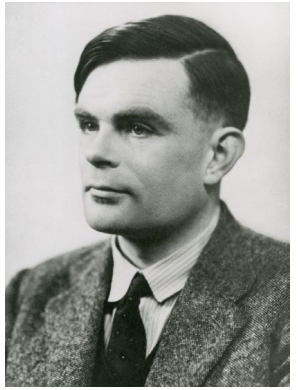
Educability

Leslie Valiant
Harvard University

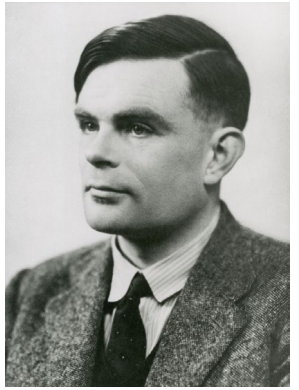
Beyond the Symbols vs Signals Debate
The Royal Society
October 28, 2024

The “Turing Gap”

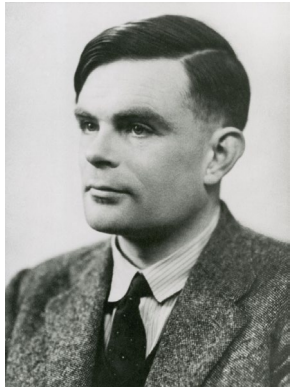


The “Turing Gap”

Can every task done by humans be done by a machine?



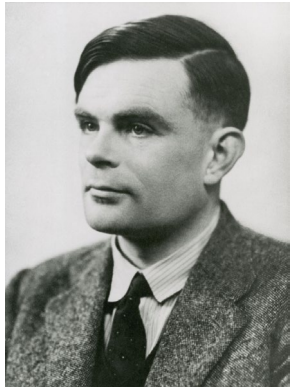
The “Turing Gap”



Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine.*

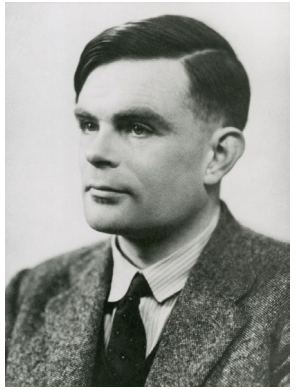
The “Turing Gap”



Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good English, or that it could not be influenced by sex-appeal or smoke a pipe.*

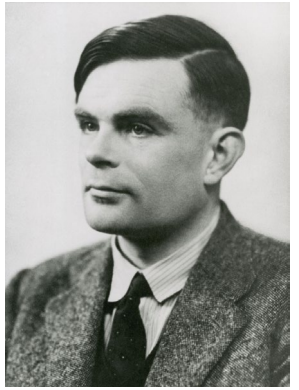
The “Turing Gap”



Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good English, or that it could not be influenced by sex-appeal or smoke a pipe. I cannot offer any such comfort, for I believe that no such bounds can be set.*

The “Turing Gap”

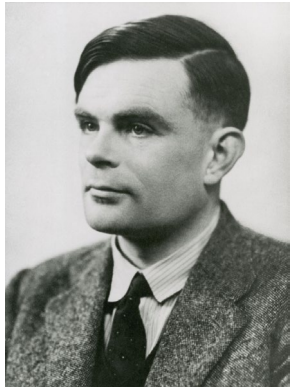


Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good English, or that it could not be influenced by sex-appeal or smoke a pipe. I cannot offer any such comfort, for I believe that no such bounds can be set.*

(Follows from Church-Turing Thesis.)

The “Turing Gap”



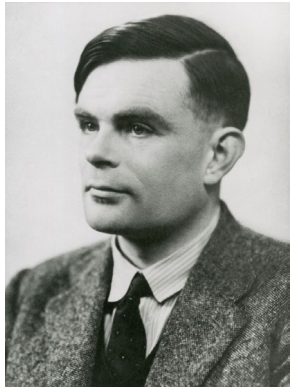
Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good English, or that it could not be influenced by sex-appeal or smoke a pipe. I cannot offer any such comfort, for I believe that no such bounds can be set.*

(Follows from Church-Turing Thesis.)

But what tasks CAN humans do?

The “Turing Gap”



Can every task done by humans be done by a machine?

Alan Turing (1951): *It is customary, in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good English, or that it could not be influenced by sex-appeal or smoke a pipe. I cannot offer any such comfort, for I believe that no such bounds can be set.*

(Follows from Church-Turing Thesis.)

But what tasks CAN humans do?

c.f. The Imitation Game

How adequate is our understanding of Human Capabilities?

How about human intelligence?

How adequate is our understanding of Human Capabilities?

How about human intelligence?

Board of Scientific Affairs of the American Psychological Association, 1994:

"... when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen, somewhat different, definitions."

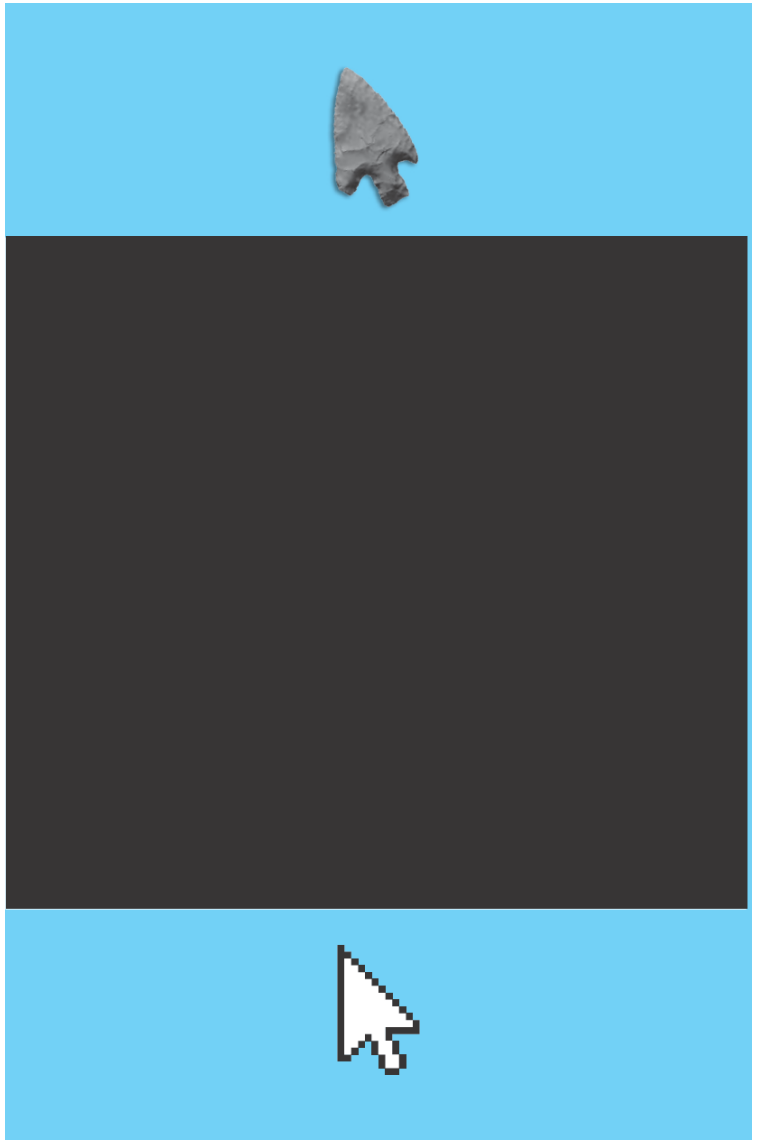
How adequate is our understanding of Human Capabilities?

How about human intelligence?

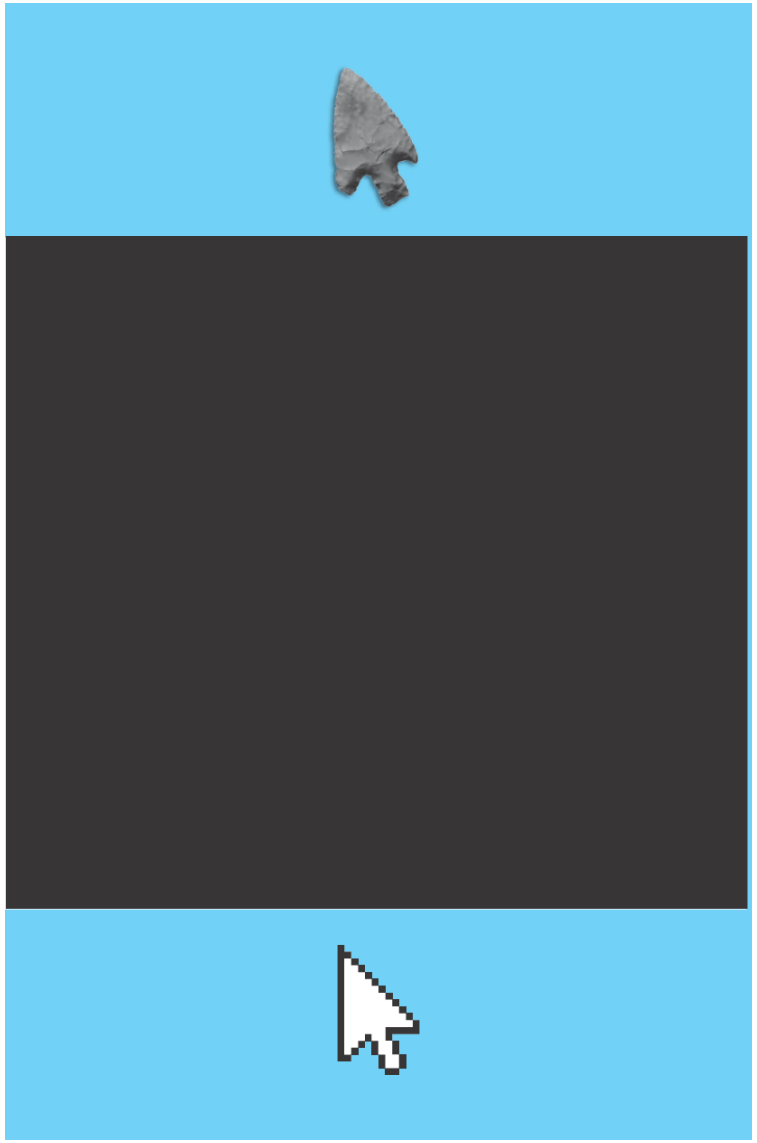
Board of Scientific Affairs of the American Psychological Association, 1994:

"... when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen, somewhat different, definitions."

