

Autonomes Ansteuern einer Vorrichtung unter Verwendung eines technischen Motivationswerts

Die vorliegende Erfindung ist auf ein Verfahren gerichtet zum autonomen Ansteuern einer Vorrichtung, wobei sich die Vorrichtung generell in einer physischen realen Welt bewegt und
5 bezüglich ihrer physischen Ausgestaltung nicht weiter spezifiziert werden muss. Das Verfahren schafft den Vorteil, dass die Vorrichtung ein autonomes Lernen durchführt und das erlernte Wissen bzw. Verhalten kontinuierlich verbessert. Generell wird der Nachteil im Stand der Technik überwunden, dass erst Trainingsdaten geschaffen werden müssen, wie dies bei gängigen Methoden der künstlichen Intelligenz der Fall ist. Generell ist das
10 Verfahren universell einsetzbar und die Vorrichtung lernt selbsttätig und korrigiert hierbei stets die eigene Wissensbasis. Darüber hinaus wird eine Vorrichtung vorgeschlagen, die zum Ausführen des Verfahrens eingerichtet ist, sowie eine Systemanordnung aufweisend mehrere der vorgeschlagenen Vorrichtungen. Darüber hinaus wird ein Computerprogrammprodukt und ein computerlesbares Speichermedium vorgeschlagen,
15 welche die Verfahrensschritte ausführen bzw. einen Computer veranlassen das Verfahren auszuführen.

Aus dem Stand der Technik sind unterschiedliche Verfahren aus der künstlichen Intelligenz KI, im Englischen auch als Artificial Intelligence AI bezeichnet, bekannt, welche es vorsehen, dass Trainingsdaten bereitgestellt werden und Algorithmen in einer Lernphase überwacht
20 Gesetzmäßigkeiten erkennen. Hierbei ist es möglich, dass spezielle Algorithmen in der Datenbasis Gesetzmäßigkeiten erkennen und diese sodann extrahieren. Somit wird also innerhalb großer Datenmengen implizites Wissen extrahiert. Ist die Trainingsphase abgeschlossen, so werden die entsprechend trainierten Algorithmen auf tatsächliche Daten angewendet und generieren hieraus wiederum implizite Abhängigkeiten zur Lösung
25 realweltlicher Probleme. Ein Nachteil ist es hierbei, dass die Trainingsdaten erst bereitgestellt werden müssen und bereits hier unerwünschter Einfluss auf die Lernphase genommen werden kann, was zudem nicht nur fehleranfällig, sondern auch aufwendig ist.

Ferner ist aus dem Stand der Technik eine Schwarmintelligenz bekannt, wobei mehrere Vorrichtungen bereitgestellt werden und diese lösen sodann kollaborativ in verteilter Weise
30 ein Problem. Auch hier ist die Ansteuerung und Koordination der einzelnen Teilnehmer aufwendig und teilweise fehleranfällig. Dies trifft insbesondere dann zu, wenn ein verteiltes Lernen stattfinden soll, welches wiederum koordiniert werden muss.

Weiter aus dem Stand der Technik sind künstliche neuronale Netze bekannt, welche in

Anlehnung an die Graphentheorie Neuronen und Verbindungen, also Knoten und Kanten, bereitstellen. Hierzu ist es bekannt, dass diese Netzwerke Funktionen des menschlichen Gehirns nachahmen und hierbei auch lernen können. So können Kantengewichtungen variiert werden und es können neue Kanten hinzugefügt bzw. alte gelöscht werden sowie bestehende Knoten können deaktiviert werden oder aber auch neue Knoten hinzugefügt werden. In dieser Weise ergibt sich ein dynamisch lernendes Gesamtsystem.

Die Verwendung von einem KNN in der oben präsentierten Weise hat allerdings mindestens vier Probleme:

1. Bei Objekten, auf die das KNN nicht trainiert wurde, liefert das KNN keine Information, um die dem Objekt zugeordneten Routinen aufzurufen.
2. Die einprogrammierten Eigenschaften oder Verhaltensweisen der erkannten und identifizierten Objekte können entweder falsch sein, sich zwischenzeitlich geändert haben oder noch unbekannt sein.
3. Ein oder mehrere Wissenschaftler, Fachleute etc. wählen die Trainingsdaten aus und geben die zu erzielenden Ergebnisse vor. Auch beim unüberwachten Lernen werden die Trainingsdaten und Hyperparameter immer noch vom Menschen vorgegeben und das kann Fehler nicht ausschließen.
4. Ein KNN kann Fehler nicht korrigieren. Jede Einzelinformation ist über alle Parameter des KNN verteilt, ähnlich einem Hologramm, in dem jeder Datenpunkt Informationen über das gesamte Bild enthält. Ein KNN muss also immer ganz gelöscht und komplett neu trainiert werden.

Die vorliegend als Ego benannte Vorrichtung versagt hingegen nicht, wenn es auf unbekannte Objekte trifft, anders als KNNs, wie zum Beispiel im Video, als das selbstfahrende Auto plötzlich zwei Kisten auf der Fahrbahn sieht. KNN steht generell für ein Künstliches Neuronales Netz.

Im Stand der Technik besteht ein Bedarf ein autonomes Ansteuern von Vorrichtungen zu schaffen, derart, dass aufwendige Vorarbeiten wie das Bereitstellen von ausgewählten und interpretierten Trainingsdaten und das Einlernen verhindert bzw. minimiert werden können. Es besteht also ein Bedarf an einem selbstlernenden System, welches auch in der Lage ist, kollektiv zu lernen bzw. abschätzen zu können, was andere Teilnehmer voraussichtlich tun werden bzw. überhaupt tun können. Somit besteht ein Bedarf an einem System, welches mehrere Teilnehmer umfasst, wobei jeder Teilnehmer tatsächlich auch Wissen über die anderen Teilnehmer vorhält und dieses weiter entwickelt, was vorliegend als Empathie

bezeichnet wird.

Dementsprechend ist es eine Aufgabe der vorliegenden Erfindung ein Verfahren zum autonomen Ansteuern einer Vorrichtung vorzuschlagen, welches lediglich anhand von bereitgestellten Zielinformationen selbstständig agiert und hierbei Handlungswissen aufbaut.

- 5 Das vorgeschlagene Verfahren soll Handlungsspielräume von Effektoren jeglicher Vorrichtung erkennen können und ein verbessertes Verfahren oder zumindest ein alternatives Verfahren zum autonomen Ansteuern der Vorrichtung bzw. der Vorrichtungen schaffen. Darüber hinaus ist es eine Aufgabe der vorliegenden Erfindung eine Vorrichtung zum Ausführen des Verfahrens vorzuschlagen sowie eine Systemanordnung aufweisend
- 10 mehrere der vorgeschlagenen Vorrichtungen. Darüber hinaus ist es eine Aufgabe ein Computerprogrammprodukt und ein computerlesbares Speichermedium vorzuschlagen, welche Befehle vorhalten, welche das Verfahren ausführen.

Die Aufgabe wird gelöst mit den Merkmalen des Patentanspruchs 1. Weitere vorteilhafte Ausgestaltungen sind in den Unteransprüchen angegeben.

- 15 Demgemäß wird ein Verfahren zum autonomen Ansteuern einer Vorrichtung vorgeschlagen, aufweisend ein Initialisieren mittels randomisiertem Betätigen von Effektoren und Auslesen einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen Vorrichtungseigenschaften und einem Auslesen von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften; ein
- 20 Abspeichern von Wechselbeziehungen, welche zwischen dem Betätigen der Effektoren und den ausgelesenen Vorrichtungseigenschaften und Umgebungseigenschaften bestehen als Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von Betätigen, Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt; und ein Ansteuern der Vorrichtung gemäß mindestens einer vordefinierten
- 25 Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer abgespeicherten Handlungsanweisung, welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und/ oder die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird.
- 30 Das vorgeschlagene Verfahren dient dem autonomen Ansteuern, das heißt also dem Erstellen von Handlungsanweisungen, einer Vorrichtung, welche generell physisch ausgeprägt ist. Hierbei kann es sich um ein Automobil handeln oder aber auch um einen Roboter oder eine Produktionsanlage. Bezüglich der Mobilität sind vorliegend keinerlei

Einschränkungen vorgesehen, sodass sich die Vorrichtung an Land, in der Luft oder auf dem Wasser oder sogar unter dem Wasser bewegen kann. Autonom heißt in diesem Zusammenhang, dass das Verfahren dafür sorgt, dass die Vorrichtung selbsttätig lernt und auch selbsttätig erkennt, welche Handlungsmöglichkeiten zur Verfügung stehen. Diese
5 können erlernt werden und es wird mittels der Sensorik erlernt, welche Handlung von welchem Ausgangspunkt zu welchem Ergebnis führt. Generell ist es möglich, dass sich die Vorrichtung eigene Zielvorgaben bereitstellt oder aber die Zielvorgaben werden extern übermittelt. Bei eigenen Zielvorgaben kann es sich beispielsweise um die Aufrechterhaltung der Betriebsfähigkeit handeln. Eine interne Zielsetzung kann es beispielsweise sein bei
10 einem niedrigen Batteriestand eine Ladestation aufzusuchen. Eine externe Zielsetzung kann dadurch übermittelt werden, dass angegeben wird, welche Aufgabe bzw. welche Tätigkeit die Vorrichtung erledigen soll.

In einem vorbereitenden Verfahrensschritt erfolgt ein Initialisieren mittels randomisierten Betätigens von Effektoren. Ein Effektor ist generell eine physische Einheit, welche einen
15 Einfluss auf die reale Welt hat. Dies kann sich beispielsweise auf die Vorrichtung selbst auswirken, wie Lenken, Beschleunigen, Bremsen etc., und/ oder auf die gezielte Manipulation von Objekten durch zum Beispiel Roboterarme, direkt oder mit Werkzeugen, oder durch Instrumente mit Fernwirkung, seien sie Geschosse oder Strahlen oder zum Beispiel Mittel zur Brandbekämpfung wie Wasser oder Löschschäume. Bei dem Beispiel des
20 Automobils kann es sich um die Bremse, dem Gaspedal etc. aber auch um Scheinwerfer, Blinker usw. handeln. Die Effektoren werden betätigt und sodann wird anhand der internen und externen Sensorik gemessen, welche Auswirkungen diese Handlungen haben. Dies wird mitprotokolliert und kann sodann im weiteren Verfahrensverlauf verwendet werden. Auf diese Art und Weise werden Handlungen erlernt und diese in späteren Verfahrensschritten
25 zielführend eingesetzt.

Da die vorgeschlagene Vorrichtung bzw. das vorgeschlagene Verfahren gänzlich ohne initialem Wissen durchgeführt bzw. angesteuert werden kann, handelt es sich bei dem vorbereitenden Schritt um ein randomisiertes Betätigen, also ein willkürliches Betätigen, da erst erlernt werden muss, welche Handlung überhaupt möglich sind und was sie bewirken
30 können, ohne dass sie bereits im Hinblick auf ein Ziel durchgeführt bzw. trainiert werden, damit das Erlernte nicht an bestimmte externe Ziele gebunden ist und womöglich später bei anderen externen Zielen nicht angewendet werden könnte.. In weiteren Iterationen des Verfahrens erfolgt dann ein zielgerichtetes Betätigen der erlernten Handlungen. Bei dem randomisierten Betätigen werden die Systemparameter der Effektoren abgeprüft und

beispielsweise ein Roboterarm wird in alle möglichen Positionen verbracht. Dies wird jeweils mitprotokolliert und es wird erkannt, welche Aktion des Effektors sich wie auf die reale Welt auswirkt. Bei dem Beispiel eines Scheinwerfers ist es möglich, dass dieser nur die Parameter an und aus bereitstellt. Sodann erfolgt also ein An- und Ausschalten und es wird
5 mittels der externen Sensorik gemessen, welchen Einfluss der Scheinwerfer auf die Umwelt hat und mittels der internen Sensorik werden Betriebsparameter des Scheinwerfers, wie beispielsweise eine Temperaturentwicklung, gemessen. Bei weiterentwickelten LED-Scheinwerfern ist es zudem möglich, sowohl die Farbe als auch die Intensität einzustellen. Dies wird mittels randomisierten Betätigens durchprobiert und sodann interne und externe
10 Resultate protokolliert.

Erfindungsgemäß sind interne Sensorikeinheiten und externe Sensorikeinheiten vorgeschlagen. Intern und extern bezieht sich hierbei nicht auf die Anordnung der Sensorikeinheiten, sondern vielmehr wird gemäß der vorliegenden Erfindung damit bezeichnet, dass die Sensorikeinheiten bezüglich der Vorrichtung externe Gegebenheiten
15 messen oder analog dazu interne Gegebenheiten. Interne Gegebenheiten sind alle Systemparameter der Vorrichtung an sich. Externe Gegebenheiten sind Umgebungsvariablen, also von Gegenständen, die die Vorrichtung umgeben. Somit werden also mittels der internen Sensorikeinheiten Vorrichtungseigenschaften gemessen und mittels externen Sensorikeinheiten Umgebungseigenschaften. Weist beispielsweise die Vorrichtung
20 einen bestimmten Batteriezustand auf, so handelt es sich hierbei um einen internen Parameter. Wird mittels eines Roboterarms ein Gegenstand von A nach B transferiert, so handelt es sich hierbei um einen externen Zustand, wobei der Zustand des Roboterarms wiederum ein interner Zustand ist. Typischerweise sind also interne und externe Parameter in Wechselwirkung und Aktionen bewirken beispielsweise immer einen verringerten
25 Batteriezustand, während diese Aktionen wiederum auf externe Eigenschaften einwirken. In anderer Richtung ist es eine Wechselwirkung, dass mittels Effektoren die Vorrichtung an eine Ladesäule herangeführt werden kann und sodann wird ein Ladevorgang eingeleitet, welcher wiederum den internen Batteriezustand beeinflusst.

