知的システム演習　レポート2

T17CS015　帯津 勇斗

状態：センサーから読みとったものを用いて状態を以下のように定義した

読み取ったセンサーは左から順に　ライトA、B、C

　値は　白の時に０、黒の時に１、青の時に４とした

報酬とともに以下の表に列挙した

報酬：上記の表のそれぞれの状態に対応する報酬の値を追加し、定義した

|  |  |  |
| --- | --- | --- |
| 状態 | 読み取ったセンサー | 報酬 |
| 0 | 000 | -100 |
| 1 | 001 | 20 |
| 2 | 010 | 70 |
| 3 | 011 | 30 |
| 4 | 100 | 20 |
| 5 | 101 | 10 |
| 6 | 110 | 30 |
| 7 | 111 | 20 |
| 8 | 040 | 50 |
| 9 | 041 | 50 |
| 10 | 140 | 50 |
| 11 | 141 | 10 |

マルコフ決定過程との関係性

Map8まではある程度の分岐はあるものの単調なため、少しの学習により解決するものであるがMap9は似たような分岐がいくつかあるため、現在の状態だけでは解決することはとても困難だと考える。

工夫した点、苦労した点

* 状態の定義の仕方の点で２進数を利用することによってデバックのしやすいものとした。
* 試行回数の設定に手間取った
* εの値によってクリアできないときがあったのでそれの判断が大変だった。
* map9はかなりの時間をかけて改変を行ったが、クリアできなかった。
* 安定してクリアできないことがあったので修正を加えた。
* できる限りまっすぐに直線を進むように修正。
* クリア率の向上
* 青いマークを通ったときの状態定義を同じように簡素化させること
* 学習の際にウィンドウが出ないように修正
* デバック用のコメントアウトも用意