On identification and nonconstructivity

Iļja Kucevalovs

Faculty of Computing
University of Latvia

This research

- Applying **nonconstructive** computation methods to **identification**
  - Identification in the limit: Gold, 1967
  - Nonconstructive computation: Freivalds, 2009
- Definition on the most general level
- Both function and language learning are studied

Identification

- Also known as:
  - Identification in the limit
  - [Computational, machine] inductive inference
  - Algorithmic learning

  …et cetera
- Introduced by E.Mark Gold in 1967 as a model for human first language acquisition

Identification as the model for human language acquisition

- A newborn child does not speak any language
- So (s)he cannot be taught the language in terms of another language
- But eventually, (s)he learns some words
- Then some more
- Then some more…

Nonconstructive computation

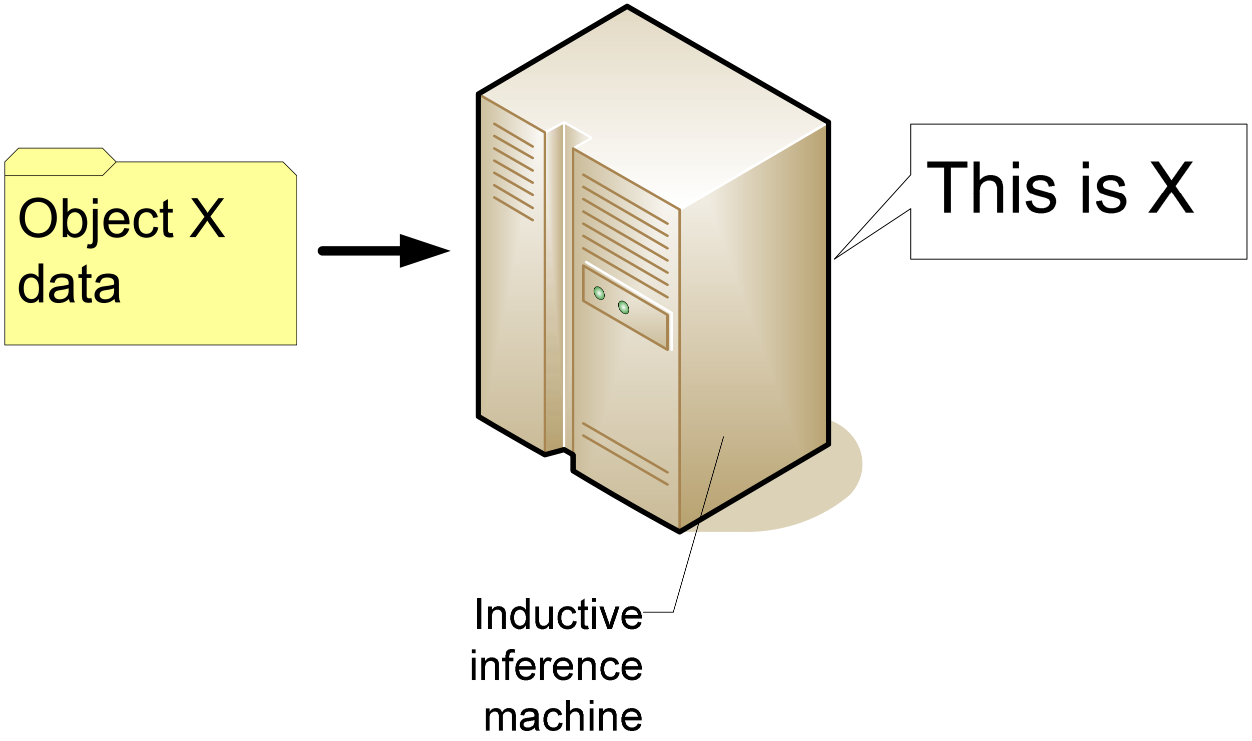
- Computation with additional information
- Defined so that trivial help is not allowed
- Based on Freivalds' observation of nonconstructive proofs

Nonconstructive identification

- Why?
  - Many classes are not identifiable
  - *R*: Class of all the total recursive functions
    - (from function graph)
  - Class of languages that contains all the finite and one infinite cardinality languages
    - (from positive data)

Inductive inference (general case)

