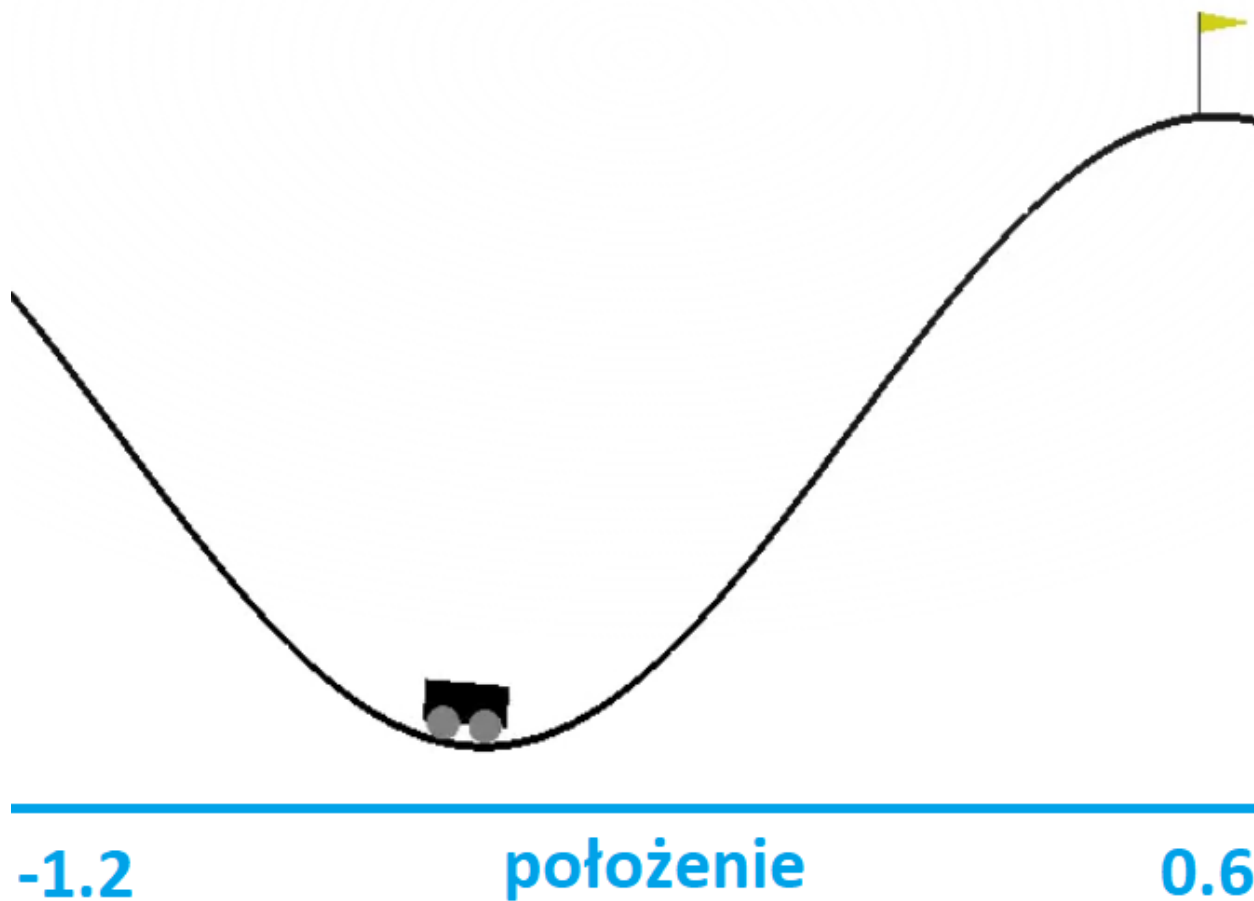


▼ Środowisko MountainCar

Pojazd znajduje się na **jednowymiarowym torze** (patrz rysunek poniżej), umieszczonym pomiędzy dwoma wzniesieniami. **Celem jest wjechać pojazdu na górę po prawej stronie** (miejsce oznaczone chorągiewką). Silnik samochodu nie jest wystarczająco mocny, aby bezpośrednio wjechać na górę. **Możliwym rozwiązaniem jest jazda "tam iz powrotem", w lewo i prawo i nabranie rozpędu.**



```
import numpy as np
import gym
import matplotlib.pyplot as plt
```

