Artificial networks with exponential representational power

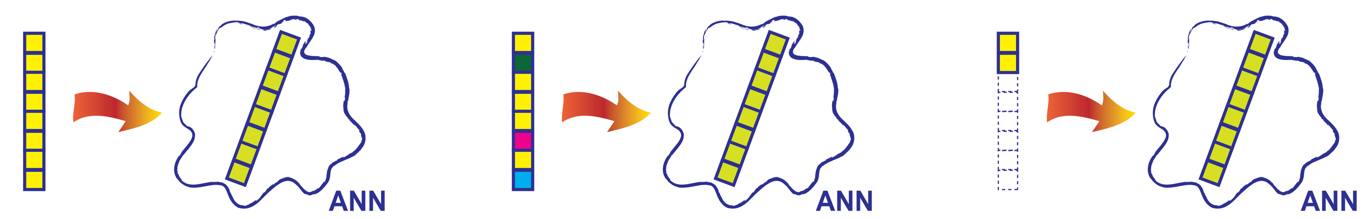
Demian Battaglia[1,2]

[1]Bernstein Center for Computational Neuroscience, [2]Max Planck Institute for Dynamics and Self-organization, Göttingen (Germany)

demian@nld.ds.mpg.de

Intro: associative networks

An associative attractor network performs three basic operations:

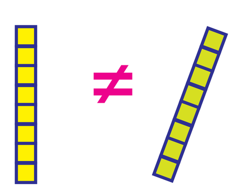


Converges to an attractor correlated with the input pattern, allowing **pattern representation** *and* **working memory**

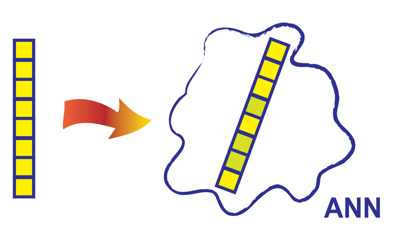
Collapses similar patterns into a same internal category, allowing **pattern restoration**

Can associate a meaningful attractor to an uncomplete pattern, allowing **pattern reconstruction**

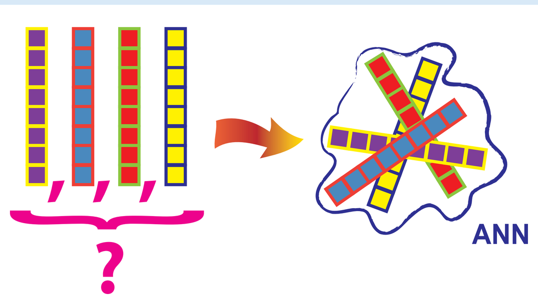
Distortion and learning



The internal state of the network does not correspond perfectly to the pattern. **Learning** *can be used to reduce distortion.*



Representational capacity



How many *(classes of) patterns can be faithfully represented? A* **Hopfield network of size *N* neurons** *can represent* **almost without distortion**:

$$C = \alpha N$$

where the critical loading is α ~1.14. Beyond critical capacity attractors with small overlaps with the learned patterns do exist, but are usually considered useless.

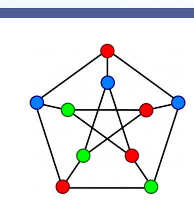
Can we devise a *small size* network associating patterns to an *exponential* number of internal categories?

