

Réseaux de cliques neurales

Vincent Gripon

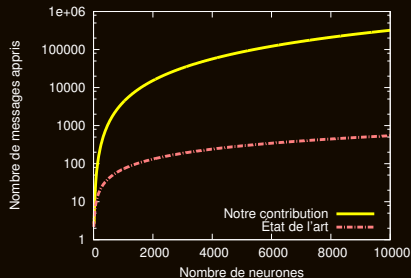
Télécom Bretagne, Lab-STICC

École doctorale SICMA

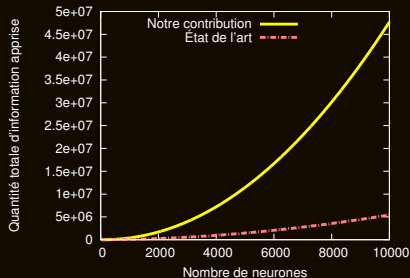
20 juillet 2011

Apprentissage de messages dans des réseaux de neurones récurrents

Diversité d'apprentissage



Capacité d'apprentissage

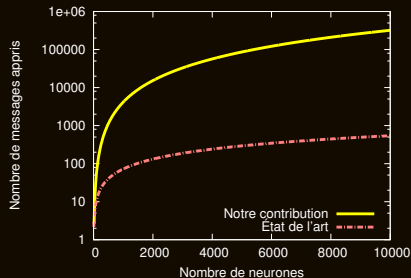


Notre contribution

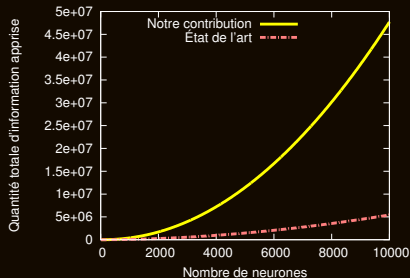
État de l'art

Apprentissage de messages dans des réseaux de neurones récurrents

Diversité d'apprentissage



Capacité d'apprentissage



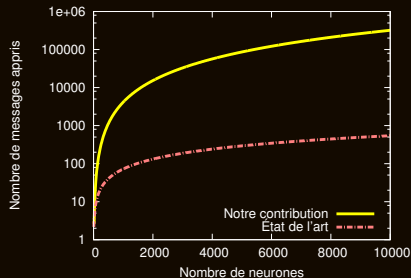
Notre contribution

parcimonie
codage correcteur

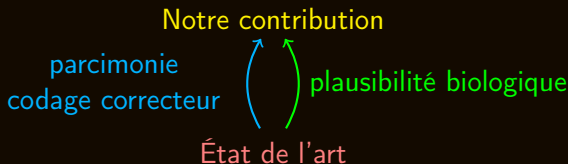
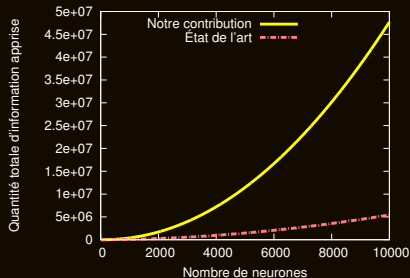
État de l'art

Apprentissage de messages dans des réseaux de neurones récurrents

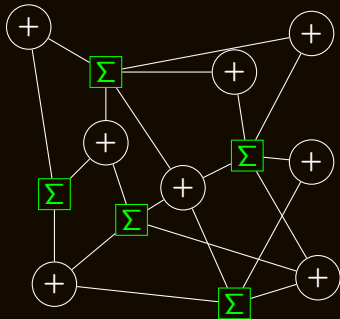
Diversité d'apprentissage



Capacité d'apprentissage



Décodeur LDPC



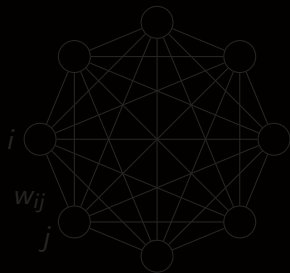
Mémoires associatives et modèle de Hopfield

Mémoire associative

Deux opérations :

- **Apprendre** des messages,
- **Retrouver** un messages appris effacé ou erroné.

L'état de l'art : le modèle de Hopfield



- Apprendre : M messages binaires d^m :

$$w_{ij} = \sum_{m=1, i \neq j}^M d_i^m d_j^m,$$

- Retrouver : itérer

$$v_i = \operatorname{sgn} \left(\sum_{j \neq i} v_j w_{ij} \right)$$

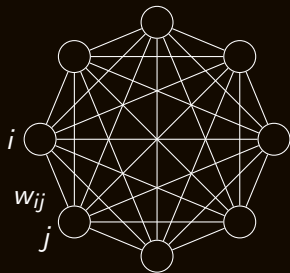
Mémoires associatives et modèle de Hopfield

Mémoire associative

Deux opérations :

- **Apprendre** des messages,
- **Retrouver** un messages appris effacé ou erroné.

L'état de l'art : le modèle de Hopfield



- Apprendre : M messages binaires \mathbf{d}^m :

$$w_{ij} = \sum_{m=1, i \neq j}^M d_i^m d_j^m,$$

- Retrouver : itérer

$$\forall i, v_i \leftarrow \operatorname{sgn}\left(\sum_{j \neq i} v_j w_{ij}\right).$$

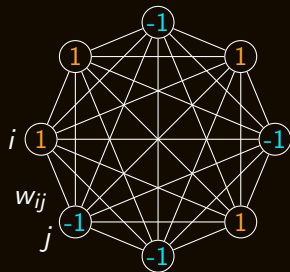
Mémoires associatives et modèle de Hopfield

Mémoire associative

Deux opérations :

- **Apprendre** des messages,
- **Retrouver** un messages appris effacé ou erroné.

L'état de l'art : le modèle de Hopfield



- Apprendre : M messages binaires \mathbf{d}^m :

$$w_{ij} = \sum_{m=1, i \neq j}^M d_i^m d_j^m,$$

- Retrouver : itérer
 $\forall i, v_i \leftarrow \text{sgn}\left(\sum_{j \neq i} v_j w_{ij}\right).$

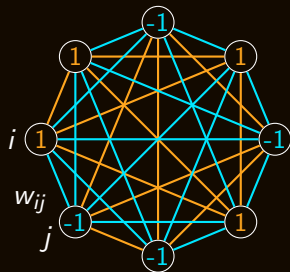
Mémoires associatives et modèle de Hopfield

Mémoire associative

Deux opérations :

- **Apprendre** des messages,
- **Retrouver** un messages appris effacé ou erroné.

L'état de l'art : le modèle de Hopfield



- Apprendre : M messages binaires \mathbf{d}^m :

$$w_{ij} = \sum_{m=1, i \neq j}^M d_i^m d_j^m,$$

- Retrouver : itérer
 $\forall i, v_i \leftarrow \text{sgn}\left(\sum_{j \neq i} v_j w_{ij}\right).$

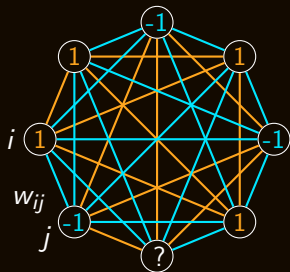
Mémoires associatives et modèle de Hopfield

Mémoire associative

Deux opérations :

- **Apprendre** des messages,
- **Retrouver** un messages appris effacé ou erroné.

