Представление знаний и проблема верификации

Андрей Родин

20 марта 2018 г.

О проекте

Мы называем знание *конструктивным*, если это знание включает в себя явно выраженную спецификацию по крайней мере некоторых эпистемических процедур, таких как:

- получение,
- верификация (обоснование),
- представление,
- распространение,
- ревизия,
- применение (использование)

данного знания.

Комментарий

Такое понятие конструктивного знания не исключает его реалистической интерпретации и тем более не предполагает точку зрения социального и/или когнитивного конструктивизма, согласно которой всякое знание является социальной (соотв. когнитивной) конструкцией, независимой от реального отношения к своему предмету.

О проекте

Целью предлагаемого исследовательского проекта является изучение формально-логических и эпистемологических аспектов конструктивных эпистемических процедур работы со знаниями, относящимися к широкому спектру научных и прикладных областей.

О проекте

Специальное внимание в этом проекте уделяется анализу современных и перспективных электронных технологий представления знаний и управления знаниями.

Участники:

- Владимир Васюков (ИФ РАН): категорная логика, философская логика, формальная эпистемология
- Сергей Ковалёв (ИПУ РАН): представление знаний в инженерии (в том числе с использованием категорных методов)
- Даниил Рогозин (магистрант МГУ): категорная логика, ГТТ, программирование на Хаскеле и Агде
- Андрей Родин (ИФ РАН, руководитель): формальная эпистемология
- Константин Шишов (аспирант МГУ): квантовые логики

Что такое знание?

Популярное определение:

знание это обоснованная истинная вера
(justified true belief)

Истинность (при классическом подходе): объективный аспект
Вера: субъективный аспект
Обоснование (доказательство, верификация): связывает истинность и веру

Проблема Гетье: корректность доказательства

Edmund Gettier 1963

Imagine that someone is standing outside a field looking at something that looks like a sheep but in fact, it is a dog disguised as a sheep. They believe there is a sheep in the field, and in fact, they are right because there is a sheep behind the hill in the middle of the field. Hence, they have a justified (?) true belief that there is a sheep in the field. But is that belief knowledge?

Проблема Гетье: корректность доказательства

Мой ответ: в данном случае вера НЕ обоснована (хотя и истинна), поскольку используемое доказательство (обоснование) некорректно. Формулировка проблемы Гетье использует чисто субъективное понятие об обосновании.

Открытая проблема логики: отсутствие хорошей теории доказательств применимой как в математике, так и в эмпирических науках.

Пропозициональное и операциональное знание

Знание высказываний (предложений) и знание правил (knowing what and knowing how)

Общее понятие о знании как обоснованной истинной вере предполагает что <u>всякое</u> знание является пропозициональным (либо сводимо к пропозициональному).

Знание логики (= знание логических правил и умение их применять) является операциональным, а не пропозициональным. Всякое научное включает такие операциональные элементы как методы обоснования (доказательства) и верификации предложений.

в математике 20-21 века

David Hilbert:

- *Grundlagen der Geometrie* (1899)
- *Axiomatishces Denken* (1918)
- *Neubegründung der Mathematik* (1922)
- *Grundlagen der Mathematik I-II* (1934-39, mit Paul Bernays)

две цели аксиоматизации в стиле Гильберта:

- представление теорий (стандартное представление готовых результатов)
- метаматематические исследования и основания

В 20-м веке эти два направления далеко разошлись. Аксиоматический метод в стиле Гильберта доказал свою эффективность в области метаматематики (и метаматематических "оснований"), но оказался мало пригодным для представления знаний в "практических" целях таких как научная коммуникация, математическое и научное образование и верификация результатов.

