Проблема обоснования в формальном представлении знаний.

1) Введение

Выражение “представление знаний” (Knowledge Representation, KR) - в компьютерных науках используется как название раздела исследований в области искусственного интеллекта (AI); в последние годы для этой же области исследований также широко используется расширенный названия “представление знаний и рассуждения” (Knowledge Representation and Reasoning).

“Представление знаний и рассуждения это область исследований искусственного интеллекта (AI), которая изучает способы символического представления знаний и автоматической символической манипуляции знаниями с помощью рассуждающих программ. Говоря менее формально, это область AI, которая имеет дело с мышлением и с ролью мышления в разумном поведении.” (Brachman & Levesque, Foreword ??)

Для философа исследования в области представления знаний и автоматизированных рассуждений представляют особый интерес, поскольку мышление, знание и рассуждения являются такой предметами традиционной философской дисциплины как эпистемология (которую также иногда называют теорией познания). Поэтому представляется естественной попытка сравнить, что говорят о знаниях, рассуждениях и их символических представления представители компьютерных наук, с одной стороны, и философы, с другой стороны.

2) Понятие знания в философской эпистемологии.

В центре современной эпистемологической дискусси (по крайней мере в Аналитической философской традиции) находится восходящая к Платону так называемая JTB концепция знания, согласно которой знание - это обоснованное истинное мнение (Justified True belief). Это определение расшифровывается так: эпистемический агент *A* знает, что *p* (где *p* это некоторая пропозиция) если и только если выполнены следующие три условия:

- пропозиция p истинна (является истинным высказыванием)
- агент A верит, что p
- A имеет достаточное основание g, для того, чтобы верить, что p.

Если речь идет об общем знании, то агент A может специально не указываться: в этом случае говорят, что “p известно”.

Отметим особо три важных для дальнейшего момента этого определения.
а) В рамках JTB теории предполагается, что предмет знания в общем случае - это пропозиция. Такое знание называют пропозициональным знанием или знанием-что (knowledge that).

б) Согласно JTB теории знание пропозиции p отличается от самой этой пропозиции. В формальных системах эпистемической логики, которые ставят своей целью представить рассуждения о знаниях в символической форме, знание пропозиции p обычно рассматривается в качестве особой (а именно, эпистемической) модальности.

в) В рамках JTB теории знание пропозиции p включает в себя два условия, которые не могут быть определены, если пропозиция рассматривается в абстракции от любых познавательных процессов, как это делается в классической логике. А именно, знание пропозиции p, во-первых, предполагает наличие эпистемического субъекта (индивидуального или коллективного) и, во-вторых, требует, чтобы эта вера (мнение) была обоснованной.

Последнее требование, а именно требование обоснования, является самым проблематичным моментом JTB теории. Широко обсуждаются так называемые “проблемы типа Гетье” названные по имени современного американского философа Эдмунда Гетье, Edmund Gettier. Пусть Петр утверждает, что у него есть магический кристалл, который позволяет ему (и никому другому) предсказывать будущее. Желая испытать действие кристалла Иван просит Петра предсказать исход бросания монеты, которая находится в его руках. Петр консультируется с кристаллом и говорит, что выпадет решка. Иван бросает монету и выпадает решка. Можно ли утверждать, что Петр действительно <u>знал</u> заранее на какую сторону упадет монета?
Здравый смысл говорит о том, что речь в описанном случае может идти о случайном совпадении: если называть один из двух возможных исходов (орел или решка) наугад, то вероятность правильного предсказания составит пятьдесят процентов. Поэтому без дальнейшей аппробации прогностических возможностей Петра и его кристалла говорить о том, что Петр обладал знанием о будущем событии в данном случае кажется некорректным. Однако условия, которые устанавливает JTB теория могут быть в данном случае удовлетворены: Петр делает истинное утверждение о будущем событии, верит, что оно произойдет (допустим это) и, наконец, приводит обоснование для своей веры (мнения). Таким образом JTB теория оказывается не в состоянии отличить знание от догадки, которая случайно оказалась верной.

Проблема Гетье указывает на то, что понятие обоснования нуждается в дальнейшем уточнении; в частности, необходимы какие-то критерии по которым можно было бы оценивать предложенное обоснование данного утверждения и либо признавать его валидным, либо отбрасывать как негодное. Это проще сделать в формальном контексте, где можно построить точное понятие доказательства, которое здесь играет роль обоснования. Впрочем, как мы увидим, сама по себе формализация не решает проблемы, поскольку формализовать понятие доказательства можно, вообще говоря разными способами, получая при этом, строго говоря, различные формализованные понятия. Польза от такой формализации как раз и состоит в том, что она позволяет проводить различия более точно.

