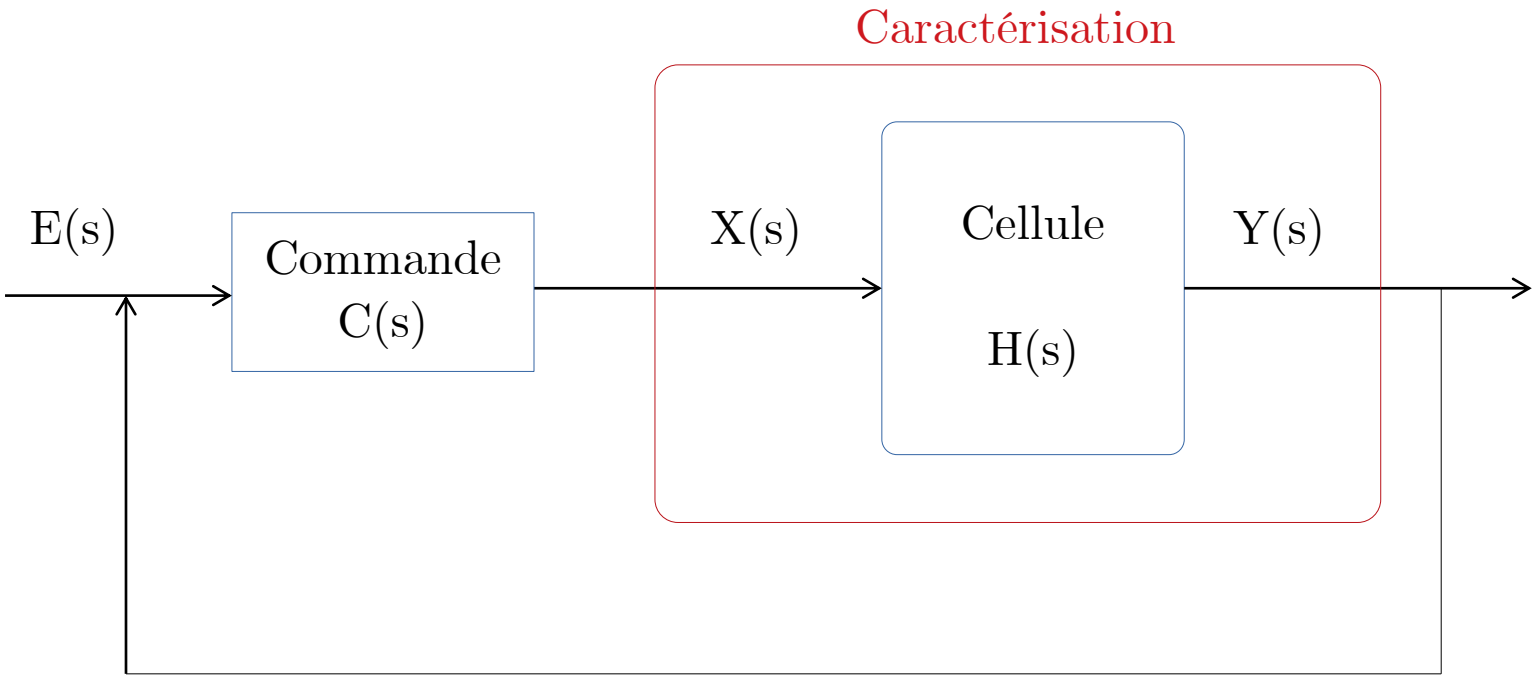
Plan de la présentation

Partie 1, Thèse : Caractérisation de la voie de signalisation NODAL lors de la différenciation des cellules souches embryonnaires.

Compétence : Physique, Biologie, Théorie du signal, Analyse d'image, Électronique et Microfluidique.

Partie 2, Reconversion IA : Un réseau de neurones artificiels fonctionnalisé par l'évolution.

Principe : Une cellule comme système

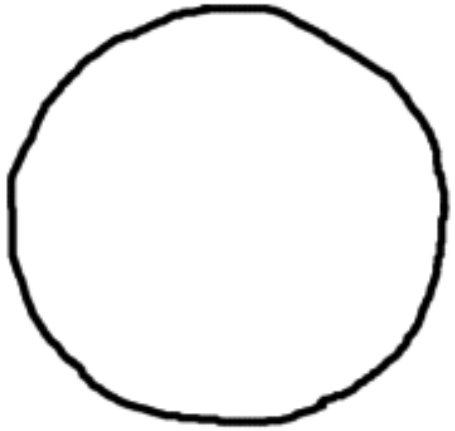


E(s) : Consigne (type cellulaire)
H(s) : Fonction de Transfert
X(s) : Entrée (morphogène)
Y(s) : Sortie (marqueur de différenciation)

Idée générale : Comment contrôler la *différenciation* des cellules souches embryonnaires ?

→ Besoin de caractériser le système biologique

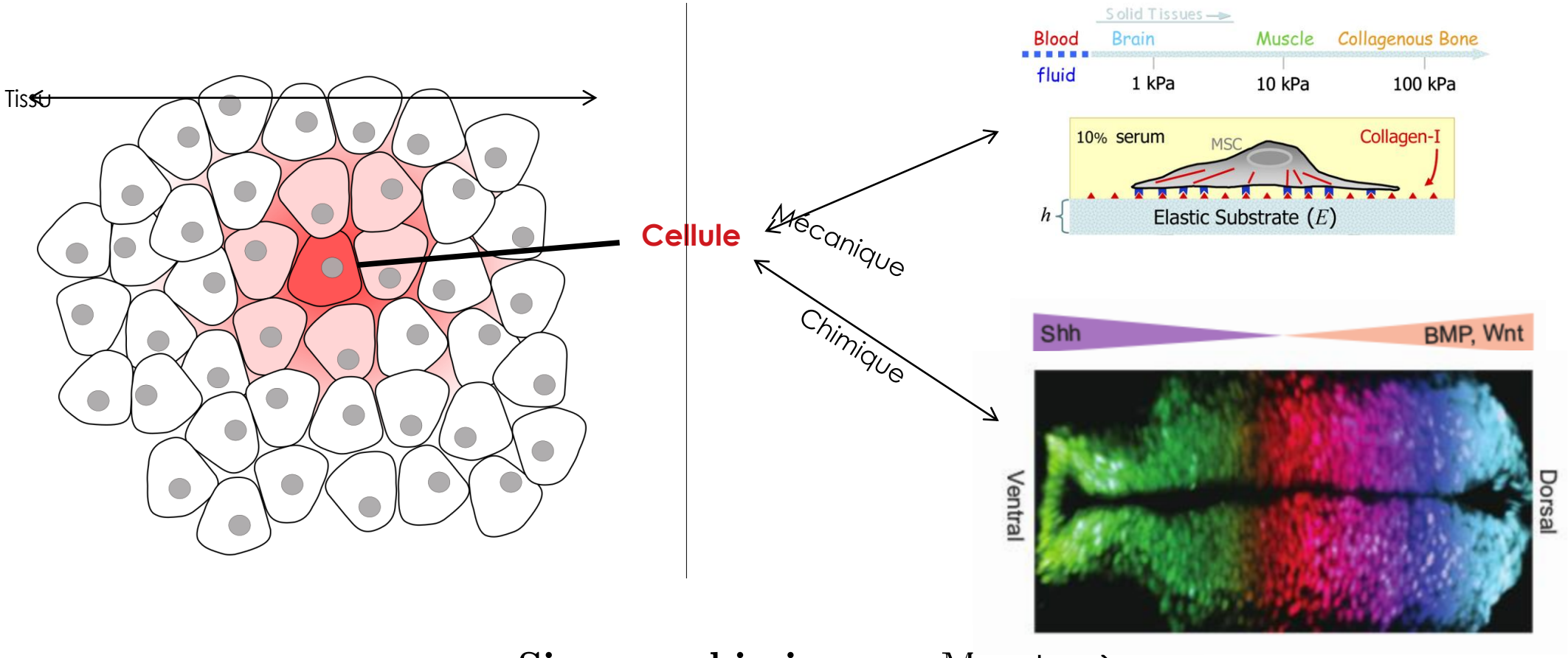
Qu'est ce que la différenciation cellulaire ?



Cellules souches embryonnaires → Pluripotence

Différenciation → Processus donnant une identité spécifique

Comment les cellules se différencient ?



