Facultade de Informática da Coruña

UNIVERSIDADE DA CORUÑA

TRABALLO DE FIN DE GRAO
GRAO EN INXENARÍA INFORMÁTICA
MENCIÓN EN COMPUTACIÓN

IMPLEMENTATION OF AN AUTOMATIC CAMERA OPERATOR USING DEEP REINFORCEMENT LEARNING

Student: Rodriguez Louzán, Adrián
Directors: Francisco Javier Taibo Pena
Emilio José Padrón González
Luis Omar Álvarez Mures

A Coruña 13th November of 2020

To my family and my brother

Acknowledgements

First of all, I would like to thank my family for supporting me unconditionally and placing all their trust in me.

Thanks also to my friends, those that I made during this 4 years, but even more to those who I met during my internation experience in South Korea.

Finally, I want to give my gratitude to my project managers Francisco Javier Taibo Pena, Emilio José Padrón Gonzalez and Luis Omar Álvarez Mures, without whom I could not have carried out this project. Thank you for all your attention.

Abstract

Reinforcement learning (RL) is an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward.

The problems of interest in reinforcement learning have also been studied in the theory of optimal control, which is concerned mostly with the existence and characterization of optimal solutions, and algorithms for their exact computation, and less with learning or approximation, particularly in the absence of a mathematical model of the environment. Reinforcement algorithms that incorporate deep learning can beat world champions at the game of Go as well as human experts playing numerous Atari video games.

In this project, we will leverage Deep Reinforcement Learning and game engines to model a typical sports scene and teach an agent to capture the action on camera. First, a scene which resembles a sports match will be created using the chosen game engine. Next, we will test out different Deep Reinforcement Learning algorithms.

Resumen

El aprendizaje por refuerzo (AR) es un área del aprendizaje automático que se ocupa de cómo los agentes software deben tomar acciones en un entorno para maximizar alguna noción de recompensa acumulativa.

Los problemas de interés en el aprendizaje por refuerzo también se han estudiado en la teoría del control óptimo, que se ocupa principalmente de la existencia y caracterización de soluciones óptimas y algoritmos para su cálculo exacto, y menos con aprendizaje o aproximación, particularmente en ausencia de un modelo matemático del medio ambiente. Los algoritmos de refuerzo que incorporan el aprendizaje profundo pueden vencer a los campeones mundiales en el Juego de Go, así como expertos humanos que juegan numerosos videojuegos de Atari.

En este proyecto, aprovecharemos el aprendizaje de refuerzo profundo y los motores de juego para modelar una escena deportiva típica y enseñarle a un agente a capturar la acción con la cámara. Primero, una escena que se asemeja a un partido deportivo será creado utilizando el motor de juego elegido. A continuación, probaremos diferentes algoritmos de aprendizaje por refuerzo profundo.

Keywords

- Artificial Intelligence
- Deep Reinforcement Learning
- Camera operator
- DQN

- Python
- Tensorflow
- Keras

Index

1   Introduction.

Deep Reinforcement Learning (DRL), a very fast-moving field, is the combination of Reinforcement Learning and Deep Learning and it is also the most trending type of Machine Learning at this moment because it is being able to solve a wide range of complex decision-making tasks that were previously out of reach for a machine to solve real-world problems with human-like intelligence.

In this introduction I will try to explain the fundamentals and concepts around Artificial intelligence and Machine Learning, in order to then deeply dive into our program and how I implemented it.

1.1   Artificial Intelligence and Machine Learning

We can divide Artificial Intelligence into two categories **Narrow AI** and **General A.I.**
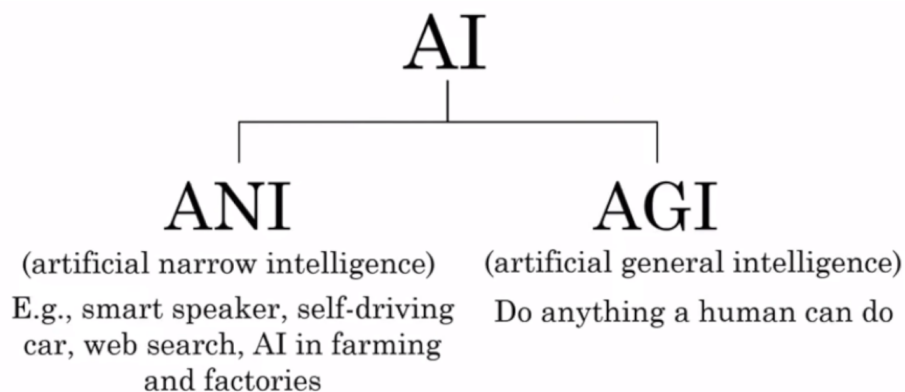

