Finding a Needle in a Haystack: from Baldwin Effect to Quantum Computation

Akira Imada

Brest State Technical University, Belarus

Alora's proposition (2007)

Brilliance = Ability to find a needle in a haystack.

"Art is to create a satisfying appreciation in a huge universe of our brain."

(Cont'd)

(Artificial) *INTELLIGENCE*
can search, classify and recognize normal objects
while
only *BRILLIANCE*
can search for a needle in a hay.

Could an endlessly typing monkey eventually create Shakespeare by chance?

Hopefully not!

Then what about *haiku*?

Might be more likely.

うぐいすの

かさおとしたる

つばきかな

How many possible different candidates?

$$50^{17} \approx 10^{19}$$

It's like a search for *a-needle-in-a-haystack* for a monkey to create a meaningful one.

Computational analogue of a-needle-in-a-haystack

A specific rare cases in a huge database;
A real necessary information from world-wide-web;
Oil spills in the ocean from satellite image;
Cause of failures in diagnosing a huge code;
A collision of a hash function;

......

$\Downarrow$

Such searches are very important for us.

Associative memory by Spiking neurons

Fitness landscape

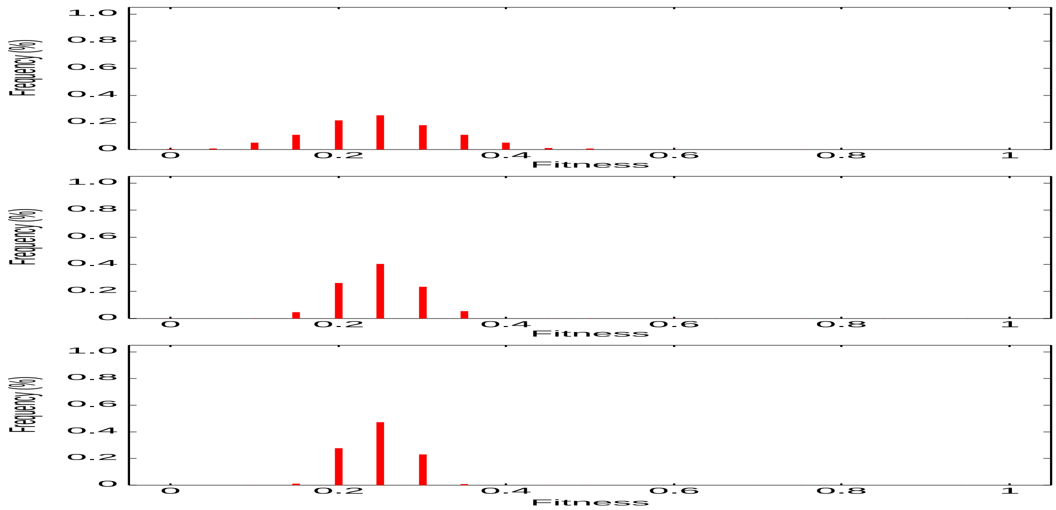
$\Downarrow$

Hopfield network ... many peaks in a rugged land

vs.

Spiking neurons ... like a flat island in a huge lake

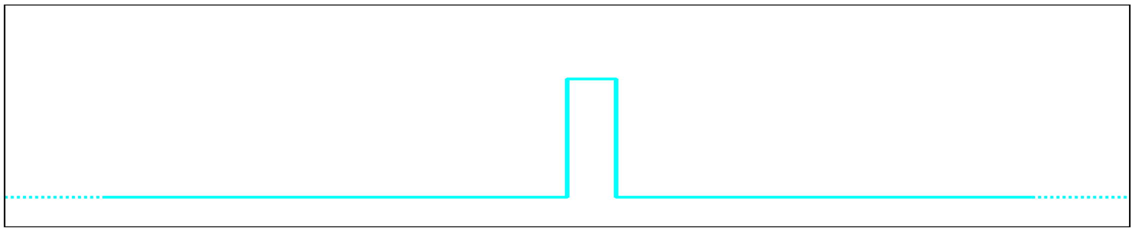
A standard initial distribution of fitness values