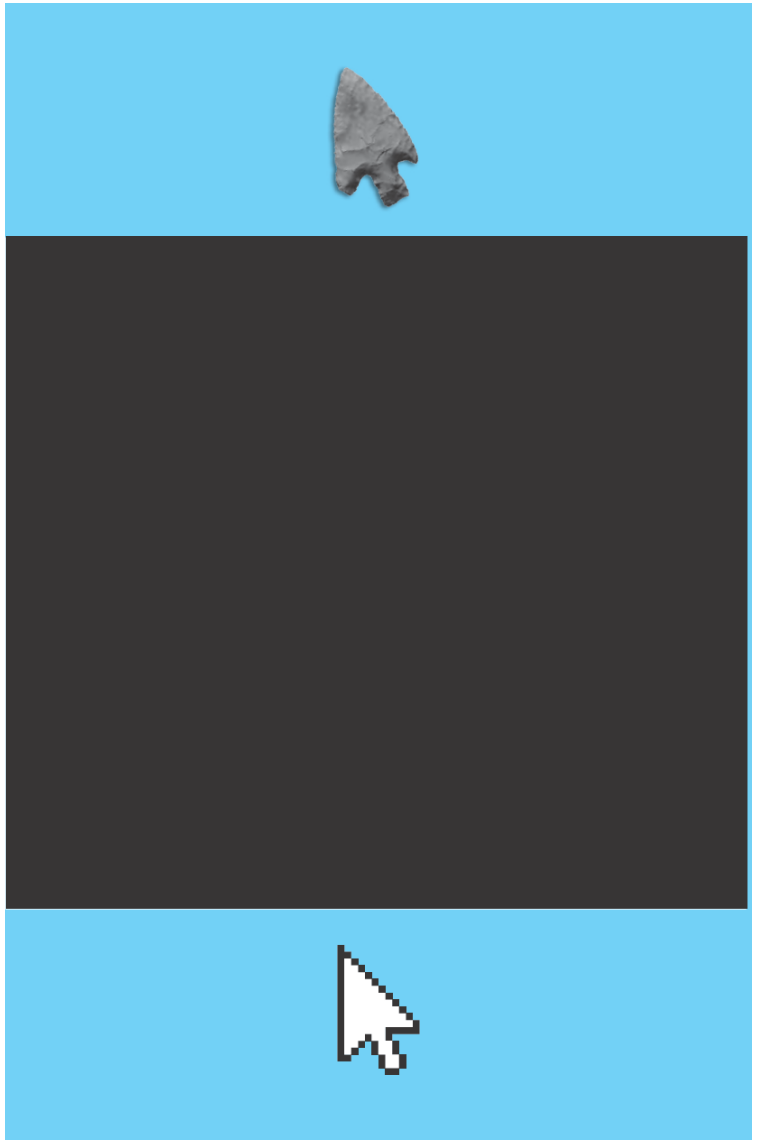
Defined only implicitly by Spearman [1904].



What question to ask about human capabilities?

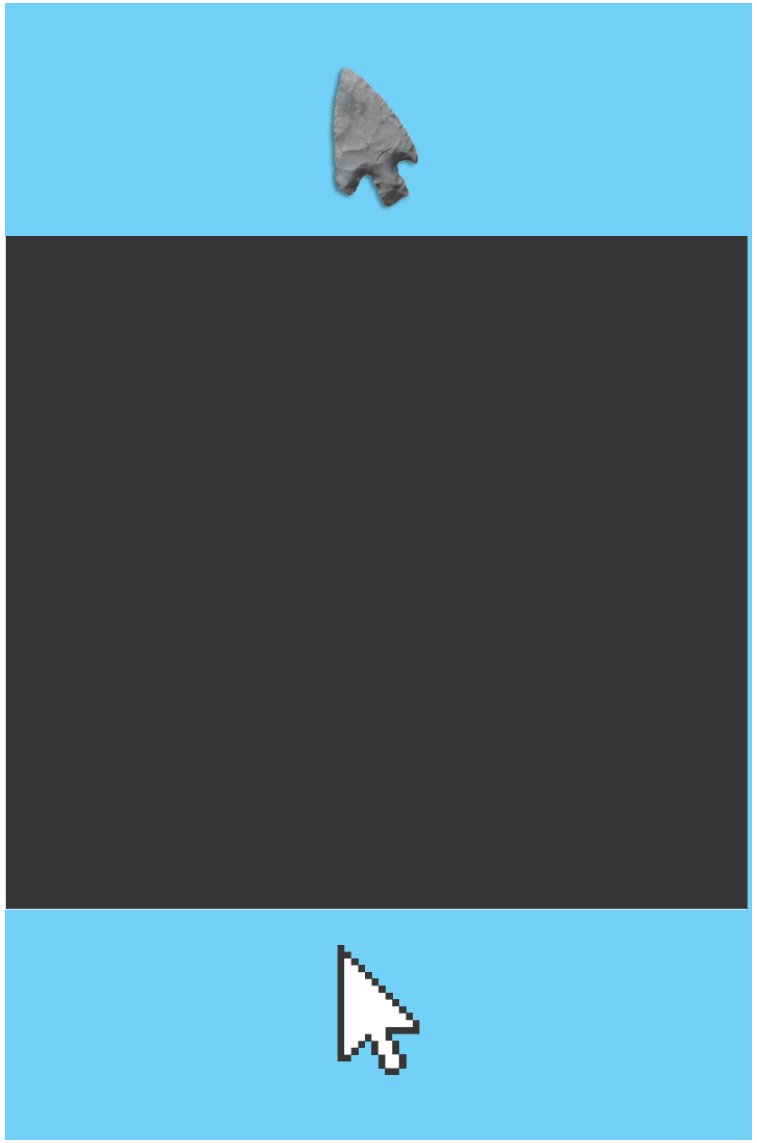


What is the **cognitive capability** that enabled humans to go from one to the other?



What is the **cognitive capability** that enabled humans to go from one to the other?

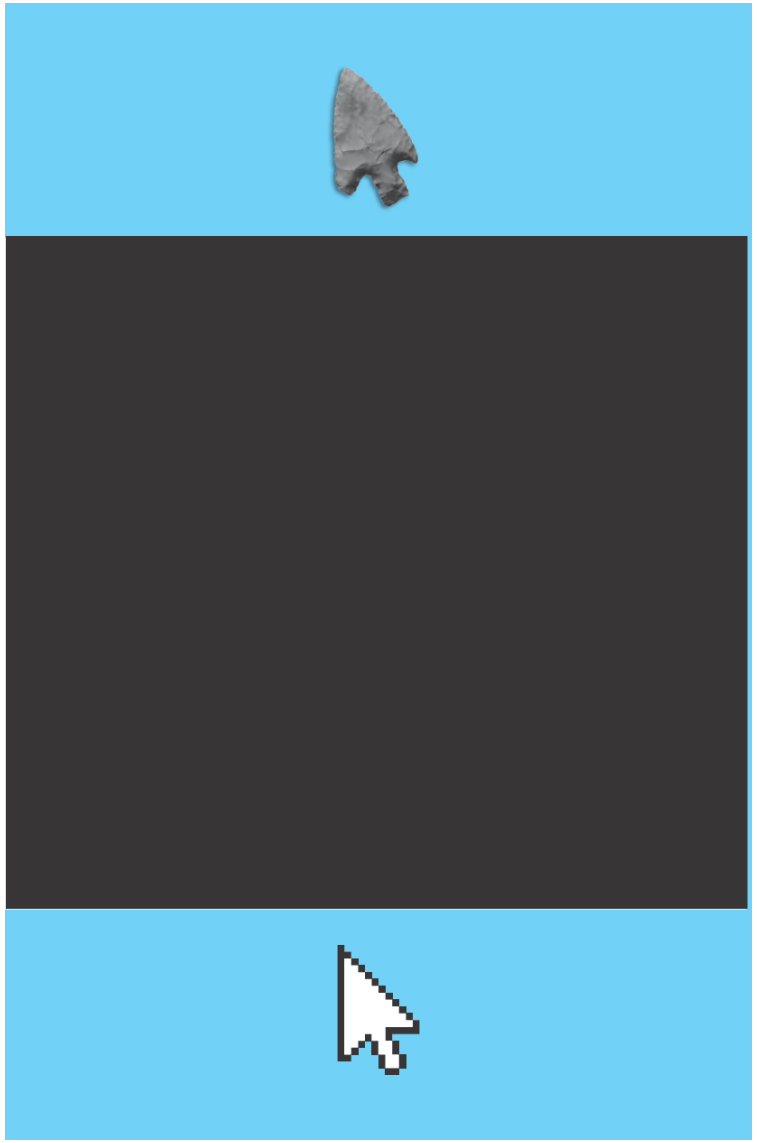
Call it the “Civilization Enabler”.



