# Formula 1

Alessandro Tenaglia

Machine and Reinforcement Learning in Control Applications

April 4, 2022

# Problem

Terraces
Ste Devote on Track Terrace
Ermano
Caravelle
Heracles
A1
K
Beaurivage
Z1
B
C
M
L
N
O
P
X2
X1
T
V
TV
General Admission «Rocher»

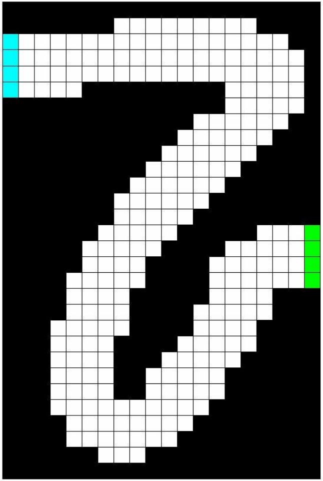
Learn to follow a track.

# Problem

Learn to follow a track.

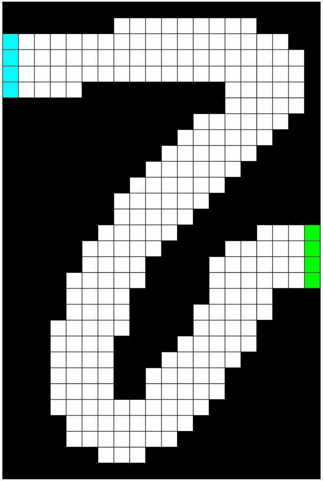
# Problem formulation

- The car starts from the starting line (cyan);
- The car must reach the finish line (green);
- The goal is to find the shortest path to the finish line;



# Problem formulation

- The car can move in 8 directions: **N, S, E, W, NE, NW, SE, SW**;
- Reward:
  - $-dist$ for each step
  - $-10^6$ for hitting the boundary



# Model

- The **state** is the position of the car in the track
  - we have $X \cdot Y$ states.
- The **action** is the direction along which the car moves
  - we have $8$ actions.

# Planning: Dynamic Programming

- Model the problem as an **MDP**;
- Compute the transition matrix **P**;
- Compute the reward matrix **R**;
- Find the optimal policy $\pi_\star$ using **Dynamic Programming** method (PI or VI).

# Learning: Montecarlo

- **Model-free**: no knowledge of MDP transitions and rewards;
- Simulate episodes:

$$S_0, A_0, R_0, S_1, A_1, R_1 ..., S_T, A_T, R_T$$

- Use experience to estimate $q_\star$:

$$\pi_\star(s) = \arg\max_a q_\star(s, a)$$

# Assignment #3

- Learn to play Blackjack.
  - The goal is to obtain cards whose sum is the closest to 21;
  - All face cards count as 10, an ace counts as 1 or as 11;
  - Player and dealer start with two cards;
  - The player can request additional cards (hit) or stop (stick);
  - The dealer plays according to a fixed strategy:
    - Stick on any sum of 17 or greater, and hit otherwise.
  - If the player exceeds 21, he goes burst an loses the game.
  - If the dealer goes burst, the player wins; otherwise, the outcome is determined by whose final sum is closer to 21;
  - Reward:
    - $+1$ for winning;
    - $-1$ for losing;
    - $0$ for drawing;
  - Do not discount ($\gamma = 1$).