Is Artificial Intelligence really intelligent?

Akira Imada

Brest State Technical University (Belarus)

Is machine intelligence a fantasy?

Fantasy to create Humanoid I

Pigmarion (In ancient Rome or in 'My Fair Lady'),

HAL (in '2001: A Space Odyssey' by Stanley Kubrick), Coppelius, etc ..,

Fantasy to create Humanoid II

Karakuri Ningyo 200 yeas ago in Japan

What we need is Definition of Intelligence

for a departure from fantasy

W. V. Bingham

"ability to solve new problems"

R. Pinter

"ability to adapt oneself adequately to relatively new situations"

D. Wechsler

"ability (i) to act purposefully, (ii) to think rationally, and (iii) to deal effectively with the environment"

U. Neisser et al.

"ability (i) to understand complex ideas, (ii) to adapt effectively to the environment, (iii) to learn from experience, (vi) to engage in various forms of reasoning to overcome obstacles."

The term 'Intelligence' is everywhere!

Lots of papers claim an intelligence

E.g. Intelligent decision in stock market

Sewell

"The Application of intelligent systems to financial time series analysis."

Tsang

"Forecasting - where computational intelligence meets the stock market."

Kuoa

"An intelligent stock trading decision support system."

Other examples we want to know its intelligence

From
Deep Blue: beat Kasparov
to
Watson: won Jeopardy
Dr. Fill: joined the human crossword tournament
Siri: incorpolated in i-Phone 4

Essay-scoring software

How intelligent?

Which agent is more intelligent?

Or, not intelligent at all?
We'd better ask monkey to throw darts.

Burton G. Malkiel "A Random Walk Down Wall Street"

"A monkey throwing darts at the WSJ to select a portfolio might be better than the one carefully selected by experts."

⇓

Are those machines less intelligent than a monkey?

(The key question of this talk.)

Turing Test

tests the ability to fool ordinary people, not foolish people, about whether

the dialogue via teletype is with a human or with a computer.

Opposit Turing Test – CAPTCHA

to prove

you're a human to a computer.

(Completely Automated Public Turing-Test to Tell Computers and Humans Apart)

Turing-like Test

Whether computer can fool human that 'I am a human not a computer!'

Is Turing Test after 50 years still valid?

We can easily cheat the test with a trick!

Turing's Chess Machine

The game was with three chess players A, B and C.
A operates the paper machine.
C plays chess from outside with either A or B both in the separate room
C should guess whether he is playing, with human or the paper machine.

Turk

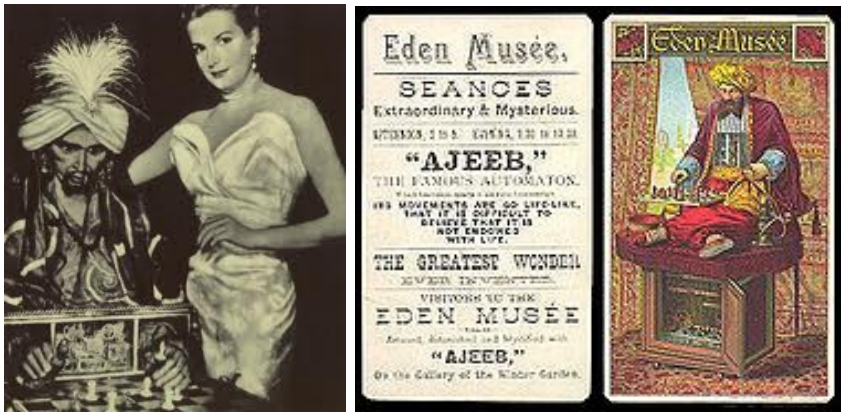
A legendary chess automaton in 18th century.
Wolfgang von Kempelen created it and claimed it plays chess like human to impress Maria Theresa (1770).

