

Aktivität 2



Die Arena

Verstärkendes Lernen

Handbuch für Lehrkräfte,	Seite 1
Handbuch für Schüler:innen,	Seite 11
Präsentation im Klassenzimmer,	Seite 21

 Author
Joël RIVET

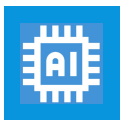
 Grafiker
David COHEN



Aktivität 2

Die Arena

Verstärkendes Lernen



Verstärkendes Lernen



Schwierigkeit
Mittel



Benötigte Zeit
60 min



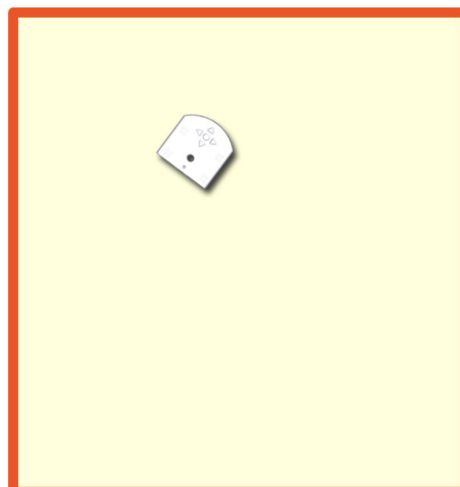
Voraussetzungen

- Lernen wir Thymio kennen.
- Wir haben die Aktivität 1 « Erste Schritte » mit ThymioAI abgeschlossen.



Materialvorbereitung

- Wir benötigen ein rechteckiges Gehäuse aus festen Wänden oder einen Raum, der durch Gegenstände begrenzt ist, die so schwer sind, dass Thymio sie nicht bewegen kann.
- Mindestabmessungen: 50cm x 50 cm.
- Wir stellen den Thymio-Roboter im Inneren der Arena auf.



Aktivität 2



Die Aufgabe

- Thymio befindet sich in einer Arena. Sein Ziel ist es, diese Arena zu erkunden, ohne die Wände zu berühren. Wie üblich weiss Thymio am Anfang nicht, was oder wie er es tun soll.
- Wir beschliessen, ihm eine neue Methode namens **verstärkendes Lernen** (*reinforcement learning*) beizubringen. Diese Methode gehört ebenfalls in den Bereich der künstlichen Intelligenz.
- Dazu werden wir wieder ein neuronales Netzwerk verwenden.
- Bei dieser Methode geben wir Thymio nicht, welche der verschiedenen Aktionen er ausführen soll. Es ist dem Roboter überlassen, eine zu wählen.
- Aber wie kann er wissen, welche Aktion er ausführen soll? Nun ... durch eine Belohnungsbewertung, die er erhält oder nicht [1].

[1] Ähnlich wie das Suchspiel: Kalt, wärmer, heiss!



Verstärkendes Lernen

- Wir trainieren nun Thymio so, indem wir ihm mitteilen, ob die Aktion, die er für eine bestimmte Sensorsituation gewählt hat, eine gute oder schlechte Wahl ist. Dafür vergeben wir eine Belohnung oder eine Strafe.
- Das neuronale Netz erkennt diese Belohnungen oder Strafen für jede ausgeführte Aktion. Wie wird das Netz reagieren? Wie ein Mensch oder ein Tier wird es versuchen, so viele Belohnungen wie möglich zu erhalten und Bestrafungen so weit wie möglich zu vermeiden.
- Das Netz nutzt die Mathematik, um voranzukommen. Falls jemand an der Nützlichkeit der Mathematik gezweifelt hat, hat hier ein gutes Beispiel für ihre Anwendung.
- Schritt für Schritt führt Thymio mehr und mehr "gute" Aktionen aus und immer weniger "schlechte".
- Um herauszufinden, welche Aktionen gut oder schlecht sind, muss der Roboter alle Handlungsoptionen erkunden.
- Von Zeit zu Zeit wird der Roboter also neue Aktionen ausprobieren, unabhängig von der Belohnung.



Aktion: vorwärts bewegen
Sensor: vorne nichts erkannt
Belohnung wird gegeben!



Aktion: vorwärts bewegen
Sensor: Wand vorne
Strafe wird gegeben!