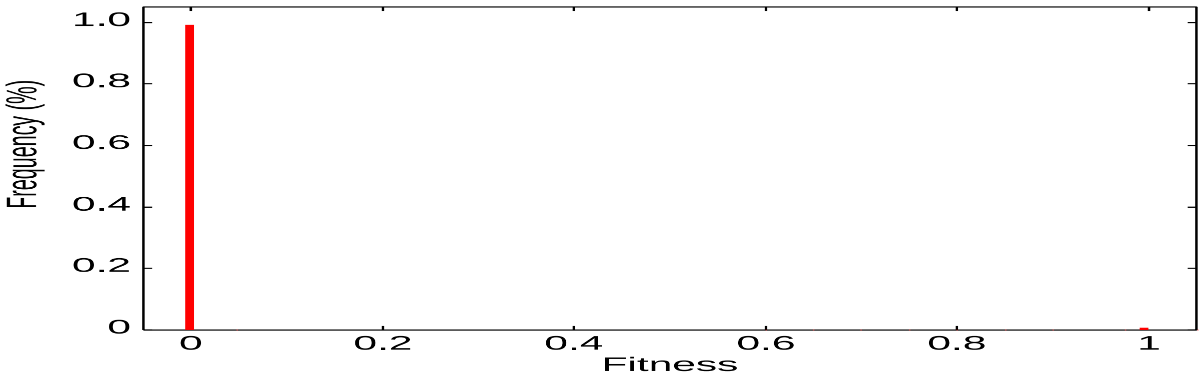
A-tiny-flat-island-in-a-huge-lake

A fictitious 1-D fitness-landscape

of Hebbian-learning by spiking neurons



Then, fitness distribution is ...



It's not possible to hillclimb if not a hill to climb!

Network intrusion data among normal transaction data

E.g. $\Rightarrow$ 4 attack types in KDD-cup-99 dataset

The winner's result

Probe	DoS	U2R	R2L
83.3%	97.1%	13.2%	8.4%

Detection rate by Sabhnani et al. (2003)

	Probe	DoS	U2R	R2L
Multi-layer Perceptron	88.7	97.2	13.2	5.6
Gaussian Classifier	90.2	82.4	22.8	9.6
K-mean Clustering	87.6	97.3	29.8	6.4
Nearest Cluster Algorithm	88.8	97.1	2.2	3.4
Radial Basis Function	93.2	73.0	6.1	5.9
Leader Algorithm	83.8	97.2	6.6	1.0
Hypersphere Algorithm	84.8	97.2	8.3	1.0
Fuzzy Art Map	77.2	97.0	6.1	3.7
C4.5 Decision Tree	80.8	97.0	1.8	4.6

Some Intrusions are immune to be detected.

(Joshi et al. 2001)

"If a class is very rare, e.g., 0.5%, prediction of non-target will be 99.5%."

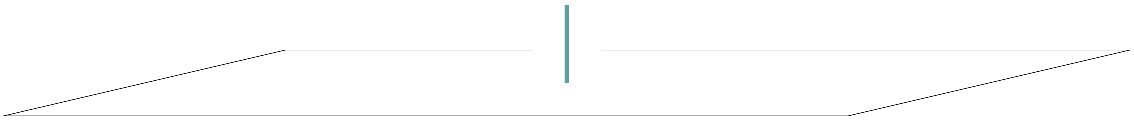
(Nevertheless, we have still many innocent optimistic reports of success.)

We now look at the other 4 cases more in detail!

- Hinton & Nowlan's experiment
- Reduced even-parity problem as a benchmark
- A robot navigation
- Quantum computation

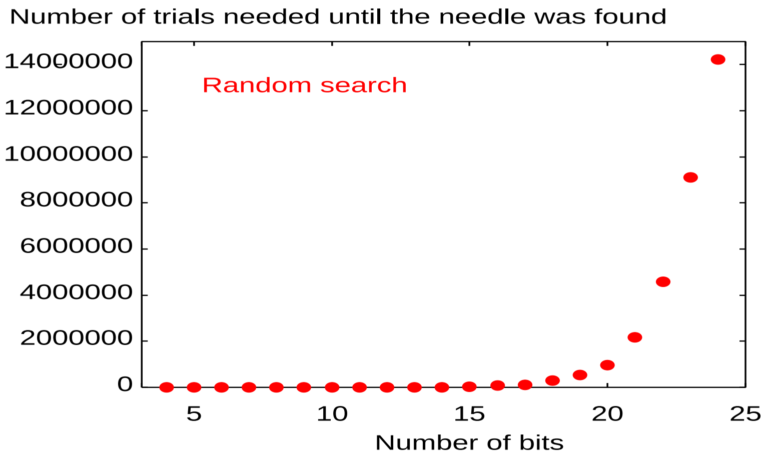
Hinton & Nowlan's Simplest Needle in a Hay.