What is the **cognitive capability** that enabled humans to go from one to the other?

Call it the “Civilization Enabler”.

(Not vision, locomotion, emotions, …)

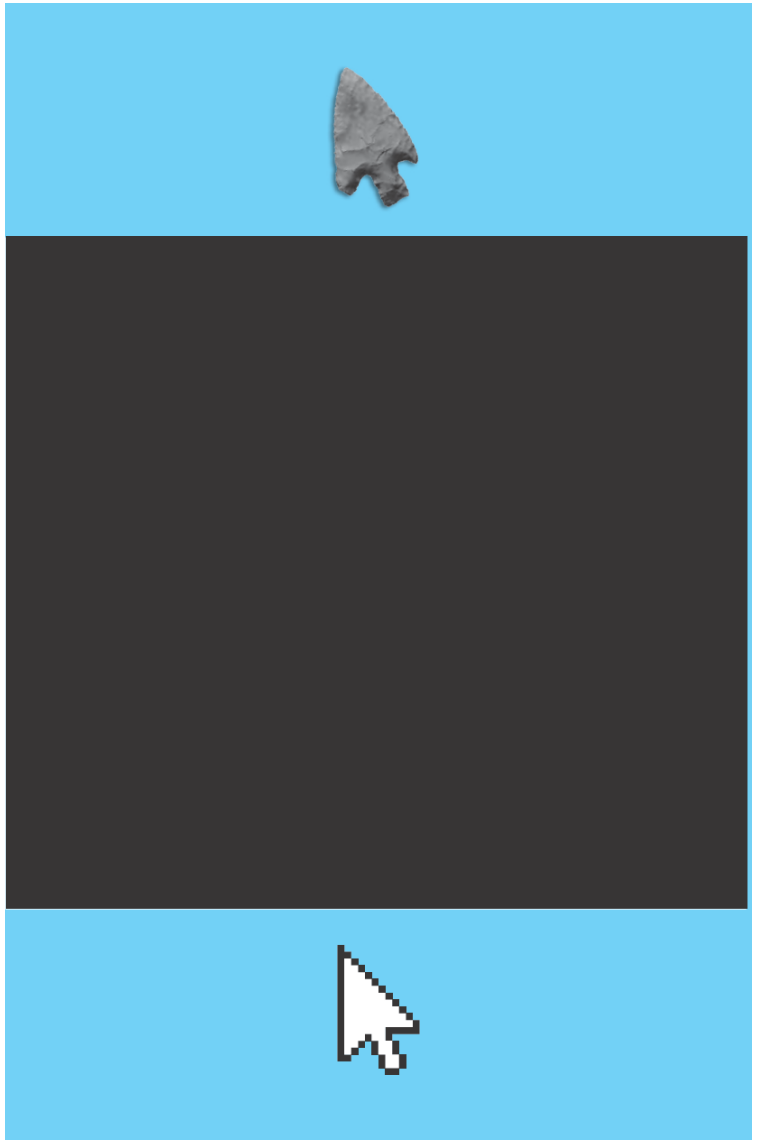


What is the **cognitive capability** that enabled humans to go from one to the other?

Call it the "Civilization Enabler".

(Not vision, locomotion, emotions, …)

Methodology: Robust Models of Computation:



What is the **cognitive capability** that enabled humans to go from one to the other?

Call it the "Civilization Enabler".

(Not vision, locomotion, emotions, …)

Methodology: Robust Models of Computation:
(i) functionally well-defined and
(ii) computationally feasible

Previous Approaches to Human Uniqueness

Hundreds of differences between humans and non-human hominids have been studied by scientists.

Previous Approaches to Human Uniqueness

Hundreds of differences between humans and non-human hominids have been studied by scientists.

Here seeking the special difference, the **Civilization Enabler,** the *cognitive capability* that enabled humans to achieve the civilization we have?

Previous Approaches to Human Uniqueness

Hundreds of differences between humans and non-human hominids have been studied by scientists.

Here seeking the special difference, the **Civilization Enabler,** the *cognitive capability* that enabled humans to achieve the civilization we have?

Not: eyebrows, animal domestication, brain lateralization, collaboration, social learning, ……., intelligence, ……, symbols, language, …..

Previous Approaches to Human Uniqueness

Hundreds of differences between humans and non-human hominids have been studied by scientists.

Here seeking the special difference, the **Civilization Enabler,** the *cognitive capability* that enabled humans to achieve the civilization we have?

Not: eyebrows, animal domestication, brain lateralization, collaboration, social learning, ……., intelligence, ……, symbols, language, …..

?Darwin

Previous Approaches to Human Uniqueness

Hundreds of differences between humans and non-human hominids have been studied by scientists.

Here seeking the special difference, the **Civilization Enabler,** the *cognitive capability* that enabled humans to achieve the civilization we have?

Not: eyebrows, animal domestication, brain lateralization, collaboration, social learning, ……., intelligence, ……, symbols, language, …..

?Darwin

?Models of Computation make complex hypotheses expressible.

Proposed Civilization Enabler:

Proposed Civilization Enabler: Educability

Informal Definition of Educability

Educability consists of capabilities to

Informal Definition of Educability

Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,

Informal Definition of Educability

Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,

(ii) *apply* beliefs in *combination* to specific situations,

Informal Definition of Educability

Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,

(ii) *apply* beliefs in *combination* to specific situations,

(iii) *acquire* beliefs *explicitly described* to us by others,

Informal Definition of Educability

Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,

(ii) *apply* beliefs in *combination* to specific situations,

(iii) *acquire* beliefs *explicitly described* to us by others,

all *integrated* with *(iv) multi-object scene analysis*

and *(v) symbolic naming*.

Informal Definition of Educability

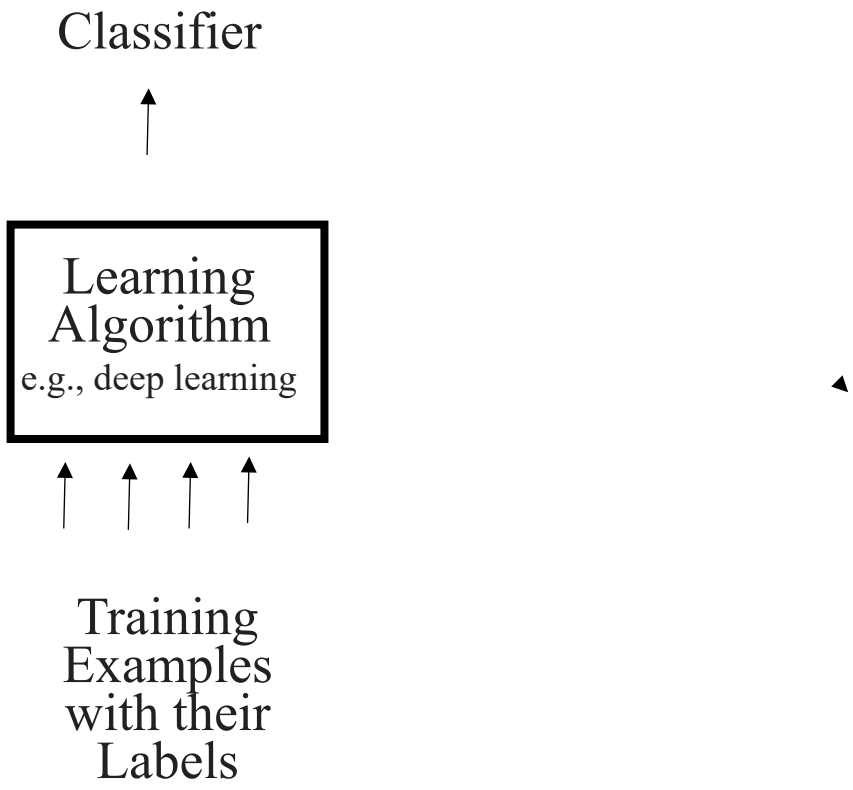
Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,
(ii) *apply* beliefs in *combination* to specific situations,
(iii) *acquire* beliefs *explicitly described* to us by others,
all *integrated* with *(iv) multi-object scene analysis*
and *(v) symbolic naming*.