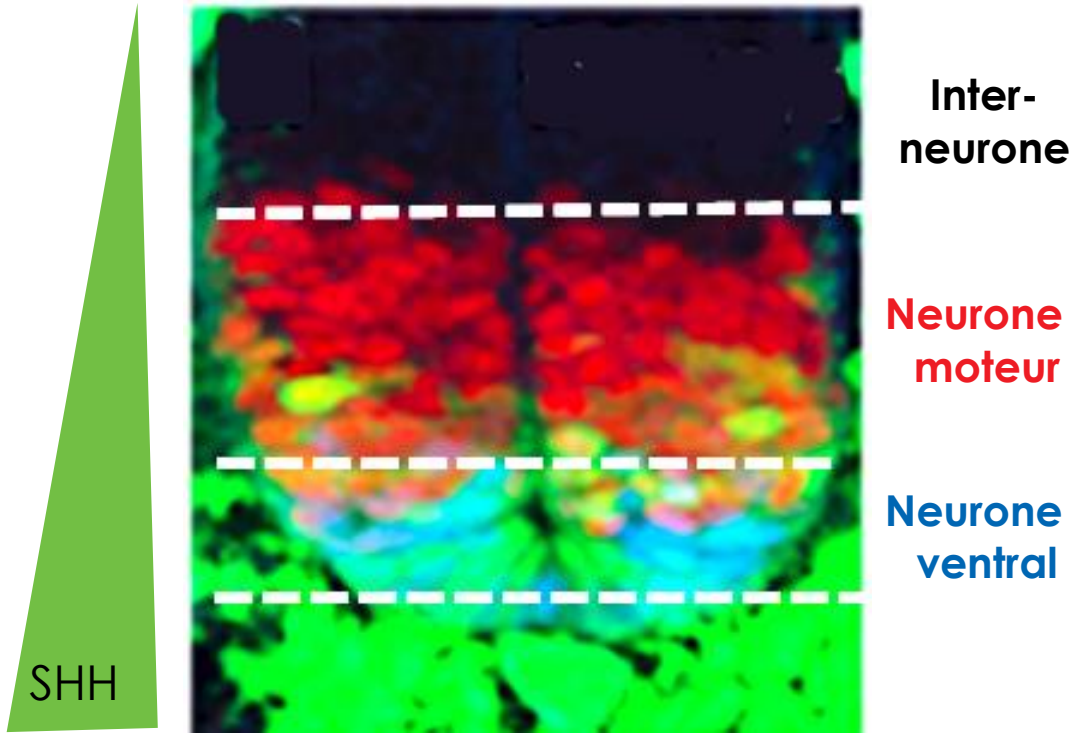
Signaux chimiques = Morphogènes

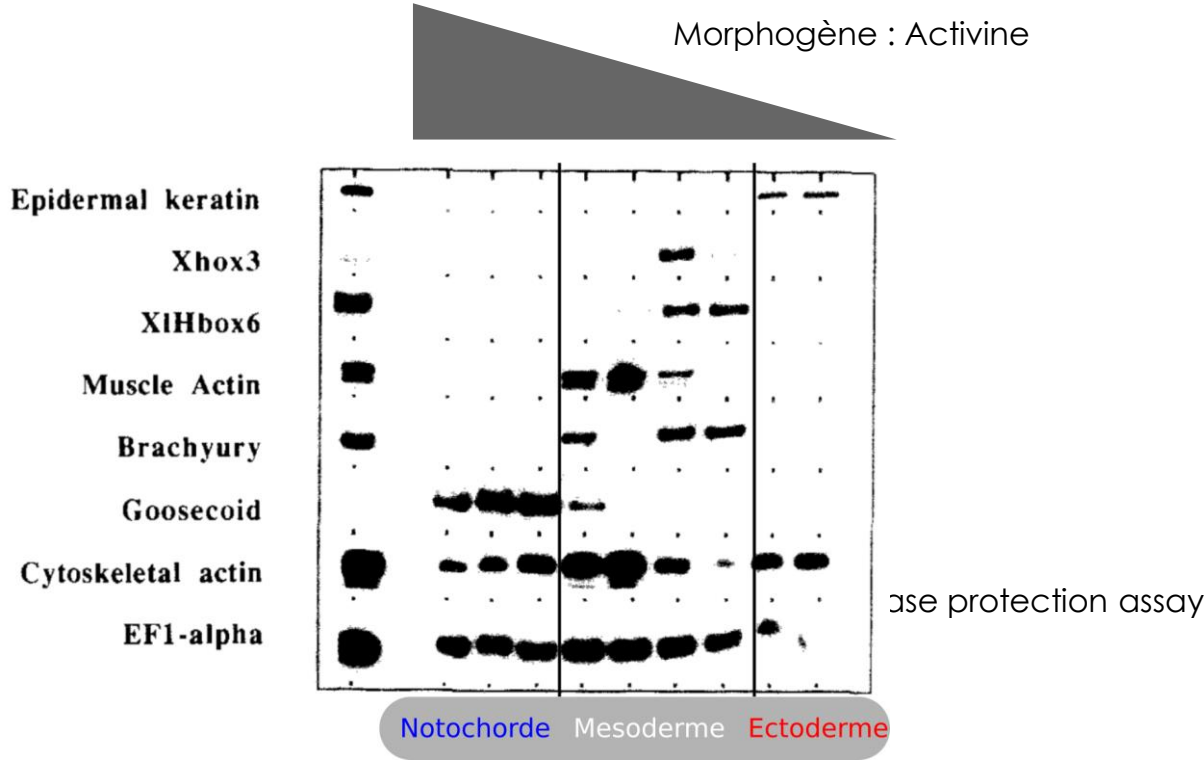
Morphogène : Sécrétés, diffusible et contrôle la différenciation

Source : Adam J. Engler, 2006 ; James Briscoe, 2015

Les tissus s'organisent spatialement

Tube neural de souris (stage 12)





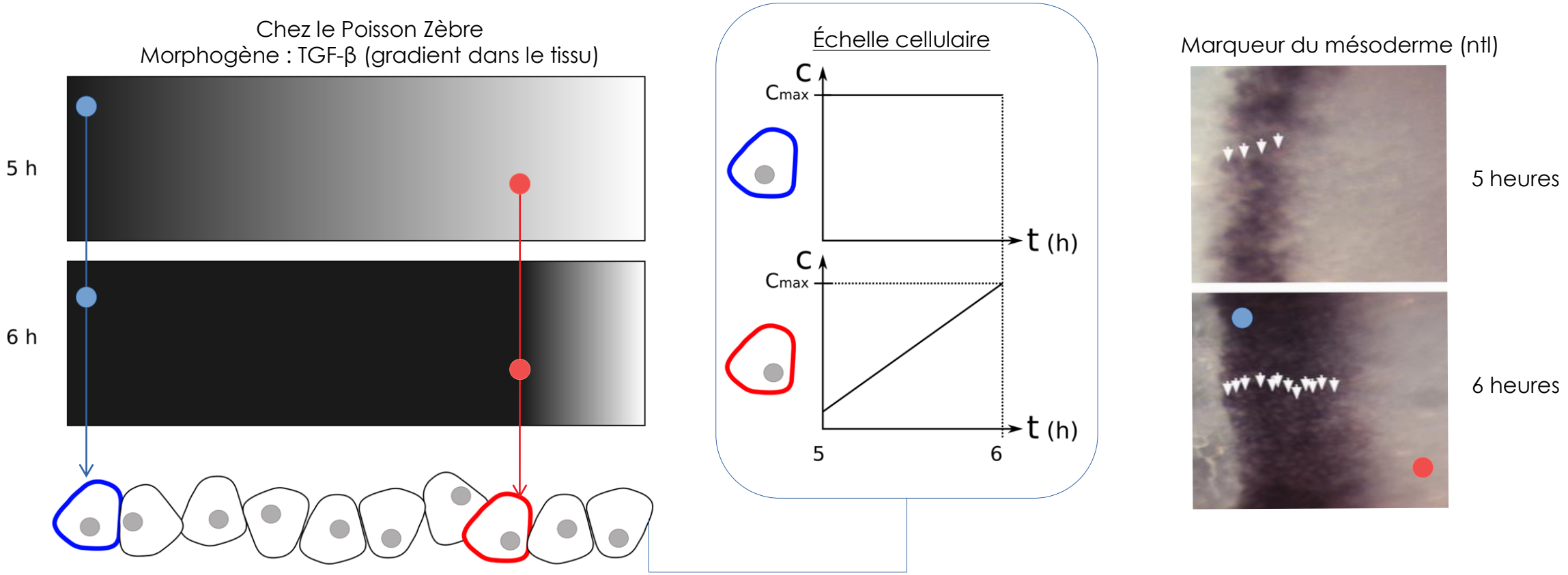
ase protection assay

1 Morphogène → 3 types cellulaires

In vitro, une cellule interprète sa position par rapport à deux seuils de concentration

Source : Balaskas et al. 2012, Green et al. 1992

In vivo, la concentration est dynamique

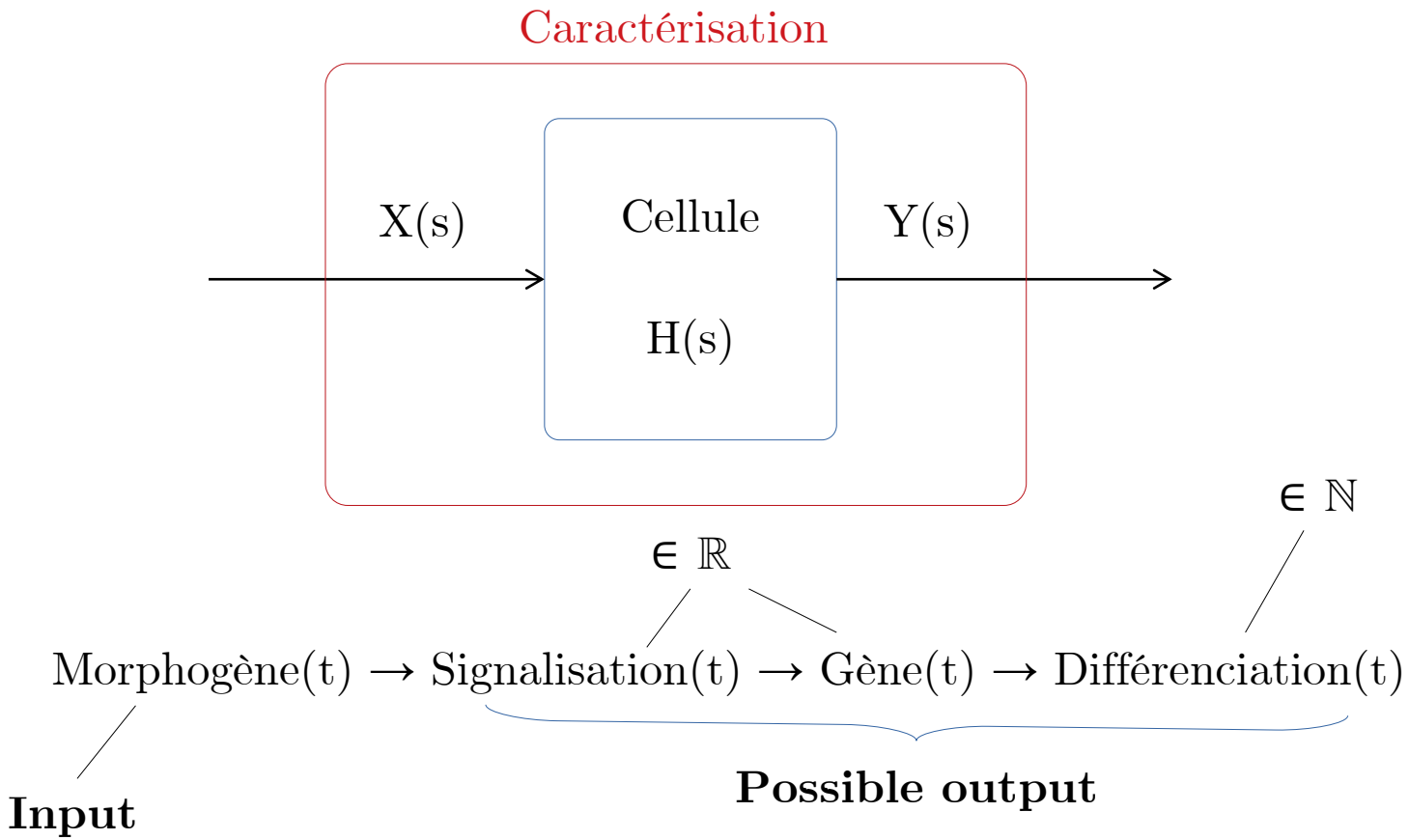


La concentration de morphogène augmente avec le **temps** durant le développement

→ Le gradient de morphogène **n'est pas statique**

Source : Harvey & Smith 2009

Comment mesurer l'effet d'un signal ?

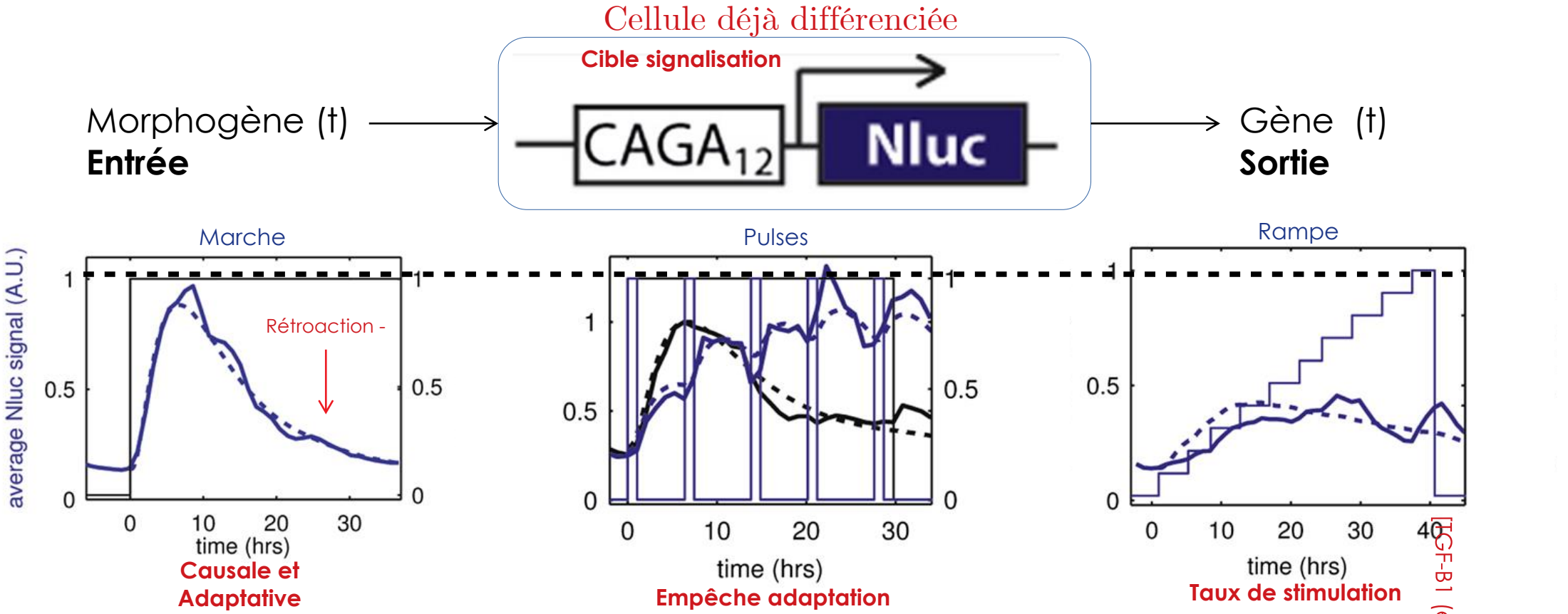


Besoin d'une mesure de l'effet d'une stimulation en **temps réel**

\+ Mesure rapide et appartenant à $\mathbb{R}$

Les fondements de nos travaux

(µfluidique et rapporteur fluorescent)

