



Санкт-Петербургский
Государственный
Политехнический
Университет

Институт прикладной
математики и механики

КАФЕДРА ТЕЛЕМАТИКА

Семинар по специальности на английском языке

тема

Verity: explanations and understanding of
reality

Workshop

Лекция 11

18 November
2020 г.

Previous lecture:

Trust in science: separating the explanations and understanding from the models

What are the difference between “trust”, “truth” and “verity”

(в чем разница между понятиями “доверие”, ”истина" и " истинностью”)

What difference between essences with “core” truth

- истина это **соответствие** фактам - truth is **conformity** to facts

Аристотель

- истинно, если **согласовано** с остальным знанием - true if **agreed** with the rest of the knowledge

Теория соответствия

- истинно, если **полезно** - true if **helpful**

Пирс

- истинно, если это результат **соглашения** - true if it is the result of an **agreement** .

Пуанкаре

- истина как **антиномия** - truth as **antinomy** .

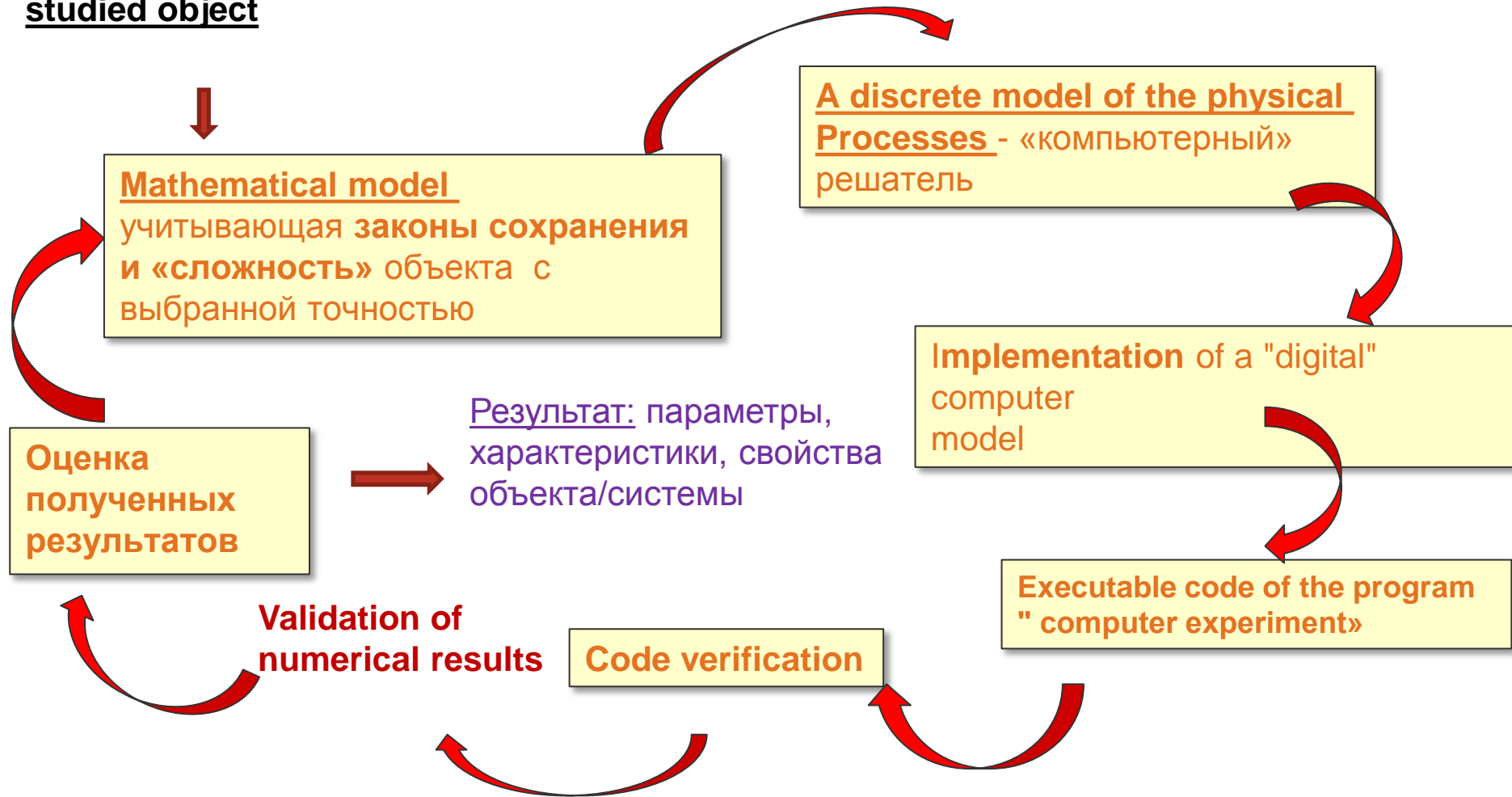
Флоренский

- истина и ложь это всего лишь **имена** - truth and falsehood are just **names** .