Hinton & Nowlan's Needle (1987)



- A-needle $\Rightarrow$ A configuration of 20-bit binary strings
- Haystack $\Rightarrow$ $2^{20} - 1$ search points

Their 20-bit was a good choice



Evolution under Baldwin effect – Lifetime learning of phenotype

A genotype:

$$(109010990110010010 91)$$

$$\Downarrow$$

Its phenotype:

$$(101010010110010010 01)$$
$$(100010110110010010 11)$$

$$\ldots$$

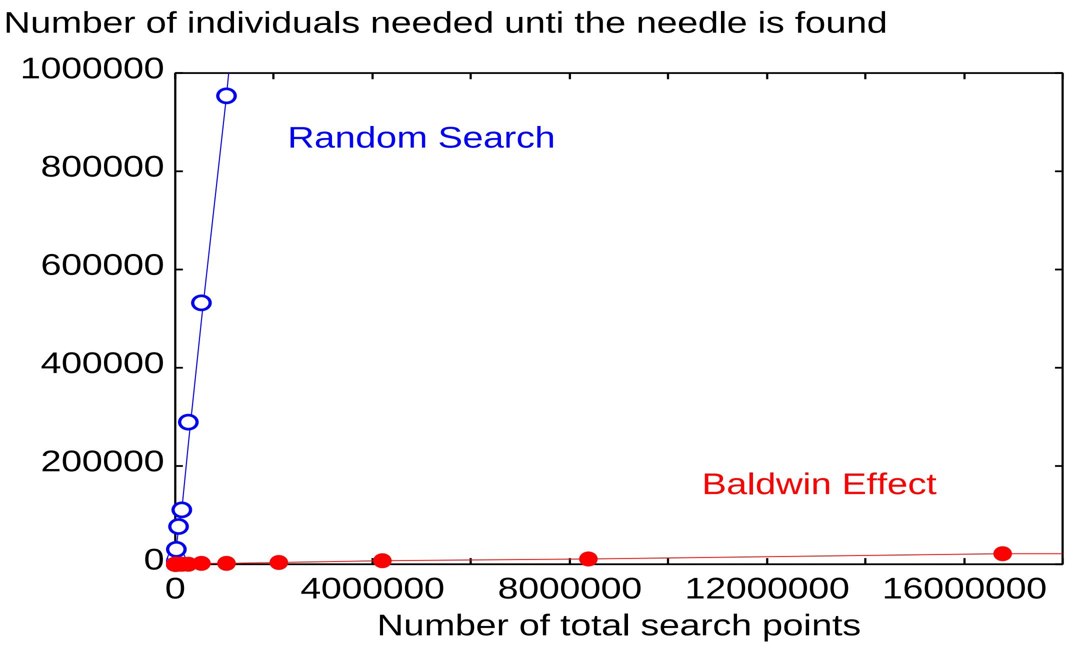
(Cont'd)

(Their assumption)
The closer the genotype to the needle,
the faster the learning of phenotype.
$\Downarrow$
This makes the needle-like-peak smoother.

Their GA implementation

1000 genotypes with 25% flexible genes,
each allowed 1000 lifetime learnings;
Roulett wheel selection;
One-point cross over;
Mutation;
......

Baldwin effect looks great!



(Cont'd)

Note, anyway, $O(N)$ steps are necessary!

Are we happy with this Baldwin effect?

Why should we continue
when lifetime-learning already finds needle?

