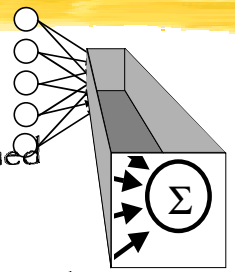
(Learning Categorization with a Race Model)

Brain: Democracy or stock market?

Denis Cousineau
Sébastien Hélie
Christine Lefebvre
Université de Montréal

Vision Integration Cognition

Neural Networks: A paradigm with an unquestioned assumption: The sum.



Why the sum?

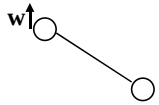
All current neural networks are built around a central operator: the sum. Is this operator absolutely necessary? Or is it there simply for convenience (in the inner product)? As an historical remnant?

In other words, can we replace it with another operator and still have a functional network?

The answer is Yes.

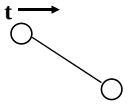
Alternatives to build networks on

a- Strength-based networks: A Democracy
(or "weighted-sum model")



Increasing strength helps concentrate on the relevant inputs. The "Strength"-based network uses the standard inner product.

b- Time-based network: A Stock market
(or "winner-take-all" model)



Increasing availability helps concentrate on the relevant inputs. The "Time"-based network uses a redefined dot product.

$$\mathbf{I} \cdot \mathbf{W} = \mathbf{I}_{(\times,\Sigma)}\mathbf{W} \quad \text{Sum}$$

$$\mathbf{I} \sim \mathbf{T} = \mathbf{I}_{(\times,\vee)}\mathbf{T} \quad \text{Min}$$

• What is a Race Network?

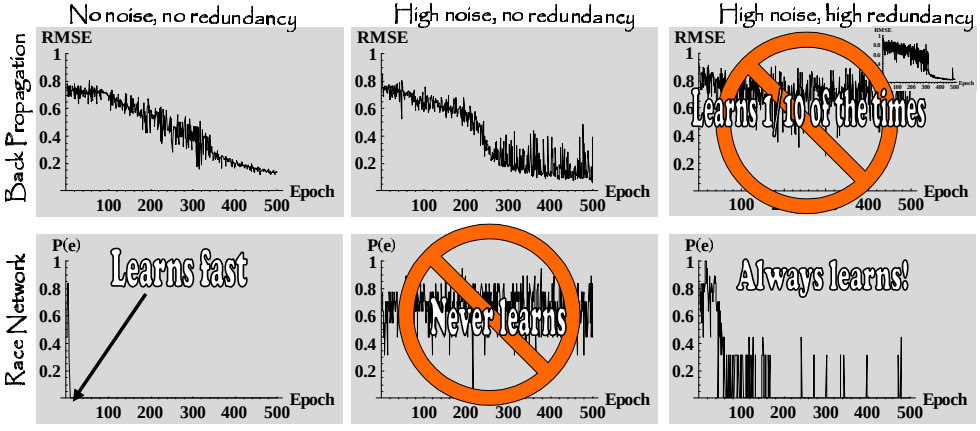
Race networks are the time-based counterpart to Neural networks. They don't wait for all the inputs before making a decision. Instead, the outputs react on the first few inputs, the fastest. As such, it is a Race Model. It has a threshold that says how much evidence must be collected. As such, it is an Accumulator Model.

Its learning rule, the redefined Delta Rule, or $\tilde{\Delta}$ rule, distributes the blame on an error by:
Reducing the sensibility of a connection that contributed to fill the faulty accumulator. As a result, this connection will take longer to fire if an input is presented again.
Increasing the size of the faulty accumulator.
This simple learning rule is very efficient (see below).

Future Works

We plan to test an alternative to the $\tilde{\Delta}$ rule, the (redefined) Hebbian learning rule. Such a Race Network will become equivalent to an Unsupervised Neural Networks and identical to a Kohonen Network.

Comparing Back Propagation Networks to Race Networks on a XOR problem

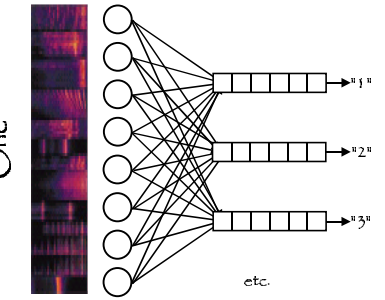


Note: 1 epoch = 10 trials; RMSE: Squared error; P(e): Percent of error

We compared a classic Back Propagation network (with the Δ rule) to a Race network (with the $\tilde{\Delta}$ rule above).
Because imperfect transmission of information is highly likely, we manipulated noise (none, high) by adding random values to each input. In addition, because the human brain is likely to involve a large amount of redundant connections, we manipulated redundancy (none, 8 times) by duplicating the input channels and the connections prior to learning.
The Back Propagation network can accommodate high noise or high redundancy (not shown) but not both. Their joint effects are catastrophic.
The Race network can accommodate high redundancy (not shown) but not high noise. However, their joint effects are canceling each other: high noise is not a problem if compensated by high redundancy.

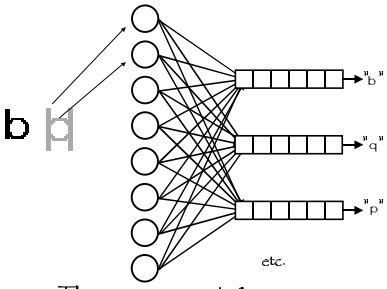
Applications of the Race Network

a- Recognizing utterances of the digits "0" to "9"



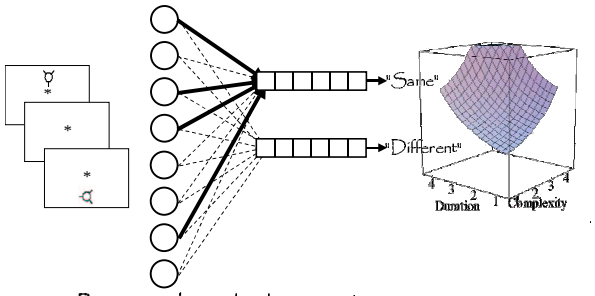
Using audiograms, the inputs are the moment at which a given frequency occurs for the first time.

b- Recognizing visual letters



The inputs are simple features. The final size of the accumulator indicates how many features are sufficient to recognize a letter. As such, this system supports Reduction of information

c- Priming of responses



Race networks predict that prime duration, prime energy and prime complexity all have an impact on Response Times.

One model, with no change, does it all!

Modeling psychological phenomena is easy with the Race network. Maybe the neural networks are a local minima in the search for the foundations of human cognition?
According to Estes, to have a model, you need to prove the necessity of the assumptions. We showed here that the weighted sum assumption is not necessary.