Insper

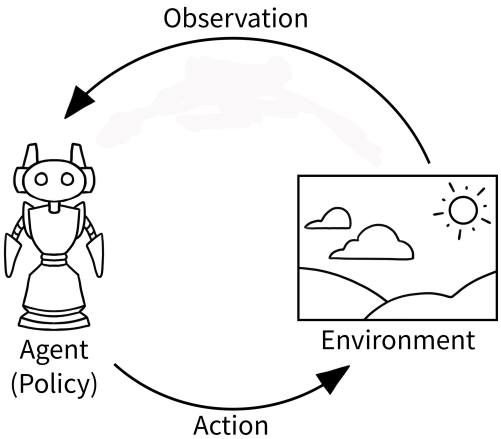
Introduction to Reinforcement Learning

February 2023

Reinforcement Learning
Fabricio Barth

AI Definitions and key concepts

Autonomous Agent



"Autonomous agents are computational systems. that inhabit some **complex dynamic environment**, **sense** and **act** autonomously in this environment, and by doing so realize a set of **goals** or **tasks** for which they are designed." Pattie Maes

IBM Deep Blue versus Kasparov (1997)

- **environment**: chess board;
- **task**: play and win a chess game;
- **actions**: move chess pieces;
- **implementation**: Min-Max algorithm + heuristics + knowledge base + dedicated hardware.

Handwritten digits recognition

The **MNIST database** of handwritten digits (1998)

- **environment**: an image
- **task**: classify an image without error
- **actions**: classify an image
- **implementation**: there are a lot of implementations. However, the best approach is always using a neural network model

In 2012 a solution achieved a very good result - a test error equals to 0.23% [http://yann.lecun.com/exdb/mnist/]

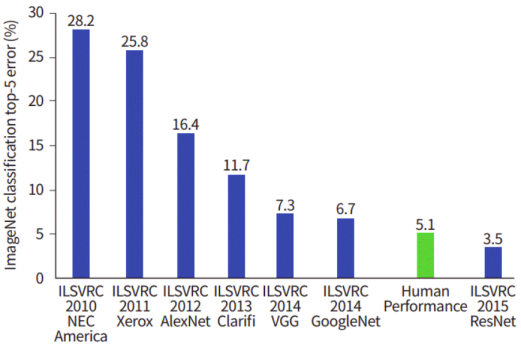
Autonomous car

This is the car from the Stanford team, which won the DARPA competition in 2005.

- **environment**: a road in a desert;
- **task**: drive through a desert and get to a specific point;
- **actions**: speed up, brake, turn left, turn right;
- **implementation**: a very complex implementation with different sensors and actuators.

The Stanford team was the first team that won this competition (2005).

ImagineNet dataset (2009)



This dataset has more than 14 million images annotated according to WordNet taxonomy.

This dataset has been used in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) since 2010. This competition is a benchmark for **image classification** tasks and **object recognition** tasks.

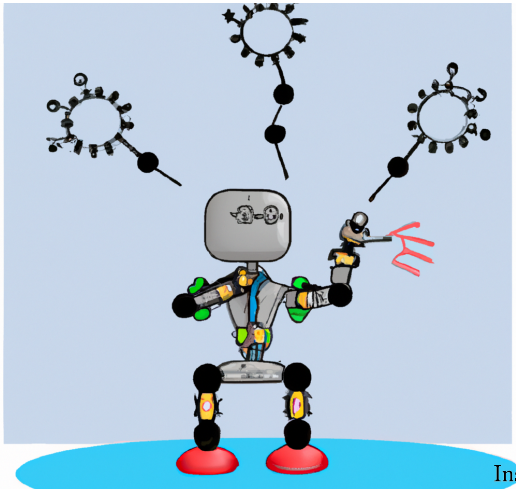
AlphaGO playing GO (2016)

- **environment**: GO board;
- **task**: play GO game and win;
- **actions**: move pieces;
- **implementation**: reinforcement learning implementation.

Generative models (DALL-E and chatGPT)

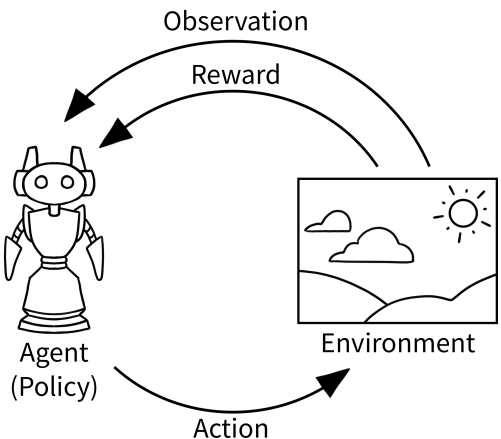
DALL-E input: "an autonomous robot solving a problem"

- **environment**: a form where the user can input a text;
- **task**: generate images that are related to the text informed;
- **actions**: generate images;
- **implementation**: a Deep Neural Network



Reinforcement Learning: definition and key concepts

Reinforcement Learning (RL)



- RL is a popular approach to AI where an agent learns to take sequential actions in an environment through trial and error.
- Each action changes the environment and has a **reward** (positive, negative or neutral).
- An Agent **learns** a **policy** that maximizes the reward.

Examples (1/2)

- Playing Atari with Deep Reinforcement Learning, 2013.
- Human-Level control through deep reinforcement learning, 2015. $\rightarrow$ *Overpassed human results*
- A General reinforcement learning algorithm that masters chess, shogi, and Go through self-play, 2018.

Examples (2/2)

- **Self-driving cars**: there are some components in self-driving cars that are optimized through RL.
- **Industry automation**: Google Data Centers are using RL to reduce energy spending.
- **Trading and finance**: there are a lot of companies saying that they are using RL to create robots for trading.
- **Natural Language Processing (NLP)**: chatGTP is using RL to improve its training.
- **Recommmendation**: there are a lot of papers about RL in recommendation systems.

Differences between other machine learning and AI techniques

Differences

- **Supervised learning**:
  - the main goal of supervised learning is to create a predicted model.
  - All the model construction is based on training and test datasets and both datasets must have a **label** attribute.
  - All the training process is in **batch** mode.
- **Unsupervised learning**:
  - the main goal of unsupervised learning is to summarize data. Usually, through clustering or rule models.
  - All the model construction is based on training datasets without a label attribute.
- **Reinforcement Learning**:
  - All the learning process is **interactive**.
  - There is no training data. However, there is an **environment**.