Другая трудность JTB теории как общей теории знании состоит в том, что эта теория предполагает, что всякое знание является пропозициональным. Между тем слово “знание” в естественном языке также используется в контекстах другого рода. В предложении “Иван знает Петра” речь не идет о том, что Иван знает какие-то истинные пропозиции, а скорее о том, что он встречался с Петром и что у него есть какой-то опыт общения с этим человеком. Знание такого рода называют “знанием по знакомству” (knowledge by aquaintance) и мы не будем его тут рассматривать. Предложение “Петр знает как делать яичницу” можно перефразировать “Петр умеет делать яичницу” речь тоже не идет о пропозициях. Такое знание в эпистемологии называют процедурным знанием или знанием-как (knowledge how). Знания этого типа играют существенную роль в формальных контекстах. Например, утверждение о том, что Петр *умеет* доказывать теорему Пифагора (а не только знает математический факт, который утверждается в этой теореме) содержит, prima facie, ссылку на знание-как. В современной эпистемологии ведется полемика по вопросу о том, можно ли знание-как в общем случае свести к знанию-что или нет. Мы предполагаем, что ответ на этот вопрос отрицательный. Если это так, то JTB следует отбросить как общую теорию знания, но ее можно сохранить как специальную теорию пропозиционального знания. В последнем разделе этой статьи мы опишем формальную символическую систему представления знаний, которая следует JTB в своем пропозициональном сегменте, но в которой процедурное знание также играет фунаментальную роль.

3) Понятия знания и рассуждения в компьютерных науках.

Компьютерная наука в отличие от философской эпистемологии, разумеется, не ставит своей целью построить общую теорию знаний и рассуждений. Тем не менее во многих монографиях по представлению знаний, особенно тех, которые задуманы как введение в эту область знаний, делаются попытки неформального определения базовых терминов включая “знание” и “рассуждение”. В частности, в монографии “Интеллектуальные системы “ 2011-го года эти базовые понятия определяются так:

“*Что такое знание?* В очень общем смысле знание можно определить как информацию (которую можно представить в виде пропозиций) об окружающей среде.
Что такое представление знаний?
Попросту говоря, представление знаний это символы, которые используются для того, чтобы представить эти пропозиции.
Что такое представление знаний и рассуждения? Это можно определить как манипуляцию символами, которые кодируют пропозиции, с целью построить представления новых пропозиций. “

Как мы видим, это определение знания вовсе не включает в себя упоминания эпистемического агента и по сути не проводит никакого различия между пропозицей p и знанием пропозиции p (каким-либо агентом). Некоторые другие подобные определения, которые нам удалось обнаружить, вводят понятие об эпистемическом агенте указывают на различие между процедурным и пропозициональным знанием (Брахман и Левеск ??), но ни одно из них не включает требование обоснования.

Вместе с тем проблема обоснования знаний не является чисто теоретической и тем более чисто философской. Ненадежность распространяемой и получаемой через электронные средства коммуникации информации в последние годы стала социальной и даже политической проблемой. Публицисты в этой связи говорят о наступлении “эпохи пост-правды” и ряд философов склонны воспринимать такие слова всерьез и ставить под вопрос понятие истины.

Ревизия фундаментальных понятий философской логики включая понятие истины всегда полезна и составляет важную и необходимую часть философской работы независимо от мотиваций, но в данном случае, на наш взгляд, речь идет о недостатке существующих компьютерных и коммуникационных технологий, который уже хорошо осознан сообществом и должен быть исправлен. Хотя в рамках настоящей статьи мы не предлагаем никакого конкретного технического решения проблемы обоснования, мы анализируем истоки этой проблемы и показываем в следующем разделе статьи, что эта проблема имеет фундаментальный аспект, связанный с характером стандартной теории логической семантики, которая широко используется в качестве теоретического прототипа для архитектуры существующих систем представления знаний. Далее мы описываем новые подходы в логической семантике, которые, на наш взгляд, могут служить теоретической основой для будущих технических решений.

4) Онтология, эпистемология и формальная логическая семантика

В своей последней статье с говорящим названием “Пренебрежение эпистемологическими соображениями в логике” ?? Горан Зундолм показывает, что в мэйнстриме развития философской логики 20-го века эпистемологические соображения систематически игнорировались Одновременно Зундольм указывает на логическую традицию, для которой эпистемологические соображения, напротив, всегда были центральными. Речь идет о конструктивной логике, которая представляет для нас в этой статье особенный интерес, поскольку она более тесно, чем классическая логика, связана с вычислительной проблематикой. Начиная по крайней мере с 1950х годов исследователи развивали конструктивную логику и математику имея в виду перспективу приложений данной теории к разработке вычислительных машин и разработке компьютерных программ. В течении последнего десятилетия интерес логически ориентированных философов к конструктивному направлению резко усилился в связи с открытием новых формальных подходов, о которых мы скажем чуть ниже.