Napoleon vs. the Turk at Schönbrunn

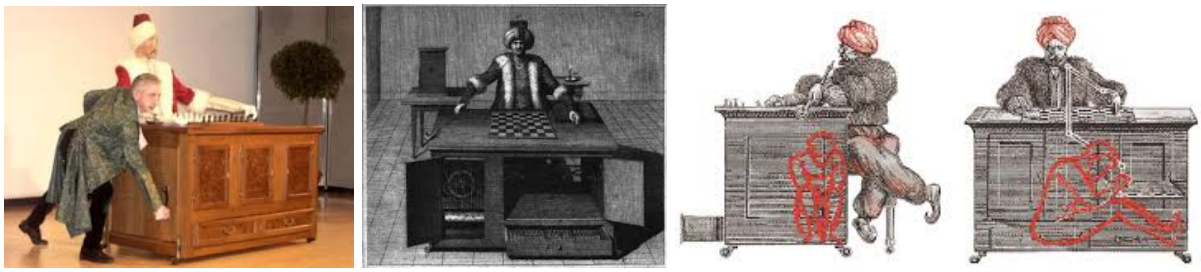
It debuted in Schönbrunn Palace then toured in Europe after 1783.

Märzel bought the Turk (1811)

In 1826, he opened an exhibition in New York.



It was human not the machine who played.



The secret had been perfectly kept for more than 100 years.

It was not until Dr. Silas Mitchell fully revealed the secret in the book "The Last of a Veteran Chess Player" (1857).

$\Rightarrow$ No one had claimed it was human!

IBM's Deep Blue

Garry Kasparov vs. Deep Blue in 6 game match (1996)

Deep Blue beat Kasparov

1996:	Deep Blue	Kasparov	even	even	Kasparov	Kasparov
1997:	Kasparov	Deep-Blue	even	even	even	Deep Blue

Most likely it would pass the Turing Test for chess.

Is Deep Blue Intelligent?

Everyone knows "it's not intelligence but just with a brute force."

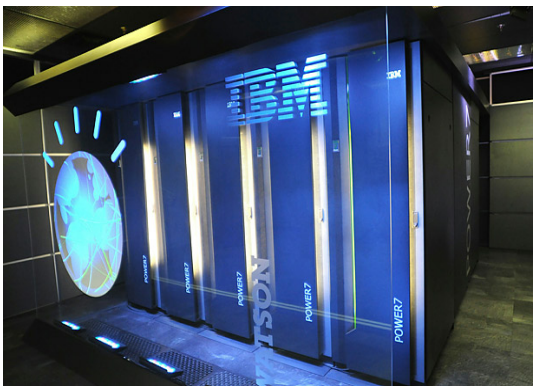
⇓

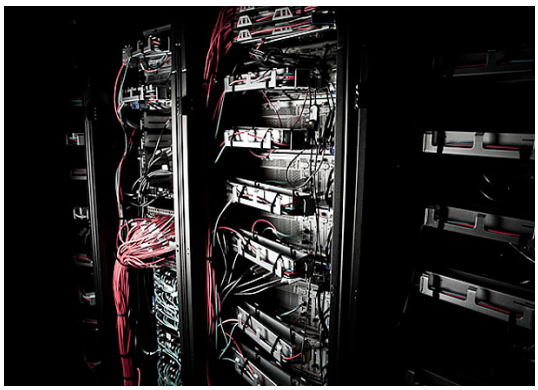
What if we change the rules of chess a little?
Kasparov would remain super,

while Deep Blue would confuse and not be so any more.

Watson

I.B.M's ground-breaking question-answering-machine





Watson challenged 'Jeopardy!' (2011)

America's Favorite Quiz Show on TV

⇓ e.g.

"Who is the 19th-century painter whose name means police officer?"

"What is the city whose largest airport is named for a World War II hero; and its second largest for a World War II battle."

Watson vs. two human ex-champions

⋆ The holder of the longest unbroken winning streak.

⋆ An undefeated champion and the game's biggest money winner.

