

Apprentissage par Réseaux de Neurons Profonds



COLLÈGE
DE FRANCE
—1530—

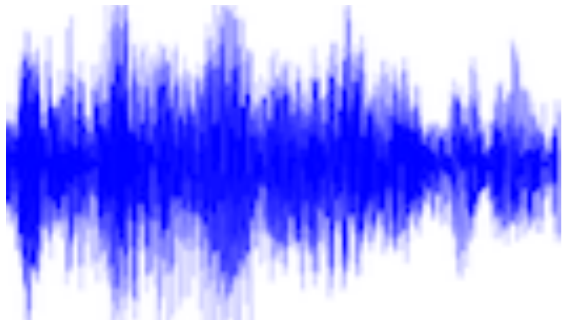
Cours 2

Applications des réseaux de Neurons Profonds

Stéphane Mallat

Données diverses mais avec un grand nombre d de variables

$$d \sim 10^6 / mn$$



Audio

$$d \sim 10^6$$

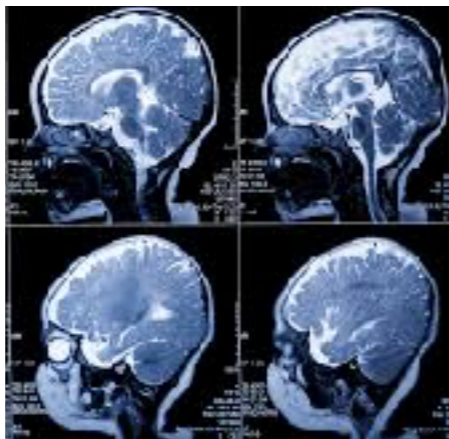


Image 2D/3D

$$d \sim 10^6$$



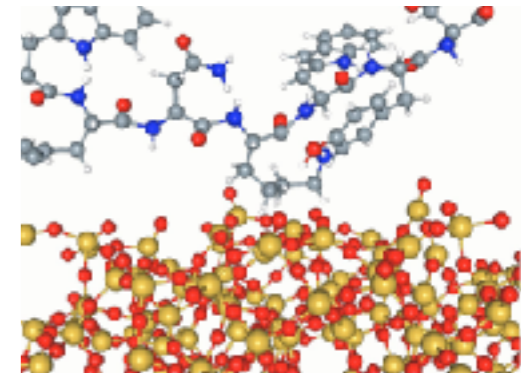
Texte

$$d \sim 10^9$$



Reseau social

$$d \sim 10^{24}$$



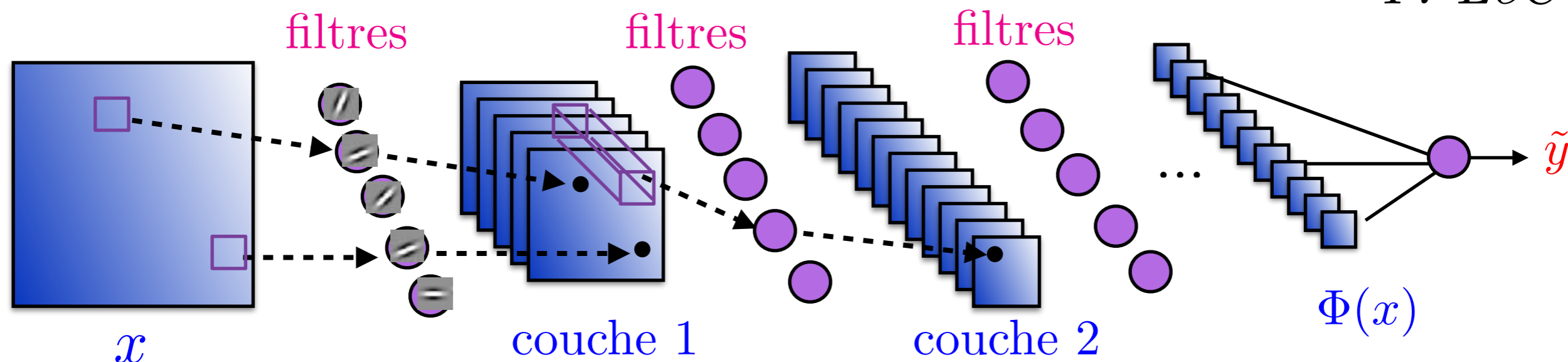
mole de matière

- *Modéliser*: capturer la nature et la variabilité des données
 - compression, restauration, synthèse
- *Prédire*: estimer la réponse d'une question à partir de données
 - Reconnaissance d'images ou de sons, diagnostics médicaux, analyse de textes ou traductions, prédire la physique...

Apprentissage statistique \longrightarrow Intelligence artificielle

- Les paramètres sont invariants par translations: **filtres**

Y. LeCun



- Apprentissage supervisé: $\{x_i, y_i\}_{i \leq n}$

Centaines de millions de paramètres

Prédictions exceptionnelles et *génériques*: pour l'image, parole, diagnostiques, traduction, en physique...

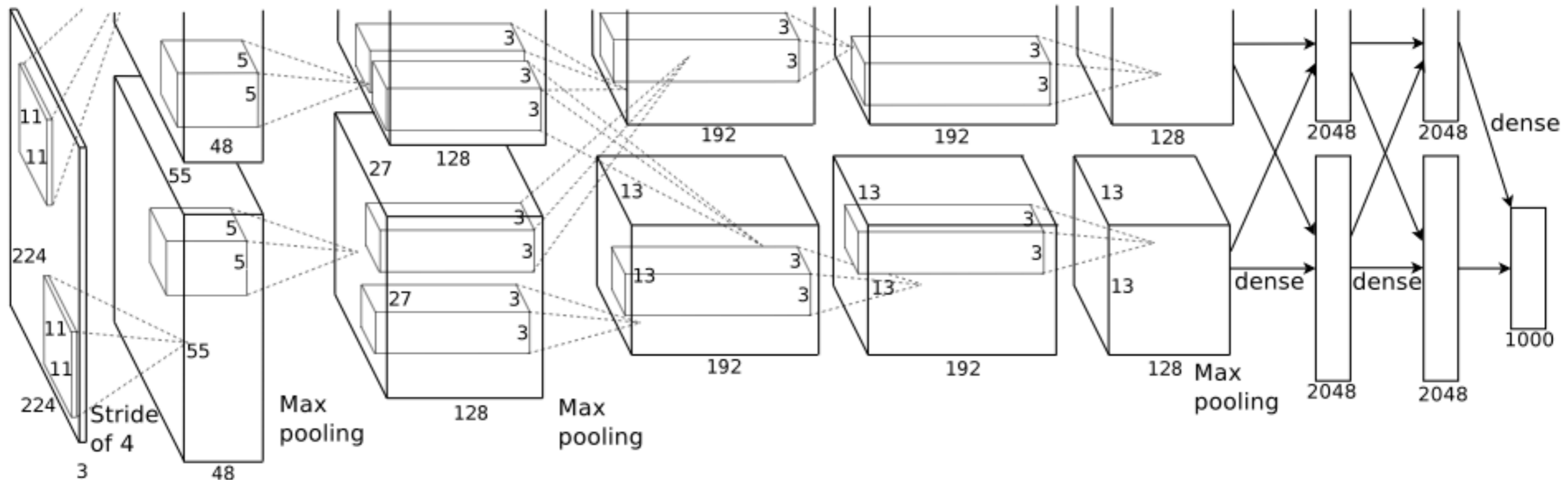
Le $\Phi(x)$ a aplati les frontières. Comment ?

Importance de l'architecture

Alex Deep Convolution Network

A. Krizhevsky, Sutsver, Hinton

- Imagenet supervised training: $1.2 \cdot 10^6$ examples, 10^3 classes
15.3% testing error in 2012



New networks with 3% errors.

Up to 150 layers!