Фреге

The problem of "truth" through the cycle of "digital transformation" of knowledge

studied object



Knowledge as a fact of “truth”

Знание как факт “истины”

“truth “ includes:

- understanding
- perception
- knowledge itself (собственно познание)

Truth as a function: set of essence that can deliver the solution to **the problem**.

(Истина как функция: совокупность сущностей, которые могут дать решение проблемы)

Main aspects of trust

Model: The formal interpretation of facts.

Understanding : понять невозможно - формализовать / it is impossible to understand - formalize

Explanation : The truth should be able to explained.

Humans - consumers of the explanations and knowledgrs . Why we trust in

Interpretability Methods layer, which helps human deal with the opacity of machine learning models (how machine calculate explanations)

Black Box Model layer - algorithms using data from the real world to make predictions, find structures or invariants

The Data layer contains ‘digital twins” anything from images, texts, tabular data and so on in order to make it process able for computers and also to store information.

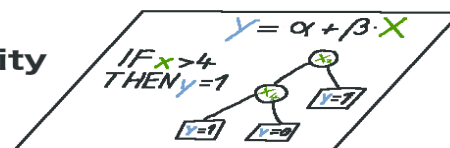
The World layer contains everything that can be observed and is of interest.

Humans



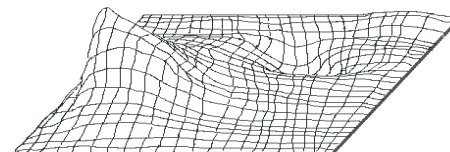
↑ inform

Interpretability Methods



↑ extract

Black Box Model



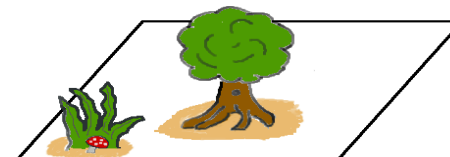
↑ learn

Data

X_1	X_2	X_3	...	X_n
10	2	0		
5	4	0		
1	-1	0		
				$0.70 \cdot X_1 + X_2$

↑ capture

World

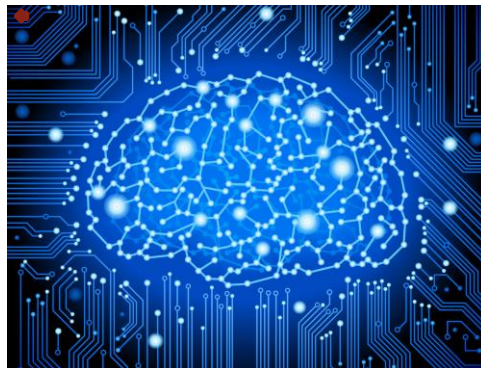


deal with the opacity - solving “inverse” problems in the space of “augmented reality”

Direct problem _ прямая задача вычислений – моделирование объектов с помощью компьютеров с заданной архитектурой (АО+ПО):



Invers problem – выбор такого алгоритма вычисления, «генерирующий» данные, на которых функция полезности достигает максимума :



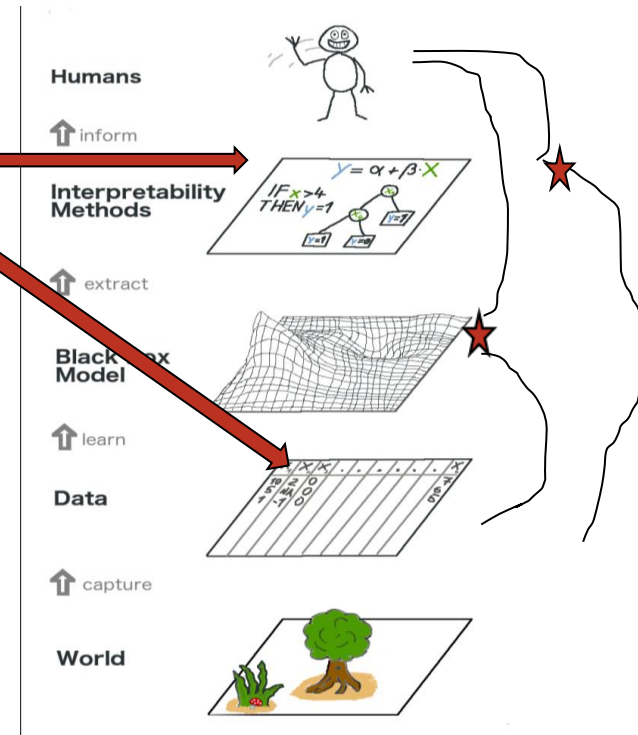
This multi-layered abstraction

We need to understand the differences in approaches between statisticians and machine learning approaches.