Stan reprezentowany jest za pomocą wektora **[położenie, prędkość]**. **Stan początkowy** ma losowe **położenie od -0,6 do -0,4** (położenie w dolinie) i **prędkość=0**:

```
env = gym.make("MountainCar-v0")
State_0 = env.reset()
State_0

array([-0.53757159,  0.          ])
```

Wartości **min** i **max** położenia oraz prędkości:

```
print("położenie min=", env.observation_space.low[0], " max=", env.observation_space.high[0])
print("prędkość min=", env.observation_space.low[1], " max=", env.observation_space.high[1])

położenie min= -1.2   max= 0.6
prędkość min= -0.07  max= 0.07
```

▼ Akcje

Możliwe do wykonania są **3 akcje**, oznaczone **0,1,2**:

```
env.action_space

Discrete(3)
```

Przetestujmy **akcję 0** - wykonajmy ją **10 razy**:

```

env.reset()
for _ in range(10):
    action = 0
    obs, reward, done, info = env.step(action)
    print("stan=",obs)

    stan= [-0.54402818 -0.00085326]
    stan= [-0.54572832 -0.00170014]
    stan= [-0.54826261 -0.00253429]
    stan= [-0.55161209 -0.00334948]
    stan= [-0.55575172 -0.00413963]
    stan= [-0.56065057 -0.00489885]
    stan= [-0.56627211 -0.00562154]
    stan= [-0.57257447 -0.00630236]
    stan= [-0.57951083 -0.00693636]
    stan= [-0.58702981 -0.00751899]

```

Jak widać powyżej akcja 0 powoduje **pchnięcie pojazdu w lewą stronę** (zmniejszeniu ulega pierwsza współrzędna) i **zwiększeniu ulega prędkość pojazdu w lewą stronę** (druga współrzędna - ujemna wartość prędkości).

W przypadku pozostałych akcji: **1 - brak pchnięcia pojazdu, 2 - popchnięcie w prawo**.

Przetestujmy działanie środowiska - **wykonajmy pewną ilość (dokładnie 50) akcji 0** (pchnięcie pojazdu w lewą stronę):

```

env.reset()
for _ in range(50):
    action = 0
    obs, reward, done, info = env.step(action)
    print("położenie: ",obs[0]," prędkość: ",obs[1])
    if obs[1]<0:
        print("pojazd porusza się w LEWO")
    else:
        print("pojazd porusza się w PRAWO")

    położenie: -0.4923972796370638 prędkość: -0.001242949307119431
    pojazd porusza się w LEWO
    położenie: -0.4948738985869782 prędkość: -0.0024766189499144022

```

pojazd porusza się w LEWO
położenie: -0.49856568757238695 prędkość: -0.003691788985408749
pojazd porusza się w LEWO
położenie: -0.5034450482876252 prędkość: -0.004879360715238208
pojazd porusza się w LEWO
położenie: -0.509475469883213 prędkość: -0.0060304215955878145
pojazd porusza się w LEWO
położenie: -0.5166117845816361 prędkość: -0.0071363146984231764
pojazd porusza się w LEWO
położenie: -0.5248004978754922 prędkość: -0.008188713293856094
pojazd porusza się w LEWO
położenie: -0.5339801982717932 prędkość: -0.009179700396300955
pojazd porusza się w LEWO
położenie: -0.5440820505844589 prędkość: -0.010101852312665704
pojazd porusza się w LEWO
położenie: -0.5550303749996074 prędkość: -0.010948324415148522
pojazd porusza się w LEWO
położenie: -0.5667433116069046 prędkość: -0.011712936607297247
pojazd porusza się w LEWO
położenie: -0.5791335669413071 prędkość: -0.012390255334402474
pojazd porusza się w LEWO
położenie: -0.5921092355184381 prędkość: -0.012975668577130978
pojazd porusza się w LEWO
położenie: -0.60557468563474 prędkość: -0.01346545011630193
pojazd porusza się w LEWO
położenie: -0.6194314951363321 prędkość: -0.013856809501592036
pojazd porusza się w LEWO
położenie: -0.6335794197416663 prędkość: -0.014147924605334176
pojazd porusza się w LEWO
położenie: -0.6479173741133288 prędkość: -0.014337954371662479
pojazd porusza się w LEWO
położenie: -0.6623444044417404 prędkość: -0.014427030328411531
pojazd porusza się w LEWO
położenie: -0.6767606309783275 prędkość: -0.014416226536587088
pojazd porusza się w LEWO
położenie: -0.6910681397936717 prędkość: -0.014307508815344178
pojazd porusza się w LEWO
położenie: -0.7051718049926392 prędkość: -0.014103665198967564
pojazd porusza się w LEWO
położenie: -0.7189800255548765 prędkość: -0.01380822056223728
pojazd porusza się w LEWO
położenie: -0.7324053646695429 prędkość: -0.013425339114666326
pojazd porusza się w LEWO
położenie: -0.7453650836299575 prędkość: -0.012959718960414606

```

pojazd porusza się w LEWO
położenie: -0.7577815667565132  prędkość: -0.01241648312655573
pojazd porusza się w LEWO
położenie: -0.7695826381409941  prędkość: -0.011801071384480942
pojazd porusza się w LEWO
położenie: -0.7807017750007231  prędkość: -0.011119136859728995
pojazd porusza się w LEWO
położenie: -0.7910782258961198  prędkość: -0.010376450895396608
pojazd porusza się w LEWO
położenie: -0.8006570448612556  prędkość: -0.0095788189651358