Опираясь на анализ Зундольма, мы утверждаем, что “пренебрежение” проблемой обоснования в теории и практике компьютерного представления знаний связана с тем, что создатели концепции искусственного интеллекта и их последователи до сих пор ориентировались на системы классической логики и логической семантики, в которых эпистемологические соображения являются второстепенными по сравнению с онтологическими соображениями. Сильным свидетельством в пользу этого тезиса является тот факт, что методы формальной *онтологии* успешно перекочевали из философии в компьютерную науку (и во многом потеряли всякую связь с философией), но подобного не произошло с методами формальной *эпистемологии*. Впрочем, нужно иметь в виду, что формальная эпистемология является более молодой дисциплиной, чем формальная онтология, и поэтому до недавнего времени разработчики систем искусственного интеллекта и представления знаний просто не могли заимствовать такого рода методы у философов.

4.1) *Отношение логического следования и логический вывод*

Поясним теперь чуть более подробно, в каком именно смысле можно говорить о том, что классические логические теории созданные в 20-м веке пренебрегают эпистемологической проблематикой. Простое отождествление рассуждения с некоторой синтаксической процедурой (манипуляцией символами) будет абсурдным и бессодрежательным, если не объяснить, как и почему данная синтаксическая процедура символическим образом выражает некоторое рассуждение. Семантическая теория, которая отвечает на этот вопрос, была предложена Альфредом Тарским в классической работе “О понятии логического следствия” 1936-го года ??. Вместо того, чтобы попытаться придать синтаксической процедуре порождения новых формул из данных формул значение логического вывода напрямую, Тарский определяет отношение *логического следствия* следующим образом:

Формула B логически следует из совокупности формул $A_{1}, \dots, A_{n}$, \\ в символах $A_{1}, \dots, A_{n} \models B$, \\

если всякая интерпретация теории, при которой формулы $A_{1}, \dots, A_{n}$ интерпретируются как истинные пропозиции, также интерпретирует формулу B как истинную пропозицию.

Напомним, что интерпретацией формулы в этом контексте называют такой способ приписать входящим в эту формулу *нелогическим* символам некоторых значений (которые могут быть определены в рамках другой теории или рассматриваться неформально), который превращает данную формулу в предложение, имеющее определенное истинностное значение. Интепретация, которая интерпретирует формулу как истинное предложение, называется моделью этой формулы. Схематическая интерпретация всех формул данной формальной теории, которая превращает эти формулы в истинные предложения, называется *моделью* этой теории.

Определенное выше понятие логического следствия (которое также называют семантическим следствием) дает частичный ответ на вопрос о семантике логических символов теории. Полностью этот вопрос мы рассматривать в рамках данной статьи не будем, но заметим, что другие логические понятия как и понятие логического следствия определяются при рассматриваемом подходе в терминах истинностных значений предложений. Поскольку интерпретация и истинностная оценка формул данной теории, вообще говоря, определяется вне рамок этой теории, отношение логического следования в смысле приведенного выше определения является мета-теоретическим.

Тарский использует понятие логического следования в качестве эрзаца понятия логического *вывода* , не определяя это последнее понятие формально. Стандартные логические исчисления обладают свойством *корректности* (soundness): если формула B выводится (то есть порождается в согласии с синтаксическими правилами дедукции) из формул $A_{1}, \dots, A_{n}$, в символах это принято записывать так: $A_{1}, \dots, A_{n} \vdash B$, то B логически следует из и формул $A_{1}, \dots,

A_{n}. Таким образом, получается, что всякая формула B выведенная из формул (предпосылок) $A_{1}, \dots, A_{n}$ по чисто синтаксическим правилам (“манипуляция символами”) оказывается логическим следствием этих предпосылок. Это позволяет думать о синтаксическом выводе (дедукции) как о логическом выводе, то есть процедуре, которая позволяет установливать логические следствия из данных предпосылок. Такую семантику вывода (основанную на отношении семантического следствия) называют *теоретико-модельной*, чтобы отличить ее от *теоретико-доказательной* семантики, о которой мы скажем чуть позже.

В случае когда синтаксический вывод осуществляется алгоритмически с помощью вычислительного устройства, это позволяет говорить метафорически об “автоматическом рассуждении” и “рассуждающих программах”, как это часто делается в компьютерной научной и публицистической литературе. Однако заметим, что соответствующие понятия логического вывода и рассуждения напрямую никак не связаны с проблематикой знания. Отношение логического (семантического) следствия по Тарскому - это отношение между формулами, которое формулируется в терминах истинных предложений и их классов. Классическое понятие истины, которое здесь используется (и которое может в свою очередь уточняться различными формальными способами) предполагает, что истинность предложения никак не зависит от того, знает ли какой-либо это предложение или нет. Вслед за Зундольмом мы можем сказать, что такая логико-семантическая схема систематически игнорирует эпистемологическую проблематику.