- Generating hypotheses about some rule from examples



Computational inductive inference

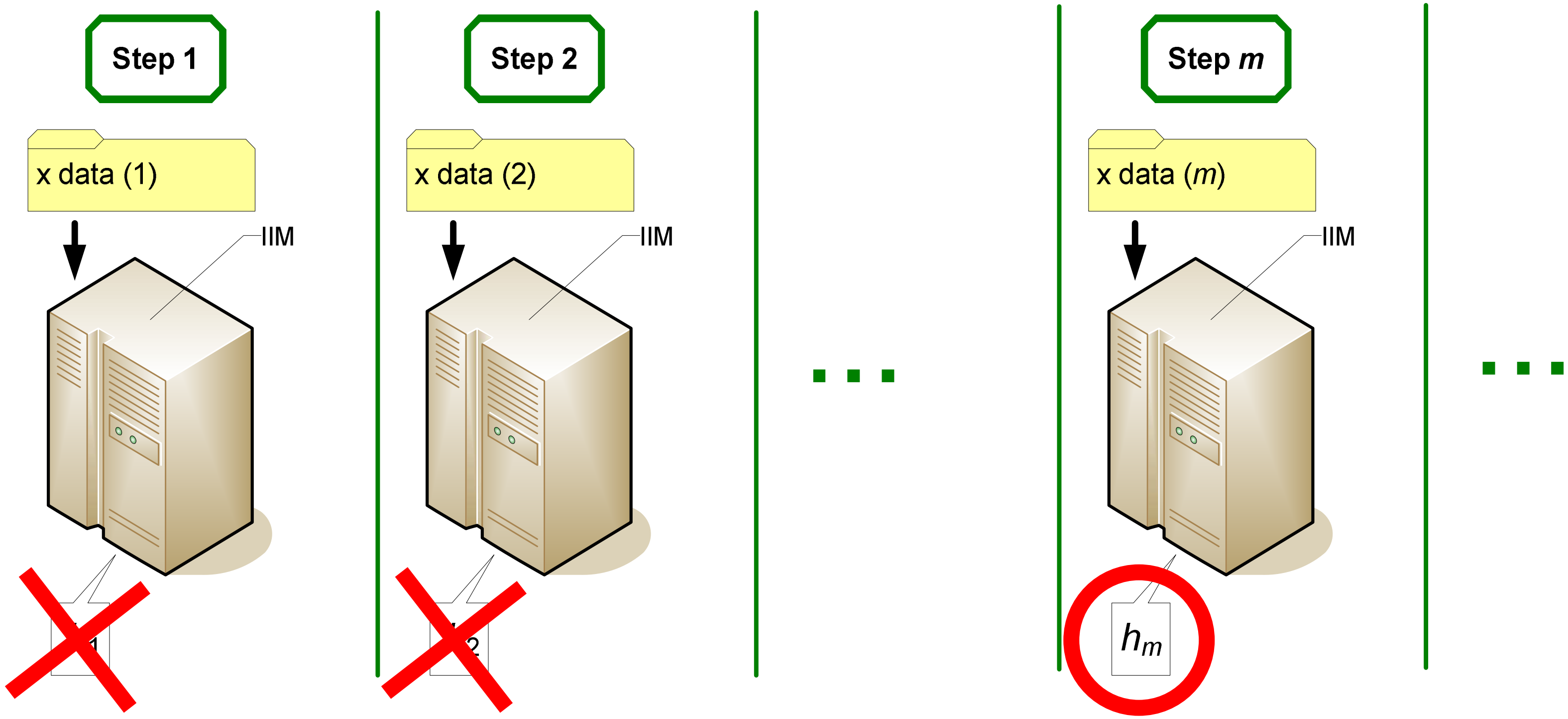
- All we work with is natural numbers
  - Information presentation is numbers
  - Objects are numbered
  - IIM is supposed to guess a number
- Time is quantized
- IIM may work for an infinitely long time

Computational inductive inference: Topics of study

- Objects of inference
  - Typically: (Formal) languages or (recursive) functions
- Types of information presentation
  - Typically: Positive or complete
  - **Additional information**
- **Successful inference criteria**
  - BC, EX, FIN and their variations
- IIM (inductive inference machines)
  - Deterministic, probabilistic, quantum
- **Inferable classes**

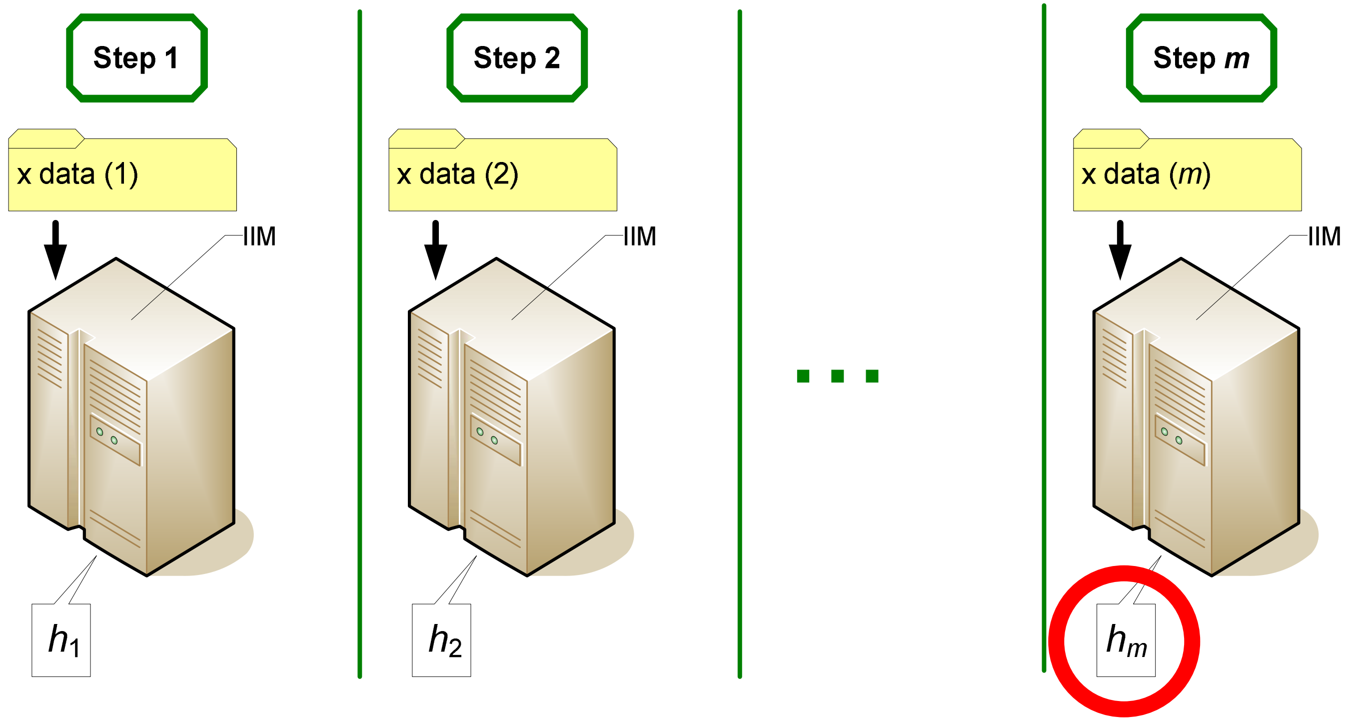
Criterion: BC (“Behaviourally correct”)

- Inference is successful, iff there is an infinite number of hypotheses and only a finite number of them is incorrect



