STATING OF INDUCTION PROBLEM USING SECOND LEVEL LAWS OF NATURE

Evgenii Vityaev[1], Irina Khomicheva[2]

[1]Institute of Mathematics, Russian Academy of Sciences, Novosibirsk 630090, Russia, vityaev@math.nsc.ru
[2]Novosibirsk State University, Novosibirsk 630090, Russia

1. Introduction.

The problem of induction appears to be the most old and at the same time the most tangled one. It captured attention of famous philosophers from the time of Socrates. The paper [7] is devoted to general formal definition of induction methods considering it linked with the set of necessary requirements that according to their essence induction methods must satisfy to. The necessary requirements mean that failure to fulfill them leads to the loss of the sense of the inductive problem. But what is more, it has been proved that there are no induction methods satisfying them (except some vacuous case). This states a negative result about the existence of induction methods satisfying necessary requirements. Analogous negative results were obtained by different authors. The paper [8], for example, demonstrates negative result in a more usual and visual language in which the observation results are presented by points in the features space R^n.

Among the necessary requirements there is a requirement of linguistic invariance – the invariance of induction methods with respect to the formulating language (language to represent hypotheses and experiment results). The reinforcement of the hypothesis must depend on the observation results and initial hypothesis but not on the language it has been formulated in. Despite the requirement being obvious there are no induction methods satisfying it.

N. Goodman in his “new riddle of induction” was the first who had noticed the dependence of induction methods on their languages. However, he didn’t and, to be honest, he couldn’t come to the conclusion out of his own paradox only, that is the one obtained from negative result - about impossibility of invariant (satisfying to the linguistic invariance requirement) induction methods. Nevertheless, many authors devoted to the paradox have realized that it possesses community and radical character. In section 2 we are going to expound Goodman’s paradox in its traditional form and present corresponding opinions in order to find the roots of the problem blocking some branch roads that sufficiently lead to nowhere.

What is more, Goodman’s paradox is tuned out to be omnipresent, emerging in each step of hypothesis reinforcement. Analyzing the negative results we found out that they follow from contradictions, which appear to be just the generalizations of Goodman’s paradox. See section 3 for details.

From our point of view, the new stating of induction problem may be the following: induction is meaningless without the frame of some language/ontology. The language/ontology defines the world. Induction methods and predictions depend on this world.

From our side, we propose the positive way out of the Goodman’s paradox and negative results. Exactly the stating of induction problem was obtained for the class of universally quantified hypotheses. In [9] we formulated just such a property of experiment - inheritance of experimental results - and prove the theorem that if experiment procedure satisfies this property then experimental dependence is logically equivalent to the set of universal formulas. The property of experimental results inheritance is obviously fulfilled for the “world” of the most part of physics (the inheritance property means that if we have the observation result for the set of objects A as a model M, then the observation result of the subset $B \subset A$ will be submodel of the model M), but not for any other domain theories. The inheritance property – may be considered as the “second level” law of physics, which explain us why physical laws are universally quantified. This property of experiment, from one side, restricts the possible transformations of the language that allows to avoid the negative result and, from the other side, represents a class of hypotheses, which express the experimental dependence. Only hypotheses from this class need to be tested by the induction method. This makes the problem of inductive method’s development well founded.

In particular, inductive method “Discovery” developed [2,3], which may discover any set of universal formulas and has found successful application in different areas (see http://www.math.nsc.ru/LBRT/logic/vityaev Scientific Discovery website).

2. Goodman’s paradox and its discussion.

The following principle of inductive inference is considered to be the most incontestable: “If the predicate F applies to all objects $a_1, \dots, a_n$ tested so far then for sufficiently great n we have to count more on F applying to the next object a_{n+1} than on F not applying to it” [4]. By contrast, Goodman’s paradox appears to be the counter - example to this rule.

Consider emerald. It is significant, that sometimes emerald might look green, but sometimes – blue. Suppose that for all moments of time t_i, $i=1,\dots,n$ up to this moment it have been green. Let’s define a language L consisted of a single predicate Green. Thereby, predicate Green is applied to all the objects $a(t_1),\dots,a(t_n)$ up to day. This fact can be realized

with the formal protocol of observations $pr_L(a(t_1), \ldots ,a(t_n)) = \{Green(a(t_1)), \ldots ,Green(a(t_n))\}$. Thus according to the formulated inductive rule we might suppose that emerald will appear to be green at the next moment of time t_{n+1} too so as the predicate $Green(a(t_{n+1}))$ is true.