L'état de l'art : le modèle de Hopfield



- Apprendre : M messages binaires \mathbf{d}^m :

$$w_{ij} = \sum_{m=1, i \neq j}^M d_i^m d_j^m,$$

- Retrouver : itérer

$$\forall i, v_i \leftarrow \operatorname{sgn}\left(\sum_{j \neq i} v_j w_{ij}\right).$$

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, $\begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$, $\begin{smallmatrix} \blacksquare & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, $\begin{matrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{matrix}$
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$, $\begin{matrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{matrix}$
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, $\begin{array}{|c|c|c|c|} \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \end{array}$
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$, $\begin{array}{|c|c|} \hline \cdot & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array}$
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, \blacksquare
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$ · \blacksquare
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, \blacksquare
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$ \blacksquare
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Réseaux de Hopfield (n neurones \longleftrightarrow)

- **Diversité** : $M = \frac{n}{2\log(n)}$, \leftrightarrow
- **Capacité** : $\frac{n^2}{2\log(n)}$, $\blacksquare = \blacksquare$
- Quantité d'information binaire requise : $\binom{n}{2} \log_2(M + 1)$, \blacksquare
- \Rightarrow **Efficacité** $\approx \frac{1}{\log(n)\log_2(M+1)}$ \blacksquare
- Connexions sensibles, longueur des messages = taille du réseau, les messages et leurs opposés sont appris en même temps...

Exemple avec $n = 790$: 

Exemple : le code économe

- Code des mots binaires à un seul "1"

Exemple : les mots binaires à un seul "1" sont codés par :

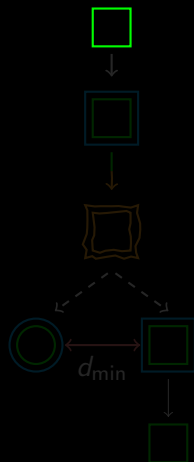
- Handicap : $d_{\min} = 2$

Exemple : les mots binaires à un seul "1" sont codés par :

- Mais pas de redondance et donc pas de correction

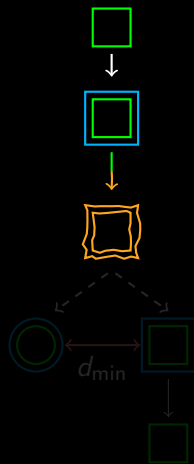
Exemple : les mots binaires à un seul "1" sont codés par :

Exemple : les mots binaires à un seul "1" sont codés par :



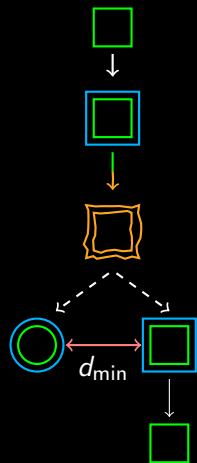
Exemple : le code économe

- Code des mots binaires à un seul "1"
- $d_{min} = 2$
- Handicap $= d_{min} - 1 = 1$
- Mais : code à longueur variable et non binaire
- $d_{min} = 2$ (si on ne considère que les mots de longueur 2)



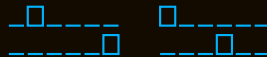
Exemple : le code économe

- Code des mots binaires à un seul "1"
- Les mots autorisés sont ceux qui ont un 1 à la position 1, 2, 4 ou 8.
- Handicap : $d_{\min} = 2$
- Le code est optimal (pour $n=8$)
- Mais il n'est pas systématique
- Exemple de la correction de la lettre "E" :



Exemple : le code économe

- Code des mots binaires à un seul "1" :



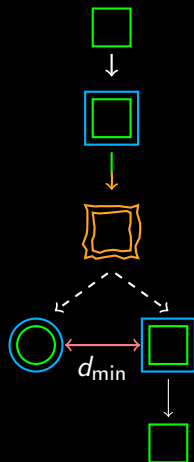
- Handicap : $d_{\min} = 2$:



- Mais facile à décoder et minimise l'énergie :



- Ces codes peuvent être associés à la manière des codes distribués. . .



Exemple : le code économe

- Code des mots binaires à un seul "1" :



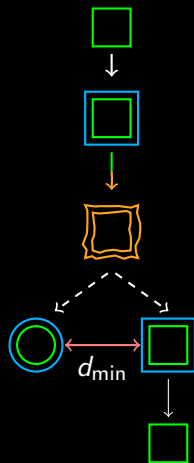
- Handicap : $d_{\min} = 2$:



- Mais facile à décoder et minimise l'énergie :



- Ces codes peuvent être associés à la manière des codes distribués...



Exemple : le code économe

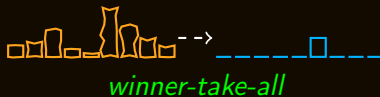
- Code des mots binaires à un seul "1" :



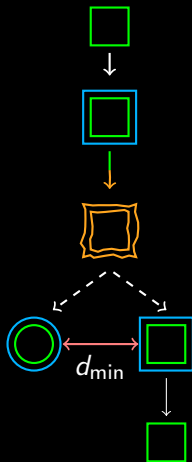
- Handicap : $d_{\min} = 2$:

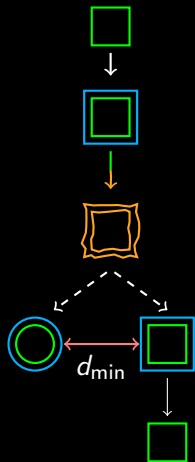


- Mais facile à décoder et minimise l'énergie :



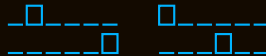
- Ces codes peuvent être associés à la manière des codes distribués. . .





Exemple : le code économe

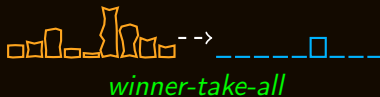
- Code des mots binaires à un seul "1" :



- Handicap : $d_{\min} = 2$:



- Mais facile à décoder et minimise l'énergie :



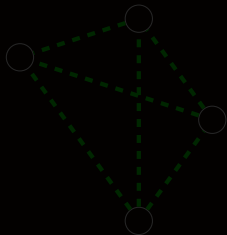
- Ces codes peuvent être associés à la manière des codes distribués...