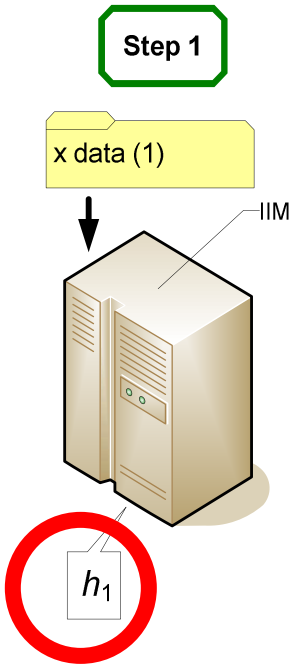
Criterion: EX ("Identification in the limit")

- Inference is successful, iff there is only a finite number of hypotheses and the last of them is correct



Criterion: FIN (“Finite identification”)

- Inference is successful, iff there is only one hypothesis, which is correct

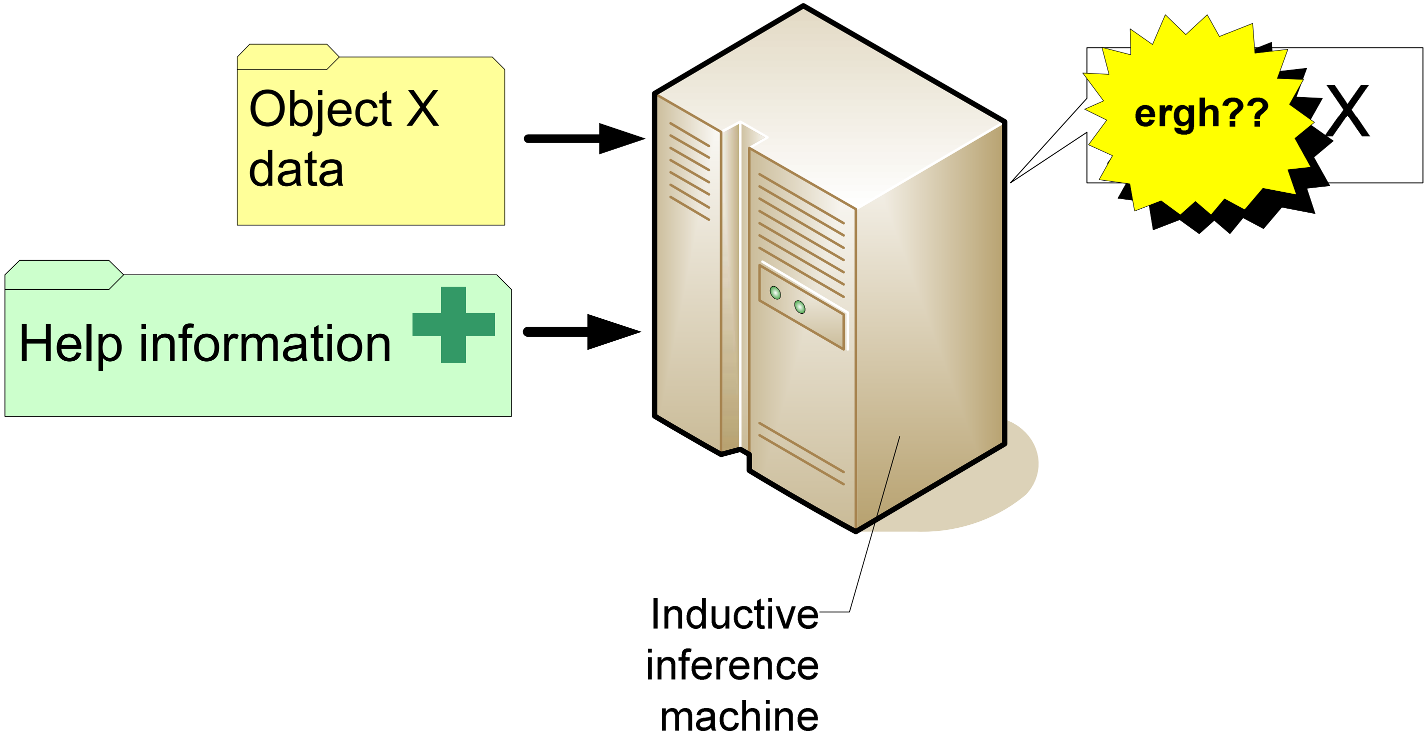


BC, EX, FIN versions

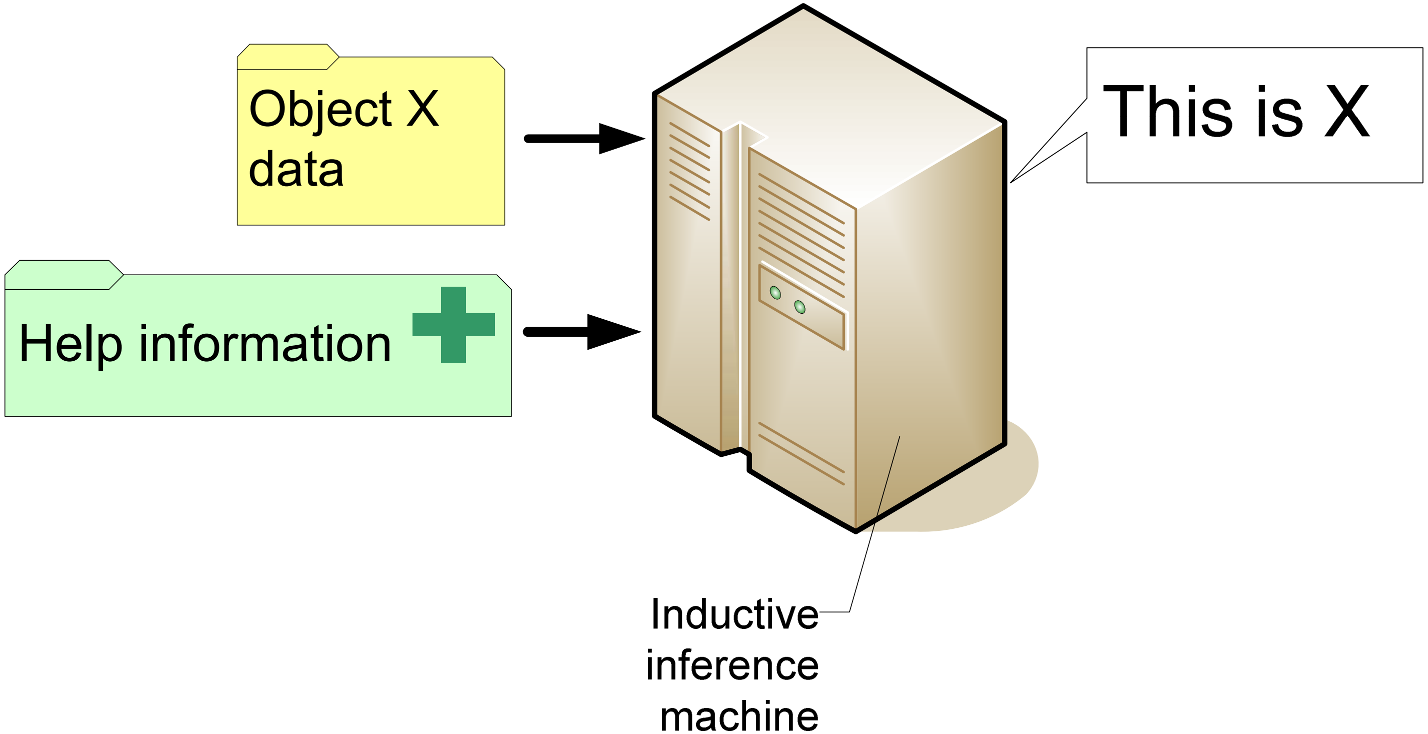
- X^n ≡ "*X* except on at most *n* anomalous inputs"
