



Санкт-Петербургский
Государственный
Политехнический
Университет

Институт прикладной
математики и механики

Введение в профессиональную деятельность

Лекция 4

**Формально-математическое
vs информационно-
вычислительное описания
Деонтическая логика**

3 марта 2020 г.

Что обсуждали на прошлой лекции – логические ошибки при создании ПО компьютеров неизбежны!?

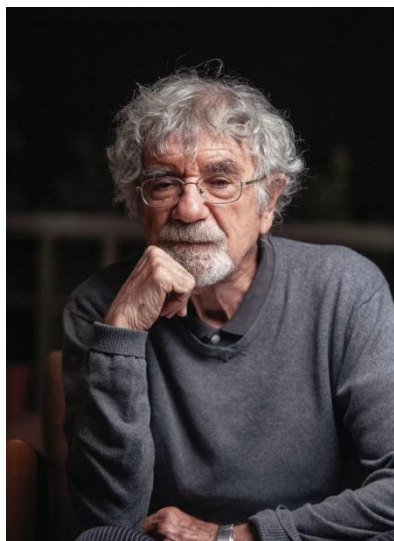
Компьютерные системы «ведут» себя в точности так, как их запрограммировали, **но не так, как предполагали программисты!** $A \vee \bar{A} = 1$

А также:

- Computer science should be about concepts, not languages.
Лесли Лэмпорт
- **Истина** не формализуема, математика – это метафора. основанная на **информационном сравнении** объектов, т.е. «вычислимом» или «семантическом» подобии.
- «К.П.Д.» математической формализации < 100% из-за теоремы Геделя, а именно какова бы ни была совокупность аксиом теории, содержащей арифметику, если теория непротиворечива, существует такое утверждение A , что ни A , ни его отрицание ($\neg A$) – в рамках этой теории не **доказуемы**.

Логика законов, описывающих реальность - непротиворечивая модель «ПО-ЗНАНИЯ», «СО-ЗНАНИЯ», «О-СОЗНАНИЯ»

Умберто Матурана: жизнь как процесс представляет собой процесс познания. Познано то, что логически доказуемо....или вычислимо.



«Биология познания» (1970)

Модель знаний основана на понятии

«**доказуемо**», которое следует законам логики:

- если высказывание доказуемо, оно истинно (доказать можно только истину, доказательств лжи не существует);
- логические следствия доказуемого также являются доказуемыми;
- логическое противоречие недоказуемо и т.п.
- если высказывание истинно, то неверно, что его отрицание также истинно («Если истинно, что Земля круглая, то неверно, что истинно, будто Земля плоская» и др.

Логика «законов компьютерной реальности»

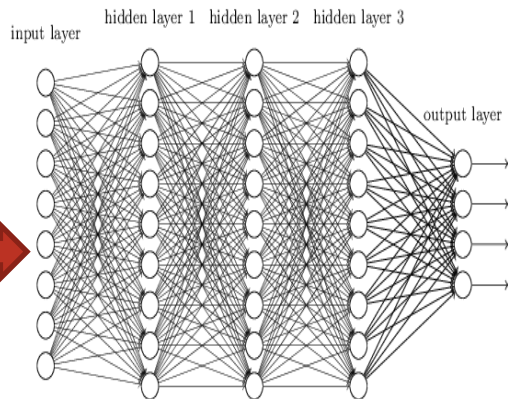
Исходим из того, что **«по-знание»** есть результат работы **«сознания»**.

- Поэтому , если мы хотим **моделировать** «сознание», включая различные **модальные** аспекты с помощью компьютерных систем, мы должны начать с метафорического описания выбранных свойств реальности, а затем перейти к конструированию структур, в которых эти описания наиболее информативны и физически воспроизводимы.

Пример: законы бинарной логики (редукция «мира» к Булевой алгебре) .

Эволюция компьютерных наук: от «логико-алгоритмических» к «нейроморфным» вычислениям.