Die erfassten Wechselbeziehungen werden sodann abgespeichert und somit wird
30 protokolliert, wie das Betätigen der Effektoren die internen und externen Parameter beeinflusst. Die Vorrichtungseigenschaften werden also mitsamt den Umgebungseigenschaften abgespeichert und somit werden Handlungsanweisungen definiert. Das Betätigen der Effektoren ist also eine Handlung die sowohl die Vorrichtung an sich als auch die Umgebung beeinflusst. Führt beispielsweise die Vorrichtung von einem

ersten geographischen Punkt zu einem zweiten geographischen Punkt so ändert dies die Umgebung der Vorrichtung, was ein externer Parameter ist und dies ändert auch den internen Zustand der Vorrichtung, beispielsweise durch eine Temperaturentwicklung oder einer Veränderung des Batteriezustands. Somit kann also ein Tripel abgespeichert werden, 5 welches angibt, wie der interne Zustand war, wie der externe Zustand war und welche Handlung sodann ausgeführt wird. Dies wird überführt in einen neuen internen Zustand und einen neuen externen Zustand. Somit besteht also ein Ausgangstripel vom Betätigen der Effektoren, Vorrichtungseigenschaften und Umgebungseigenschaften und dies wird in ein Zieltripel überführt. In dem Zieltripel können der neue interne Zustand, der neue externe 10 Zustand sowie mögliche Handlungen spezifiziert werden.

Das Konzept der Tripel soll lediglich den allgemeinen Übergang der Systemzustände verdeutlichen. Generell ist es alternativ auch möglich, dass interne Zustände, externe Zustände mittels einer Funktion in einen neuen internen Zustand und einen neuen externen Zustand überführt werden. Somit besteht also die Handlung in einer Abbildungsfunktion von 15 einem ersten Tupel in einen zweiten Tupel. Der Fachmann erkennt hierbei, dass die Mächtigkeit beider Beschreibungen gleichwertig ist. Entweder es wird also ein Ausgangstupel in einen Zieltupel mittels einer Funktion überführt oder aber es wird ein Eingangstripel gebildet, was interne Zustände, externe Zustände und eine Handlung spezifiziert, wobei diese Handlung sodann zu einem Zieltripel führt, welches einen neuen 20 internen Zustand, einen neuen externen Zustand und ein weiteres Feld aufweist. Das weitere Feld kann entweder leer bleiben, wobei es jedoch bevorzugt ist, dass hier mindestens eine Handlung ausgeführt wird, die nunmehr in diesem Zustand möglich ist. Das letzte Feld kann auch derart belegt werden, dass hier beispielsweise ein Vektor eingeführt wird, der Kennungen weiterer Handlungen beinhaltet.

25 Um dies an einem Beispiel zu verdeutlichen, kann die Vorrichtung den Effektor namens Motorantrieb betätigen und vorwärtsfahren. Was nunmehr mitprotokolliert wird, ist ein Ausgangstripel eines internen Zustands, nämlich einem Ladestand der Batterie, einem externen Zustand, nämlich ein Bildsignal einer physischen Umwelt und einer Handlungsanweisung, nämlich „Fahren“. Dieser Ausgangsvektor bzw. das Ausgangstripel 30 wird nunmehr überführt in ein Zieltripel, welches einen neuen Batterieladezustand aufweist, ein neues Bildsignal der Umgebung sowie Handlungen, die nunmehr möglich wären, beispielsweise Vorwärtsfahren oder Rückwärtsfahren. Während der Handlung der Vorwärtsbewegung ist es auch möglich zu spezifizieren, dass hierbei beispielsweise ein Bremsen oder ein Betätigen der Scheinwerfer möglich wäre.

In einem alternativen Beispiel wird der Ausgangstupel des Batteriezustands und des Bildsignals der Umgebung mittels der Funktion „Bewegen“ übergeführt in ein Zieltupel, welches den neuen Batteriezustand beschreibt sowie ein neues Bildsignal der Umgebung. Somit erfolgt also ein mitprotokollieren, was bei einer bestimmten Handlung intern und
5 extern erfolgt. Dieses Wissen kann in weiteren Handlungsschritten in Richtung eines vordefinierten Ziels wiederverwendet werden.

Im weiteren Verfahrensschritten kann ein Bereitstellen von Zielvorrichtungseigenschaften oder vordefinierten Zielumgebungseigenschaften erfolgen. Somit kann die Vorrichtung derart angesteuert werden, dass sich die mindestens eine vordefinierte Zielvorrichtungseigenschaft
10 und/ oder die mindestens eine vordefinierten Zielumgebungseigenschaft einstellt. Aus den vorhergehenden Verfahrensschritten sind die Handlungsanweisungen bekannt, und somit ist also bekannt, wie ein interner Zustand und ein externer Zustand in einen weiteren internen Zustand und einen weiteren externen Zustand überführt werden kann. Hierzu dienen die Zielvorgaben, also die Zielvorrichtungseigenschaft und/ oder die Zielumgebungseigenschaft.
15 Diese Zieleigenschaften bestimmen also was erreicht werden soll und dies erfolgt dadurch, dass die Handlungsanweisungen ausgeführt werden. Es wird also ausgelesen, in welchem Zustand sich die Vorrichtung befindet und sodann werden abgespeicherte Handlungen ausgewählt, welche ausgehend von dem Ist- Zustand den Soll- Zustand erreichen. In einem bevorzugten Fall reicht hierzu das Ausführen einer Handlungsanweisung, welche sofort die
20 Zielvorgaben erreicht. In einem typischen Fall wird es jedoch so sein, dass mittels mehrerer Handlungseigenschaften sukzessive der Ausgangs-Zustand in den Soll-Zustand überführt wird. Es werden also mehrere Handlungsanweisungen derart verknüpft, dass sich letztendlich die Zielvorgaben einstellen.

Um dies in einem Beispiel zu verdeutlichen, kann das Verfahren identifiziert haben, dass die
25 Vorrichtung sich von A nach B nach C nach D bewegen kann. Somit werden also die Handlungsanweisungen abgespeichert, die vorsehen, dass das Auto von A nach B fahren kann, von B nach C sowie von C nach D. Es wird implizit auch abgespeichert, dass das Auto von A nach C fahren kann und von A nach D fahren kann. Darüber hinaus wird abgespeichert, dass das Auto von C nach D fahren kann. Diese Möglichkeiten werden
30 nunmehr verwendet und wenn eine Zielvorgabe vorliegt, die besagt, dass das Auto von A nach C fahren soll, hat die Vorrichtung zwei Optionen zur Auswahl, nämlich von A nach B nach C zu fahren oder von A nach C zu fahren. Generell wäre es auch möglich von A nach D zu fahren und sodann von D nach C zu fahren. Welche Auswahl erfolgt, kann ebenfalls in den Zielvorgaben definiert werden. Beispielsweise kann eine Maximaldauer der Fahrt in

Zeitangaben definiert werden. Darüber hinaus kann ein maximaler Energieverbrauch definiert werden. In vorbereitenden Verfahrensschritten werden also randomisiert Handlungen ausgeführt und diese liegen dann zur Laufzeit vor. Generell kann das Verfahren ein randomisiertes Betätigen der Effektoren vorsehen es kann aber auch additiv vorsehen, dass die bereits gelernten Handlungsanweisungen wiederverwendet werden. Somit kann eine Lernphase des randomisierten Betätigens durchgeführt werden, sowie eine Ausführungsphase, welche auf bereits gespeicherte Handlungen zurückgreift. Dies kann auch in beliebiger Reihenfolge alterniert werden. Darüber hinaus wird auch in der Ausführungsphase mittels der internen Sensorik und der externen Sensorik das abgespeicherte Handlungswissen verfeinert und es werden somit neue Handlungsanweisungen generiert, die bezüglich einer anderen Zielvorgabe bevorzugt sein können. Darüber hinaus können sich auch während dem Ausführen einer Handlung die Zielvorgaben derart ändern, dass beispielsweise ein Batteriezustand kritisch wird und insofern wird die Zielvorgabe der eigenen Operabilität derart erhöht, dass das Verfahren erkennt, dass es nunmehr notwendig ist, eine Ladestation anzufahren. Somit ergibt sich zumindest temporär eine überlagernde Zielvorgabe, nämlich die eine Ladestation anzufahren, auch wenn dadurch nicht das Ziel erreicht wird die ursprüngliche Zielkoordinate anzufahren. Sobald ein gewisser Ladezustand erreicht ist, wird die Zielvorgabe der eigenen Operabilität wieder zurückgestuft und die Zielvorgabe des Anfahrens des Zieles wird derart erhöht, dass die Fahrt nunmehr fortgesetzt wird.

Gemäß einem Aspekt der vorliegenden Erfindung wird das Verfahren lernend iteriert, derart, dass stets neue Handlungsanweisungen erkannt werden, welche jeweils ein Ausgangstripel in ein Zieltripel überführen und diese Handlungsanweisungen zum Ansteuern der Vorrichtung gespeichert werden. Dies hat den Vorteil, dass das Verfahren ganz oder zumindest teilweise derart etabliert wird, dass stets neue Situationen erkannt werden und Handlungsanweisungen derart geschaffen werden, dass Effektoren betätigt werden. So kann das Verfahren derart variiert werden, dass das Betätigen der Effektoren nicht mehr randomisiert erfolgt, sondern vielmehr können bestehende Handlungsanweisungen kombiniert werden oder es ist auch möglich, dass Teile vorhandener Handlungsanweisungen randomisiert werden, sodass neue mögliche Handlungsanweisungen geschaffen werden bezüglich derer auch klar ist, welche Wechselwirkungen sie auslösen. Darüber hinaus kann das Verfahren auch derart etabliert werden, dass ausgehend von dem Abspeichern von Wechselbeziehungen bis zum Ansteuern der Vorrichtung etabliert wird. Somit werden auch bei einem Ausführen von bekannten Handlungsanweisungen Wechselbeziehungen abgespeichert und es können

neue Parameter identifiziert werden. So kann beispielsweise erkannt werden, dass bei einem zweimaligen Ausführen einer gleichen Handlungsanweisung andere Zielparameter entstehen. Dies kann zum Beispiel dann der Fall sein, wenn bei einer Autofahrt ein Seitenwind aufkommt. Dies wird abgespeichert und sodann wird mittels des externen
5 Sensors erkannt, dass hier ein Wind aufgetreten sein muss. Somit wird ein neues Ausgangstripel geschaffen und es kann mittels der Effektoren randomisiert ausprobiert werden, wie in einer solchen Situation gegengelenkt werden soll. Somit wurden also neue Wechselbeziehungen erkannt und wird nochmals ein derartiges Ausgangstripel identifiziert, ist nunmehr klar, welche Handlung ausgeführt werden muss, um das gewünschte Zieltripel
10 zu erreichen.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung erfolgt das Auslesen von internen und/ oder externen Effektoren derart, dass einzelne Stützwerte abgespeichert werden und Zwischenwerte interpoliert und/ oder extrapoliert werden. Dies hat den Vorteil, dass die Stützwerte entsprechend ihrer Verfügbarkeit gewählt werden können und dass
15 diese auch derart geschätzt werden können, dass nicht an jedem möglichen Wert eine Messung vorliegen muss. Vielmehr kann anhand von vorhandenen Messungen geschätzt werden, welche Werte sich nach normalen Umständen ergeben würden. Somit können weitere Stützwerte mathematisch errechnet werden.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird die
20 Zielvorrichtungseigenschaft und/ oder die Zielumgebungseigenschaft mindestens temporär von einer Vorrichtungseigenschaft und/ oder einer Umgebungseigenschaft beeinflusst. Dies hat den Vorteil, dass die Zielvorgaben auch geändert werden können bzw. einer Priorisierung unterliegen. Soll beispielsweise von einem ersten geographischen Punkt an einen zweiten geographischen Punkt gefahren werden, so kann zwischenzeitlich das
25 Anfahren einer Ladestation priorisiert werden, da ansonsten das Endziel nicht erreicht werden würde. Somit ergibt sich also der Vorteil, dass stets garantiert wird, dass die Zielvorgaben erreicht werden und es wird ein ausfallsicheres Verfahren geschaffen.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung werden mehrere Handlungsanweisungen zu einer Choreografie kombiniert, welche ein Ausgangstripel in ein
30 Zieltripel überführt. Dies hat den Vorteil, dass auch komplexe, das heißt zusammengesetzte Handlungen ausgeführt werden können und mit Fortschreiten des Verfahrens beziehungsweise dem Etablieren des Verfahrens werden stets neue Handlungsanweisungen geschaffen, die zu immer effizienteren Choreografien

zusammengefügt werden. Somit wird also das vorgeschlagene Verfahren iterativ verbessert.

Die Vorrichtung sei gemäß einem Aspekt der vorliegenden Erfindung ein Fertigungsroboter mit zwei Roboterarmen mit jeweils einem Effektor (Greifhand), die im Sinne der Vorrichtung ausgebildet sind, d. h. sie „wissen“, wie sie Gegenstände ergreifen, bewegen,

5 manipulieren etc. kann.

Der Vorrichtung wird nun nur einmal demonstriert, wie die neuen Werkstücke Werkstück-1 und Werkstück-2 ergriffen, gewendet und dann zusammengesteckt werden sollen. Das Ergebnis wird dann der Vorrichtung als Ziel markiert. Jetzt kann die Vorrichtung das ab sofort jederzeit wiederholen: Einmal zeigen, statt umfangreich trainieren.

- 10 Erscheinen jetzt die zwei Werkstücke auf einem Band, werden jeweils zwei Alter-Egos erstellt, und aufgrund der Situation, dem Ausgangstripel, und dem damit verbundenen Ziel, fällt der Motivationswert ins Negative, wodurch die Vorrichtung eine Handlungsaufforderung erkennt und beide zusammenfügt, was zu einem Motivationswert führt, der die Vorrichtung in einen Stillstand oder Ruheposition führt. Dies ist ein Beispiel, wie die Möglichkeit der
- 15 Vorrichtung, aus der Beobachtung zu lernen, eingesetzt werden kann.

- Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird das Verfahren auf einer ersten Vorrichtung zur Ausführung gebracht und anhand von externen Sensorikeinheiten der ersten Vorrichtung wird eine zweite Vorrichtung erkannt, welche das Verfahren ebenfalls ausführt und deren erzeugte Wechselwirkungen auf der ersten Vorrichtung abgespeichert
- 20 werden oder die Wechselwirkungen werden entfernt abgespeichert und mittels einer Schnittstelle mindestens für die erste Vorrichtung und/ oder die zweite Vorrichtung verfügbar gemacht. Dies hat den Vorteil, dass das Verfahren auf mehreren Instanzen zur Ausführung gebracht wird und ein Beobachten der Fähigkeiten der zweiten Vorrichtung durch die erste Vorrichtung ausgeführt wird. Es wird also das Verhalten der zweiten Vorrichtung analysiert
- 25 und auf der ersten Vorrichtung oder aber auch zentral, für die erste Vorrichtung zugänglich, abgespeichert. Die Ergebnisse des Durchführens des Verfahrens auf der zweiten Vorrichtung werden also von der ersten Vorrichtung ausgelesen und/ oder ausgemessen und als potentielle eigene Handlungen auf der ersten Vorrichtung oder beschrieben zentral abgespeichert.

