

UWAGA: Wczytaj do Colab plik **frozen_lake.py** (instrukcja w pliku **COLAB_instrukcja.pdf**)

FrozenLake 1

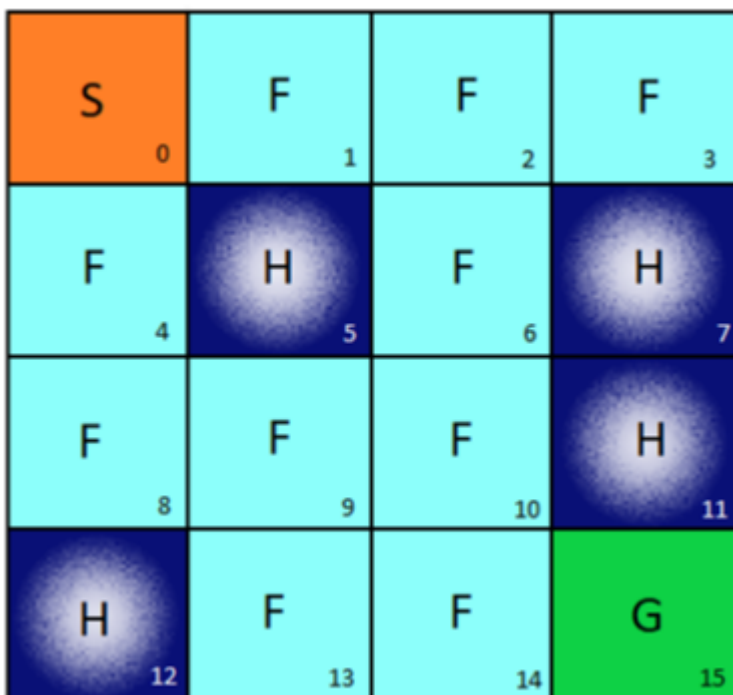
▼ Wprowadzenie

Agent porusza się w świecie złożonym z **16 pól (stanów)**. Stany są ponumerowane od 0 do 15.

Niektóre pola siatki są dostępne do chodzenia (**F**-frozen), a inne są przerębiami (**H**-hole).

Możliwe są 4 akcje: **0 - LEFT, 1 - DOWN, 2 - RIGHT, 3 - UP**

Agent jest nagradzany (**R=1**) za dotarcie do pola **G**. W pozostałych przypadkach **R=0**.



Łaadowanie biblioteki (wcześniej konieczne załadowanie pliku **frozen_lake.py** do Colaba - instrukcja)

```
from frozen_lake import FrozenLakeEnv
```

Wczytanie środowiska:

```
env = FrozenLakeEnv()
```

Sprawdzamy ilość możliwych stanów (16) i akcji (4)

```
print(env.nS)
```

```
print(env.nA)
```

```
16  
4
```

▼ Dynamika

Dynamika opisana jest za pomocą: `env.P[s][a]`

gdzie: **s** to **stan** (0,1,2,...,15), **a** to **akcja** (0,1,2,3).

Rozważmy przykład: w stanie 0 agent wykonuje akcję 1 (porusza się w dół):

```
env.P[0][1]
```

```
[(1.0, 4, 0.0, False)]
```

Powyższą czwórkę interpretujemy jako: (**prawdopodobieństwo**, **nowy stan**, **nagroda**, **czy koniec?**)

Czyli w powyższym przykładzie: po wykonaniu w stanie 0 akcji 1 prawdopodobieństwo przejścia c wpadł do przerębli ani nie dotarł do pola G.

