

Arbeitsblatt 2 Reinforcement Learning: Die Q-Map berechnen

Wie findet die Maus den besten Weg zum Käse?



Nachdem die Schülerinnen und Schüler das Prinzip der Wertausbreitung in der Unplugged-Aktivität körperlich erfahren haben, vertiefen sie es hier auf Papier:

In zwei Aufgaben mit steigendem Schwierigkeitsgrad berechnen sie von Hand eine Q-Map – zuerst in einem kleinen 3×3-Labyrinth, dann in einem komplexeren 4×4-Labyrinth mit mehreren Hindernissen.

Die Rechenregel bleibt dabei dieselbe wie in der Gerüchteküche: „Bester Nachbar minus 1“. Zu jeder Aufgabe zeichnen die Schülerinnen und Schüler anschließend die zugehörige Policy-Map mit Richtungspfeilen und verfolgen den Weg der Maus zum Käse.

Das Arbeitsblatt schließt mit einer Reflexion, die den Bogen zur MouseMaze-App schlägt. Ein separates Lösungsblatt für die Lehrkraft liegt bei.

💡 Die Lernregel der Maus

Die Maus lernt, indem sich das Wissen über den Käse Schritt für Schritt ausbreitet:

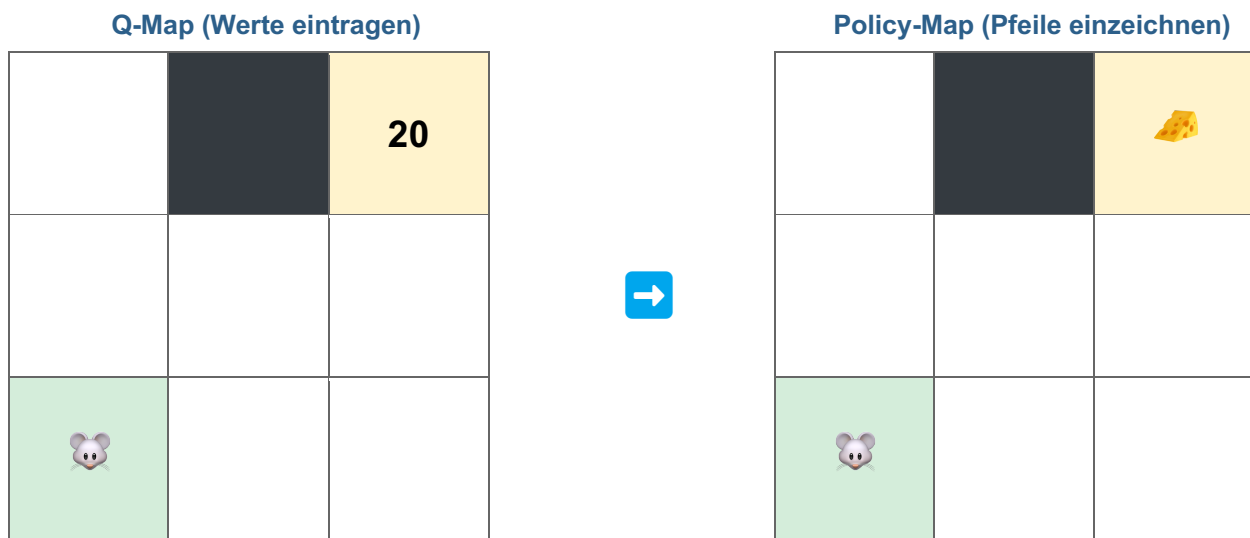
1. Der Käse 🧀 ist **20 Punkte** wert. Alle anderen Felder starten bei **0**.
2. Ein Feld bekommt den Wert: **bester Nachbar – 1** (nur $\updownarrow\leftarrow\rightarrow$, nicht diagonal).
3. Hindernisse (■) haben keinen Wert und blockieren den Weg.
4. Die Maus geht immer zum Nachbarfeld mit dem **höchsten Wert** – das ist ihre Strategie (Policy)!

Aufgabe 1: Das kleine Labyrinth (3×3)

Die Maus 🐭 steht unten links. Der Käse 🧀 (= 20) liegt oben rechts. Ein Hindernis versperrt den direkten Weg.

a) Berechne die Werte: Beginne beim Käse (20) und arbeite dich wellenförmig nach außen. Trage die Werte in die linke Tabelle ein.

b) Zeichne die Pfeile: Zeichne in der rechten Tabelle auf jedem Feld einen Pfeil (↑ ↓ ← →) in Richtung des Nachbarn mit dem höchsten Wert.



👉 Tipp

Arbeite in Wellen vom Käse aus: **Welle 1:** Nachbarn des Käses → 19. **Welle 2:** Deren Nachbarn → 18. Und so weiter ...

Hindernisse überspringen – sie geben keinen Wert weiter und man kann nicht durch sie hindurch!