- 30 Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird die erste Vorrichtung anhand der erfassten Wechselwirkungen der zweiten Vorrichtung angesteuert. Dies hat den Vorteil, dass ein Ansteuern der ersten Vorrichtung gemäß mindestens einer vordefinierten

Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens der abgespeicherten Handlungsanweisung der zweiten Vorrichtung erfolgen kann.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung werden Wechselbeziehungen
5 der erfassten zweiten Vorrichtung zwischen dessen Parameter und dessen Handlungen zur Erstellung von Handlungsanweisungen der ersten Vorrichtung verwendet. Dies hat den Vorteil, dass ein Verbund von Vorrichtungen geschaffen wird, die untereinander mittels randomisierten Betätigens von Effektoren und Auslesen einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen Vorrichtungseigenschaften und
10 einem Auslesen von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften Handlungsanweisungen bzw. Wechselwirkungen erstellen bzw. abspeichern und diese sodann untereinander zur jeweiligen oder gemeinsamen Nutzung teilen. Auch können die Überführungen der Ausgangstripel in Zieltripel kollektiv durchgeführt werden. So sind entsprechende Choreographien nicht auf
15 eine Vorrichtung, welche das Verfahren ausführt, beschränkt, sondern vielmehr können die Choreographien arbeitsteilig durch mehrere Vorrichtungen durchgeführt werden.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird anhand von externer Sensorik ein Objekt erfasst und dessen erfasste Parameter mit Vorrichtungseigenschaften und/ oder Handlungsanweisungen der Vorrichtungen durch die Abbildung auf eine Kopie
20 des zentralen Templates, seien es von der Vorrichtung selbst trainiert oder von einer vergleichbaren Vorrichtung übernommen und implementiert, verglichen. Dies hat den Vorteil, dass sich die Vorrichtung quasi selbst klonen kann beziehungsweise von den eigenen Eigenschaften auf die Eigenschaften des Objekts schließen kann. Dies kann auch als Empathie bezeichnet werden. Wird mittels der externen Sensorik beispielsweise eine
25 Vorrichtung identifiziert, die ähnliche Charakteristika aufweist, wie die ausführende Vorrichtung, so wird davon ausgegangen, dass diese neue erkannte Vorrichtung ähnliche Fähigkeiten aufweist wie die ausführende Vorrichtung. Um dies in einem Beispiel zu verdeutlichen, wird auf ein Fahrzeug abgestellt, welches die vorgeschlagene Erfindung ausführt. Dieses erste Fahrzeug führt also das vorgeschlagene Verfahren aus und erkennt
30 mittels eines externen Sensors, hier einer bildgebenden Einheit, eine weitere, also zweite Vorrichtung. Die erste Vorrichtung bzw. das erste Fahrzeug hat gelernt, dass es bei einer bestimmten Geschwindigkeit nur bedingt bremsen kann und nur bedingt quer lenken kann. Das erste Fahrzeug erkennt nunmehr, dass das zweite Fahrzeug eine ähnliche Abmessung aufweist und eine ähnliche Geschwindigkeit fährt, wie das erste Fahrzeug. Sodann werden

die Handlungsanweisungen in das zweite Fahrzeug konzeptionell projiziert und das erste Fahrzeug erkennt, dass das zweite Fahrzeug nur bedingt bremsen kann und nur sehr bedingt einen Richtungswechsel durchführen kann. Wird nunmehr auf einer Autobahn einen Überholvorgang eingeleitet, so erkennt das erste Fahrzeug, dass das zweite Fahrzeug
5 bremsen könnte oder aber auch ebenfalls einen Spurwechsel durchführen könnte. Somit wird also von dem eigenen Verhalten beziehungsweise der möglichen Ausgangstriplets und der möglichen Zieltriplets gefolgert, dass dieses Objekt ähnliche Eigenschaften aufweist. Darüber hinaus kann mittels der externen Sensorik das Verhalten des erkannten zweiten Fahrzeugs überwacht werden und sodann kann der eigene Datenspeicher aktualisiert
10 werden, der Wechselbeziehungen abgespeichert. Somit können also auch Ausgangstriplets und Zieltriplets erkannt werden und es kann wiederum von dem zweiten Fahrzeug auf das erste Fahrzeug geschlossen werden.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung werden die erfassten Parameter zur Prädiktion eines Verhaltens des erfassten Objekts verwendet. Dies hat den Vorteil, dass
15 erkannt wird, dass die eigene Vorrichtung in einer bestimmten Situation eine bestimmte Handlungsanweisung ausführt, sodass das weitere erkannte Objekt in derselben Situation sehr wahrscheinlich auch die gleiche Handlung ausführen würde. Es werden also die eigenen Handlungen bzw. ausgeführten Handlungsanweisungen mitprotokolliert und sodann wird davon ausgegangen, dass sich das erfasste Objekt gleich oder zumindest ähnlich
20 verhalten könnte.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung werden Wechselbeziehungen des erfassten Objekts zwischen dessen Parameter und dessen Handlungen zur Erstellung von Handlungsanweisungen der Vorrichtung verwendet. Dies hat den Vorteil, dass auch externe erkannte Wechselbeziehungen bezüglich der vorgeschlagenen Vorrichtung bzw. der
25 Vorrichtung, welche mittels des Verfahrens angesteuert wird, verwendet werden können. Somit schafft also die Vorrichtung nicht alleine Wechselbeziehungen, welche ausgewertet werden, sondern vielmehr kann auch die externe Welt beobachtet werden und sodann Rückschlüsse auf das Überführen von Ausgangstripel in Zieltripel erkannt werden.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung liegt ein Effektor als ein
30 Prozessor, ein Speicher, ein Motor, ein Greifarm, eine Fortbewegungseinheit, ein Antrieb, eine Extremität, ein Scheibenwischer, ein Blinker, eine Leuchte, eine Lenkung und/ oder eine ausführende Einheit vor. Dies hat den Vorteil, dass anhand der vorgeschlagenen Vorrichtung bzw. der verfahrensgemäß anzusteuernenden Vorrichtung alle möglichen

Hardwareeinheiten vorliegen können. Die Ausführung ist nur beispielhaft und nicht abschließend, sodass jegliche physische Einheit gemäß dem vorgeschlagenen Verfahren autonom angesteuert werden kann.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung erfassen interne

- 5 Sensorikeinheiten einen Effektorzustand, eine Effektoreinstellung, eine Effektorposition, eine Größe, ein Gewicht, eine Abmessung, eine Geschwindigkeit, eine Beschleunigung, ein Abbremsen, einen Batteriestand, eine Reichweite, eine Position, eine Prozessorauslastung, eine Speicherauslastung und/ oder einen Systemparameter der Vorrichtung. Dies hat den Vorteil, dass die internen Sensorikeinheiten alle Parameter und Zustände der Vorrichtung
- 10 ausmessen können, die das Verfahren ausführt bzw. die durch das Verfahren ausgeführt bzw. angesteuert wird.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung erfassen externe

- Sensorikeinheiten einen Abstand, ein Objekt, ein Lidarsignal, ein Bildsignal, eine Position, eine Größe, eine Abmessung, ein Akustiksignal und/ oder einen externen Parameter. Dies
- 15 hat den Vorteil, dass die komplette Umwelt der Vorrichtung analysiert und erfasst werden kann. Hierzu sind alle möglichen Sensoren einzeln oder in Kombination möglich, die die Umwelt in irgendeiner Art und Weise erkennen.

Da gemäß einem Aspekt der vorliegenden Erfindung trainierte neuronale Netze

- vorgeschlagen werden, ist es vorteilhaft, dass das erfindungsgemäße zentrale Template die
- 20 Funktionen trainierter neuronaler Netze übernimmt, aber in seinen Möglichkeiten und Flexibilität die der NN weit überschreitet.

Das zentrale Template kann entweder selbst trainiert werden oder von einer vergleichbaren Vorrichtung kopiert werden. Im Falle einer künstlichen Intelligenz eines Smart Home kann

- 25 ein zentrales Template niemals selber sich bewegen lernen, um z.B. die Handlungen einer älteren Person im Rahmen einer gesundheitlichen Überwachung (Stichworte Empathie, Sturzkontrolle etc.) zu überwachen, es muss im Stand der Technik von anderer Stelle kopiert und implementiert werden.

- 30 Das erste Training oder Ausbildung des zentralen Templates ist dem Spielen von Kindern vergleichbar und ist eben nicht bestimmt durch das Erreichen extern festgelegter Ziele. Es dient nur dem Erlernen des eigenen Handels: Was kann ich alles machen. Später können dann durch die Kombinationen der so gelernten Aktionsmöglichkeiten die externen Ziele

erreicht werden, wobei der Weg dazu selbst entworfen wird oder werden muss. Andernfalls würde das Handeln von Anfang an auf bestimmte Ziele fixiert werden (Tür öffnen), um dann zu versagen, wenn andere Ziele (Staub saugen) zu erreichen wären.

- 5 Er erfolgt gemäß einem Aspekt der vorliegenden Erfindung das Ausprobieren der eigenen Möglichkeiten wie Gas geben, lenken u.a. – ähnlich dem „babbling“. Aber die Vorrichtung erkennt und speichert zugleich auch (und das geht über das „babbling“ hinaus) die Änderungen (Abstände, Geschwindigkeiten, Beschleunigung, ..), die dadurch bewirkt werden, sodass sie beginnt nicht mehr das eigene Handeln (Lenken) sich selbst als Zieltripel
- 10 zu setzen, sondern die Änderungen, die sie so erzielen kann und das geht über das „babbling“ hinaus. So wie ein Kind erst das Krabbeln lernt (lernen muss) und dann beginnt gezielt hierhin und dorthin zu krabbeln. So werden für Kinder wie für die Vorrichtung die Zieltripel immer „abstrakter“ im Vergleich zu den eigenen, „ursprünglichen“ Handlungsmöglichkeiten, die so zu reinen Mitteln für immer komplexere Zieltripel werden.
- 15 NN hingegen sind und bleiben beschränkt auf die antrainierten Ausgangs- und Zieltripel.

- Die Vorrichtung gemäß einem Aspekt der vorliegenden Erfindung setzt sich in der juvenilen, der Ausbildungsphase, so wie kleine Kinder selbst Ziele bzw. Zieltripel wie: ich will dorthin, ich will da anstoßen. Dieses zufällige eigenständige Zielsetzen beeinflusst den internen
- 20 Motivationswert negativ, anders ausgedrückt, er wird „ausgelenkt“. Jetzt versucht sie durch eigene Handlungen diesen Wert wieder gegenteilig, also positiv, zu verändern. Ein Kind würde das wahrscheinlich spüren als (positive) Aufregung oder Anspannung (also nicht als ein negatives Gefühl wie Angst z.B.). Technisch gesprochen kann und wird so jedes Zieltripel auf einen internen Wert abgebildet, über den die Vorrichtung ihr Verhalten selber
- 25 steuern kann. Bei NN muss der Mensch in der Trainingsphase konkret das Erreichen angeben, ohne dass der Backpropagation-Prozess nicht funktionieren würde

- In der juvenilen Ausbildungsphase versucht jetzt die Vorrichtung gemäß einem Aspekt der vorliegenden Erfindung durch Betätigen seiner Handlungsmöglichkeiten wie „Gas geben
- 30 (vorwärts bewegen) , lenken, bremsen“ zum Ziel zu gelangen. Über die Sensorik merkt sie, dass z.B. „Gas geben“ aber ein Lenken in die falsche Richtung den Abstandswert nicht so beeinflusst, dass sich der Motivationswert positiv verändert. Jede Handlung führt also sofort zu einer Änderung des Motivationswertes, aus der die Vorrichtung erkennt, ob sie dem Ziel

näher kommt oder nicht, sodass sie die Handlung fortsetzt (bei einer positiven Änderung) oder ändert (bei keiner oder einer negativen Änderung). Über die Matrix der Stützwerte und der Interpolation kann die Vorrichtung auch erkennen, ob eine kleine Änderung der Handlung (etwas mehr oder etwas weniger Gase geben) die Motivationswertes positiv oder negativ verändert, sodass sie wiederum selbstständig die Handlungen für das Erreichen des Ziels selbstständig optimiert. In diesem Sinne kann sie die schnelleren oder kürzeren Weges erkennen und wählen. Wenn z.B. das Ziel schräg rechts vor der Vorrichtung (erkennbar über die Sensorik und dem internen Polarkoordinatensystem) liegt, das sie erreichen möchte, kann sie erkennen, dass ein Lenken nach links (weg vom Ziel) die Änderung des Motivationswertes negativ verändert, ein Lenken nach rechts hingegen positiviert.

Mit den gespeicherten Stützwerten kann die Vorrichtung im Rahmen einer Simulationsrechnung vor der Entscheidung für eine bestimmte Handlung (lenken nach links oder rechts) die jeweilige Wirkung auf den Motivationswert berechnen und dann gleich die optimale Handlung (nach rechts lenken im obigen Szenario) selbst wählen, ohne dass das, wie das bei einem NN notwendig wäre, vorher antrainiert werden müsste. Das geht natürlich erst, wenn eine ausreichend große Datenbasis vorhanden ist.

Es können gemäß einem Aspekt der vorliegenden Erfindung auch z. B. einschränkende Nebenbedingungen berücksichtigt werden. Beispiel: Eine Straße führt nicht direkt zum Ziel, das hinter einem frisch gepflügten Feld liegt, sondern nur im Winkel von 45°. Jetzt führt die Simulation des geraden Weges (größte Motivationsänderung) über das Feld mit der Gefahr des Steckenbleibens (massive negative Motivationswertänderung) oder weiter die Straße entlang (kleine aber doch positive Motivationsänderung)

25

So lernt die Vorrichtung bzw. das Verfahren gemäß einem Aspekt der vorliegenden Erfindung oder ist danach fähig:

1. Handlungen verändern den Motivationswert, wobei genuin angelegt ist, diesen Wert auf einen maximal positiven Wert zu bringen (ob das in einer Realisation 0, ein größt- oder kleinstmöglicher Wert ist, ist für das Prinzip nebensächlich).
2. Manche Handlungen beeinflussen die Veränderung des Motivationswertes stärker, andere weniger stark, sodass die Handlung(en) mit dem stärksten Einfluss technisch erkennbar und wählbar sind.