Aktivität 2



AlphaAI-Einstellungen

Erinnerung daran, wie man Thymio verbindet

- Wir schliessen Thymio mit einer der folgenden zwei Methoden an:
 - Entweder schliessen wir den Roboter mit einem USB-Kabel an den Computer an,
 - oder wir schliessen einen Wireless-USB-Stick an und schalten den Roboter ein.
- Dann starten wir die Software « Thymio Suite » und warten ein paar Augenblicke.
- Nun starten wir die AlphaAI-Software.
- In AlphaAI wählen wir den Roboter aus, der auf dem blauen Bildschirm erscheint, um eine Verbindung herzustellen.

AlphaAI-Einstellungen

- Die Einstellungen befinden sich in einer Datei, die wir laden müssen.
- Im Menü **Settings** (Einstellungen) wählen wir die Option **Load sample settings ...** (Beispieleinstellungen laden).
- Es erscheint nun eine Box mit der Option **obstacle avoidance** (Hindernisse meiden), auf welche wir doppelklicken.
- Unter dem neuronalen Netzwerk erscheinen zwei schwarze Fortschrittsbalken.
- Der linke Balken misst die Belohnungen oder Bestrafungen als Zahl.
 - Eine Belohnung wird durch eine positive Zahl dargestellt. Ein negativer Wert entspricht einer Strafe. Mögliche Werte für die Belohnungen sind bereits durch AlphaAI festgelegt.
 - Der rechte Balken zeigt das Lernniveau an, d.h. den Durchschnitt der bisher erreichten Belohnung- und Strafpunkte. Das Niveau entspricht dem Lernzustand von Thymio.
- Die Schaltflächen **Learning** (Lernen) und **Exploration** (Erkunden) lassen wir aktiv.

Aktivität 2



Erste Übungsaufgabe

- Wir beobachten Thymio in der Arena.
- Wir starten den Selbststeuerungsmodus (*autonomous*). Thymio beginnt sich nun zu bewegen. Seine erste Bewegung wird zufällig ausgewählt. Beobachten wir seine Bewegungen und Farbveränderungen. Nach ca. 20 Sekunden klicken wir erneut auf die Schaltfläche *autonomous*, um den Lernprozess zu beenden.

Frage 1: Wir versuchen eine Verbindung herzustellen, zwischen dem, was der Roboter tut, und den Farben, die er verwendet. Schreiben wir unsere Antworten auf.

Wenn Thymio sich in der Arena vorwärts bewegt, ist er grün. Wenn er einer Wand zu nahe kommt oder sie berührt, wird er rot. Grün bedeutet, dass Thymio eine Belohnung erhält. Seine Aktion steht im Einklang mit seiner Mission. Rot bedeutet, dass Thymio eine Strafe erhält. Seine Aktion steht im Widerspruch zu seinem Auftrag.

Setzen wir das Training nun für 5-10 Minuten fort.

Frage 2: Bemerken wir nun eine Veränderung in Thymios Verhalten? Wenn ja, beschreiben wir sie. Schreiben wir unsere Antworten auf.

Wir beobachten, dass Thymio anfangs recht häufig gegen die Wände stößt und Mühe hat, von ihnen wegzukommen. Mit der Zeit gelingt es ihm immer öfter, sie zu meiden. Wir können sogar mehrere Stufen im Lernprozess feststellen:

Am Anfang entdeckte Thymio sehr schnell, dass es nicht nötig war, rückwärts zu fahren. Danach hat er sich schnell eine von zwei Verhaltensweisen angeeignet (die Schüler:innen werden die eine oder andere an ihrem Thymio beobachten):

Entweder dreht er sich im Kreis (das ist das häufigste Verhalten), oder er fährt geradeaus und dreht sich ab, wenn er auf eine Wand trifft. Dabei lernt er allmählich, abwechselnd geradeaus zu fahren und sich abzdrehen, wobei er immer häufiger geradeaus fährt, wenn kein Hindernis vor ihm ist, und nach links abbiegt, wenn rechts eine Wand steht, und nach rechts, wenn links eine Wand steht.

Aktivität 2



Die Rolle des Entdeckens

Wir klicken auf die Schaltfläche **explore** (*Entdecken*), um sie zu deaktivieren.

