Spike timing dependent plasticity (STDP)
to make robot navigation intelligent

Akira Imada

Brest State Technical University (Belarus)

This is not a success report

but

a guidline of our ongoing project

or

a report of what we are currently planning.

To start with

a benchmark to know

if what we are planning will work or not.

Path Planning

or

Robot Navigation

$\Downarrow$ where

aim is usually to minimize the route from start to goal, but...

Task of minimizing navigation on its own is not so difficult

$$\Downarrow$$

A-star, Genetic Algorithm, Reinforcement Learning, Neural Network,

Ant Colony Optimization, etc...

Genetic Algorithm (GA)

$\Downarrow$

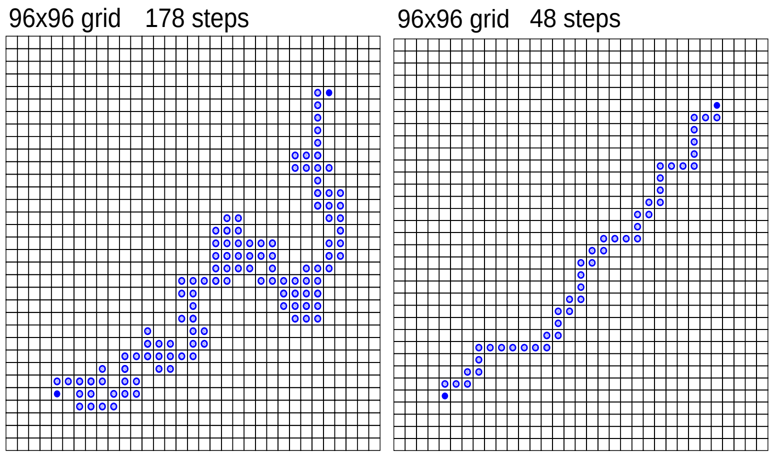
Agent decides actions following its chromosome

e.g.

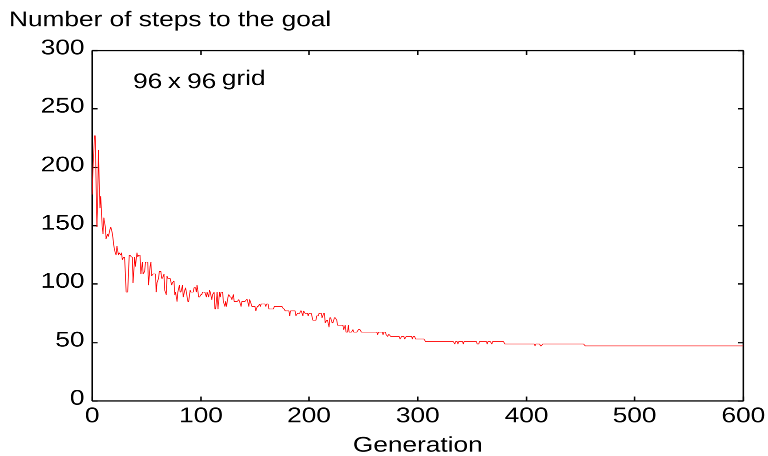
with chromosomes like

(4 1 1 2 1 3 2 3 3 4 2 1 3 4 3 2 3 3 2 1 1 3 2 1 3 ...)

Random walk evolved to be minimized



Task was easy



Let's extend it

From Minimization to Maximization

$$\Downarrow$$

We don't care goal points but to maximize an exploration

A Camel in a Desert

The 52nd problem

in the

"Propositiones ad acuendos inventes" (in Latin)

by

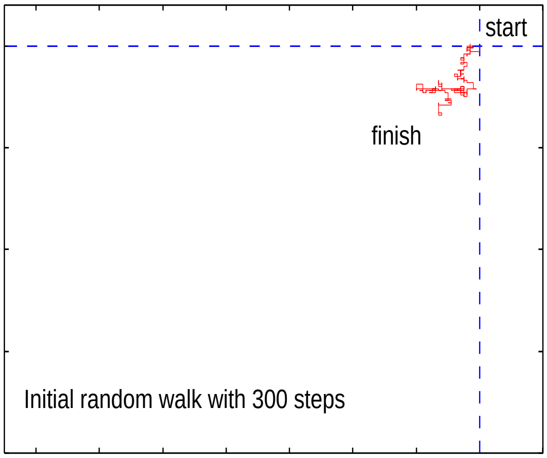
Alcuin of York (732–804)