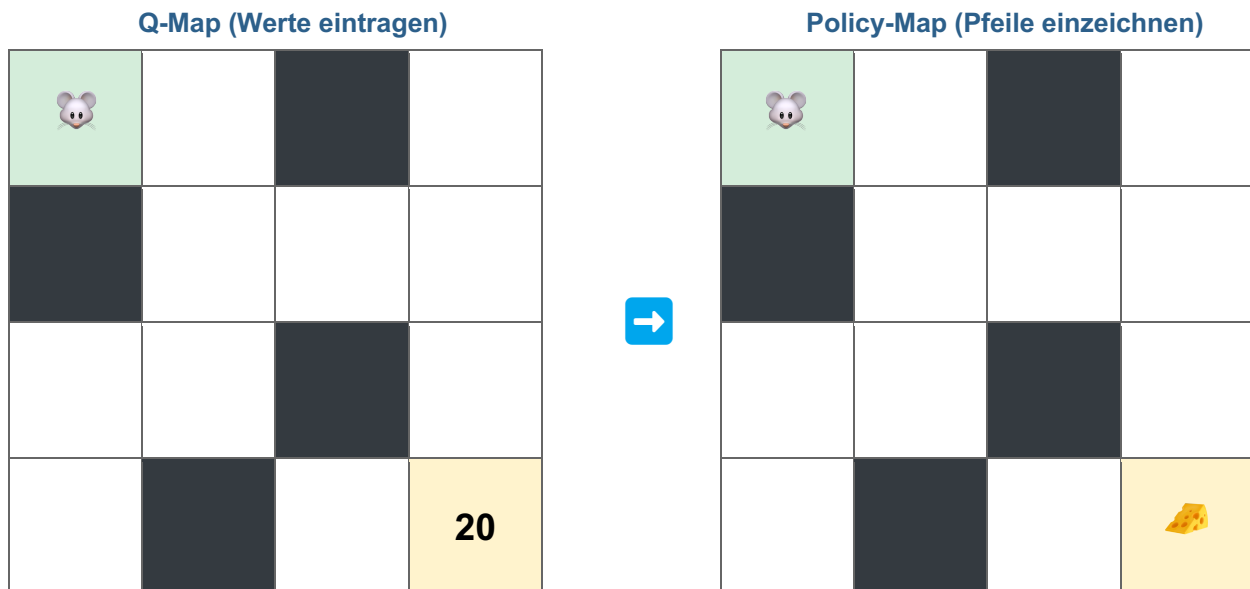
c) Verfolge die Pfeile von der Maus zum Käse. Beschreibe den Weg:

d) Gibt es noch einen zweiten Weg, der genauso gut ist? Wenn ja, welchen?

Aufgabe 2: Das große Labyrinth (4×4)

Jetzt wird es kniffliger! Das Spielfeld ist größer und es gibt mehrere Hindernisse. Aber die Regel bleibt dieselbe: **bester Nachbar – 1**

- Berechne alle Werte. Arbeite wieder in Wellen vom Käse aus.
- Zeichne die Pfeile ein.



- Schreibe den Weg der Maus zum Käse auf:

- Wie viele Schritte braucht die Maus mindestens? _____

- Stell dir vor, das Hindernis in der Mitte (Zeile 3, Spalte 3) würde entfernt. Was würde sich an der Q-Map ändern?

Zum Nachdenken

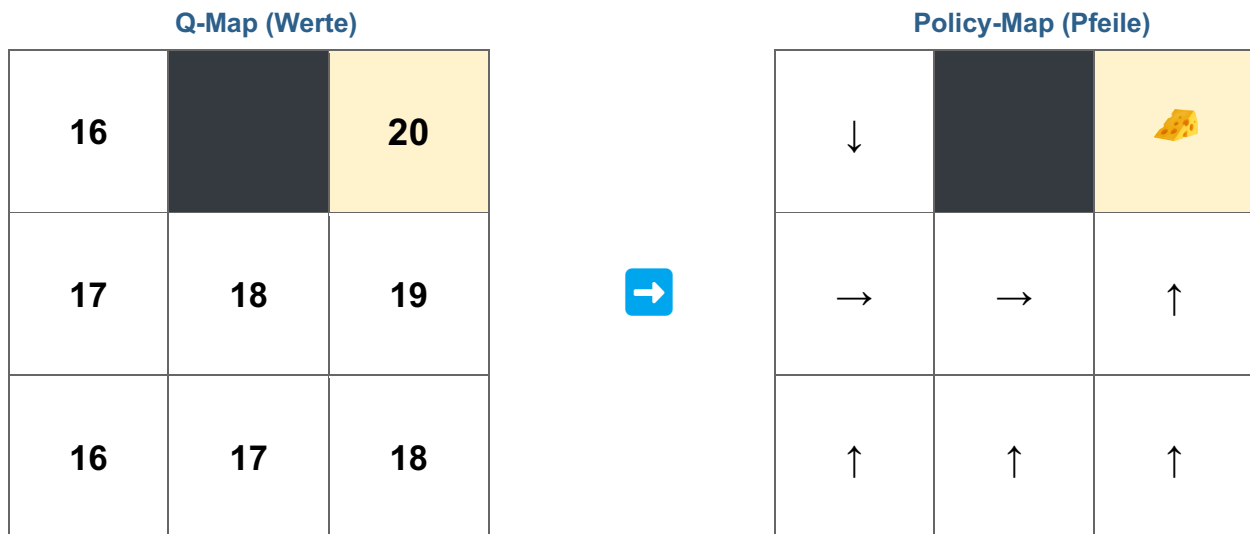
Was hat das mit Künstlicher Intelligenz zu tun? Du hast gerade von Hand gemacht, was ein Computer beim **Reinforcement Learning** automatisch tut: Er berechnet für jede Position einen Wert und folgt dann immer dem stärksten Anstieg zum Ziel. Deine Wellenausbreitung entspricht dem Training der KI – je mehr Wellen (= Trainingsdurchläufe), desto besser kennt sie den Weg.

