

Learning And Understanding Strategies On Zero-Sum Games

José Lucas De Melo Costa^{1,2}, Jonathan Poli¹, Manh-Bao Nguyen¹, Aymeric Ducatez¹, Maxime Bouscary¹, Benjamin Boucher¹

¹ CentraleSupélec, Université Paris-Saclay

² Université de São Paulo

jose-lucas.de-melo-costa@student-cs.fr, jonathan.poli@student-cs.fr

Results given by artificial intelligence (AI) and neural networks are currently hard to fully explain. This is the case, for example, of learning algorithms applied to zero-sum games, where one player's gain is equivalent to another's loss, like Chess, Go and TicTacToe. For these games, the AIs get impressive results without clearly explaining their decisions. We propose methods and algorithms to explain given strategies.

One of the best AI at Chess and Go is still AlphaZero [1], developed by DeepMind in late 2017. We have adapted the algorithm to play other zero-sum games (Figure 1). This way, there is a very complex, dense neural network, i.e. several artificial neurons in successive layers, which enables to learn many statistical models by applying very basic operations. It can take an input, such as a board of a game, and return an output, such as the best move to play. Therefore, an analysis of the network's structure and parameters will enable us to infer what is important in the game for the program to return a decision. [2]

In addition, we use a model called *Transformer* [3] to propose both a new learning method for those games and a way to interpret the results. Indeed, Transformers, designed in 2017, use an attention mechanism, which enables to explicitly choose relevant information and to understand relations inside the data. This model, mainly used for Natural Language Processing, i.e. process of textual information, is then adapted to the field of learning to play games.

Keywords : Zero-Sum Games, Learning, Interpretability, Neural Networks

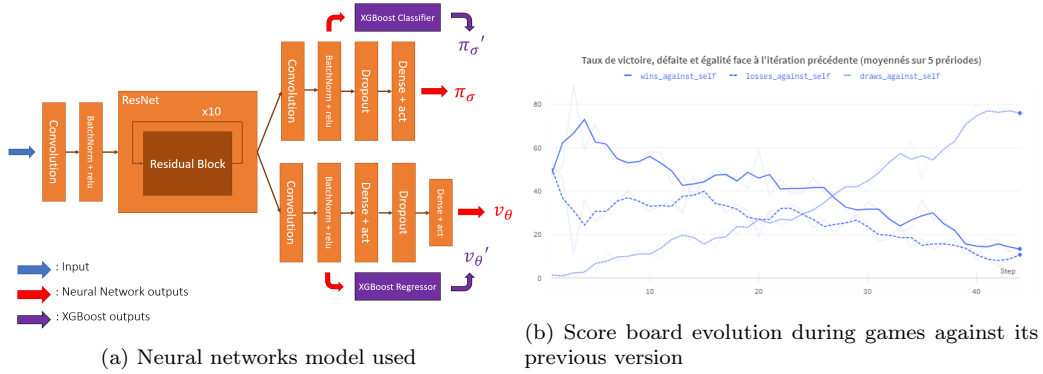


FIGURE 1 – Training on Connect 4 of a deep complex neural network which plays against its previous version to assess its progress.

Supervisors : Arpad Rimmel (Researcher at LRI), Johanne Cohen (Research Director at LRI)

References

- [1] D.Silver, T.Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan and D. Hassabis *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*, 2017, DeepMind.
- [2] T. McGrath, A. Kaphishnikov, N. Tomašev, A. Pearce, D. Hassabis, B. Kim, U. Paquet and V. Kramnik *Acquisition of Chess Knowledge in AlphaZero*, 2021.
- [3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin *Attention is All You Need*, 2017.