- X_n ≡ "*X* with at most *n* mindchanges"
- *MinX* (converges to the minimal possible number)
- ...et cetera.

Nonconstructive inductive inference (general case)

- An IIM is allowed to get some additional ("help") information about the object being identified

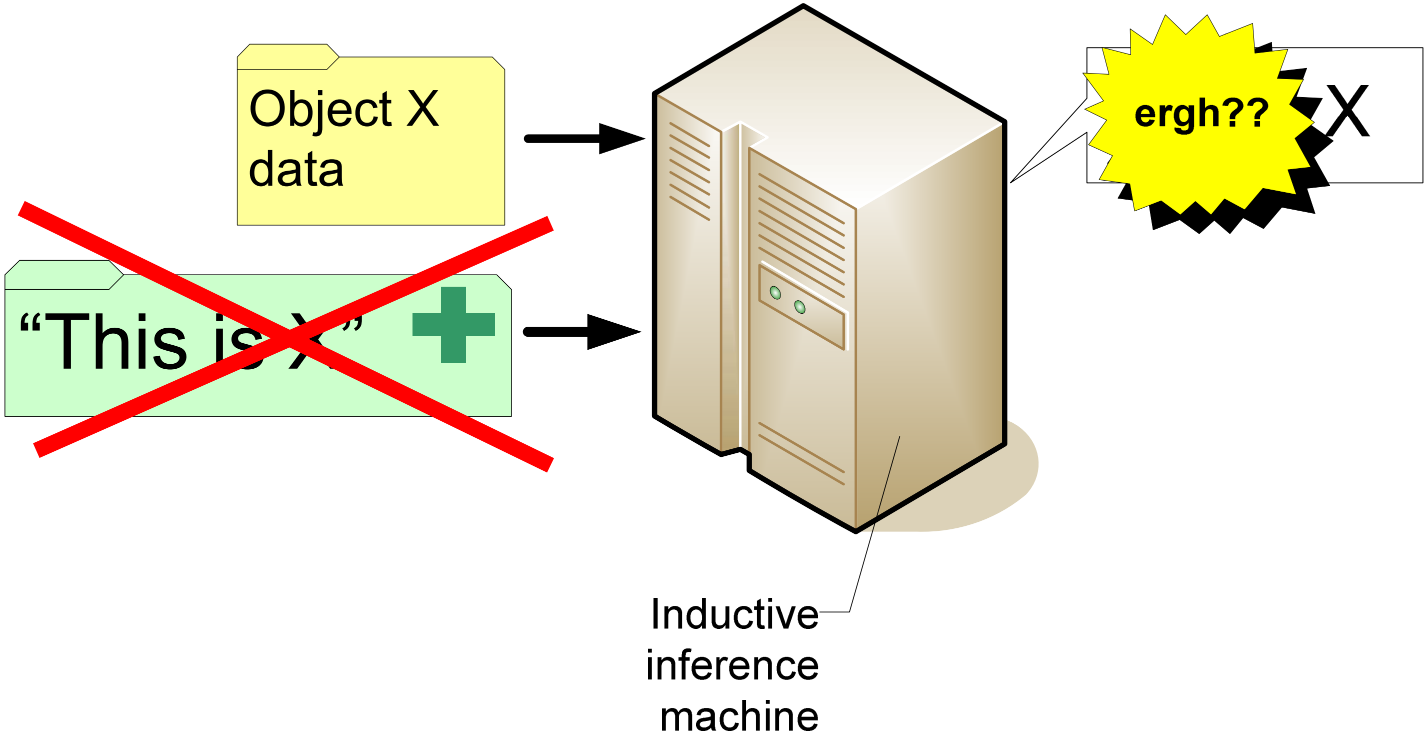


Nonconstructive inductive inference: situations to avoid (1)



Nonconstructive inductive inference: situations to avoid (2)

- If we don't put any restrictions on nonconstructive information...