3. Aus den gespeicherten Stützwerten und einer neuen Situation (Ausgangstripel) kann die Vorrichtung im Sinne einer Simulationsrechnung die für das Erreichen des Ziel jeweils optimalen Handlungen auswählen bzw. „sich dafür entscheiden“, ohne speziell auf das Ausgangs- und das Zieltripel trainiert worden zu sein.
- 5 4. Durch die Fähigkeit zur Simulation kann nicht nur eine direkte Handlung, die das Ausgangs- direkt in das Zieltripel überführt, sondern es ermöglicht, das Erreichen des Zieltripels in mehrere Einzelschritte zu unterteilen, um z. B. bestimmte Nebenbedingungen nicht zu verletzen.
- 10 5. Wenn die Vorrichtung die Konsequenzen ihre Handlungen durch eine technische Simulation selbst berechnen kann (wann werde ich wo sein und welche Geschwindigkeit haben), können damit auch die Konsequenzen der Handlungen der anderen, neuen, unbekannten Objekte über die Simulation technisch berechnet werden. Die Vorrichtung kopiert die eigene Objekt-Instanz (das Ego) mit allen Daten, Methoden und Funktionen und aktualisiert die jeweilige Kopie (das Alter-Ego) mit den
15 aktuellen Werten der Sensorik an.

Eine interne Handlungsanweisung ist z. B. das eigenständige Betätigen der Lenkung durch die Vorrichtung, um die Vorrichtung z. B. auf das aktuelle Ziel auszurichten – sie ergibt sich aus der Simulation und der sich daraus ergebenden optimalen Trajektorie für das Erreichen
20 des Zieltripels.

Eine externe Handlungsanweisung ist gemäß einem Aspekt der vorliegenden Erfindung das externe Setzen eines (neuen) Zieltripels.

Ein Ausgangstripel ist gemäß einem Aspekt der vorliegenden Erfindung eine aktuelle
25 Situation bestehend aus der Vorrichtung, den weiteren dynamischen Objekten, den Nebenbedingungen des Szenarios ..., die die Vorrichtung „veranlasst“ durch die Simulation, also quasi dem „theoretischen“ Ausprobieren verschiedener Entwicklungen einen optimalen Weg (Trajektorie) zum Zieltripel zu finden.

30 Vorrichtungseigenschaften sind gemäß einem Aspekt der vorliegenden Erfindung dass die eigenen physikalischen Eigenschaften (Masse, Größe, „PS“,...), die Fähigkeit selbst aktiv zu werden (bremsen, lenken, beschleunigen, ...), die Fähigkeit die Umgebung zu ändern (anstoßen, schubsen, schieben,...)

35 Umgebungseigenschaften sind gemäß einem Aspekt der vorliegenden Erfindung, siehe

auch Ausgangstripel: eine aktuelle Situation bestehend aus der Vorrichtung, den weiteren dynamischen Objekten, den Nebenbedingungen des Szenarios ...

Das Zieltripel kann gemäß einem Aspekt der vorliegenden Erfindung wie folgt beschrieben werden: in der juvenilen Ausbildungsphase wären das zunächst reine zufällige direkte Handlungen, deren Auswirkungen (Abstände, Zusammenstöße, Beschleunigung, ...) gespeichert werden, dergestalt, dass diese dann die Auswirkungen zu direkten positiven (zu erreichen) und negativen (Zusammenstöße : unbedingt zu vermeiden) Zieltripel werden. Sobald eine ausreichend große Datenbasis existiert, kann die Vorrichtung sich Ziele setzen, die nicht mehr direkt erkennbar sind, sondern z. B. erst nach dem Umfahren eines Hindernisses erreicht werden kann: wieder etwas, was Kinder nach einer gewissen Zeit erlernen: Das Überwinden des „Aus den Augen, aus dem Sinn“

Es erfolgt also gemäß einem Aspekt der vorliegenden Erfindung kein Versagen bei Unbekannten. Das ist eine technische Lösung eines technischen Problems. Das Ansteuern der Vorrichtung gemäß mindestens einer vordefinierten Zielvorrichtungseigenschaft und mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer abgespeicherten Handlungsanweisung erfolgt gemäß einem Aspekt der vorliegenden Erfindung wie folgt: Die mindestens eine Handlungsanweisung gibt an, wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird

Die mindestens eine Handlungsanweisung überführt das Ausgangstripel in das Zieltripel. Es gibt eine Ausgangskonfiguration und diese wird anhand von einer oder mehrerer Handlungsanweisungen, im letzteren Fall also eine Kette von Handlungsanweisungen die einzelne Ausgangstripel in Zieltripel überführen, in eine Zielkonfiguration überführt.

Nunmehr hat das Verfahren bzw. die Vorrichtung noch nicht gelernt, wie die vordefinierte Zielvorrichtungseigenschaft und die mindestens eine vordefinierte Zielumgebungseigenschaft anhand der gelernten Überführungen bzw. Handlungsanweisungen erreicht werden kann. Die Vorrichtung muss somit aus dem maschinell erlerntem Erfahrungsschatz und den Entwicklungsprognosen einer dynamischen Situation diejenige(n) Handlungsanweisungen auswählen, welche die Zielkonfiguration auf dem besten Weg (optimale Trajektorie) erreicht. Nähert sich die Vorrichtung der Zielkonfiguration an, so kann diese Handlungsanweisung verwendet werden; der Motivationswert verändert sich positiv. Entfernt sich die Vorrichtung stattdessen von der

Zielkonfiguration, so kann diese Handlungsanweisung von ihr nicht verwendet werden; der Motivationswert verändert sich negativ. Dies kann aufgrund der abgespeicherten Handlungsanweisungen evaluiert bzw. simuliert werden. Somit ergibt sich die Möglichkeit erfindungsgemäß selbstständig diejenige(n) Handlungsanweisung(en) auszuwählen, welche
5 am besten funktionieren, also den besten Weg (optimale Trajektorie) zur Zielkonfiguration erwarten lassen.

So kann zum Beispiel bei einem Annähern der vorgeschlagenen Vorrichtung unter Verwendung des vorgeschlagenen Verfahrens sensorisch ein Abstand ermittelt werden. Gelernt wurde das ein Ausgangstripel in ein Zieltripel überführt. Das Ausgangstripel im
10 vorliegenden Beispiel ist Betätigen (Fortbewegungsmittel einschalten), Vorrichtungseigenschaften (die Vorrichtung kann fahren, lenken, ...) und Umgebungseigenschaften (die Vorrichtung befindet sich bei geographischem Punkt A). Das Zieltripel ist Betätigen (die Vorrichtung dreht um), Vorrichtungseigenschaften (die Vorrichtung kann aufgrund eines Hindernisses nicht fahren) und Umgebungseigenschaften
15 (die Vorrichtung befindet sich bei geographischem Punkt B, jedoch vor einem Hindernis). Nunmehr merkt die Vorrichtung, dass der Motivationswert sich negativ verändert hat, da ein Hindernis angefahren wurde. Es wird diejenige mindestens eine Handlungsanweisung verwendet, welche einen Motivationswert maximal verbessert und der Motivationswert angibt wie nahe das Überführen von Ausgangstripel in das Zieltripel die vordefinierten
20 Zielvorrichtungseigenschaft und der mindestens einen vordefinierten Zielumgebungseigenschaft erreicht.

Da diese Tripel abgespeichert sind, kann diese Bewegung also simuliert werden und es wird die Handlungsanweisung verwendet die die Zielkonfiguration am besten erreicht. Somit erfolgt ein geplantes und vorrausschauendes Auswählen/ Verwenden der
25 Handlungsanweisung(en).

Der Motivationswert kann auch technische Konfigurationen der Vorrichtung berücksichtigen. So kann auch ein hoher Energieverbrauch den Motivationswert verschlechtern. Erfolgt eine Beschädigung der Vorrichtung kann dies ebenfalls den Motivationswert negativ verändern. Weitere Beispiele sind ein zu hoher Zeitbedarf, es wird also der kürzeste Zeitbedarf
30 verwendet, oder es entstehen unerwünschte Zielkonfigurationen wie die Beschädigung von externen Objekten. Es kann empirisch eine Metrik bereitgestellt werden, welche die Einflussfaktoren auf den Motivationswert, korreliert und empirisch ein Optimum erstellt wird, welches sodann gemäß der Handlungsanweisung umgesetzt wird. Dementsprechend werden die Effektoren angesteuert.

- Die Aufgabe wird auch gelöst durch eine Vorrichtung eingerichtet zum Ausführen eines Verfahrens nach einem der vorhergehenden Ansprüche, aufweisend eine Initialisierungseinheit eingerichtet zum Initialisieren mittels randomisiertem Betätigen von Effektoren und Auslesen einer Vielzahl von internen Sensorikeinheiten eingerichtet zur
- 5 Messung von internen Vorrichtungseigenschaften und einem Auslesen von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften; eine Speichereinheit eingerichtet zum Abspeichern von Wechselbeziehungen, welche zwischen dem Betätigen der Effektoren und den ausgelesenen Vorrichtungseigenschaften und Umgebungseigenschaften bestehen als
- 10 Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von Betätigen, Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt; und eine Steuerungseinheit eingerichtet zum Ansteuern der Vorrichtung gemäß mindestens einer vordefinierten Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer
- 15 abgespeicherten Handlungsanweisung, welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und/ oder die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird.

Die Aufgabe wird auch gelöst durch eine Anordnung aufweisend mehrere Vorrichtungen, welche kommunikationstechnisch gekoppelt sind und Handlungsanweisungen austauschen.

- 20 Die Aufgabe wird auch gelöst durch ein Computerprogrammprodukt mit Steuerbefehlen, welche das vorgeschlagene Verfahren implementieren bzw. die vorgeschlagene Vorrichtung betreiben.

- Erfindungsgemäß ist es besonders vorteilhaft, dass das Verfahren zum Betreiben der vorgeschlagenen Vorrichtungen und Einheiten verwendet werden kann. Ferner eignen sich
- 25 die vorgeschlagenen Vorrichtungen und Einrichtungen zur Ausführung des erfindungsgemäßen Verfahrens. Somit implementiert jeweils die Vorrichtung strukturelle Merkmale, welche geeignet sind, das entsprechende Verfahren auszuführen. Die strukturellen Merkmale können jedoch auch als Verfahrensschritte ausgestaltet werden. Auch hält das vorgeschlagene Verfahren Schritte zu Umsetzung der Funktion der
- 30 strukturellen Merkmale bereit. Darüber hinaus können physische Komponenten gleicherweise auch virtuell bzw. virtualisiert bereitgestellt werden.

Weitere Vorteile, Merkmale und Einzelheiten der Erfindung ergeben sich aus der nachfolgenden Beschreibung, in der unter Bezugnahme auf die Zeichnungen Aspekte der

Erfindung im Einzelnen beschrieben sind. Dabei können die in den Ansprüchen und in der Beschreibung erwähnten Merkmale jeweils einzeln für sich oder in beliebiger Kombination erfindungswesentlich sein. Ebenso können die vorstehend genannten und die hier weiter ausgeführten Merkmale je für sich oder zu mehreren in beliebigen Kombinationen

- 5 Verwendung finden. Funktionsähnliche oder identische Bauteile oder Komponenten sind teilweise mit gleichen Bezugszeichen versehen. Die in der Beschreibung der Ausführungsbeispiele verwendeten Begriffe „links“, „rechts“, „oben“ und „unten“ beziehen sich auf die Zeichnungen in einer Ausrichtung mit normal lesbarer Figurenbezeichnung bzw. normal lesbaren Bezugszeichen. Die gezeigten und beschriebenen Ausführungsformen sind
- 10 nicht als abschließend zu verstehen, sondern haben beispielhaften Charakter zur Erläuterung der Erfindung. Die detaillierte Beschreibung dient der Information des Fachmanns, daher werden bei der Beschreibung bekannte Schaltungen, Strukturen und Verfahren nicht im Detail gezeigt oder erläutert, um das Verständnis der vorliegenden Beschreibung nicht zu erschweren. Die Figuren zeigen:

- 15 **Figur 1:** ein schematisches Ablaufdiagramm eines Verfahrens zum autonomen Ansteuern einer Vorrichtung gemäß einem Aspekt der vorliegenden Erfindung;

- Figur 2:** ein Diagramm, anhand dessen interne und/oder externe Eigenschaften geschätzt werden können. So können interne und/ oder externe Parameter ausgemessen werden und zwischen den Stützstellen kann interpoliert werden bzw. es können
- 20 weitere Stützstellen extrapoliert werden;

- Figur 3:** eine Darstellung einer realweltlichen Situation, bei der ein Kind eine Straße überqueren will und die vorgeschlagene Vorrichtung die Situation erkennt und daraus belehrend;

- Figur 4:** eine realweltliche Situation in einem Straßenverkehr, in der die vorgeschlagene Vorrichtung bzw. das vorgeschlagene Verfahren gemäß einem Aspekt der
- 25 vorliegenden Erfindung Anwendung findet; und

- Figur 5:** eine weitere realweltliche Situation in einem Straßenverkehr, wobei die Vorrichtung eine Kurvenfahrt gemäß einem weiteren Aspekt der vorliegenden Erfindung einübt.

- 30 **Figur 6A:** ein schematisches Ablaufdiagramm eines Verfahrens zum autonomen Ansteuern einer Vorrichtung gemäß einem Aspekt der vorliegenden Erfindung in der

juvenilen Phase mit dem vorrangigen Ziel der Präzisierung der eigenen Handlungen; und

Figur 6B: ein schematisches Ablaufdiagramm eines Verfahrens zum autonomen Ansteuern einer Vorrichtung gemäß einem Aspekt der vorliegenden Erfindung in der adulten, 5 der Anwendungsphase für das Erreichen der eigenen Ziele unter Berücksichtigung und Vorhersage der Verhaltensweisen der in einer Situation relevanten weiteren Objekte.

Figur 1 zeigt in einem schematischen Ablaufdiagramm ein Verfahren zum autonomen Ansteuern einer Vorrichtung, aufweisend ein Initialisieren 100 mittels randomisiertem 10 Betätigen 101 von Effektoren und Auslesen 102 einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen Vorrichtungseigenschaften und einem Auslesen 103 von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften; ein Abspeichern 104 von Wechselbeziehungen, welche zwischen dem Betätigen 101 der Effektoren und den 15 ausgelesenen Vorrichtungseigenschaften und Umgebungseigenschaften bestehen als Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von Betätigen 101, Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt; und ein Ansteuern 105 der Vorrichtung gemäß mindestens einer vordefinierten Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten 20 Zielumgebungseigenschaft unter Verwendung von mindestens einer abgespeicherten 104 Handlungsanweisung, welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und/ oder die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird.