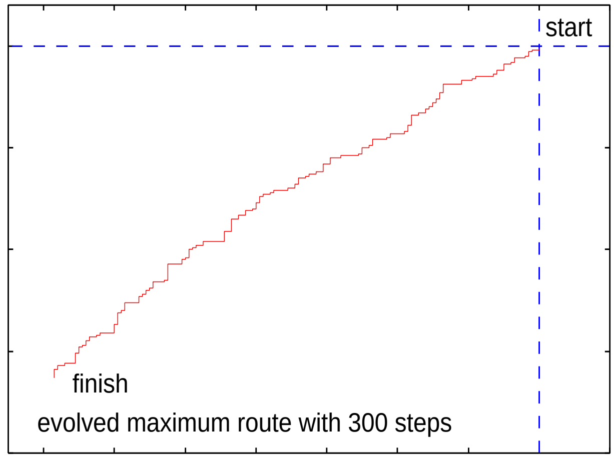
Can a camel maximize its exploration in a desert surviving with grains on his back?

Modern version: From Camel to Jeep in a Desert

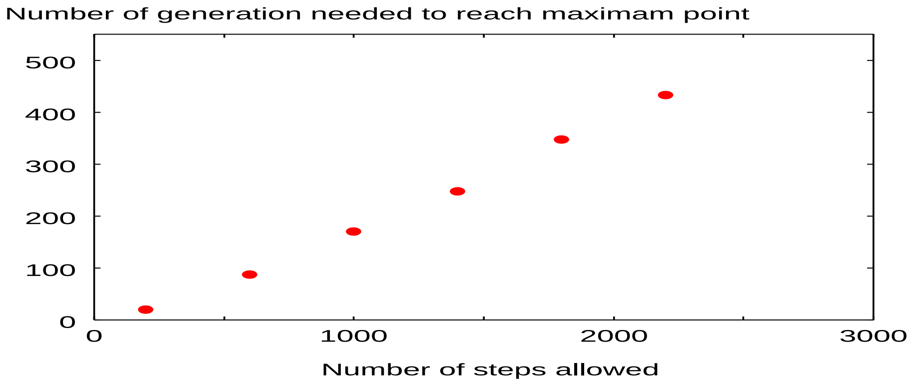
Can a jeep maximize its exploration in a desert with a tank of fuel?

Still it's not so difficult by GA for example





Task was rather easier



The complexity is $O(N)$

Why not extend it further

Let's make it return to the point it started

$$\Downarrow$$

while also maximizing an exploration

Planet Land-rover Problem

From A (start) to B (goal) minimizing its route

$\Downarrow$

From A to A maximizing its route

(This is our proposal here as a benchmark)

Can the rover return to the base with its exploration being maximum spending a limited energy given when it starts?

Much more demanding

$$\Downarrow$$

Taking GA as an example, what might be a fitness?

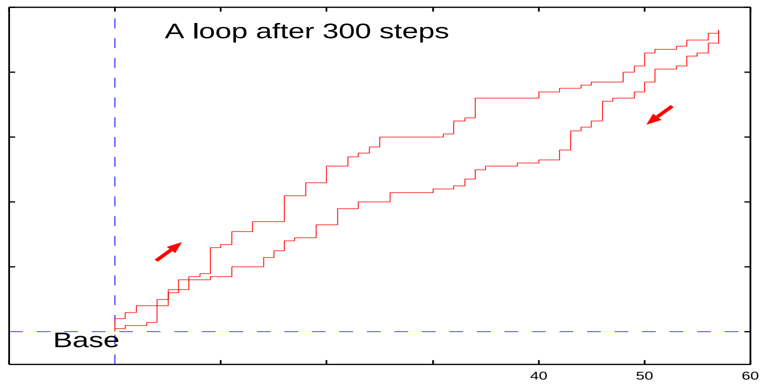
e.g.

$\{$*total length*$\}$ *&* $\{$*how often it has crossed previous route?*$\}$

with

Multi Objective GA (MOGA)

A Heuristic create such a route



What heuristic do you guess?

Topic of today's talk:

What is intelligence?

$\Downarrow$

Intelligence should be spontaneous, flexible, or unpredictable more or less

Stolle et al. (2002)

"... every day we might be cooking a different breakfast, but the kitchen layout is the same from day to day."

“I beg your pardon?”

Do they change the expression for an easier understanding?

or

Do they repeat the expression, maybe with a bigger voice?

What if your canary stop singing?