▼ Polecenie 1 (do uzupełnienia)

Sprawdź dynamikę dla następujących przypadków:

W **stanie 1** agent **przechodzi w dół**:

```
env.P[1][1]
```

```
[(1.0, 5, 0.0, True)]
```

W **stanie 10** agent **przechodzi w lewo**:

```
env.P[10][0]
```

```
[(1.0, 9, 0.0, False)]
```

W **stanie 14** agent **przechodzi w prawo**:

```
env.P[14][2]
```

```
[(1.0, 15, 1.0, True)]
```

▼ Poruszanie i wizualizacja

W świecie FrozenLake możemy się poruszać wykonując 4 akcje (omówione powyżej). Podgląd pc `env.render()` (wcześniej resetujemy położenie agenta).

```
env.reset()
env.render()
```



```
SFFF
FHFH
FFFF
HFFG
```

Wykonajmy dwa ruchy w prawo i jeden w dół:

```
env.reset()
env.step(2)
env.step(2)
env.step(1)
env.render()
```



```
(Down)
SFFF
FHFH
FFFF
HFFG
```

Metoda `step` zwraca krotkę (**nowy stan, nagroda, czy koniec ruchu,**_{...}). Koniec następuje wtedy gdy 15 - GOAL). Sprawdźmy to.

Z pola początkowego 0 agent rusza się w prawo (akcja - 2) na pole 1 i zdobywa nagrodę 0:

```
env.reset()
env.step(2)
```



```
(1, 0.0, False, {'prob': 1.0})
```

Agent kontynuuje ruch: rusza się w prawo (akcja - 2) na pole 2 i zdobywa nagrodę 0:

```
env.step(2)
```



```
(2, 0.0, False, {'prob': 1.0})
```

▼ Polecenie 2 (do uzupełnienia)

Przeprowadź agenta dowolną drogą z pola 0 do pola 15 (GOAL). Sprawdź czy nagroda po wejściu

```
env.reset()
env.step(2)
env.step(2)
env.step(1)
env.step(1)
env.step(1)
env.step(1)
print(env.step(2)[1])
env.render()
```

```
1.0
(Right)
SFFF
FHFH
FFFH
HFFG
```

▼ Ruch agenta w pętli

Ruch agenta można zapętląć. Na razie akcja w każdym stanie generowana jest losowo (wykorzyst Agent wykona 10 akcji.

UWAGA: kiedy agent jest na polu oznaczonym **H** (stany 5,7,11,12) **dowolna akcja pozostawia go n**

```
env.reset()
for i in range(10):
    action = env.action_space.sample()
    obs, rew, fin, _ = env.step(action)
    print("Action=",action,"State =",obs,"Reward =",rew,"End =",fin)
```

```
Action= 3 State = 0 Reward = 0.0 End = False
Action= 2 State = 1 Reward = 0.0 End = False
Action= 1 State = 5 Reward = 0.0 End = True
Action= 3 State = 5 Reward = 0 End = True
Action= 0 State = 5 Reward = 0 End = True
Action= 1 State = 5 Reward = 0 End = True
Action= 1 State = 5 Reward = 0 End = True
Action= 3 State = 5 Reward = 0 End = True
Action= 1 State = 5 Reward = 0 End = True
Action= 3 State = 5 Reward = 0 End = True
```

▼ Polecenie 3 (do uzupełnienia)

Sprawdź czy możliwe jest dotarcie agenta do pola G w przypadku gdy akcje są generowane losowo (z pomocą odpowiedniej pętli). Zawsze gdy agent wpadnie do przerębli przerwij pętlę.

Poniżej wpisz kod:

```
for j in range(500):
    env.reset()
    for i in range(52):
        action = env.action_space.sample()
        obs, rew, fin, _ = env.step(action)
        #print("Action=",action,"State =",obs,"Reward =",rew,"End =",fin)
        if(rew==1):
            print("udalo sie dotrzec do pola G")
```

```
udalo sie dotrzec do pola G
udalo sie dotrzec do pola G
udalo sie dotrzec do pola G
udalo sie dotrzec do pola G
udalo sie dotrzec do pola G
udalo sie dotrzec do pola G
```

TWOJE PODSUMOWANIE TESTÓW:

Na 500 przeprowadzonych testów do pola G agent dotarł 6 razy