

знания и их представление в компьютерную эпоху:

истина, обоснование, интерпретация, ревизия

Андрей Родин

ИФРАН и СПбГУ

Вологда, 3 июня 2017

план:

представление знаний: формальные и социальные аспекты

логика и онтология

истина и обоснование: конструктивный подход

объективация свидетельств

заключение

wiki: eng

Knowledge representation and reasoning (KR) is the field of artificial intelligence (AI) dedicated to representing information about the world in a form that a computer system can utilize to solve complex tasks such as diagnosing a medical condition or having a dialog in a natural language. Knowledge representation incorporates findings from psychology [?] about how humans solve problems and represent knowledge in order to design formalisms that will make complex systems easier to design and build. Knowledge representation and reasoning also incorporates findings from logic to automate various kinds of reasoning, such as the application of rules or the relations of sets and subsets.

Представление знаний - вопрос, возникающий в когнитологии (науке о мышлении), в информатике и в исследованиях искусственного интеллекта. В когнитологии он связан с тем, как люди хранят и обрабатывают информацию. В информатике - с подбором представления конкретных и обобщенных знаний, сведений и фактов для накопления и обработки информации в ЭВМ. Главная задача в искусственном интеллекте (ИИ) - научиться хранить знания таким образом, чтобы программы могли осмысленно обрабатывать их и достигнуть тем подобия человеческого интеллекта.

а также:

- ▶ wiki
- ▶ arXiv.org
- ▶ Open Journal Systems
- ▶ MOOCs
- ▶ crowd-sourcing science (aka citizen science)
- ▶ ...

В настоящее время вики и другие перечисленные выше платформы и технологии как правило НЕ пользуются технологиями представления знаний в том смысле, в каком этот термин используется в ИИ.

Однако это вопрос времени! Системы управления обучением (LMS) уже активно используют KR (включая школьное образование). Проект построения унивалентных оснований математики (Воеводский) направлен на построение универсального формального стандарта записи математических текстов, допускающего эффективную реализацию в виде компьютерного кода (Coq, Agda). Можно ожидать, что в обозримом будущем технологии представления знаний будут в той или иной степени использоваться в любых эпистемических практиках.

логическая формализация и социальные институты

Вики, arXiv.org, MOOCs плохо совместимы с традиционными эпистемическими институтами (образовательные, исследовательские, экспертные институты) и способствуют их реформированию, а также созданию независимых эпистемических институтов нового типа.

Практический пример: внедрение Open Journal Systems в ИФРАНе

В контексте современных информационных технологий связи между формальными и социальными аспектами представления знаний имеют прагматический характер и являются очевидными!

ОНТОЛОГИЯ

Knowledge Representation

- **What to Represent?**

Let us first consider what kinds of knowledge might need to be represented in AI systems:

- **Objects**

- -- Facts about objects in our world domain. *e.g.* Guitars have strings, trumpets are brass instruments.

- **Events**

- -- Actions that occur in our world. *e.g.* Steve Vai played the guitar in Frank Zappa's Band.

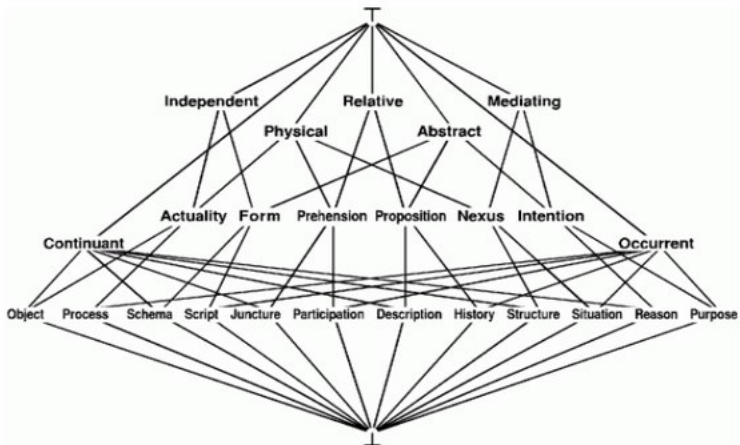
SUMO

Suggested Upper Merged Ontology (Adam Pease at www.adampease.org/OP/)



SOWA

Алмаз Сова: John Sowa at www.jfsowa.com/ontology/



использование онтологий разных уровней

(неполный список)

- ▶ инвариантная база для переводов (в другие формальные рамки) и интерпретаций знаний
- ▶ возможность многоуровневой ревизии знаний (онтологическая ревизия высокого уровня - самая глубокая)

Хотя в эпистемологии есть ряд влиятельных неформальных теорий ревизии и роста знаний (Гегель, Лакатош, Кун), такие теории до сих пор остаются слабым местом *формальной* эпистемологии. Аналогичное замечание может быть сделано по поводу теорий перевода и интерпретации (герменевтики).

Однако для задач компьютерного представления знания эти две задачи приходится решать в любом случае.