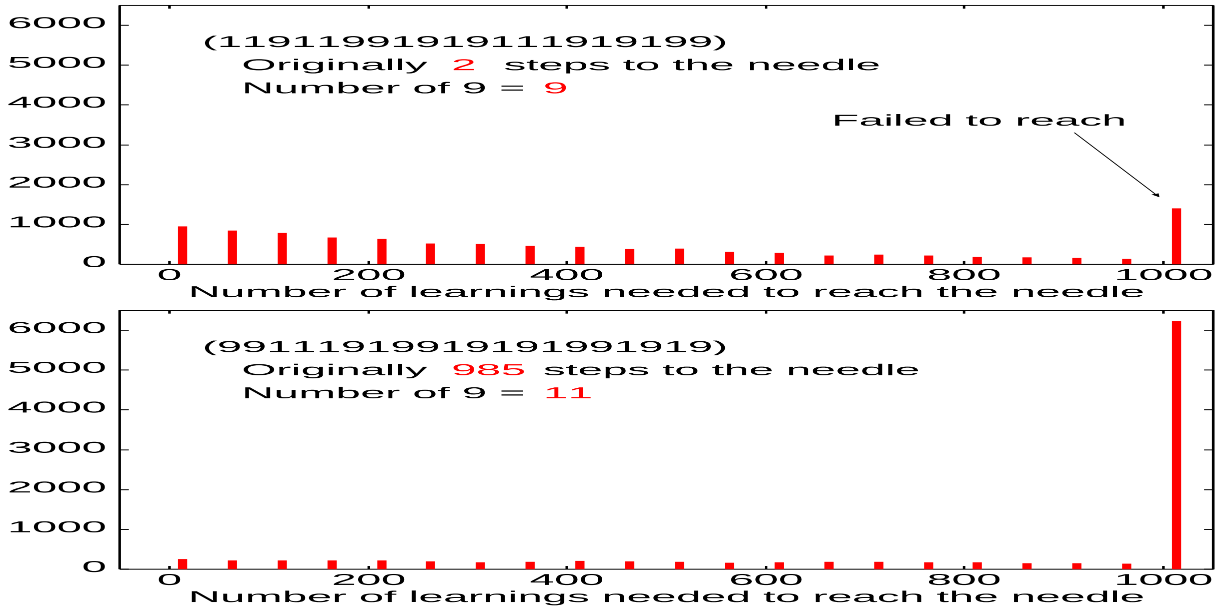
$\Downarrow$

(Turney 1987)

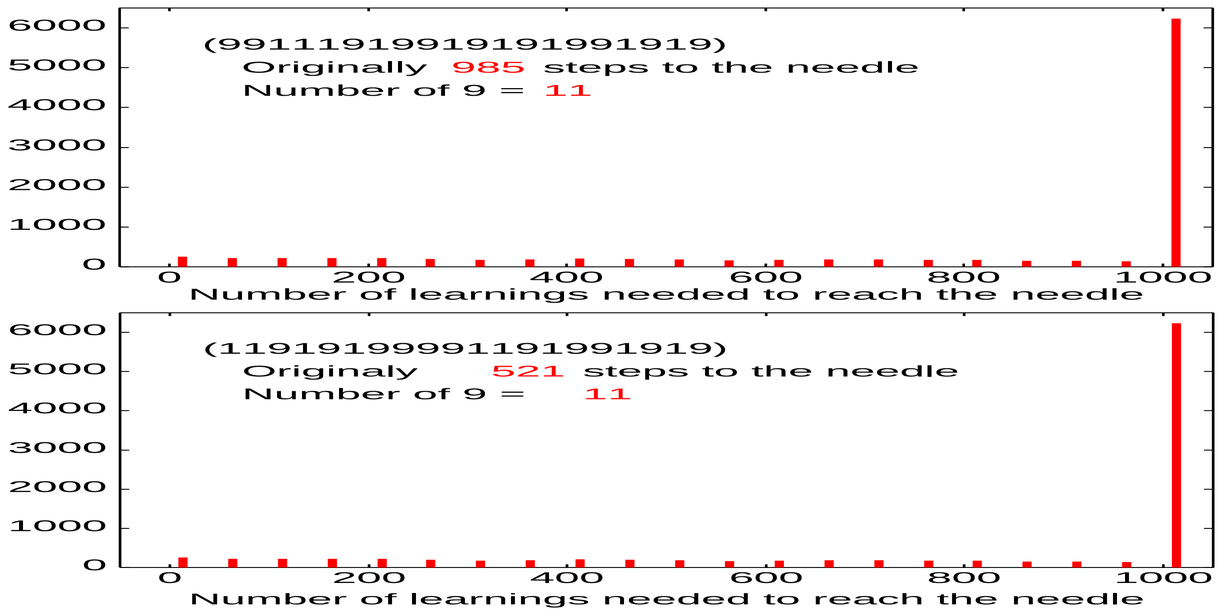
"Not from an engineering but a biological interest."

Is a high fitness gene a good gene?

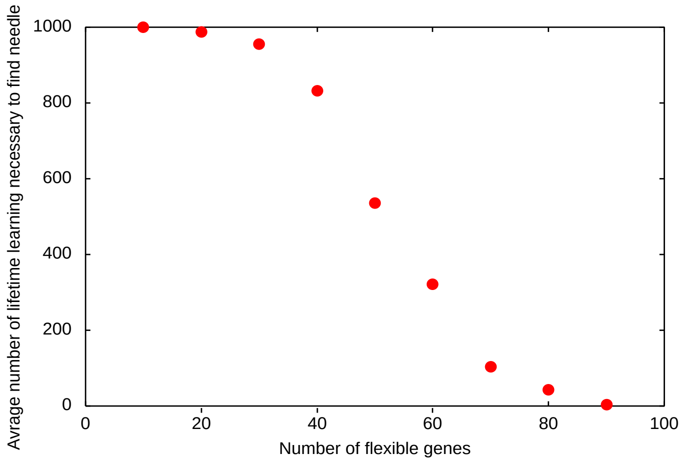
(How many successes out of 1000 learnings of 1000 phenotypes)



Goodnes depends on number of 9?