нейронная сеть - **индуктивный** классификатор входных данных

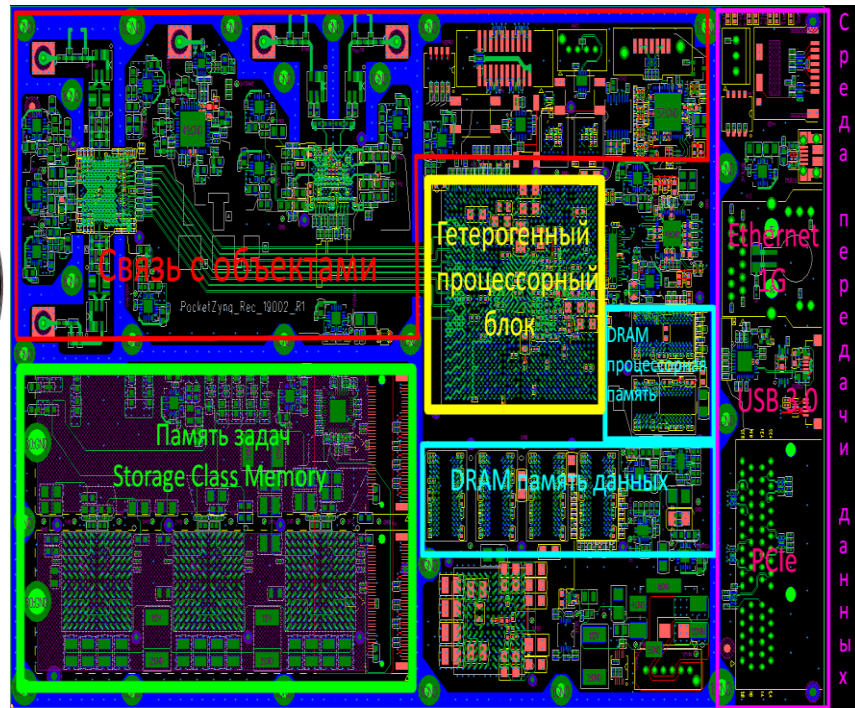


MT

$$\nabla \cdot \nabla \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$$
$$= \frac{1}{r^2 \sin \theta} \left[\sin \theta \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right]$$



Гетерогенный вычислитель:



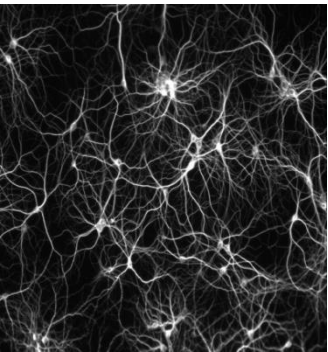
Суть нейроморфного подхода:

- Непроцедурное отображение множества входных данных на **множество заданных классов**
- Программирование без явно заданного алгоритма с помощью «обучения» на примерах

Суть гетерогенного подхода:

- Реализация возможностей **прямого алгоритмического** решения и индукции с помощью «обучения» **неявного** алгоритма на **заданном классе примеров**