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



2 noeuds différents
 $\Rightarrow d_{\min} = 2$

Codes sur cliques de taille $c \ll n$

$$n \cdot d_{\min} = 2(c-1) \approx 2c$$

$$n = F \cdot d_{\min} \approx 2F$$

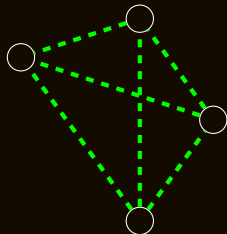
→ On peut donc encoder ≈ 2 bits par clique de taille constante

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

$$n \cdot d_{\min} = 2(c-1) \approx 2c$$

$$n = F \cdot d_{\min} \approx 2F$$

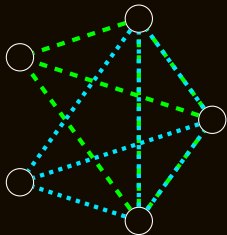
$$n = 2F \approx 2 \cdot \frac{1}{2} \log_2 \frac{1}{\epsilon} = \log_2 \frac{1}{\epsilon}$$

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différent
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

$$n \cdot d_{\min} = 2(c-1) \cdot 2c$$

$$n = \frac{2c(c-1)}{d_{\min}}$$

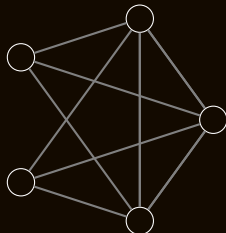
$$n = \frac{2c^2}{d_{\min}}$$

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

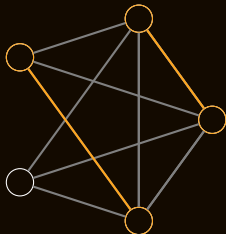
- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

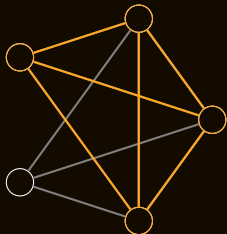
- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2} \binom{c}{2}^{-1}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

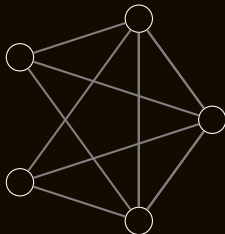
- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2} \binom{c}{2}^{-1}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

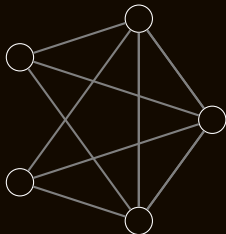
- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2} \binom{c}{2}^{-1}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

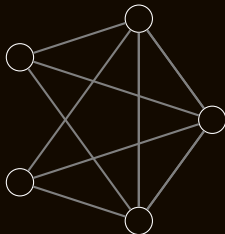
- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2} \binom{c}{2}^{-1}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...et gratuit!

Des codes sur des cliques de taille constante

Exemple : mots de codes = cliques à 4 noeuds

Clique

Ensemble de noeuds d'un graphe entièrement interconnectés.



Symboles = arêtes

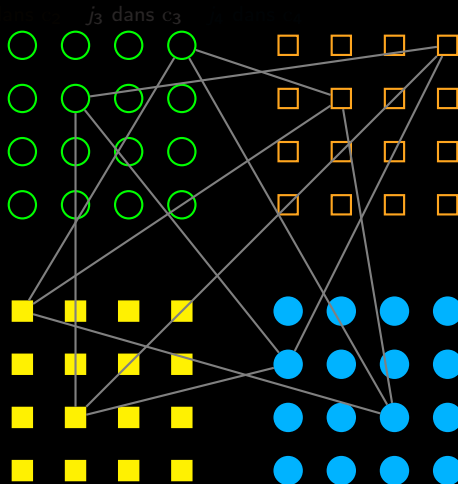
2 noeuds différents
 $\Rightarrow d_{\min} = 6$ arêtes

Codes sur cliques de taille $c \ll n$

- $d_{\min} = 2(c - 1) \approx 2c$, rendement $r \approx \frac{c}{2} \binom{c}{2}^{-1}$
- $\Rightarrow F = rd_{\min} \approx 2$,
- Les cliques sont les mots de code d'un code correcteur d'erreurs très intéressant...**et gratuit!**

Notre modèle : apprentissage

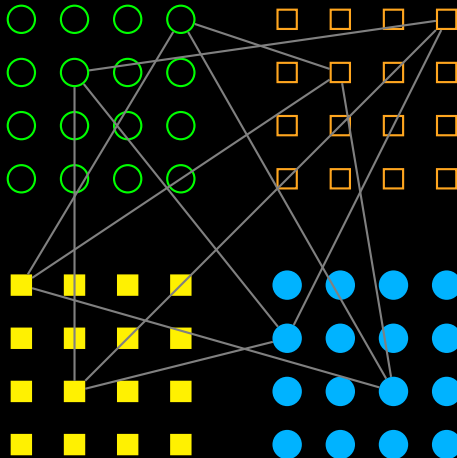
- Exemple : $c = 4$ clusters de $l = 16$ neurones chacun,
- $\underbrace{1000}_{= 8} \underbrace{0011}_{= 3} \underbrace{0010}_{= 2} \underbrace{1001}_{= 9}$,



Notre modèle : apprentissage

- Exemple : $c = 4$ clusters de $l = 16$ neurones chacun,

- $\underbrace{1000}_{j_1 \text{ dans } c_1} = 8 \underbrace{0011}_{j_2 \text{ dans } c_2} = 3 \underbrace{0010}_{j_3 \text{ dans } c_3} = 2 \underbrace{1001}_{j_4 \text{ dans } c_4} = 9,$



Notre modèle : apprentissage

- Exemple : $c = 4$ clusters de $l = 16$ neurones chacun,

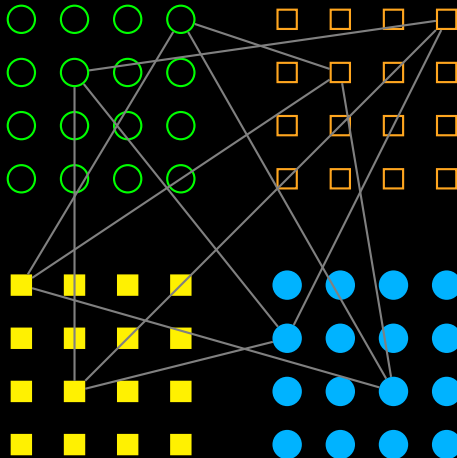
- $\underbrace{1000}_{j_1 \text{ dans } c_1} = 8$ $\underbrace{0011}_{j_2 \text{ dans } c_2} = 3$ $\underbrace{0010}_{j_3 \text{ dans } c_3} = 2$ $\underbrace{1001}_{j_4 \text{ dans } c_4} = 9$,

j_1 dans c_1

j_2 dans c_2

j_3 dans c_3

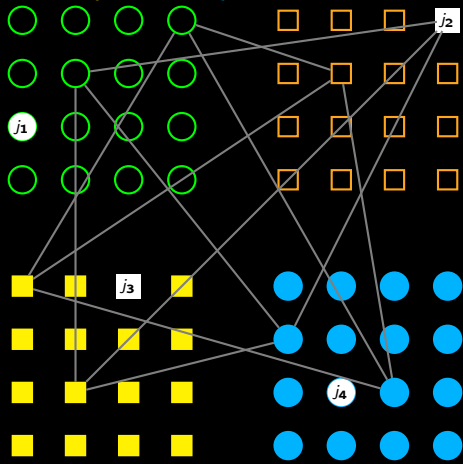
j_4 dans c_4



Notre modèle : apprentissage

- Exemple : $c = 4$ clusters de $l = 16$ neurones chacun,

- $\underbrace{1000}_{j_1 \text{ dans } c_1} = 8$ $\underbrace{0011}_{j_2 \text{ dans } c_2} = 3$ $\underbrace{0010}_{j_3 \text{ dans } c_3} = 2$ $\underbrace{1001}_{j_4 \text{ dans } c_4} = 9$,