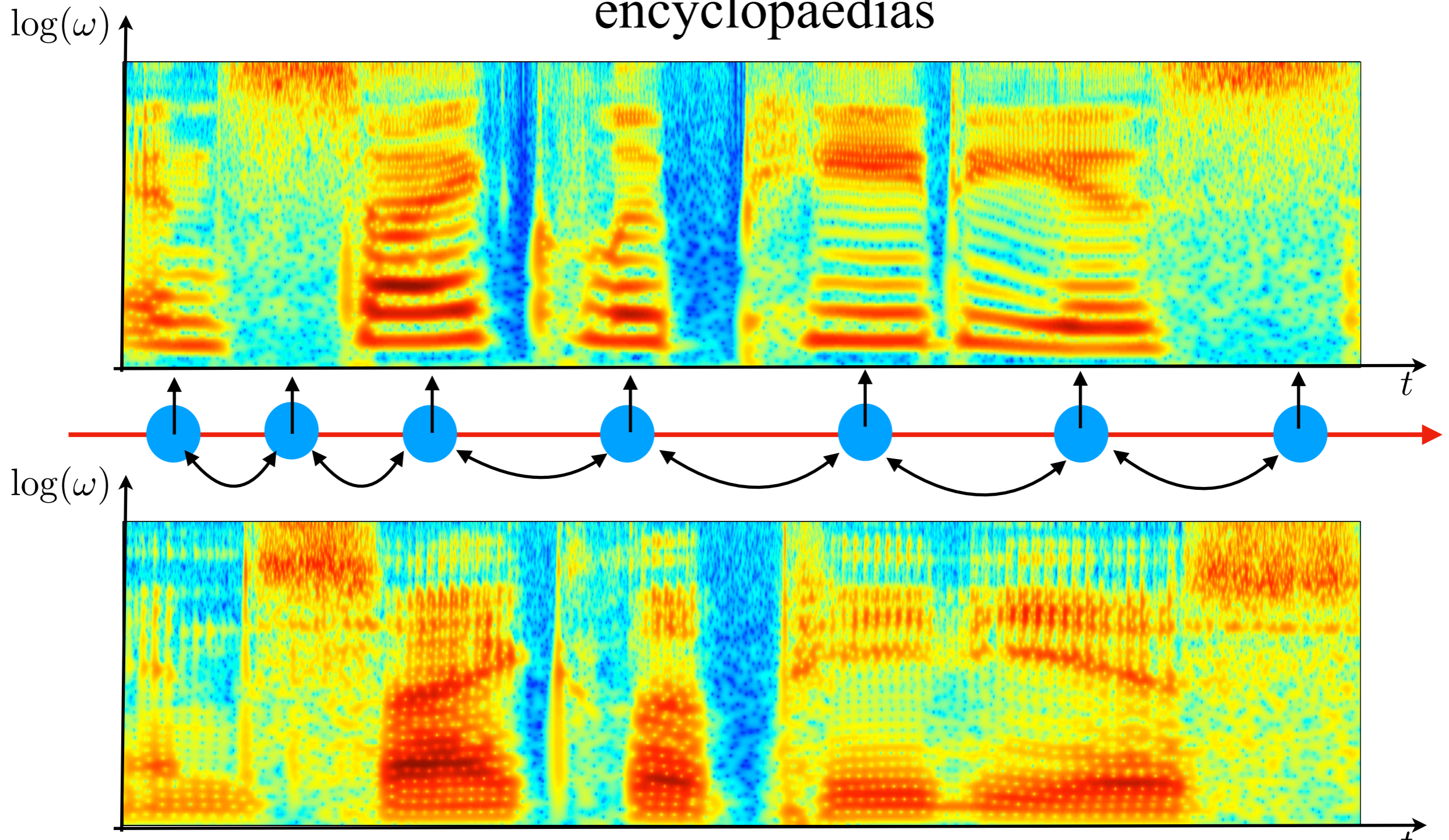
Pourquoi cela marche ?



Reconnaissance de la parole

- Depuis les années 60.
- Jusqu'en 2010, basé sur le calcul de spectrogrammes et modèles de chaînes de Markov.

encyclopaedias

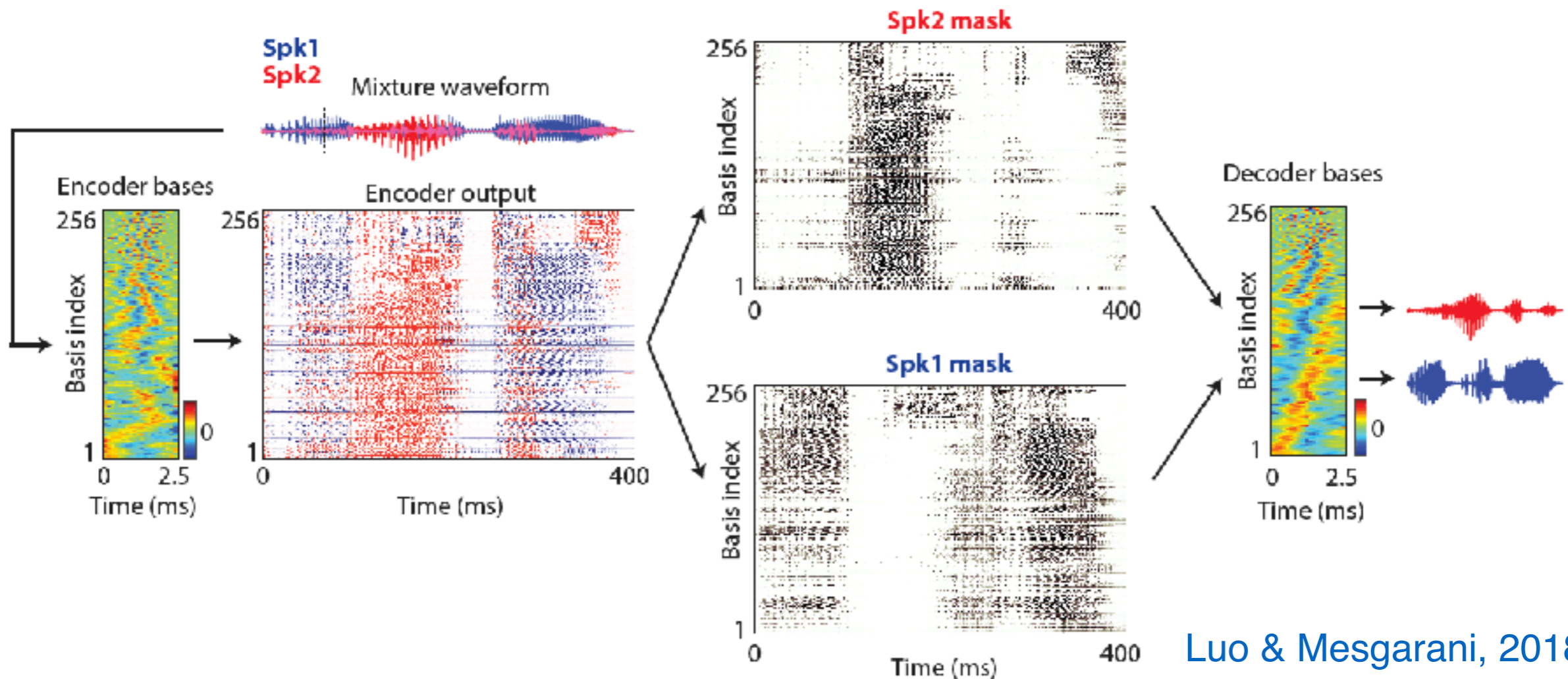


Reconnaissance Audio

- Reconnaissance de la parole, de la musique:



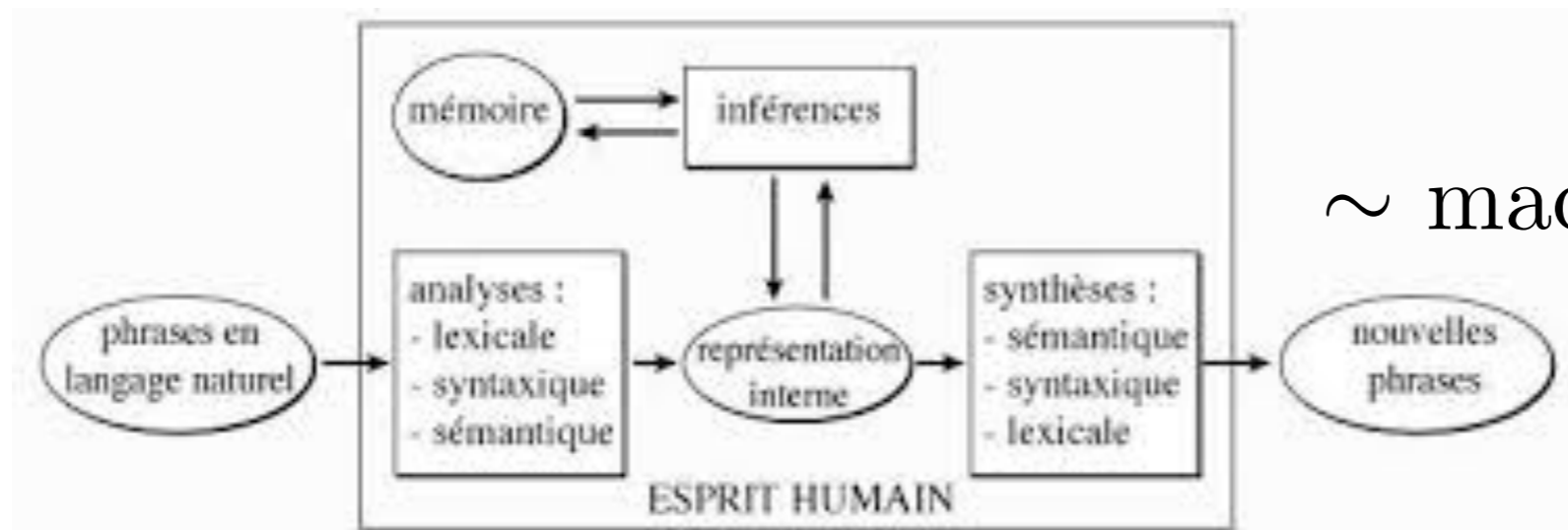
- Séparation de sources:





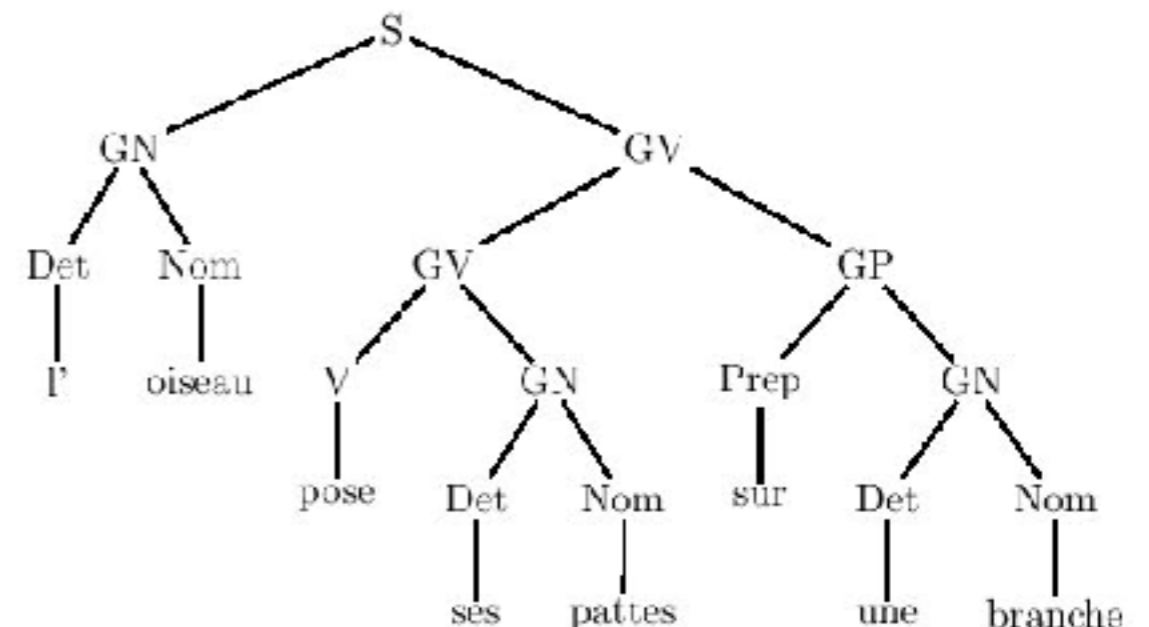
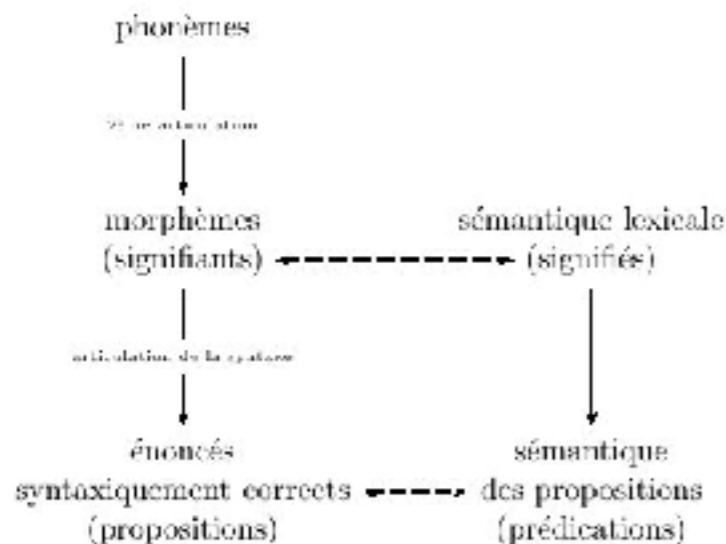
Traitement du Langage Naturel

- Traduction, analyse de textes, conversation d'agents, réponse à des questions, commentaires descriptifs... : *linguistique*
- Structuralisme: système de signes (signifiant, signifié)
- 1960: grammaires formelles de Chomsky, récursivité.
- Sémantique formelle: représentations propositionnelles, logique



~ machine de Turing

niveaux de composition sémantique associée



- Échec partiel: difficulté de représenter la « connaissance » d'un **monde complexe**: « John aime Marie » et ses « règles ».
- Un aller-retour entre rationalisme et empirisme:

Rationalisme

Empirisme



Platon, Leibnitz, Descartes, Kant, Wittgentstein, Hume, Locke

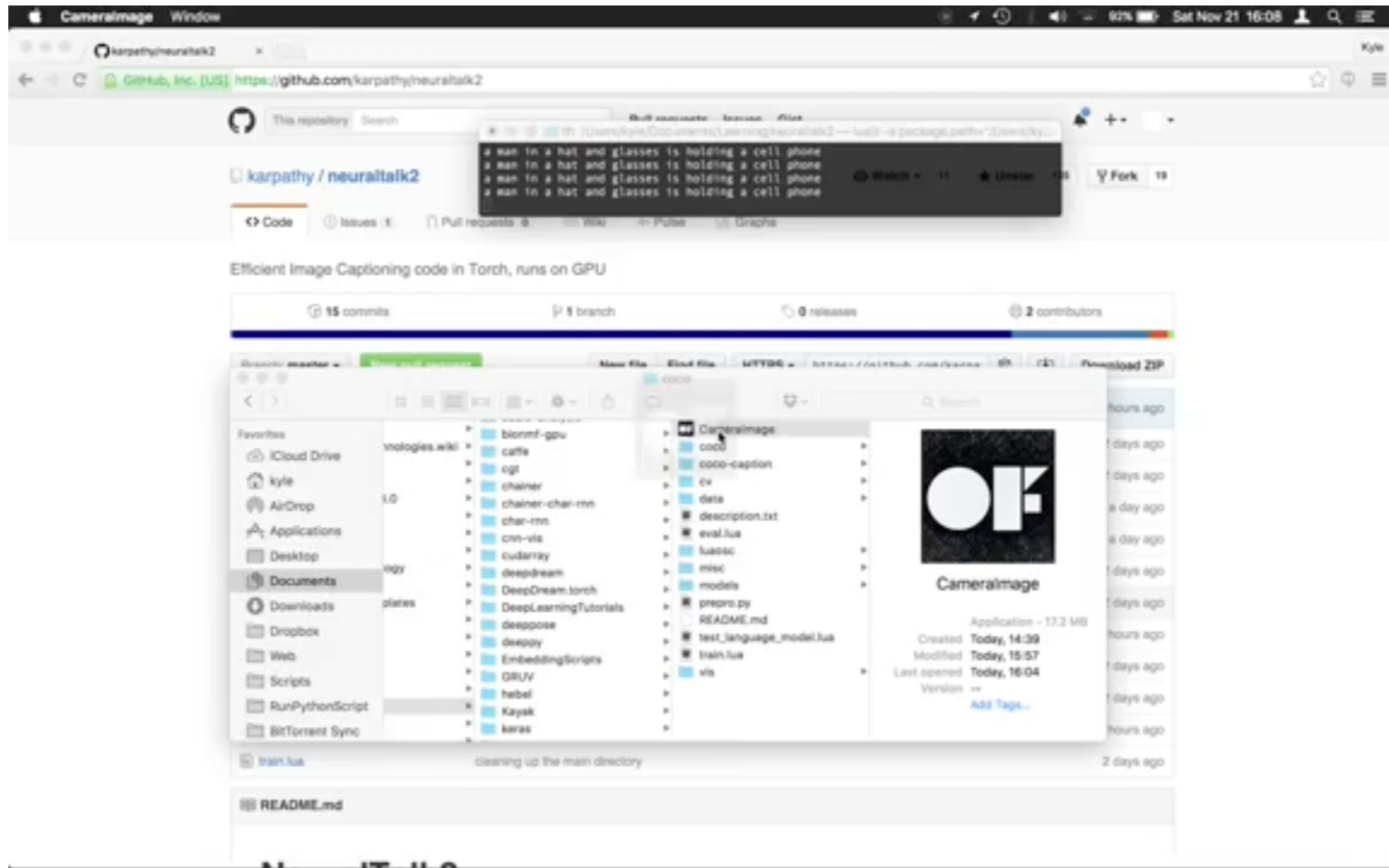
- Renouveau de l'IA par l'empirisme: apprentissage statistique.
- Pour les mathématiques: de la **logique** à la **géométrie**.
- Intégration avec les autres « sciences des données » : physique,...