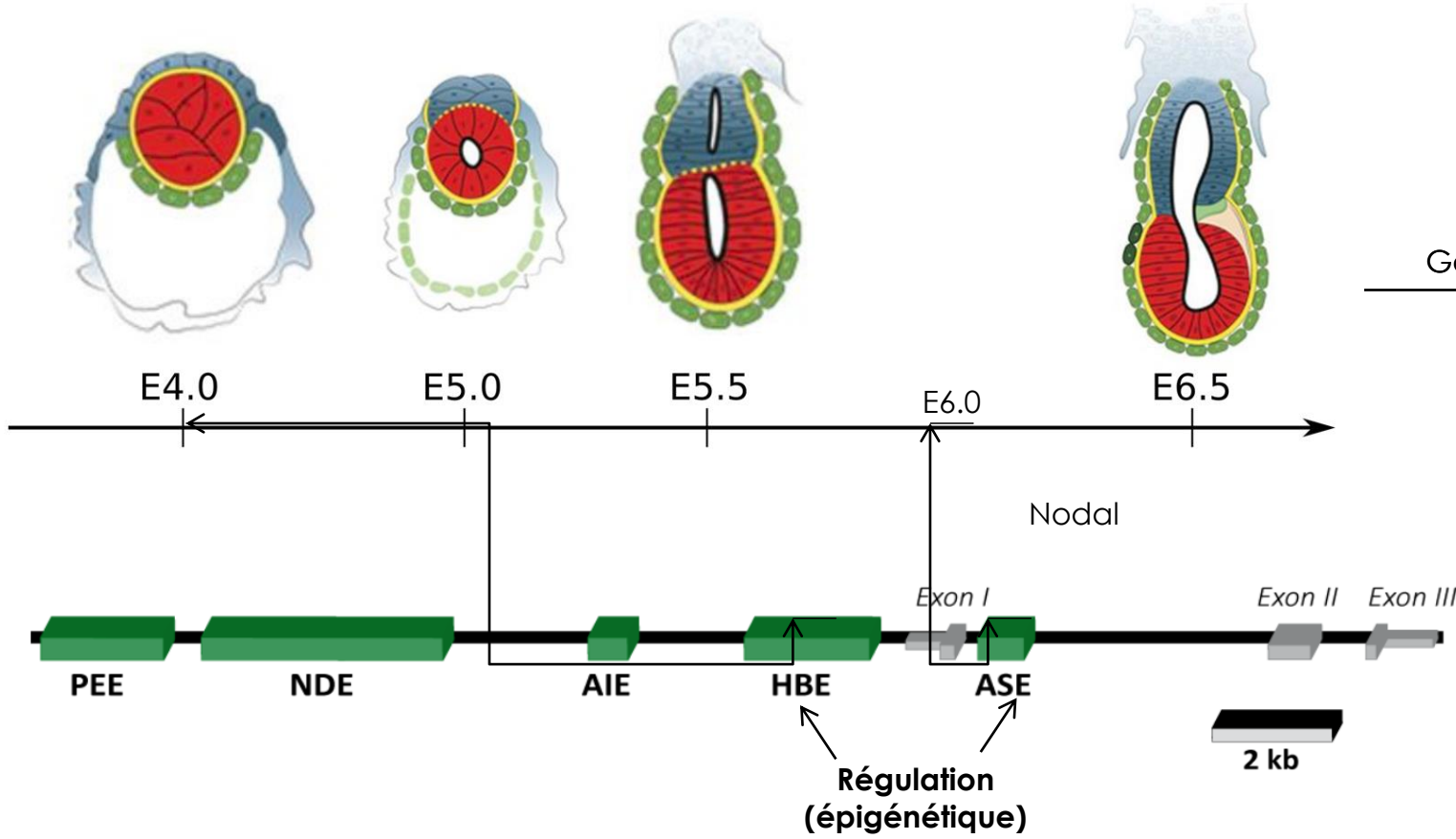


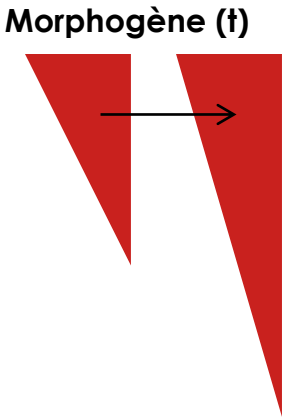
Contexte : Avant 2010, on était obligé de **fixer** les cellules → pas d'étude dynamique possible

Quel modèle utiliser pour étudier la différenciation ?

Source : Sorre 2014

Modèle : Les cellules souches de souris pendant le développement précoce





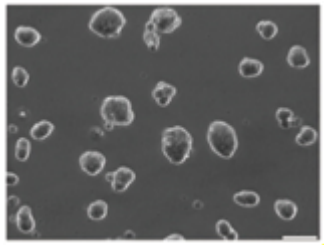
Pattern : 3 feuillets
(Ectoderme, Mésoderme, Endoderme)

NODAL s'auto-active

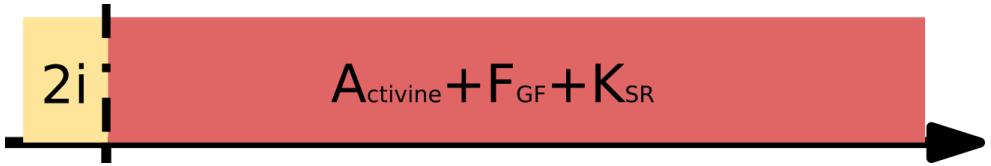
Comment la voie de signalisation (TGF-β) **contrôle** l'expression d'un gène de différenciation (NODAL) ?

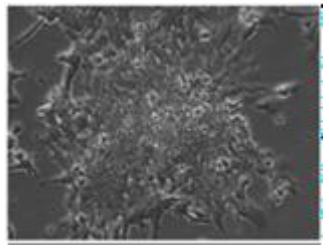
Source : Harrison 2017 ; Vieira 2018 ; Papanayotou 2014

Approche microfluidique *in vitro* pour émuler les signaux dans un embryon

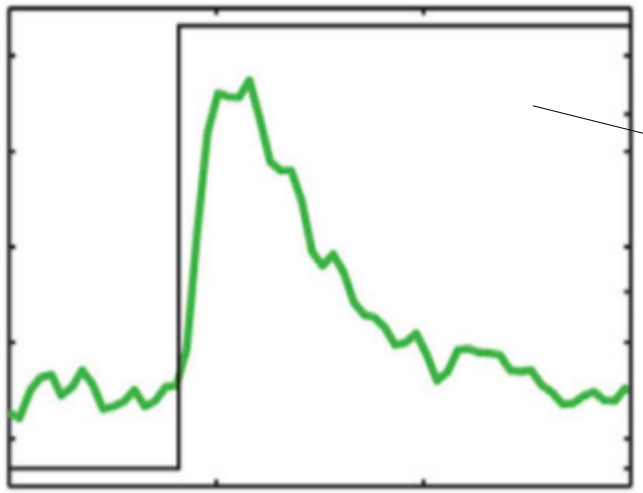


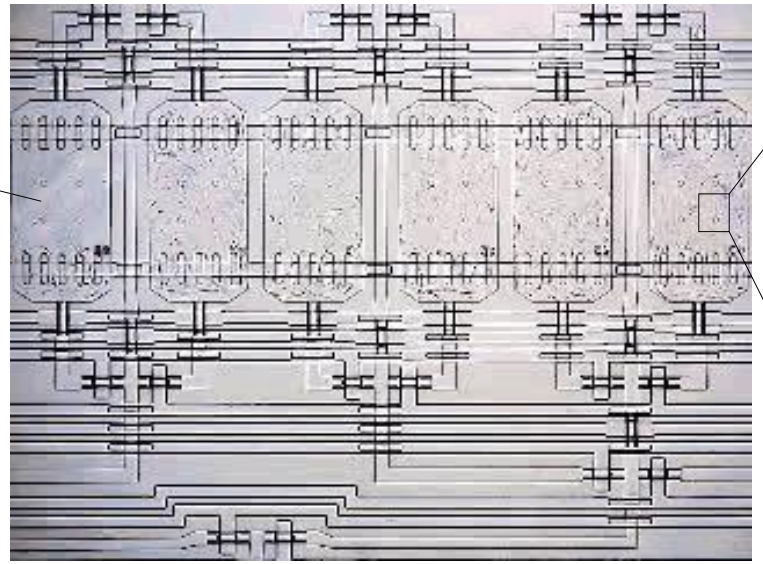
mESC (~E3.5)

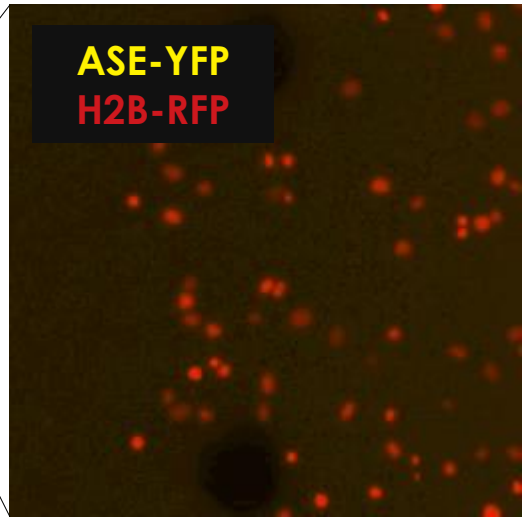




mEpiLC (~E6.5)



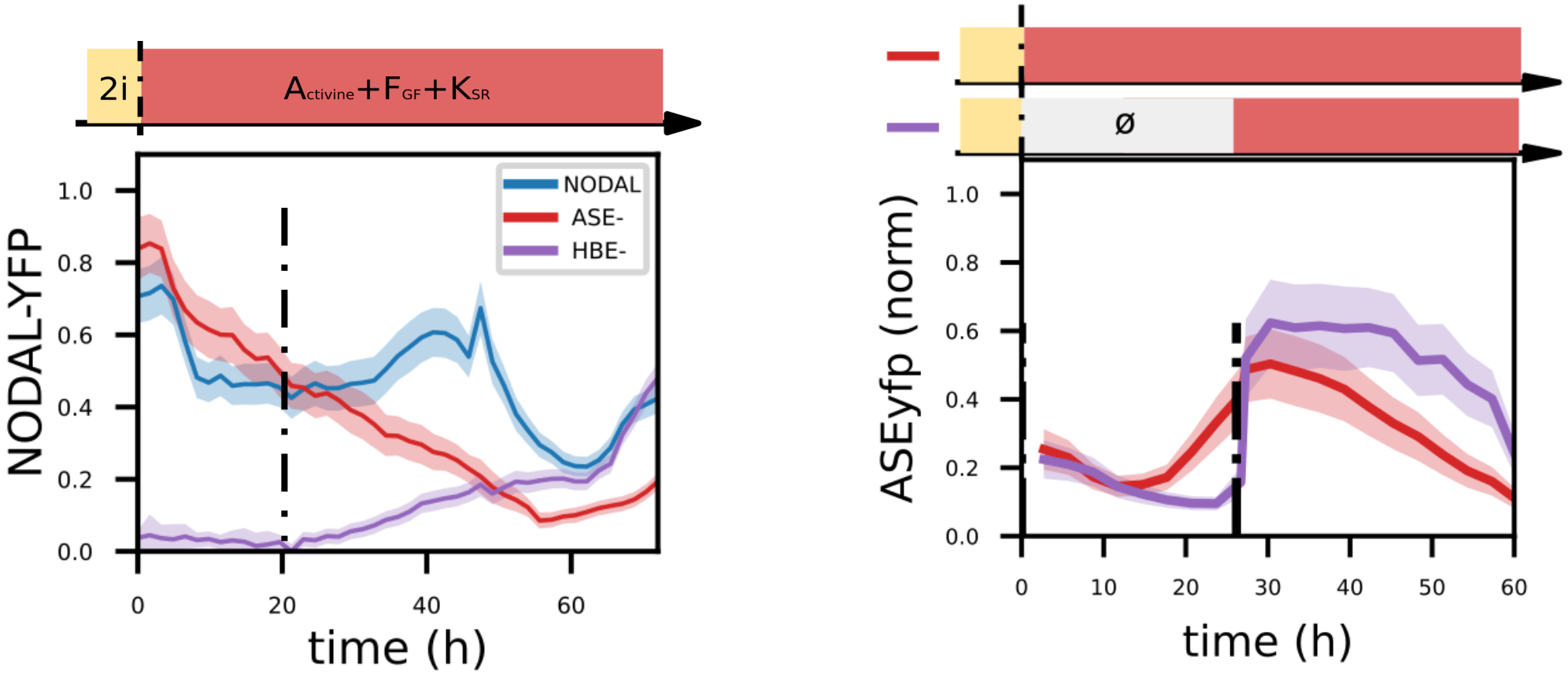




Génération de signaux par simple **changement de milieu** de culture et mesure par **analyse d'image**