Notre modèle : apprentissage

- Exemple : $c = 4$ clusters de $l = 16$ neurones chacun,

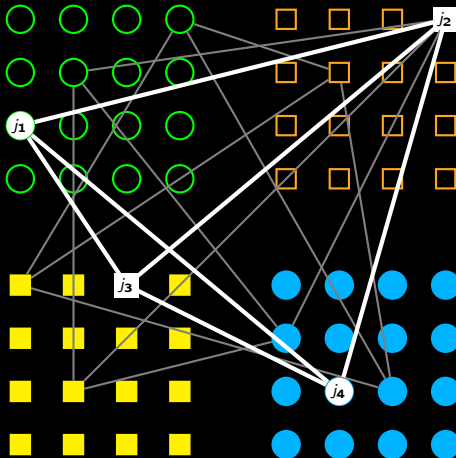
- $\underbrace{1000}_{j_1 \text{ dans } c_1} = 8$ $\underbrace{0011}_{j_2 \text{ dans } c_2} = 3$ $\underbrace{0010}_{j_3 \text{ dans } c_3} = 2$ $\underbrace{1001}_{j_4 \text{ dans } c_4} = 9$,

j_1 dans c_1

j_2 dans c_2

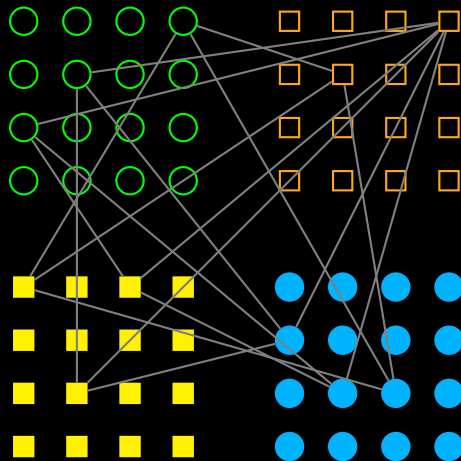
j_3 dans c_3

j_4 dans c_4



Notre modèle : remémoration

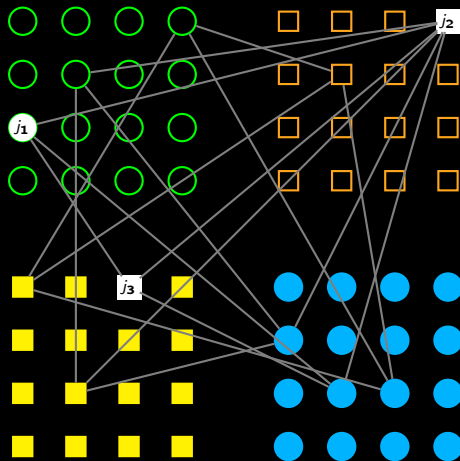
$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $????,$



- Correspondance locale,
- Décodage global :
somme,
- Décodage local :
winner-take-all,
- Itération possible des
deux décodages.

Notre modèle : remémoration

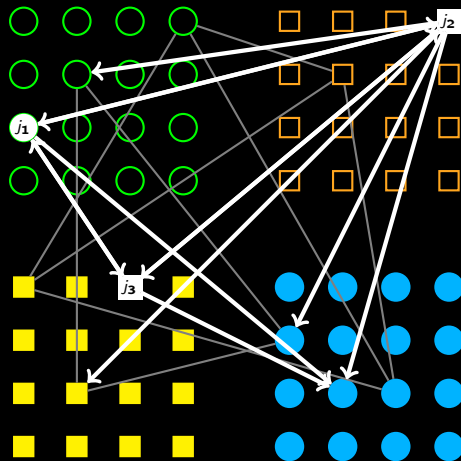
$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $????,$



- Correspondance locale,
- Décodage global :
somme,
- Décodage local :
winner-take-all,
- Itération possible des
deux décodages.

Notre modèle : remémoration

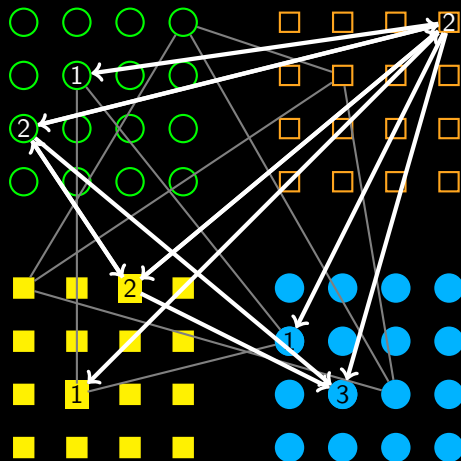
$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $????,$



- Correspondance locale,
- Décodage global : somme,
- Décodage local : *winner-take-all*,
- Itération possible des deux décodages.

Notre modèle : remémoration

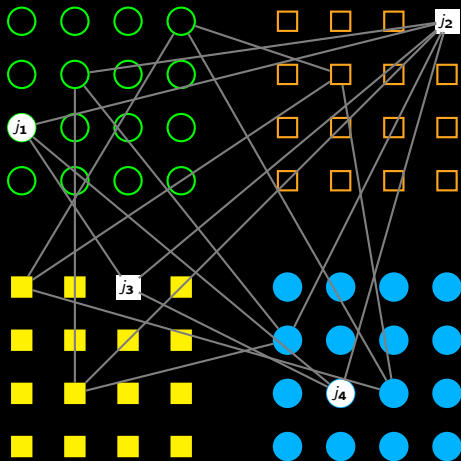
$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $????,$



- Correspondance locale,
- Décodage global :
somme,
- Décodage local :
winner-take-all,
- Itération possible des
deux décodages.

Notre modèle : remémoration

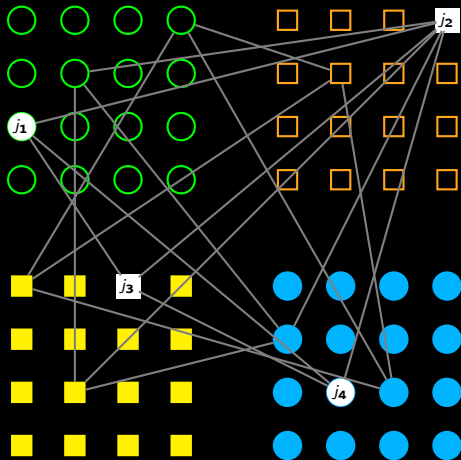
$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $\underbrace{1001}_{j_4 \text{ dans } c_4}$,



- Correspondance locale,
- Décodage global :
somme,
- Décodage local :
winner-take-all,
- Itération possible des
deux décodages.

Notre modèle : remémoration

$\underbrace{1000}_{j_1 \text{ dans } c_1}$ $\underbrace{0011}_{j_2 \text{ dans } c_2}$ $\underbrace{0010}_{j_3 \text{ dans } c_3}$ $\underbrace{1001}_{j_4 \text{ dans } c_4}$,



- Correspondance locale,
- Décodage global :
somme,
- Décodage local :
winner-take-all,
- Itération possible des
deux décodages.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre le nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{I^2}\right)^M$.

Courbes

Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non,
- $d = f(I, M)$ indépendant de c ,
- $d \approx \frac{M}{I^2}$ pour $M \ll I^2$.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre le nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{l^2}\right)^M$.