Могут возразить, что в нашем кратком описании классической логической семантики мы просто забыли упомянуть о таком важном эпистемологически нагруженном понятии как *доказательство.* Этим словом Гильберт и Бернайс стали называть синтаксические выводы, то есть цепочки формул, которые включают в себя ряд предпосылок $A_{1}, \dots, A_{n}$ (которые могут иметь статус *аксиом* теории) из которых в соответствии с правилами дедукции данной теории выводится некоторая искомая формула B: такую цепочку вслед за Гильбертом и Бернайсом часто называют “доказательством формулы B на основании предпосылок (аксиом) $A_{1}, \dots, A_{n}$.

Такие формальные доказательства представляют собой сложные синтаксические объекты, иследование которых математическими методами составляет предмет классической *теории доказательств* (Beweisstheorie). Однако как убедительно, на наш взгляд, аргументирует Правиц ?? и его последователи, понятие доказательства как синтаксического объекта на самом деле является техническим, и его связь с эпистемологией - а также с тем, что принято называть доказательствами или обоснованиями в математике за пределами математической логики и тем более в юриспуденции и других областях человеческой практики связанных с аргументацией - остается по меньшей мере неясной. Аксиматизация теории при таком подходе представляет собой способ удобного представления этой теории в компактном виде, который включает в себя небольшой набор аксиом и формальные правила дедукции, которые позволяют получить из этих аксиом все остальные предложения данной теории (число которых обычно потенциально бесконечно). Формальные синтаксические “доказательства” теорем теории при этом не выполняют (или во всяком случае не обязаны выполнять) функцию обоснования или доказательства в обычном смысле; эти синтаксические конструкции являются ключевыми элементами более общего синтаксического механизма упорядочивания и структуризации теории, которая превращает ее в более компактное и более обозримое целое.

Несмотря на свое название, теория доказательств в духе Гильберта и Бернайса не ставит и не решает никаких эпистемологических задач. Поэтому указание на синтаксическую теорию доказательств в духе Гильберта-Бернайса не отменяет вывода Зундолма о “пренебрежении эпистемологическими соображениями в логике” 20-го века; сама по себе такого рода теория не может помочь нам решить проблему обоснования в компьютерном представлении знаний.

4.2) *Истина, существование и реализм факторов истины*

Теоретику-модельную логическую семантику в духе Тарского можно простым образом связать с онтологией, то есть теорией о том, какого рода сущности существуют (“бывают”) в мире вообще или по крайней мере в некоторой интересующей нас части мира. Для этого достаточно принять следующий тезис:

Для всякого истинного высказывания *существует* некоторая вещь (или вещи), которая делает это высказывание истинным.

Такую вещь по отношению к данному высказыванию в современной философской логике и аналитической метафизике принято называть *фактором истины* (truth-maker) этого высказывания, а приведенный выше тезис - тезисом *реализма факторов истины* (truth-maker realism, TMR).

TMR не решает вопроса о том, каков фактор или факторы истины каждого конкретного истинного предложения, например математического предложения 2+3=5. Однако, простым (хотя и не обязательно правильным) ответом на вопрос о факторах истины предложения 2+3=5 будет такое: это предложение истинно, поскольку существуют числа, которые мы обозначаем символами “2”, “3” и “5” и эти три вещи действительно находятся в трехместном отношении, которое выражено с помощью привычных математических символов. Если речь идет о формулах первопорядковой теории, а не о формулах элементарной неформальной арифметики, как в предыдущем примере, то подобные “естественные” онтологические допущения включают в себя утверждение о существовании (в онтологическом смысле) домена (универсума) индивидов и связанных с ним конструкций (под-домены и другие) подходящих для интерпретации нелогических символов этой теории. Например, стандартная первопорядковая аксиоматическая теория множеств ZF может быть таким образом дополнена и фундирована утверждением о том, что аксиомы и теоремы этой теории истинны благодаря тому, что в мире существуют такие вещи как множества, которые обладают именно теми свойствами и находятся именно в тех отношениях, о которых нам говорят предложения этой теории.

Такой общий подход позволяет заимстововать из логической семантики и использовать в онтологии формальные символические методы, что и делается в *формальной* онтологии, которая в настоящее время яляется не только философской, но и компьютерной дисциплиной, имеющей приложения в искусственном интеллекте вообще и в первую очередь в области (компьютерного) представления знаний. Разумеется, в задачи компьютерной онтологии не входит выяснение различия между бытием и существованием и другие подобные сугубо философские сюжеты; задачей таких онтологий состоит в том, чтобы схематическим образом фиксировать и на базовом уровне описать ту предметную область, к которой относится некоторое представляемое знание.

Использование философских подходов наработок в теории и практике компьютерного представления знаний можно только приветствовать, но в данном случае вызывает недоумение тот факт, что для решения задач представления знаний компьютерная наука заимствует из философии именно онтологические, но не эпистемологические подходы. Как мы постарались показать, это обстоятельство не является случайным; оно связано с особенностями исторического развития логики в 20-м веке и с тем, что знание в компьютерной науке отождествляется со структурированной определенным образом информацией. Как мы уже говорили, такое отождествление не только представляется неправомерным с точки зрения философской эпистемологии, но и имеет вполне ощутимые негативные практические последствия.