Source : Hayashi 2011, Gómez-Sjöberg, 2007

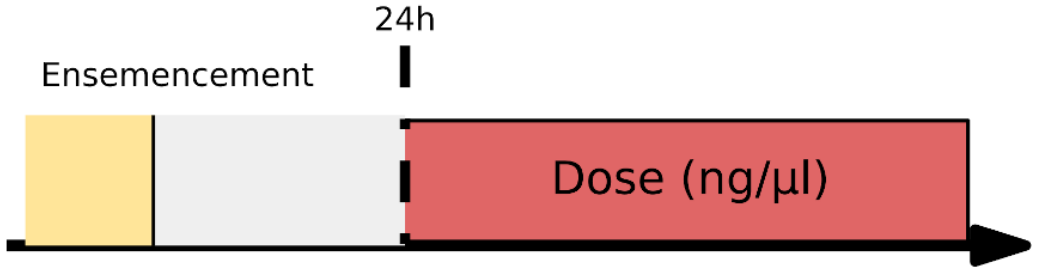
Résultat 1 : Changement de régulation au cours de la différenciation précoce

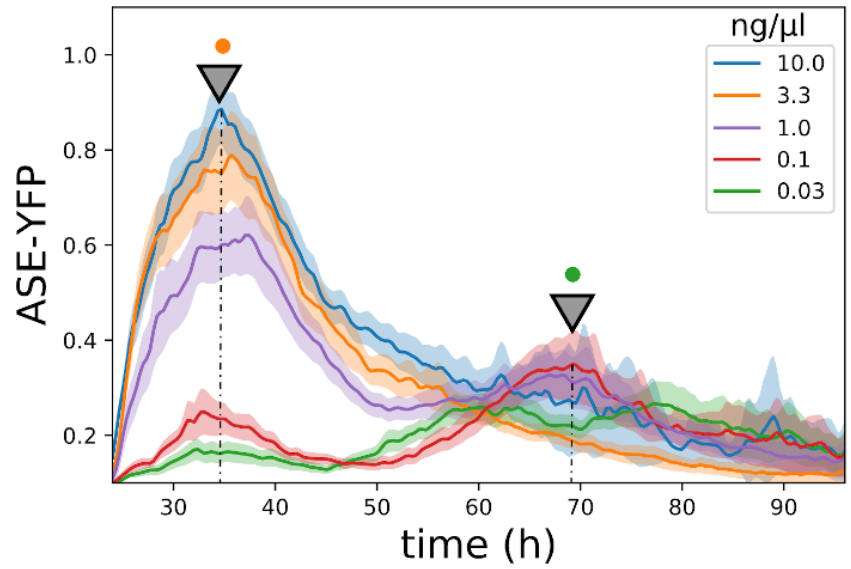


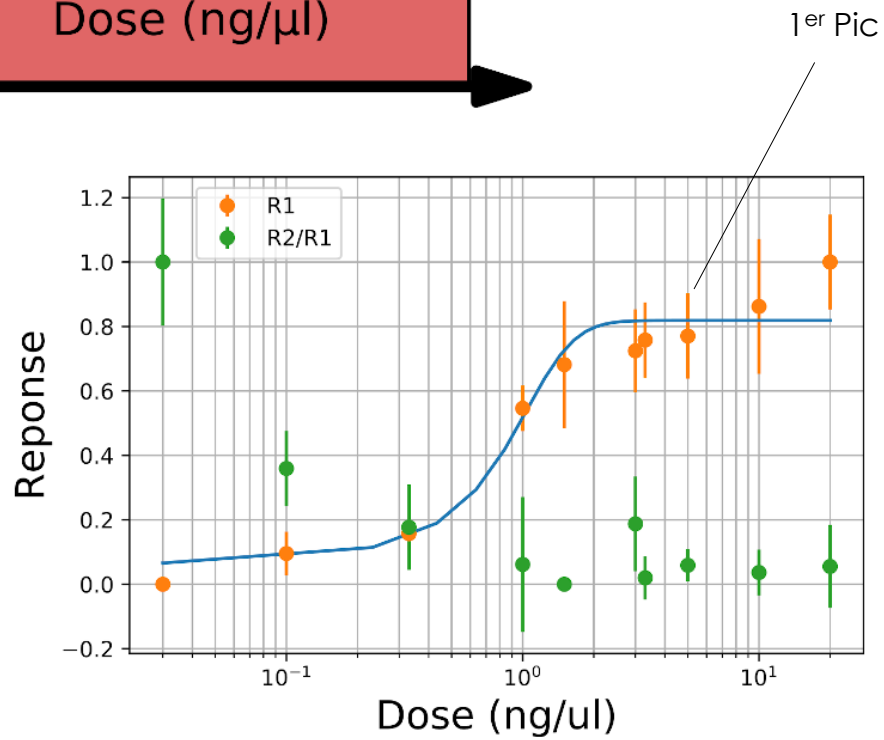
Observation : Balance de réponse **et** Délai de réponse de 20h

Hypothèse : Transition HBE→ASE *in vitro*

Résultat 2 : Phénomène de mémoire à la stimulation



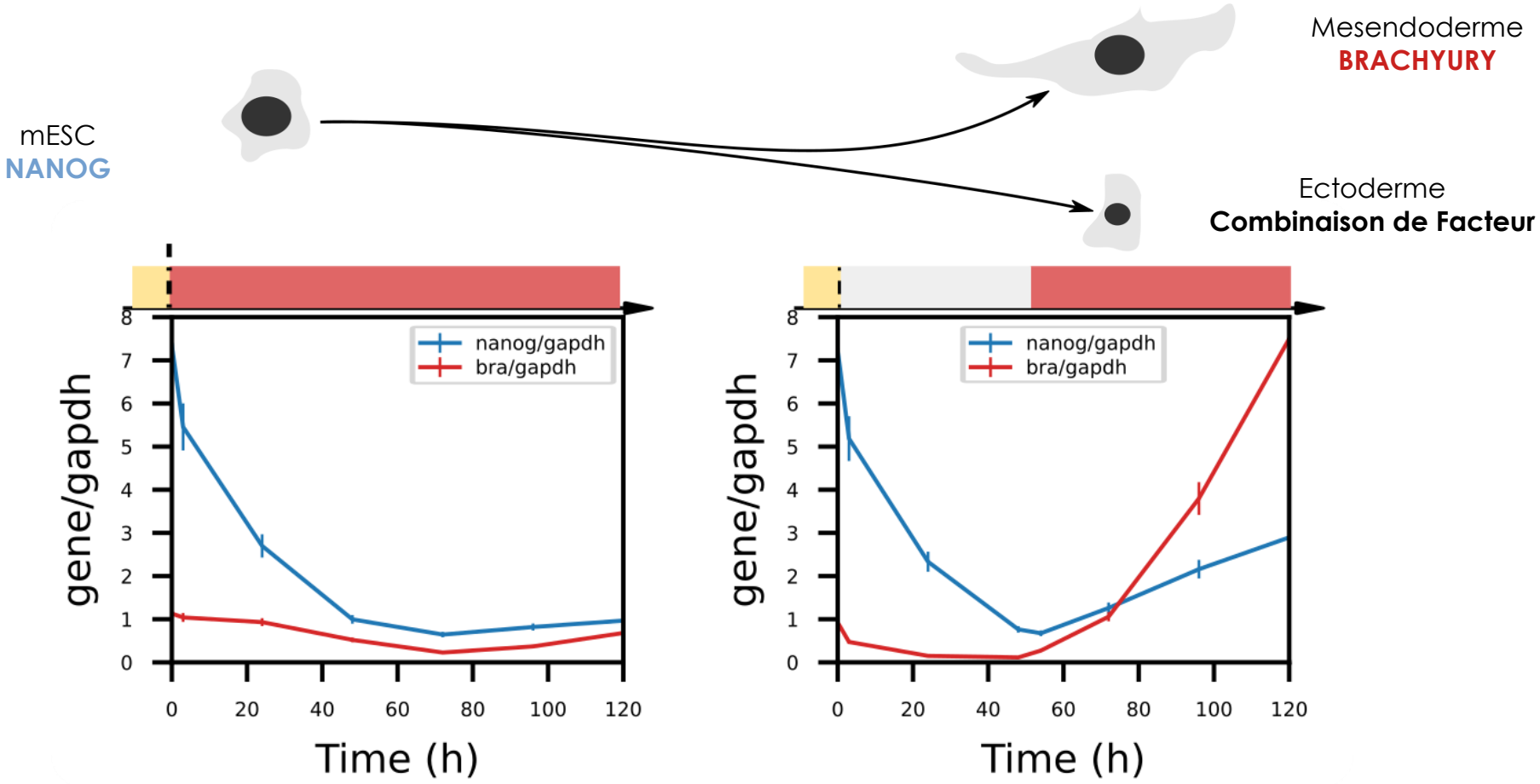




Observation : 1er Réponse linéaire et adaptative **et** balance de réponse (mémoire+densité)