Courbes

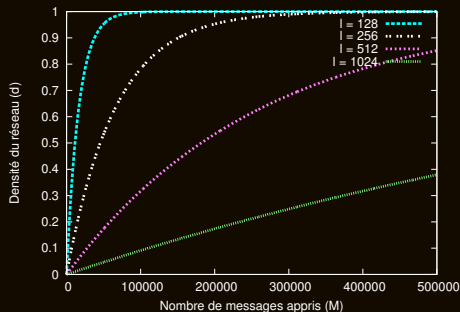
Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non,
- $d = f(1, M)$ indépendant de c ,
- $d \approx \frac{M}{l^2}$ pour $M \ll l^2$.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre les nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{l^2}\right)^M$.

Courbes



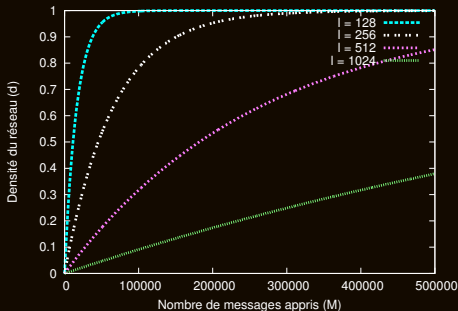
Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non, $d = 1 - \left(1 - \frac{1}{l^2}\right)^M$ indépendant de l pour $M \gg l^2$, $d \approx \frac{M}{l^2}$ pour $M \ll l^2$.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre le nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{l^2}\right)^M$.

Courbes



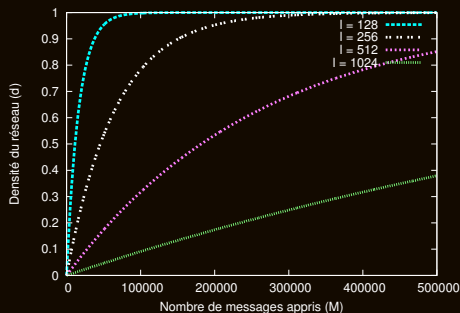
Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non,
- $d = f(l, M)$, indépendant de c ,
- $d \approx \frac{M}{l^2}$, pour $M \ll l^2$.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre les nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{l^2}\right)^M$.

Courbes



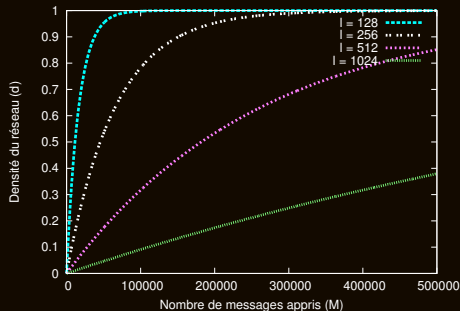
Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non,
- $d = f(l, M)$, indépendant de c ,
- $d \approx \frac{M}{l^2}$, pour $M \ll l^2$.

Introduction d'un nouveau paramètre

- La densité d est le rapport entre le nombre de connexions utilisées et disponibles au total,
- Si messages i.i.d. : $d \approx 1 - \left(1 - \frac{1}{l^2}\right)^M$.

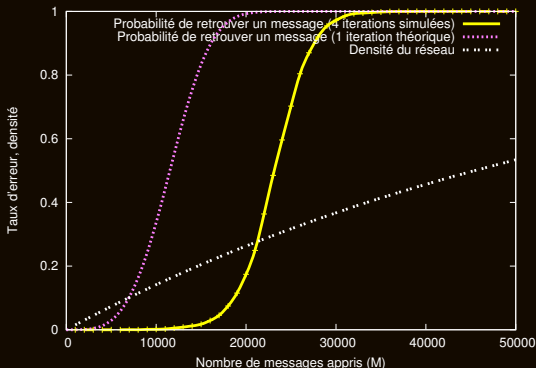
Courbes



Remarques

- $d = 1$: plus de distinction possible entre messages appris ou non,
- $d = f(l, M)$, indépendant de c ,
- $d \approx \frac{M}{l^2}$, pour $M \ll l^2$.

Mémoire associative



$c = 8$ clusters de $l = 256$ neurones (\sim messages de longueur 64 bits),
Probabilité d'erreur lors de la remémoration de messages effacés à moitié.

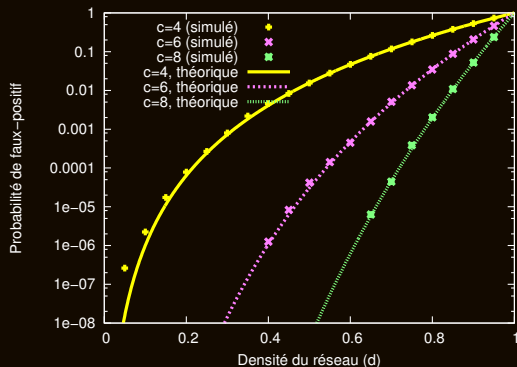
Réseau de Hopfield ($n = 790$)



Notre réseau



Classification



Probabilité de faux-positif pour divers nombres de clusters c et $l = 512$ neurones par cluster.

Réseau de Hopfield ($n = 740$)

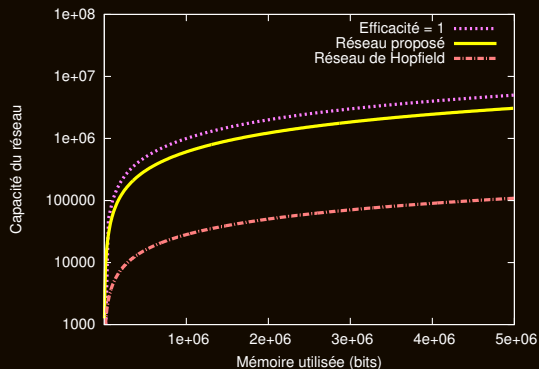


Notre réseau



Comparaison des capacités de notre réseau et de celui de Hopfield

Performances (3/3)



Comparaison des capacités du modèle de Hopfield et de nos réseaux (dans le cas de mémoire associative) pour une même quantité d'information utilisée.

Un mot sur la plausibilité

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

- Nécessité de posséder une information parfaite (bien que partielle)
- Les messages ne doivent pas être corrélés
- Les clusters doivent être grands et peu nombreux
- Taille fixe des messages appris
- Utilisation systématique de tous les clusters

Un mot sur la plausibilité

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

- Nécessité de posséder une information parfaite (bien que partielle),
- Les messages ne doivent pas être corrélés,
- Les clusters doivent être grands et peu nombreux,
- Taille fixe des messages appris,
- Utilisation systématique de tous les clusters.

Un mot sur la plausibilité

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

- Nécessité de posséder une information parfaite (bien que partielle),
- Les messages ne doivent pas être corrélés,
- Les clusters doivent être grands et peu nombreux,
- Taille fixe des messages appris,
- Utilisation systématique de tous les clusters.

Un mot sur la plausibilité

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

- Nécessité de posséder une information parfaite (bien que partielle),
- Les messages ne doivent pas être corrélés,
- Les clusters doivent être grands et peu nombreux,
- Taille fixe des messages appris,
- Utilisation systématique de tous les clusters.