**Frage 3: Bemerken wir eine Entwicklung in Thymios Verhalten?
Wenn ja, beschreiben wir sie. Schreiben wir die Antwort auf.**

Thymio macht weniger Fehler. Er unterbricht seine geraden Linien nicht mehr durch unerwartete Bewegungen.

Wir setzen nur die künstliche Intelligenz mit der Schaltfläche *reset AI* (KI zurücksetzen) zurück. Dadurch vergisst Thymio alles, was er zuvor gelernt hat. Er fängt wieder von vorne an zu lernen. (Den Modus *autonomous* lassen wir aktiviert, aber den Modus *exploration* deaktiviert.)

Frage 4: Fallen uns Unterschiede zwischen dem neuen Lernen und dem vorherigen auf? Wenn ja, beschreiben wir sie. Wenn wir keinen Unterschied feststellen, starten wir einen neuen Lernprozess, indem wir auf *Reset AI* drücken. Schreiben wir die Antwort auf.

Thymio bleibt im Verhalten «im Kreis gehen» stecken, ohne die gerade Linie zu finden. (Dies geschieht nicht systematisch und Thymio lernt manchmal auch richtig, wenn der Modus **exploration** deaktiviert ist).

Erkenntnis

Erkundungen sind für das Lernen unerlässlich.

- Die Künstliche Intelligenz versucht gelegentlich andere Aktionen aus, als die, die sie für die beste hält (wenn das passiert, leuchtet das Aktionssymbol rechts auf dem Bildschirm blau statt schwarz). Dadurch wird vermieden, dass sie in einem schlechten Verhalten stecken bleibt.
- Andererseits ist der Erkundungsmodus **exploration** nach Abschluss des Lernprozesses nicht mehr nützlich, so dass es sich lohnt, ihn auszuschalten, um ein möglichst perfektes Verhalten zu erzielen.

Aktivität 2



Das Verhalten des neuronalen Netzes

Beobachten wir nun das Verhalten des neuronalen Netzwerks

Wir werden das Verhalten zu Beginn des Lernprozesses in einigen Schritten detailliert beobachten und beschreiben.

1. Wir setzen die Künstliche Intelligenz mit der Schaltfläche **reset AI** zurück und platzieren Thymio in der Mitte des Spielfeldes. Es ist zu beachten, dass Thymio seinen ersten Zug zufällig wählen wird. Um sicher zu gehen, können wir mehrmals hintereinander auf **reset AI** und **autonomous** klicken.
2. Wir füllen nun die erste Zeile der Tabelle aus. Ein kleiner Strich bedeutet, dass die vorderen Sensoren von Thymio nichts erkennen, weil nichts da ist.
3. Wir beobachten Thymio nun genau und klicken auf die Schaltfläche **step by step** (Schritt für Schritt).

«Schritt für Schritt» Phase	Sensoren vorne	Nach rückwärts links bewegen	Nach vorwärts links bewegen	Vorwärts bewegen	Nach rückwärts rechts bewegen	Nach vorwärts rechts bewegen	Belohnungs-niveau	Lern-niveau

Aktivität 2

Frage 5: Welche Bewegungen hat Thymio ausgeführt? Schreiben wir in die Tabellenzeile, die wir gerade ausgefüllt haben, den Wert, der dieser Bewegung entspricht. Vergleichen wir diese Zahl mit den Werten der anderen Aktionen. Was stellen wir fest? Schreiben wir die Antworten auf.

Dieser Wert wird der höchste sein.

Das bedeutet, dass Thymio die Bewegung ausführt, die dem höchsten Wert entspricht.

Frage 6: Wir sehen auch, dass der Roboter eine Belohnung erhalten hat. Steht diese Belohnung im Einklang mit dem Zweck der Mission? Erklären wir.

Mögliche Antwort: Thymio ist nach links abgebogen und hat eine Belohnung von +55 erhalten. Das ist zu erwarten, denn es ist kein Hindernis vor ihm, so dass er abbiegen kann.

Sobald die erste Belohnung gegeben wurde, werden die Ausgabewerte vom neuronalen Netz neu errechnet.

Wir füllen nun die 2. Zeile der Tabelle aus und raten, was die nächste Bewegung von Thymio sein wird.