Legendary three strategies in Japan.

(1) Just wait until she sings again.

(2) Try something so that she sings again.

(3) Kill her if she doesn't sing any more?

To be intelligent,
action should be different more or less
even in an exactly identical situation

Performance is not intelligent

if we use

deterministic A-star, GA, NN with fixed weights?

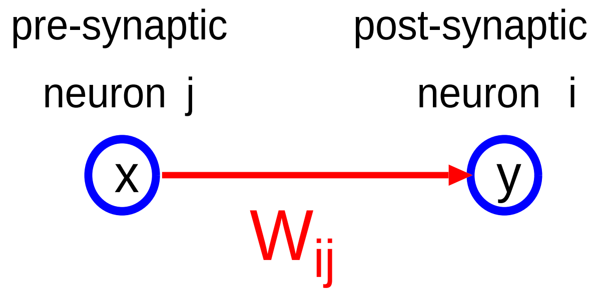
Let's make a learning occur during behavior

$$\Downarrow$$

Evolution of Learning

A notation

Hebbian Learning of NN



$$w_{ij}(t+1) = w_{ij}(t) + \eta x_j(t) y_i(t) \quad (x_j, y_i \in [0,1])$$

Floreano's approach (2000)

$$\Downarrow$$

Modification of w_{ij} during exploration

with either one of four Hebbian and Hebbian-like rules

Hebbian learning

$$(1) \qquad \Delta w = (1 - w)xy$$

Hebbian type learning

Weaken if the post-synaptic is active, while the pre-synaptic is not

$$\Delta w = w(-1 + x)y + (1 - w)xy \tag{2}$$

Weaken if the pre-synaptic is active, while the post-synaptic is not

$$\Delta w = wx(-1 + y) + (1 - w)xy \tag{3}$$

And

Strengthen if the two have similar activity and weaken otherwise.

$$\Delta w = \begin{cases} (1-w)F(x,y) & \text{if} \quad F(x,y) > 0 \\ wF(x,y) & \text{otherwise} \end{cases} \tag{4}$$

where

$$F(x,y) = \tanh(4(1-|x-y|)-2)$$

Evolution of leaning

Which rule and what parameter's value should be assigned to each of the synapses?

$\Downarrow$

Evolution leads the combination to the optimum.

Stanley's implementation (2003)

$$\Delta w = \begin{cases} \eta_1(1-w)xy + \eta_2 wx(y-1) & \text{for excitatory neuron} \\ -\eta_1(1-w)xy + \eta_2(1-w)x(1-y) & \text{for inhibitory neuron} \end{cases}$$

$$\Downarrow$$

Evolution is on (η_1, η_2) of each synaptic weight

Their result

Starting with random weights every time anew
the weights are modified step by step by the rule it learned

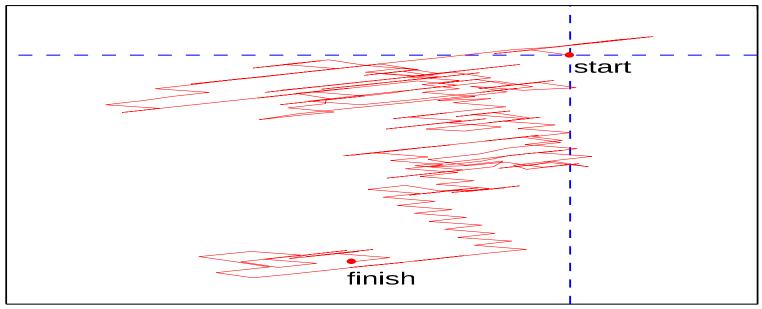
$$\Downarrow$$

Every run is different depending on the initial random weights

More general learning (Durr et al. 2008)

$$\Delta w = \eta(Axy + Bx + Cy + D)$$

An example of exploration by a recurrent NN with random weights



Can it learn so that path will finish at the start point?

Our Aim

$$\Downarrow$$

Learning should occur during a random exploration

with a spiking neural network seeking its biological plausibility

Learning by STDP

(Spiking neuron's version of Hebbian Learning)

$$\Downarrow$$

The topic is still open!

Meunier et al. (2005)

$\Downarrow$

"Up to now, nobody has been able to show how it is possible to learn with STDP..."