5) Формальная эпистемология и конструктивная логика

Как нужно изменить стандартную логическую архитектуру, описанную выше, чтобы учесть эпистемологическую сторону дела и использовать эту новую архитектуру в системах компьютерного представления знаний? Вспомним три элемента JTB теории знания: истинные пропозиции, вера эпистемического агента (пользователя системы представления знаний) в эти пропозиции и обоснование этой веры. Можно рассуждать так: концептуальную схему для обработки истинных пропозиций мы уже имеем в стандартной архитектуре, поэтому нужно добавить к ней недостающие два элемента. Такой подход отвечает философской идее о том, что эпистемология является второстепенной дисциплиной по отношению к онтологии: онтология и связанная (посредством тезиса о реализме факторов истины) с ней логика дает нам общую теорию истинных пропозиций, а эпистемологии остается разобраться с тем, какие из этих истин известны или могут быть известны, кому именно известны и каким образом известны.

В рамках такого классического подхода строятся различные системы эпистемической логики как модального расширения классической логики, которые мы не будем более подробно рассматривать в этой статье. Фундаментальная проблема, которая делает подобные подходы мало применимыми на практике, состоит в том, что принимая тезис о первичности онтологии по отношению к эпистемологии, мы оказываемся вынуждены оценивать ситуацию с точки зрения Всезнающего Существа, которому заранее известны все истинные пропозиции. В компьютерных системах представления знаний роль такого Высшего Существа может играть Инженер такой системы: именно он должен гарантировать, что исходные данные используемые в данной системы надежны, а обработка этих данных устроена так, что Пользователь, не имея к этим данным прямого доступа и не зная деталей их обработки, тем не менее всегда получает по своему запросу достоверную информацию. Однако существующие системы представления знания ни имеют единого создателя: они являются плодом работы больших распределенных коллективов людей, которые организованы скорее сетевым, а не строго иерархическим способом. Это лишний аргумент в пользу того, что подобный подход к знаниям не адекватен существующим информационным технологиям и коммуникационным практикам: возможность обоснования и проверки знаний должна быть предоставлена каждому пользователю, а не быть эксклюзивным правом какой-то узкой группы технических специалистов.

В научных исследованиях, юридической практике и многих других областях человеческой деятельности имеет место обратная ситуация: никакое высказывание не признается истинным без досточного для этого основания. Такой ситуации отвечает иная концепция логики, в рамках которой логика, по словам Мориса Когена и Эрнста Нагеля, яляется инструментом для определения “лучшего доступного свидетельства” (best available evidence) ??. В рамках такой концепции логики логический вывод, истина и другие базовые логические понятия имеют эпистемологическую нагрузку. Говоря о “пренебрежении эпистемологическими соображениями в логике”, Зундольм в качестве альтернативы текущей ситуации указывает именно на эту логическую традицию, которую сегодня по ряду причин (которые станут ясными из дальнейшего) называют *интуиционистской* или *конструктивной*. Чисто формально классическую и интуицинистскую логику различают по следующему критерию: в классической логике действует правило исключенного третьего (то есть дизъюнкция $P \vee \neg P$ является тавтологией), а в интуиционистской логике нет. Однако мы здесь говорим не только и не столько о формальных, сколько о семантических, концептуальных и даже “идеологических” различиях, и прежде всего - о разных отношениях к эпистемологической проблематике.

Сделаем теперь важное терминологическое уточнение, без учета которого любое сравнение конструктивной и классической логики немедленно приводит к путанице. Термины “истина” и “доказательство” понимаются в этих двух случаях по-разному. В классической логике (снабженной описанной выше стандартной семантикой) истинностное значение высказывания задается вместе с интерпретацией соответствующей формулы и никак не связано в доказательством (даже если под доказательством понимать синтаксический вывод). В конструктивной логике и математике высказывание считается истинным, если у него есть доказательство, и ложным в противном случае. Доказательство в этом случае понимается более широко, чем в классическом случае: точный смысл понятия доказательства необходимо каждый раз уточнять. В частности, доказательствами (свидетельствами) при таком подходе могут считаться напрямую факторы истины (в частности, материальные предметы). Чтобы отличить теорию доказательств в духе Гильберта и Бернайса, которая изучает только синтаксическую структуру дедукций, от теории доказательств в более общем смысле, которая принимает в расчет и проблематизирует эпистемологические аспекты понятия доказательства (и понимает доказательство шире), теорию доказательств в последнем смысле называют сегодня *общей* (general proof theory).

