

# **TOUTELIA 2021 : Statistical Physics, Probability and AI**

## **Rapport sur les contributions**

ID de Contribution: 1

Type: **Non spécifié**

## Talk by Yann Ollivier

*jeudi 16 décembre 2021 14:00 (1h 30m)*

This is the conference talk. Yann Ollivier will also give a colloquim talk the next day.

Abstract: Markov decision processes are a model for several artificial intelligence problems, such as games (chess, Go...) or robotics. At each timestep, an agent has to choose an action, then receives a reward, and then the agent's environment changes (deterministically or stochastically) in response to the agent's action. The agent's goal is to adjust its actions to maximize its total reward. In principle, the optimal behavior can be obtained by dynamic programming or optimal control techniques, although practice is another story.

Here we consider a more complex problem: learn all optimal behaviors for all possible reward functions in a given environment. Ideally, such a "controllable agent" could be given a description of a task (reward function, such as "you get +10 for reaching here but -1 for going through there") and immediately perform the optimal behavior for that task. This requires a good understanding of the mapping from a reward function to the associated optimal behavior.

We will present our recent theoretical and empirical results in this direction. There exists a particular "map" of a Markov decision process, on which near-optimal behaviors for all reward functions can be read directly by an algebraic formula. Moreover, this "map" is learnable by standard deep learning techniques from random interactions with the environment.

ID de Contribution: 2

Type: **Non spécifié**

## COLLOQUIUM by Yann Ollivier

*vendredi 17 décembre 2021 14:00 (1 heure)*

Résumé : “Les problèmes de raisonnement inductif ou d’extrapolation comme “deviner la suite d’une série de nombres”, ou plus généralement, “comprendre la structure cachée dans des observations”, sont fondamentaux si l’on veut un jour construire une intelligence artificielle. On a parfois l’impression que ces problèmes ne sont pas mathématiquement bien définis. Or il existe une théorie mathématique rigoureuse du raisonnement inductif et de l’extrapolation, basée sur la théorie de l’information. Cette théorie est très élégante, mais difficile à appliquer.

En pratique aujourd’hui, ce sont les réseaux de neurones qui donnent les meilleurs résultats sur toute une série de problèmes concrets d’induction et d’apprentissage (vision, reconnaissance de la parole, récemment le jeu de Go ou les voitures sans pilote...) Je ferai le point sur quelques-uns des principes mathématiques sous-jacents et sur leur lien avec la théorie de l’information.”

ID de Contribution: 3

Type: **Non spécifié**

## Talk by Marc Lelarge

*vendredi 17 décembre 2021 10:30 (1h 30m)*

Abstract: Geometric deep learning is an attempt for geometric unification of a broad class of machine learning problems from the perspectives of symmetry and invariance. In this talk, I will present some advances of geometric deep learning applied to combinatorial structures. I will focus on various classes of graph neural networks that have been shown to be successful in a wide range of applications with graph structured data.

ID de Contribution: 4

Type: **Non spécifié**

## Talk by Franck Gabriel

*vendredi 17 décembre 2021 09:00 (1h 30m)*

In recent years, randomness has become more important in machine learning. Through two examples, we will see that it can be used to [1] “select” well-behaved regions of parameters and [2] provide an easier optimization problem.

[1] In deep learning, I will present the NTK regime, where, by considering a “wide” random initialization, it can be shown that neural networks with large width converge, and the dynamics of the output function can be described.

[2] In kernel methods, instead of looking for an optimal function in the RKHS, one can look for an optimal function in a random vector space: this is the random feature method. After explaining why it provides an approximation to kernel methods, I will present the implicit bias that finite sampling induces on the output function.

ID de Contribution: 5

Type: **Non spécifié**

## Talk by Sandrine Péché

*jeudi 16 décembre 2021 15:30 (1 heure)*

We consider a random matrix model  $M = YY^*$  where  $Y = (f(WX)_{ij})$  and  $W$  and  $X$  are large rectangular matrices with iid entries. The function  $f$  is called the activation function in certain neural networks.

Pennington and Worah have identified the empirical eigenvalue distribution of such random matrices in the Gaussian case ( $W$  and  $X$ ). We extend their result to a wider class of distributions for a certain class of activation functions.

This is joint work with Lucas Benigni.