Wir klicken noch ein paar Mal auf die Schaltfläche *step by step* und beobachten dabei die Entwicklung des Belohnungsniveaus und des Lernniveaus.

Aktivität 2



Die Rolle der Belohnungen

Beobachten wir nun das Verhalten des neuronalen Netzwerks

Wir werden das Verhalten zu Beginn des Lernprozesses in einigen Schritten detailliert beobachten und beschreiben.

1. Wir setzen die Künstliche Intelligenz mit der Schaltfläche **reset AI** zurück und platzieren Thymio in der Mitte des Spielfeldes. Es ist zu beachten, dass Thymio seinen ersten Zug zufällig wählen wird. Um sicher zu gehen, können wir mehrmals hintereinander auf **reset AI** und **autonomous** klicken.
2. Wir füllen nun die erste Zeile der Tabelle aus. Ein kleiner Strich bedeutet, dass die vorderen Sensoren von Thymio nichts erkennen, weil nichts da ist.
3. Wir beobachten Thymio nun genau und klicken auf die Schaltfläche **step by step** (Schritt für Schritt).

«Schritt für Schritt»-Phase	Sensoren vorne	Nach rückwärts links bewegen	Nach vorwärts links bewegen	Vorwärts bewegen	Nach rückwärts rechts bewegen	Nach vorwärts rechts bewegen	Belohnungs-niveau	Lern-niveau

Aktivität 2

Frage 7: Wie ändert sich das Niveau, wenn der Roboter eine Belohnung oder im Gegenteil eine Strafe erhält? Wie verhält sich das Niveau? Schreiben wir die Antwort auf.

- Wenn die Belohnung positiv ist, steigt das Niveau.
- Wenn die Belohnung negativ ist, sinkt das Niveau.

Das Niveau bildet die Fähigkeit von Thymio ab, positive Belohnungen zu erhalten. Es wird als Durchschnitt der Belohnungs- und Strafpunkte berechnet, die Thymio in der letzten Minute erhalten hat.

Drücken wir nun auf den Selbststeuerungsmodus *autonomous*, um Thymio weiter lernen zu lassen.

Frage 8: Wie verändert sich das Niveau während des Lernprozesses? Warum? Erklären wir.

Das Niveau steigt im Laufe des Lernprozesses. Thymio erhält nämlich immer mehr hohe Belohnungen (vor allem, wenn er geradeaus fährt) und immer weniger Strafen (da er immer seltener anstößt). Tatsächlich besteht das Ziel des Lernens genau darin, das Niveau von Thymio zu steigern.

Aktivität 2



Der Einfluss einer Strafe auf das erlernte Verhalten

Wenn wir die verschiedenen Werte anschauen, die in der Fortschrittsanzeige erscheinen, stellen wir Folgendes fest.

- **100:** Wenn Thymio geradeaus fährt und kein Hindernis vor sich hat. Dies ist der höchste Wert.
- **55:** Wenn Thymio abdreht, ohne ein Hindernis vor sich zu haben.
- **-50:** Wenn Thymio eine der zahlreichen "schlechten" Aktionen ausführt, wie z. B. gegen eine Wand zu fahren oder rückwärts zu fahren, wenn nichts vor ihm ist etc.

Wir können den Wert der Strafe ändern. Wir öffnen dazu die Registerkarte **reward** (*Belohnung*) und setzen die Strafe auf einen kleinen Wert, zum Beispiel null. Dann setzen wir die Künstliche Intelligenz mittels **reset AI** zurück und starten das Lernen für ein paar Minuten neu. Wir beobachten dabei das Verhalten von Thymio:

Ist er eher wagemutig oder vorsichtig? Fährt er mehr oder weniger oft gegen die Wände?

Legen wir nun eine höhere Strafe fest, zum Beispiel 1.5 Punkte. Wir setzen die Künstliche Intelligenz erneut zurück und starten den Lernprozess für ein paar Minuten neu. Es stellt sich die selbe Frage zu Thymios Verhalten wie zuvor.

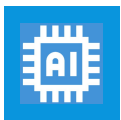
Frage 9: Fassen wir nun zusammen, wie sich das Verhalten von Thymio ändert, wenn wir den Strafwert ändern. Schreiben wir unsere Antwort auf.