5.1) *Теоретико-доказательная семантика*

Идея теоретико-доказательной семантики синтаксической дедукции как альтернативы для (или дополнения к) теоретико-модельной семантики, которая была описана выше, состоит в том, чтобы синтаксическим процедурам и правилам вывода одних формул из других поставить в соответствие подробно эксплицированные эпистемологические процедуры, которые отвечают неформальному полнятию логического вывода. Если в рамках теоретико-модельной семантики понятие истины (в модели) принимается в качестве базового, а понятие логического вывода (посредством метатеоретического отношения логического следования) определяется с помощью условий истинности (truth conditions), то в рамках теоретико-доказательной семантики понятие логического вывода определяется путем формальной спецификации требования эпистемической прозрачности. Поскольку истинность при таком подходе понимается как существование доказательства, а логический вывод является элементом доказательства, нельзя сказать, что при таком подходе понятие логического вывода никак не зависит от истинности. Однако в этом случае понятие истины не является заранее предположенным; можно сказать, что при построении теоретико-доказательной семантики логического вывода понятие истины включает в себя понятие вывода, а не наоборот.

Теоретико-доказательный подход в формальной семантике накладывает новые условия на синтаксис теории, аналогов которым нет при теоретико-модельном подходе. Технические подробности можно найти в ?? Использование такого рода формальной семантики в компьютерном представлении знаний можно представить себе в следующем виде. Представим себе информационную систему, которая не просто выдает пользователю в удобном виде некоторую полезную информацию по его запросу, осуществляя при этом поиск и обработку доступных для нее данных, но и предоставляет пользователю доступную для его понимания и анализа запись этой процедуры, благодаря которому пользователь сможет (1) отождествить первоначальный источник нужной ему информации и (2) увидеть, по крайней мере в общих чертах, как именно информационная система обрабатывает исходные данные, чтобы ответить на его запрос. В терминах JTB теории знания можно сказать, что реализация такой функции позволит пользователям не просто получать достоверную информацию по индивидуальному запросу, но и получать соответствующее запросу индивидуальное *обоснование* знания, содержащемся в этой информации.

5.2) *Формальные системы гильбертовского и генценовского типа*.

Современное понятие об аксиоматической теории и аксиоматическом методе представления теорий восходит к работе Давида Гильберта 1899 года “Основания геометрии” и последующей совместной работе Гильберта и Бернайса над формализацией математике. Помимо чистой математики Гильберт планировал применять аксиоматический метод в физике и других областях научного знания (??наша статья). Его работы в области логики и оснований математики оказали огромное влияние на развитие символической логики в 20-м веке. Формальная аксиоматическая теория “в стиле Гильберта” (Hilbert-style axiomatics, теория Гильбертовского типа) содержит выделенное множество формул, которые называют *аксиомами*. Идея состоит в том, что удачным образом подобранные формальные аксиомы при подходящей интерпретации будут выражать некоторые элементарные истины данной теории, из которых, можно было бы формально вывести (если не все, то “почти все”) другие утверждения этой теории, опираясь при этом на стандартную фиксированную интерпретацию синтаксической дедукции как сохраняющего истинность логического вывода. Никаких явных требований эпистемологического характера Гильберт к аксиомам не предъявлял, но из его попытки аксиоматизировать Евклидову геометрию и других примеров ясно, что он все-таки подразумевал, что аксиомы должны быть по возможности простыми (в синтаксическом смысле), а их истинность при заданной интерпретации - легко проверяемой элементарными средствами (хотя и не обязательно “самоочевидной”). Что касается правил вывода, то стратегия Гильберта состояла в том, чтобы сократить число таких правил до минимума и использовать один и тот же набор правил вывода во всех случаях. Это отвечает идее о том, что правила логики являются универсальными и формальными в том смысле, что они никак не зависят от содержания конкретной теории.

Студент Бернайса и ассистент Гильберта Герхард Генцен предложил альтернативный подход к формализации теорий. Такие формальные системы, которые сегодня называют построенными “в стиле Генцена” (или системами генценовского типа) содержат много правил и вовсе не содержат аксиом (или содержат мало аксиом). Памятуя о том, что термин “аксиома” в истории имел значения отличающиейся от современного - и в частности тот факт, что Аристотель называет аксиомами логические *правила*, включая правило *совершенного силлогизма* - мы будем называть построенные в стиле Генцена формальные системы аксиоматическими в широком смысле слова (даже если в них нет аксиом в обычном современном смысле).

Генцен использовал этот подход в логике и в формальной арифметике, но в отличие от Гильберта он на протяжении своей короткой жизни и научной карьеры не делал попыток использовать свой формальный подход в других областях математики или в естественных науках. Тем не менее, на наш взгляд, системы в стиле Генцена представляют интерес для компьютерного представления знаний по крайней мере по следующим причинам:

(а) Формальные системы генценовского типа проще реализуются в виде программного кода, чем системы гильбертовского типа. Правило синтаксической дедукции можно реализовать в виде простого программного шага, а реализация любой пропозиции и, в частности, аксиомы, требует программной спецификации на языке более высокого уровня. Если теория представлена формальной системой генценовского типа, ее вычислительные характеристики гораздо лучше поддаются анализу, чем аналогичные характеристики систем гильбертовского типа.