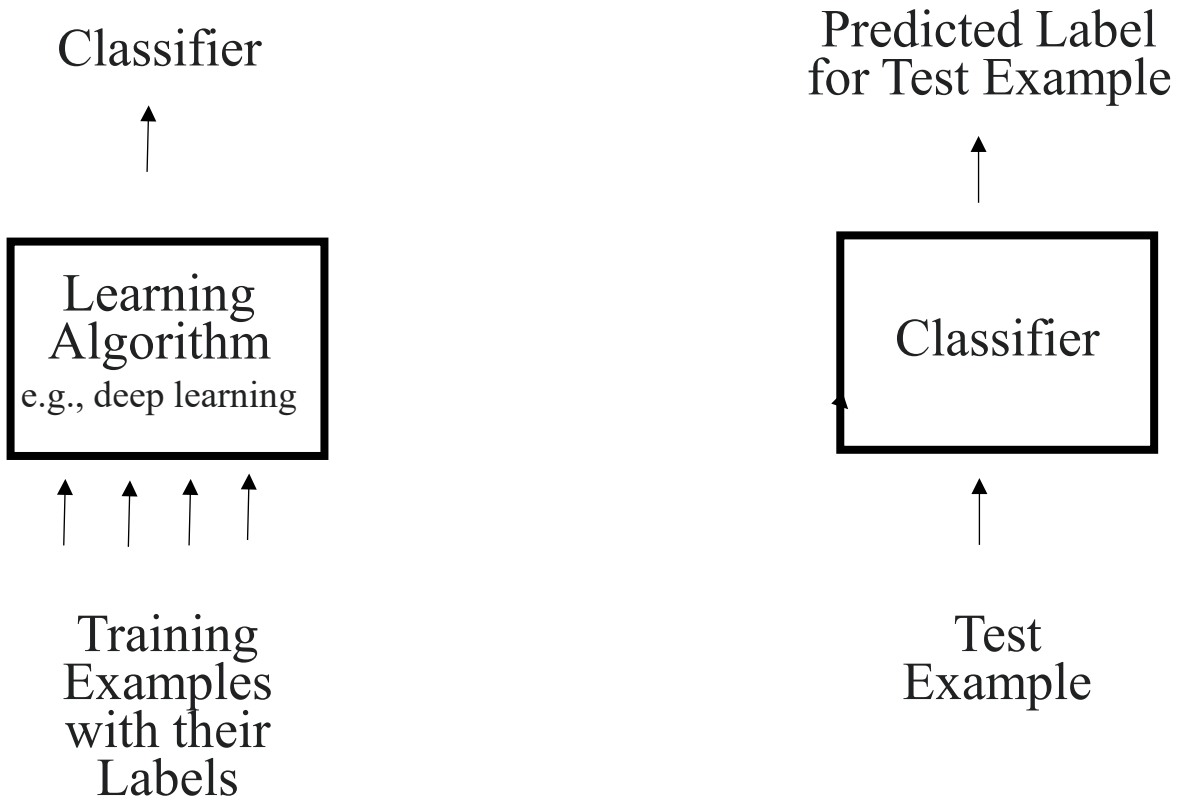
(c.f., educational philosophy)