аксиоматический метод в современной математической практике:

- Бурбаки (1939 -)
- Унивалентные Основания (2010 -)

в естественных науках

- Шестая проблема Гильберта (1900)
- Эйнштейн указывает на аксиоматический метод Гильберта как на мотивировку теории относительности (1921)
- von Neumann, Mathematische Grundlagen der Quantenmechanik (1932)
- Woodger, Axiomatic Method in Biology (1937)

в философии науки

Ошибочно думать, что фон Нейман действительно построил аксиоматические основания квантовой механики. В его изложении отсутствуют все характеристики современной аксиоматики: он не предъявляет своих предпосылок, не указывает основные понятия своей теории, не перечисляет основные предположения (аксиомы), не предлагает последовательной физической интерпретации своего формализма. В его изложении много противоречий и наивных утверждений философского характера. И тем не менее, по какой-то странной причине эта работа считается образцом физической аксиоматики. (Bunge 1972)

два формальных подхода к формальной реконструкции научных теорий

- **синтаксический**: логические позитивисты: Carnap, Hempel, Nagel
- **семантический**: Beth, Suppes, van Fraassen

идея семантического подхода

Научная теория - это не аксиоматическая система, а класс моделей.

На практике: попытка использовать полу-формальный синтаксис в стиле Бурбаки за пределами чистой математики.

Суппес: математические модели в физике и других естественных науках допускают представление в виде теоретико-множественных моделей.

логика как общий метод науки

The constant and universal feature of science is its general method, which consists in the persisting search for truth, constantly asking: Is it so? To what extent is it so? Why is it so? [. . .] And this can be seen on reflection to be *the demand for the best available evidence*, the determination of which we call logic. Scientific method is thus the persistent application of logic as the common feature of all reasoned knowledge (Nagel и Cohen 1934).

формальные выводы это не всегда доказательства

The deduction makes no appeal whatsoever to experiment or observation, to any sensory elements. ... Whether anything in the world of existence conforms to this system requires empirical knowledge. ... That the world does exemplify such a structure can be verified only within the limits of the errors of our experimental procedure. (Nagel и Cohen 1934)

какую роль формальные выводы могут играть в обосновании эмпирических утверждений?

The evidence for propositions which are elements in a [hypothetico-deductive] system accumulates more rapidly than that for isolated propositions. The evidence for a proposition may come from its own verifying instances, or from the verifying instances of other propositions which are connected with the first [i.e., the given proposition] in a system. It is this schematic character of scientific theories which gives such high probabilities to the various individual propositions of a science.

промежуточный вывод

Логическая техника, которую использует Нагель, не решает тех задач, которые он ставит перед логикой в неформальном описании.

классическая vs. конструктивная логика

- истина как соответствие положению вещей (Т-схема) vs. истина как существование доказательства
- доказательство как формальный вывод из аксиом vs. доказательство как эффективно предъявляемое свидетельство
- эпистемологическая рефлексия $\Box P \to P$ vs. ко-рефлексия $P \to \Box P$

Общая теория доказательств

In model theory, one concentrates on questions like what sentences are logically valid and what sentences follow logically from other sentences. But one disregards questions concerning how we know that a sentence is logically valid or follows logically from another sentence. General proof theory would thus be an attempt to supplement model theory by studying also the evidence or the process - i.e., in other words, the proofs - by which we come to know logical validities and logical consequences. Prawitz 1974

Теоретико-модельная и теоретико-доказательная семантика

(Мета-теоретическое) отношение семантического следствия ($A_1, \ldots, A_n \models B$: во всех интерпретациях, в которых истинны $A_1, \ldots, A_n$ также истинно B) не годится в качестве семантики для правил вывода. Сохранение истинности - это необходимый, но не достаточный критерий валидности вывода (если потребовать, что всякий валидный вывод может быть шагом корректного доказательства). Обоснование правила вывода логической системы не должно зависеть от истинности каких-либо высказываний. Истинность должна определяться в терминах правил (и доказательств основанных на этих правилах), а не наоборот.

"Объяснение значения" для правил логического вывода

Интерпретация логического синтаксиса как компиляция компьютерной программы (Мартин-Лёф)

Два аксиоматических стиля

- аксиоматизация в стиле Гильберта: много аксиом и мало или одно правил(о)
- аксиоматизация в стиле Генцена: много правил и мало или ноль аксиом

С конструктивной точки зрения стиль Генцена имеет эпистемологическое преимущество. Этот стиль имеет также и практическое преимущество, поскольку такая формализация упрощает вычислительную реализацию на компьютере.