Un mot sur la plausibilité

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

- Nécessité de posséder une information parfaite (bien que partielle),
- Les messages ne doivent pas être corrélés,
- Les clusters doivent être grands et peu nombreux,
- Taille fixe des messages appris,
- Utilisation systématique de tous les clusters.

Analogies

Notre réseau		Littérature neuroscientifique
Cliques de neurones	↔	Cliques neurales
Décodage local	↔	<i>Winner-take-all</i>
Clusters	↔	Colonnes néocorticales
Code économe	↔	Neurones spécifiques

Limites

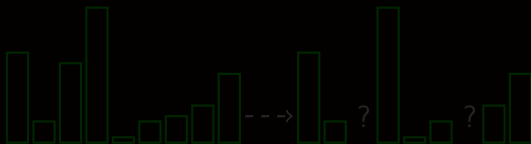
- Nécessité de posséder une information parfaite (bien que partielle),
- Les messages ne doivent pas être corrélés,
- Les clusters doivent être grands et peu nombreux,
- Taille fixe des messages appris,
- Utilisation systématique de tous les clusters.

Messages flous

Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

Modèle du bruit



Décodage *soft*



Messages flous

Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

Modèle du bruit



Décodage *soft*



Messages flous

Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

Modèle du bruit



Décodage *soft*



Messages flous

Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

Modèle du bruit



Décodage *soft*



Messages flous

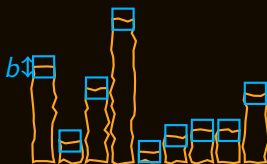
Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

Modèle du bruit



Décodage *soft*

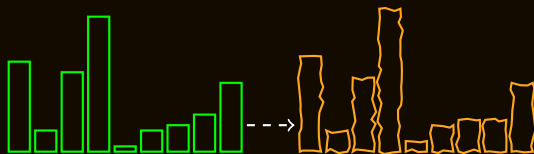


Messages flous

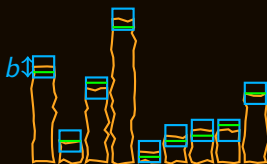
Problème

Les messages partiels doivent comporter des valeurs correctes et précises.

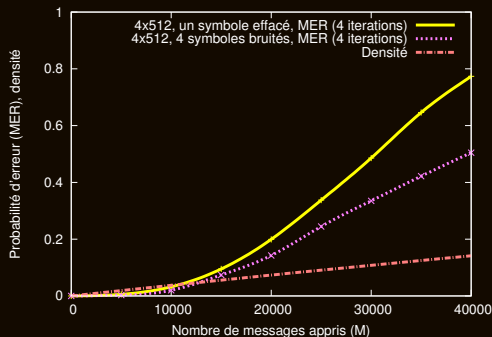
Modèle du bruit



Décodage *soft*



Simulations

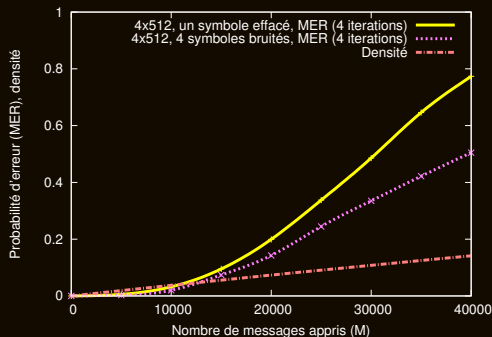


Comparaison des performances quand les messages sont partiellement effacés et lorsqu'ils sont bruités ($b = 5$).

Raison des bonnes performances

- Effacement : \searrow cliques concurrentes ($\approx l$) \nearrow probabilité ($\approx d^{c-1}$),
- Bruit : \nearrow cliques concurrentes ($\approx b^c$) \searrow probabilité ($\approx d^{\frac{c(c-1)}{2}}$).

Simulations



Comparaison des performances quand les messages sont partiellement effacés et lorsqu'ils sont bruités ($b = 5$).

Raison des bonnes performances

- Effacement : \searrow cliques concurrentes ($\approx l$) \nearrow probabilité ($\approx d^{c-1}$),
- Bruit : \nearrow cliques concurrentes ($\approx b^c$) \searrow probabilité ($\approx d^{\frac{c(c-1)}{2}}$).

Problème

La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in* ?
 - Un effet venant du réseau

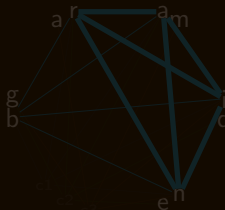


Problème

La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in*?
 - Un effet venant du réseau :



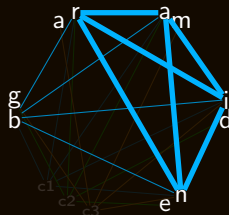
Problème

La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in*?
 - Un effet venant du réseau :

brain



Problème

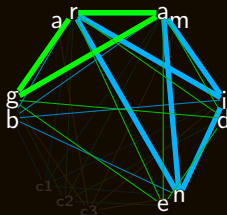
La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in*?
 - Un effet venant du réseau :

brain

grade



Problème

La corrélation des messages accentue l'apparition de fausses cliques.

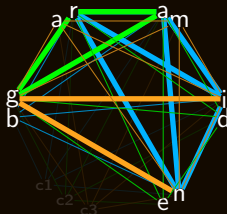
Combattre la corrélation par l'ajout de redondance aléatoire

- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in*?
 - Un effet venant du réseau :

brain

grade

gamin



Problème

La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

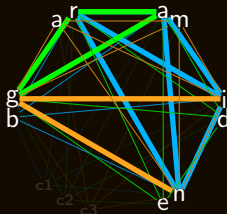
- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in* ?
 - Un effet venant du réseau :

brain

grade

gamin

grain



Problème

La corrélation des messages accentue l'apparition de fausses cliques.

Combattre la corrélation par l'ajout de redondance aléatoire

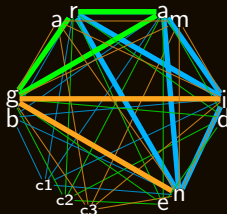
- Il y a deux effets de la corrélation :
 - Un effet incontournable : *malin* et *matin* appris \rightarrow *ma * in* ?
 - Un effet venant du réseau :

brain +c1

grade +c2

gamin +c3

grain +c?



