

# Using multistability to solve fading memory problems in reinforcement learning

**Florent De Geeter**

*Supervisor:* Prof. Guillaume Drion

MSc in Computer Science and Engineering

*Academic year 2020-2021*

## Abstract

This thesis focuses on two important topics in machine learning. First, recurrent neural networks (RNN), which are neural networks with memory, i.e. which can propagate information through the time. Second, reinforcement learning (RL), a machine learning technique in which agents have to learn to perform the best actions by interacting with environments. The goal of the thesis is to use RNNs in a reinforcement learning setting to learn to retain information which will be useful in the future for selecting the best actions.

More specifically, this thesis studies the benefits of using RNN agents whose memories are based on multistability when dealing with environments that require a long-lasting memory. Indeed, the memory of RNNs usually fades with time because it relies on their dynamics, but it is no longer true when multistability is used.

The goal of this paper is twofold: first we show the interest of the multistability-based memory by testing a new type of RNN called the nBRC built for being multistable on a RL environment specifically made to require a long-lasting memory. This new RNN shows a great generalization capability thanks to its memory: by only being trained on a small version of the environment, it is able to generalize its knowledge and to play correctly on longer versions.

The second objective of this thesis is to introduce and test a pretraining algorithm, called the *multistability warmup*, which is supposed to force a RNN to become multistable. This algorithm is applied to several types of RNNs, including the well-known GRU and LSTM. Then these RNNs are tested on the same environment as previously, and we observe that their results are improved, especially for GRU which is able to compete with the nBRC.

Finally, we discuss about what could be done next. On one hand, multistability gave us great observations, and it could be very interesting to test it on more complex environments to see what it has to offer. On the other hand, the *multistability warmup* could still be improved, as it does not work very well with each type of RNN.

# Utiliser la multistabilité pour résoudre des problèmes de mémoire qui s'estompe dans l'apprentissage par renforcement

Florent De Geeter

Promoteur: Prof. Guillaume Drion

Master en ingénieur civil en informatique

Année académique 2020-2021

## Résumé

Cette thèse se concentre sur deux sujets importants de l'apprentissage automatique. Premièrement, les réseaux de neurones récurrents (RNN), qui sont des réseaux de neurones avec mémoire, c'est-à-dire qui sont capables de propager de l'information dans le temps. Deuxièmement, l'apprentissage par renforcement (RL), une technique d'apprentissage automatique dans laquelle des agents doivent apprendre à effectuer les meilleures actions en interagissant avec des environnements. L'objectif de cette thèse est d'utiliser les RNNs dans un contexte d'apprentissage par renforcement pour apprendre à retenir des informations qui leur seront utiles dans le futur pour sélectionner les meilleures actions.

Plus précisément, cette thèse étudie les avantages de l'utilisation d'agents basés sur des RNNs dont la mémoire utilise la multistabilité dans des environnements qui nécessitent une mémoire durable. En effet, la mémoire des RNNs s'estompe généralement avec le temps car elle repose sur leur dynamique, mais ce n'est plus vrai lorsque la multistabilité est utilisée.

L'objectif de cette thèse est double: tout d'abord nous montrons l'intérêt de la mémoire basée sur la multistabilité en testant un nouveau type de RNN appelé nBRC construit pour être multistable dans un environnement RL spécifiquement conçu pour nécessiter une mémoire durable. Ce nouveau RNN montre une grande capacité de généralisation grâce à sa mémoire: en étant seulement entraîné sur une petite version de l'environnement, il est capable de généraliser et de jouer correctement sur des versions beaucoup plus longues.

Le deuxième objectif de cette thèse est d'introduire et de tester un algorithme de pré-entraînement, appelé *multistability warmup* ("échauffement de multistabilité"), qui est censé forcer un RNN à devenir multistable. Cet algorithme est appliqué à plusieurs types de RNNs, dont les bien connus GRU et LSTM. Ensuite, ces RNNs sont testés sur le même environnement que précédemment, et nous observons que leurs résultats sont améliorés, en particulier pour le GRU qui est capable de rivaliser avec le nBRC.

Enfin, nous discutons de ce qui pourrait être fait ensuite. D'une part, la multistabilité nous a donné de grandes observations, et il pourrait être très intéressant de la tester sur des environnements plus complexes pour voir ce qu'elle a à offrir. D'autre part, la *multistability warmup* pourrait encore être améliorée, car elle ne fonctionne pas très bien avec chaque type de RNN.