Apprentissage Et Interprétabilité Des Stratégies Sur Les Jeux A Somme Nulle

José Lucas De Melo Costa^{1,2}, Jonathan Poli¹, Manh-Bao Nguyen¹, Aymeric Ducatez¹, Maxime Bouscary¹, Benjamin Boucher¹

¹ CentraleSupélec, Université Paris-Saclay

² Université de São Paulo

jose-lucas.de-melo-costa@student-cs.fr, jonathan.poli@student-cs.fr

Les résultats donnés par l'intelligence artificielle (IA) et les réseaux neuronaux sont actuellement difficiles à complètement expliquer. C'est le cas, par exemple, des algorithmes d'apprentissage appliqués aux jeux à somme nulle, où le gain d'un joueur équivaut à la perte d'un autre, comme les échecs, le Go et le Morpion. Pour ces jeux, l'IA obtient des résultats impressionnants sans expliquer clairement ses décisions. Nous proposons des méthodes et des algorithmes pour expliquer les stratégies données.

L'une des meilleures IA aux échecs et au Go reste AlphaZero [1], développée par DeepMind fin 2017. Nous avons adapté l'algorithme pour jouer à d'autres jeux à somme nulle (Figure 2). On a ainsi un réseau de neurones très complexe et dense, soit plusieurs neurones artificiels en couches successives, capable d'apprendre de nombreux modèles statistiques en appliquant des opérations très basiques. Il peut prendre une entrée, comme le plateau d'un jeu, et renvoyer une sortie, comme le meilleur coup à jouer. Ainsi, une analyse de la structure du réseau et de ses paramètres nous permettra d'explicitier ce qui est important dans le jeu pour que le programme retourne une décision. [2]

De plus, nous utilisons un modèle appelé *Transformer* [3] pour proposer à la fois une nouvelle méthode d'apprentissage pour ces jeux et une façon d'interpréter les résultats. En effet, les Transformers, conçus en 2017, utilisent un mécanisme d'attention, capable de choisir explicitement les informations et les liens pertinents dans les données. Ce modèle, principalement utilisé pour le traitement d'informations textuelles, est ainsi adapté au domaine de l'apprentissage des jeux.

Mots-clés : Jeux à somme nulle, Apprentissage, Interprétation, Réseaux Neuronaux

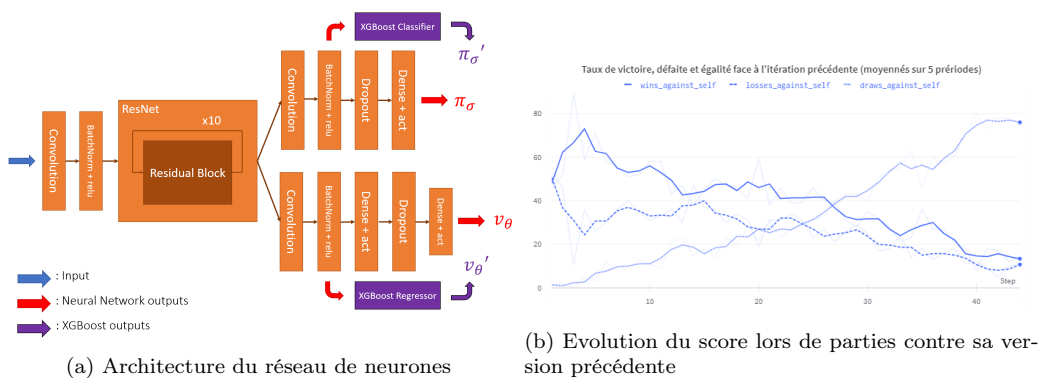


FIGURE 2 – Entraînement sur le Puissance 4 d'un réseau complexe de neurones jouant contre sa version antérieure afin d'évaluer sa progression par rapport à sa dernière version.

Encadrants : Arpad Rimmel (Chercheur au LRI), Johanne Cohen (Directrice de recherche au LRI)

Références

- [1] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan and D. Hassabis *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*, 2017, DeepMind.
- [2] T. McGrath, A. Kaphishnikov, N. Tomašev, A. Pearce, D. Hassabis, B. Kim, U. Paquet and V. Kramnik *Acquisition of Chess Knowledge in AlphaZero*, 2021.
- [3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin *Attention is All You Need*, 2017.