

## GRUNDLAGEN ADAPTIVER WISSENSYSTEME (SS2025)

Prof. Dr. Thomas Gabel

### Aufgabenblatt 2

#### Aufgabe 5: Absorbierende Zustände

- (a) Was ist ein absorbierender Zustand?
- (b) Sie wollen ein Pfadplanungsproblem mittels eines MDP modellieren. Der Zielzustand muss hierbei ein absorbierender Zustand sein. Welche Form direkter Kosten sollten hierbei dem Zielzustand und den Nicht-Zielzuständen zugeordnet sein, wenn das Ziel darin besteht, die Pfadlänge zu minimieren.
- (c) Betrachten Sie nun ein Diskontierungsproblem mit Diskontierungsfaktor  $\gamma \in [0, 1]$  das einen absorbierenden Zustand  $s$  enthält, in dem direkte Kosten  $c$  entstehen. Was ist der Wert  $V(s)$  (die Pfadkosten) jenes Zustandes?

#### Aufgabe 6: Labyrinth und Brücke

In Abbildung 1 sind zwei MDPs dargestellt. Im Labyrinth-MDP sind alle Zustandsübergänge deterministisch. Im Brücken-MDP sind nur Zustandsübergänge infolge von Aktionen, die den Agenten in das Wasser (N, S) bewegen, deterministisch. Alle anderen Aktionen (Bewegung nach W oder O) gelingen mit einer Wahrscheinlichkeit von  $1 - \varrho$ . Mit einer Wahrscheinlichkeit von  $\frac{\varrho}{3}$  ( $\varrho$  wird im Folgenden als Rauschintensität bezeichnet) hingegen verharrt der Agent auf der Stelle und mit jeweils  $\frac{\varrho}{3}$  bewegt er sich nach N oder S (und stürzt damit von der Brücke).

Der Agent startet stets im Startzustand  $S$ . Felder, in denen eine Zahl steht, sind Terminalzustände. Für Übergänge in Terminalzustände sind stets die zugehörigen direkten Kosten angegeben (z.B. Kosten von 100 dafür, von der Brücke zu fallen; Kosten von -10 dafür, die Brücke überquert zu haben, Kosten von -1 dafür, das Labyrinth zu verlassen). Für Aktionen, die den Agenten in eine Wand bewegen würden, werden Kosten von 1 vergeben (Agent bewegt sich nicht); alle anderen Transitionen sind kostenfrei.

- (a) Wie viele Iterationen des Wertiterationsverfahrens sind für den Labyrinth-MDP notwendig, bis dieser in seiner Wertfunktion (Pfadkostenvektor) für den Zustand  $S$  (Start) erstmalig einen Wert ungleich null einträgt? Warum?

Betrachten Sie nun den Brücken-MDP, in dem ein Diskontierungsfaktor von  $\gamma = 0.9$  sowie eine Rauschintensität von  $\varrho = 0.3$  zur Anwendung kommen.

- (b) Ermitteln Sie die optimalen Pfadkosten im Zustand  $S$  sowie in allen Zuständen östlich und westlich von  $S$ .

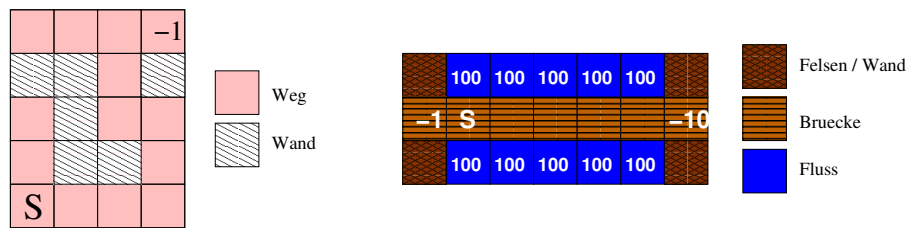


Abbildung 1: Labyrinth und Bruecke

- (c) Welcher der beiden Parameter  $\gamma$  und  $\varrho$  ist abzuändern, damit sich der Agent traut, die Brücke zu überqueren?
- (d) Auf welchen Wert muss jener Parameter zu diesem Zweck (näherungsweise) gesetzt werden?