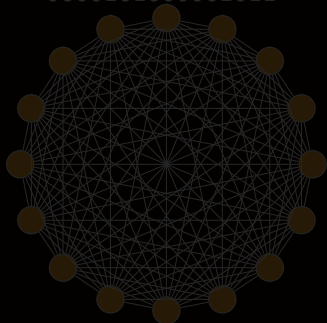
Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration

0000101000001011



Idée

- Messages plus courts,
- Clusters et codes économiques
- Réseau creux
- Messages parcimonieux

Solutions

- *Winner-take-all* global
- Synchronisation temporelle

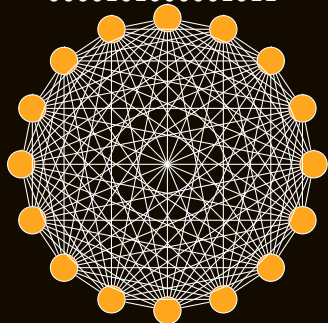
Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration

0000101000001011



Idée

- Messages plus courts,
- Clusters et codes économiques
- Réseau global

Solutions

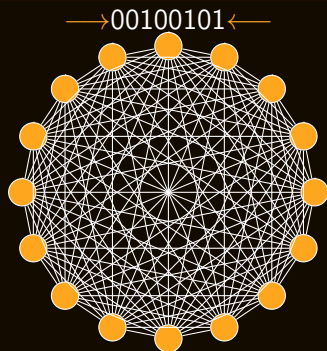
- Winner-take-all global
- Synchronisation temporelle

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

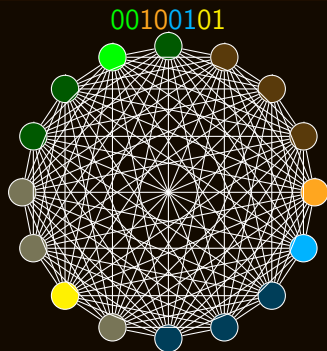
Dimensionnalité globale
Synchronisation temporelle

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

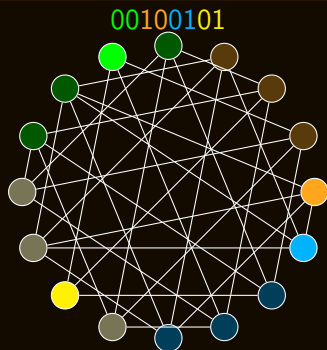
Dimensionnalité globale
Synchronisation temporelle

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

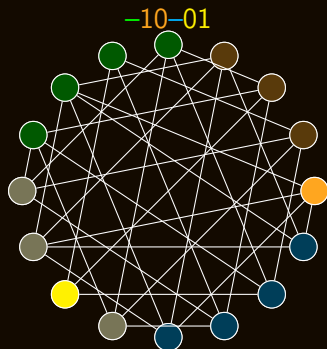
• Winner-take-all global
• Synchronisation temporelle

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

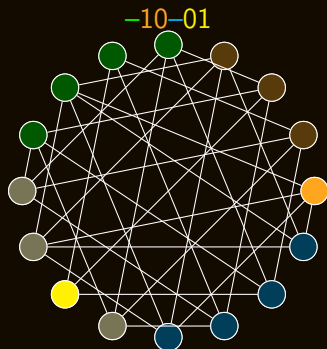
Winner-take-all global
Synchronisation temporelle

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

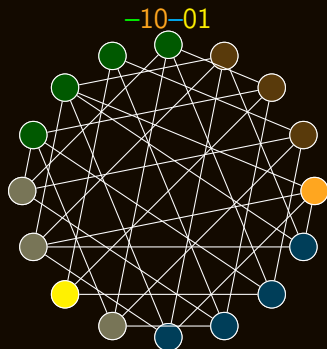
- *Winner-take-all* global,
- Synchronisation temporelle.

Vers un quatrième niveau de parcimonie

Problèmes

- Les cliques doivent rester peu nombreuses et grandes,
- Les messages appris sont tous de même taille.

Illustration



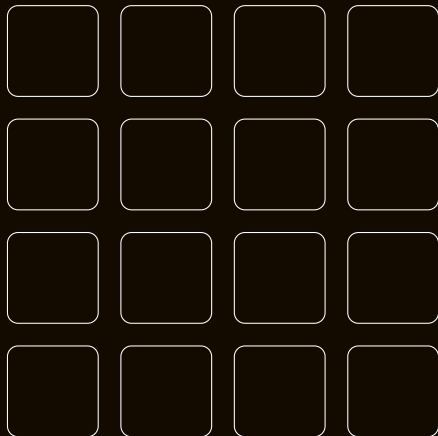
Idée

- 1 Messages plus courts,
- 2 Clusters et codes économes,
- 3 Réseau creux,
- 4 Messages parcimonieux.

Solutions

- *Winner-take-all* global,
- Synchronisation temporelle.

Illustration



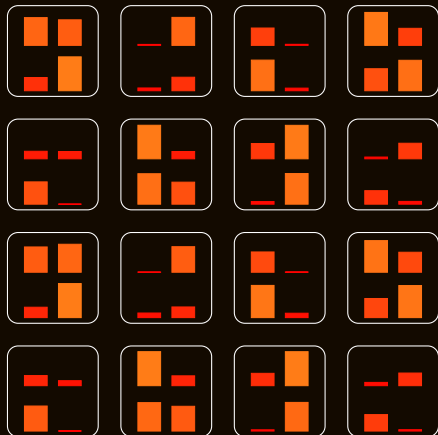
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité
- Taille des messages appris variable

Illustration



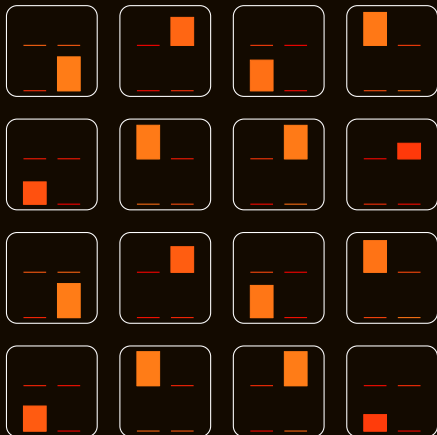
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité
- Taille des messages appris variable

Illustration



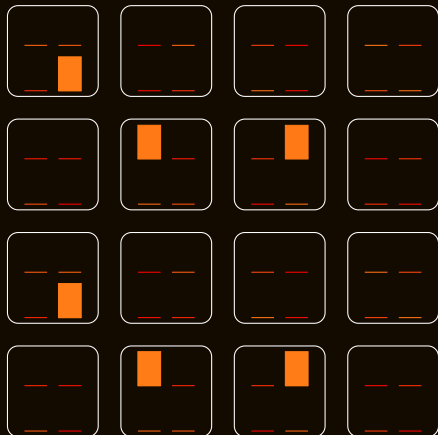
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité
- Taille des messages après variable

Illustration



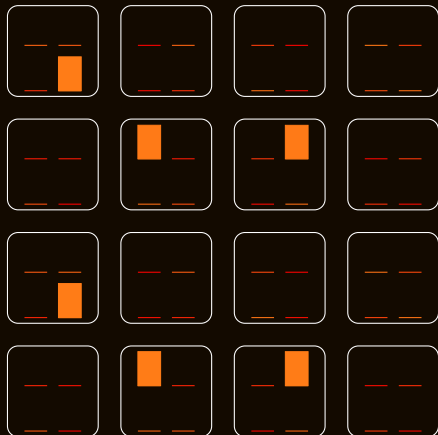
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité
- Taille des messages après variable

Illustration



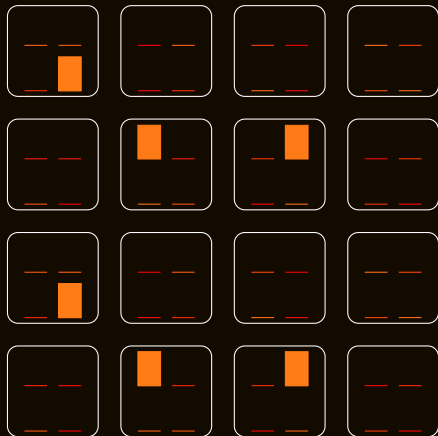
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité $\propto c^2$,
- Taille des messages appris variable.