- 1990: Approches statistiques: chaines de Markov
- 2010: Réseaux de neurones profonds
- In computer science, artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals. Computer science defines AI research as the study of "intelligent agents ».
- En informatique, l'intelligence artificielle (IA), parfois appelée intelligence machine, est une intelligence démontrée par des machines, par opposition à l'intelligence naturelle affichée par l'homme et d'autres animaux. L'informatique définit la recherche en intelligence artificielle comme l'étude des "agents intelligents".

Raconter une vidéo

Interactions multimodales



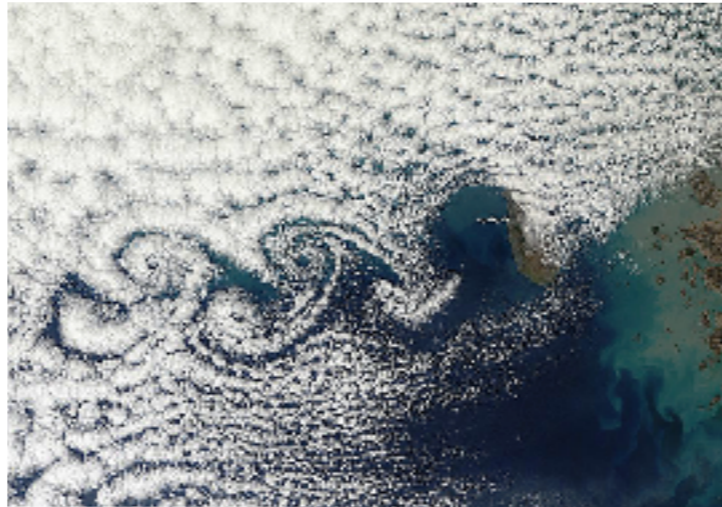
NeuralTalk and Walk, recognition, text description of the image while walking



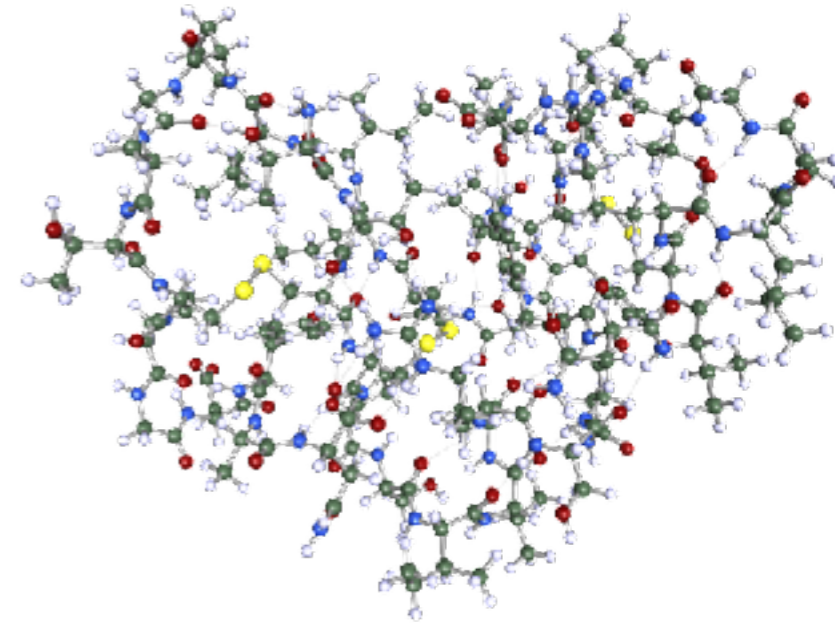
Astrophysique



Dynamique des fluides



Chimie Quantique

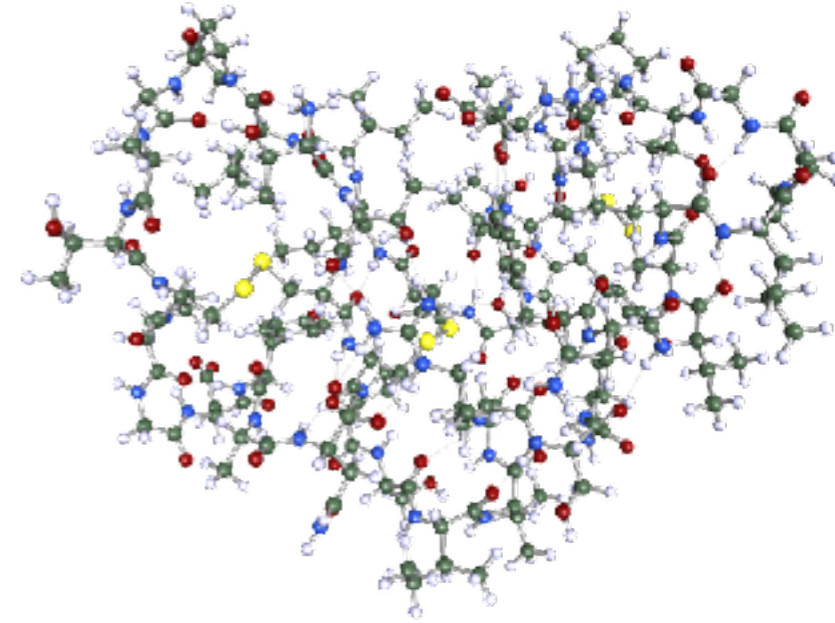


- Lois de la physique connues: équations de Newton, Boltzmann, Maxwell, Navier-Stokes, Schrodinger... mais calcul difficile des solutions résultant d'interactions d'un grand nombre de « corps élémentaires ».
- Simulations numériques: à partir des équations fondamentales.
- Peut on prédire une solution par régression à partir d'une base de données de solutions et quelques informations à priori ?
De mieux en mieux avec des réseaux de neurones profonds...

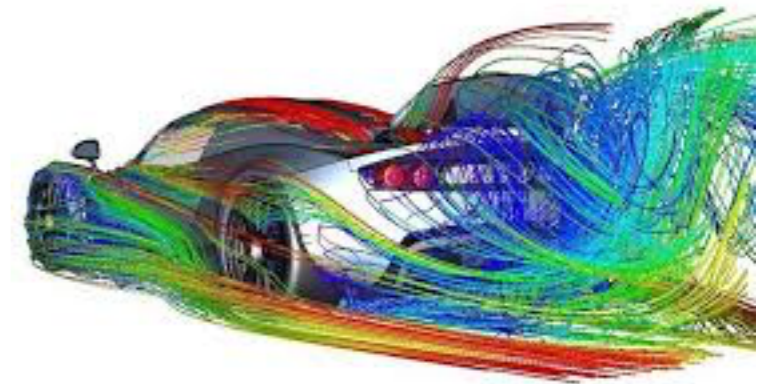
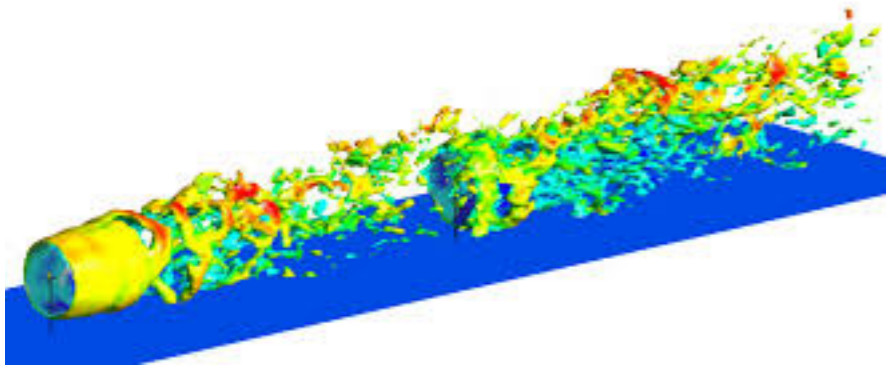