ЕСТЕСТВЕННЫЙ ЯЗЫК И ЛОГИКА

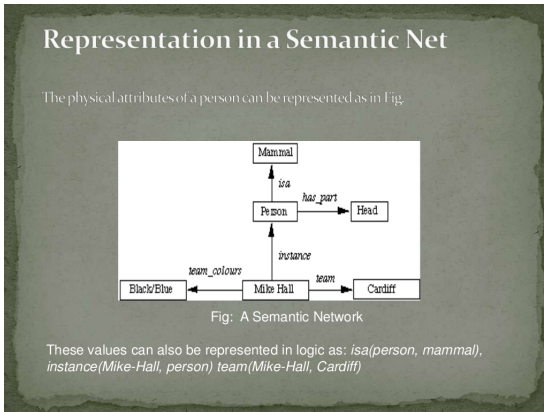
- English or natural language is an obvious way of representing and handling facts.
- Logic enables us to consider the following fact:
- *spot is a dog* as $dog(spot)$ We could then infer that all dogs have tails with:

$\forall x : dog(x) \rightarrow hasatail(x)$ We can then deduce:
 $hasatail(Spot)$

дескриптивные логики: расширения классической FOL
программная реализация: LISP, PROLOG (общие), KRL
(специализированный)

семантические сети

Идея: значение понятия это система отношения этого понятия с другими понятиями.



замечание

С точки зрения обычной логики и теории моделей здесь речь все еще идет о синтаксисе! Ср. аксиомы Гильберта для евклидовой геометрии.

представление знаний: формальные и социальные аспекты
логика и онтология
истина и обоснование: конструктивный подход
объективация свидетельств
заключение

вопрос:

Действительно ли в “представлении знаний” (в ИИ) речь идет о знании в смысле традиционной эпистемологии?

Знание - это [i] истинная [ii] обоснованная [iii] вера
(Платон, Тэетет)

замечание 1: Это определения годится только для пропозиционального знания, то есть знанием представленным в виде истинного и обоснованного высказывания, которое также называют знанием-что (knowledge-that). Знание-как (knowledge-how) aka процедурное знание не подпадает под это определение (но, как аргументируют некоторые логики и философы, может быть в определенном смысле редуцировано к пропозициональному знанию).

замечание 2: Это определение ничего не говорит о том, что пропозициональное знание может быть представлено в виде *теории*, которая содержит множество высказываний связанных друг с другом отношением логического следствия (или несколькими отношениями такого рода).

наблюдение:

Во многих KR-контекстах знаниями называют просто пропозиции (высказывания), связанные подходящими отношениями (включая отношение логического следствия) в определенные структуры (ср. семантические сети и фреймы). Вопрос об истинности и обоснованности (доказательстве) этих высказываний во всех известных мне случаях остается за рамками рассмотрения. Считается, по-видимому, что вопросы касающиеся истинности и обоснованности не входят в область компетенции и ответственности инженерии знаний. Поэтому

уточнение:

Логический вывод сложных высказываний из примитивных высказываний (“фактов”), который реализуется в KR-системах в виде алгоритмов, программных средств и действительных вычислений, согласно используемой логической теории, сохраняет истинность. Если примитивные высказывания истинны, то такие выводы будут доказательствами сложных высказываний. Но как проверить и доказать истинность примитивных высказываний (аксиом)? Как я покажу чуть ниже, этот вопрос **ТАКЖЕ** имеет прямое отношение к инженерии знаний и к информационным технологиям.

вера и убеждение

Большой прогресс (в основном независимо от KR) достигнут в исследованиях аспектов *веры*, которые имеют большое прикладное значение для медийных технологий убеждения / суггестии, которые используются, в частности, в политической пропаганде и коммерческой рекламе.

постправда?

Ср. популярные дискуссии о наступлении “эпохи постправды”. На мой взгляд, ситуация состоит в том, что существующие технологии представления знаний на сегодняшний день являются несовершенными. До вопросов связанных с истиной и обоснованием у инженеров просто еще не дошли руки (как я покажу, возможные решения проблемы с технологической точки зрения не являются простыми). Возможно, отсутствие прогресса в этой области связано с длительным доминированием в математической и философской логике теорий, в которых проблемы истины и обоснования отодвигаются на задний план (я уточню чуть позже, что я имею в виду).

постправда?

Это техническое несовершенство вряд ли может быть основанием для ревизии традиционной эпистемологии. В такой ситуации можно ожидать, что эпистемология, наоборот, может мотивировать новые технические решения. Заметим, что информационные технологии как и любые другие высокие технологии основаны на и зависят от фундаментальных физических и технологических знаний в классическом смысле слова.