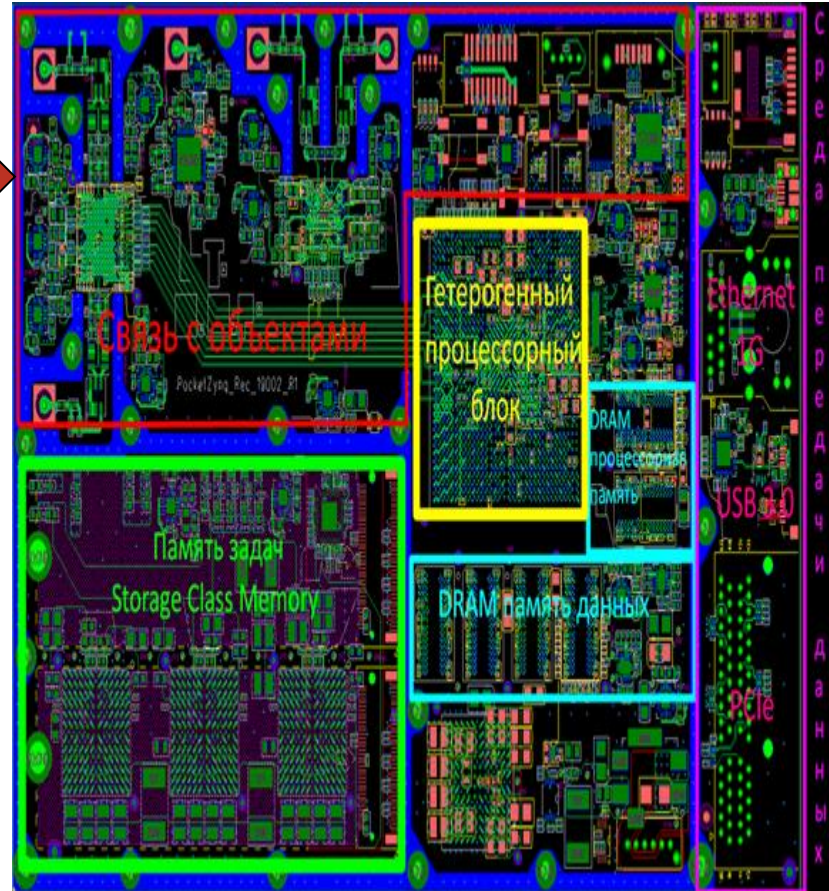
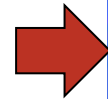
Характеристики «нейро» и электронных «вычислительных» сетей



Формирование
нейронных
колонок,
цепей,
сетей



МРТ снимок мозга



Структура вычислителя:

Число процессорных ядер 10^4

Длина соединительных линий 50 см

Объем памяти 10^{13} байт

Число нейронов (коммутаторов) 10^{11} шт.

Число синапсов 10^{14} их длина более 10^6 км

Объем памяти $> 10^{15}$ байт

Метафоры vs количественные характеристики :

- Основные логические понятия «математики и компьютерных наук»: **Точное равенство, перечислимость, вычислимость, разрешимость.**
- В максимальной степени эти понятия-метафоры информативны для объектов, которые обозначаемы символами знаками с далеко отстоящими друг от друга в общем семантическом поле свойствами (железный, ... человек, «1», «0»....) . Объекты реальности м.б. лишь схожи, но не равны.
- В **семантическом** обогащении научного знания о реальности заключается ключевая роль метафоры: «гладкая функция», «мягкое условие», «грубая оценка» и «жесткое условие». «странный аттрактор», «мнимое число»...

Итого: т.к. логическое противоречие (равенство логических несовместных переменных) недоказуемо, поэтому истина «реальности» не формализуема, а значит объекты не **перечислимы**, все их характеристики не **вычислимы**, а их совокупности не **разрешимы** , если не учитывается контекст или **модальный** аспект (законы и нормы, которым следуют **все** субъекты и объекты реальности)

Опыты математической метафоризации реальности

Приводят к сложным лингвистическим казусам и логическим нелепостям, например:

- «ребро графа» (метафоризация **омонимии** - одинаковые по написанию и звучанию, но разные по значению слова);
- «автоматы на деревьях» (двойная метафора)
- «хорошее» начальное приближение.
- «граф- «звезда»

Это говорит о том, что самое математическое знание в своем выражении в научном языке нуждается в **постоянной** образной поддержке (ибо метафора – это прежде всего образное сравнение) со стороны повседневного человеческого опыта.