Figure 1: Categories of Artificial Intelligence

Most of the movies and representations that we see in the mainstream media are around General A.I. Machines that do everything that humans do, and better. It can talk, it can learn to play games, it can communicate... etc

Right now, in our field we are working on narrow A.I. Are machines that can do only one thing and the application of the machine is narrowed to a specific goal. In this field and in the recent years, great advances were made.For instance, AlphaGo defeated the best professional human player in the game of Go. Or last year, for instance, our friend Oriol Vinyals and his team in DeepMind showed the AlphaStar agent beat professional players at the game of StarCraft II. Or a few months later, OpenAI?s Dota-2-playing bot became the first AI system to beat the world champions in an e-sports game.

Machine Learning [2] is one of the many approaches that Artificial Intelligence (AI) uses when it comes to extracting knowledge from experience, not only as a tool to achieve the objectives linked to AI but as a vehicle to reduce time and effort used for resolution. ML recognizes patterns through the use of examples in rather than addressing its implementation. The machine learns a model from examples and use this one to solve the problem. A system that continuously learns, that is capable of make decisions based on data instead of algorithms and that changes your behavior, it is a system based on ML. Given a problem the choice of the appropriate algorithm is not trivial. For this it is necessary ask yourself two fundamental questions: what do I want to do? and what information I have to achieve my goal? Both require us to carry out an exhaustive analysis of the own problem and the data, quantity and quality, that we have for its resolution.
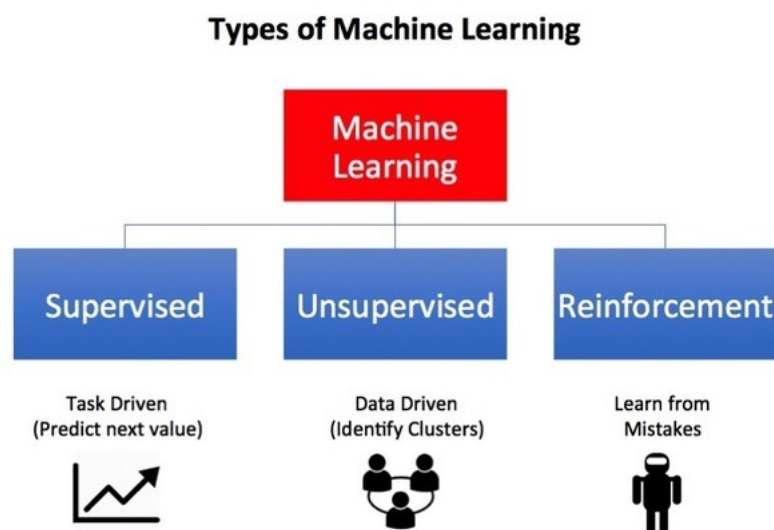

Figure 2: Types of Machine Learning

Reinforcement Learning is one of the three branches in which ML techniques are generally categorized:

- **Supervised Learning** is the task of learning from tagged data and its goal is to generalize.

- **Unsupervised Learning** is the task of learning from unlabeled data and its goal is to compress.

- **Reinforcement Learning** is the task of learning through trial and error and its goal is to act.

Orthogonal to this categorization we can consider a powerful recent approach to ML, called **Deep Learning (DL)**, topic of which we have discussed extensively in previous posts. DL is not a separate branch of ML, so it?s not a different task than those described above.

DL is a collection of techniques and methods for using neural networks to solve ML tasks, either Supervised Learning, Unsupervised Learning, or Reinforcement Learning and we can represent it graphically in the following figure:
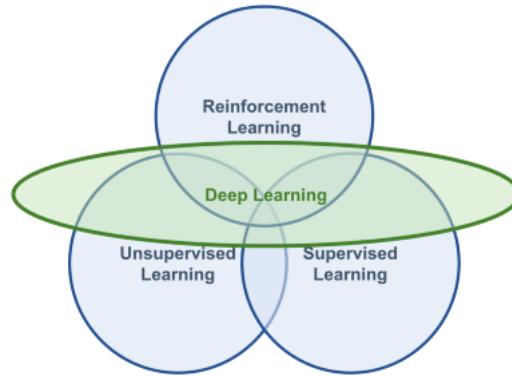

Figure 3: Visual relationship of Deep Learning with the Machine Learning categories.