- Wenn die Strafe niedrig ist, stösst Thymio oft an die Wände, wird mit der Zeit aber mutiger und erkundet das Gebiet vollständig.
- Wenn die Strafe hoch ist, stösst Thymio seltener gegen die Wände, wird aber vorsichtiger und bleibt in einem begrenzten Gebiet.

Aktivität 2

Die Arena

Verstärkendes Lernen



Verstärkendes Lernen



Schwierigkeit
Mittel

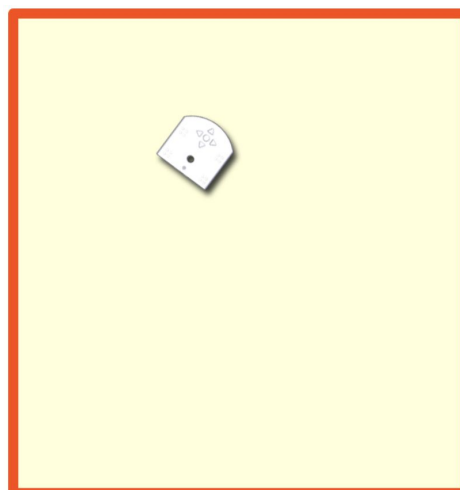


Benötigte Zeit
60 min



Die Mission von Thymio - sein Lernprinzip

- Thymio befindet sich in einer Arena. Sein Ziel ist es, diese Arena zu erkunden, ohne die Wände zu berühren. Wie üblich weiss Thymio am Anfang nicht, was oder wie er es tun soll.
- Wir beschliessen, ihm eine neue Methode namens **verstärkendes Lernen** (*reinforcement learning*) beizubringen. Diese Methode gehört ebenfalls in den Bereich der künstlichen Intelligenz.
- Dazu werden wieder ein neuronales Netzwerk verwenden.
- Bei dieser Methode geben wir Thymio nicht, welche der verschiedenen Aktionen er ausführen soll. Es ist dem Roboter überlassen, eine zu wählen.
- Aber wie kann er wissen, welche Aktion er ausführen soll? Nun ... durch eine Belohnungsbewertung, die er erhält oder nicht[1].



[1] Ähnlich wie das Suchspiel: Kalt, wärmer, heiss!

Aktivität 2



Verstärkendes Lernen

- Wir trainieren nun Thymio so, indem wir ihm mitteilen, ob die Aktion, die er für eine bestimmte Sensorsituation gewählt hat, eine gute oder schlechte Wahl ist. Dafür vergeben wir eine Belohnung oder eine Strafe.
- Das neuronale Netz erkennt diese Belohnungen oder Strafen für jede ausgeführte Aktion. Wie wird das Netz reagieren? Wie ein Mensch oder ein Tier wird es versuchen, so viele Belohnungen wie möglich zu erhalten und Bestrafungen so weit wie möglich zu vermeiden.
- Das Netz nutzt die Mathematik, um voranzukommen. Falls jemand an der Nützlichkeit der Mathematik gezweifelt hat, hat hier ein gutes Beispiel für ihre Anwendung.
- Schritt für Schritt führt Thymio mehr und mehr "gute" Aktionen aus und immer weniger "schlechte".
- Um herauszufinden, welche Aktionen gut oder schlecht sind, muss der Roboter alle Handlungsoptionen erkunden.
- Von Zeit zu Zeit wird der Roboter also neue Aktionen ausprobieren, unabhängig von der Belohnung.



Youpi !

Aktion: vorwärts bewegen
Sensor: vorne nichts erkannt
Belohnung wird gegeben!



Ouille !

Aktion: vorwärts bewegen
Sensor: Wand Vorne
Strafe wird gegeben!

Aktivität 2



AlphaAI-Einstellungen

Erinnerung daran, wie man Thymio verbindet

- Wir schliessen Thymio mit einer der folgenden zwei Methoden an:
 - Entweder schliessen wir den Roboter mit einem USB-Kabel an den Computer an,
 - oder wir schliessen einen Wireless-USB-Stick an und schalten den Roboter ein.
- Dann starten wir die Software « Thymio Suite » und warten ein paar Augenblicke.
- Nun starten wir die AlphaAI-Software.
- In AlphaAI wählen wir den Roboter aus, der auf dem blauen Bildschirm erscheint, um eine Verbindung herzustellen.