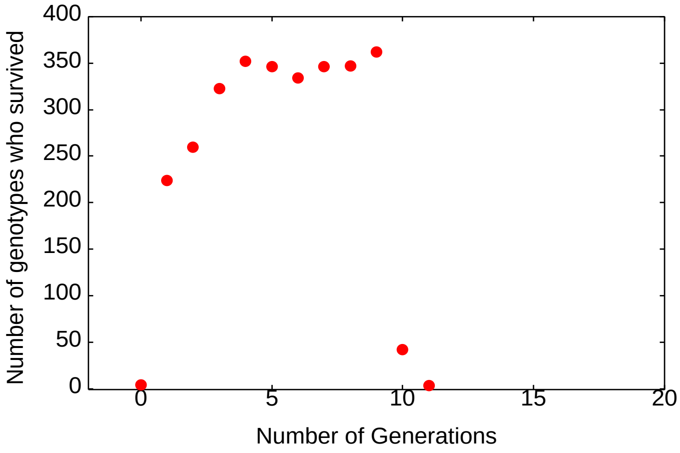


How many 9's are necessary?



Is the number of 9's decreasing?

An extinction!



What about Lamarckian Inheritance?

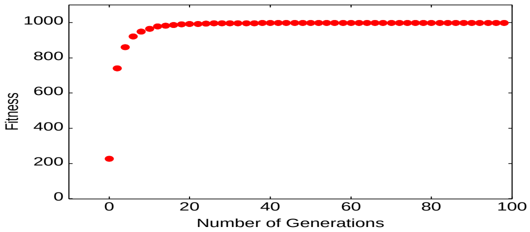
Turney (1996)

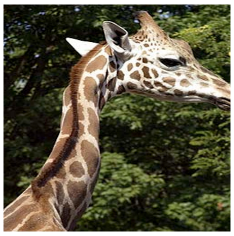
"Lamarckian evolution requires an inverse mapping from phenotype and environment to genotype ...

We believe that computing this mapping is intractable in general,"

Why Giraffe has a long neck?

Let's re-map a few of successful 9's in phenotype
to its genotype.





Eventually, all 9's disappeared to converge to the needle itself.

Alas, the number of total points searched is almost similar as a random search.

Reduced Even-Parity Boolean Function.

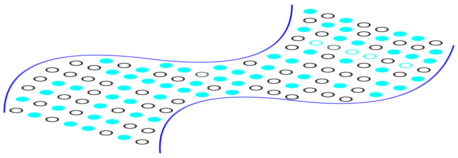
A discussion

(Yu & Miller, 2002)
"Finding Needles in Haystacks is not Hard with Neutrality."

vs.

(Collins, 2005)
"Finding Needles in Haystacks is Harder with Neutrality."

Reduced Even-n-Parity Problem

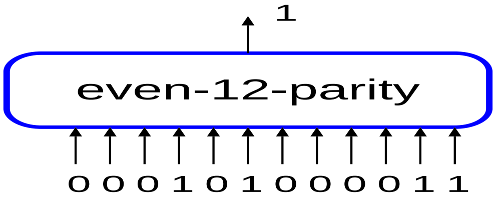


Even-parity Boolean function only with XOR and EQ

$\Downarrow$

A commonly used benchmark in GP, GA etc.

Reduced Even-n-Parity Problem (Cont'd)



with

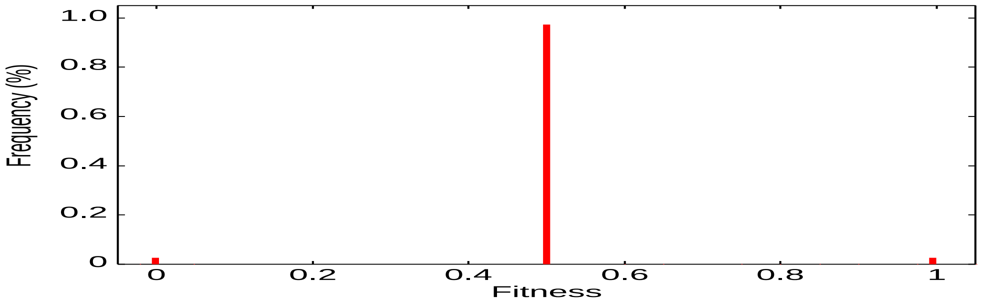
EQ

0 0	1
0 1	0
1 0	0
1 1	1

XOR

0 0	0
0 1	1
1 0	1
1 1	0

A strange fitness distribution



(1) All, (2) half, or (3) zero outputs are correct.