Applications Industrielles

Nouveaux matériaux, molécules pharmaceutiques:



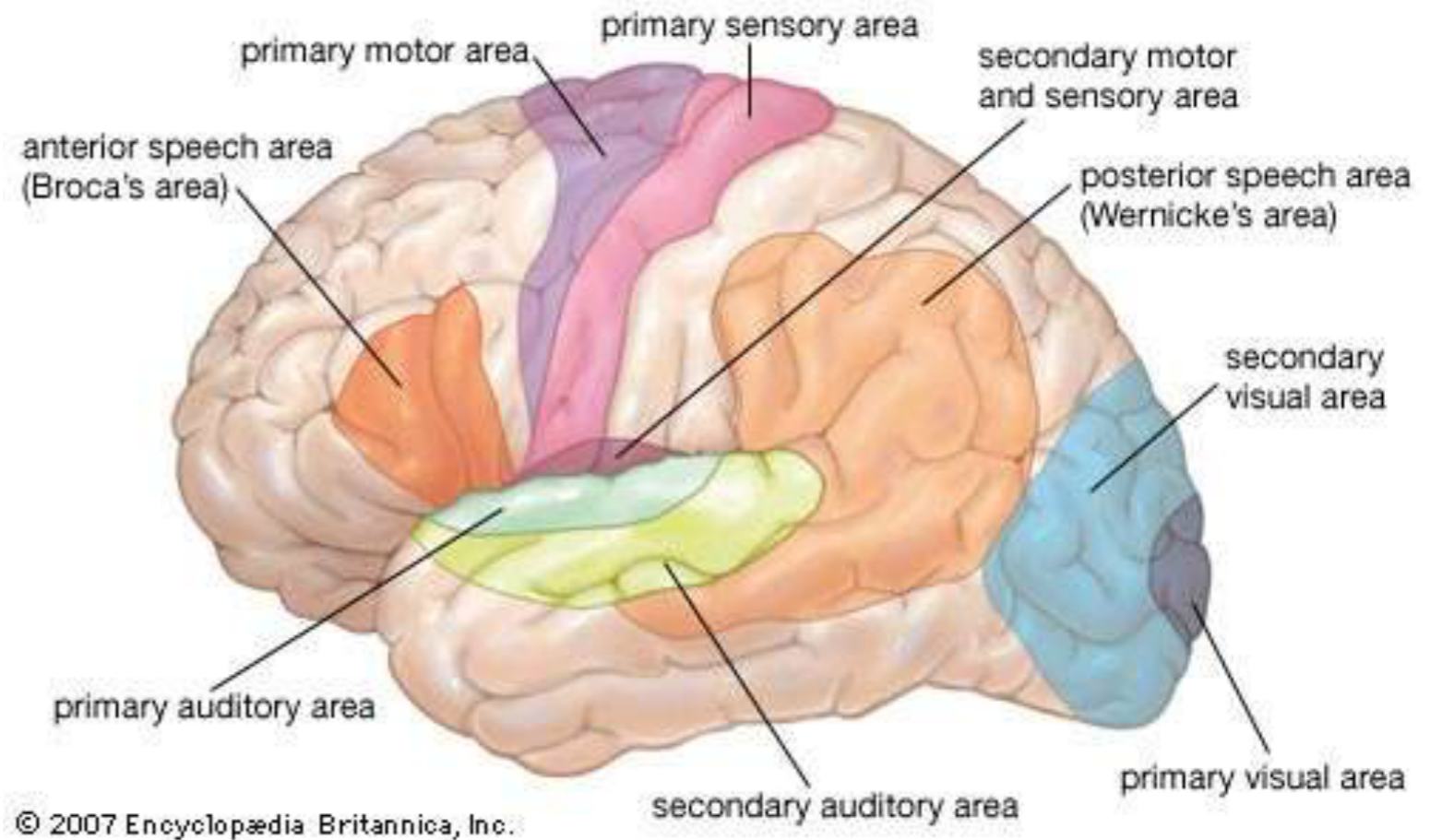
Simulations numériques rapide ?



Neurone



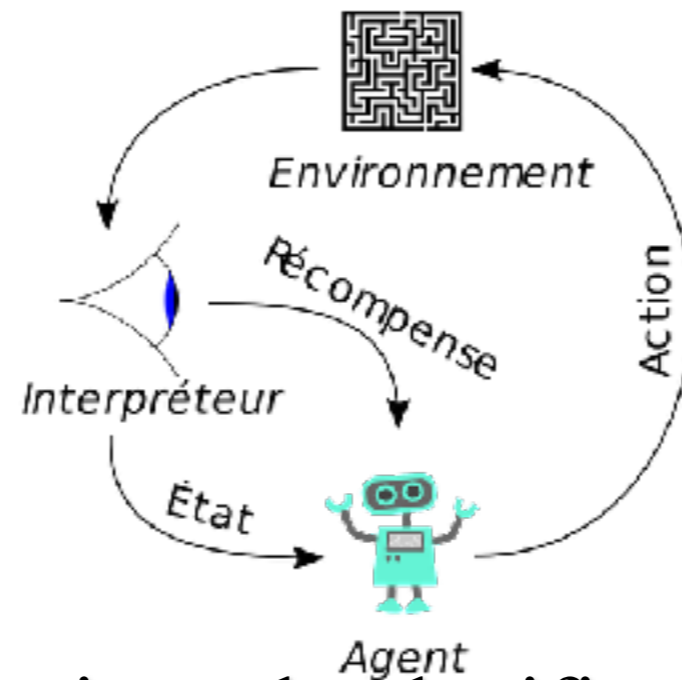
Cortex



- Quelle similarité entre réseaux de neurones artificiels et biologie ? traitements, niveaux fonctionnels, modèles...
- Retour de l'Intelligence Artificielle

Apprentissage par Renforcement

- L'agent apprend à choisir ses actions pour optimiser une récompense au cours du temps. Apprentissage par « essais et erreurs »



- Pour le contrôle en robotique, la planification, les jeux (AlphaGo)



