Zadanie 5

Załóżmy, że problem pole-balancing traktujemy jako zadanie epizodyczne ze zdyskontowaniem. Przyjmujemy, że wszystkie nagrody równe są 0, z wyjątkiem -1 po niepowodzeniu.

**Jaki będzie zwrot za każdym razem?**

Za każdym razem byłby liczbą kroków do momentu niepowodzenia.

**Jak ten zwrot różni się od przypadku ze zdyskontowaniem i ruchem ciągłym?**

W tym wypadku zwrot każdym razem byłby liczbą kroków do momentu niepowodzenia zaś w ruchu ciągłym zwrot w tym przypadku będzie równy −𝛾K , gdzie K jest liczbą kroków przed niepowodzeniem.