AlphaAI-Einstellungen

- Die Einstellungen befinden sich in einer Datei, die wir laden müssen.
- Im Menü **Settings** (Einstellungen) wählen wir die Option **Load sample settings ...** (Beispieleinstellungen laden).
- Es erscheint nun eine Box mit der Option **obstacle avoidance** (Hindernisse meiden), auf welche wir doppelklicken.
- Unter dem neuronalen Netzwerk erscheinen zwei schwarze Fortschrittsbalken.
- Der linke Balken misst die Belohnungen oder Bestrafungen als Zahl.
 - Eine Belohnung wird durch eine positive Zahl dargestellt. Ein negativer Wert entspricht einer Strafe. Mögliche Werte für die Belohnungen sind bereits durch AlphaAI festgelegt.
 - Der rechte Balken zeigt das Lernniveau an, d.h. den Durchschnitt der bisher erreichten Belohnung- und Strafpunkte. Das Niveau entspricht dem Lernzustand von Thymio.
- Die Schaltflächen **Learning** (Lernen) und **Exploration** (Erkunden) lassen wir aktiv.

Aktivität 2



Erste Übungsaufgabe

- Beobachten wir Thymio in der Arena.
- Wir starten den Selbststeuerungsmodus (**autonomous**). Thymio beginnt sich nun zu bewegen. Seine erste Bewegung wird zufällig ausgewählt. Beobachten wir seine Bewegungen und Farbveränderungen. Nach ca. 20 Sekunden klicken wir erneut auf die Schaltfläche **autonomous**, um den Lernprozess zu beenden.

Frage 1: Wir versuchen eine Verbindung herzustellen, zwischen dem, was der Roboter tut, und den Farben, die er verwendet. Schreiben wir unsere Antworten auf.

Setzen wir das Training für 5-10 Minuten fort.

Frage 2: Bemerken wir nun eine Veränderung in Thymios Verhalten? Wenn ja, beschreiben wir sie. Schreiben wir unsere Antworten auf.



Die Rolle des Entdeckens

Wir klicken auf die Schaltfläche **explore** (Entdecken), um sie zu deaktivieren.

Frage 3: Bemerken wir eine Entwicklung in Thymios Verhalten? Wenn ja, beschreiben wir sie. Schreiben wir die Antwort auf.

Wir setzen nur die künstliche Intelligenz mit der Schaltfläche **reset AI** (KI zurücksetzen) zurück. Dadurch vergisst Thymio alles, was er zuvor gelernt hat. Er fängt wieder von vorne an zu lernen. (Den Modus **autonomous** lassen wir aktiviert, aber den Modus **exploration** deaktiviert.)

Frage 4: Fallen uns Unterschiede zwischen dem neuen Lernen und dem vorherigen auf? Wenn ja, beschreiben wir sie. Wenn wir keinen Unterschied feststellen, starten wir einen neuen Lernprozess, indem wir auf Reset AI drücken. Schreiben wir die Antwort auf.

Aktivität 2



Die Rolle der Belohnungen

Beobachten wir nun das Verhalten des neuronalen Netzwerks

Wir werden das Verhalten zu Beginn des Lernprozesses in einigen Schritten detailliert beobachten und beschreiben.

1. Wir setzen die Künstliche Intelligenz mit der Schaltfläche **reset AI** zurück und platzieren Thymio in der Mitte des Spielfeldes. Es ist zu beachten, dass Thymio seinen ersten Zug zufällig wählen wird. Um sicher zu gehen, können wir mehrmals hintereinander auf **reset AI** und **autonomous** klicken.
2. Wir füllen nun die erste Zeile der Tabelle aus. Ein kleiner Strich bedeutet, dass die vorderen Sensoren von Thymio nichts erkennen, weil nichts da ist.
3. Wir beobachten Thymio nun genau und klicken auf die Schaltfläche **step by step** (Schritt für Schritt).