Digression: constraint satisfaction

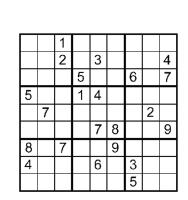
N variables $s_i = v_1, v_2, \dots v_Q$

M constraints $C_a = C_a(s_{a_1}, \dots s_{a_K}) = 0, 1$

$$\mathcal{H}[s_1, \dots s_N] = \sum_{a=1}^{M} C_a$$



Graph coloring

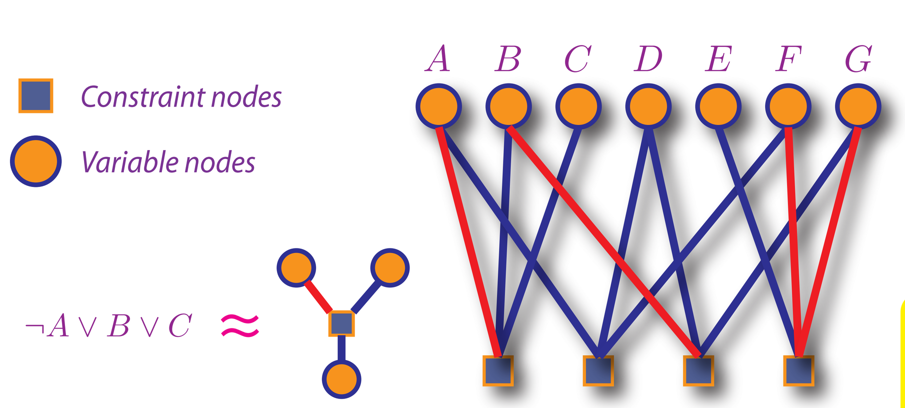


Sudoku

A **Constraint Satisfaction Problem (CSP)** *is satisfied when the configuration of the N discrete variables satisfies all the M constraints (i.e.* **minimizes the energy***)*

$$(\neg A \vee B \vee C) \wedge (A \vee D \vee F) \wedge$$
$$(\neg B \vee D \vee G) \wedge (\neg F \vee E \vee \neg G) \dots$$

Boolean Satisfiability



A **random instance** *of a CSP can be represented as a diluted bipartite* **network** *(factor graph)*

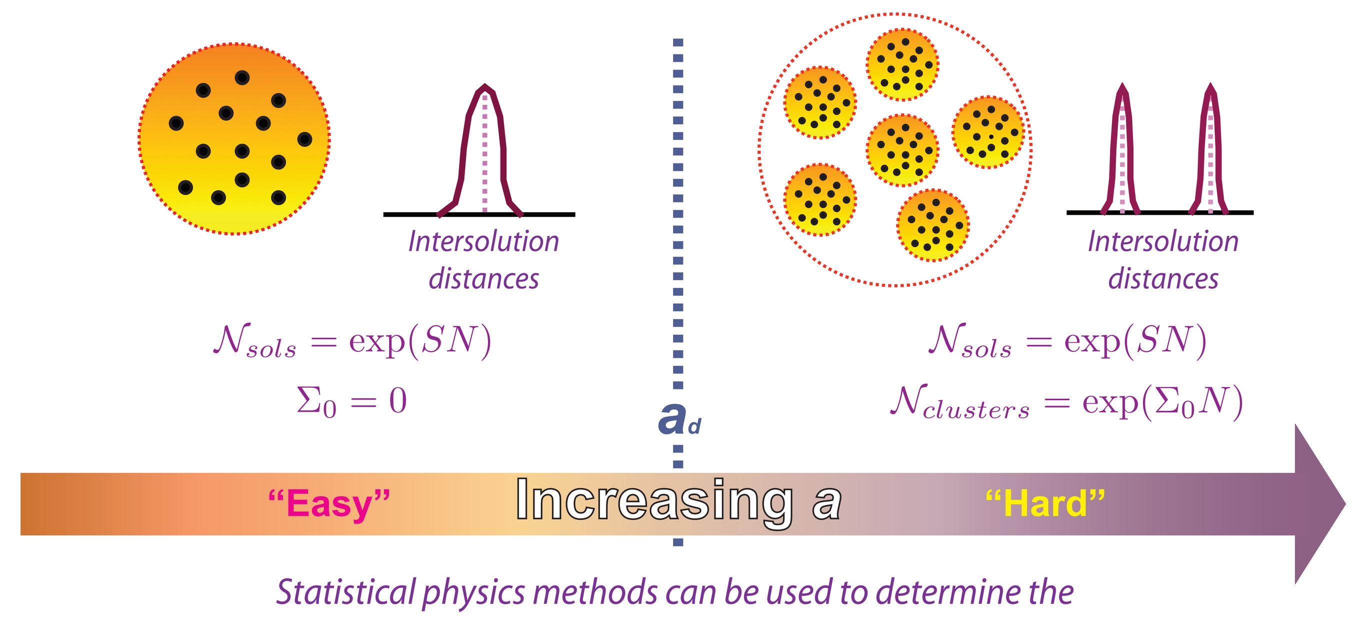
$s_i = +1, -1$ (True or False)

$$C_a(s_i, s_j, s_k) = \prod_{l=i,j,k} \frac{1 + J_{al} s_l}{2}$$

$J_{al} = \mp 1$ (directed or negated)