The final tally

The two humans won $24,000 and $21,600 while Watson $77,147.

But are we happy?

A huge look-up table based machine could, in theory, pass the Turing test

$$\Downarrow$$

but without any real intelligence.

Can we tell it's performed by computer?

False positive or not? (2)

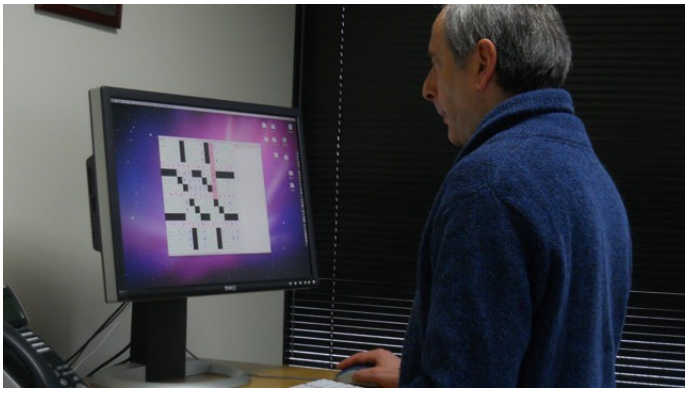
$\Downarrow$

If we can, it's because of

always exactly the same action in a same situation.

Crosswords solving algorithm

Dr. Fill designed by Matthew Ginsberg



American Crossword Puzzle Tournament

in Brooklyn (2012)

600 humans of the world's best crossword players participated.
The trophy went to a human and Dr. Fill finished 141st.

Still impressive, isn't it?

A degree to how a machine is intelligent

Gregory Chaitin (1987), Warren Smith (2006)
tried to answer it using complexity theory.

⇓

Now we have a fair amount of such definitions.

Legg & Hutter's definition

From informal

"An agent's ability to achieve goals in a wide range of environments."

$$\Downarrow$$

to formal

$$\gamma(\pi) = \sum_{\mu \in E} 2^{-K(\mu)} \cdot V_\mu^\pi$$

Agent interacts with environment

Observation $o_i \Rightarrow$ Action $a_i \Rightarrow$ Reward r_i

E.g. Agent to Invest in Stock Market

(1) observes financial environment of stock market

$\Downarrow$

(2) makes an action in the environment, i.e., to sell or to buy

$\Downarrow$

(3) gets rewards, i.e., profit or loss.

Definition of agent

Function that takes the current history as input
and
produces an action as output

$$\pi(a_k|o_1 r_1 a_1 o_2 r_2 a_2 \cdots o_{k-1} r_{k-1})$$

or probability function for indeterministic.

Definition of environment

Function which produces output $o_k r_k$ given the current history

$$\mu(o_k r_k | o_1 r_1 a_1 o_2 r_2 a_2 \cdots o_{k-1} r_{k-1} a_{k-1})$$

or probability function for indeterministic.

Expected value of sum of rewards

$$V_{\mu}^{\pi} = E(\sum_{i=1}^{\infty} r_i)$$

Definition of intelligence

Weighted sum of
expected value of sum of rewards over infinitely various environments

$$\gamma(\pi) = \sum_{\mu \in E} w_\mu \cdot V_\mu^\pi$$

How will those weights be specified?

Translate the environment into a binary string x by Turing Machine U

$\Downarrow$

Calculate Kolmogorov complexity K of x

(length of the shortest program that computes x)

$K(x) = \min_p\{l(p)|U(p) = x\}$

$\Downarrow$

$w_\mu = 2^{-K(\mu)}$

The smaller the complexity the larger the weight.

Universal Machine Intelligence

by Legg and Hutter

$$\gamma(\pi) = \sum_{\mu \in E} 2^{-K(\mu)} \cdot V_\mu^\pi$$

"An agent's ability to achieve goals in a wide range of environments."

Too conceptual or too theoretical

Goertzel:

"Universal but not practical."