«Schritt für Schritt»-Phase	Sensoren vorne	Nach rückwärts links bewegen	Nach vorwärts links bewegen	Vorwärts bewegen	Nach rückwärts rechts bewegen	Nach vorwärts rechts bewegen	Belohnungs-niveau	Lern-niveau

Aktivität 2

Frage 5: Welche Bewegungen hat Thymio ausgeführt? Schreiben wir in die Tabellenzeile, die wir gerade ausgefüllt haben, den Wert, der dieser Bewegung entspricht. Vergleichen wir diese Zahl mit den Werten der anderen Aktionen. Was stellen wir fest? Schreiben wir die Antworten auf.

Frage 6: Wir sehen auch, dass der Roboter eine Belohnung erhalten hat. Steht diese Belohnung im Einklang mit dem Zweck der Mission? Erklären wir.

Sobald die erste Belohnung gegeben wurde, werden die Ausgabewerte vom neuronalen Netz neu errechnet.

Wir füllen nun die 2. Zeile der Tabelle aus und raten, was die nächste Bewegung von Thymio sein wird.

Wir klicken noch ein paar Mal auf die Schaltfläche **step by step** und beobachten dabei die Entwicklung des Belohnungsniveaus und des Lernniveaus.

Frage 7: Wie ändert sich das Niveau, wenn der Roboter eine Belohnung oder im Gegenteil eine Strafe erhält? Wie verhält sich das Niveau? Schreiben wir die Antwort auf.

Drücken wir nun auf den Selbststeuerungsmodus **autonomous**, um Thymio weiter lernen zu lassen.

Frage 8: Wie verändert sich das Niveau während des Lernprozesses? Warum? Erklären wir.

Aktivität 2



Der Einfluss einer Strafe auf das erlernte Verhalten

Wenn wir die verschiedenen Werte anschauen, die in der Fortschrittsanzeige erscheinen, stellen wir Folgendes fest.

- 100: Wenn Thymio geradeaus fährt und kein Hindernis vor sich hat. Dies ist der höchste Wert.
- 55: Wenn Thymio abdreht, ohne ein Hindernis vor sich zu haben.
- -50: Wenn Thymio eine der zahlreichen "schlechten" Aktionen ausführt, wie z. B. gegen eine Wand zu fahren oder rückwärts zu fahren, wenn nichts vor ihm ist etc.

Wir können den Wert der Strafe ändern. Wir öffnen dazu die Registerkarte **reward** (*Belohnung*) und setzen die Strafe auf einen kleinen Wert, zum Beispiel null. Dann setzen wir die Künstliche Intelligenz mittels **reset AI** zurück und starten das Lernen für ein paar Minuten neu. Wir beobachten dabei das Verhalten von Thymio:

Ist er eher wagemutig oder vorsichtig? Fährt er mehr oder weniger oft gegen die Wände?

Legen wir nun eine höhere Strafe fest, zum Beispiel 1.5 Punkte. Wir setzen die Künstliche Intelligenz erneut zurück und starten den Lernprozess für ein paar Minuten neu. Es stellt sich die selbe Frage zu Thymios Verhalten wie zuvor.

Frage 9: Fassen wir nun zusammen, wie sich das Verhalten von Thymio ändert, wenn wir den Strafwert ändern. Schreiben wir unsere Antwort auf.

- Wenn die Strafe niedrig ist, stösst Thymio oft an die Wände, wird mit der Zeit aber mutiger und erkundet das Gebiet vollständig.
- Wenn die Strafe hoch ist, stösst Thymio seltener gegen die Wände, wird aber vorsichtiger und bleibt in einem begrenzten Gebiet.

Aktivität 2



Antworten

Name:

Klasse:

Frage 1: Wir versuchen eine Verbindung herzustellen, zwischen dem, was der Roboter tut, und den Farben, die er verwendet. Schreiben wir unsere Antworten auf.

Frage 2: Bemerkten wir nun eine Veränderung in Thymios Verhalten? Wenn ja, beschreiben wir sie. Schreiben wir unsere Antworten auf.

Frage 3: Bemerkten wir eine Entwicklung in Thymios Verhalten? Wenn ja, beschreiben wir sie. Schreiben wir die Antwort auf.