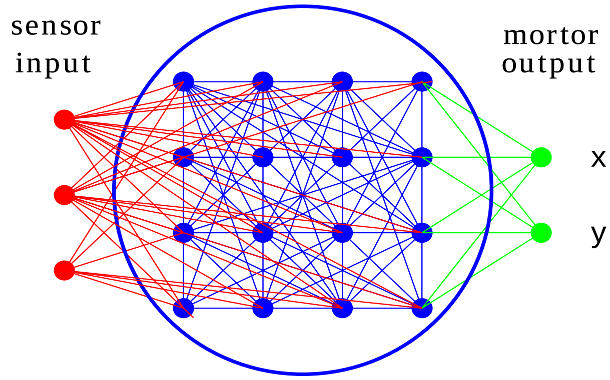
Farries et al. (2007)

$$\Downarrow$$

"Although synaptic plasticity is widely believed to be a major component of learning, it is unclear how STDP itself could serve as a mechanism for general purpose learning."

Why not challenge?

Architecture we are planning

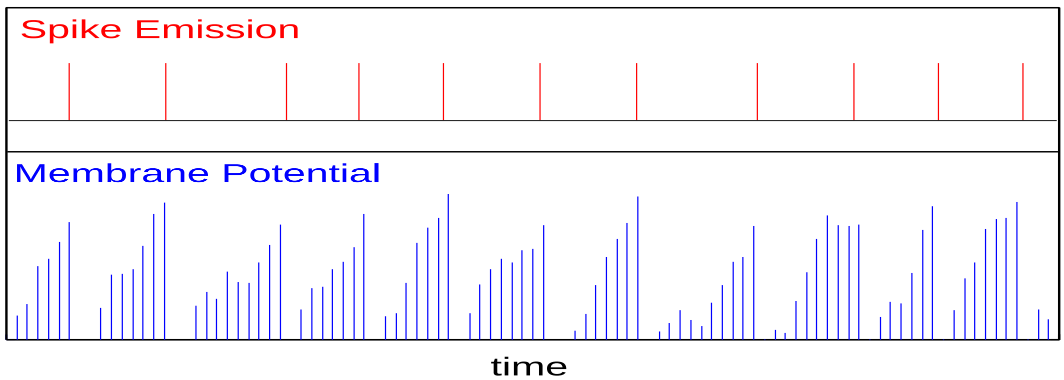


$\Downarrow$ a.k.a.

Echo State Network, Liquid State Network or Reservoir Computing

Integrate-and-Fire Model

$$u_i(t) = u_r + (u_i(t - \delta t) - u_r) \exp(-\delta t/\tau) + \sum_j w_{ij} f_j(t - \delta t)$$

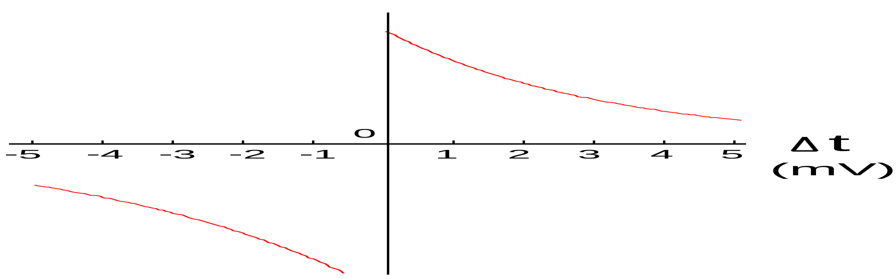


What is STDP?

$$W(\Delta t) = \begin{cases} A_+ \exp(-\Delta t/\tau_+) & \text{if} \quad \Delta t \geq 0 \\ -A_- \exp(-\Delta t/\tau_-) & \text{if} \quad \Delta t < 0 \end{cases}$$

where $\quad \Delta t = t_{post} - t_{pre}$

$W(\Delta t)$



In short

potentiation when a pre-synaptic neuron fires
shortly before a post-synaptic neuron

and

depression when the pre-synaptic neuron fires shortly after.