Many clusters :-(

For a random CSP network, the critical parameter is ***a* = *M* / *N***



Statistical physics methods can be used to determine the phase diagram of random CSPs (e.g. random 3-SAT)

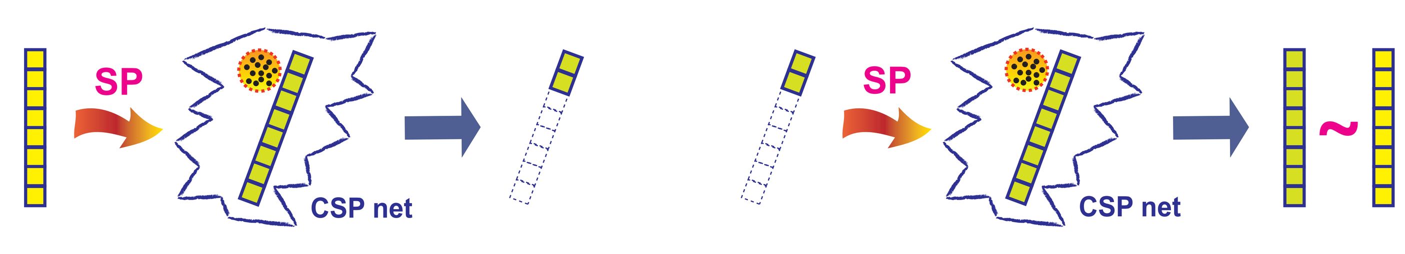
The space of solutions breaks into an exponential number of clusters — Finite Complexity $\Sigma_0 > 0$

Drop in the performance of local search algorithms: clustering is BAD for optimization

Survey Propagation selects clusters

This clustering has already been exploited as a COMPUTATIONAL RESOURCE, for **lossy data compression** applications (Battaglia et al. 2005; Ciliberti et al. 2006)

In a nutshell, solution clusters are **attractors** *for the dynamics of a bayesian* **message-passing** *algorithm,* **Survey Propagation (SP)**



Encoding — *Decoding*

A "spiking" random 3-SAT network

The full message-passing bayesian algorithm SP can be **approximated** *over time, by a* **network dynamics** *in which the nodes of the CSP network exchange* **spikes of binary information**

Three types of messages:
h_{ka} are sent by variable nodes to constraint nodes
h_{ai} are sent by constraint nodes to variable nodes
h_j are related to an external input pattern $\vec{\xi}$

For random 3-SAT the messages are:

$$h_{ai} = [J_{aj} h_{ja} + J_{ak} h_{ka} - 1]_+ = 0, 1$$

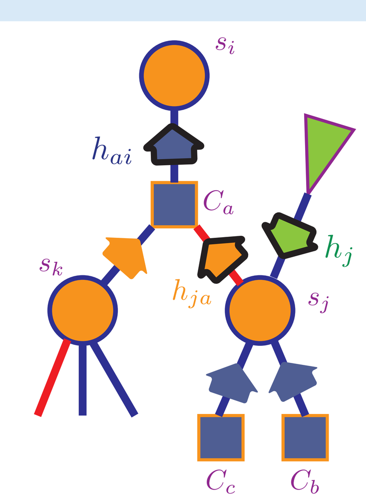
$$h_{ja} = -\text{sign}(J_{bj} h_{bj} + J_{cj} h_{cj} + h_j - \theta_j) = -1, 0, 1$$