(I) Generalization from Experience

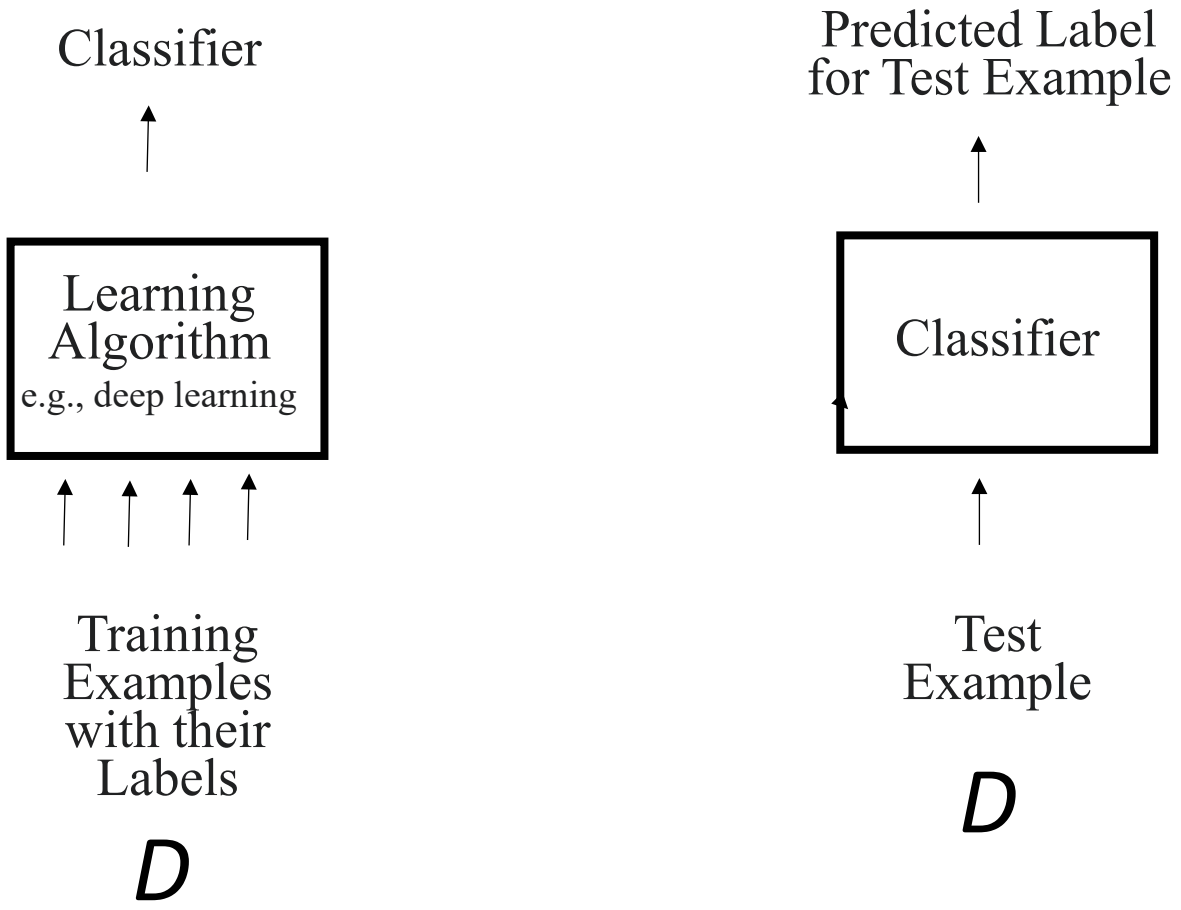
(I) Generalization from Experience



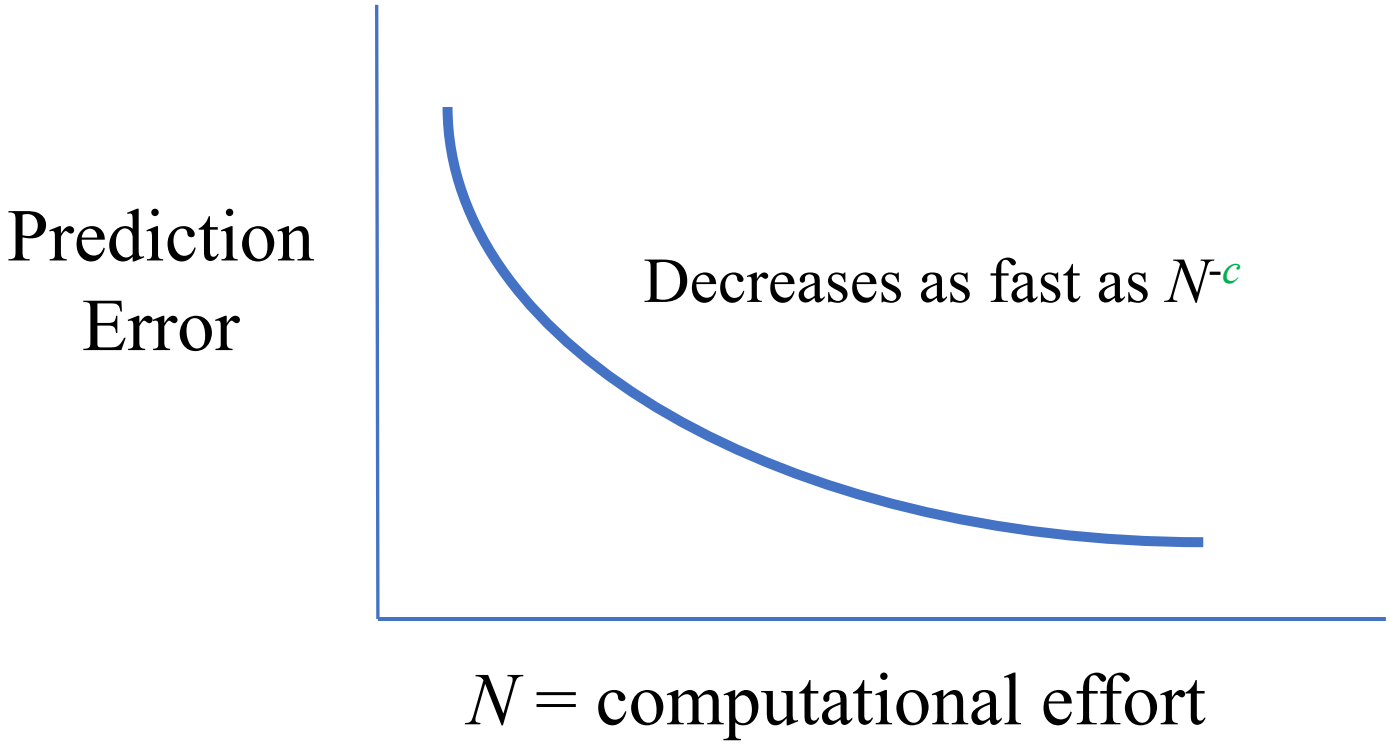
(I) Generalization from Experience



(I) Generalization from Experience

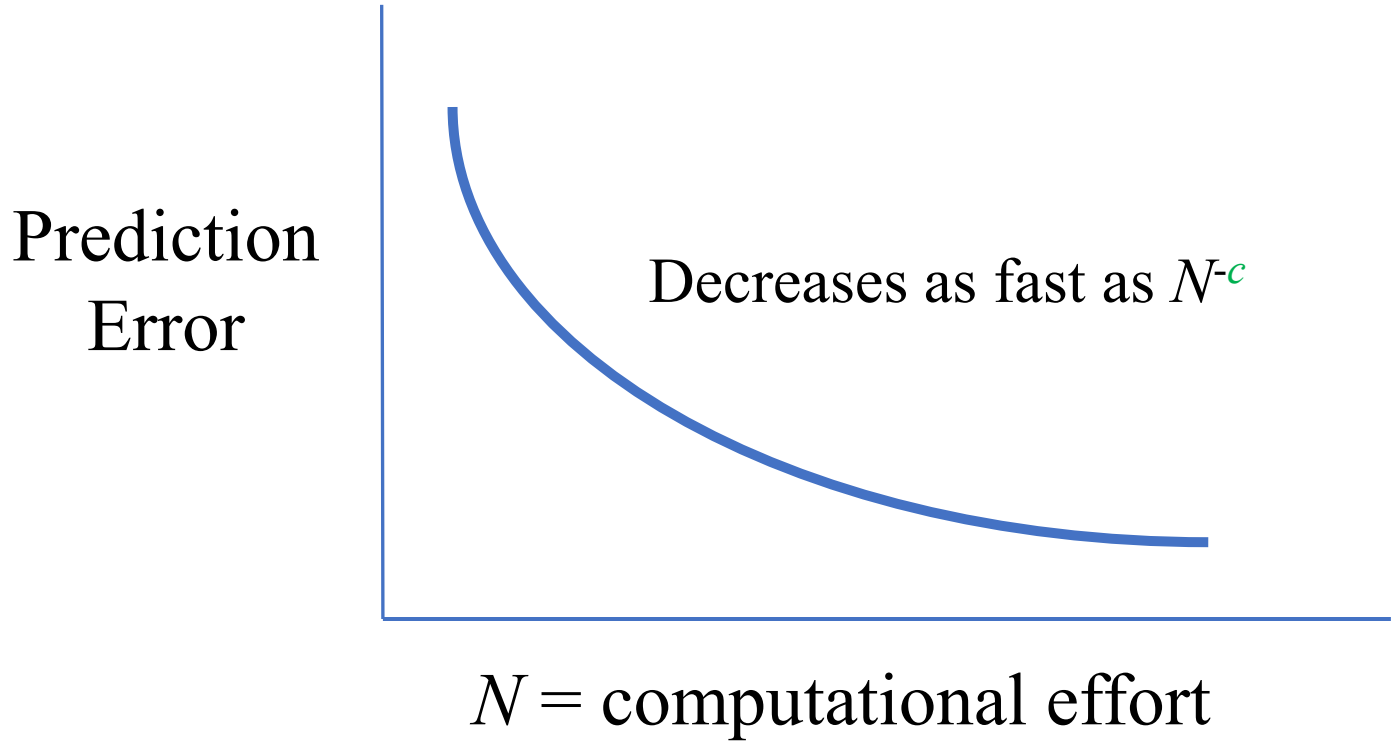


Probably Approximately Correct (PAC) Learning



PAC Learning: Better prediction with more effort, and affordably so!

Probably Approximately Correct (PAC) Learning



e.g. If with 10x more effort we halve the error,
then with another 10x more effort we halve the error again.

Precision versus Informality

“PAC Learning” versus “(Self-)Supervised Learning”

Precision versus Informality

"PAC Learning" versus "(Self-)Supervised Learning"

Generalization

Precision versus Informality

“PAC Learning”	versus	“(Self-)Supervised Learning”
Generalization		**Generalization,** **Explainability** **No Spurious Correlations** **No Adversarial Examples** **Lifelong Learning** **Out of distribution …** **……**

Precision versus Informality

"PAC Learning" versus "(Self-)Supervised Learning"

Generalization

Generalization,
Explainability
No Spurious Correlations
No Adversarial Examples
Lifelong Learning
Out of distribution …
……

Large Language Models are trained to predict next token

Precision versus Informality

"PAC Learning" versus "(Self-)Supervised Learning"

Generalization

Generalization,
Explainability
No Spurious Correlations
No Adversarial Examples
Lifelong Learning
Out of distribution …
……

Large Language Models are trained to predict next token

Do they do anything else reliably???

Think of Newton's 3 Laws, Maxwell's 4 equations.

Several principles whose interaction produces the richness of the phenomena.

(II) Chaining Beliefs (Reasoning)

(II) Chaining Beliefs (Reasoning)

Phenomenon addressed here: Chaining on uncertain beliefs.

(II) Chaining Beliefs (Reasoning)

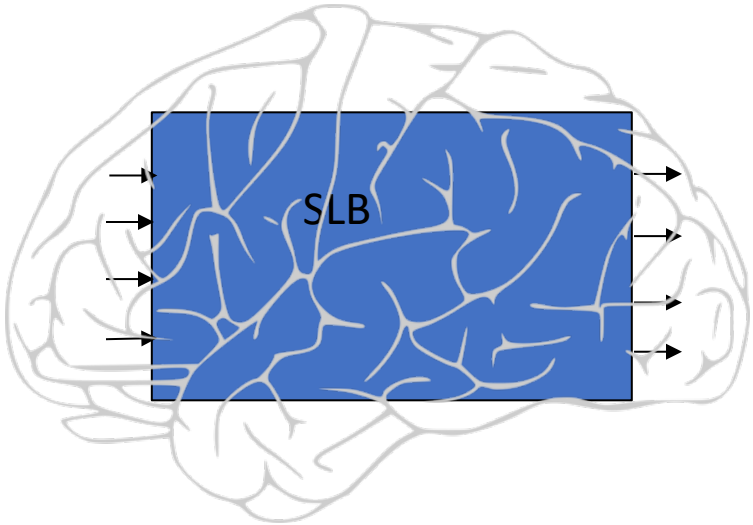
Phenomenon addressed here: Chaining on uncertain beliefs.

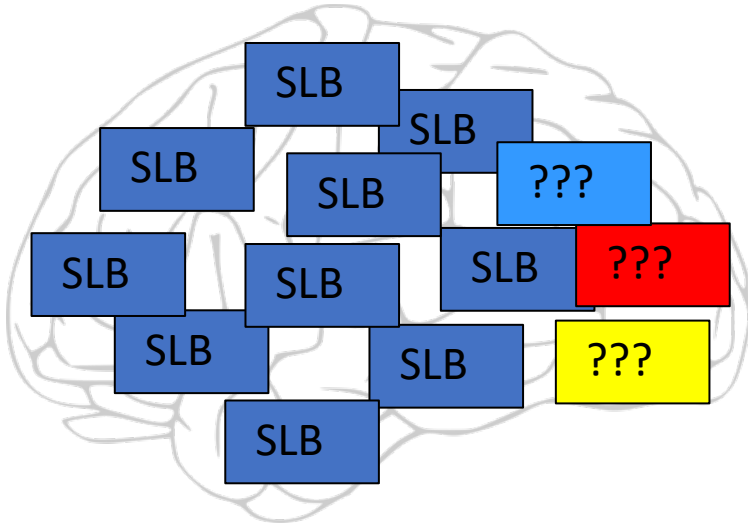
Formulation: Robust Logic

Soundness of conclusions??

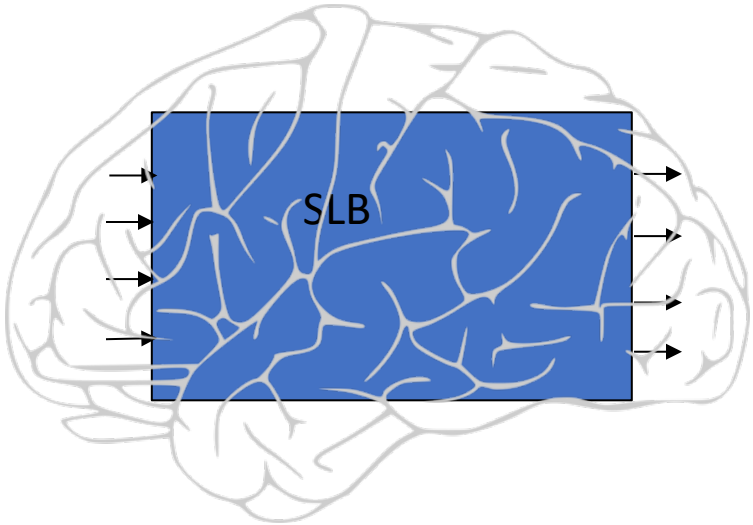
Computational feasibility??