Концепция «дуала» К. В. Анохина: Ментальное и физическое. Феномен супервентности



Супервентность (англ. Supervenience) — отношение детерминированности состояния любой системы состоянием другой системы. Набор свойств одной системы **супервентен** относительно набора свойств другой системы в том случае, если существование **различия** между двумя фактами в свойствах первой системы невозможно без существования такого же **различия** между двумя фактами в свойствах второй системы



Код физического
процесса



Интеллектуальный
субъект,
«понимающий» код



Инструмент
«расшифровки»
кода



физический
«процесс»

Супервентный характер феномена «музыка» очевиден.

Объекты образуют «дуал» - отображение в ментальную (метафорическую) и физическую реальности.

Примеры супервентности: 1) **отсутствие различий в компьютерной программе** при отсутствии различий в аппаратной конфигурации компьютера; 2) **отсутствие различий в экономике** при отсутствии различий в поведении экономических агентов.

С точки зрения информационного описания конкретного материального объекта физической реальности все возможные «равновесные» состояния не различимы (в равновесном состоянии игнорируются микроскопические сущности, например, движение молекул и атомов, из которых состоит



Гипотеза 1. Сущности, которые нельзя измерить, а их характеристики вычислить нельзя арифметизировать, т.е. знания о них не имеют **числовой меры или характеристики**.

Гипотеза 2. Знания, которые можно арифметизировать, обладают свойством **аддитивности** (их можно складывать и накапливать).

Гипотеза 3. Любые знания, которые можно арифметизировать, можно представить в форме компьютерных программ

- **Математика – цифровая (числовая) метафора свойств реальности.** В Природе действуют не только количественные, но другие математические закономерности, а это приводит к «не идеальности» наблюдаемых физических процессов и «отклонению» от строгих математических формул.
- Законы, которые непосредственно выражаются математическими формулами, применимы к объектам, которые имеют «идеальную» цифровую модель
- Такая модель формирует «двойственные» пары :

число - слово,
модель - дефиниция,
представления - ощущения

...

Информация как мера «неоднородности» :

Физическая информация - объективное свойство реальности, которое проявляется в **неоднородности** распределения материи (вещества и энергии) в пространстве, а также в неравномерности протекания динамических процессов в неживой природе, технических и биологических системах.

$$H = K \cdot \sum_{i=1}^n P_i \log P_i$$

Формула Шеннона (1949 г.)

где H — среднее количество информации в системе с выбором сообщений, с вероятностями $(P_1, P_2 \dots P_n)$, K — константа, зависящая от единицы измерения

Модели физического и информационного планов

«Модель мира» - это научная абстракция 5-ого порядка, следующая за такими абстракциями как:

- Понятие (сущность естественного языка,)
- Число (носитель количественной меры)
- Алгебра (структура и свойства операций)
- Топология и отношения порядка

Дуал «модели мира» объединяет описания

- **физического плана** – локальные и замкнутые. Такие модели описывают реальность, в которой прошлое и будущее формально «**симметрично**», действует принцип физической «относительности».
- **информационного (ментального) плана** – глобальные и открытые. В них «стрела времени» физически необратима, поэтому «**прошлое**» **физически недостижимо**, действует принцип относительности по отношению к знаниям субъекта.

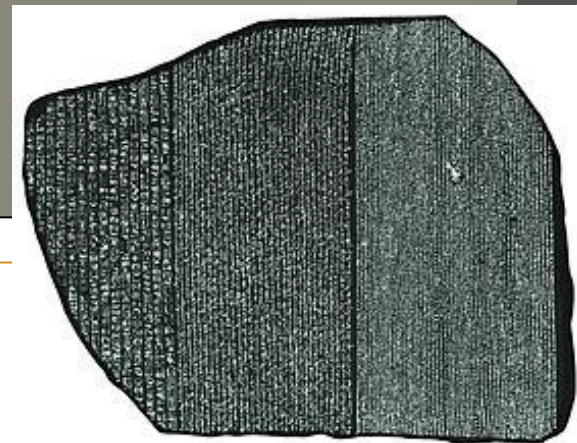
Однако остается вопрос , а достижимо ли прошлое «**информационно**» ?