Restriction #1: *K*-nonconstructivity

- Kolmogorov complexity of help information must differ more than by a constant from the correct answer
- I.e. for any $c \in \mathbb{N}$:
  $(\exists u \in U)[\ \exists p_0 \in p(u)\text{: } \forall n \in \{\ i \in \mathbb{N} \mid \varphi_i = u\ \}\ \mathrm{C}(p_0) < \mathrm{C}(n) - c\ ]$
- Or (which is equivalent)
  $(\exists u \in U)[\ \exists p_0 \in p(u)\text{: } C(p_0) < \min \{\ C(n) \mid \varphi_n = u, n \in \mathbb{N}\ \} - c\ ]$

Note on ***C*(*x*)**

- We consider **plain** Kolmogorov complexity of **natural numbers**
- C: $\mathbb{N} \rightarrow \mathbb{N}$
- $C(x)$ is the length of the minimal program that outputs x

Why *K*-nonconstructivity?

- Consider a language numbering φ_0, φ_1, φ_2, ...
- $g: \mathbb{N} \to \mathbb{N}$, $h: \mathbb{N} \to \mathbb{N}$
- $(\forall n \in \mathbb{N})\ [(g(2n) = 2n + 1) \wedge (g(2n + 1) = 2n)]$
- Class $U = \{ L \mid (\forall x \in \mathbb{N})[x \in L \Leftrightarrow g(x) \notin L] \}$
- $(\forall i \in \mathbb{N})\ [\varphi_{h(i)} = \{ x \mid g(x) \in \varphi_i \}]$
- $p(L) = p_0\ p_1\ p_2\ ...$:

  $[(\lim_{i\to\infty} p_i = j) \wedge (\varphi_j = \{ x \in \mathbb{N} \mid x \notin L \})]$
- Then for every language $\varphi_i \in U$ we have

  $C(i) \le C(h(p_\infty)) \le C(p_\infty) + C(h)$

where $p_\infty = \lim_{i\to\infty} p_i$

A simple lemma on *K*

- If the help information p: $U \to 2^{\mathbb{N}}$ is such that some $p_0 \in p(u)$ for infinitely many $u \in U$, then p is a *K*-help for *U* identification
- If some $p_0 \in p(u)$ for infinitely many $u \in U$, then these *u* have infinitely many indices
- Then $\min\{ C(n) \mid \varphi_n = u \in U, n \in \mathbb{N} \}$ is not limited from above
- Then for any *c* we have $(\exists u \in U)[\ \exists p_0 \in p(u): C(p_0) < \min\{ C(n) \mid \varphi_n = u, n \in \mathbb{N} \} - c\ ]$

But…

- Define a class U in $\varphi_0, \varphi_1, \varphi_2, \ldots$:
  $(\forall f \in U)\ [\ \exists m,n \in \mathbb{N}: \varphi_n = f,\ \varphi_n(m) = n\ ]$
- $p_u(f) = \{\ m \in \mathbb{N} \mid \varphi_n(m) = n,\ \varphi_n = f\ \}]$
- This is a K-help (from the previous lemma)
- R (total recursive function class) is K-identifiable with nonconstructivity amount $\lceil \log_2 n \rceil + 1$
- $p(f) = \text{If } (f \in U \cap R,\ p_u(f),\ i: \varphi_i = f)$

So…

- *R* is trivially *K*-identifiable
- We need something stronger

Restriction #2: *S*-nonconstructivity

- Kolmogorov complexity of help information must differ more than by a constant from the correct answer *for infinitely many objects*
- I.e. for any $c \in \mathbb{N}$:
  $(\forall^{\infty} u \in U)[\, \exists p_0 \in p(u): C(p_0) < \min \{ C(n) \mid \varphi_n = u, n \in \mathbb{N} \} - c \,]$
- Any *S*-help is a *K*-help
- Any *S*-identifiable class is *K*-identifiable

Theorem on constant *S*-nonconstructivity

- A class U is S-nonconstructively X-identifiable from presentation I with nonconstructivity n, iff U is a union

$$U = U_0 \cup U_1 \cup \ldots \cup U_{k-1}$$

- Moreover, $k \leq 2^{n+1-2}$ and each U_i is constructively X-identifiable
- X may be any constructive criterion
- U must have an infinite cardinality

S-nonconstructivity: Application #1

- There exist two classes such that each of them is identifiable, but their union is not
- (Independently discovered by Jānis Bārzdiņš and Lenore & Manuel Blum in the 1970s)

S-nonconstructivity: Application #2

- For any natural $n \geq 2$ there exist infinitely many language classes that are not *K*-*BC*-identifiable with nonconstructivity less than *n*, but are *S*-*EX*-identifiable with nonconstructivity *n*
- (Discovered by I.Kucevalovs in 2010, inspired by a ‘1988 paper by Mark Fulk)

Restriction #3: *F*-nonconstructivity

- Has appeared in literature before
- A **<u>nonconstructivity amount</u>** function $d(n)$ is defined
- Any help word of a length $d(n)$ must work for any input object having index n or less
- Essentially S.Jain and A.Sharma's generalized "learning with the knowledge on the program upper bound"

Identification in the *k*-limit

- An IIM outputs not a sequence, but a *k*-dimensional array of hypotheses
  - Always assumed to be infinite
  - Recall *EX* and *FIN*: we can always build IIMs which output infinite sequences