Illustration



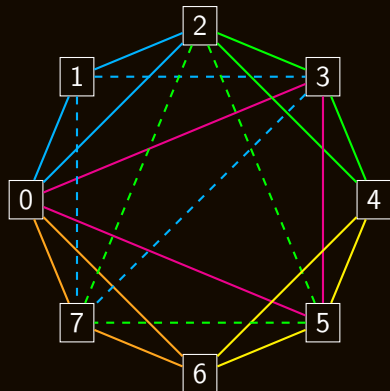
Idée

- Après passage de messages globaux. . .
- Après sélection des gagnants locaux. . .
- Sélection des gagnants globaux.

Intérêts

- Diversité $\propto c^2$,
- Taille des messages appris variable.

Illustration



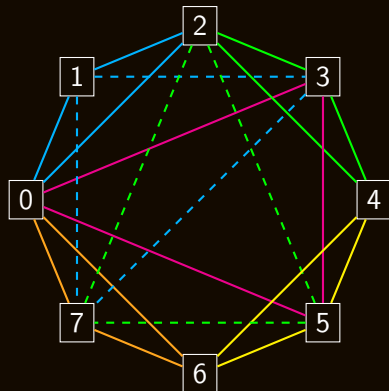
Idée

- Sous-réseaux indépendants,
- Localement : agrégation et *winner-take-all*,
- Globalement : détection de coïncidences.

Intérêt

Rapprochement possible vers la modélisation de pathologies liées à des défauts de synchronisation.

Illustration



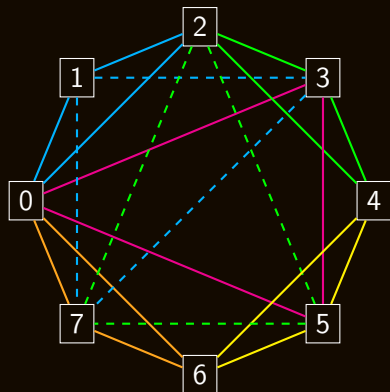
Idée

- Sous-réseaux indépendants,
- Localement : agrégation et *winner-take-all*,
- Globalement : détection de coïncidences.

Intérêt

Rapprochement possible vers la modélisation de pathologies liées à des défauts de synchronisation.

Illustration



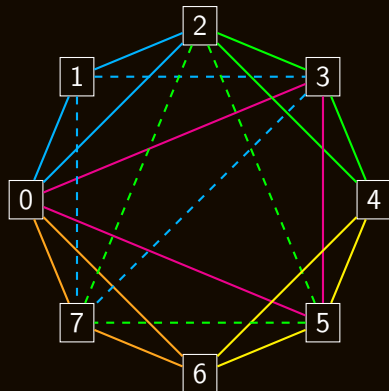
Idée

- Sous-réseaux indépendants,
- Localement : agrégation et *winner-take-all*,
- Globalement : détection de coïncidences.

Intérêt

Rapprochement possible vers la modélisation de pathologies liées à des défauts de synchronisation.

Illustration



Idée

- Sous-réseaux indépendants,
- Localement : agrégation et *winner-take-all*,
- Globalement : détection de coïncidences.

Intérêt

Rapprochement possible vers la modélisation de pathologies liées à des défauts de synchronisation.

Conclusion

Démarche

élaboration d'une mémoire associative →

↑
parcimonie

↑
codes distribués

architecture néocorticale

Résultats

- Capacités quasi-optimales, diversités considérables,

et architectures réalisables

Conclusion

Démarche

élaboration d'une mémoire associative →

↑
parcimonie

↑
codes distribués

architecture néocorticale

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation...
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Conclusion

Démarche

élaboration d'une mémoire associative →

architecture néocorticale

↑
parcimonie

↑
codes distribués

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation...
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Conclusion

Démarche

élaboration d'une mémoire associative →

architecture néocorticale

↑
parcimonie

↑
codes distribués

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation...
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Conclusion

Démarche

élaboration d'une mémoire associative →

architecture néocorticale

↑
parcimonie

↑
codes distribués

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation. . . ,
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Conclusion

Démarche

élaboration d'une mémoire associative →

architecture néocorticale

↑
parcimonie

↑
codes distribués

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation... ,
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Conclusion

Démarche

élaboration d'une mémoire associative →

architecture néocorticale

↑
parcimonie

↑
codes distribués

Résultats

- Capacités quasi-optimales, diversités considérables,
- Architecture massivement parallèle,
- Analogies avec les architectures et fonctionnements neurobiologiques,
- Robustesse, résilience, synchronisation. . . ,
- Degrés de liberté : inhibitions, temps, poids,
- Aucun compromis nécessaire entre performances et plausibilité.

Journal

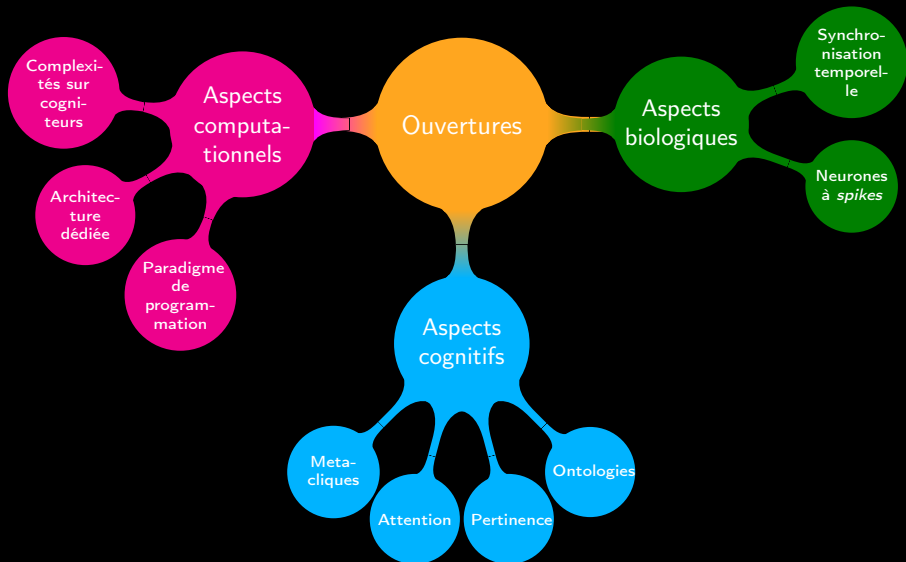
- Un article dans *IEEE Transactions on neural networks*.

Conférences

- Une communication dans *Proc. of 6th Int'l Symposium on Turbo Codes and Iterative Information Processing*,
- Une communication dans *Proc. of IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain*.

Brevets

- Un premier brevet déposé en 2010 : réseau présenté,
- Un second en cours de dépôt : apprentissage de séquences.



Merci pour votre attention. Je suis à votre disposition pour toutes questions complémentaires.

“Décodeur” néocortical

Décodeur LDPC