Hypothèse : Changement d'état cellulaire ⇒ Changement de fonction de transfert

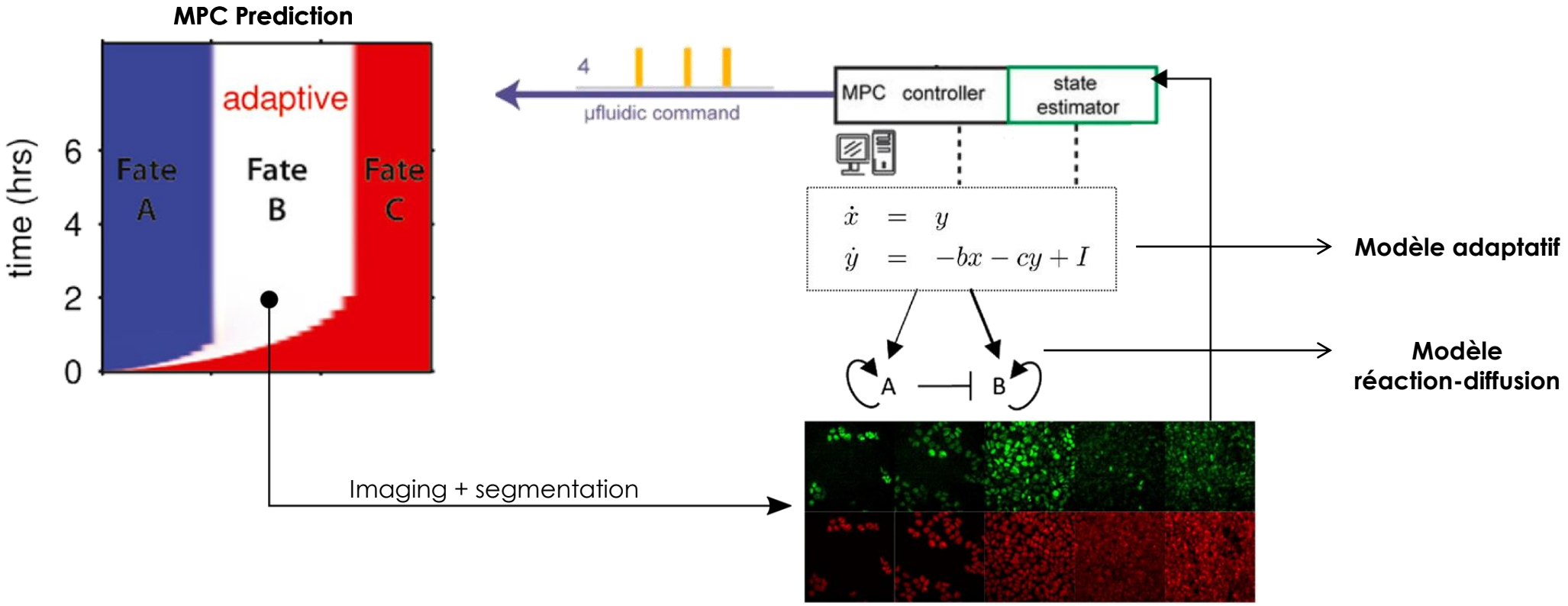
Résultat 3 : Un délai de stimulation influence la différenciation



Observation : La temporalité de stimulation influence le futur type cellulaire

Hypothèse : L'épigénétique change la fonction de transfert de la cellule.

Perspective : Vers une commande prédictive



Modélisation : Prise en compte de l'épigénétique (changement d'état) + densité cellulaire

Expérience possible : Contrôle en temps réel de la différenciation (mais problème de temps long)

Conclusion PhD

1) Transition de la régulation au cours de la différenciation.

2) Adaptation, Auto-activation et Mémoire.

3) Modélisation de la fonction de transfert par l'ajout d'épigénétique.

MSC

INSTITUT JACQUES MONOD

E I P F D

HFSP

Plan de la présentation

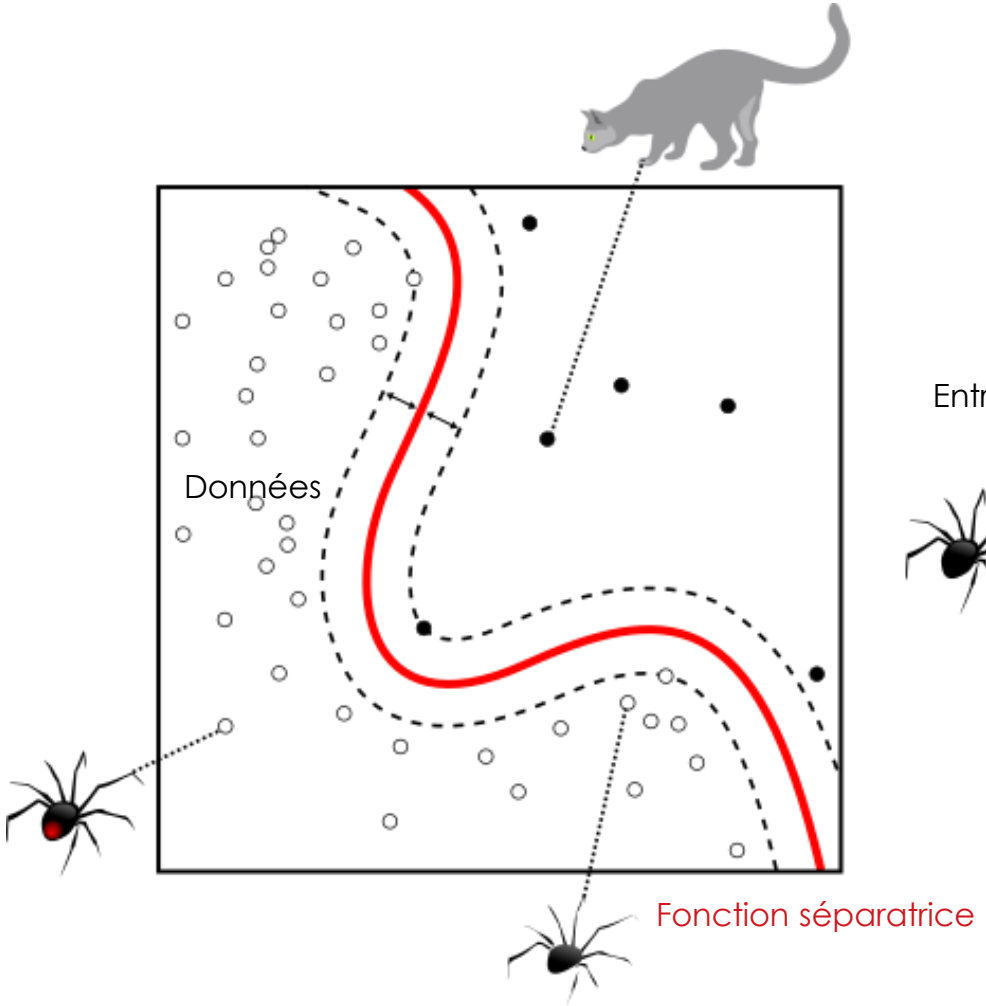
Partie 1, Thèse : Caractérisation de la voie de signalisation NODAL lors de la différenciation des cellules souches embryonnaires.

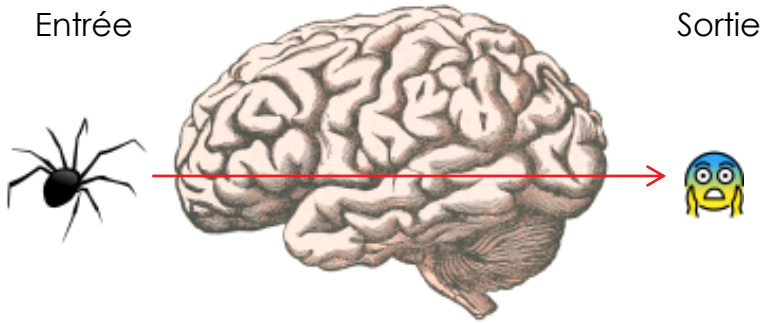
(*Bonus :* How to illustrate the beauty of mathematics? in Instagram @fabienfrfr)

Partie 2, Reconversion IA : Un réseau de neurones artificiels fonctionnalisé par l'évolution.

Compétence : Python, Mathématiques appliquées, Machine Learning, Renforcement et Data.

Principe : Le vivant comme modèle de ML*





« Fonction séparatrice »

$$\text{Cerveau} = \sum \text{neurone}$$

(traitement de l'information)

*ML : Machine Learning

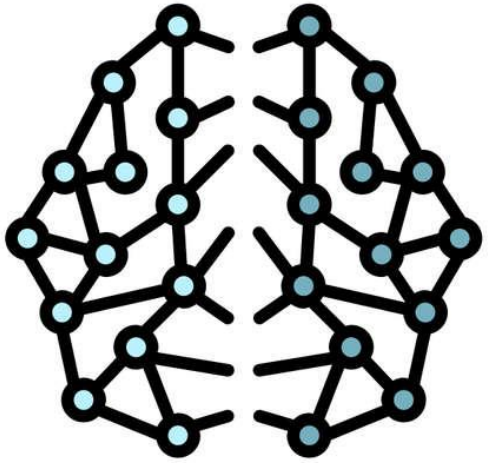
Idée générale : Le vivant n'a pas besoin de beaucoup de données pour apprendre

→ S'inspirer d'autant plus du vivant pour construire des réseaux de neurones.

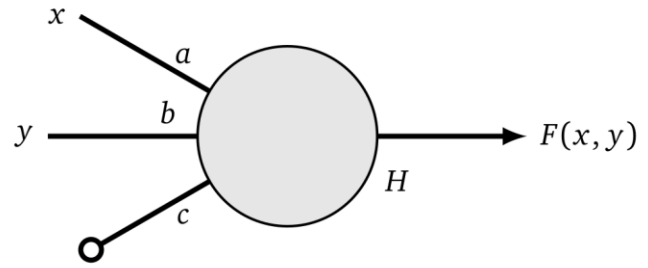
Cognitivisme vs Connexionnisme

(Théorie de l'esprit contemporaine)

Stimulus $\sum$x Comportement



Hyperconnectée & Répétition



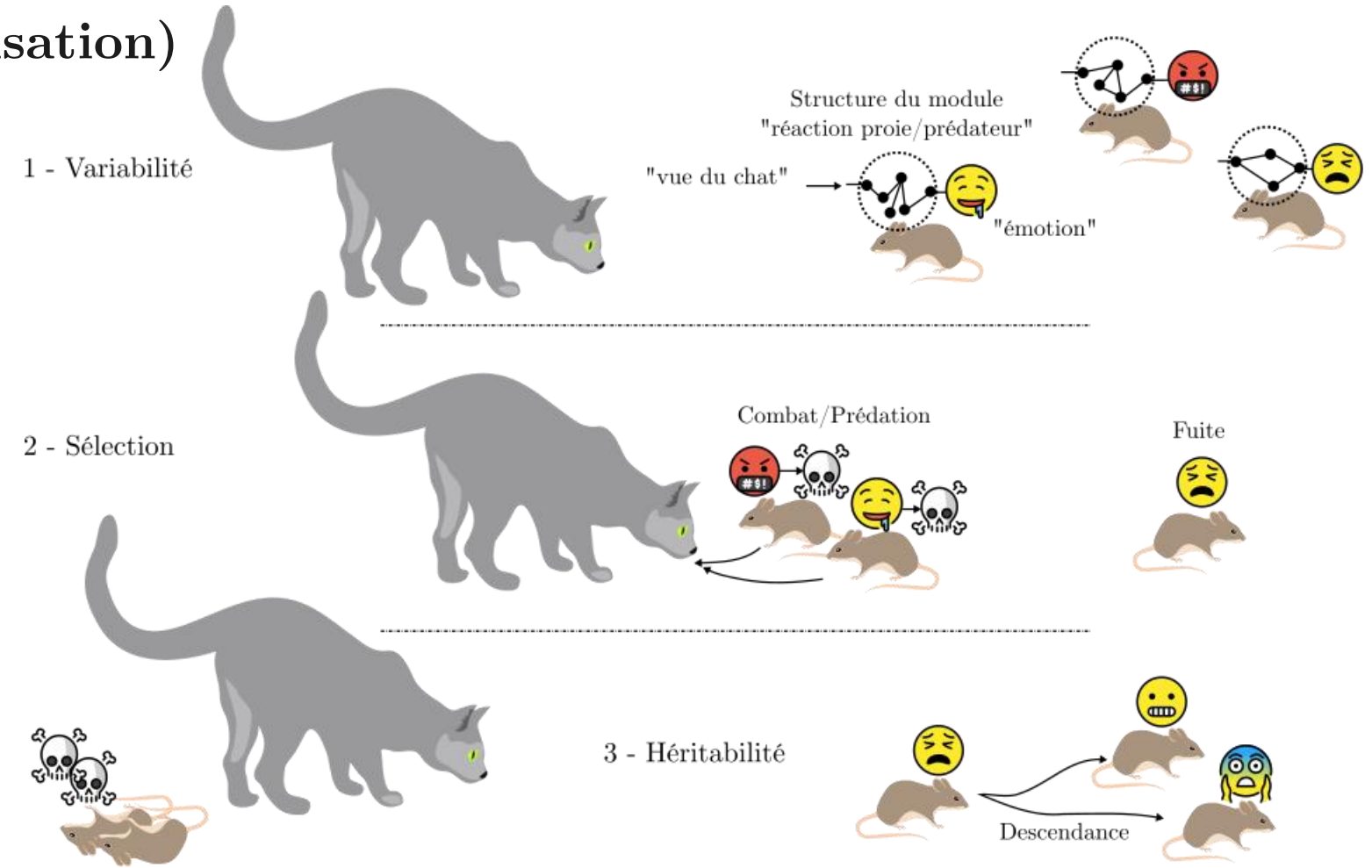
$F(x,y) = 1 \quad if \quad (x.a + y.b + c) > 0$