$$h_j = -\xi_j \text{ injected with probability } \nu$$

Transient threshold shifts
reinforce convergence and resistance to noise

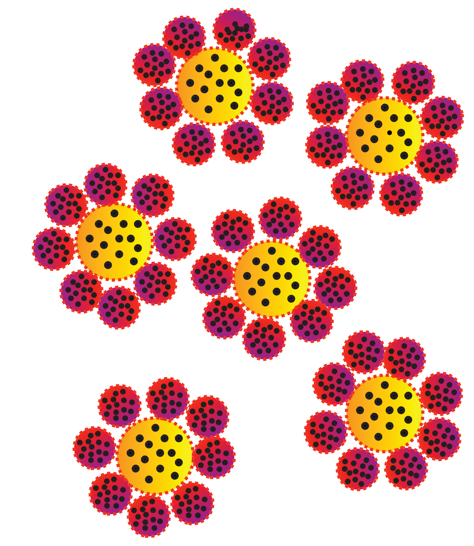
$$\theta_j = \text{sign}\ \overline{s_i} = -\text{sign}\left[\overline{\sum_a J_{aj} h_{aj} + h_j}\right]$$

They can be reset to zero when dynamics has stabilized.



The iterative update dynamics **does not freeze** *into a fixed-point, but* **stays confined** *in a small region determined by the input pattern*

Many associations :-)

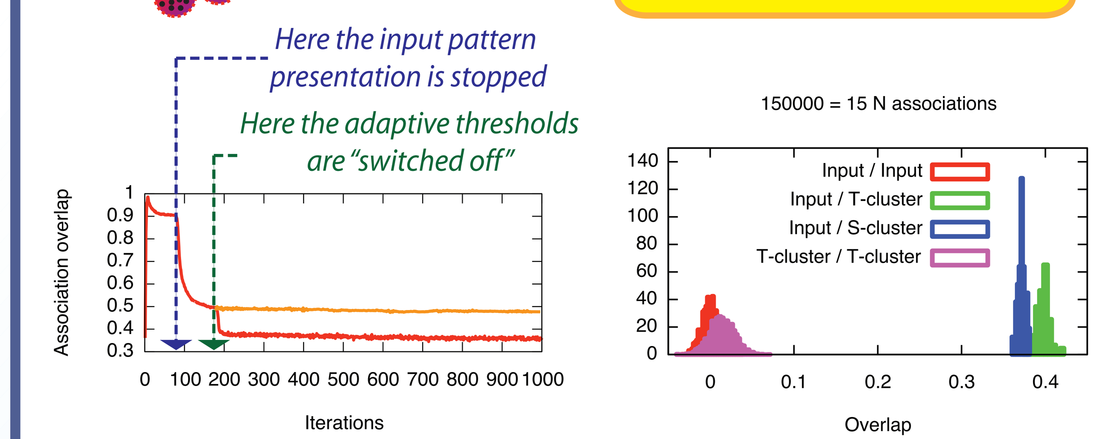


How many (classes of) patterns can be meaningfully represented?

The spiking CSP network select as internal representations clusters of weakly energetic configurations **(T-clusters)** *nearby to solution clusters* **(S-clusters)**.

We can estimate their number as:

$$\text{"}C\text{"} \sim \exp(\Sigma_0 N)$$

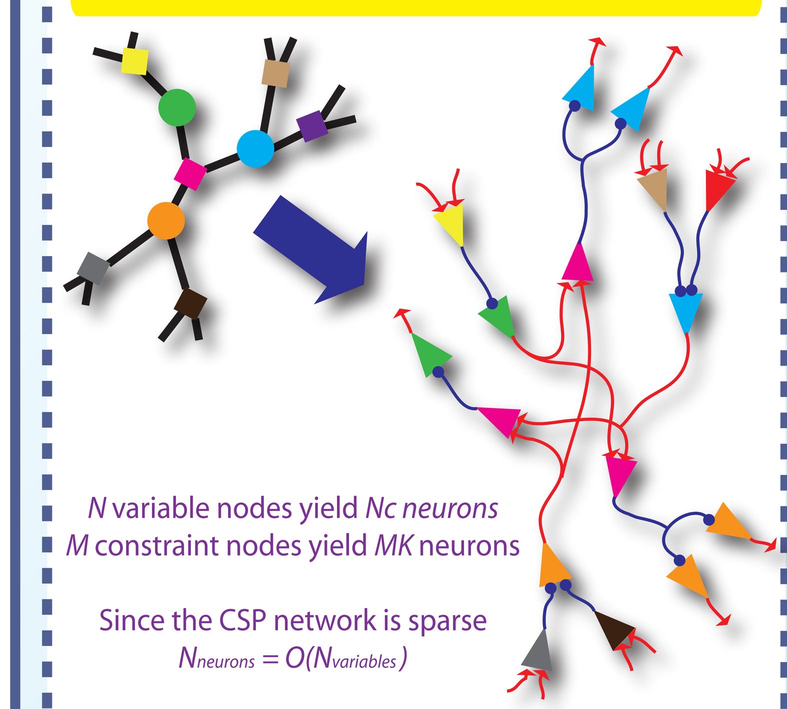


Dynamics of the association — **Orthogonal representations**

1. **Exponential number** of internal representations (*association, restoration and reconstruction*)
2. The **distortion** of these representations **is large**
3. Still **uncorrelated patterns** yield **uncorrelated clusters**
4. These representations are **unlearned**.

Neuronal 3-SAT

It is possible to build a directed **network of 0/1 binary neurons** which *emulates* the dynamics of a random 3-SAT network



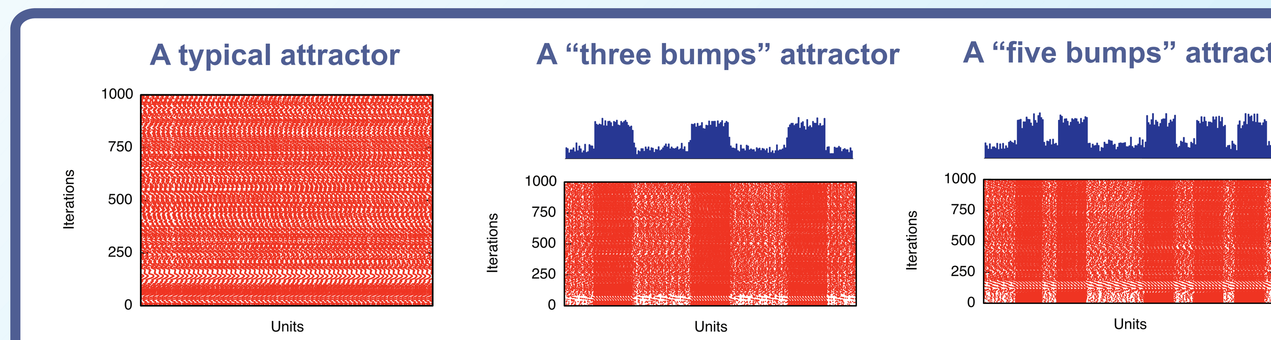
N variable nodes yield *Nc neurons*
M constraint nodes yield *MK* neurons

Since the CSP network is sparse
$N_{neurons} = O(N_{variables})$

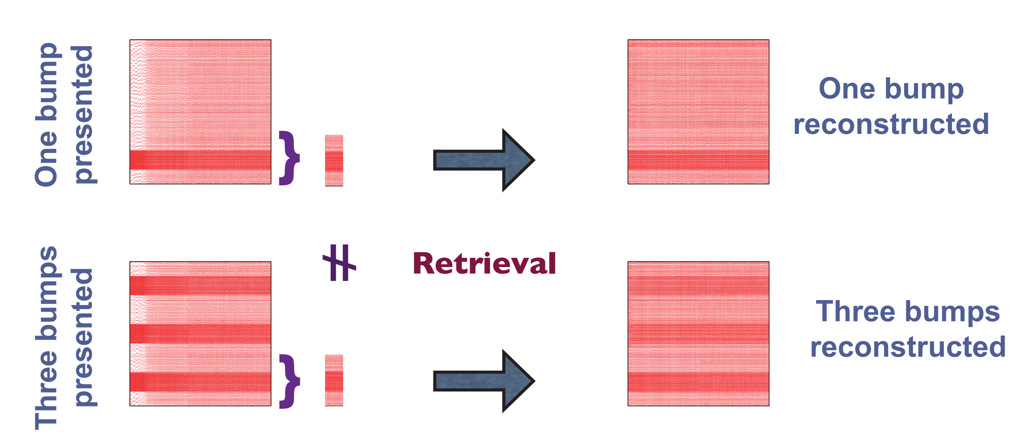
Very artificial...
Maybe have a look to *network motifs*?