(б) Системы генценовского типа естественным образом связаны с теоретико-доказательной семантикой логического вывода, поскольку семантика этого типа придает значения непосредственно синаксическим правилам. (Более того, именно на Генцена часто указывают как на изобретателя этой семантики.) Поэтому использование в системах представления знаний формальных систем генценовского (а не гильбертовского) упрощает решение проблемы обоснования.

(в) Если система генценовского типа интерпретируется не как чисто логическое исчисление, а как некоторая содержательная теория, то выводы в такой системе не будут уже чисто логическими. С прагматический точки зрения преимущество использования в системах представления знаний именно логических выводов во всех случаях не является очевидным. С одной стороны, строго логический подход позволяет использовать одни и те же схемы вывода независимо от предметной области, к которой относится данный тип знаний. Но, с другой стороны, цена такой унификации в конкретных приложениях может оказаться несоразмерно высокой: если речь идет о системе представления знаний в конкретной предметной области, то может оказаться выгоднее использовать процедуры вывода валидные только в данной области (как например вывод о том, что вероятно скоро пойдет дождь на основании единственной предпосылки о том, что небо затянули тучи).

В компьютерной науке и информационных технологиях системы использующие подобные внелогические процедуры известны по названием “экспертных систем основанных на правилах” (rule-based expert systems) (ref 2011??). Как и системы представления знаний использущие гильбертовскую архитектуру, существующие экспертные системы основанные на правилах не поддерживают функцию эпистемического обоснования выводов. Однако именно генценовская архитектура основанная на правилах, позволяет проще реализовать функцию обоснования путем предъявления к правилам вывода специальных требований, теоретическим прототипом которых могут быть разработки в области формальной теоретико-доказательной семантики и возможные расширения такой семантики на область внелогических выводов (конструкций). В заключительном разделе статьи мы опишем перспективную формальную систему, которая включает в себя внелогические процедуры, которые, однако, имеют важную логическую функцию.

6) <u>Процедурное знание, изоморфизм Карри-Говарда и гомотопическая теория типов</u>

Говоря о формальных правилах дедукции, мы до сих пор подразумевали, что формулы, представляющие посылки и заключения таких дедукций, интерпретируются как пропозиции (даже если соответствующие выводы не являются логическими, как в нашем примере с погодой). Мы будем называть такого рода дедуктивные процедуры *пропозициональными.* Понятие процедурного знания, разумеется, не сводится к знанию пропозициональных процедур: помимо знания пропозициональных процедур в приложениях большую роль играют, например, процедурные знания о технологических процессах. На самом деле, как мы сейчас увидим, нет никакой необходимости ограничивать семантику формальной дедукции именно пропозициональными (в том числе логическими) процедурами; порождение символических формул из данных формул может также напрямую интерпретироваться как та или иная содержательная процедура (например, как процедура построения геометрической фигуры циркулем и линейкой) и даже как материальная технологическая процедура.

Символическое исчисление с непропозициональной интерпретацией было предложена Андреем Колмогоровым в 1932 году: автор назвал его “исчислением задач”. С синтаксической точки зрения исчисление задач Колмогорова представляет собой вариант интуиционистского исчисления высказываний, но при этом автор интерпретирует формулы не как предложения, а как задачи. Другой важный пример непропозиционального символического исчисления - это лямбда-исчисление впервые предложенное примерно в то же время (1930 г.) Алонзо Черчем в качестве общей математической модели *вычисления* . Позже Хаскель Карри и Вильям Говард обнаружили, что просто типизированный (simply typed) вариант лямбда-исчисления имеет такую же (с точностью до изоморфизма) синтаксическую структуру, как и импликативный фрагмент построенного Генценом исчисления, который он назвал “естественным выводом” (natural deduction), и которому он придавал пропозициональную семантику. Общая структура просто типизированного лямбда-исчисления и импликативного фрагмента натурального вывода в 1970-х годах была исследована методами математической теории категорий Ламбеком и другими авторами: использование языка теории категорий позволило абстрагироваться от синтаксических деталей и описать искомую структуру в “инвариантном” (по отношению к случайным синтаксическим выборам) виде.

Изоморфизм Карри-Говарда нашел широкое применение в компьютерной науке, поскольку он указал на прямую связь между вычислениями, с одной стороны, и доказательствами (выводом) пропозиций, с другой стороны. Эти соображения породили так называемые “парадигмы” в компьютерной науке и программировании, которые называют “доказательства как программы” (proofs-as-programs) и “высказывания как типы” (propositions-as-types). Точный логический и эпистемологический смысл таких “парадигм” до сих пор остается проблематичным, хотя, как мы сейчас покажем, в самые последние годы в этой области был достигнут значительный прогресс.

