Zadanie 3

Przypuśćmy, że uczenie ze wzmocnieniem jest stosowane w celu określenia temperatury i tempa wsuwania prętów kontrolnych dla reaktora atomowego.

Przykład 1:

Akcje: docelowe temperatury i docelowe szybkości wsuwania prętów sterujących/kontrolnych(wektor)

Stany: odczyty parametrów, być może przefiltrowane i opóźnione, oraz dane wejściowe reprezentujące moc reaktora i stan zanieczyszczania Ksenonem-135(zmniejsza on moc reaktora przez to, że silnie pochłania neurony, powodując nawet jego zatrzymanie i przestój) . Stany są wektorami.

Nagroda to szybkość z jaką użyteczna moc jest wytwarzana przez reaktor(liczba, a nie wektor).

Przykład 2:

Robot sprzątający mieszkanie.

Stan opisuje poziom naładowania baterii: S = {wysoki, niski}

Akcje:umyj, czekaj, ładuj

A(niski) = {umyj, czekaj, ładuj}

A(wysoki= {umyj, czekaj}

Nagrody: każdy przejechany i posprzątany metr kwadratowy +1

rumyj > rczekaj,0,-3

Przykład 3:

Suszarka automatyczna do ubrań

Akcje: docelowe temperatury i docelowe szybkości wirowania (wektor)

Stany: odczyty parametrów, być może przefiltrowane i opóźnione, oraz dane wejściowe reprezentujące stan wilgotności ubrań. Stany są wektorami.

Nagroda to szybkość z jaką pranie jest wysuszone przez suszarkę(liczba, a nie wektor).