Le neurone formel (1957 Rosenblatt)

L'apprentissage est une propriété émergente du cerveau

Les réseaux de neurones artificiels modernes vont dans le sens du connexionnisme

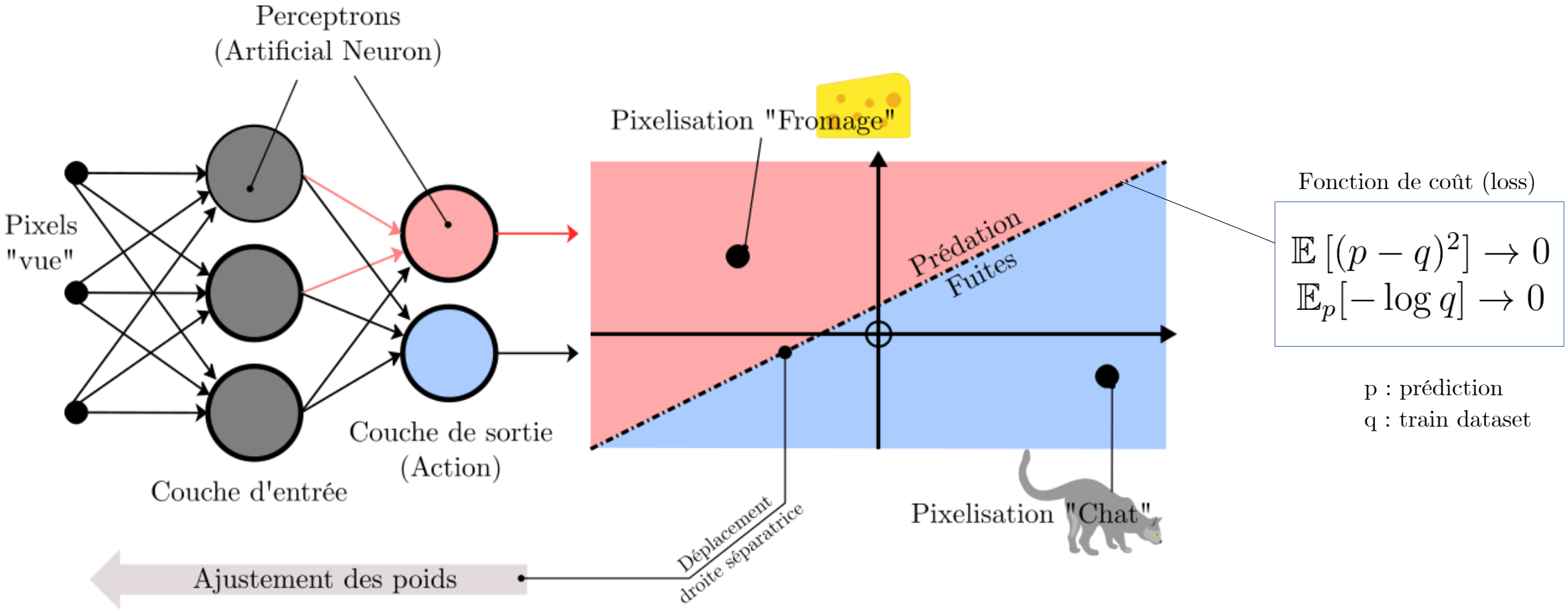
Approche évolutionniste
(Vulgarisation)



Les deux approches précédente ne sont pas incompatibles

La structure neuronale est spécifique et nécessite moins d'événement pour l'apprentissage

L'apprentissage = Minimisation d'erreur
(Vulgarisation)



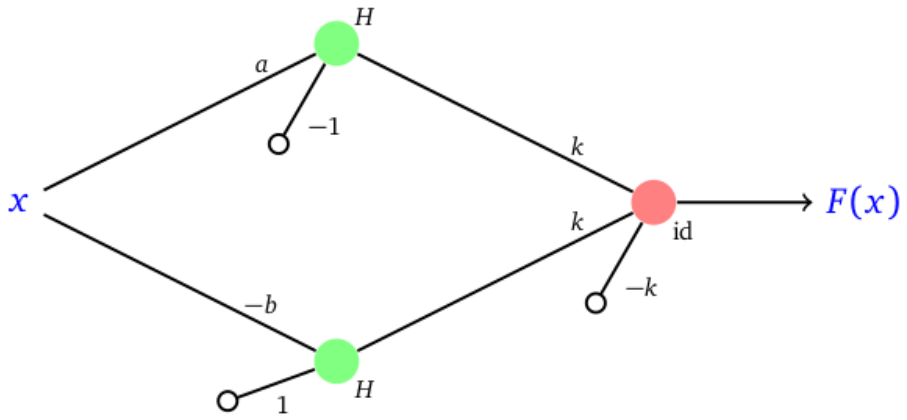
L'apprentissage consiste à optimiser la droite qui sépare correctement les données

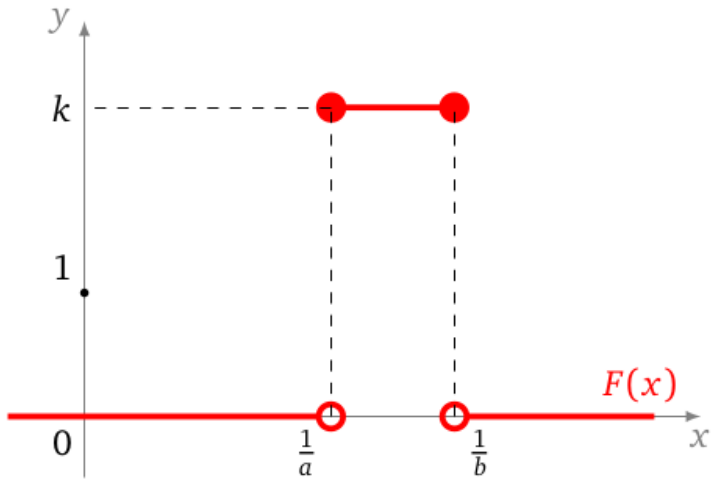
Lorsque la droite est "bien" placé, la fonction de coût est à un minimum global

Démonstration : Pourquoi ça marche ?

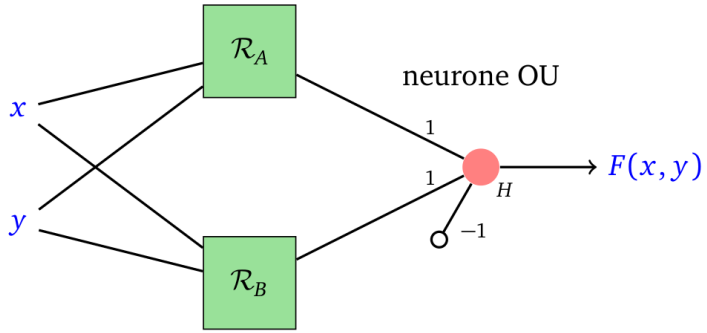
(Proposition)

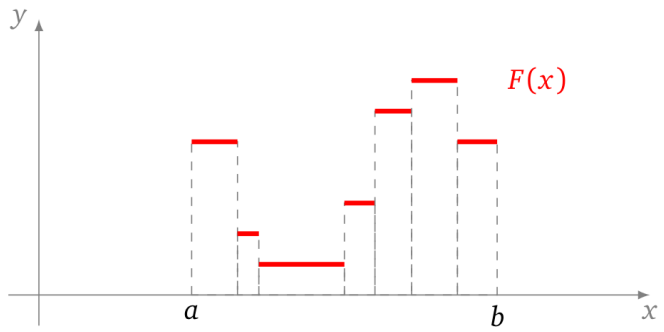
1er : Fonction créneaux





2eme : Fonction en escalier





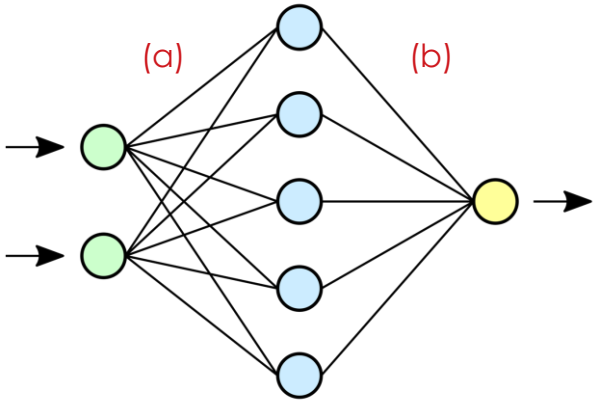
Théorème d'approximation universelle

Toute fonction continue f : [a,b] → ℝ peut être approché par une fonction F : [a,b] → ℝ réalisé par un réseau de neurone (⇔ Intégrale de Lebesgue)

Ajustement des poids :

(gradient vs évolutif)

$$\mathbb{E}\left[(p-q)^2\right] \to 0$$
$$\mathbb{E}_p[-\log q] \to 0$$

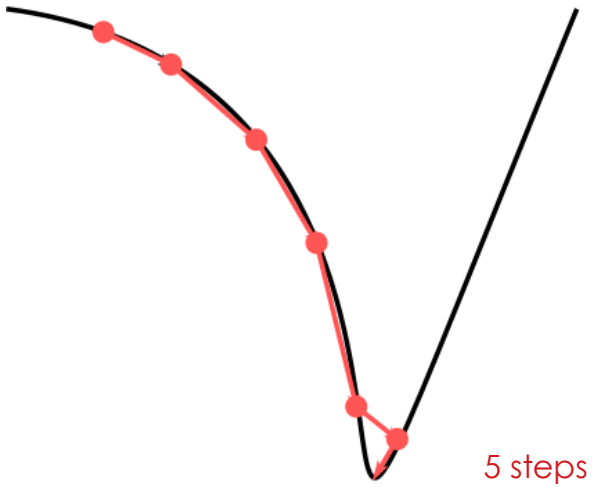


Descente de gradient

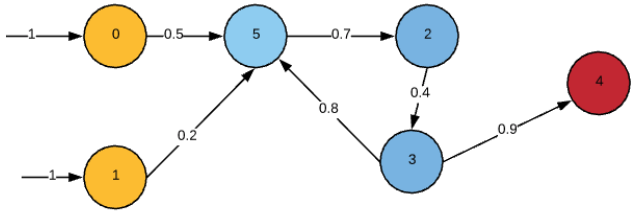
$$a = a - \alpha \frac{\partial J(a,b)}{\partial a}$$

Formule du gradient

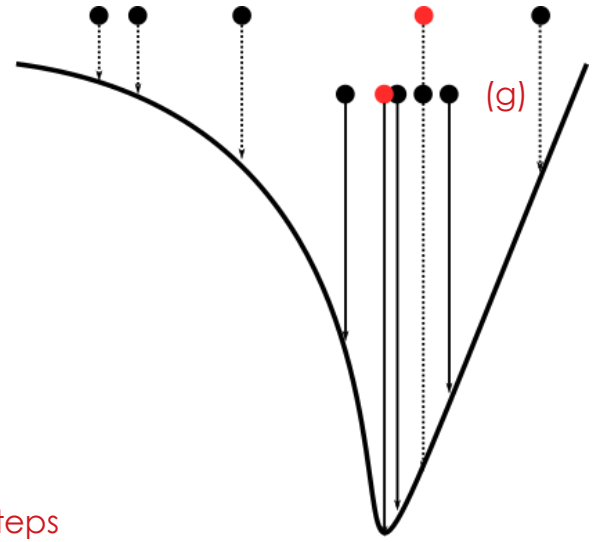
$$\frac{\partial F}{\partial a} = \frac{\partial h}{\partial a}\frac{\partial b}{\partial h}\frac{\partial F}{\partial b}$$