- Créativité et structures: nouvelles stratégies aux échecs



COLLÈGE
DE FRANCE
—1530—

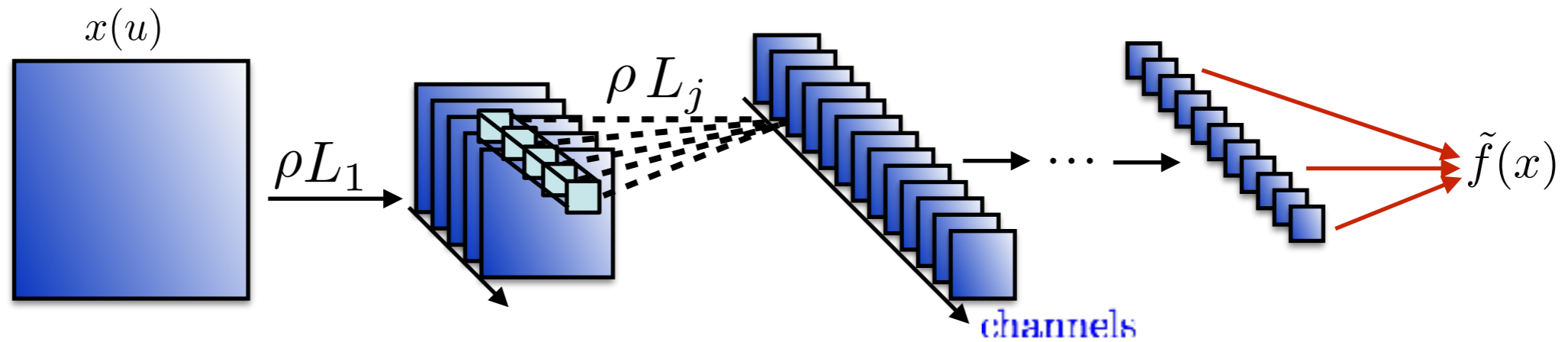
Apprentissage Robotique

Deep Sensorimotor Learning

rll.berkeley.edu/deeplearningrobotics

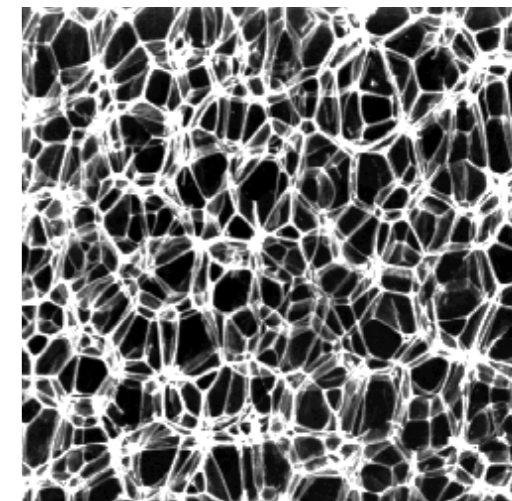
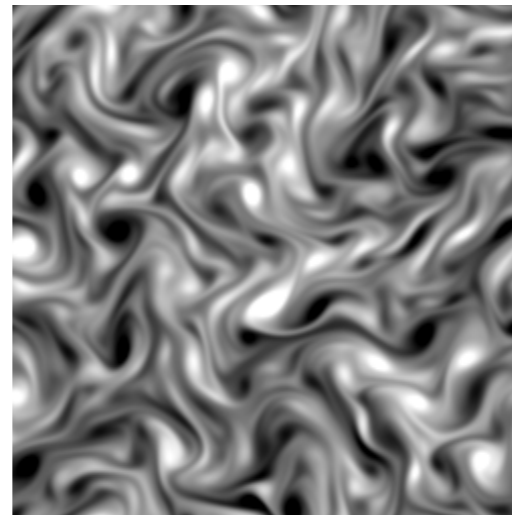
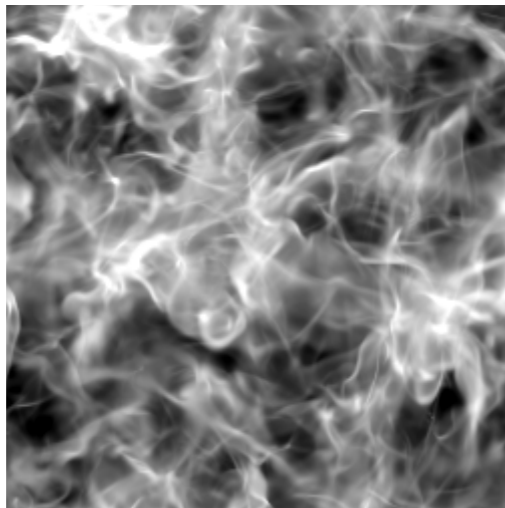
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

Apprentissage Non Supervisé

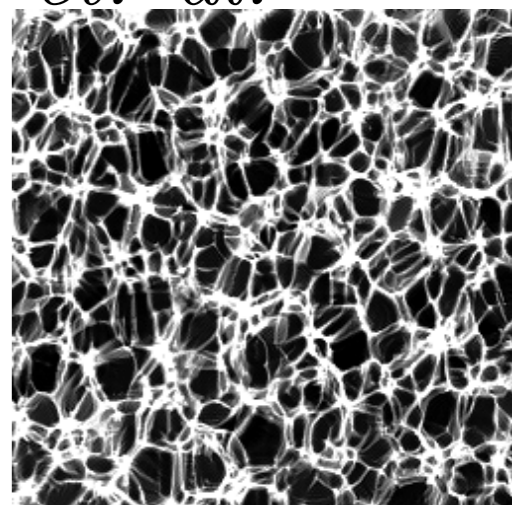
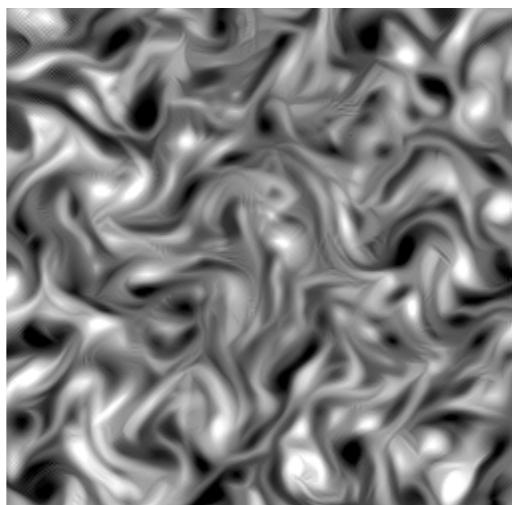
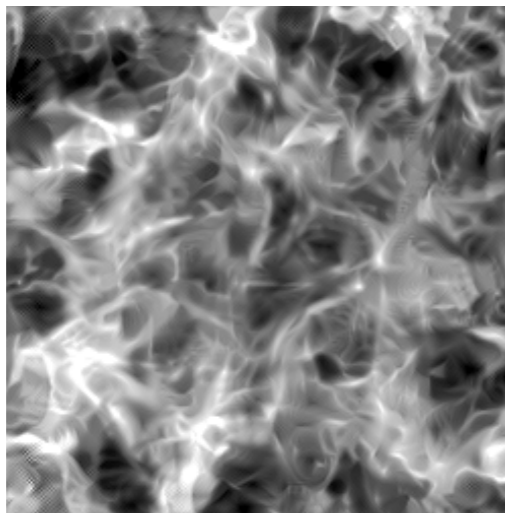


Modèles à partir de corrélations de coefficients du réseau:

x
 $d = 6 \cdot 10^4$



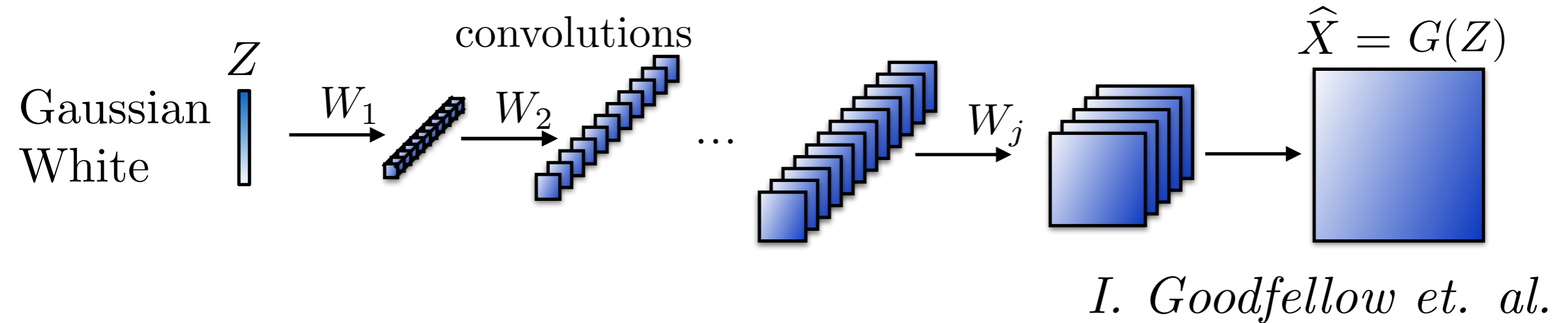
M. Bethdge et. al.



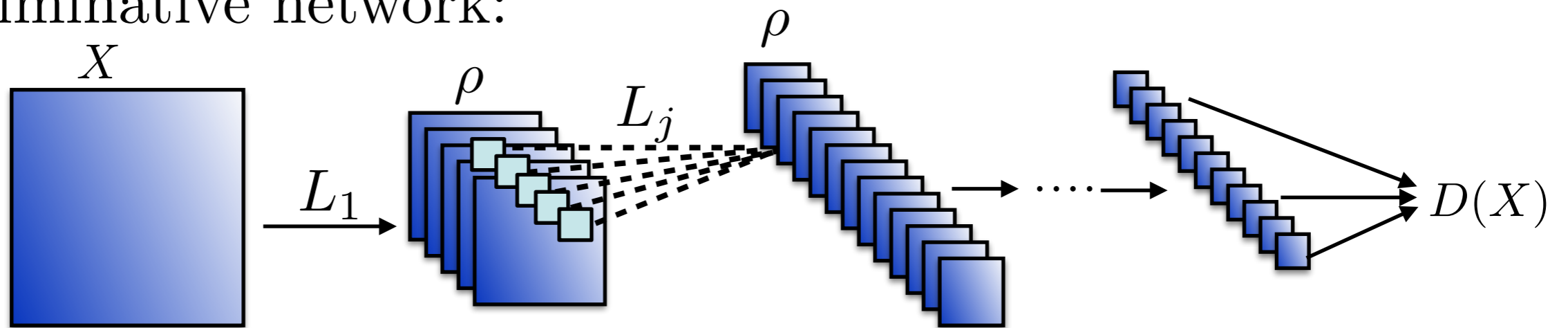
\tilde{x}
 $d' = 2 \cdot 10^5$
number of
correlation
coefficients