Das hier vorgestellte Modell verwendet gemäß einem Aspekt der vorliegenden Erfindung ein 25 zentrales Template, mit der sich die KI in der unbekannten und sich immer wieder verändernden Welt zurechtfindet. Funktional ersetzt es die mit vorab ausgewählten und interpretierten Daten trainierten NNs, wodurch es eben nicht vorab auf bestimmte externe Ziele oder auf vorab definierte externe Objekte festgelegt ist, welches zu einer viel größeren Freiheit des Erkennens und des Handelns führt. Dieser Ansatz erlaubt es, kurzfristige 30 Verhaltensprognosen der beobachteten Objekte zu erstellen, um so die Entwicklung der Situationen abzuschätzen, um das beim Entwurf der eigenen Handlungen einzubeziehen. Er ermöglicht auch das Lernen aus Beobachtung und Empathie. Aspekte der Erfindung sind:

- Zentrales Template als Zentrum der KI zur Erfassung aller Objekte einer Situation;

- Datenspeicherung, die ein Erkennen von Gesetzen ermöglicht;
- Verhaltensprognosen aller beobachteten Objekte einer Situation;
- Empathiefähigkeit, die auf dem Zentralen Template basiert;
- Fähigkeit des Lernens aus Beobachtungen;
- 5 - Erkennen von nicht direkt beobachtbaren Einflussfaktoren;
- Erkennen von Naturgesetzen und kausalen Abhängigkeiten;
- computergerechte Form von Theorien, die selbst erstellt, gespeichert, wiederverwendet und laufend verbessert werden kann; und/ oder
- Selbstbewertung und die Bewertung der selbsterstellten Theorien als Grundlage der
- 10 Erkenntnis, des Lernens und der Auswahl der jeweils besten Theorie für das Erreichen der eigenen Ziele.

Als Beispiel wird ein Auto, das sich nach diesen Ideen verhält, beschrieben. Wie beurteilt es ein Kind oder eine Kiste, wie einen Überholvorgang auf der Autobahn, oder wie erkennt es das Gravitationsgesetz und Seitenwinde (als Beispiel für das Erkennen von Kausalitäten und

15 nicht manifestierbaren Objekten). Weiter wird beschrieben, wie über Empathie das Lernen aus Beobachtung möglich ist. All das zeigt, dass die Möglichkeiten dieser KI über das Fahren von Autos hinausgeht. Die wesentliche Grundlage ist aber ein physikalisch existierender Roboter in der physikalisch existierenden Umgebung, das andere Objekte direkt oder indirekt beobachtet.

20 Im Folgenden werden einige Aspekte der vorliegenden Erfindung vorgeschlagen, welche eine beispielhafte Umsetzung des Verfahrens bzw. der Vorrichtung und/ oder der Systemanordnung ermöglichen. Die folgenden Aspekte sind lediglich beispielhaft zu verstehen und können einzeln oder in Kombination angewendet werden.

Mögliche Hardware des Roboters beziehungsweise der vorgeschlagenen Vorrichtung:

25 Hier wird die technische Ausstattung beschrieben, über die der Roboter oder das Auto gemäß einem Aspekt der vorliegenden Erfindung verfügen kann.

Effektoren:

In einfachen Fällen wie zum Beispiel einem Auto, wären die Effektoren, sieht man einmal von Scheibenwischer, Blinker, Licht, ab. Bremsen, Beschleunigen und Lenken, mit

30 diesen der Roboter auf seine Umgebung einwirkt und sei es nur die eigene Ortsveränderung ist auch eine Situationsveränderung.

Interne Sensorik:

- Die interne Sensorik zur Messung und Rückmeldung eigener Eigenschaften wie Größe, Gewicht, Geschwindigkeit, Beschleunigung etc. aufgrund der Aktionen der Effektoren. Sie sind notwendig für die Beurteilung der eigenen Maßnahmen und als Voraussetzungen für das Lernen. Damit kann es die eigene Aktion zum Beispiel Stärke des Bremsens und Wirkung (das Ausmaß der negativen Beschleunigung) erkennen, Verbindungen und um Gedächtnis speichern, um so selbstständig zu lernen, und das heißt, die eigenen Handlungen stetig verbessern.
- Als eine physikalische Entität in einer physikalischen Umgebung unterliegt der Roboter den physikalischen Gesetzen. Eine gemessene Beschleunigung, also eine Geschwindigkeitsveränderung, ist neben der Masse unter anderem davon abhängig, wie weit das Gaspedal, oder wie stark die Bremse gedrückt wird. Wenn diese Gesetzmäßigkeiten nicht fest einprogrammiert werden sollen, sondern der Roboter dieses selber entdecken, speichern und verwenden soll, benötigt es die interne Sensorik, um die Konsequenzen aus dem Einsatz der Effektoren zu erfahren und speichern zu können. Die interne Sensorik, zum Beispiel eines Autos, erfasst daher Daten wie Größe, Geschwindigkeit, Beschleunigung, etc.

Externe Sensorik:

- Die externe Sensorik erfasst gemäß einem Aspekt der vorliegenden Erfindung die Umgebung (Abstände, Objekte etc.) zum Beispiel durch Lidar, Bildbearbeitung etc. Natürlich muss der Roboter in geeigneter Weise Kenntnis über die Umgebung erhalten. Was genau benötigt wird, wird weiter unten näher beschrieben. Welche Systeme aber sich am besten eignen, soll nicht vorab festgelegt werden, da die Informationserfordernisse aufgrund der Verwendung des zentralen Templates geringer sind. Es muss nicht alles, was theoretisch an Informationen zur Verfügung steht, auch erfasst und verarbeitet werden.

Mögliche Software des Roboters:

Hier werden die Teile beschrieben, mit denen die vorgestellte KI gemäß einem Aspekt der vorliegenden Erfindung ihre Intelligenz erreicht.

- Datenspeicherung:

Die Datenspeicherung ist ein hochverfügbarer Datenspeicher für Datentupel, die sich zusammensetzen aus den Daten der internen Sensorik, der Effektoren, die Handlungen und deren Motivationswerte (s.u.). Der Speicher verbindet in geeigneter Weise ausgeführte Handlungen und die damit verbundenen, tatsächlichen, erfahrenen und beobachteten

5 Ergebnisse. Er speichert also nur interne Daten in Tabellen und zum Beispiel keine Pixel eines oder mehrerer Bilder. Es gibt also nur Daten, deren interne Bedeutung bekannt ist.

Ein weiterer Teil dieses Moduls ist gemäß einem Aspekt der vorliegenden Erfindung die wiederholte Prüfung auf eine innere Stimmigkeit der Daten. Damit ist gemeint, sie dürfen keine Widersprüche enthalten, Strukturen, die zu Schleifen oder Zirkel führen etc. Ziel ist,

10 dass aus den Daten sich immer nur eine Entscheidung ergeben darf. Die Routine, die das überprüft, soll auch gespeicherte redundante Daten wieder entfernen, nicht nur um so die Datenbasis so schlank wie möglich zu halten, sondern auch um eine Überadaptierung zu verhindern. Was als redundant angesehen wird, ergibt sich aus der Art der Theoriebildung.

Figur 2 zeigt die Theoriebildung gemäß einem Aspekt der vorliegenden Erfindung:

15 Das zentrale Template kann mit den hier beschriebenen einfachen Mitteln jede funktionelle Wechselbeziehung nachbilden, abspeichern und für die Prognose verwenden.

Das Bremsen oder Beschleunigen hängen funktional von der beschleunigten Masse und der aufgewendeten Kraft ab. Allerdings ist diese Funktion erst einmal unbekannt, denn die mathematische Funktion der Geschwindigkeitsveränderung dV ist abhängig von den

20 Voraussetzungen V_n (aktuelle Geschwindigkeit, Gewicht etc.) und der Verwendung der Effektoren E_m (Kraft auf das Gaspedal, Lenkung,...), also:

$$dV = f(V_n, E_m) \quad (1)$$

Diese Funktion soll selbst erlernt und nicht einprogrammiert werden, denn so kann das Erlernte später auf die gleiche Weise erweitert, korrigiert und verbessert werden. Die

25 effektive Beschleunigung beim Bremsen (oder Gas geben) ist über die interne Sensorik der gemessenen Beschleunigung a mit der (mathematisch) unabhängigen Variablen der Kraft F (wie stark das jew. Pedal gedrückt wird) und der Masse m nach dem Gesetz $F = m \cdot a$ verbunden. Das Prinzip, mit dem der Roboter selbst diese Gleichung findet, ist einfach. Nehmen wir zur Veranschaulichung eine (andere, beliebige) quadratische Funktion $y = -((x - 4)^2) + 9$, die nachgebildet werden soll.

30

- Es sei, dass zunächst nur die zwei Messwerte an den Punkten 11 und 12 existieren. Nun sei es, dass an der Stelle $x=3$ (Punkt 2) eine Prognose p erforderlich ist, was über die Gerade a zwischen 11 und 12 zu einem Wert von $y=3,8$ führt. Da aber der Fehler zum ex post gemessene Wert $y=8,0$ viel zu groß ist, wird dieser Punkt 2R bei $x=3,0$ und $y=8,0$ in die
- 5 Erinnerung übernommen. Daraufhin würde bei einer neuen Prognose entweder die Gerade b_1 zwischen [11,2R] oder b_2 zwischen [2R, 12] herangezogen werden. Mit der Zeit entstehen in dem obigen Beispiel viele weitere Punkte (31,32,33,...), über die der funktionale Zusammenhang von x und y durch die jeweiligen Geraden entlang der Datentupel ($c_1, c_2, c_3, \dots$) mit beliebiger Genauigkeit nachgebildet werden kann. Die Genauigkeit ist nur
- 10 beschränkt durch (noch) nicht gemachte Erfahrung und der Größe des Erinnerungsspeichers. Natürlich lässt sich das auf, so es die Rechen- und Speicherkapazitäten zulassen, auf beliebig viele Variablen erweitern, statt Geraden werden halt Hyperebenen $a_1x_1 + \dots + a_nx_n = \text{const}$ zur Prognose mittels der stückweisen Linearen Inter- oder Extrapolation herangezogen.
- 15 Um aber die Anzahl der gespeicherten Eck- oder Datentupel so klein wie möglich zu halten, können diese auch wieder entfernt werden: Wenn sich bei einer ex post Analyse herausstellt, dass ein Datentupel so nahe oder sogar auf der Geraden/Hyperebene zwischen den zwei benachbarten Tupeln liegt, dann kann dieses Tupel als redundant gelöscht werden. Dadurch verringert sich nicht nur der Speicherbedarf, sondern auch der
- 20 Rechenaufwand und das Verhalten passt sich im Laufe der Zeit immer besser an die tatsächliche, wiewohl nach wie vor unbekannte mathematische Funktion an, was dann zu einem im Laufe der Zeit immer effizienteren Verhalten führt.

- Dieses Vorgehen verhindert auch eine Überadaptierung, die man bei Auswahl und Umfang der Trainingsdaten für ein KNN unbedingt berücksichtigen muss. Redundante Erfahrungen
- 25 werden so nicht immer wieder eingespeichert, was die selten auftretenden Ereignisse, die sogenannten ‚schwarzen Schwäne‘ mit unter Umständen heftigen Konsequenzen, vollkommen überlagern würden.

- Das Speichern von funktionalen Zusammenhängen von Variablen hat noch einen weiteren Vorteil, da auf diese Weise zugleich auch jede mögliche Umkehrfunktion gespeichert ist, da
- 30 die gespeicherten Datentupel nicht unterscheiden, welche Werte bei ihrem Auftreten die ‚abhängigen‘ und welche die ‚unabhängigen‘ waren. In der obigen Funktion ist x die unabhängige Variable und der zugehörige y -Wert kann über die Funktion $y = -((x-4)^2)+9$ ermittelt werden. In der Datenbank müsste nach dem x -Wert gesucht werden und, wenn der

x-Wert nicht explizit gespeichert ist, kann näherungsweise der y-Wert über die beiden nächsten beiden x-Werten ermittelt werden. Nun kann man aber in dem Speicher auch einfach nach dem y-Wert suchen und über die stückweise Lineare Inter- oder Extrapolation einen x-Wert ermitteln – das ist sehr viel einfacher als zu versuchen, die Umkehrfunktion
5 zum Beispiel der obigen Gleichung zu ermitteln.

Auf diese sehr einfache Weise kann jeder funktioneller Zusammenhang beliebig genau nachgebildet werden, Voraussetzung ist nur ein ausreichendes Handlungstraining, wie bei einer komplexen Sportart. Natürlich kann auch der akzeptierte Fehler ϵ im Laufe der Zeit geändert und angepasst werden, sei es, dass er verkleinert werden muss, damit sich die
10 benötigte Präzision erhöht, oder sei es, dass er vergrößert werden kann, um den Speicher- und der Rechenaufwand zu verringern, da das ja wieder die Anzahl der gespeicherten Eck- und Datenpunkte beeinflusst, es ist eine stetige Optimierung im laufenden Prozess.

Im Rahmen dieser Erfindung definiert sich gemäß einem Aspekt der vorliegenden Erfindung eine Theorie über die Regeln und Gesetze einer aktuellen Situation daher als ein in sich
15 geschlossener, konsistenter Datensatz, aus dem sich mittels der stückweisen Linearen Inter- oder Extrapolation die Prognosen dieser Theorie von dem Roboter errechnen lassen. Der Datensatz besteht aus eigenen Messergebnissen und eventuell auch aus den Parametern erstellter Alter Egos. So kann der Roboter jede seiner Theorien selbst ermitteln, abspeichern, wiederverwenden und verbessern. Im Rahmen der Selbstbewertung (s.u.) ist
20 er auch in der Lage, in einer Situation die spezifisch beste Theorie (Datensatz), die mit der höchsten Kompetenz (s.u.) auszuwählen und anzuwenden.