Reinforcement learning [3] [1] is a form of machine learning in which an agent attempts to learn a policy that maximizes a numeric reward signal [1]. In reinforcement learning, an agent learns by trial and error and discovers optimal actions through its own experiences. Unlike supervised learning, the agent does not learn by comparing its own actions to those of an expert; everything it learns is from its own interactions with the environment. Reinforcement learning attempts to solve optimization problems that are defined by a Markov Decision Process (MDP) [1]. A A Markov Decision Process defines the behavior of the environment by mathematically defining the environment´s one step dynamics.

1.2   Elements in Reinforcement Learning

In Reinforcement Learning there are two core components:

- **An Agent**, that represents the "solution" , which is a computer program with a single role of making decisions (actions) to solve complex decision-making problems under uncertainty.

- **An Enviroment**, that is the representation of a "problem", which is everything that comes after the decision of the Agent. The environment responds with the consequences of those actions which are observations or states and rewards also sometimes called costs.

For example, let´s take the game of PacMan where the goal of the agent(PacMan) is to eat the food in the grid while avoiding the ghosts on its way. In this case, the grid world is the

interactive environment for the agent where it acts. Agent receives a reward for eating food and punishment if it gets killed by the ghost (loses the game). The states are the location of the agent in the grid world and the total cumulative reward is the agent winning the game.
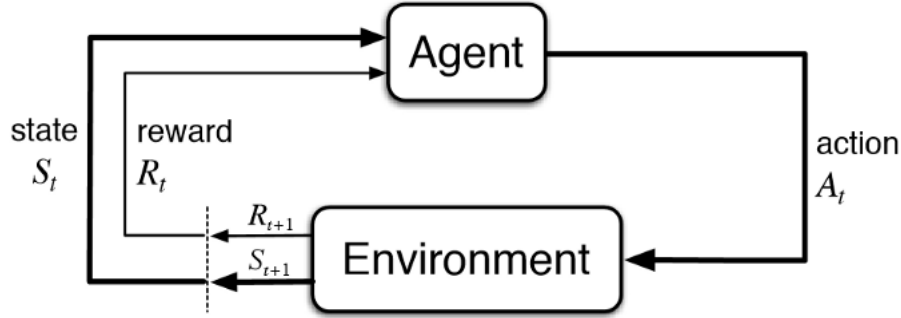

Figure 4: Reinforcement learning cycle

These two core components interact constantly in a way that the Agent attempts to make some changes in the Environment through actions, and the Environment changes due to the Agent?s actions. How the environment changes to certain actions is defined by a model which may or may not be known by the Agent.

When the Agent does not know the model, it needs to make decisions with incomplete information; do model-free RL or try to learn the model explicitly as part of the algorithm.

State The Environment is represented by a set of variables related to the problem. This set of variables and all the possible values that they can take are referred to as the state space. A state is, a snapshot of the current situation in the enviroment, with all the maximum information about it. In our case, since we had to make our own enviroment, it will be define by the frame of the video and the position of the ball and camera.

Observation Due that we are considering that the Agent doesn?t have access to the actual full state of the Environment, it is usually called observation the part of the state that the Agent can observe. We also followed this convention in order to make it as close as possible to openAI Gym enviroments

1.3 Q-Learning

Q-Learning is a model-free form of machine learning, in the sense that the AI "agent" does not need to know or have a model of the environment that it will be in. The same algorithm can be used across a variety of environments.

For a given environment, everything is broken down into "states" and "actions." The states are observations and samplings that we pull from the environment, and the actions are the choices the agent has made based on the observation. For the purposes of the rest of this tutorial, we'll use the context of our environment to exemplify how this works.

2 Bibliografía

[1] Sutton, R.S. and Barto, A.G. (1998) Reinforcement Learning: An Introduction, Vol. 1. MIT Press, Cambridge.

[2] J. Hurwitz and D. Kirsch, *Machine Learning For Dummies, IBM Limited Edition*, 1st ed. John Wiley Sons, Inc., 2018.

[3] Kaelbling, L. P., Littman, M. L., Moore, A. W., 1996. Reinforcement learning: A survey. Journal of artificial intelligence research 4, 237?285.

[4] Juan Diego Polo. (2014). *EyeEm compra tecnología para reconocer, clasificar y ordenar fotos. Lugar de publicación: whatsnew.*
`http://www.hatsnew.com/2014/08/15/eyeem-compra-tecnologia-para-reconocer-clasificar-y-ordenar-fotos/`