Generative Adversarial Networks

- Generative network for non-stationary processes:



- Discriminative network:



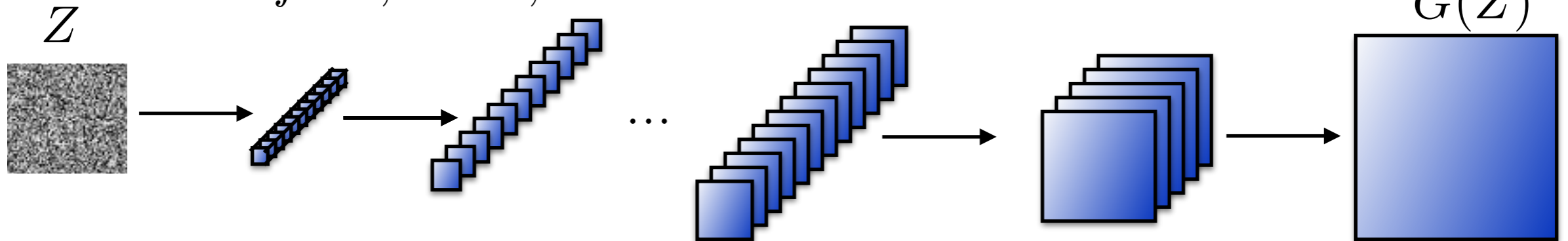
The generator G and discriminator D are optimized by:

$$\min_G \max_D \mathbb{E}[\log D(X)] + \mathbb{E}[\log(1 - D(G(Z)))]$$



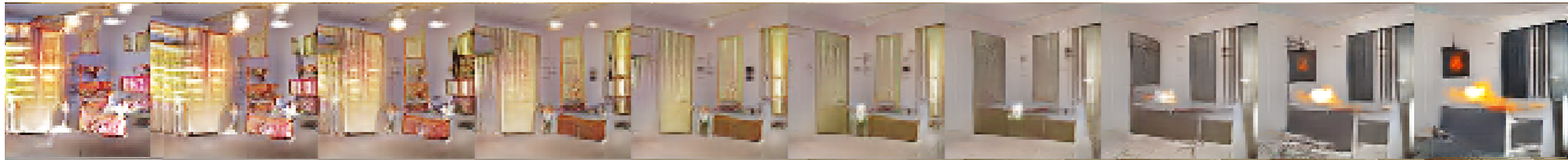
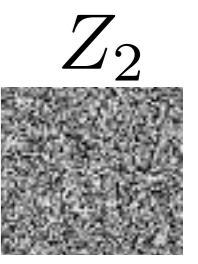
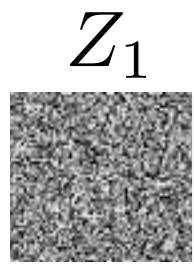
Generative Networks

Radford, Metz, Chintala



Network trained on bedroom images:

$$Z = \alpha Z_1 + (1 - \alpha) Z_2$$

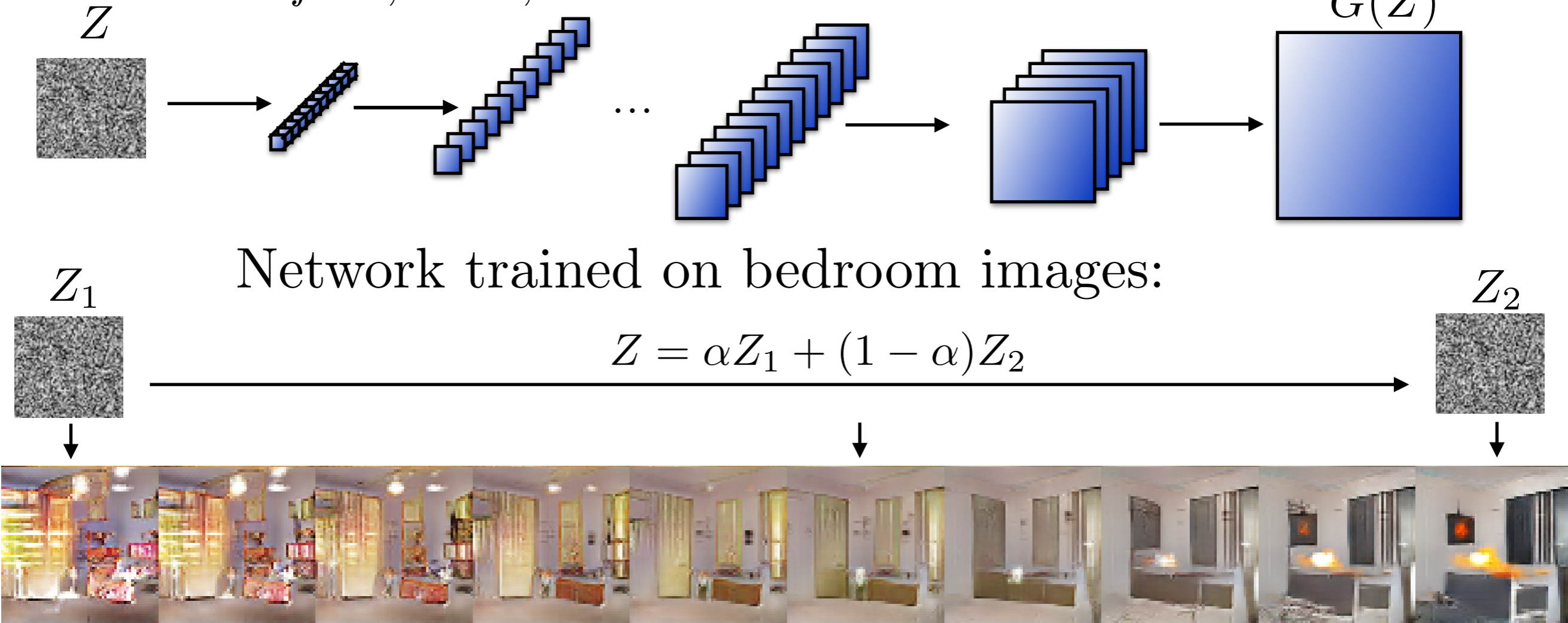


$G(Z_1)$ Linear interpolations are mapped to deformations $G(Z_2)$



Generative Networks

Radford, Metz, Chintala

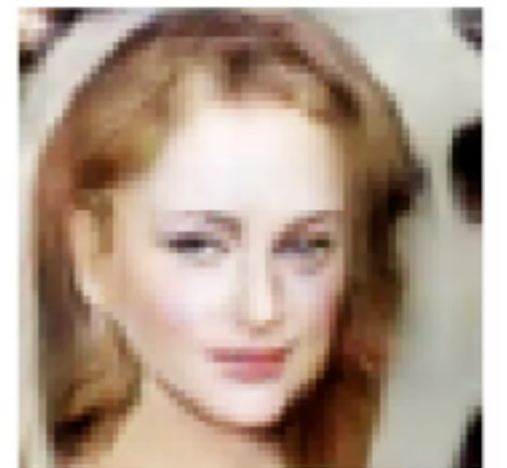


Bedrooms are mapped in a linear space

Network trained on faces of celebrities:

$G(Z)$

What kind of memory ?