Стиль Генцена за пределами чистой логики

Сам Генцен использовал свой аксиоматический стиль только в чистой логике. Унивалентные Основания используют стиль Генцена для формализации математики. Мы исследуем возможности использования этого стиля для широкого круга задач представления знаний, рассматривая системы формальных правил в качестве инструмента для представления операционального знания. Гомотопическая семантика ГТТ дает ключ к пониманию связи операционального знания с пропозициональным.

MLTT: Syntax

- 4 basic forms of judgement:
  (i) $A : TYPE$;
  (ii) $A \equiv_{TYPE} B$;
  (iii) $a : A$;
  (iv) $a \equiv_A a'$
- Context : $\Gamma \vdash$ judgement (of one of the above forms)
- no axioms (!)
- rules for contextual judgements; Ex.: dependent product :
  If $\Gamma, x : X \vdash A(x) : TYPE$, then $\Gamma \vdash (\Pi x : X)A(x) : TYPE$

MLTT: Semantics of $t : T$ (Martin-Löf 1983)

- t is an element of set T
- t is a proof (construction) of proposition T ("propositions-as-types")
- t is a method of fulfilling (realizing) the intention (expectation) T
- t is a method of solving the problem (doing the task) T (BHK-style semantics)

Sets and Propositions Are the Same

If we take seriously the idea that a proposition is defined by lying down how its canonical proofs are formed [. . .] and accept that a set is defined by prescribing how its canonical elements are formed, then it is clear that it would only lead to an unnecessary duplication to keep the notions of proposition and set [. . .] apart. Instead we simply identify them, that is, treat them as one and the same notion. (Martin-Löf 1983)

MLTT: Definitional aka judgmental equality/identity

$x, y : A$ (in words: x, y are of type A)

$x \equiv_A y$ (in words: x is y by definition)

MLTT: Propositional equality/identity

$p : x =_A y$ (in words: x, y are (propositionally) equal as this is evidenced by proof p)

Definitional eq. entails Propositional eq.

$$\frac{x \equiv_A y}{p : x =_A y}$$

where $p \equiv_{x =_A y} refl_x$ is built canonically

Equality Reflection Rule (ER)

$$\frac{p : x =_A y}{x \equiv_A y}$$

ER is <u>not</u> a theorem in the (intensional) MLTT (Streicher 1993).

Extension and Intension in MLTT

- MLTT + ER is called *extensional* MLTT
- MLTT w/out ER is called *intensional*
  (notice that according to this definition intensionality is a negative property!)

Higher Identity Types

- $x', y' : x =_A y$
- $x'', y'' : x' =_{x =_A y} y'$
- $\ldots$

HoTT: the Idea

Types in MLTT are (informally!) modeled by spaces (up to homotopy equivalence) in Homotopy theory, or equivalently, by higher-dimensional groupoids in Category theory (in which case one thinks of n-groupoids as higher homotopy groupoids of an appropriate topological space).

Обратите внимание на *внелогическую* интерпретацию логического понятия тождества! В аксиоматических теориях в стиле Гильберта различение логических и внелогических символов жестко фиксировано заранее и интерпретации подлежат *только* внелогические символы: ср. понятие сигнатуры у Бурбаки!

Homotopical interpretation of Intensional MLTT

- $x, y : A$
  x, y are points in space A
- $x', y' : x =_A y$
  x', y' are paths between points x, y; $x =_A y$ is the space of all such paths
- $x'', y'' : x' =_{x =_A y} y'$
  x'', y'' are homotopies between paths x', y'; $x' =_{x =_A y} y'$ is the space of all such homotopies
- ...

Point

Definition

Space S is called contractible or space of h-level (-2) when there is point $p : S$ connected by a path with each point $x : A$ in such a way that all these paths are homotopic (i.e., there exists a homotopy between any two such paths).

Homotopy Levels

Definition

We say that S is a space of h-level $n+1$ if for all its points x, y path spaces $x =_S y$ are of h-level n.

Cummulative Hierarchy of Homotopy Types

- -2-type: single point pt;
- -1-type: the empty space $\emptyset$ and the point pt: truth-values aka (mere) propositions
- 0-type: sets: points in space with no (non-trivial) paths
- 1-type: flat groupoids: points and paths in space with no (non-trivial) homotopies
- 2-type: 2-groupoids: points and paths and homotopies of paths in space with no (non-trivial) 2-homotopies
- ...