Yu & Miller wrote (2002)

Cartesian Genetic Programming

for $n = 12$

48 successes out of 100 runs each with 10,000 iterations
while
none of 4,000,000 random candidates failed

Yu & Miller's CGP

(1) Create 100 gates of XOR or EQ at random.

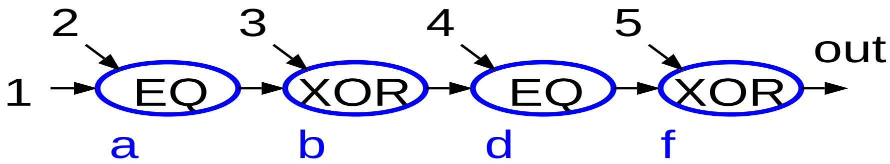
(2) Select randomly one output gate out of those 100 created.

(3) From one gate to the next, set the two input connections to either of the output line previously set gate or one of the n input lines at random.

An example genotype of CGP for Even-5-Parity

(a EQ 1 2)(b XOR 3 a) (c EQ a a) (d EQ 4 b)(e XOR 4 c) (c EQ a a) (f XOR d 5)

⇓



Neutral mutation on introns

Some genes do not contribute to phenotype $\Rightarrow$ *intron*.

And, so does a Mutation on intron $\Rightarrow$ *neutral*

Hence
Size of function is flexible!
But
it enhances the search?

Collins' doubt (2005)

"Unable to repeat the same performance as reported."

"We also implement and test a trivial random sampling algorithm which uses a fully expressed genotype. We show that this algorithm considerably exceeds expectations unexpectedly performing better than random sampling. An analysis of how simply removing introns from the representation can produce such a result is provided."

Collins' random sampling

(1) Create 100 random gates either from XOR or EQ.

(2) Set the output to the 100-th gate.

(3) For each gate:

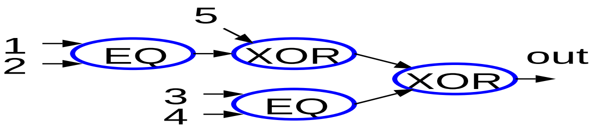
   (i) Set the type of the gate to either XOR or EQ at random.

   (ii) Set one of the input connection of the gate to the previous gate.

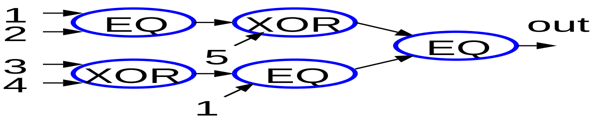
   (iii) Set the other input connection to a randomly selected input line.

(4) Repeat (3) up to N.

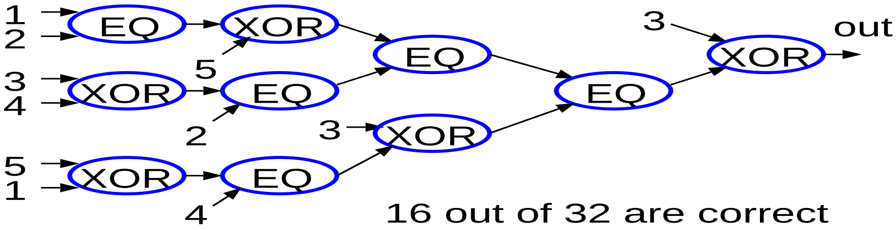
Collins' algorithm excludes many candidates such as...



16 out of 32 are correct



16 out of 32 are correct



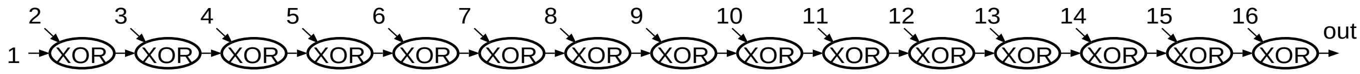
16 out of 32 are correct

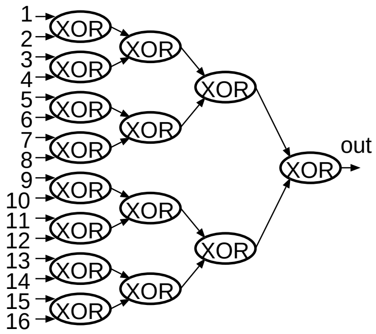
(Cont'd)

Anyway their solution is up to $n = 12$

On the other hand...