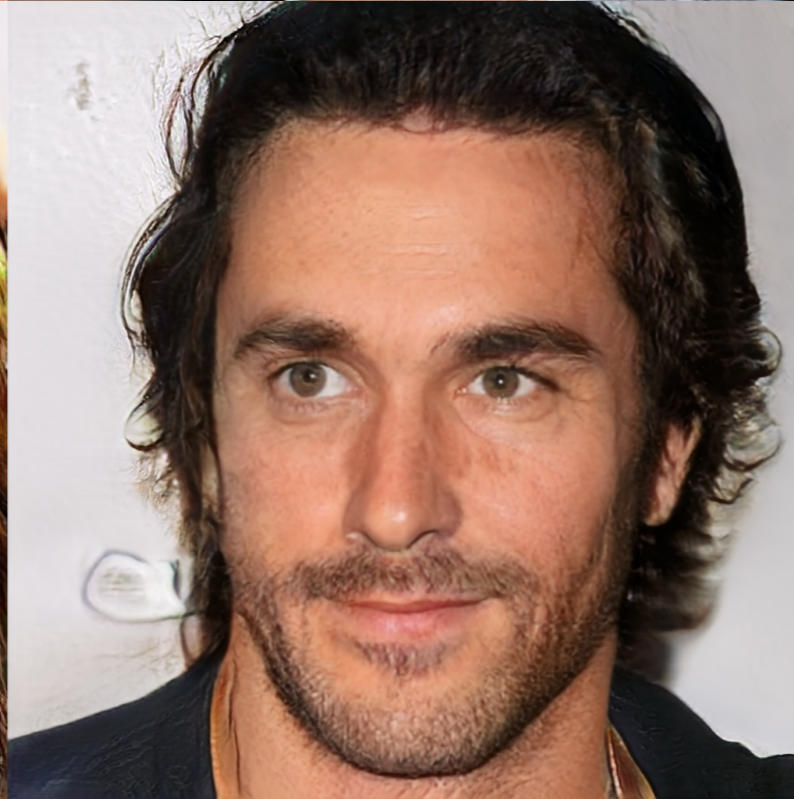
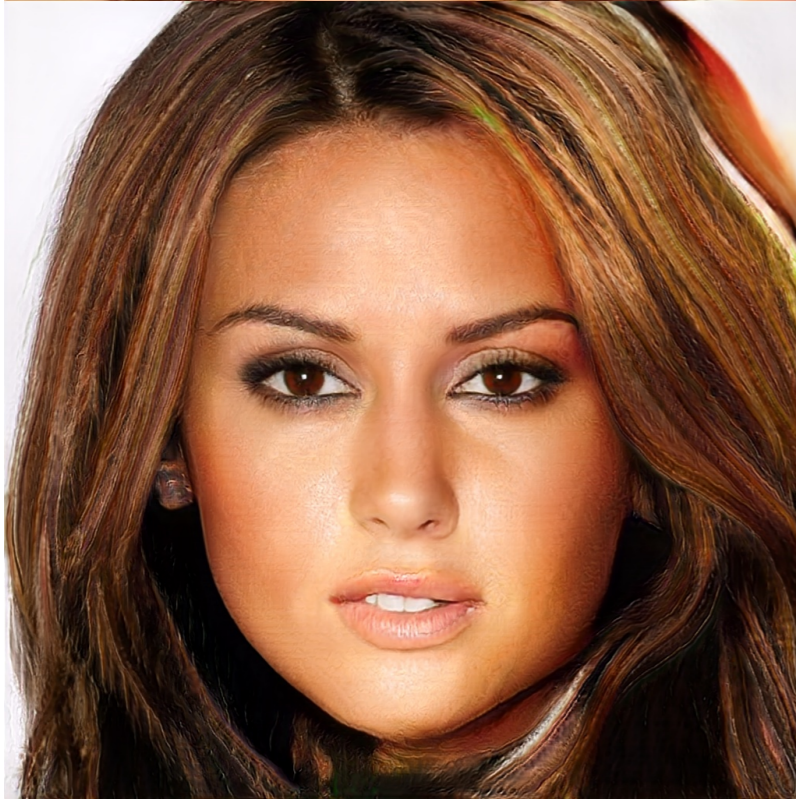




High Resolution Generation

T. Karras, T. Aila, S. Laine, J. Lehtinen

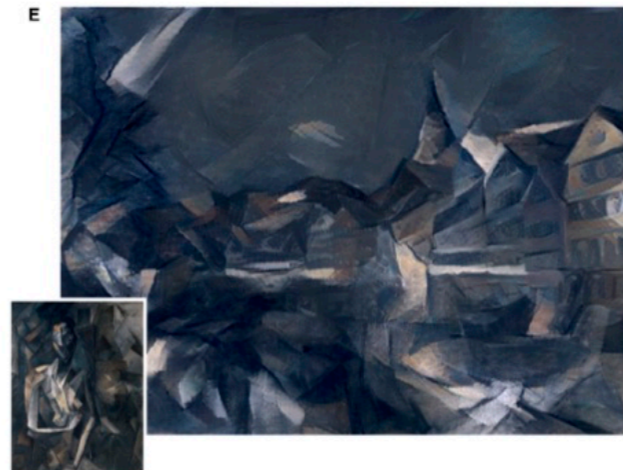
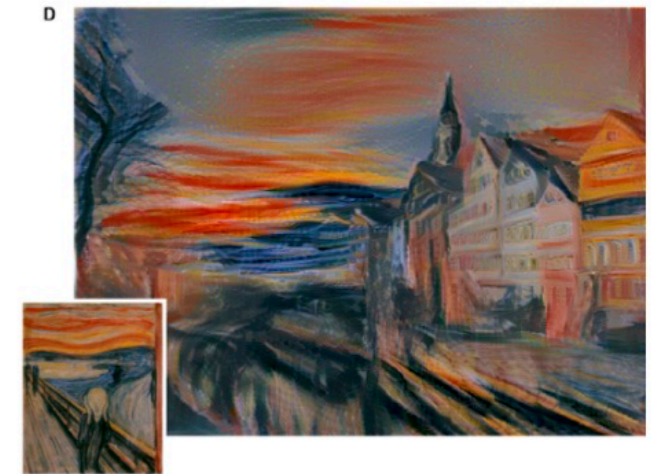
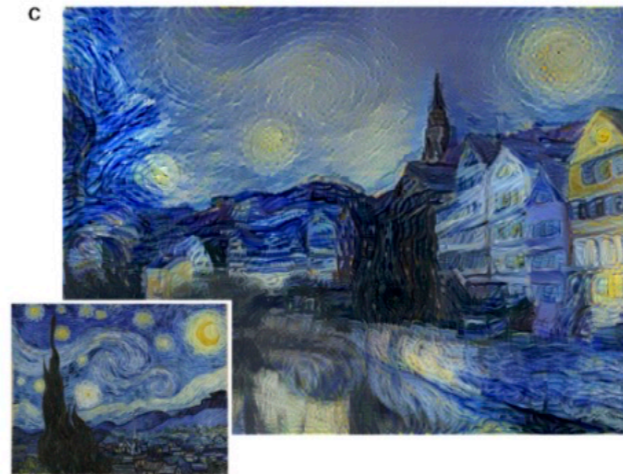
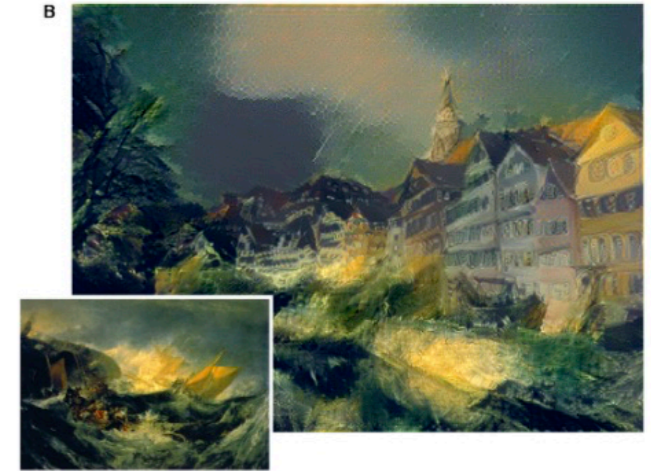
Generated from Hollywood celebrities data basis





Modifications de Styles

Gatys, Ecker, Bethdge, 2015



- Liens avec le monde de l'art
- Questions de créativité

B The Shipwreck of the Minotaur by J.M.W. Turner, 1805. **C** The Starry Night by Vincent van Gogh, 1889. **D** Der Schrei by Edvard Munch, 1893. **E** Femme nue assise by Pablo Picasso, 1910. **F** Composition VII by Wassily Kandinsky, 1913.



Limites et Opportunités

- Domaine algorithmique et essentiellement empirique
- Besoin de beaucoup de données labélisées (x_i, y_i) pour apprendre
- On ne comprend pas les performances de ces algorithmes, leurs capacité de généralisation et leurs limites
- Quelles « structures » sont apprises, quelles régularités ?
- Pas de contrôle « a priori » sur les erreurs: tests statistiques souvent biaisés.
- Architectures optimisées empiriquement: long et couteux

MAIS

- Approche générique de problèmes très différents
- Perspectives scientifiques et d'applications considérables
- Beaucoup d'opportunités de recherche: cela va très vite