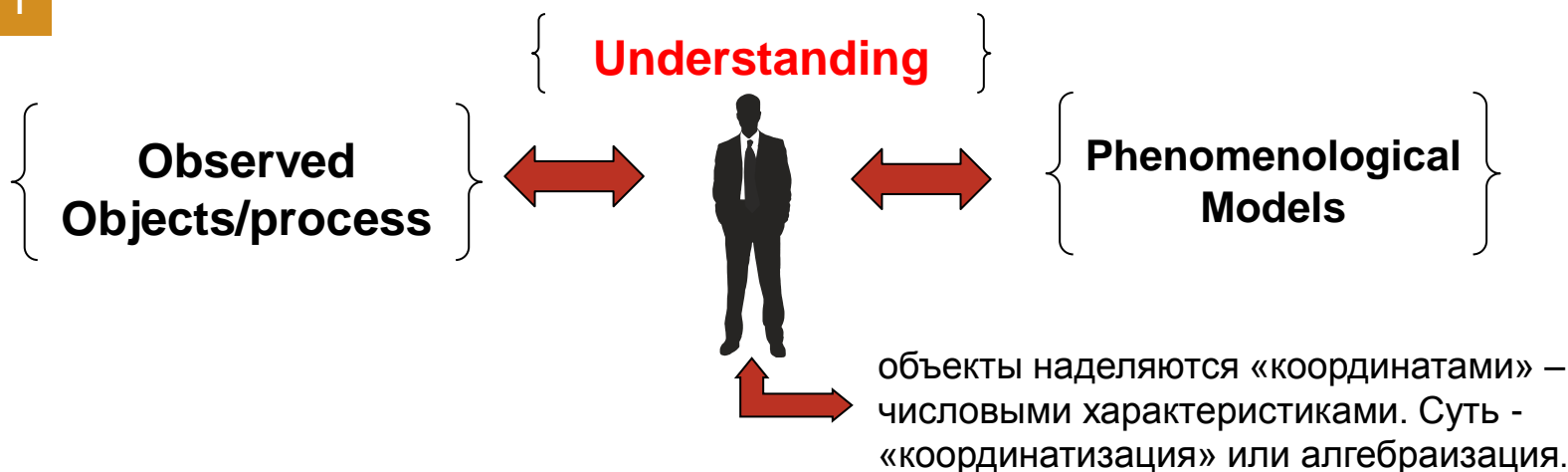
Statisticians deal with the Data layer, such as planning, estimation, predictions, skip the Black Box Model layer and go right to the Interpretability Methods layer.

Machine learning specialists also deal with the Data layer, train a black box machine learning model and skip the Interpretability Methods layer, so Humans directly deal with the Black Box model predictions.

It's great that interpretable machine learning merge (unite, join) the work of statisticians and machine learning specialists.

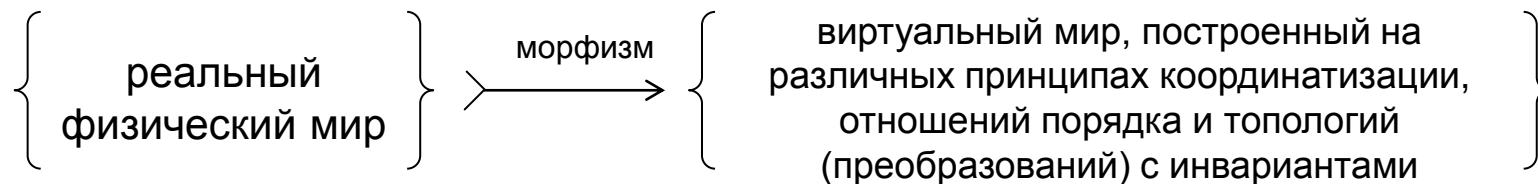


шаг 1



шаг 2

На множестве объектов с координатами задается алгебра кардиналов множеств: операции сложения и умножения, вычисления характеристических функций – предикатов: $ND \rightarrow 2D: f(x_1, x_2, \dots, x_n) = \begin{cases} 1 \\ 0 \end{cases}$





Так, **чтение или письмо** – есть тренировка для головного мозга, в особенности если при этом вы узнаете или выражаете нечто новое.

- A change in consciousness in the process of thinking leads to changes in the physical body of an intellectual subject.
- "Machine" will acquire the ability to think "if it acquires the properties of a" data-driven processor "