Опираясь на свою совместную работу с Колмогоровым, на философский подход Правица (общая теория доказательств и теоретико-доказательная семантика логического вывода), на изоморфизм Карри-Говарда и на свой интерес к компьютерной науке, Пер Мартин-Леф опубликовал в 1983 году ?? типизированное формальное исчисление (с зависимыми типами), которое он называл “интуиционистской” теорией типов. Задуманная семантика этого исчисления предполагает, что базовые формулы этого исчисления вида $a : A$ (где A называется типом, a - термом данного типа, а сама формула называется *суждением*) допускает разные интерпретации из следующего списка:

- a является доказательством пропозиции A
- a является элементом множества A
- a является решением задачи A
- a является реализацией намерения A

Согласно Мартину-Лефу 1983 года возможность интерпретировать один и тот же синтаксис такими разными способами указывает на то, что понятия пропозиции, множества, задачи и намерения (это последнее понятие Мартин-Леф, по всей видимости, заимствует у Гуссерля) являются в каком-то смысле тождественными. Понятия истиныи доказательства, с которыми связаны пропозиции в теории типов Мартина-Лефа (ТТМЛ) являются конструктивными: пропозиция (тип) считается истинной, если у нее есть доказательство (хотя бы один терм) и ложна в противном случае (пустой тип). Добавим, что ТТМЛ это исчисление генценовского типа (без аксиом), фрагменты которого уже реализованы в виде языка программирования AGDA и программы-прувера Coq.

Существенный прогресс в исследовании ТТМЛ и смежных формальных теорий, который заставляет пересмотреть некоторые аспекты их первоначально задуманной семантики, был достигнут в течении последнего десятилетия после того, как Владимир Воеводский установил связь между ТТМЛ и геометрической теорией гомотопий. ТТМЛ и аналогичные теории с новой гомотопической семантикой сегодня принято называть *гомотопической теорией типов* (ГТТ). В рамках такой семантики типы интерпретируются как гомотопические типы (абстрактые пространства, в которых определены основные понятия теории гомотопий), а термы - как точки таких пространств. Не имея возможности обсуждать здесь детали, укажем только на один принципиальный момент: ГТТ позволила установить, что “не все типы одинаковы”: только типы специального вида (гомотопического уровня), а именно типы содержащие самое большее единственный терм, можно отождествить с пропозициями; типы другого специального вида отождествляются с множествами; типы более выскоих уровней мы оставим в стороне. Таким образом, первоначальная идея о том, что всякий тип допускает интерпретации как пропозиция, множество и т.д. в контексте ГТТ не выглядит более убедительной. Заметим, однако, что если данный тип представляет собой множество, то из него можно получить пропозицию, отождествив все элементы этого множества (если они есть). Такую процедуру в ГТТ называют *обрезанием* и она может быть определена также в случае высших типов. Имея в виду это процедуру, можно сказать, что хотя не всякий тип в ГТТ является пропозицией, любой тип более выского порождает пропозицию с помощью обрезания. Если думать о множестве как множестве доказательств, то можно сказать, что отождествляя эти доказательства, мы перестаем интересоваться их деталями и фиксируем только наличие или отстуствие доказательства у данного высказывания, то есть, говоря другими словами, фиксируем только истинностное значение этого высказывания (в конструктивном смысле).

Показажем теперь, каким образом ГТТ, на наш взгляд, может быть теоретической моделью представления знаний, включая пропозициональные и процедурные знания. Если понятие доказательства утверждения понимается достаточно широко (как это и делается в конструктивной логике и общей теории доказательств в духе Правица) и включает в себя, в частности, материальные свидетельства и результаты физических экспериментов, то описание экспериментальной процедуры, которая позволяет получить нужное свидетельство, тоже можно рассматривать как описание процедуры доказательства соответствующего физического утверждения. В эпистемологических терминах можно сказать, что знание *как* провести эксперимент используется в такой ситуации в качестве элемента знания *что* физический мир устроен некоторым образом, а именно элемента, который обеспечивает *обоснование* данного пропозиционального знания. Подобного рода структуру мы как раз и имеем в ГТТ: синтаксис этой теории представляет собой систему правил, которые на пропозициональном гомотопическом уровне применяются как логические правила; на более высоких уровнях те же самые схематические правила используются как правила непропозициональных процедур, которые играют роль процедур верификации (доказательства) некоторых утверждений. Таким образом, мы имеем здесь теоретическую схему, с помощью которой может быть представлено как пропозициональное знание, так и процедурное знание, которое является знанием о том, *как обосновать* соответствующее пропозициональное знание. Таким образом, в рамках данного формализма обоснование не сводится к чисто пропозициональным процедурам, что дает возможность представлять экспериментальное знание, технологические знания и другие типы знаний, которые могут иметь большое прикладное значение. В случае технологических знаний это открывает возможность не только формально представлять технологические процедуры в программном виде, но и формально доказывать, что изготовленное с помощью таких процедур техническое изделие будет обладать требуемыми характеристиками.