```

Jak widać powyżej pojazd **początkowo porusza się w lewo**, ale **po pewnym czasie zaczyna opadać, czyli poruszać się w prawo**. W tym momencie prędkość pojazdu zmienia wartość z ujemnej (prędkość w lewo) na dodatnią (prędkość w prawo).

Zobaczmy co się dzieje dalej. Spróbujmy wykonać teraz pewną ilość razy **akcję 1 czyli brak pchnięcia** (UWAGA: **nie resetujemy stanu środowiska!!!**):

```

for _ in range(50):
    action = 1
    obs, reward, done, info = env.step(action)
    print("położenie: ",obs[0]," prędkość: ",obs[1])
    if obs[1]<0:
        print("pojazd porusza się w LEWO")
    else:
        print("pojazd porusza się w PRAWO")
    położenie: -0.5220183474302387  prędkość: 0.017846909115466082
    pojazd porusza się w PRAWO
    położenie: -0.30559281147276396  prędkość: 0.016425535957494717
    pojazd porusza się w PRAWO
    położenie: -0.2906882257434215  prędkość: 0.014904585729342428
    pojazd porusza się w PRAWO
    położenie: -0.2773917577175691  prędkość: 0.013296468025852427
    pojazd porusza się w PRAWO
    położenie: -0.26577846209757827  prędkość: 0.011613295619990789
    pojazd porusza się w PRAWO
    położenie: -0.25591170575406935  prędkość: 0.009866756343508924
    pojazd porusza się w PRAWO
    położenie: -0.24784366313960526  prędkość: 0.00806804261446408
    pojazd porusza się w PRAWO
    położenie: -0.24161582816671529  prędkość: 0.006227834972889976