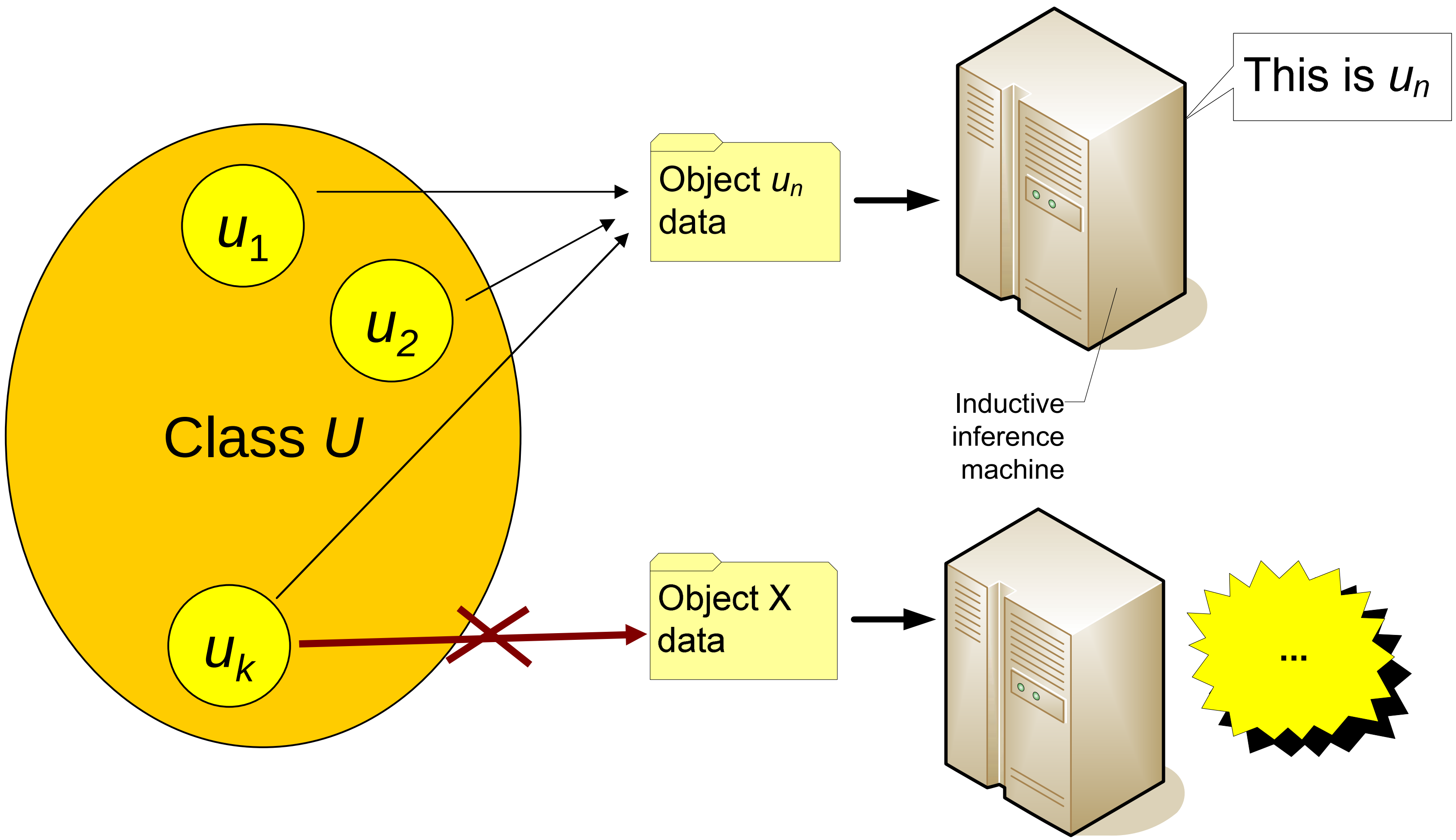
Identification in the *k*-limit (ctd.)

- Criteria are written in the form
  $(X_0 \times X_1 \times \ldots \times X_k)$
- E.g. $(BC \times EX)$