We human can create an even-n-parity for however large n.





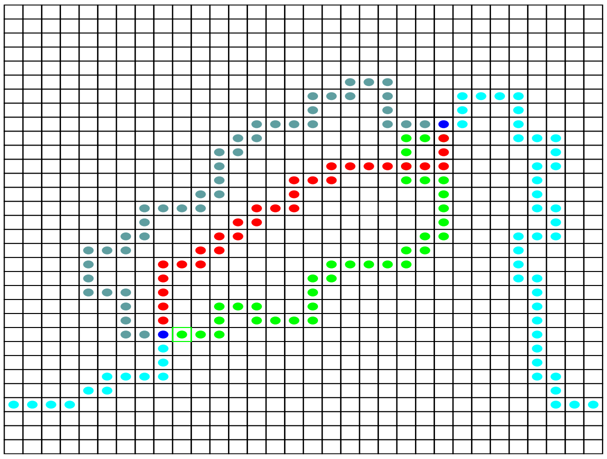
Yet another Bench Mark – A Robot Navigation in 2-D grid.

Can random-walker find a needle in a 2-dimensional grid?

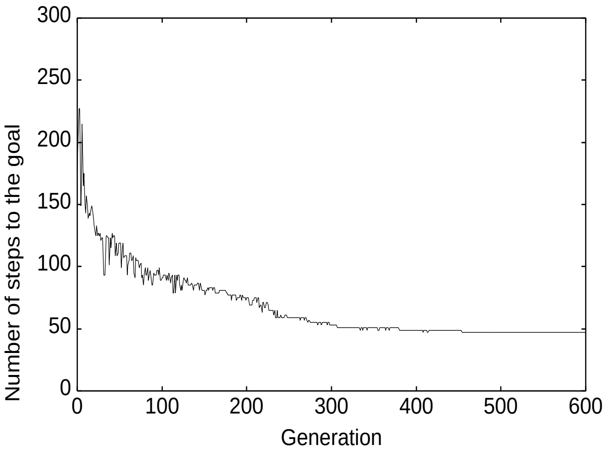
A scale limit of search with random walk.

A robot navigation

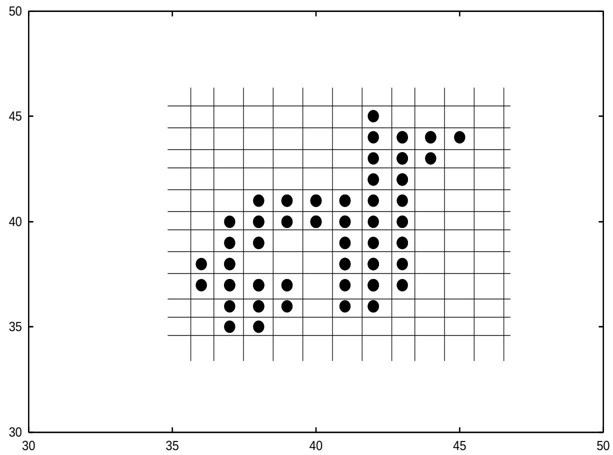
Can a learning enhance efficiency?