Goertzel

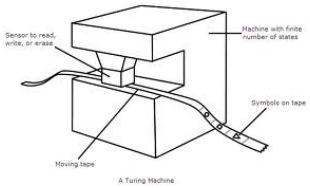
pragmatic intelligence

$$\Pi(\pi) = \sum_{\mu \in E, g \in G, T} \nu(\mu)\gamma(g,\mu)\tau_{g,\mu}(T)V^{\pi}_{\mu,g,T}$$

$$\Downarrow$$

It's not very practical yet, isn't it?

Problem is, translation of environment by Turing machine



⇓

Are there easier alternatives to the Turing machine?

Hernándes-Orallo

The other representations of environment

$\Downarrow$

(i) λ-calculus, (ii) combinatory logic, (iii) abstract state machines,
(iv) register machines, (v) Markov algorithms,
(vi) term-rewriting systems, ...

to generate environments and calculate complexity automatically.

"still Turing-complete, but
more natural and easy to work with than the Turing machine."

Hernándes-Orallo's emphasis

in order for the environments to be balanced

$\Downarrow$

a random agent's expected accumulated reward should be zero

assuming rewards range from -1 to 1

Hernandes's target was strings

He measured the ability of finding the shortest explanation for some strings of different difficulty in a fixed time

e.g.

$$x = aaaaaaabaaaaaaabaaaaaaabaaaaaaa$$

$$x = aabbbccdddeefffgghhhiiijjj$$

$$x = 1,2,3,5,7,11,13,17,19,23$$

$$\Downarrow$$

Still not so practical for our purpose.

Hibbird used finite state machine (FSM) to be a little more realistic

Environment and agent are both modeled as FSMs.

⇓

Create a hierarchy of increasingly difficult environments.

⇓

Agent intelligence is measured as the highest level of environment against which it can win the game.

Hibbird's game

Evader e produces the sequence by

$$x_{n+1} = e(y_1 y_2 y_3 \cdots y_n),$$

and predictor p produces the sequence by

$$y_{n+1} = p(x_1 x_2 x_3 \cdots x_n).$$

Then p wins round $n+1$ if $y_{n+1} = x_{n+1}$

and e wins if $y_{n+1} \neq x_{n+1}$.

Finite State Machine instead of Turing Machine

No need to be universal (at least currently)

Stock market forecasting as a benchmark

transactions

$$\Downarrow$$

observation $\rightarrow$ action $\rightarrow$ reward

Representation of stock market by FSM

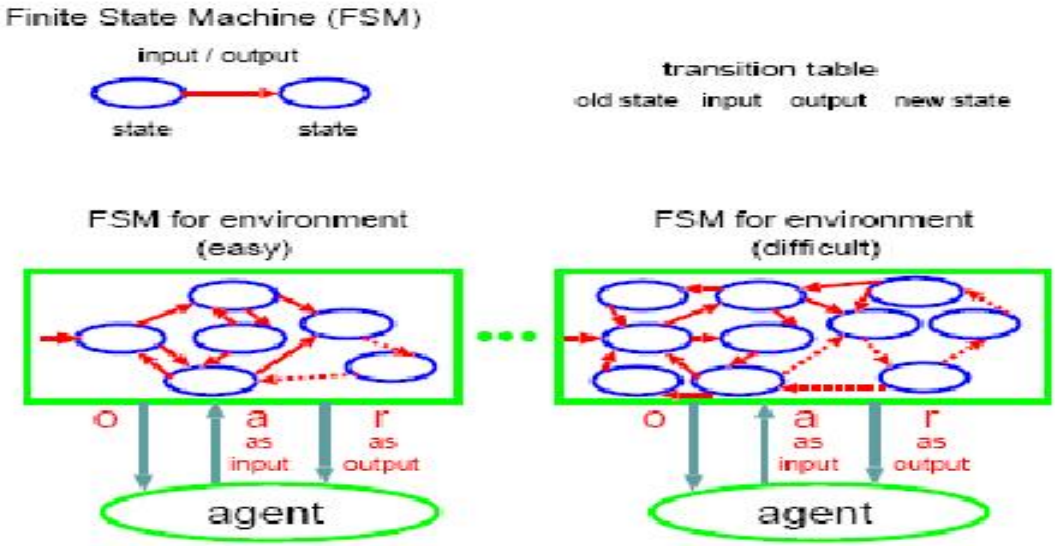
e.g.

