and consciousness as the ability of a complex hierarchical reflection of a «natural» classification of the outer world.

Theoretical results on consistency of inference and consistency of fixed points of our formal model are suppose that a probability measure of events is known. However, if we discover causal relationships on the training set, and intend to predict properties of a new object out of the training set and belonging to a wider general population, or to recognize a new object as a member of some «natural» concept, there might be inconsistencies. Here, a certain criterion of maximum consistency is employed, which is based upon information measure, close in meaning to an entropic measure of integrated information [19].

3 Discussion

Theoretical results obtained in the paper suggest that it is possible to create a mathematically precise system of the reality reflection, based on the most specific causal relationships and the fixed points. It can be shown that the reflection of causal relationships is able to model a multitude of cognitive functions in accordance with existing physiological and psychological theories. The organization of purposeful behavior is modeled by causal relationships between actions and their results [20], which fully correspond to the theory of functional systems [21]. The fixed points adequately model the perception [19]. A set of causal relationships models expert knowledge [22]. Therefore, the verification of this formal model for compliance with the actual processes of the brain seems to be an important task.

References

1. Carnap, R. Philosophical Foundations of Physics. Basic Books (1966).
2. Vityaev, E. The logic of prediction. In Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference, Novosibirsk, Russia, August 16-19, 2005; Goncharov, S., Downey, R., Ono, H., Eds.; World Scientific: Singapore, 2006; pp. 263-276.

3. Vityaev, E., Odintsov, S. How to predict consistently? Trends in Mathematics and Computational Intelligence. Studies in Computational Intelligence. V. 796, 2019, Pages 35-41.
4. Hempel, C. Aspects of Scientific Explanation. In: Aspects of Scientific Explanation and other Essays in the Philosophy of Science, The Free Press: New York, 1965.
5. Hempel, C. Maximal Specificity and Lawlikeness in Probabilistic Explanation. Philosophy of Science 1968, 35, 16–33.
6. Mill, J. System of Logic, Ratiocinative and Inductive. L. (1983).
7. Smirnof, E. Constructions of forms from the taxonomic view. Zool. Jour. 1938, v.17(3), 387-418. (in Russian).
8. Ganter, B. Formal Concept Analysis: Methods, and Applications in Computer Science. TU: Dresden, Germany (2003).
9. Vityaev, E.; Demin, A.; Ponomaryov, D. Probabilistic Generalization of Formal Concepts. Programming and Computer Software 2012, V. 38:5, 219–230.
10. Vityaev, E.; Martinovich, V. Probabilistic Formal Concepts with Negation. In: PCI 2014, Voronkov, A., Virbitskaite, I., Eds.; LNCS 8974, 2015; pp. 385-399.
11. Vityaev, E.; Martynovich V. "Natural" classification and systematic formalization as a fix-point of predictions. Siberian Electronic Mathematical Reports 2015, 12, 1006-1031.
12. Vityaev, E. A formal model of neuron that provides consistent predictions. In Biologically Inspired Cognitive Architectures 2012. Proceedings of the Third Annual Meeting of the BICA Society (Chella, A., et al eds.). Advances in Intelligent Systems and Computing, V.196, Springer: Heidelberg, New York, Dordrecht, London, 2013; pp. 339-344.
13. Rosch, E. Natural categories. Cognitive Psychology 1973, 4, 328-350.
14. Rosch, E. Principles of Categorization. In Cognition and Categorization, Rosch, E., Lloyd, B., Eds.; Lawrence Erlbaum Associates, Publishers: Hillsdale, 1978; pp. 27–48.
15. Rehder, B. Categorization as causal reasoning. Cognitive Science 2003, 27, 709–748.
16. Rehder, Bob; Martin, J. Towards A Generative Model of Causal Cycles. In 33rd Annual Meeting of the Cognitive Science Society 2011, (CogSci 2011), Boston, Massachusetts, USA, 20-23 July 2011, V.1; pp. 2944-2949.
17. Ozumi, M.; Albantakis, L.; Tononi, G. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0 // PLOS Computational Biology 2014 May, V.10. Issue 5.
18. Tononi G. An information integration theory of consciousness. BMC 2004. Neurosci 5:42.
19. Vityaev, E.; Neupokoev, N. Perception formal model based on fix-point of predictions. In Approaches to mind modeling 2014. Ed. Red'ko, V., URSS Editorials: Moscow, p. 155-172.
20. Vityaev, E. Purposefulness as a Principle of Brain Activity. In Anticipation: Learning from the Past 2015. Ed. Nadin, M., Cognitive Systems Monographs, V.25, Chapter No.:13, Springer, 231-254.
21. Anokhin, P.K.: Biology and neurophysiology of the conditioned reflex and its role in adaptive behaviour1974, Oxford etc.: Pergamon press, pp. 574.
22. Vityaev, E.; Perlovsky, L.; Kovalerchuk, B.; Speransky S. Probabilistic dynamic logic of cognition. Biologically Inspired Cognitive Architectures 2013. Special issue: Papers from the Fourth Annual Meeting of the BICA Society (BICA 2013), v.6, 159-168.