```

pojazd porusza się w PRAWO

położenie: -0.23725949457305998 prędkość: 0.004356333593655307

pojazd porusza się w PRAWO

położenie: -0.23479616448315874 prędkość: 0.0024633300899012294

pojazd porusza się w PRAWO

położenie: -0.2342378536089337 prędkość: 0.0005583108742250305

pojazd porusza się w PRAWO

położenie: -0.23558727085877704 prędkość: -0.0013494172498433356

pojazd porusza się w LEWO

położenie: -0.23883785974908883 prędkość: -0.0032505888903117896

pojazd porusza się w LEWO

położenie: -0.2439736987587622 prędkość: -0.005135839009673387

pojazd porusza się w LEWO

położenie: -0.2509692675366515 prędkość: -0.006995568777889266

pojazd porusza się w LEWO

położenie: -0.2597890955829743 prędkość: -0.008819828046322824

pojazd porusza się w LEWO

położenie: -0.27038731955402645 prędkość: -0.010598223971052133

pojazd porusza się w LEWO

położenie: -0.2827071844648852 prędkość: -0.012319864910858744

pojazd porusza się w LEWO

położenie: -0.2966805324277101 prędkość: -0.013973347962824925

pojazd porusza się w LEWO

położenie: -0.31222732964542416 prędkość: -0.015546797217714087

pojazd porusza się w LEWO

położenie: -0.3292552875138456 prędkość: -0.0170279578684214

pojazd porusza się w LEWO

położenie: -0.3476596360988327 prędkość: -0.018404348584987105

pojazd porusza się w LEWO

położenie: -0.36732310716345457 prędkość: -0.01966347106462184

pojazd porusza się w LEWO

położenie: -0.3881161786510248 prędkość: -0.020793071487570237

pojazd porusza się w LEWO

położenie: -0.4098976226760657 prędkość: -0.021781444025040904

pojazd porusza się w LEWO

położenie: -0.432515384650606 prędkość: -0.022617761974540306

pojazd porusza się w LEWO

położenie: -0.45580780273600685 prędkość: -0.023292418085400807

pojazd porusza się w LEWO

położenie: -0.4796051555194483 prędkość: -0.023797352783441476

pojazd porusza się w LEWO

położenie: -0.5037315033842221 prędkość: -0.02412634786477381

pojazd porusza się w LEWO

Po zmianach położenia widać, że **pojazd początkowo kontynuuje ruch w prawo** (wartość położenia ulega zwiększeniu) czyli **podjeżdża pod wzniesienie z prawej strony**. Po pewnym czasie jednak **prędkość zmienia wartość z dodatniej na ujemną** co oznacza, że **pojazd zaczyna opadać** czyli **porusza się w lewo**.

▼ Nagrody

Nagroda wynosi **-1 za każdy krok czasowy**, aż do osiągnięcia **docelowej pozycji 0.5** (położenie chorągiewki).

▼ Zakończenie epizodu

Zakończenie epizodu następuje w chwili **osiągnięcia położenia 0.5** lub **po przejściu 200 kroków**.

Przetestujmy działanie środowiska w przypadku **losowego wyboru akcji** (`action = env.action_space.sample()`).

```
R = 0
env.reset()
for e in range(300):
    #env.render()
    action = env.action_space.sample()
    obs, reward, done, info = env.step(action)
    print("Epizod=",e," akcja=",action," R=",reward," koniec epizodu?=",done," stan=",obs)
    Epizod= 242 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.32040317 -0.00307143]
    Epizod= 243 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.33095034 -0.00446517]
    Epizod= 244 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.33578127 -0.00483093]
    Epizod= 245 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.34094747 -0.0051662 ]
    Epizod= 246 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.34741603 -0.00646856]
    Epizod= 247 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.35614529 -0.00872926]
    Epizod= 248 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.36607829 -0.009933 ]
    Epizod= 249 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.3781492 -0.01207092]
    Epizod= 250 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.39027671 -0.0121275 ]
    Epizod= 251 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.40437768 -0.01410097]
```

```
Epizod= 252 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.42035386 -0.01597619]
Epizod= 253 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.43609206 -0.0157382 ]
Epizod= 254 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.45147905 -0.01538699]
Epizod= 255 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.46840272 -0.01692368]
Epizod= 256 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.48473848 -0.01633576]
Epizod= 257 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.50236503 -0.01762655]
Epizod= 258 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.52115073 -0.0187857 ]
Epizod= 259 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.53895478 -0.01780406]
Epizod= 260 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.55764371 -0.01868893]
Epizod= 261 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.57707774 -0.01943403]
Epizod= 262 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.5951124 -0.01803466]
Epizod= 263 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.61261481 -0.01750241]
Epizod= 264 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.63045751 -0.0178427 ]
Epizod= 265 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.64851242 -0.01805491]
Epizod= 266 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.66565226 -0.01713984]
Epizod= 267 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.68275865 -0.01710639]
Epizod= 268 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.69971619 -0.01695754]
Epizod= 269 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.71641345 -0.01669726]
Epizod= 270 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.7327439 -0.01633045]
Epizod= 271 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.74860668 -0.01586278]
Epizod= 272 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.76390717 -0.01530049]

Epizod= 273 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.77855746 -0.01465029]
Epizod= 274 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.79047663 -0.01191917]
Epizod= 275 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.7996013 -0.00912468]
Epizod= 276 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.80788452 -0.00828321]
Epizod= 277 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.81428482 -0.00640031]
Epizod= 278 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.81777099 -0.00348617]
Epizod= 279 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-8.18326313e-01 -5.55319726e-04]
Epizod= 280 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.81694814 0.00137817]
Epizod= 281 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.81264304 0.0043051 ]
Epizod= 282 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.80643175 0.00621129]
Epizod= 283 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.79934474 0.00708701]
Epizod= 284 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.79141757 0.00792717]
Epizod= 285 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.781691 0.00972657]
Epizod= 286 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.77121643 0.01047457]
Epizod= 287 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.75805088 0.01316555]
Epizod= 288 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.74426837 0.0137825 ]
Epizod= 289 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.72794912 0.01631926]
Epizod= 290 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.70919144 0.01875767]
Epizod= 291 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.68811264 0.02107881]
Epizod= 292 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.66584946 0.02226318]
Epizod= 293 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.64255149 0.02329797]
Epizod= 294 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.61838019 0.0241713 ]
Epizod= 295 akcja= 2 R= -1.0 koniec epizodu?= True stan= [-0.59250758 0.02587262]
```



```
Epizod= 296 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.56612182  0.02638576]
Epizod= 297 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.540418      0.02570382]
Epizod= 298 akcja= 1 R= -1.0 koniec epizodu?= True stan= [-0.51458809  0.02582991]
Epizod= 299 akcja= 0 R= -1.0 koniec epizodu?= True stan= [-0.48982575  0.02476234]
```