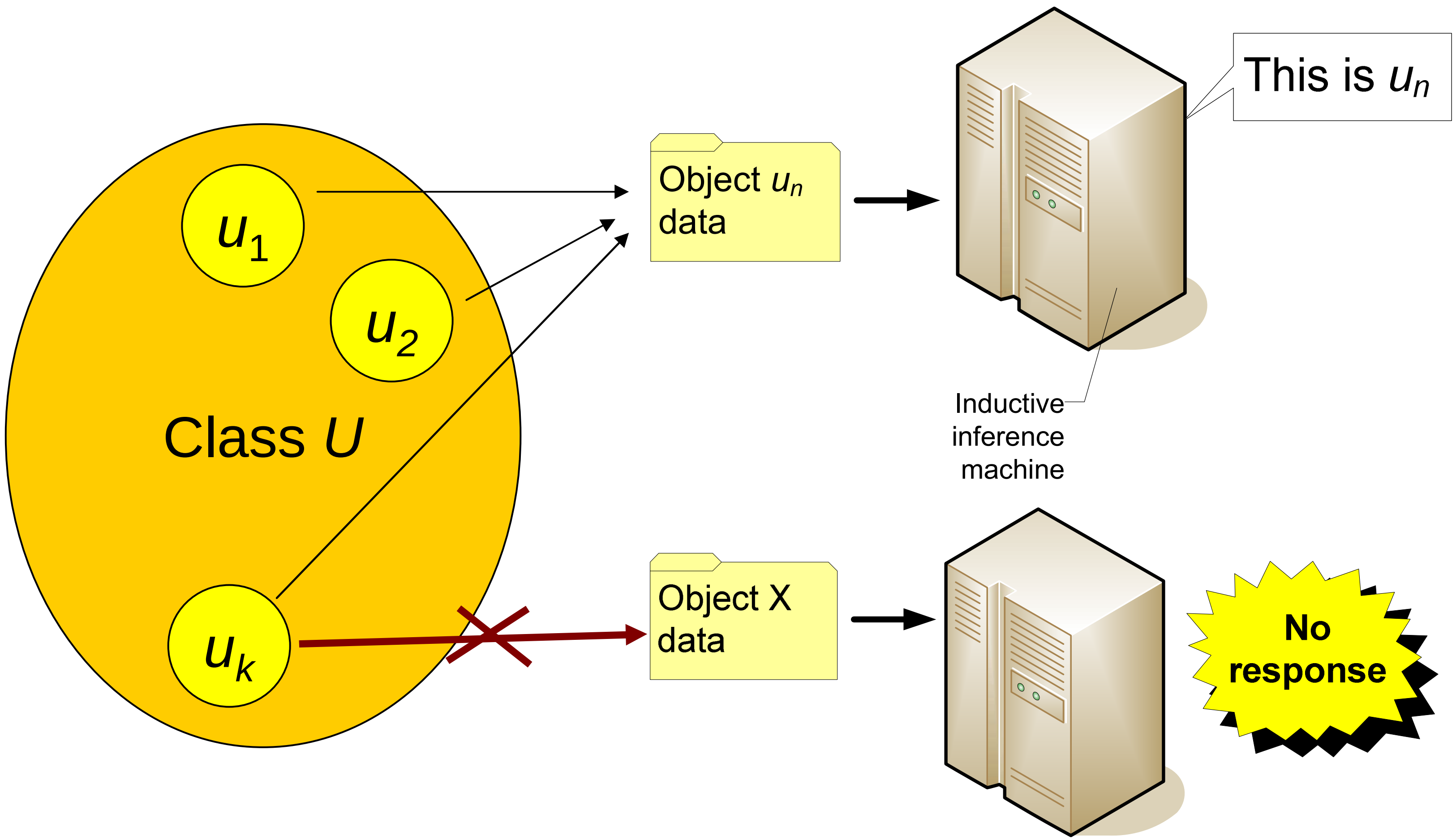
Basic lemma on the *k*-limit

- A class U is $(BC \times X_0 \times \ldots \times X_k)$-identifiable from presentation I in a numbering φ, iff there exists an infinite recursive sequence of *IIM* M s.t. for every $u \in U$:
  $\forall^\infty i \in \mathbb{N}: M_i(I(u)) \in (X_0 \times \ldots \times X_k)(u, \varphi)$
- Analogously for $(EX \times X_0 \times \ldots \times X_k)$
- In the case of recursive functions, it means that U is F-nonconstructively X_k-identifiable

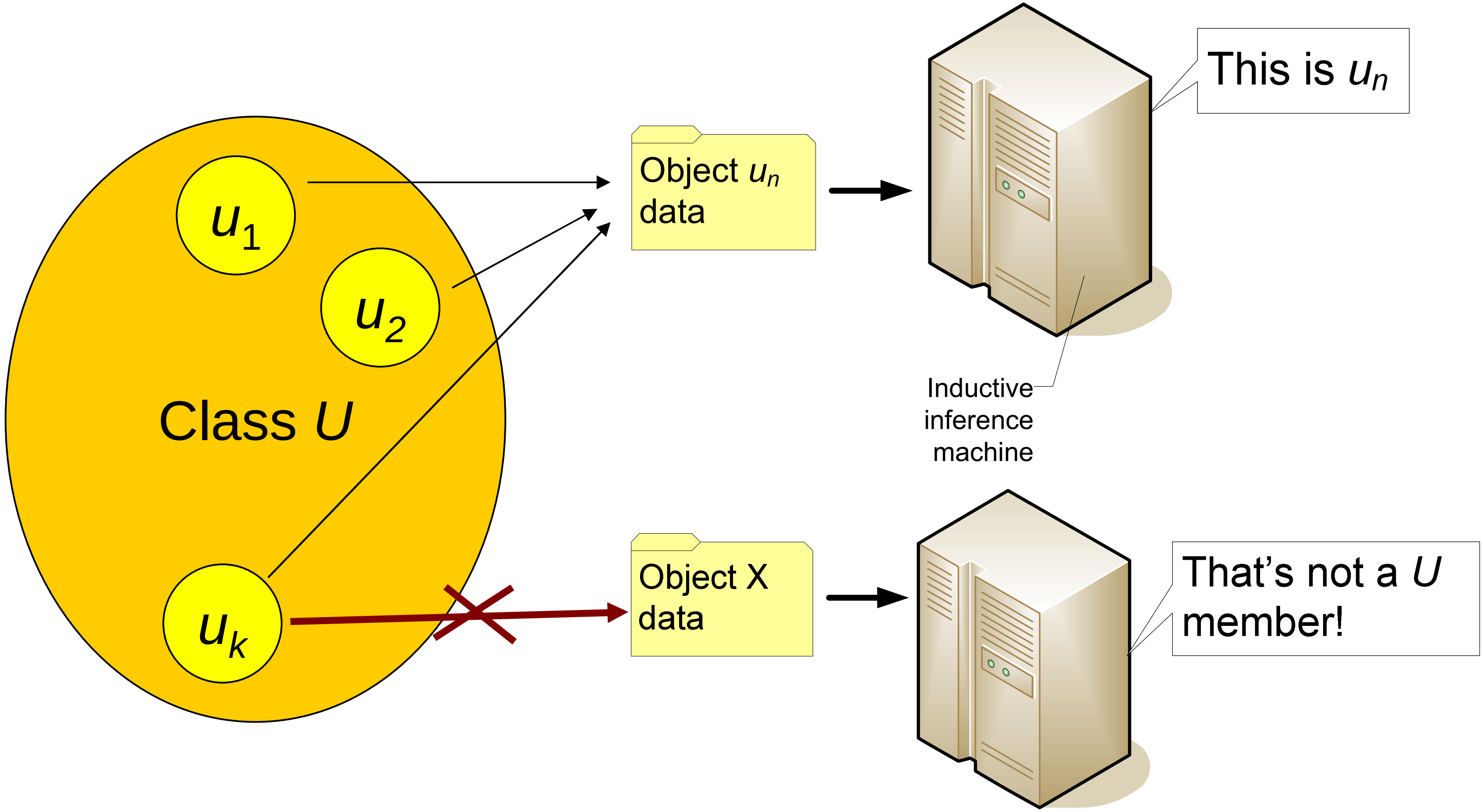
Reliable and refutable identification
(I) Non-reliable identification