Let's define another language Q consisted of a predicate Grue (Green+Blue) that is true if emerald a(t) is green up to this moment, and blue at the subsequent moments of time, particularly at the time t_{n+1}. Observation protocol of the predicate Grue(a) is just like the one of the predicate Green(a), namely $pr_Q(a(t_1),\ldots,a(t_n)) = \{Grue(a(t_1)), \ldots , Grue(a(t_n))\}$. According to the rule been formulated relatively to this predicate we can conclude that it would be true at the next moment of time t_{n+1} and thus the predicate $Grue(a(t_{n+1}))$ would be true for the object $a(t_{n+1})$. But owing to the definition of the predicate Grue emerald $a(t_{n+1})$ ought to be blue, and thus, not to be green that is at variance with the predicate $Green(a(t_{n+1}))$ true for the emerald $a(t_{n+1})$. So, following to the simplest inductive rule we come to the contradicting predictions about the color of an emerald at the moment of time t_{n+1}. Green and Grue predicates are absolutely symmetric as in their syntactic features so in the observation protocols, and thus nobody is entitled according to any formal or objective reasons to prefer one predicate to another. Such preference is the preference of one language to interpret experiment and hypothesis to another.

Goodman's paradox or "the new riddle of induction" challenged too much philosophical debate that centered on Grue and other predicates providing the similar paradoxes. There have been forming several broad standpoints that had different opinions about the roots of the problem, about what is wrong with the paradoxical predictions. They led to different, sometimes conflicting, conclusions of the riddle and understanding the induction problem as the whole.

In the first considered standpoint people tried to discredit the paradox as a problem, i.e. to "dissolve" it by coming down that paradox is nothing but the play of words, or even the work of unscientific imagination that artificially introduced the predicate Grue. As the example you can take a paper [6]. But, you know, their conclusions seem to be rather unpersuasive to us.

Another school reasoned approximately in the same way. They also did not realize that Goodman's paradox is a kind of experimental embodiment of the complicated induction problem. So they suggested the "obvious solution" of the problem. People said that, of course, we must throw aside the Grue predicate, as it is not natural for us in a sense of our terminology and conceptions, as it does not satisfy our cognitive processes, our attitude of mind and thus it cannot be proclaimed as the "vital" predicate. On the other hand, we must accept the predicate Green as it is natural for us, as it satisfies the way we think and appears to be rather "vital", essential for our life.

Nevertheless, people have been trying to find the scientific reasons unlike the intuitive notion for the objective preference. Introducing the definitions of simplicity, stability and some kind, they tried to ground why Green predicate is the simplest, when Grue is "instable", "locational"…But Anthony Matthew in his "Prediction and Predication" [5] have strictly proved that "the paradox cannot be avoided simply by adopting rules for making predictions which exclude the use of Grue and similar artificial predicates." There is no such criterion that recognizes the simpler predicate.

Goodman himself within his "theory of projection" was trying to find the positive way out of the paradox. His idea is to associate with each hypothesis its degree of entrenchment and thus to find the predicate that is better entrenched (in our language and culture). The third school followed Goodman's theory in their proposals. Unfortunately, analyzing the theory of projection they have concluded that there is no "logical criterion" for defining the required feature of the hypothesis, what is more, there is no empirical criterion. Nevertheless, people accept Goodman's theory, but consider its need in thorough examination and further development.

At last, we present the fourth school that does not try to choose the predicate and propose that paradox solution lies in its prevention. The fundamental idea discussed by Mary Hesse [1] was that Grue seems so unbelievable to us unless we have "no radically new theory different from our physics… but in such a way common and usual." We do not have such Grue – theory as we do not need it or just think like this. Marry Hesse concludes that "there can be the cases of alternative sub-theories with equal inductive support, where we do not know which prediction to consider most reasonable."

Another example of the fourth conception is the paper of F. Von Kutschera [4]. Author proposes to extend all the existing models by introducing a-priori an agreement about the choice of "language to describe the world." He states, "Our beliefs about the world are founded on a priori assumptions that cannot be established by experience." That is why it is quite important to agree upon the language and thus to fix a-priori assumptions about the inductive method itself. Author concludes that the language must not possess invariant ability when transition from one world to another as "it is quite possible for several language systems to be equally successful but to lead to different theories about the world" and thus "there are no empirical reasons for choosing one language rather than another". We follow this company.

3. Goodman's paradox generalization.

Generally, inductive method is a function F that starting from the initial observation protocol pr_0 and hypothesis h_0 (possesses any general description of the experiment or admits everything) generates a new hypothesis h_1 that is consistent with pr_0 but it more (at any rate, not less) informative than h_0 in the sense of classifying protocols different from pr_0 as consistent or inconsistent with $h_1 : F(\langle h_0,\ pr_0 \rangle) = h_1$.

Definition of the inductive method is linked with the following necessary requirements: non – triviality, non – contradictoriness of the generated hypothesis to the initial data and universal application to the empirical research. The induc-

tion method f is said to satisfy the requirement of linguistic invariance with the respect to the class of homomorphism Φ and initial data $\langle h_0,\ pr_0 \rangle$, if for any homomorphism $\varphi \in \Phi$ the following condition is satisfied: $\varphi(F(\langle h_0,\ pr_0 \rangle)) = F(\langle \varphi(h_0),\ \varphi(pr_0) \rangle)$.