Jak widać **po 200 krokach następuje zakończenie epizodu.**

▼ Rozwiązanie zadania 1 z zestawu 11

```
max_right = -1.0
for _ in range(100):
    action = 0
    obs, reward, done, info = env.step(action)
    print("stan=", obs)
    if obs[1] > 0:
        action = 2
        print("Right")
    if obs[0] > max_right:
        max_right = obs[0]
```

```
stan= [-0.45744573 -0.02147493]
stan= [-0.48041356 -0.02296783]
stan= [-0.50470437 -0.02429081]
stan= [-0.53013681 -0.02543244]
stan= [-0.55652022 -0.02638341]
stan= [-0.58365712 -0.0271369 ]
stan= [-0.61134602 -0.0276889 ]
stan= [-0.63938439 -0.02803837]

stan= [-0.66757172 -0.02818733]
stan= [-0.69571252 -0.0281408 ]
stan= [-0.72361906 -0.02790654]
stan= [-0.75111387 -0.02749481]
stan= [-0.77803175 -0.02691789]
stan= [-0.80422137 -0.02618961]
stan= [-0.82954626 -0.02532489]
stan= [-0.85388542 -0.02433916]
stan= [-0.87713334 -0.02324792]
stan= [-0.89919963 -0.02206629]
```

```

stan= [-0.89919989 -0.82288829]
stan= [-0.92000831 -0.02080868]
stan= [-0.93949679 -0.01948848]
stan= [-0.95761462 -0.01811784]
stan= [-0.9743222 -0.01670758]
stan= [-0.98958929 -0.01526709]
stan= [-1.00339363 -0.01380433]
stan= [-1.01571951 -0.01232589]
stan= [-1.02655654 -0.01083703]
stan= [-1.03589836 -0.00934182]
stan= [-1.04374161 -0.00784325]
stan= [-1.050085 -0.00634339]
stan= [-1.05492848 -0.00484348]
stan= [-1.05827263 -0.00334415]
stan= [-1.06011817 -0.00184553]
stan= [-1.06046558e+00 -3.47411703e-04]
stan= [-1.05931497 0.00115061]
Right
stan= [-1.05666601 0.00264896]
Right
stan= [-1.05251807 0.00414795]
Right
stan= [-1.04687044 0.00564763]
Right
stan= [-1.03972281 0.00714763]
Right
stan= [-1.03107581 0.008647 ]
Right
stan= [-1.02093173 0.01014408]
Right
stan= [-1.00929541 0.01163632]
Right
stan= [-0.99617524 0.01312017]
Right
stan= [-0.98158429 0.01459095]
Right
stan= [-0.96554162 0.01604267]
Right
stan= [-0.94807359 0.01746803]
Right

```

```
print("Max right: ", max_right)
```

Max right: -0.3304226619208586

✓ 0 s ukończono o 17:32