Reliable and refutable identification
(II) Reliable identification



Reliable and refutable identification (III) Refutable identification



R-NK-identification models

Reliability of identification

The given object

1) must be a member of the class in question

2) can be a non-member

Reliability of nonconstructivity

The given help

a) must be correct

b) can be incorrect

	1)	2)
a)	**NK-X**	**NK-R-X**
b)	**R-NK-X**	***R-NK-R-X***

Big question (R-NK-X)

- Can it be?
  - Some class is not constructively identifiable
  - But if we get some help, it is identifiable
  - *Even if the help is incorrect*

The answer

- Yes, with certain restrictions on the error
- The error, however, may grow to infinity

The construction (part 1)

- Consider the following functions:

  $h(0) = C(1024)$

  $h(x) = \min\{ n \mid n>h(x\text{-}1) \wedge C(n)>C(x\text{-}1) \}$

  $m(x) = \min\{ C(n) \mid n \geq x \}$

- Both do exist
- Neither is computable

The construction (part 2)

- $p_0, p_1, \ldots$ is a growing sequence of primes starting from 3
- For every natural k, define $f_k(x) \equiv h((p_k)^{x+1})$
- Define the numbering $w_i = f_k$ for such $j \geq k$ that $f_k(n) = i$ for some n in $h(j) \pm \lfloor m(j)/2 \rfloor$

The construction (idea)

k \ x	$f_k(0)$	$f_k(1)$	$f_k(2)$	$f_k(3)$	$f_k(4)$	…
f_0	3	9	27	81	243	…
f_1	5	25	125	625	3125	…
f_2	7	49	343	2401	16807	…
f_3	11	121	1331	14641	161051	…
…	…	…	…	…	…	…

The construction (idea, ctd)

k \ x	…	n	n+1	n+2	n+3	n+4	n+5	n+6	…
f_0	…	p_0^{n+1}	p_0^{n+2}	p_0^{n+3}	p_0^{n+4}	p_0^{n+5}	p_0^{n+6}	p_0^{n+7}	…
f_1	…	p_1^{n+1}	p_1^{n+2}	p_1^{n+3}	p_1^{n+4}	p_1^{n+5}	p_1^{n+6}	p_1^{n+7}	…
f_2	…	p_2^{n+1}	p_2^{n+2}	p_2^{n+3}	p_2^{n+4}	p_2^{n+5}	p_2^{n+6}	p_2^{n+7}	…
f_3	…	p_3^{n+1}	p_3^{n+2}	p_3^{n+3}	p_3^{n+4}	p_3^{n+5}	p_3^{n+6}	p_3^{n+7}	…
f_4	…	p_4^{n+1}	p_4^{n+2}	p_4^{n+3}	p_4^{n+4}	p_4^{n+5}	p_4^{n+6}	p_4^{n+7}	…
f_5	…	p_5^{n+1}	p_5^{n+2}	p_5^{n+3}	p_5^{n+4}	p_5^{n+5}	p_5^{n+6}	p_5^{n+7}	…
f_6	…	p_6^{n+1}	p_6^{n+2}	p_6^{n+3}	p_6^{n+4}	p_6^{n+5}	p_6^{n+6}	p_6^{n+7}	…
f_7	…	p_7^{n+1}	p_7^{n+2}	p_7^{n+3}	p_7^{n+4}	p_7^{n+5}	p_7^{n+6}	p_7^{n+7}	…
…	…	…	…	…	…	…	…	…	…

if these are some *h* values

k \ x	…	n	n+1	n+2	n+3	n+4	n+5	n+6	…
f_0	…	p_0^{n+1}	p_0^{n+2}	p_0^{n+3}	p_0^{n+4}	p_0^{n+5}	p_0^{n+6}	p_0^{n+7}	…
f_1	…	p_1^{n+1}	p_1^{n+2}	p_1^{n+3}	p_1^{n+4}	p_1^{n+5}	p_1^{n+6}	p_1^{n+7}	…
f_2	…	p_2^{n+1}	p_2^{n+2}	p_2^{n+3}	p_2^{n+4}	p_2^{n+5}	p_2^{n+6}	p_2^{n+7}	…
f_3	…	p_3^{n+1}	p_3^{n+2}	p_3^{n+3}	p_3^{n+4}	p_3^{n+5}	p_3^{n+6}	p_3^{n+7}	…
f_4	…	p_4^{n+1}	p_4^{n+2}	p_4^{n+3}	p_4^{n+4}	p_4^{n+5}	p_4^{n+6}	p_4^{n+7}	…
f_5	…	p_5^{n+1}	p_5^{n+2}	p_5^{n+3}	p_5^{n+4}	p_5^{n+5}	p_5^{n+6}	p_5^{n+7}	…
f_6	…	p_6^{n+1}	p_6^{n+2}	p_6^{n+3}	p_6^{n+4}	p_6^{n+5}	p_6^{n+6}	p_6^{n+7}	…
f_7	…	p_7^{n+1}	p_7^{n+2}	p_7^{n+3}	p_7^{n+4}	p_7^{n+5}	p_7^{n+6}	p_7^{n+7}	…
…	…	…	…	…	…	…	…	…	…

The construction (idea, final)

- Now, take the help equal to *h*
- We get *F*-nonconstructive *FIN*-identifiability
- If we take function values from the argument equal not to a single value of *h*, but to a interval bounded by *m*, we can allow an error
- Moreover, this error grows to infinity
- ...but incomputably slowly

Literature

- I.Kucevalovs. “*Nekonstruktivitātes daudzums induktīvajā izvedumā*”, master thesis, University of Latvia, 2010
- I.Kucevalovs. “*Randomization vs. Amount of nonconstructivity in learning of recursive functions*”, “Randomized and Quantum Computation”, MFCS+CSL, 2010.
- I.Kucevalovs. “*On reliability and refutability in nonconstructive identification*”, to appear in Proceedings of MEMICS 2010 (holds in Mikulov, CZ on 2010.10.22-24)

Thank you for your attention