Handlungen:

Eine Handlung stellt gemäß einem Aspekt der vorliegenden Erfindung eine geordnete Abfolge von Effektoreneinsätze dar, um ein Ziel zu erreichen. Mithilfe der dem Roboter
25 bekannten funktionellen Zusammenhänge zwischen Effektoren, den internen Parametern und der bekannten Auswirkungen können Handlungen entworfen werden. Zunächst, in der juvenilen Lernphase (s.u.), werden einfache Handlungen gesetzt. Die KI lernt, Gas geben und bremsen, dann Gas geben, warten und an einer Begrenzung anstoßen (extern gebremst mit einer ‚Verletzungsgefahr‘), Gas geben, Lenken, Bremsen und so weiter. Diese
30 einfachen Handlungen werden dann zu immer komplexeren ‚Choreographien‘ zusammengesetzt und diese werden dann weiter optimiert, in dem zum Beispiel das Bremsen durch einen schwächer werdenden Pedaldruck bei abnehmender Geschwindigkeit weicher wird und zum Beispiel das Kurvenfahren flüssiger gestaltet wird. Erreicht wird die

Optimierung durch einen Motivationswert (s.u.), der eine Art interne Bewertung der ausgeführten Handlung darstellt. So wird von den gespeicherten (Teil-) ‚Choreographien‘ diejenige mit dem höchsten Motivationswert gewählt. Auf diesem Weg kann die KI selbstständig den Einsatz seiner Effektoren laufend verbessern. Außerdem können über den

5 Bezug zurzeit und die gemachten Handlungserfahrungen die eigene Situation in zukünftigen Zeitpunkten prognostiziert werden, welches die Grundlage für die eigene Handlungsplanung ist.

Motivationswerte:

Motivationswerte spiegeln eine Art Belohnung wider. Einerseits werden sie intern ermittelt,

10 nachdem ein selbstgestecktes Ziel erreicht wurde (gelingen), andererseits kann ein Motivationswert auch extern übermittelt werden (gut gemacht). Wenn zum Beispiel eine Fahrstrecke mit starken Lenkbewegungen, geringer Geschwindigkeit und trotzdem hohen Seitbeschleunigungen zurückgelegt wird, erhält diese Handlungsabfolge einen geringeren Motivationswert als das Resultat des ‚Trainings‘, nach dem dieselbe Strecke schneller und

15 trotzdem mit geringeren Seitbeschleunigungen durchquert wird. Das muss natürlich auch im Gedächtnis gespeichert werden.

Die Motivationswerte stellen somit eine Form von nicht festgelegten, intrinsischen Zielen dar. Es können Fernziele wie der Endpunkt einer Reise verbunden werden oder kurzfristige Zwischenziele wie der Besuch einer Tankstelle, wenn der Tank leer ist: Mit abnehmender

20 Tankfüllung steigt langsam der Motivationswert für ‚Tanken‘, bis er größer ist als der des Fernziels. Das Verfolgen des Fernziels wird unterbrochen zugunsten des Tankens und danach wieder aufgenommen.

Zielsystem:

Der Roboter ist gemäß einem Aspekt der vorliegenden Erfindung immer ‚angeschaltet‘,

25 obwohl es natürlich über einen Ausschalter verfügen kann. Es führt daher immer eine Tätigkeit aus. Allerdings gehört auch das Nichtstun (aufladen) oder in einem Ruhezustand den eigenen Langzeitspeicher der Handlungsoptionen optimieren (schlafen) dazu. Als Handlung wird immer diejenige gewählt, die aktuell den höchsten Motivationswert hat. Erreicht eine aktuell nicht ausgeführte Handlung einen höheren Motivationswert als die

30 aktuell ausgeführte, wird diese abgebrochen, beendet oder unterbrochen, damit die mit der höchsten Motivation ausgeführt werden kann. Wenn die Motivation eines derartigen Autos auf dem Weg von München nach Hamburg zu tanken oder die Akkus aufzuladen größer

wird als das Erreichen von Hamburg, wird es bei nächster Gelegenheit die Fahrt unterbrechen und eine Tankstelle ansteuern.

In der juvenilen Phase (s.u.), also in der Lernphase, eines derartigen Roboters bzw. Autos werden die Motivationen das Fahren, das Bremsen, das Beschleunigen und das

- 5 Kurvenfahren auf kurzen Strecken hohe Motivationswerte erhalten, ganz einfach, weil der interne Langzeitspeicher signalisiert, es soll gelernt bzw. der Handlungsfehler reduziert werden, um das eigene Handeln zu präzisieren. Später, in der adulten Phase (s.u.), wenn solches Verhalten keine oder nur noch unwesentliche Veränderungen im Datenspeicher verursacht, die Fehlerreduktion nähert sich der Null, wird ein solches Verhalten ‚langweilig‘,
10 es erhält nur geringen Motivationswerte und dann wird zum Beispiel das Energiesparen bevorzugt.

Damit vermeidet dieses Zielsystem gemäß einem Aspekt der vorliegenden Erfindung eine Festlegung auf die Wirklichkeit, welches ja eine Interpretation derselben bedeuten würde, die möglicherweise in der Zukunft sich als falsch erweisen könnte.

- 15 Das zentrale Template, das Ego:

Das zentrale Template stellt den Kern und die Basis dieser KI gemäß einem Aspekt der vorliegenden Erfindung dar. Es wird Ego genannt, weil es sich selbst in das Zentrum seiner Datenverarbeitung stellt und erst einmal alles aus der eigenen Sicht betrachtet und bewertet. Daher müssen dem Ego keine Daten mit deren Bedeutung (zum Beispiel Identifikation von
20 Bildern) vorgegeben oder antrainiert werden, es erschafft sich alles selbst. Erfindungsgemäß kann das zentrale Template oder das Ego in der Vorrichtung durch das selbstständige Erlernen der eigenen Fähigkeiten selbst entstanden sein oder von einer vergleichbaren Vorrichtung kopiert und implementiert worden sein.

Subjektives Polarkoordinatensystem:

- 25 Alle Daten der Beobachtungen und der Erfahrungen dieser KI sind subjektiv bzw. werden gemäß einem Aspekt der vorliegenden Erfindung aus der Perspektive des Egos gemacht. Dieser Idee folgt auch das Raumkonzept. Es wird kein kartesisches Koordinatensystem mit einem irgendwo angenommenen Ursprung der beobachteten Umgebung entworfen, in dem jedem erkannten Objekt die jeweiligen Koordinaten zugewiesen werden. Stattdessen kann
30 ein Polarkoordinatensystem verwendet. Dadurch wird nicht nur der mathematische Ursprung automatisch im Ego festgelegt, sondern auch die subjektive Orientierung des Egos mit

- vorne-hinten, oben-unten, rechts-links einschließlich des Bewegungsvektors. Der umgebende (leere) Raum wird als gegeben vorausgesetzt. Die beobachteten Objekte werden nur über den Abstand zum eigenen Standpunkt, also dem Ursprung des Polarkoordinatensystems, und dem Winkel erfasst. Der leere Raum im physikalischen Sinn
- 5 wird also nicht als eine eigene Entität oder Dimension angenommen. Er ist da und ist einfach nur eine Option sich dorthin zu bewegen, es sei denn, dort ist, steht oder bewegt sich ein anderes Objekt (hin), dann wäre der ‚Platz‘ besetzt bzw. der Raum nicht leer. Ein weiterer Vorteil der subjektivistischen Polarkoordinaten ergibt sich im Folgenden für das Konzept der Alter Egos (s.u.).
- 10 Juvenile Phase:
- Innerhalb unbedingter Vorgaben, wie die Unversehrtheit der beobachteten Objekte und des Roboters selbst, der Erhaltung der Einsatzbereitschaft (Ladezustand der Batterien), Zielsetzungen wie sanftes Beschleunigen und Bremsen und optimierte Kurvenfahrten und anderen, soll/kann der Roboter den Einsatz seiner Effektoren und die damit hervorgerufenen
- 15 Konsequenzen selbst erlernen vor allem, um seine Effektoren mit seiner internen und externen Sensorik zu koordinieren und durch das steige Training 100-105, zeitweise unterbrochen von Ruhephasen, um zum Beispiel die Batterien aufzuladen, und die interne Datenbasis zu pflegen (Redundanz, Konsistenz,...), zu präzisieren, bis der erkannte Fehler aus Planung und Ergebnis ausreichend klein geworden ist. Bei einem Auto wären das vor
- 20 allem der Einsatz von Gas geben, Lenken, Bremsen, einen Zielpunkt erreichen (Erfolg) oder verfehlen (Fehler) etc. Zuerst setzt das Ego sich selbst einfache Ziele für einfache Handlungen, die ausgeführt und mit Ergebnis abgespeichert werden, die später dann zu immer komplizierteren ‚Choreografien‘ zusammengesetzt werden. Diese Ausbildungsphase kommt zum Ende, wenn die Handlungen der selbst- oder fremdgesetzten Ziele („fahr mal
- 25 dahin“) nicht weiter optimiert werden können, also, wenn das mathematische Modul den Fehler nicht oder nur marginal noch verringern kann und/oder die Datenbasis kaum noch optimiert (Anzahl und Lage der Punkte in der Erinnerung) werden kann. Wenn sich die Fehlerreduktion immer langsamer verringert, neigt sich die juvenile Phase immer mehr ihrem Ende zu. Ein KNN hingegen benötigt gezielt ausgewählte und interpretierte Trainingsdaten
- 30 in hohem Umfang, das Ego hingegen nur eine reale Umgebung, man könnte sagen ein Spielzimmer, wo es keine Schäden anrichten könnte, um die eigenen Fähigkeiten auszuprobieren und auszubilden. Es folgt intrinsischen Zielen und nicht extern vorgegebenen wie ein KNN und damit keiner Interpretation der Umgebung oder Welt.

Adulte Phase:

In der juvenilen Phase hat die Vorrichtung die geordnete Abfolge seiner Effektoreneinsätze für das Erreichen der gesetzten Ziele erlernt. Welchen Wert haben die in der juvenilen Phase gewonnenen Erkenntnisse und Erfahrungen? Alle Daten über die Welt, die das Ego in der Welt, in der es sich bewegt, kennengelernt hat, sind eigene Daten, eigene Beobachtungen und eigene Erfahrungen. In einem übergeordneten wissenschaftstheoretischen oder philosophischen Sinne muss man feststellen, diese Daten sind aus der Sichtweise des Egos wahr, im Sinne eines Das-Ist-So oder Das-War-So. Damit stellt dies eine natürliche und feste Ausgangsbasis der Erkenntnis dar. Natürlich muss das Ego seine internen Daten pflegen. Neben einer oben bereits angedeuteten Optimierung auch dahingehend, dass sie widerspruchs- und fehlerfrei sind. Andernfalls könnte man als Hersteller oder Betreiber des Roboters nicht davon ausgehen, dass der Roboter mit seinen entworfenen Handlungen auch die angestrebten Ziele erreichen wird.

Den selbst erfahrenen und selbst verursachte Veränderungen in der Umgebung, wie zum Beispiel ein Ortswechsel durch Gas geben, lenken und bremsen, wird durch das Verbinden der internen und externen Daten erlebt und gespeichert. Es gibt, und es braucht dazu keine weiter hinter die eigenen Effektoren gehende Ursachenkenntnis im Sinne von Fragen wie, warum führt das getretene Gaspedal zu einer Beschleunigung, das Lenken zu einer Kurvenfahrt, das Bremsen zu einer negativen Beschleunigung. Dem Ego eines Autos reicht ein Es-gibt-die-drei-Optionen, um gezielt den Ort zu wechseln. Auf dieser Ebene ist das So-Ist-Es das Wichtige, nicht das Warum. Eine Taube auf der Straße hat sicher keine Ahnung, wie ein Auto oder ein Radfahrer funktioniert. Wenn aber so ein ‚Verkehrsteilnehmer‘ weit genug an ihr vorbeifahren würde, bleibt sie sitzen, wird aber die Bewegungsrichtung in ihre Richtung geändert, fliegt sie davon. Sie weiß zwar nicht, warum, aber sie weiß um die Verhaltensmöglichkeiten und passt daher genau auf.

Alter Ego:

Wenn das (adulte) Ego über die externe Sensorik in der beobachtbaren Umgebung ein Objekt erkennt, erstellt das Ego einfach eine Kopie seiner selbst und passt die Parameter der Kopie gemäß den beobachteten Eigenschaften an. Das Ego ist somit das Zentrale Template für alle beobachtbaren und nicht beobachtbaren Objekte und Phänomene. Daher auch die Namen Ego und, als Konsequenz, Alter Ego für die Kopien.

Durch die Verwendung des Egos als zentrales Template gibt es keine unbekannten Objekte

mehr. Es gibt nur unbekannte Parameter, die aber über die externe Sensorik gemessen, aus der eigenen Erfahrung abgeschätzt (Stichwort Vorurteil) und deren mögliche Bandbreite durch weitere Beobachtungen eingeschränkt werden können.

So wie das Ego sein Verhalten mit seinen Methoden, gespeicherten Daten und Parametern berechnet und die eigene Situation prognostizieren kann, kann auch das Verhalten der Alter Egos mit denselben Methoden, Daten und den adaptierten Parametern berechnet werden. Die damit den Alter Egos unterstellte Annahmen beschränken sich auf dieselben physikalischen Gesetze und darauf, dass die kurzfristigen Ziele aus der beobachteten Orientierung und Bewegungsrichtung ableitbar sind. Sei das Beobachtete nun eine Kiste auf der Straße, ein Känguru in Australien oder eine alte Dame mit Rollator.

Programmtechnisch sollte das Ego als Objekt vorliegen. Dann kann davon für ein neu aufgetretenes Objekt ganz einfach eine Kopie des Egos erstellt werden, um danach die Parameter der Kopie des neuen Objektes mit Hilfe der externen Sensorik und der eigenen Erfahrung anzupassen. Es könnte eine Liste mit den Alter Egos der augenblicklichen Situation geben, in die die neue Kopie eingehängt oder erstellt wird:

```
Class CEgo { ... };

CEgo Ego( ... );

// Beginn von Ausbildung & Lernen, juvenile Phase

...

20 // Beginn der adulten Phase

CEgo AlterEgos[];

...

// neues Objekt erkannt

AlterEgos[i] = Ego.clone(); // Klonen des Egos für das Objekt

25 AlterEgos[i].adjust(...); // Parameteranpassung des neuen Objektes

...
```

Damit hat das Ego intern ein Alter Ego als berechenbares Abbild eines beobachteten Objektes in der Umgebung erstellt und es kann, wie in Figur 6B gezeigt, zur kontinuierlichen Fortschreibung der beobachteten Situation die innere Schleife bilden aus Objekterkennung 311, Aktualisierung der beobachtbaren Objektparameter 312, Prognose des Verhaltens der
 5 Objekte 313, Anpassen der Abfolge der Effektoreneinsätze 314 und eine zweite, äußere Schleife nach dem Erreichen des Zieles und dem Eintritt in eine Ruhe- und (auch) Ladephase 315, bis es mit der Initialisierung 310 für ein neues Ziel weitergeht.