Linguistic invariance requirement is the necessary one directly linked with the definition as it means independence of the induction method on the language it is formulated in. The particular case of the linguistic transformations for induction methods is the invariance with the respect to the choice of measurement units.

Despite the requirement being obvious it is too strong to be accepted, as any induction method F does not satisfy it. Goodman's paradox demonstrates such negative result for the case of two languages, Green and Grue. What is more valuable, Goodman's paradox is tuned out to be omnipresent, emerging in each step of hypothesis reinforcement. Analyzing the negative results we found out that they follow from contradictions, which appear to be just the generalizations of Goodman's paradox. These contradictions have the following matter: for some language L and any initial hypothesis h_0, and protocol pr_0, and for any inductive method F there always can be constructed a language Q and translation $\varphi : L \leftrightarrow Q$, which induce one-to-one transformation of the protocol sets for these languages, such that diagram on fig 1 is fulfilled. In this diagram we have absolutely symmetric applications of inductive method F in different languages L and Q such that initial hypotheses are empirically equivalent, but the reinforced hypotheses are not.

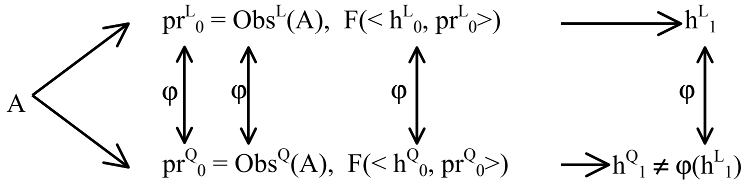

Fig. 1. Obs^L, Obs^Q – are experimental procedures, which for an object set A produce a protocol of experiment in two different languages L,Q; h^L_0,h^Q_0 - are empirically equivalent hypotheses; pr^L_0,pr^Q_0 - protocols obtained by applying the experimental procedure Obs^L to set of objects A. This protocols of experiment must satisfy hypotheses h^L_0, h^Q_0; h^L_1,h^Q_1 – reinforcements of hypotheses h^L_0,h^Q_0 by the inductive method F; $\varphi(h^L_1)$ – transformation of hypothesis h^L_1.

Despite this symmetry the resulting hypotheses h^Q_1, h^L_1 are different (when compared with respect to φ) and hypotheses h^Q_1, $\varphi(h^L_1)$ provide different predictions about experimental results.
Moreover, different predictions could be obtained for any protocol pr, which differentiate the hypotheses h^L_0 and h^L_1. More precisely, for reinforcement $F(< h^L_0, pr^L_0>) \rightarrow h^L_1$ of any hypothesis h^L_0 by the inductive method F in the language L ($h^L_1 \neq h^L_0$) and for any protocol pr, which differentiate hypotheses h^L_0 and h^L_1 (satisfy h^L_0 and falsify h^L_1) there can be constructed such language Q and translation φ that protocol φ(pr) would falsify hypothesis h^L_1,$\varphi(h^L_1)$, but not falsify h^Q_1. This statement clarifies the omnipresence of Goodman's paradox.
Traditionally induction and prediction are considered to be objective, but fig. 1 illustrates that they entirely depend on our own choice of the language, the point where it turns out to be subjective. Moreover, by choosing the language and experimental procedure for a protocol basically we may choose any prediction that we'd like. Goodman's paradox omnipresence judges us to a conclusion that the induction problem stating doesn't complete.

References

1. Hesse, M. [1969]. Ramifications of "grue". *The British journal for the philosophy of science*, 20:13-15.
2. Kovalerchuk B., Vityaev E. [2000]. Data Mining in Finance: Advances in Relational and Hybrid methods. (Kluwer international series in engineering and computer science; SECS 547), Kluwer Academic Publishers, p.308.
3. Kovalerchuk, B., Vityaev E., Ruiz J.F. [2001]. Consistent and Complete Data and "Expert" Mining in Medicine, In: Medical Data Mining and Knowledge Discovery, Springer, pp. 238-280.
4. Kutschera F. Von. [1973]. Induction and empiricist model of knowledge. In: Logic, Methodology and Philosophy of Science IV, eds.: P.Suppes et al, North-Holland Pub.Co., pp.345-356.
5. Matthew, A. [1971] Prediction and Predication. *The British journal for the philosophy of science,* 22:171–182.
6. Mulhall, S. [1989] No smoke without fire: The meaning of grue. *The Philosophical Quarterly*, 39:166-89.
7. Samokhvalov K. F. [1987]. Logics of Discovery. In: Inductive logic and scientific knowledge formation. Moscow, Science, pp.173-184.
8. Vityaev E.E., Novikov V.F. [1988]. Induction method paradoxicality. *International Journal of Pattern Recognition and Artificial Intelligence*, v.3, #1:147-157.
9. Vityaev E.E., Khomicheva I.V., Goodman's paradox generalizations. Proceedings of the 12th International Congress of Logic Methodology and Philosophy of Science, (12 International Congress of Logic, Methodology and Philosophy of Science Oviedo (Spain), August 7-13 2003), 2003, pp.146-147.