Истинность или доказанность ... услышанного, произнесенного и вычисленного.

Ключевая идея логики знаний - истина «относительна» и зависит от

- Авторитета того, чьи слова услышаны
- Убежденности в истинности того говорит
- Правильности формул, которые используются для вычисления

Наука 2.0 – ЭТО Rosetta Stone Описания природы возможно различных «языках»

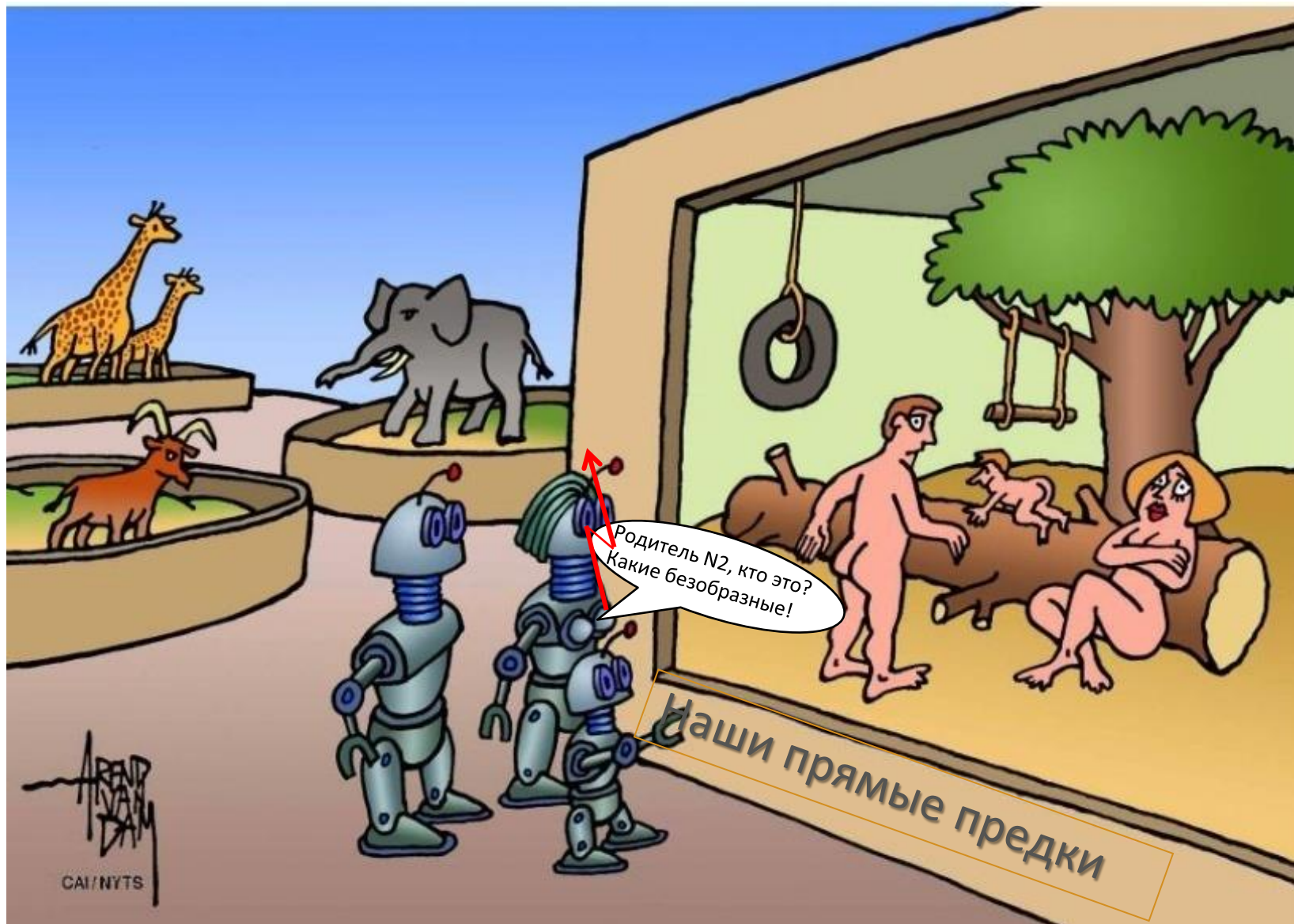


5

В настоящее время эти «языки» математические:

- метафора физики **существующего** (И. Ньютон) – классическая математика или обратимая зависимость между причиной и следствием. **Сила и инерция есть причины движения**, результат – траектории движения есть функции начальных условий (Принцип достаточного основания)
- Метафора физики **возникающего** (И. Пригожин) – появления новых макроструктур или объектов более высокого порядка. **Причина «движения» к порядку - возрастание энтропии** внешней среды (диссипативное становление)
- Метафора физики **управляемого**:
 - кибернетика (Н. Винер) **Цель - причина движения, результат – алгоритм управления «движением», основанный на передачи информации (Bit from It)**,
 - киберфизика - **причина движения - передача информация, результат движения - процесс вычисления цели (It from Bit)**

Что мы будем делать дальше, если ...



Проблема, которую надо решить: современный человек не понимает структуры и правил функционирования среды, в которой живет

- Человек уже сегодня оказался в новом «цифровом» мире, в котором принятие основных решений не является ТОЛЬКО его прерогативой.
- Среда вокруг нас заполнена сложнейшими техническими системами, правила функционирования которых не могут быть поняты ДО КОНЦА ни одним человеком.
- Решения по управлению сложными системами вокруг нас принимаются все в большей степени автономно, без непосредственного участия ЧЕЛОВЕКА.
- **Это порождает новые вызовы и проблемы: как существовать, как выжить человеку в этих новых «цифровых» условиях?**

Нормативная (деонтическая) логика - разновидность модальной логики

Логика	Модальность необходимого $\Box p$	Модальность возможного $\Diamond p$
Алетическая alethic	необходимо, что p истинно (p is necessary true)	возможно, что p истинно (p is possible true)
Доксастическая doxastic	p обязательно истинно (p is believed to be true)	предположительно, p истинно (p is supposed to be true)
Темпоральная temporal	всегда в будущем p истинно (p is always true)	иногда в будущем p истинно (p is sometimes true)
Деонтическая deontic	p обязательно выполнить (p is obligatory)	p разрешено выполнить (p is permitted)
Эпистемическая epistemic	знаю, что p истинно (p is known to be true)	думаю (полагаю), что p истинно (p is believed to be true)

Деонтическая логика – это ветвь модальной логики, которая формализует рассуждения о нормах, обязательствах, запретах и разрешениях

Модальная логика – надстройка над логикой высказываний

- Простейшая модальная логика – расширение обычной логики высказываний.
- Грамматика (рекурсивное определение ППФ):

$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid []\varphi$

Как обычно, выводимые операции:

$\varphi \wedge \psi$ для $\neg(\neg\varphi \vee \neg\psi)$;

$\varphi \rightarrow \psi$ для $\neg\varphi \vee \psi$;

$\varphi \leftrightarrow \psi$ для $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$;

True для $\varphi \vee \neg\varphi$;

False для $\neg\text{True}$;

$\diamond\varphi$ для $\neg[]\neg\varphi$

Пример **правильно** построенной формулы: $[](p \rightarrow \diamond q) \vee []\neg\diamond(p \wedge \neg q)$

Пример **НЕправильно** построенной формулы: $[]\vee \rightarrow (p \diamond \rightarrow q)$

- Семантика операторов $[]$ и \diamond в разных **ЛОГИКАХ** различается.