истина и доказательство: “классика”

Доказательство = формальный (синтаксический) вывод из аксиом; истинность - семантическое свойство символических выражений (Тарский 1944), которое возникает при подходящей интерпретации выражений подходящего вида (включая аксиомы); аксиомы имеют тривиальные доказательства (= “принимаются без доказательства”). (Аристотель, Фреге, Гильберт, Гедель, Тарский)

истина и доказательство: конструктивизм

Истинность высказывания = предъявление (хотя бы одного) доказательства; примитивные истинные суждения (“аксиомы”) тоже нужно доказывать! Примитивное доказательство = свидетельство. Доказательство \neq формальному выводу. (Лейбниц, Брауэр, Колмогоров, Мартин-Леф, Правиц)

Конструктивная версия Первой теоремы Геделя о неполноте: в любой формальной системе включающей в себя арифметику натуральных чисел существуют доказательства, которые невозможно алгоритмически построить из аксиом (= примитивных доказательств) этой системы. Поэтому доказательство \neq формальному выводу!

философские интерпретации конструктивной истины

- ▶ субъективная (Брауэр)
- ▶ объективная (Колмогоров)
- ▶ реалистическая : truth-maker realism in constructive setting (Мартин-Леф и Смит)

утверждение:

Конструктивный подход к истине и обоснованию лучше подходит для инженерии знаний, поскольку он заставляет инженеров заниматься этими проблемами, а не оставлять их решение каким-то другим людям.

пример: исследования климата

Каким образом можно удостовериться, что эмпирические данные, на которые опираются знания о климате, НА САМОМ ДЕЛЕ были получены с таких-то метеостанций определенном месте и в определенное время, а не были подтасованы или изменились без чьего-либо намерения их изменить? А если речь идет о дорогостоящих данных полученных с космических аппаратов и дорогих экспериментальных установок вроде БАКа?

Насколько мне известно, современные информационные технологии не в состоянии давать такого рода гарантии. Проблема становится более острой, если речь идет о старых данных, полученных десятки, сотни и тысячи лет назад.

DISIDERATUM: Инженерия знаний должна научиться решать эту проблему, чтобы ответственно говорить о “фактах”.

возобновляемые и уникальные свидетельства

В некоторых случаях проблему отсутствия долгой и надежной памяти в системах представления знаний можно решить за счет воспроизведения свидетельств. Например, с базой знаний “Ньютонова Механика” можно интегрировать простые экспериментальные установки демонстрирующие принципы ньютоновой механики, вроде тех, которые используются (или по крайней мере ранее использовались) на школьных уроках физики.

Знание-как (провести нужный эксперимент) в таких случаях используется для производства свидетельств (доказательств) для *знания-что* (например, что ускорение тела прямо пропорционально действующей на это тело силе и обратно пропорциональна его массе).

объективация уникальных свидетельств

В терминах компьютерной науки можно сказать, что проблема недостатка памяти может быть решена за счет дополнительных вычислений. Однако такое решение не может быть универсальным, поскольку уникальные невоспроизводимые свидетельства, если они объективированы (как, например, материальные исторические свидетельства в публично доступных архивах), также играют фундаментальную роль в науке. Каждый отдельный эксперимент и наблюдение является уникальным свидетельством, даже если эти эксперименты и наблюдения воспроизводимы.

объективация уникальных свидетельств

DISEDERATUM:

Инженерия знаний должна предоставить эффективные технические средства для объективации уникальных свидетельств разного рода. Формат для представления *факта* должен включать в себя форму *подтверждения* этого факта.

пространство и время памяти

Характерно, что базовая идеальная модель вычислений, которую использует современная компьютерная наука, а именно универсальная машина Тьюринга, заставляет ученых оценивать объем памяти вычислительных устройств в пространственных единицах измерения. Баланс между объемом вычислений и объемом памяти, которые необходимы для решения данной вычислительной задачи, иначе называют пространственно-временным балансом (space-time tradeoff). Идея о том, что память помимо пространства еще как-то связана со временем, которая кажется естественной для человека с улицы, выходит за рамки базовых идеализаций современной компьютерной науки.

пространство и время памяти

Можно сказать, что эта базовая идеализация описывает только относительно короткие вычисления и делает допущение о том, что во время данного вычисления содержимое памяти полностью контролируется вычислительным устройством. Очевидно, что такая модель не подходит для информационных систем, которые допускают возможность поступления в систему информации извне этой системы (а таких систем сегодня уже большинство).

Floridi 2014

Современные ИТ технологии возвращают нас в доисторический период?

Our digital memory seems as volatile as our oral culture was but perhaps even more unstable, because it gives us the opposite impression. This paradox of a digital 'prehistory' [to be contrasted with the new era of "hyper-history" promised by the same author - A.R.] - ICTs are not preserving the past for future consumption because they make us live in a perennial present - will become increasingly pressing in the near future. [...] The risk is that differences are erased, alternatives amalgamated, the past constantly rewritten, and history reduced to the perennial here and now.

заключение:

Развитие информационных технологий вообще и технологий представления знаний в частности представляет собой вызов для традиционной эпистемологии. Философы должны реагировать на технологические изменения. Эта реакция НЕ должна быть чисто пассивной, как в случае историсофской теории постправды. Реакция должна быть скорее активной и критической. В этом случае у философов есть шанс внести свой вклад в дальнейший технологический и гуманитарный прогресс.

представление знаний: формальные и социальные аспекты
логика и онтология
истина и обоснование: конструктивный подход
объективация свидетельств
заключение

СПАСИБО!