Es spielt dabei keine Rolle, ob das beobachtete Objekt bekannt oder ganz neu ist. Wenn es neu ist, ist die Bandbreite der möglichen Merkmalsausprägungen größer, was eine erhöhte
 10 Aufmerksamkeit (Abtastfrequenz) bedingt. Dies hat folgende Vorteile:

- Jedes beobachtete Objekt ist vom Grunde her bekannt!
- Verhaltensprognose aller externen Objekte mit eigenen Methoden und Daten!
- Der Mensch muss nicht mehr eingreifen, um Fehlendes einzuprogrammieren.
- Das Ego lernt und verbessert laufend sein Verhalten.
- 15 - Das Ego ist empathiefähig und ist sich der aktuellen Situation bewusst (s.u.).
- Das Ego kann aus der Beobachtung der Alter Egos lernen (s.u.).
- Alter Egos können unbekannte Kausalitäten und Regeln verkörpern (s.u.).

Figur 3 zeigt Szenario 1 „ein Kind“ gemäß einem Aspekt der vorliegenden Erfindung.

Wenn das Ego, zum Beispiel die KI eines einfachen Autos, einem anderen
 20 Verkehrsteilnehmer, im Beispiel hier ein Kind, begegnet, erstellt es eine Kopie seiner selbst und passt die Parameter (Entfernung d zum Ego, Orientierung, Geschwindigkeit v_K , ...) an:

So kann das Alter Ego mit den Methoden des Egos prognostizieren, wo es sich wann befinden wird. Und so kann das Ego über den Vergleich mit der eigenen Berechnung, wann es sich wo befinden wird, feststellen, ob es zu einer möglicherweise gefährlichen Situation
 25 kommen könnte, um entsprechen zu reagieren, also zu bremsen. So kann mit den eigenen Methoden und Erfahrungen das kurzfristige Verhalten des Kindes prognostiziert werden. NICHTS muss extra für das Kind programmiert oder trainiert werden, das dann nicht vorläge, wenn diese KI einem Kind das erste Mal begegnet.

Figur 4 zeigt Szenario 2 „Autobahn“ gemäß einem Aspekt der vorliegenden Erfindung.

Ein anderes Beispiel, das deutlicher zeigt, wie mächtig diese so einfache Idee der Alter Egos ist, ist eine Situation auf einer Autobahn mit einem LKW und zwei PKW dahinter und einem PKW auf der Überholspur, von denen der zweite PKWEgo unser Ego ist. Das Ego kann

5 zunächst einmal aufgrund der eigenen Kenntnis vom Kraftgesetz ($F=m \cdot a$) von Bremsstärke, Masse mit den Parametern der beobachteten PKWs und des LKWs, also mit vermuteter Masse, gemessener Geschwindigkeit und vermuteter Bremswirkung allein mit den eigenen Methoden die Bremswege aller Verkehrsteilnehmer in einer möglichen Unfallsituation prognostizieren, um so den eigenen Sicherheitsabstand zu bestimmen. Behandeln wir aber
10 zunächst die ‚Gedanken‘ bzw. die Unterstellungen des Egos über die anderen Verkehrsteilnehmer für die Prognose ihrer Handlungen:

- Der LKW0 fährt eine bestimmte Geschwindigkeit, über den das Ego (PKWEgo) gelernt hat (ohne speziell darauf trainiert werden zu müssen), dass derart große Verkehrsteilnehmer sie nur ganz selten überschreiten. Das Alter Ego des LKWs wird
15 also keine Geschwindigkeitsveränderung prognostizieren.
- Der PKWEgo, das Zentrum der Betrachtung, wäre in der Lage schneller zu fahren und würde den LKW0 unter Berücksichtigung der anderen Verkehrsteilnehmer überholen. Das Überholen würde den PKWEgo schneller an sein Ziel bringen und erhält damit einen im Moment höheren Motivationswert, er bereitet den
20 Überholvorgang vor.
- Der PKW1, ein weiteres Alter Ego, denkt systembedingt dasselbe wie der PKWEgo, weil das Ego an seiner Stelle so handeln würde. Das Ego unterstellt also dem PKW1, LKW0 auch überholen zu wollen. Natürlich ‚sieht‘ das Alter Ego des PKW1 PKWEgo und wird seinen (vom Ego) unterstellten Überholvorgang unter Berücksichtigung der Existenz des PKWEgo durchführen, also zum Beispiel Blinker setzen und unter besonderer Aufmerksamkeit auf den PKWEgo und möglicher weiterer Verkehrsteilnehmer auf der Überholspur zu lenken. Das wäre das, was das Ego des
25 PKWEgo ‚denkt‘ und wie es das Verhalten des PKW1 prognostizieren würde.
- Ein möglicher weiterer Verkehrsteilnehmer, der PKW2, nähert sich auf der Überholspur mit hoher Geschwindigkeit. Auch dafür erstellt das Ego ein Alter Ego, entwirft dessen Handlungsprognose aufgrund der freien Fahrbahn.

35

- Wie oben dargestellt, kann das Programm, das Ego, die gesamte Situation mit allen relevanten Verkehrsteilnehmern prognostizieren und sie eigene Abfolge der Effektoreneinsätze entsprechend planen, sodass es zu keiner gefährlichen Entwicklung kommt.

5 Alter Ego, Bewusstsein:

Somit hat das Ego intern ein vollständiges Bild der beobachtbaren Umgebung mit allen erkannten Objekten und die Fähigkeit, kurzfristig deren Verhalten zu prognostizieren. Das wäre eine mögliche und programmtechnische umsetzbare Definition von Bewusstsein, einem echten, kein vorgetäushtes oder imitiertes.

- 10 Die kurzfristige Verhaltensprognose aufgrund der Fähigkeit, sich an die Stelle der beobachteten Objekte zu versetzen, heißt hier, die Situation mit allen für die Alter Egos umgerechneten Daten betrachten, und über das berechnete wahrscheinliche Verhalten der beobachteten Objekte, die Entwicklung der Situation zu berechnen.

Das Ego in der Autobahn-Situation oben kann für alle beobachteten Objekte, ohne dass

- 15 dafür extra Routinen programmiert werden müssen:

- deren Verhalten prognostizieren mit der unterstellten Bedingung, alle Verkehrsteilnehmer wollen so schnell wie möglich und ohne Unfall vorankommen,
- laufend die Prognosen mit dem beobachteten tatsächlichen Verhalten (bremsen, 20 beschleunigen, lenken, Blinker setzen, Lichthupe betätigen etc.) anpassen und
- selbst in einem geeigneten Moment den LKW zu überholen.

Das mögliche Verhalten aller Beteiligten in dieser Situation ist durch das Ego ohne Probleme abzuschätzen, in dem es die eigenen Methoden und Erfahrungen (Daten) mit den jeweils angepassten Parametern auf sie anwendet. Macht der Mensch das nicht genauso?

- 25 Die Beispiele machen klar, die Basis, die Grundlage und der Ausgangspunkt dieser KI ist immer das Ego, das Zentrale Template. Je differenzierter seine Möglichkeiten bei der internen und der externen Sensorik und je besser die Ausbildung in der juvenilen Phase, desto intelligenter sein Verhalten und desto größer sein Potential, das Verhalten beobachteter Objekte zu prognostizieren. Wenn man dem Ego auch einen Parameter für die 30 eigene Verwundbarkeit, bei einem Auto würde man da wohl eher von Verformbarkeit

sprechen, könnte das Ego auch die Risiken eines starken oder schwachen Kontaktes abschätzen. ABER, es muss immer NUR (einmal) das Ego gebaut, programmiert und ausgebildet werden, wobei er letzte Teil natürlich auch von einem ausgebildeten Ego übertragen werden kann. Beobachtete Objekte müssen weder identifiziert (KNNs und ihre Trainingsdaten), noch muss das Verhalten der identifizierten Objekte programmiert werden. Man könnte natürlich einwenden, das Ego reicht nicht aus, das Verhalten anderer Verkehrsteilnehmer zu prognostizieren. Aber da muss man sich nur überlegen, wie der Mensch das Verhalten anderer Verkehrsteilnehmer versucht abzuschätzen. Auch der Mensch kennt nicht die Fernziele der Verkehrsteilnehmer auf einer Autobahn, auch er kann nur die kurzfristigen Ziele der anderen Verkehrsteilnehmer (so schnell wie möglich unfallfrei vorankommen) errahnen und sich darauf vorbereiten.

Die Fähigkeiten des Roboters:

Empathie:

Oben wurde gezeigt, wie sich das Ego mit Hilfe der Alter Egos in die Situationen der beobachteten Objekte seiner Umgebung versetzen kann, um deren Verhalten aus deren Sicht zu berechnen. Man kann es Empathie nennen, wenn man den Begriff nicht ausschließlich für den Menschen reserviert.

Lernen aus der Beobachtung:

Lernen aus Beobachtung bedeutet, man beobachtet ein Verhalten, das im Vergleich zum eigenen möglicherweise besser ist und übernimmt es in dem Fall und falls möglich. Die Basis ist der Vergleich zwischen fremdem und eigenem Verhalten und dieser Vergleich wird ermöglicht durch die oben definierte Empathie.

Empathie des Egos wird hier über Alter Egos realisiert, über die das Ego sich in die Situation anderer beobachteter Objekte versetzt, um aus deren Sicht auf die Umgebung deren Verhalten zu prognostizieren. Angenommen, es ergibt sich jetzt eine Differenz zwischen prognostiziertem und beobachtetem Verhalten. Und wenn jetzt das Alter Ego die Einsatzfolge der Effektoren so verändert, dass das beobachtete Verhalten nachgebildet wird, hat das Ego die Möglichkeit das beobachtete Verhalten zu kopieren, wenn es denn einen Vorteil bietet.

Figur 5 zeigt Szenario 3 unterschiedliches „Kurvenverhalten“.

Stellen wir uns vor das Ego hat gelernt, eine Kurve immer mit dem gleichen Abstand zum rechten Fahrbahnrand zu fahren, die Linie LE. Nun beobachtet es ein vorausfahrendes Auto, das die Kurve abschneidet, indem es früher aber weniger stark einlenkt, wodurch sich auch die Seitbeschleunigung verringert und somit schneller um die Kurve fährt, auf der
5 gepunkteten Kurve LB.

So gesehen ist die hier beschriebene Empathie Voraussetzung für das Lernen aus der Beobachtung und dem Nachmachen des Verhaltens anderer. (Was spricht dagegen, ähnliches auch für den Menschen zu vermuten, also Lernen durch Imitation auf Basis der Empathie.)

- 10 Wenn nun die Daten der externen Sensorik, umgerechnet auf die Situation und Position des Alter Egos, so mit den gespeicherten, einfachen und komplexeren Handlungsabfolgen des Egos verknüpft werden, die ja in das Alter Ego kopiert wurden, ist das Ego in der Lage aus der Beobachtung zu lernen: Es kann die Differenz der selbst geplanten Handlungsfolge der eigenen Effektoren (immer gleicher Abstand zum Fahrbahnrand) zur beobachteten (früheres
15 bremsen und einlenken, kleinere Lenkbewegung und höhere Kurvengeschwindigkeit und früheres beschleunigen) erkennen und feststellen, dass es so schneller durch die Kurven fahren könnte. Diese Form des Lernens ist sicher schneller als das übliche Versuch-und-Irrtum.

Erkennen von Kausalgesetzen:

- 20 Im Abschnitt Theoriebildung wurde erläutert, wie jeder beobachtete, funktionale Zusammenhang nachgebildet werden kann. Verwenden wir diese Fähigkeit für das Erkennen von Naturgesetzen, wofür auch, falls nötig und geeignet, Alter Egos verwendet werden. Natürlich liegen oder fahren die nicht auf einer Straße, aber sie bewirken Veränderungen, die von einer externen und/oder internen Sensorik erfasst werden können.
- 25 Stellen wir uns ein Experiment vor, bei dem unser KI-Ego einen fallenden Apfel beobachtet. Es erstellt ein Alter Ego und bemerkt die beschleunigte Bewegung Richtung Erde. Nun wird die erste Unterstellung des Egos sein, der Apfel hat selbst die Beschleunigung verursacht, es verfügt wie das Auto-Ego über ein eigenes ‚Gaspedal‘, um sich zu beschleunigen. Dann wird das Ego aber beobachten, dass sich kein Apfel auf dem Boden selbst bewegt und dass
30 andere Dinge ebenfalls zur Erde fallen und sich dann dort nicht weiterbewegen. Außerdem dürfte das Ego selbst das Gravitationsgesetz kennengelernt haben. Auf einer abschüssigen Straße würde es, ohne das Gaspedal zu betätigen, eine Beschleunigung bemerkt und

gelernt haben, aber immer nur ‚bergab‘, während es ‚bergauf‘ mehr Gas geben muss. Je steiler der Winkel α , desto größer der die Kraft gemäß dieser uns bekannten Formel (mit g =Erdbeschleunigung, $9,81 \text{ m/sec}^2$):

$$F = m * g * \sin(\alpha) \quad (2)$$

- 5 Wir erinnern uns, das Ego kennt natürlich nicht explizit diese Gleichung oder die Gleichung für den freien Fall, also das Gravitationsgesetz ($\sin(\alpha=90^\circ) = 1$), aber über die Datenpunkte und der stückweisen Linearen Inter- oder Extrapolation hat es eine implizite Kenntnis über diesen Zusammenhang der relevanten Größen.

- Das Ego verhält sich also so, als ob es wüsste, dass es, wie ein Physiker sagen würde, als
10 Körper mit schwerer Masse der Massenanziehung dem obigen Gesetz unterworfen ist. Wenn es nun ein anderes Objekt sieht und dessen Alter Ego erstellt, ist auch das Alter Ego dieser Massenanziehung unterworfen. Auf diese Weise erhält diese Erkenntnis quasi den Status eines Naturgesetzes, das auf alle Körper wirkt, ohne dass es extra einprogrammiert werden muss. Wenn das Verhalten eines solchen Egos von außen betrachtet werden
15 würde, könnte man nicht sagen, ob es das Gravitationsgesetz tatsächlich kennt, wie es ein Physiker tut, oder ob es diese Einsicht nur vortäuscht. Selbst, wenn das Ego eine Taube auf der Straße sitzen sieht, die wegfliegt, sobald sich das Ego nähert, kann es erkennen, dass Flügelschlag und Aufwärtsbeschleunigung korrespondieren.