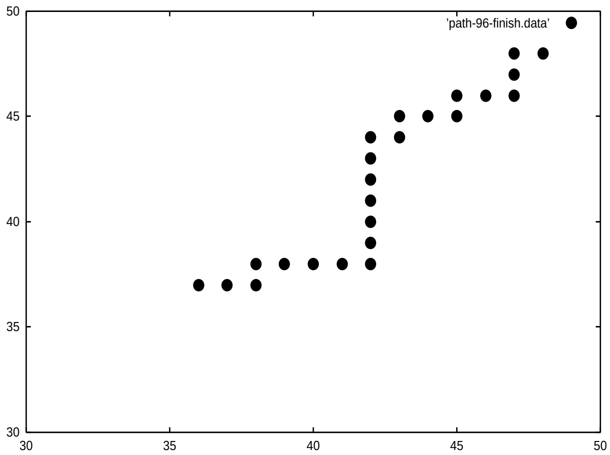


An evolution of routes



An evolution of routes



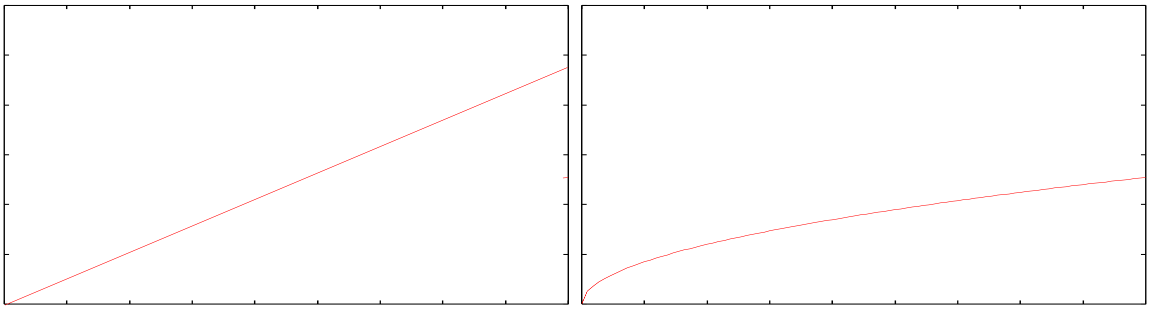


Quantum Computation

Quicker search with a quantum computation

Grover's search algorithm (1997)

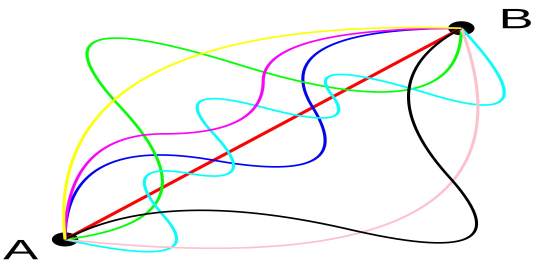
$O(\sqrt{N})$ instead of $O(N)$

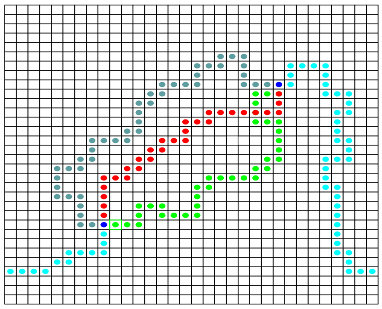


A strange tenet of quantum computation

When a particle goes from A to B,

it takes all possible paths all at the same time.





N/2 queries on average N-1 on worst case

A quantum walk on the line

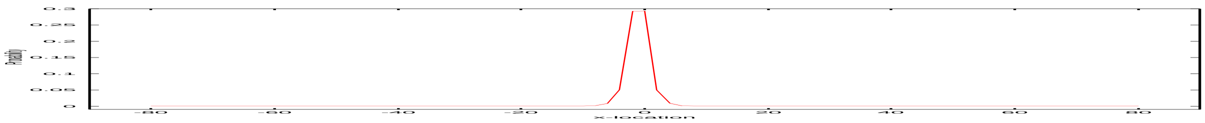
$$p(n,t) = \|\psi_L(n,t)\|^2 + \|\psi_R(n,t)\|^2$$

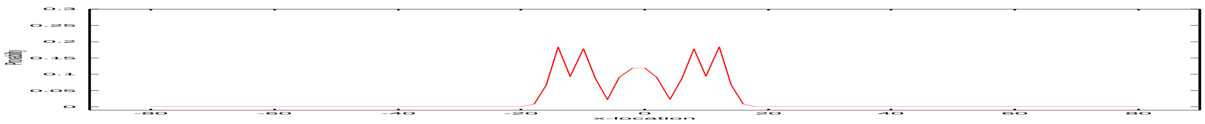
where

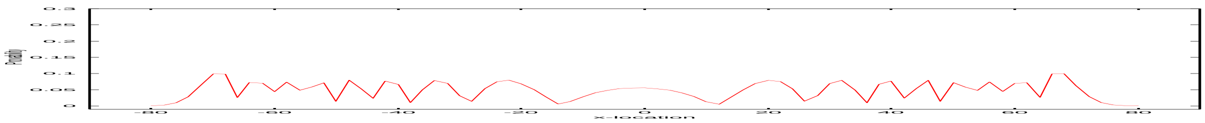
$$\psi_L = \frac{1+(-1)^{n+t}}{2}\int_{-\pi}^{\pi}\frac{dk}{2\pi}(1+\frac{\cos k}{\sqrt{1+\cos^2 k}})\exp\{-i(\omega_k t + kn)\}$$

$$\psi_R = \frac{1+(-1)^{n+t}}{2}\int_{-\pi}^{\pi}\frac{dk}{2\pi}\frac{\exp(ik)}{\sqrt{1+\cos^2 k}})\exp\{-i(\omega_k t + kn)\}$$

A quantum walk on the line







Concluding Remarks

- We have to avoid an effect of *like-to-hear-what-we-would-like-to-hear.*