 **Öffne jetzt die App** und vergleiche deine Lösung mit der Heatmap und den Pfeilen! Erkennst du deine Zahlen in den Farben wieder?

Lösungsblatt

(Nur für die Lehrkraft)

Lösung Aufgabe 1 (3×3)

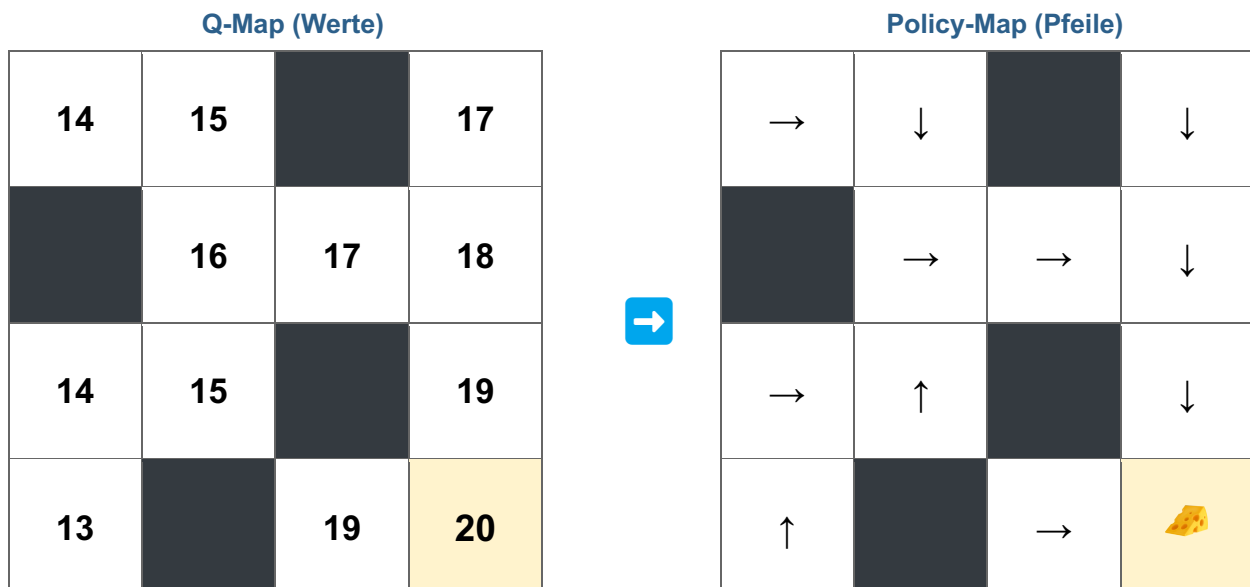


Wellenausbreitung: Käse=20 → Welle 1: 19 → Welle 2: 18 → Welle 3: 17 → Welle 4: 16 (4 Wellen)

c) **Weg 1:** 🐭 → hoch → rechts → rechts → hoch → 🧀 (4 Schritte)

d) **Weg 2:** 🐭 → rechts → rechts → hoch → hoch → 🧀 (4 Schritte) – Beide Wege sind gleich lang!

Lösung Aufgabe 2 (4×4)



Wellenausbreitung: 20 → 19 → 18 → 17 → 16 → 15 → 14 → 13 (7 Wellen – die Hindernisse verlängern den Weg!)

c) **Weg:** 🐭 → rechts → runter → rechts → rechts → runter → runter → 🧀

d) 6 Schritte.

e) Ohne das Hindernis bei (Zeile 3, Spalte 3) könnte das Wissen direkt von (3,2)=19 nach oben fließen. Mehrere Felder bekämen höhere Werte und die Maus könnte einen kürzeren Weg finden.