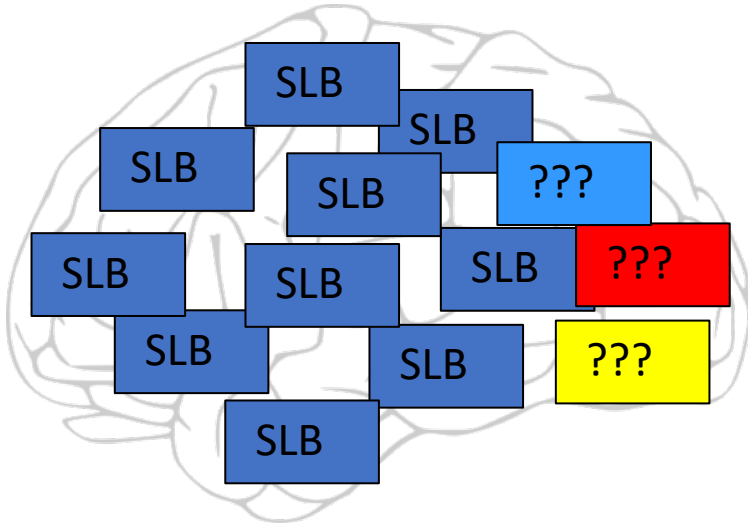
Single Learning Box v. Robust Logic





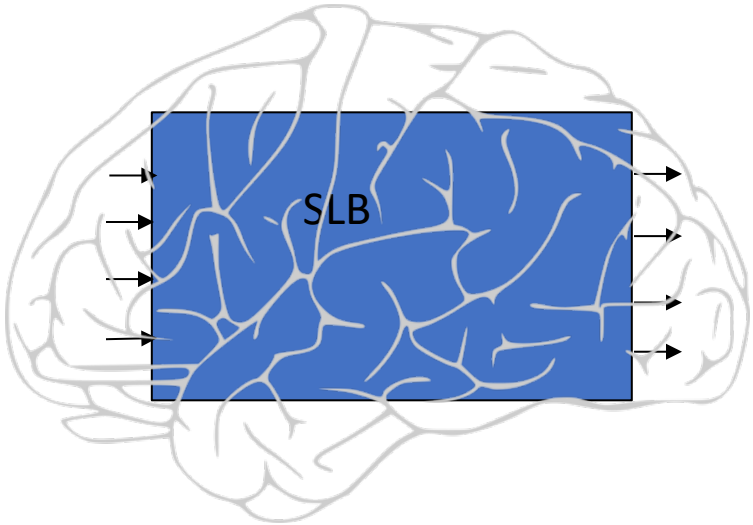
Single Learning Box v. Robust Logic



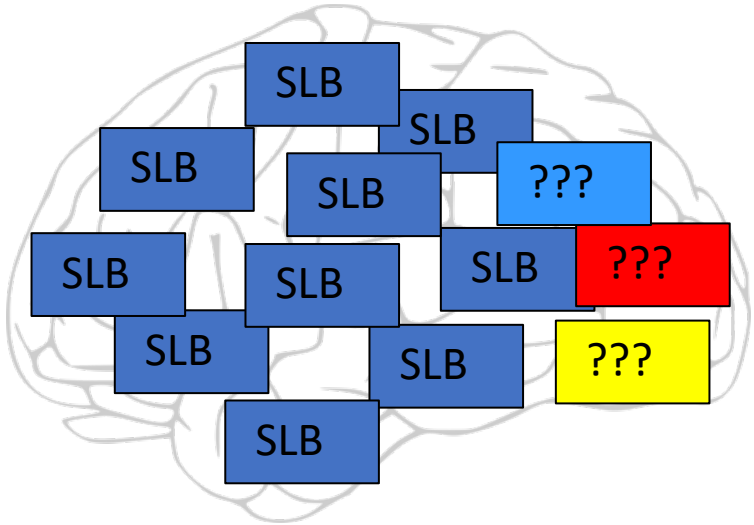


Standard PAC Learning guarantees for generalization.

Single Learning Box v. Robust Logic

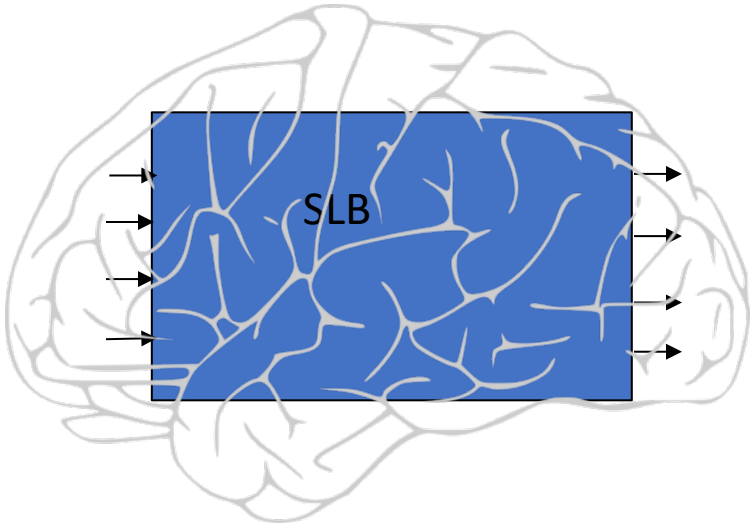


Standard PAC Learning guarantees for generalization.

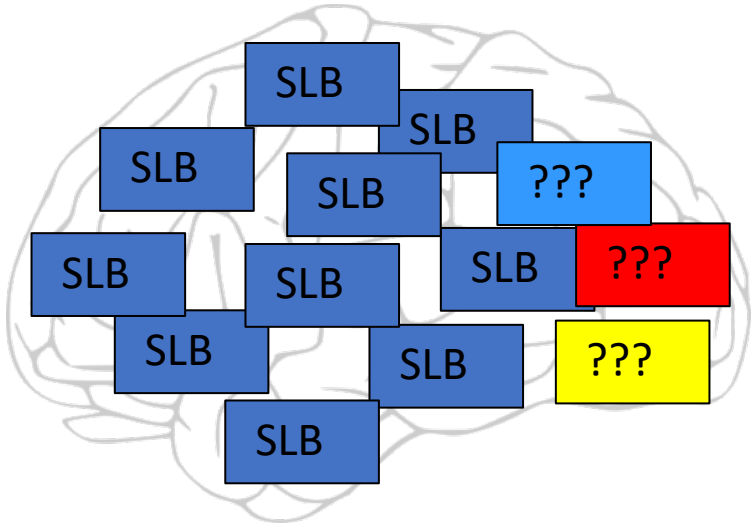


+ Soundness guarantees for chaining:

Single Learning Box v. Robust Logic

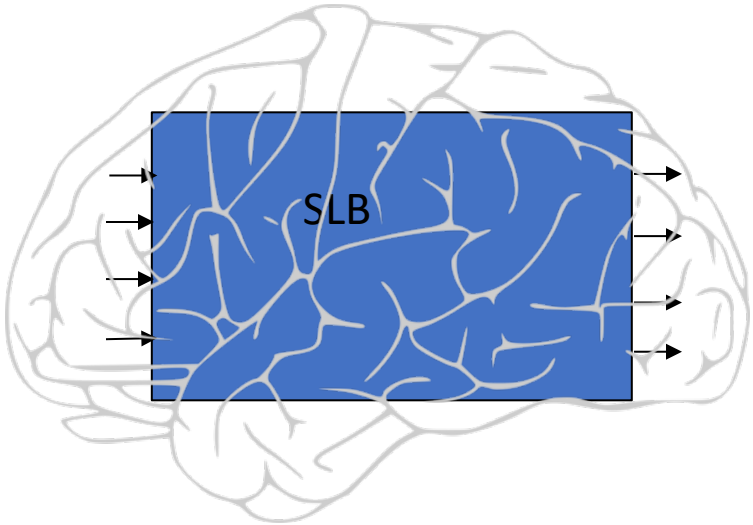


Standard PAC Learning guarantees for generalization.

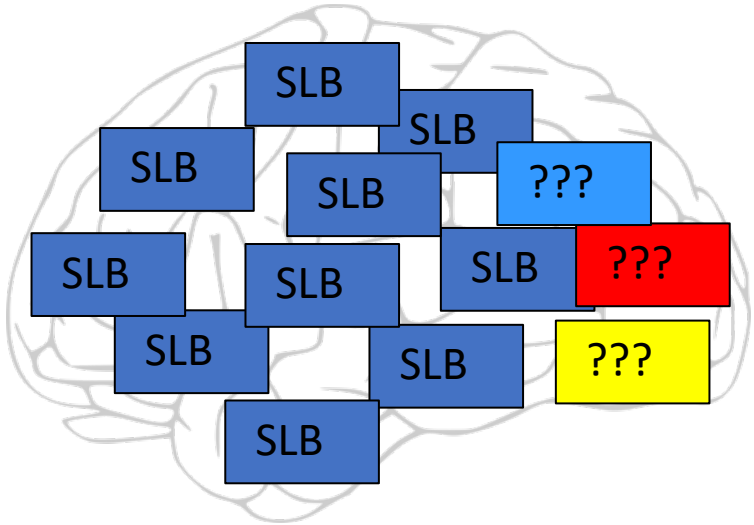


+ Soundness guarantees for chaining: e.g. chaining two 99%-rules gives 98% accuracy.