Not better than Hopfield... but certainly different!

Toy-model: a network that build symbols



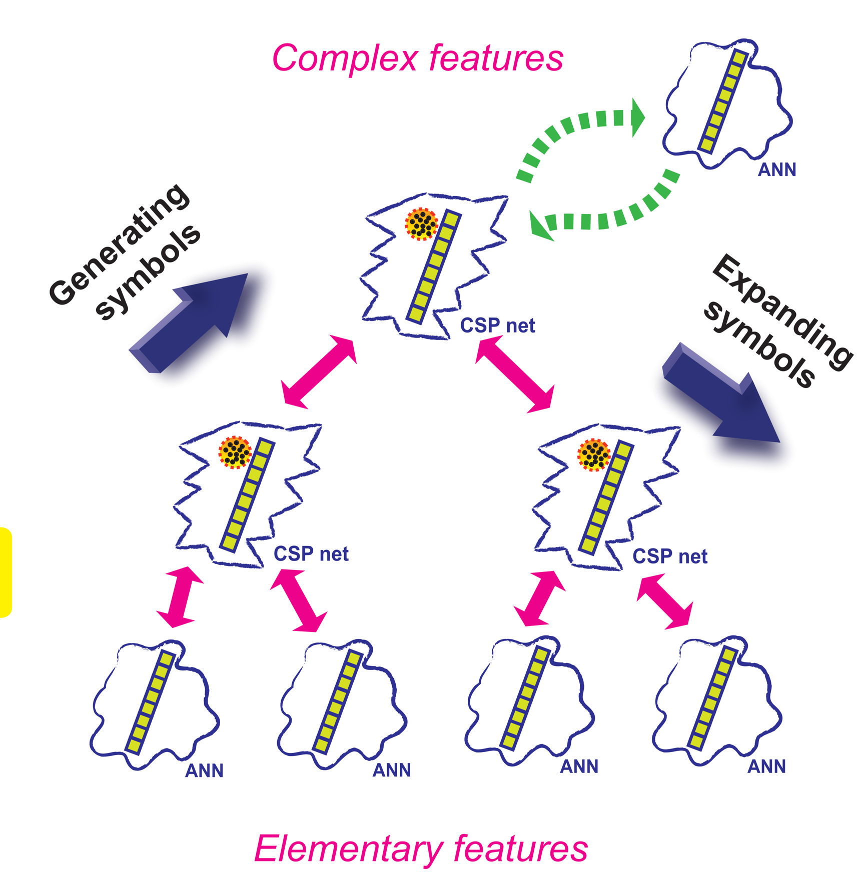
Among the 10^{131} attracting clusters that our $N=10^4$ variables 3-SAT net has, there are also some attractors that seem to have a spatial structure. Let play with them for the sake of visualization.



The second output bump is a **symbol** *for three input bumps*

Use a single CSP network of size *NQ* to represent **combinations** of outputs of *Q* feature networks of size *N* each

NB: a Hopfield network for the same task would have size $O(N^Q)$



References

[1] Battaglia D, Braunstein A, Chavas J, Zecchina R (2005) Source coding by efficient probing of ground state clusters. Physical Review E 72, 015103

[2] Battaglia D (2009) Neuron-less neural-like networks with exponential association capacity at tabula rasa, Proc. IWINAC 2009, Lecture Notes in Computer Science 5601/2009, 184-194

The author wishes to acknowledge Riccardo Zecchina, Erez Maharshak and Jozsef Fiser for inspiring discussions.