It's too basic to be used

Reward-modulated STDP Learning
(Florian 2007)

$$w_{ij}(t+\delta t) = w_{ij}(t) + \gamma r(t+\delta t)\zeta_{ij}(t)$$

where

$$\zeta_{ij}(t) = P_{ij}^{+}(t)f_i(t) + P_{ij}^{-}(t)f_i(t)$$

$$P_{ij}^{+}(t) = P_{ij}^{+}(t-\delta t)\exp(-\delta t/\tau_{+}) + A_{+}f_j(t)$$

$$P_{ij}^{-}(t) = P_{ij}^{-}(t-\delta t)\exp(-\delta t/\tau_{-}) + A_{-}f_j(t)$$

Reinforcement Learning (RL)

What an agent learns is a *Policy*

and

Policy is how to select an *Action* in a given *Situation (State)*

maximizing

total *Rewards* it occasionary will receive from the environment

Policy indicates agent which action should be chosen in any possible situation

$\Downarrow$

While in GA
Agent decides actions following its chromosome

in RL
Agent decides actions following its policy

However

our world is with no obstacle, no wall, no corridor, only one goal

$\Downarrow$

Everywhere no reward except for one point (goal)

$\Downarrow$

A needle in a Haystack

(Rewards are not likely to be encountered by a random exploration.)

We have not succeeded yet

$\Downarrow$

Nevertheless we think of a further extention of the problem as a bigger challange

$\Downarrow$

What if the rover carries container(s) for fuel?

Jeep Problem

Where jeep explore is 1-D desert.

Jeep can unload its fuels anywhere in the desert.

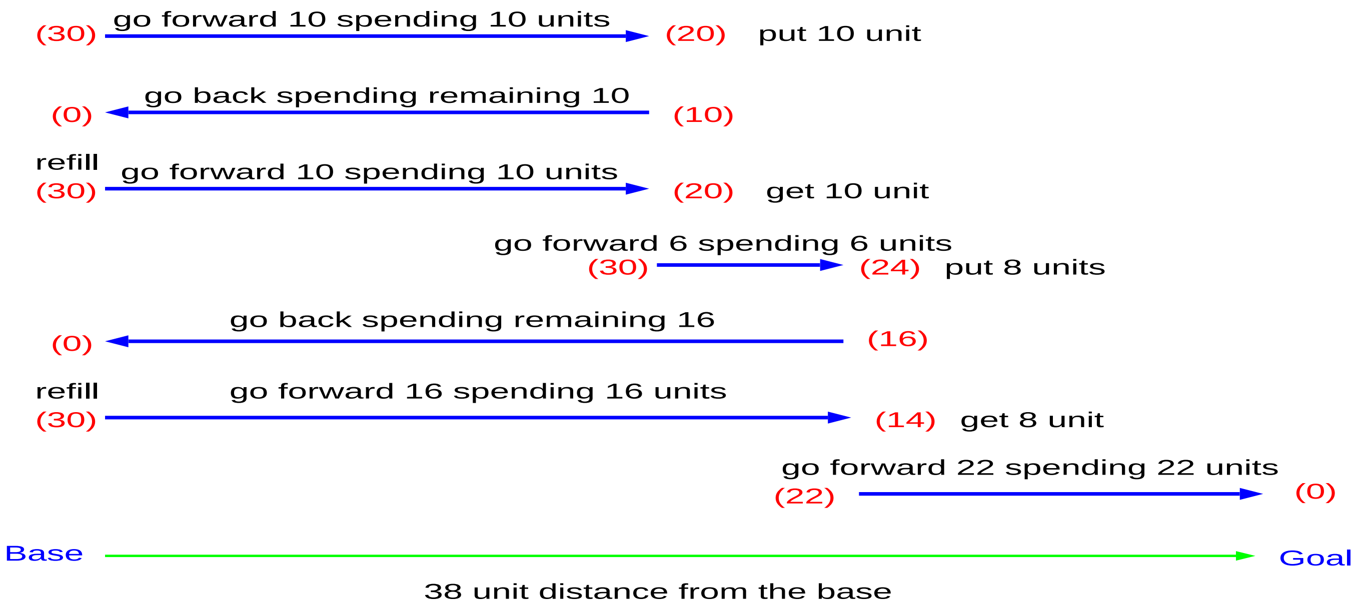
Fuels can be filled only at the base.

Jeep can go back to the base n times to re-fill its tank.

$\Downarrow$ thus

The jeep should maximize its penetration

An example of a success



Let's give the rover the same condition (though in 2-D)

$\Downarrow$

A very tough benchmark and infinitely large number of solutions

$\Downarrow$

Good as a benchmark to check *so-far-mentioned* intelligence

Let me conclude

To claim artificial NN to be intelligent

a different action more or less should be made under identical situation.

$\Downarrow$

We want to, or we are going to realize it
by
Spiking NN with learning by STDP

also

to be more biologically plausible.

SONY's AIBO

can excellently learn but still repeats same action in the same situation

Hoping collaborations

to design

an agent like real human intelligence

Thank You!