Single Learning Box v. Robust Logic



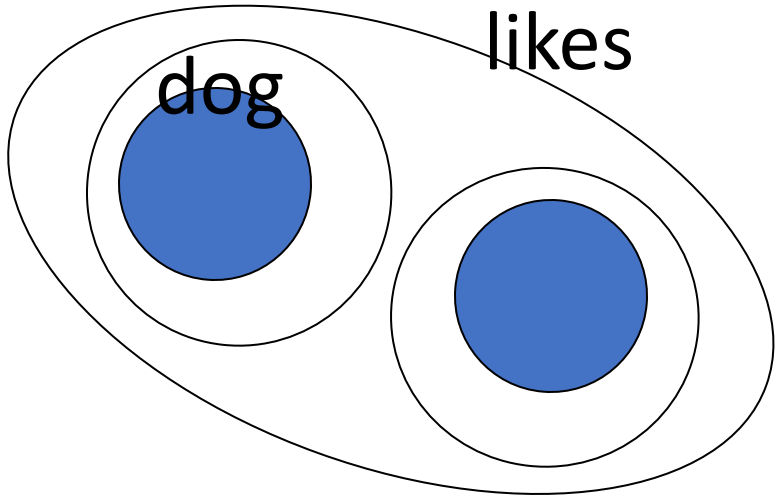
Standard PAC Learning guarantees for generalization.



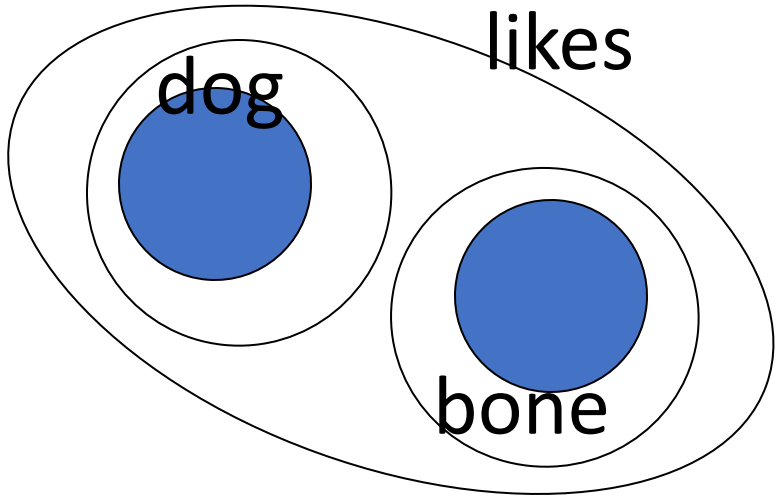
+ Soundness guarantees for chaining: e.g. chaining two 99%-rules gives 98% accuracy.

Did Aristotle own a coffee making machine?

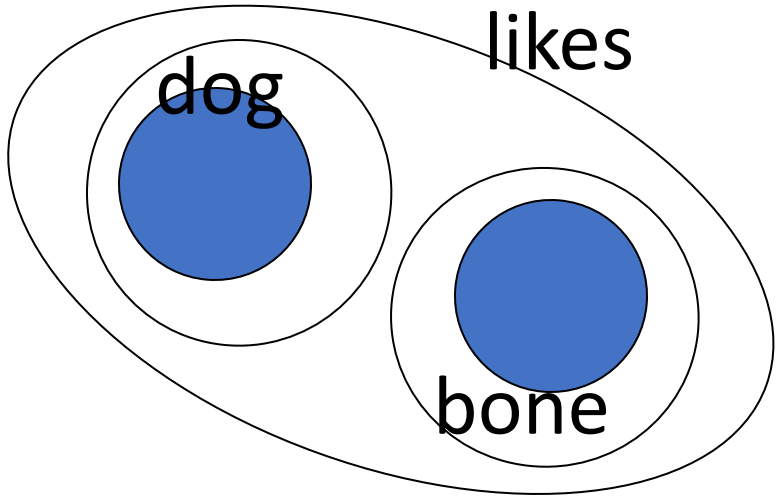
Robust Logic Feature: Mind's Eye



Robust Logic Feature: Mind's Eye



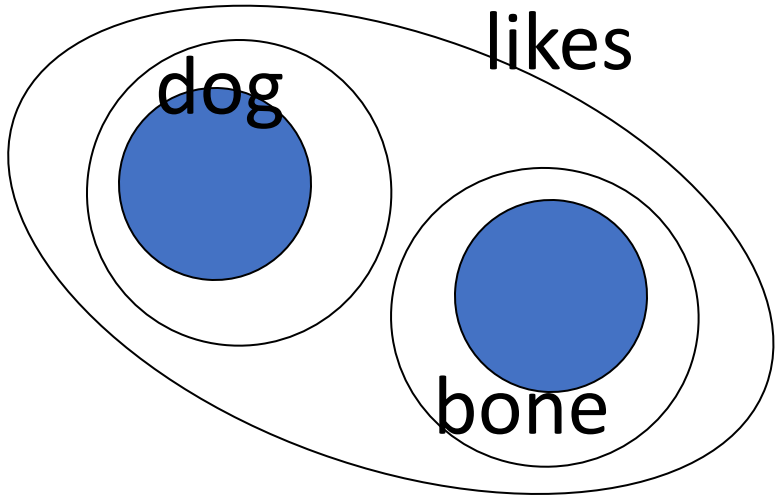
Robust Logic Feature: Mind's Eye



A Robust Logic Rule is a Classifier, e.g.,

"Expression from learnable class" ≡ Bone(x_2)

Robust Logic Feature: Mind's Eye

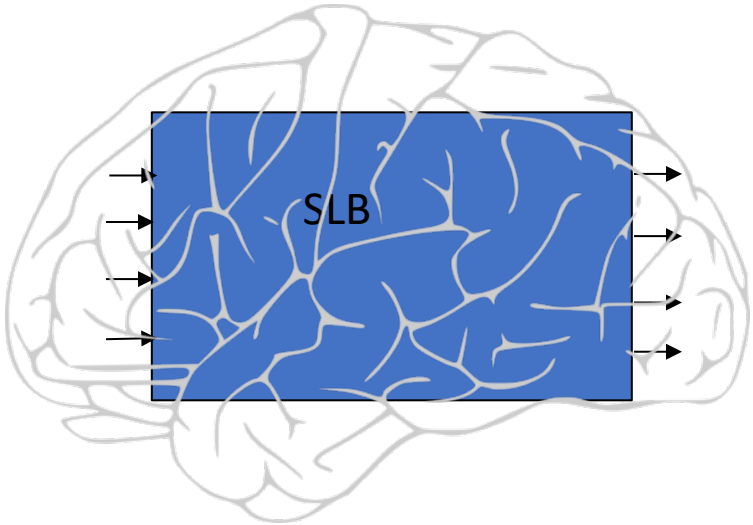


A Robust Logic Rule is a Classifier, e.g.,

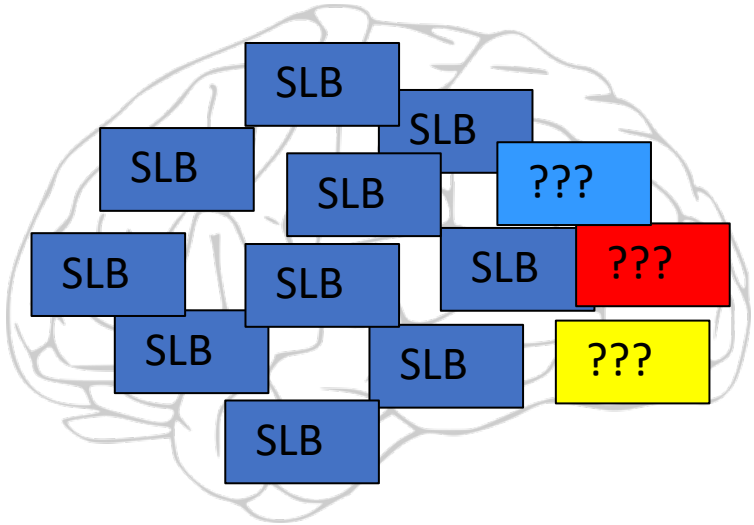
"Expression from learnable class" ≡ Bone(x_2)

For a situation in Mind's Eye chain the applicable learned rules.

Single Learning Box v. Robust Logic



Standard PAC Learning guarantees for generalization.



+ Soundness guarantees for chaining: e.g. chaining two 99%-rules gives 98% accuracy.

Did Aristotle own a coffee making machine?

Back to Civilization Enabler: Hypothesis is that the complex architecture needed for supporting learning, chaining, and Mind's Eye had evolved well before humans.

(III) Taking Instruction from Others

(III) Taking Instruction from Others

Also allow this architecture to execute programs (recipes, methods, instructions) ***received explicitly*** (like a universal Turing machine.)

(III) Taking Instruction from Others

Also allow this architecture to execute programs (recipes, methods, instructions) ***received explicitly*** (like a universal Turing machine.)

Educability is the realization and integration of the three capabilities (i), (ii) and (iii).

Educability

Upside:

Educability

Upside:

- Allows discoveries made with painful and costly experience to be shared.

Educability

Upside:

- Allows discoveries made with painful and costly experience to be shared.
- These discoveries can be new ways of reasoning.

