**Zadanie 1**

**Pytanie 1:**

Wartość gamma jest to wartość rabatu przyznawanego za otrzymanie nagrody czyli wykonanie poprawnego ruchu. Gamma pozwala na zwiększenie wartości dla kolejnej akcji. W przypadku ustawienia gammy na 0 nie ma wpływu na wartość akcji w zależności od poprzedniego ruchu. Wartość gamma ustawiona na 1 pozwala na zwiększenie wartości akcji o całą wartość następnego stanu.

**Pytanie 2:**

Metoda Policy Iteration pozwala zapętlić działanie programu tak aby uzyskać ocenę polityki (wartości stanu) i funkcję poprawy polityki aż do uzyskania optymalnej polityki czyli trasy przejścia na zamarzniętym jeziorze.

Metoda Value\_iteration polega na określeniu trasy poprzez przejście przez pętla tak aby otrzymać jednorazową politykę czyli trasę przejścia. Wykonywanie tras odbywa się tak długo aż otrzyma się wynik zbliżony do założenia.

**Pytanie 3:**

Value iteration znajduję lepszą trasę ponieważ sprawdza wszystkie dostępne trasy, niezależnie od rozkładu Police iteration. Sprawdzając każdą trasę jest możliwość odnalezienia najlepszej trasy. Natomiast police iteration może w początkowym okresie wybrać ścieżkę, która będzie nie optymalna i nie pozwoli osiągnąć najlepszych rezultatów.

**Zadanie 2.**

**Pytanie 1:**

Parametr poślizg powoduje zwiększenie skomplikowania trasy co może spowodować brak możliwości odnalezienia odpowiedniej trasy przez zamarznięte jezioro.

**Pytanie 2:**

W obliczonych przypadkach dla gammy równej 0, 0.2 i 0.6 parametr gamma nie wpłyną na wybór trasy. We wszystkich przypadkach trasa była tak sama.

**Zadanie 3.**

W funkcji value\_iteration wartość q\_s jest obliczana dla wszystkich akcji, a w przypadku value\_iteration2 obliczamy q\_s z jednego kroku naprzód od v. Funkcja value\_iteration jest to rozszerzona wersja funkcji do funkcji value\_iteration.