⇓

"Agent Based Economic Modeling with Finite State Machines: Modeling the market

by Gnilomedov & Nikolenko

Environment by finite state machine



Our goal is to measure intelligence of a proposed machine

mostly not the one controlled by FSM.

What we now are thinking of (1)

Let's be specific not universal!

⋆ "He is an intelligent football player,"
while we know he is not good at Mathematics,
which we don't care.

⋆ This conductor always makes an intelligent interpretation of symphony,
but very bad at football.

⋆ Did Einstein play football intelligently?

⇓

Intelligence doesn't need to be universal!

Intelligence in a specific environment

$$\gamma(\pi) = \sum_{\mu \in E} 2^{-K(\mu)} \cdot V_\mu^\pi \qquad \Rightarrow \qquad \gamma(\pi) = 2^{-K(\mu)} \cdot V_\mu^\pi$$

Note: Deep Blue is very intelligent in this formula.

What we now are thinking of (2)

Let's be more unpredictable!

Human-intelligence is spontaneous, flexible, and/or unpredictable, more or less.

Or even erroneous sometimes.

NN with fixed weights

$\Downarrow$

never be intelligent despite their claim,

but only repeats exactly the same action in the same situation.

A possible trick to the Turing Test

might be

to give a same question repeatedly.

Can we tell it's a computer, not a human?

If we can, it's because of too precise reaction,

Is computer intelligent enough to pretend to be a little foolish?

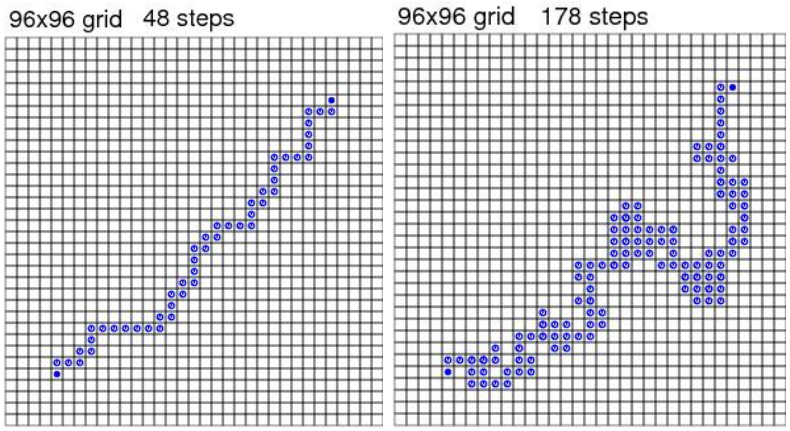
To mimic a human

$\Downarrow$

Do not always exactly the same action even in a same situation.

E.g. which route is more intelligent?

A business person & a philosopher going for a walk in Manhattan



Let's try for a walk in Manhattan?

Even if we want a shortest route, we want a different one than before.

Is it easy to make an agent take a different path from run to run?

⇓

Yes, by using random numbers

⇓

But it's not an intelligence.

What about from (0,0) to (m,0)?

⇓

only one unique solution.

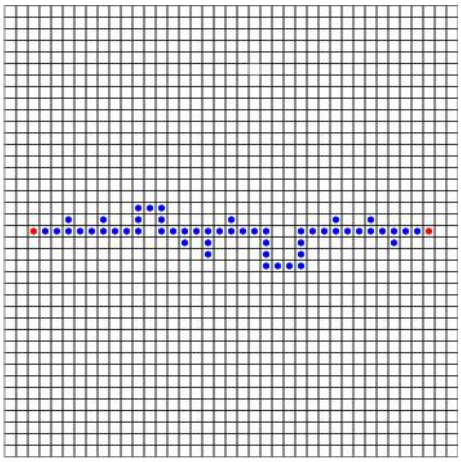
⇓

Can an agent still be flexible from run to run?