5 steps



meilleur score



2 steps

Algorithme évolutif

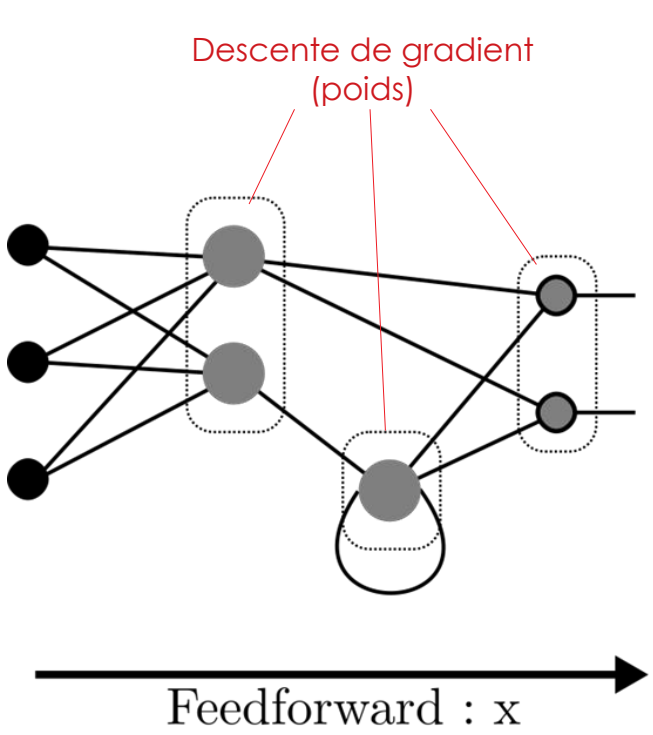
$$(a_g) = a_{g-1} + (a_m)$$

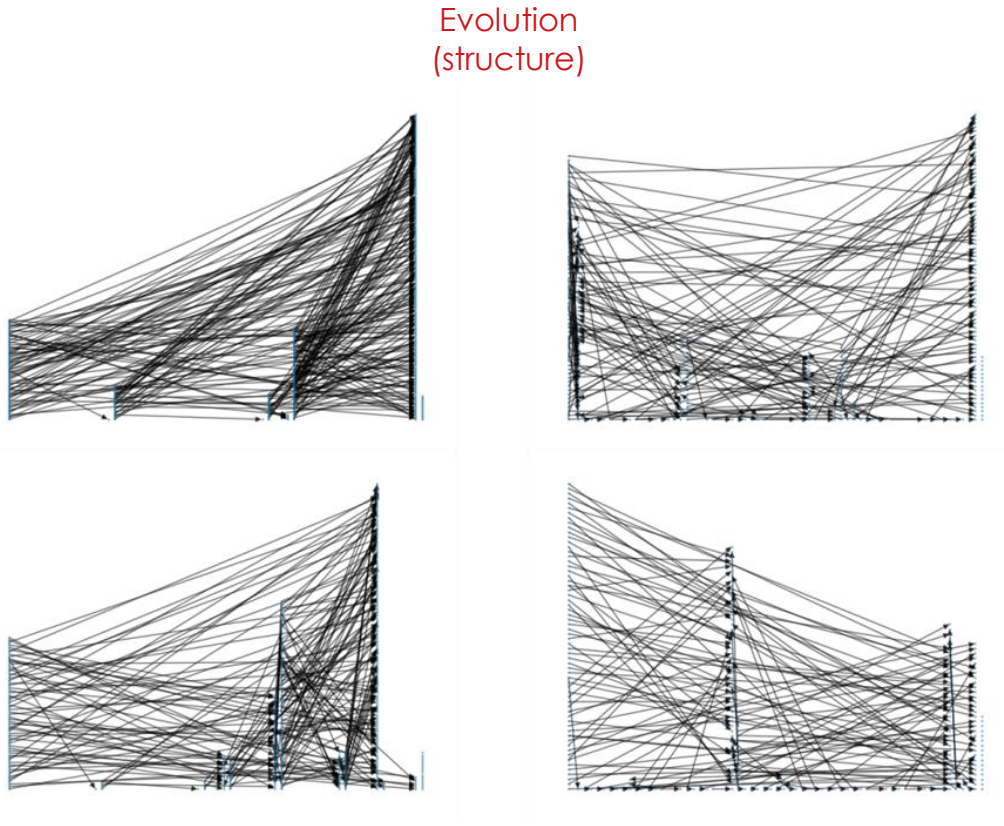
Les algorithmes de descente de gradient ont pris de l'ampleur avec l'amélioration du calcul tensoriel par GPU

Les algorithmes évolutifs convergent plus vite, mais ne sont pas GPU-like

Vers une approche hybride

(Le modèle)

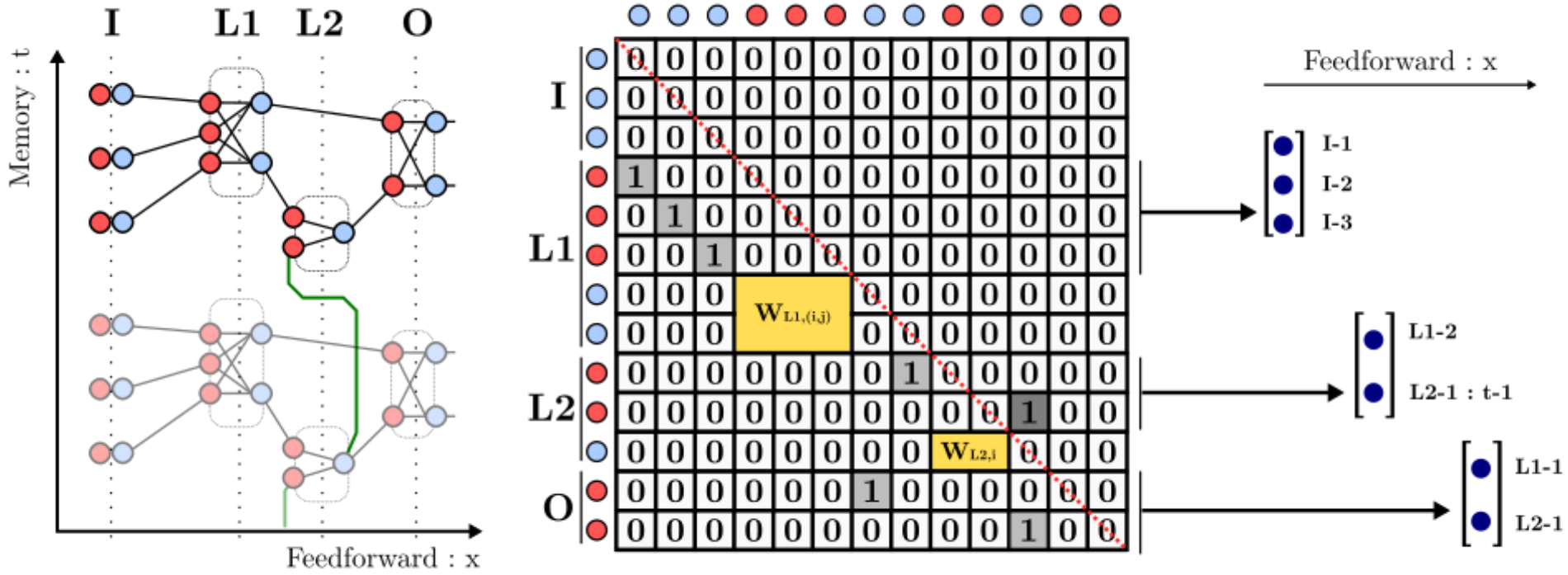




Les couches de neurones sont GPU-like et la structure est sélectionné par l'évolution

Hypothèse : Le structure est pré-fonctionnalisé pour accélérer la convergence par descente de gradient

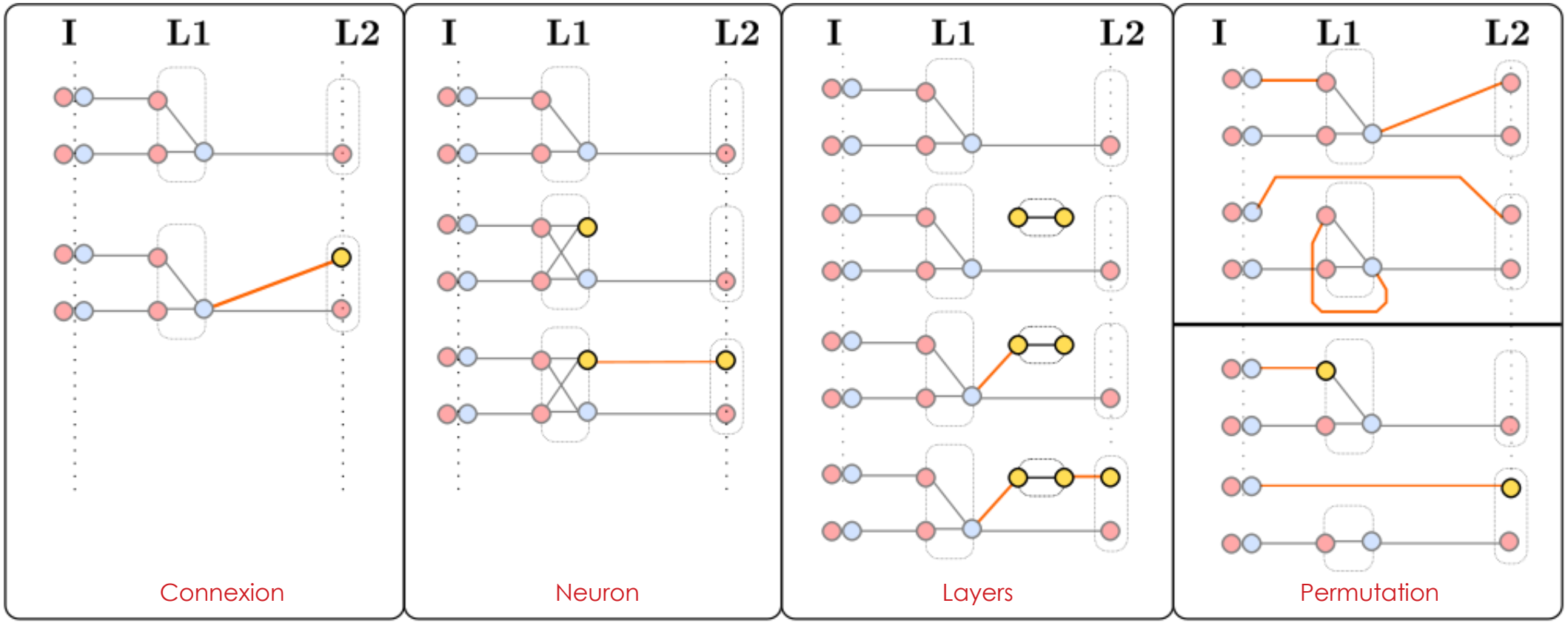
Construction d'un réseau par des graphes



La construction de la propagation avant se fait par liste de connexion de nœud

Il permet les phénomènes de mémorisation par l'apparition de lien vers l'arrière à t-1 (récurrence)