Формулы модальной логики

= формулы логики

высказываний **+**

один унарный оператор :

$[]$ – необходимо **ИЛИ** \diamond – возможно

$[]$ - бокс (*box*) - необходимо,

\diamond - ромб (*diamond*) - возможно

Деонтическая логика – основа для построения «электронных» институтов цифровой экономики

- Понятия законов и норм, ответственность за их нарушение – все это было **СОЦИАЛЬНЫМИ** понятиями, эти понятия были применимы только к людям, функционирующим в **ОБЩЕСТВЕ** людей.

Цифровая трансформация и создание активных автономных адаптивных агентов придало новое направление **ДЕОНТИЧЕСКОЙ МОДАЛЬНОЙ ЛОГИКЕ** – создание **коллективов ИНТЕЛЛЕКТУАЛЬНЫХ агентов**, в которых гарантированно выполняются “социальные” нормы и правила поведения

Деонтическая логика формализует нормативные высказывания **независимо от того, относятся они к биологическим или искусственным агентам.**

Поэтому на основе ДЛ можно создать базис для построения коллективов автономных агентов, **гарантированно** функционирующих в соответствии с заданным набором норм, законов и правил.

Нормы для построения 'электронных сообществ'

- Рационально построенная нормативно–правовая система «новой цифровой реальности» должна удовлетворять следующему минимальному набору требований:
 - **непротиворечивость**: в системе права отсутствуют несовместимые нормы.
 - **сбалансированность**: для всякой правопредоставляющей нормы предусмотрена соответствующая ей правообязывающая норма.
 - **полнота**: все возможные в данной предметной области действия в системе регулируются (т.е. в каждой ситуации можно определить, является ли конкретное действие конкретного агента необходимым, допустимым, разрешенным или запрещенным), нерегулируемых действий нет.

Проблема: как строить новое цифровое общество , в котором *определены и выполняются* непротиворечивые, сбалансированные системы норм?

Дополнительные операторы

- Кроме модальностей ‘ромб’ и ‘бокс’ нужны дополнительные модальные операторы для выражения модальностей.

Пусть: q – “студенты учатся”

$O q$ – “студенты **обязаны (должны)** учиться” ($[] q$) ‘сильная’ истина

$P q$ – “студентам **разрешено** учиться” ($\Diamond q$) ‘слабая’ истина

$F q$ – “студентам **запрещено** учиться” $F q = O \neg q$

$W q$ – “студентам **необязательно** учиться” $W q = \neg O q$

- Все модальности выражаются через ‘сильную’ модальность:
 - $P q = \neg O \neg q$ “**разрешено** спать” = “**неверно, что обязан не** спать”
 - $F q = O \neg q$ “**запрещено** сидеть” = “**обязан не** сидеть”
 - $W q = \neg O q$ “**необязательно** петь” = “**неверно, что обязан** петь”

Формальные определения модальностей должны согласовываться с интуицией

Заключение: «в начале был закон»

Надо

- создать формальный метафорический язык для спецификации норм, которым следуют как люди, так и интеллектуальные агенты-роботы . Семантика такого языка должна , позволяющая строить структуры, используемые для синтеза планов действий, которые следуют этим нормам.
- формировать нормы на основе законов модельных логик, которые, в свою очередь, можно транслировать в операционную семантику планов поведения групп интеллектуальных агентов. На основе этой семантики агенты САМИ могут строить и анализировать свои планы действий.
- Создавать только таких интеллектуальных агентов, которые действуют и планируют свои действия на основе заранее заданных и формально описанных модельных норм.

Домашнее задание: Семантика Крипке

- Семантика Крипке соответствует интуитивному пониманию смысла утверждений нормативной логики
- $[]q$: “обязательно выполнение a ” Oq – obligation q
- $\Diamond q$: “разрешено выполнение a ” Pq – permission q

“Возможные миры” Крипке следует можно понимать как возможные альтернативы документа или альтернативные варианты реализации системы.

Интуитивное понимание смысла утверждений согласуется с семантикой Крипке: для действия q его выполнение обязательно (Oq) в мире w , iff утверждение о выполнении q присутствует во всех альтернативах мира w .