Propositions-as-**Some**-Types !

Which types are propositions?

Def.: Type P is a *mere proposition* if $x, y : P$ implies $x = y$ (definitionally).

Truncation

Each type is transformed into a (mere) proposition when one ceases to distinguish between its terms, i.e., *truncates* its higher-order homotopical structure.

Interpretation: Truncation reduces the higher-order structure to a single element, which is **truth-value**: for any non-empty type this value is **true** and for an empty type it is **false**.
The reduced structure is the structure of **proofs** of the corresponding proposition.
To treat a type as a proposition is to ask whether or not this type is instantiated without asking for more.

- Thus in HoTT "merely logical" rules (i.e. rules for handling propositions) are instances of more general formal rules, which equally apply to non-propositional types.
- These general rules work as rules of building models of the given theory from certain basic elements which interpret primitive terms (= basic types) of this given theory.
- Thus HoTT qualify as *constructive* theory in the sense that besides of propositions it comprises non-propositional objects (on equal footing with propositions rather than "packed into" propositions as usual!) and formal rules for managing such objects (in particular, for constructing new objects from given ones). In fact, HoTT comprises rules with apply *both* to propositional and non-propositional types.

Univalence

$$(A =_{TYPE} B) \simeq (A \simeq B)$$

Словами: эквивалентность типов эквивалентна их равенству. Структурализм?

For PROPs: $(p = q) \leftrightarrow (p \leftrightarrow q)$ (propositional extensionality)

For SETs: Утверждения об изоморфных множествах логически эквивалентны (изморфизм-инвариантность)

Univalence implies *functional extensionality*: if for all $x \quad X$ one has $fx =_Y gx$ then $f =_{X \to Y} g$ (the property holds at all h-levels).

Models of HoTT after Voevodsky

(1) Construct a general model of given type theory **T** (MLTT or its variant) as a category $\mathcal{C}$ with additional structures which model **T**-rules. For that purpose the authors use the notion of *contextual category* due to Cartmell [?]; in later works Voevodsky uses a modified version of this concept named by the author a *C-system*.

Models of HoTT after Voevodsky

(2) Construct a particular contextual category (variant: a C-system) $\mathcal{C}(\mathbf{T})$ of syntactic character, which is called *term model*. Objects of $\mathcal{C}(\mathbf{T})$ are MLTT-contexts, i.e., expressions of form

$$[x_1 : A_1, \ldots, x_n : A_n]$$

taken up to the definitional equality and the renaming of free variables and its morphisms are substitutions (of the contexts into $\mathbf{T}$-rule schemata) also identified up to the definitional equality and the renaming of variables). More precisely, morphisms of $\mathcal{C}(T)$ are of form

Models of HoTT after Voevodsky

$$f : [x_1 : A_1, \dots, x_n : A_n] \to [y_1 : B_1, \dots, y_m : B_m]$$

where f is represented by a sequent of terms $f_1, \dots, f_m$ such that
$x_1 : A_1, \dots, x_n : A_n \vdash f_1 : B_1$
$\vdots$
$x_1 : A_1, \dots, x_n : A_n \vdash f_m : B_m(f_1, \dots, f_m)$
Thus morphisms of $\mathcal{C}(T)$ represent derivations in **T**.

Models of HoTT after Voevodsky

- Define an appropriate notion of morphism between contextual categories (C-systems) and form category $CTXT$ of such categories.
- Show that $\mathcal{C}(\mathbf{T})$ is initial in $CTXT$, that is, that for any object $\mathcal{C}$ of $CTXT$ there is precisely one morphism (functor) of form $\mathcal{C}(\mathbf{T}) \to \mathcal{C}$.

The last item is the **Initiality Conjecture** that presently stands open.

- Оказывают ли исследования в области философской логики и формальной эпистемологии какое-либо влияние на теорию и практику компьютерного представления знаний? Предварительный ответ положительный. Формальные онтологии прочно вошли в арсенал компьютерных наук и, на наш взгляд, теперь очередь за формальной эпистемологией.
- Как совместить конструктивный подход в логике и математике с научным реализмом? Предварительный ответ: с помощью теории факторов истины (truth-maker realism).