Educability

Upside:

- Allows discoveries made with painful and costly experience to be shared.
- These discoveries can be new ways of reasoning.
- Accumulation of science with contributions from thousands of individuals.

Educability

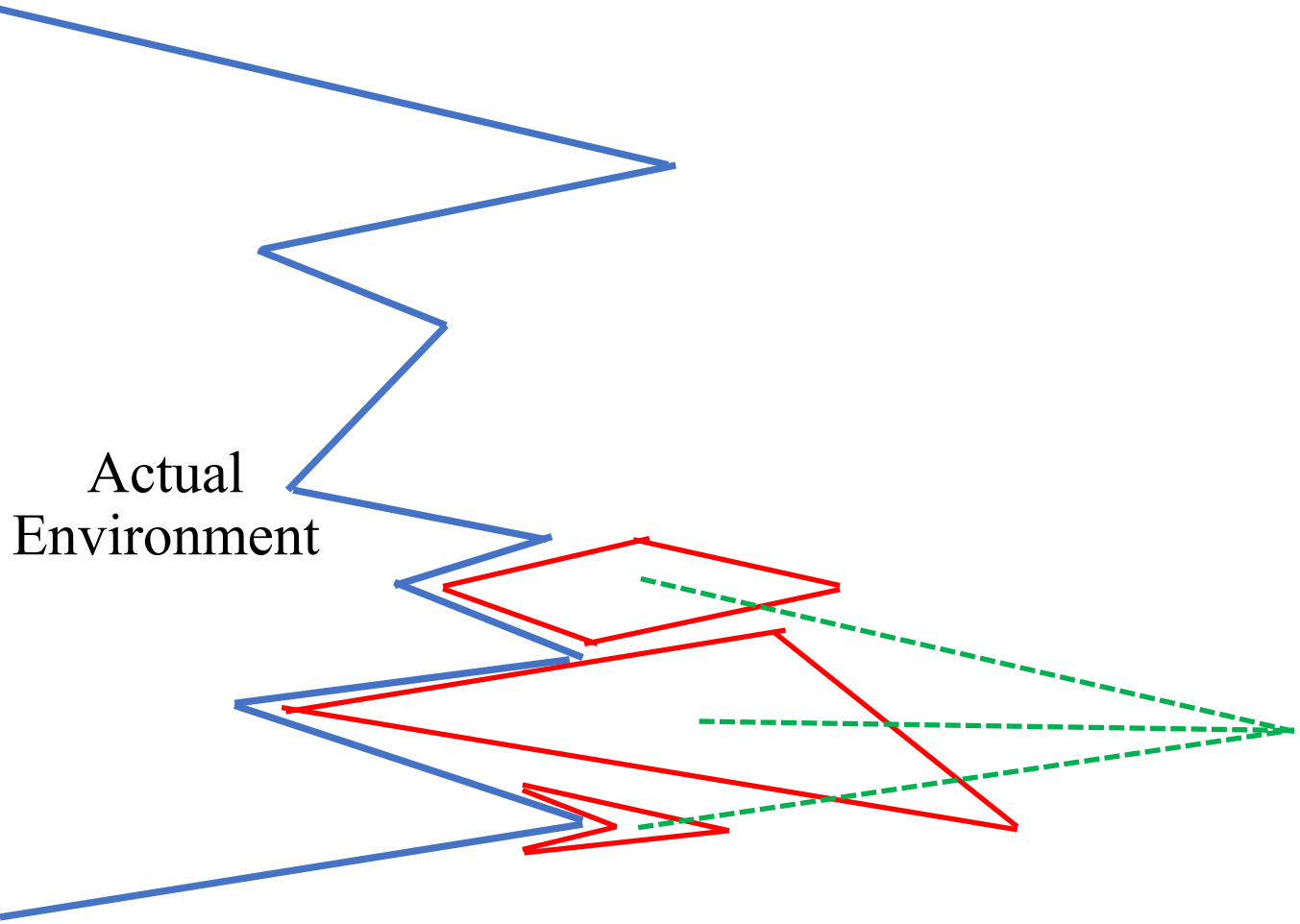
Upside:

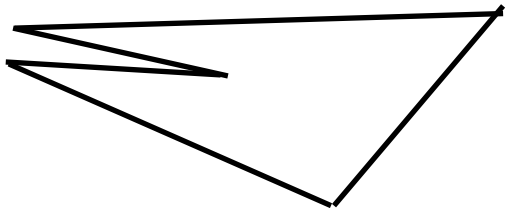
- Allows discoveries made with painful and costly experience to be shared.
- These discoveries can be new ways of reasoning.
- Accumulation of science with contributions from thousands of individuals.
- Allows humans to go to the moon, build computers, cure disease.

Educability

Upside:

- Allows discoveries made with painful and costly experience to be shared.
- These discoveries can be new ways of reasoning.
- Accumulation of science with contributions from thousands of individuals.
- Allows humans to go to the moon, build computers, cure disease.
- Complex cultures





PAC
Learning

Robust Logic

Program
Acquired by
Instruction

Seven Differences all this may Make

1. A Downside of being Educable

1. A Downside of being Educable

- Educability includes capabilities for absorbing instruction **without** accompanying capability for verifying its truth/validity/reality.

1. A Downside of being Educable

- Educability includes capabilities for absorbing instruction **without** accompanying capability for verifying its truth/validity/reality.
- Hence educable beings are easy victims to arbitrary belief systems, ideologies, conspiracy theories.

1. A Downside of being Educable

- Educability includes capabilities for absorbing instruction **without** accompanying capability for verifying its truth/validity/reality.
- Hence educable beings are easy victims to arbitrary belief systems, ideologies, conspiracy theories. (The fault is also in ourselves!)

2. How do I Improve my Educability?

2. How do I Improve my Educability?

Subquestion: What would a test for educability look like?

2. How do I Improve my Educability?

Subquestion: What would a test for educability look like?

“It would test only for knowledge acquired after the beginning of the test”

2. How do I Improve my Educability?

Subquestion: What would a test for educability look like?

“It would test only for knowledge acquired after the beginning of the test”

If apple is to seed, then ….

2. How do I Improve my Educability?

Subquestion: What would a test for educability look like?

“It would test only for knowledge acquired after the beginning of the test”

If apple is to seed, then …. **NO**

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

- **Management strategies for the Mind's Eye:** e.g. How readily to refresh the information.

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

- **Management strategies for the Mind's Eye:** e.g. How readily to refresh the information.
- **Belief Choice:** What source to trust for examples, explicit beliefs, etc?

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

- **Management strategies for the Mind's Eye:** e.g. How readily to refresh the information.
- **Belief Choice:** What source to trust for examples, explicit beliefs, etc?
- **New Concept Formation:** e.g. conjunctions of what?

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

- **Management strategies for the Mind's Eye:** e.g. How readily to refresh the information.
- **Belief Choice:** What source to trust for examples, explicit beliefs, etc?
- **New Concept Formation:** e.g. conjunctions of what?
- **Teaching to Reason:** Which reasoning methods? E.g. arithmetic

3. Educability Has Many Parameters

(Even conventional ML systems have parameters: choice of data and choice of learning algorithm.)

- **Management strategies for the Mind's Eye:** e.g. How readily to refresh the information.
- **Belief Choice:** What source to trust for examples, explicit beliefs, etc?
- **New Concept Formation:** e.g. conjunctions of what?
- **Teaching to Reason:** Which reasoning methods? E.g. arithmetic

Parameters are not a bug of the educability model but a feature of cognition.

4. No Singularity?

4. No Singularity?

I. J. Good (1965): *Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. …*

4. No Singularity?

I. J. Good (1965): *Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. …*

What if the capability can be expressed in terms of explicit prosaic understandable components?

4. No Singularity?

I. J. Good (1965): *Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. …*

What if the capability can be expressed in terms of explicit prosaic understandable components? What if educability is all there is,

4. No Singularity?

I. J. Good (1965): *Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. …*

What if the capability can be expressed in terms of explicit prosaic understandable components? What if educability is all there is, **with all the parameters**?

5. Pitfalls of Studying Human Behavior

5. Pitfalls of Studying Human Behavior

Why is it so hard to draw conclusions about human behavior that are transferrable in time and space?

5. Pitfalls of Studying Human Behavior

Why is it so hard to draw conclusions about human behavior that are transferrable in time and space?

How can social science theories come in different political flavors?

5. Pitfalls of Studying Human Behavior

Why is it so hard to draw conclusions about human behavior that are transferrable in time and space?

How can social science theories come in different political flavors?

Where did the eugenicists go wrong? (Galton, Pearson, Fisher)

5. Pitfalls of Studying Human Behavior

Why is it so hard to draw conclusions about human behavior that are transferrable in time and space?

How can social science theories come in different political flavors?

Where did the eugenicists go wrong? (Galton, Pearson, Fisher)

Answer to all: EDUCABILITY

6. Education

It would be nice to contribute some to
“a more scientific basis for education.”

7. Like PAC Learning, a Technological Proposal

THANK YOU

Informal Definition of Educability

Educability consists of capabilities to

(i) *acquire* beliefs by *generalizing* from experience,

(ii) *apply* beliefs in *combination* to specific situations,

(iii) *acquire* beliefs *explicitly described* to us by others,

all *integrated* with *(iv) multi-object scene analysis*

and *(v) symbolic naming*.