⇓

This could be a benchmark.

How about this?



Again by random numbers if we don't care efficiency

It's as Papert wrote (1965)

"A very simple algorithm can sometimes obtain the same results as the holistic, intuitive human mind,"

But it's not intelligent either

Intelligence tends to avoid a similar behavior

"I beg your pardon?"

Intelligent people try a different explanation for an easier understanding

while

others just repeat the same expression, maybe louder.

Do you agree that always 'being same' is not intelligent?

Excellent but always slightly different

(One of the reason I like this.)


Does intelligence imply preciseness, efficiency, or effectiveness?

If so

We have already lots of intelligent machines

Why spontaneous?

We sometimes need spontaneous and unpredictable intelligence rather than efficiency or effectiveness like in case of SONY's AIBO.

It learns excellently and acts differently in different situation but repeats the same action in a same situation.

A measure of similarity

Choi et al.'s similarity of two binary strings $\mathbf{x}$ and $\mathbf{y}$

calculate, first $a = \mathbf{x} \cdot \mathbf{y}$, $b = \bar{\mathbf{x}} \cdot \mathbf{y}$, $c = \mathbf{x} \cdot \bar{\mathbf{y}}$ and $d = \bar{\mathbf{x}} \cdot \bar{\mathbf{y}}$
then $sim(\mathbf{x}, \mathbf{y}) = f(a, b, c, d)$
e.g.

$$\frac{(b+c)^2}{(a+b+c+d)^2}$$

Or

Lin et al.'s information theoretic definition of similarity

$$sim(X,Y) = \frac{\log p(common(X,Y))}{\log p(description(X,Y))}$$

What we now are thinking of (3)

"Occam's Razor Principle"

Complexity of series of action (not of environment).

$\Downarrow$

The simpler the actions, the higher the intelligence.

Simple measure of complexity

Crutchfield et al.: "Comment on simple measure for complexity"

$$\Gamma_{\alpha\beta} = (S/S_{max})^{\alpha}(1-(S/S_{max}))^{\beta}$$

Fioretti: "A subjective measure of complexity"

$$C_O(S) = \sum_{q=0}^{Q} \frac{q+1}{s_q}$$

Lloyd: "A survey - Measures of complexity a non-exhaustive list"

What we now are thinking of (4)

"Intelligence learns from previous runs"

Our proposal

$$\gamma(\pi) = \sum_{\mu \in E} 2^{-K(\mu)} \cdot V_\mu^\pi \qquad \Rightarrow \qquad \gamma(\pi) = V_\pi$$

$\Rightarrow$ repeat a run in a same condition to see similarity, simplicity, learning

$$\gamma(\pi) = \sum_i \frac{\{\text{IMPROVEMENT}\}_i \cdot \{\text{EXPECTED REWARD}\}_i}{\{\text{COMPLEXITY}\}_i \cdot \{\text{SIMILARITY}\}_i}$$

Still a thought experiment but ...

Stock market investment as a benchmark

Input: Daily prices of all stocks of previous one year
Run the algorithm we want to measure intelligence multiple time
Measurements are:
(1) total reward
(2) complexity of series of actions in each run
(3) similarity of n series of actions
(4) degree of improvement of result of n run

Summary

We incorporate to our definition of machine intelligence:
(1) total reward
(2) simplicity of actions
(3) more or less different actions than before
(4) learning

Occam's razor principle vs. Intelligence

Once Kluger wrote in the TIME Magazine
"intelligent individuals are more difficult to learn to know."
⇓
So does a Machine Intelligence?

Thank you for your interest!