Aktivität 2



Antworten

Name:

Klasse:

Frage 4: Fallen uns Unterschiede zwischen dem neuen Lernen und dem vorherigen auf? Wenn ja, beschreiben wir sie. Wenn wir keinen Unterschied feststellen, starten wir einen neuen Lernprozess, indem wir auf *Reset AI* drücken. Schreiben wir die Antwort auf.

Frage 5: Welche Bewegungen hat Thymio ausgeführt? Schreiben wir in die Tabellenzeile, die wir gerade ausgefüllt haben, den Wert, der dieser Bewegung entspricht. Vergleichen wir diese Zahl mit den Werten der anderen Aktionen. Was stellen wir fest? Schreiben wir die Antworten auf.

Frage 6: Wir sehen auch, dass der Roboter eine Belohnung erhalten hat. Steht diese Belohnung im Einklang mit dem Zweck der Mission? Erklären wir.

Aktivität 2



Antworten

Name:

Klasse:

Frage 7: Wie ändert sich das Niveau, wenn der Roboter eine Belohnung oder im Gegenteil eine Strafe erhält? Wie verhält sich das Niveau? Schreiben wir die Antwort auf.

Frage 8: Wie verändert sich das Niveau während des Lernprozesses? Warum? Erklären wir.

Frage 9: Fassen wir nun zusammen, wie sich das Verhalten von Thymio ändert, wenn wir den Strafwert ändern. Schreiben wir unsere Antwort auf.

Die Arena



Die Mission

- Thymio befindet sich in einer geschlossenen Arena.
- Sein Ziel ist es, diese Arena zu erkunden, ohne die Wände zu berühren.



Verstärkendes Lernen

- Jedes Mal, wenn Thymio eine Aktion ausführt, wird ihm mitgeteilt, ob die Aktion, die er in einer gegebenen Sensorsituation ausgeführt hat, eine gute oder schlechte Wahl ist, indem Belohnungs- oder Strafpunkte vergeben werden.
- Thymio wählt seine Aktionen so, dass er die maximale Belohnung erhält.

Die Arena

Verstärkendes Lernen



Youpi !

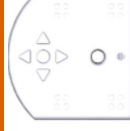


Aktion: vorwärts bewegen
Sensor: vorne nichts erkannt

Belohnung wird gegeben!



Ouille !



Aktion: vorwärts bewegen
Sensor: Wand vorne

Strafe wird gegeben!



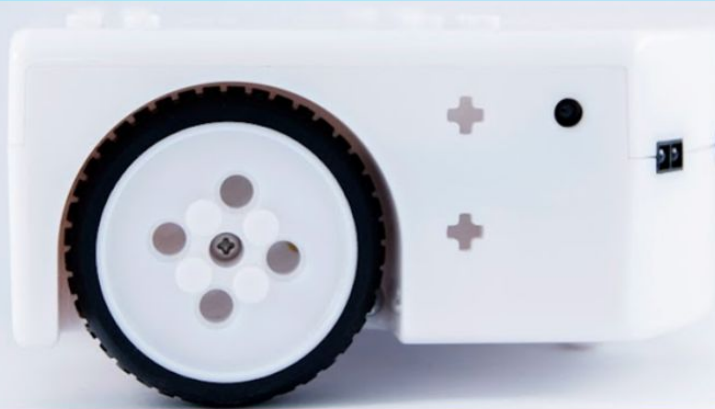
Erinnerung daran, wie man Thymio verbindet

- Wir schliessen Thymio mit einer der folgenden zwei Methoden an:
- Entweder schliessen wir den Roboter mit einem USB-Kabel an den Computer an,
 - oder wir schliessen einen Wireless-USB-Stick an und schalten den Roboter ein.
- Dann starten wir die Software « Thymio Suite » und warten ein paar Augenblicke.
- Nun starten wir die AlphaI-Software.
- In AlphaI wählen wir den Roboter auf, der auf dem blauen Bildschirm erscheint, um eine Verbindung herzustellen.



«Schritt für Schritt»-Tabelle

Step by Step phase	Front sensors	Move backward to the left	Move forward to the left	Move forward	Move backward to the right	Move forward to the right	Reward value	Level value



www.thymio.org

thymio
by MÖBSYA

Chemin du Closel 3, 1020 Renens - Switzerland
info@thymio.org