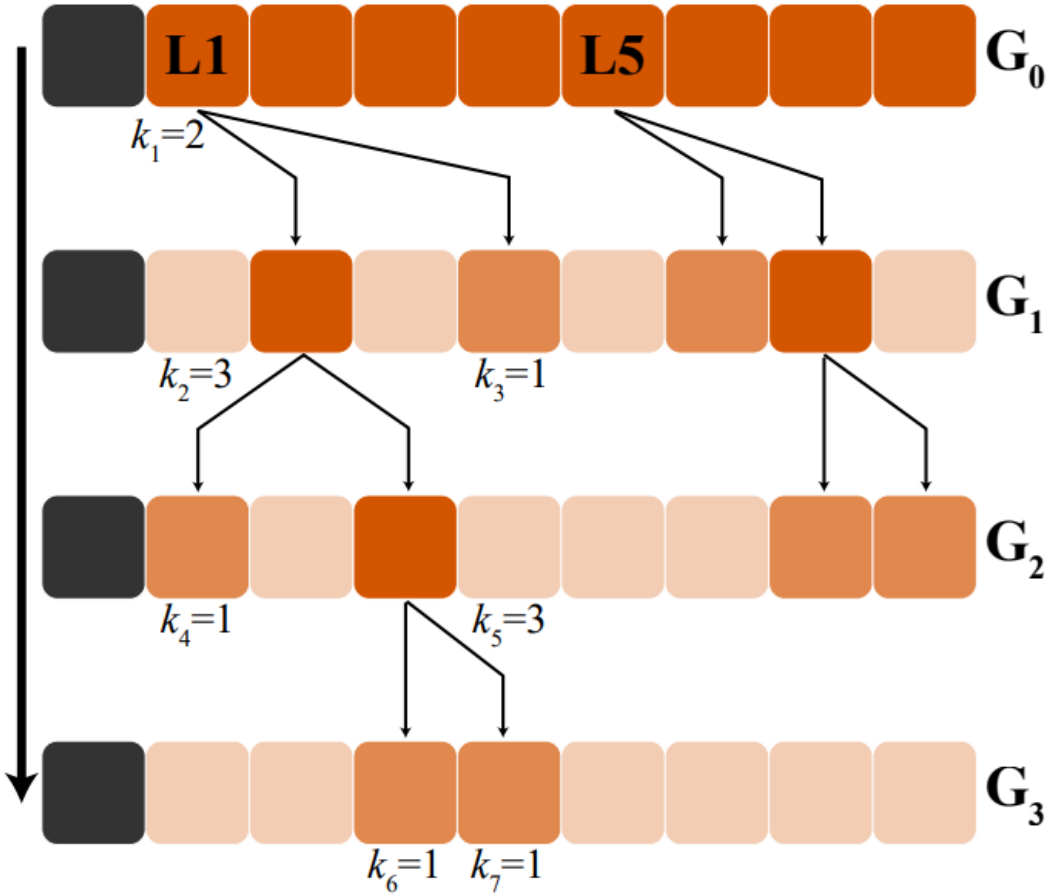
Il existe 7 types de mutations possibles



La topologie initiale est conservé et le graphe doit être complet

Problème : Possible structure vestigiale au cours du processus

Processus de sélection pseudo-Darwinien



L'évolution du réseau est caractérisé par un arbre "phylogénétique"

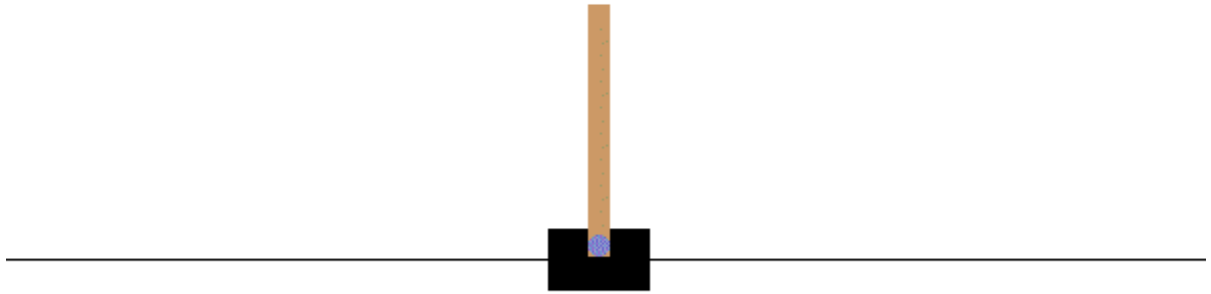
Hypothèse : Plus on avance dans les générations, plus le réseaux est pré-fonctionnalisé

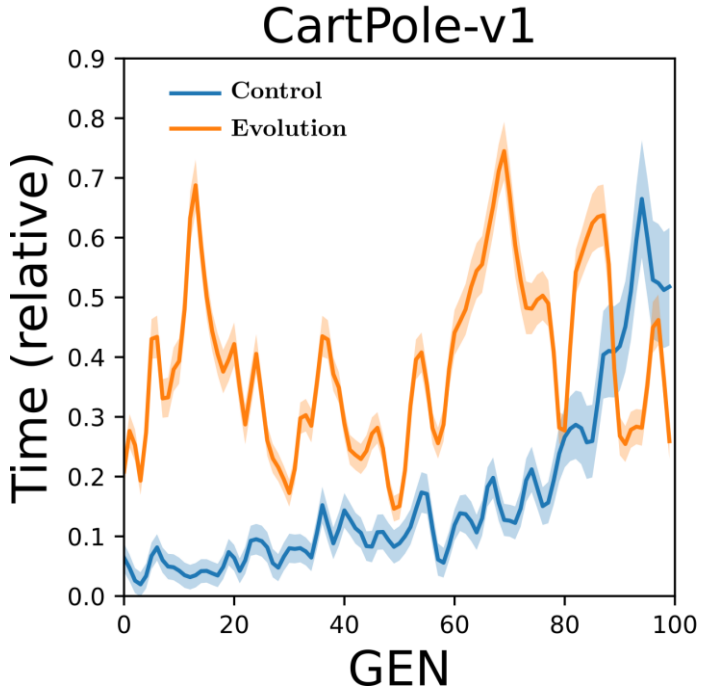
Résultat 1 : Pour des réseaux de petite taille, la convergence est instable

$$\mathbb{E}\left[(p-q)^2\right] \to 0$$

Classic control (agent ↔ environnement) :
Contrôle d'équilibre

Entrée : Position, Vitesse, Angle, Vitesse Angulaire
Sortie : Aller à gauche / Aller à droite





Le score du processus de décision markovien est défini par le temps de maintient en équilibre

Observation : La convergence est rapide, mais une simple perturbation dans le réseau le rend moins efficace

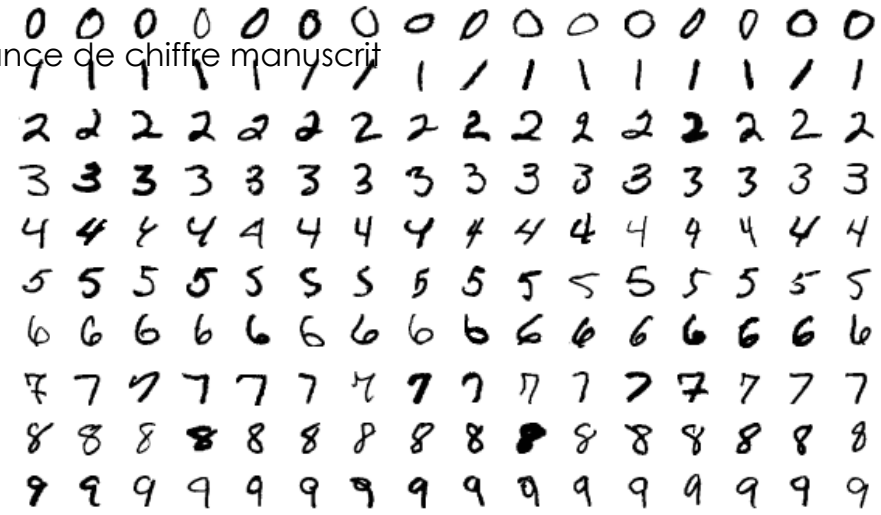
Résultat 2 : Pour des réseaux de grande taille, la structure est stabilisante

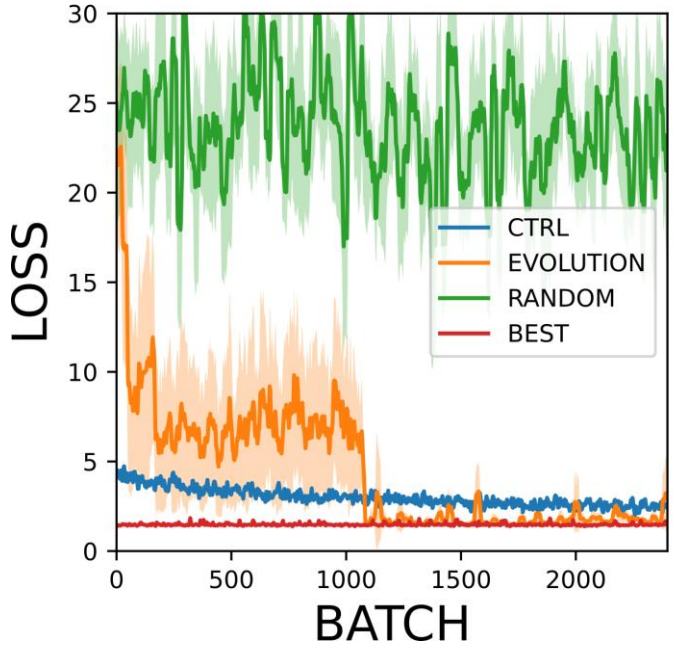
$$\mathbb{E}_p[-\log q] \to 0$$

Classic classification: Reconnaissance de chiffre manuscrit

Entrée : Image 28*28

Sortie : 0 → 9





La stabilité du réseau évolutif est déterminée par le nombre de nœuds dans le réseau

Observation : Lorsque la fonction de coût est stabilisé, la génération suivante hérite 20k événement équivalent

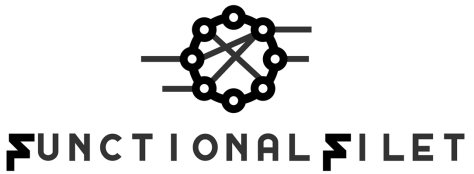
Comment utiliser mon modèle ?

Installation (bash) :

```
python3 -m pip install functionalfilet
```

Utilisation (python) :

```
# package
from functionalfilet import model as ff
# init model
model = ff.FunctionalFilet()
# train
model.fit(X,y)
# test
y_pred = model.predict(X, index=seeder_idx)
```



Une documentation est présente sur GitHub sous forme de « Learning by Example »

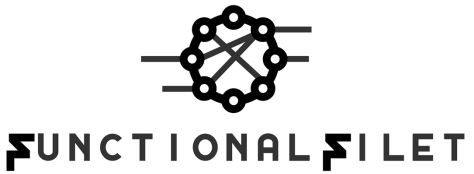
https://github.com/fabienfrfr/functionalfilet (Toute contribution est la bienvenue)

Suite : La construction du réseau en C++


```python
class pRNNFunction(Function):
    @staticmethod
    def forward(ctx, input, weights, bias):
        """ Calculation """
        return

    @staticmethod
    def backward(ctx, grad_h, grad_cell):
        """ Calculation """
        return
```

Convert to **C Extension**




```python
class pRNN(nn.Module):
    def __init__(self, input_features, state_size):
        super(pRNN, self).__init__()
        self.weights = nn.Parameter(
            torch.Tensor(3 * state_size, input_features + state_size))
        self.bias = nn.Parameter(torch.Tensor(1, 3 * state_size))

    def forward(self, input, state):
        return pRNNFunction.apply(input, self.weights, self.bias, *state)
```

Le projet est open source CC 4.0

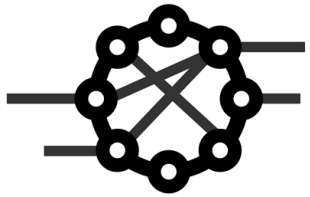
Une citation est apprécié pour l'utilisation public (https://doi.org/10.48550/arXiv.2205.10118)

Conclusion IA

1) Approche hybride possible

2) Vers une pré-fonctionnalisation

3) Amélioration par extension C++



FUNCTIONAL FILET