- Ein anderes Beispiel wäre ein starker Seitenwind. Das Ego sei ein großer leerer
20 Kastenwagen. Bei einer normalen Geradeausfahrt ist die Seitbeschleunigung gleich Null. Dann aber wird der Kastenwagen während der Geradeausfahrt plötzlich seitversetzt und der Sensor für die Seitbeschleunigung schlägt an. In diesem Fall kann das Ego für die unerwartete Kraft und die unbekannte Ursache, einfach ein neues Alter Ego erstellen und das mit dieser Seitbeschleunigung verknüpfen. Wenn das Ego dann auch noch die Daten
25 der Umgebung, wie im Wald oder der Ebene, oder die Bewegung von Ästen erfassen und dem Datenspeicher zufügen kann, hat das Ego nicht nur eine bisher unbekannte Größe (Seitenwind) erkannt, es hat auch einen möglichen ursächlichen Zusammenhang erkannt. Wenn das Ego auch noch das Ausmaß der Bewegungen der Äste sehen und beurteilen könnte, könnte es sogar die Stärke der auf die Seite wirkenden Kraft in etwa abschätzen und
30 so den wahrscheinlichen Einfluss auf des Fahrverhaltens. Das Ego hat auf diese Weise selbst eine Theorie mit einer neuen Variablen erstellt. Ähnlich den Physikern, die Dunkle Materie und Dunkle Energie, wiewohl nicht sicht- und beobachtbar, einführten, oder wie der Nobelpreisträger Peter Higgs, der in den 1960er Jahren ein Teilchen konzipierte, das dann

2012 im Cern erstmalig nachgewiesen wurde. Das Alter Ego, also der Kopie des Selbst, kann mit diesem Vorgehen im Sinne einer Wahrheitssuche natürlich niemals die wahre Natur eines Phänomens erkennen, aber es reicht für das kompetente Erfassen der Situation.

Selbstbewertung:

- 5 Die erste Form des Erkenntnisgewinns wurde bereits beschrieben. Im Abschnitt über das „Erkennen von Kausalgesetzen“ wurde das Beispiel des Seitenwindes entwickelt. Eine dem Ego (zunächst) unbekannte Kraft erzeugt plötzlich auf gerader Strecke eine Seitbeschleunigung, eine messbare Wirkung auf den Roboter bzw. dem Ego. Dafür erstellt es ein Alter Ego. Allerdings fehlen zunächst weitere Informationen, die dem Ego
- 10 signalisieren, wann das Phänomen mit welcher Stärke auftritt. Erst mit der Beobachtung des Egos über einen zeitlichen Zusammenhang von den Bewegungen der Ästen, mal stärker und mal schwächer, und der eigenen Seitbeschleunigung, mal keine und mal doch, kann das Ego beides verbinden und für beides dieselbe Ursache vermuten. Was passiert hier? Das Ego kann den Seitenwind selbst nicht erkennen, aber die eigene Reaktion darauf
- 15 (Seitbeschleunigung) und die Reaktion der Äste. Damit erzeugt es ein Element, das mal eine Auswirkung auf das Ego hat und mal nicht, das aber nicht nur Auswirkungen auf das Ego hat, sondern auf andere beobachtbare Objekte auch.

- Die zweite Form der Erkenntnisgewinnung basiert auf dem Vergleich des eigenen Verhaltens mit dem von anderen Objekten. Voraussetzung ist die Annahme, dass das
- 20 Verhalten der beobachteten Objekte ‚rational‘ und nicht randomisiert ist, obwohl auch das als solches erkannt werden könnte, wodurch dann der Vergleich für einen möglichen Erkenntniszugewinn keine Maßgabe mehr wäre. Es werden zwei unterschiedliche Verhaltensvergleiche durchgeführt, wodurch das Ego feststellen kann, ob die eigene Kompetenz (s.u.) in einer Situation denen der beobachteten Objekte unterlegen, gleichwertig
- 25 oder überlegen ist. Das heißt, dass die beobachteten Objekte mehr, gleichviel oder weniger von der Situation wahrnehmen können. Ein beobachteter Kastenwagen verringert plötzlich auf einer geraden Straße die Geschwindigkeit. Er scheint etwas zu sehen, das dem beobachtenden Ego unbekannt ist, es erkennt weniger als das Objekt. Wurde eine unterlegene Kompetenz festgestellt, kann das als Anlass verstanden werden, diese
- 30 Situationen gezielt zu untersuchen, um die eigene Kompetenz zumindest auf das Niveau der anderen zu heben. Verglichen wird nicht, ob das Verhalten richtig oder falsch ist, das würde nur zu den bekannten Problemen der Wahrheitstheorien führen, sondern nur, ob und wann ein Verhalten sich ändert und ob sich Verhaltensweisen wiederholen. Der erste

Verhaltensvergleich betrifft unterschiedliche Situationen: Zeigen die beobachteten Objekte in für das Ego unterschiedliche Situationen auch unterschiedliche Verhaltensweisen? Der zweite Vergleich betrifft gleiche Situationen: Wiederholt sich das Verhalten der beobachteten Objekte in für das Ego gleichen Situationen oder variiert es? Aus diesen Beobachtungen von
5 mehreren Situationen über einen längeren Zeitraum kann das Ego schließen,

- dass die eigene Kompetenz unterlegen ist, wenn wiederholt in für das Ego gleichen Situationen beobachtete Objekte unterschiedliche Verhalten zeigen und in für das Ego unterschiedlichen Situationen auch ihr Verhalten variieren,
- 10 - dass die eigene Kompetenz überlegen ist, wenn wiederholt in für das Ego gleichen Situationen beobachtete Objekte ihr Verhalten wiederholen aber auch in für das Ego unterschiedlichen Situationen ihr Verhalten nicht variieren,
- dass die eigene Kompetenz äquivalent ist, wenn wiederholt in für das Ego gleichen
15 Situationen beobachtete Objekte ihr Verhalten wiederholen aber in für das Ego unterschiedlichen Situationen ihr Verhalten variieren.

Mit diesem einfachen Vergleichen kann das Ego erkennen, wie groß die eigene Kompetenz in bestimmten Situationen im Vergleich zu anderen Objekten ist. Wohlgemerkt, dies ist kein Vergleichsversuch mit der „Wahrheit der Realwelt“.

- 20 Aus diesen Vergleichen kann das Ego sich selbst bewerten und zum Beispiel ableiten, ob es sich in bestimmten Situationen vorsichtiger verhalten sollte und versuchen sollte das (noch) nicht Erkennbare zu identifizieren. Es kann als ein interner Auftrag verstanden werden, gezielt zu forschen als Teil des „stetigen Lernens und der fortgesetzte, Vergrößerung der Erkenntnis und des Wissens“ (s.o.).

- 25 Kompetenz:

Aus dem Ergebnis dieser Vergleiche lässt sich aber auch extern die Kompetenz der Fähigkeiten des Roboters beurteilen, um über seine Einsatzmöglichkeiten zu entscheiden. Im Gegensatz zur Wahrheit des Wissens über die Realwelt, dass das Ego sich erarbeitet hat, gibt es bei der Kompetenz keine sprachlichen Probleme, Kompetenzen zu vergleichen
30 und bei der Feststellung von größerer oder kleinerer Kompetenz.

Der Begriff der Wahrheit in Bezug auf Theorien hat aber noch einen wichtigen Aspekt, den

Kompetenz ebenfalls erfüllen muss. Theorien ermöglichen Prognosen. Daher ist es wichtig zu wissen, welche Qualität eine Theorie und in der Folge ihre Prognosen haben. Allgemein dient die Zuschreibung von Wahrheit zu einer Theorie als Kriterium, sie anwenden zu können, anwenden zu dürfen, mit dem Vorteil, dass das Kriterium Wahrheit alle

- 5 konkurrierenden Theorien automatisch ausschließt und sich so ein Entscheidungsproblem, welche Theorie zu verwenden wäre, nicht stellt. Der Begriff der Kompetenz muss Vergleichbares leisten.

- Kompetenz wird hier definiert und verwendet als das Erfassen der relevanten Einflussfaktoren einer Situation oder gleicher Situationen. Damit ist sie zunächst einmal ein
- 10 beschränktes Kriterium, dessen Höhe aber in einem Vergleichsverfahren ermittelt werden kann, ganz im Gegensatz zur Wahrheit einer Theorie, die eine zeitlich und räumlich unbeschränkte Gültigkeit unterstellt. Dies kann aber weder bewiesen werden noch erlaubt das Kriterium, unterschiedliche Theorien miteinander vergleichen.

- Damit ist Kompetenz zur Leistungsbeschreibung der Fähigkeiten eines Roboters besser
- 15 geeignet als der Begriff der Wahrheit und er kann, auch anders als Wahrheit, zur Selbstbewertung von dem Roboter in einem formalen Verfahren selbst ermittelt werden.

- Figur 6A zeigt in einem schematischen Ablaufdiagramm ein Verfahren zum autonomen Ansteuern einer Vorrichtung in der juvenilen bzw. der Ausbildungsphase. Ziel und Ende der Phase ist die ausreichend präzise Verwendung und Einsatz der Effektoren, ein möglichst
- 20 kleiner Fehler. Nach dem Initialisieren 200 kommt es zum randomisiertem Betätigen 201 von Effektoren und dem Auslesen 202 einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen Vorrichtungseigenschaften und einem Auslesen 203 von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften; ein Abspeichern 204 von Wechselbeziehungen, welche
- 25 zwischen dem Betätigen 201 der Effektoren und den ausgelesenen Vorrichtungseigenschaften und Umgebungseigenschaften vorliegen als Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von Betätigen 201, Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt ein Ansteuern 205 der Vorrichtung gemäß mindestens einer vordefinierten
- 30 Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer abgespeicherten 204 Handlungsanweisung, welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und/ oder die mindestens eine vordefinierte

Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird.

Figur 6A zeigt die Juvenile Phase:

- 200 Initialisierung als Vorbereitung zum Lernen
- 201 randomisiertes Betätigen der Effektoren
- 5 202 Auslesen 202 einer Vielzahl von internen Sensorikeinheiten
- 203 internen Vorrichtungseigenschaften und einem Auslesen 203 von einer Vielzahl von externen Sensorikeinheiten
- 204 Abspeichern 204 von Wechselbeziehungen
- 205 Feststellen des Fehlers der Bewegungspräzision, sofort, wenn zu groß weiter bei
- 10 201

Optional: Ruhephase zum Aufladen und ex Post zur Datenaufbereitung:

Fehler, Redundanz, Konsistenz,...

Figur 6B zeigt die Adulte Phase:

- 310 Initialisierung als Vorbereitung ein längerfristiges Ziel zu erreichen
- 15 311 Erkennen der Objekte 311 in der aktuellen Situation
- 312 Erstellen oder aktualisieren der Alter Egos bzw. deren Parameter 312
- 313 Verhaltensprognose 313 der beobachteten Objekte mittels der aktuellen Alter Egos
- 314 Berechnung des optimalen Effektoreneinsatzes 314 zur weiteren eigenen
- 20 Zielverfolgung
- 315 Ruhephase 315 zum Aufladen und der ex Post Datenaufbereitung:
- Fehler, Redundanz, Konsistenz,...

Vorliegend werden Roboter , Vorrichtungen, Systeme und Ego synonym verwendet.

Patentansprüche

1. Verfahren zum autonomen Ansteuern einer Vorrichtung, aufweisend:

- ein Initialisieren (100) mittels randomisierten Betätigens (101) von Effektoren und Auslesen (102) einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen
 - 5 Vorrichtungseigenschaften und einem Auslesen (103) von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur Messung von externen Umgebungseigenschaften;
 - ein Abspeichern (104) von Wechselbeziehungen, welche zwischen dem Betätigen (101) der Effektoren und den ausgelesenen Vorrichtungseigenschaften und
 - 10 Umgebungseigenschaften bestehen als Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von Betätigen (101), Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt; und
 - ein Ansteuern (105) der Vorrichtung gemäß mindestens einer vordefinierten Zielvorrichtungseigenschaft und mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer abgespeicherten (104) Handlungsanweisung,
 - 15 welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird, wobei diejenige mindestens eine Handlungsanweisung verwendet wird, welche einen Motivationswert maximal erhöht und der Motivationswert angibt wie nahe das Überführen von Ausgangstripel in das Zieltripel die mindestens ein
 - 20 vordefinierte Zielvorrichtungseigenschaft und die mindestens eine vordefinierte Zielumgebungseigenschaft erreicht.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass das Verfahren lernend randomisiert iteriert wird, derart, dass stets neue Handlungsanweisungen erkannt werden, welche jeweils ein Ausgangstripel in ein Zieltripel überführen und diese
- 25 Handlungsanweisungen zum Ansteuern (105) der Vorrichtung gespeichert werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass das Auslesen (102, 103) von internen und/ oder externen Effektoren derart erfolgt, dass einzelne Stützwerte abgespeichert werden und Zwischenwerte interpoliert und/ oder extrapoliert werden.
4. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass
- 30 die Zielvorrichtungseigenschaft und/ oder die Zielumgebungseigenschaft mindestens temporär von einer Vorrichtungseigenschaft und/ oder einer Umgebungseigenschaft

beeinflusst wird.

5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass mehrere Handlungsanweisungen zu einer Choreografie kombiniert werden, welche ein Ausgangstripel in ein Zieltripel überführt.
- 5 6. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Verfahren auf einer ersten Vorrichtung zur Ausführung gebracht wird und anhand von externen Sensorikeinheiten der ersten Vorrichtung eine zweite Vorrichtung erkannt wird, welche das Verfahren ebenfalls ausführt und deren erzeugte Wechselwirkungen auf der ersten Vorrichtung abgespeichert werden .
- 10 7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, dass die erste Vorrichtung anhand der erfassten Wechselwirkungen der zweiten Vorrichtung angesteuert (105) wird.
8. Verfahren nach einem der Ansprüche 6 oder 7, dadurch gekennzeichnet, dass Wechselbeziehungen der erfassten zweiten Vorrichtung zwischen dessen Parameter und dessen Handlungen zur Erstellung von Handlungsanweisungen der ersten Vorrichtung
15 verwendet werden.
9. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass Handlungsanweisungen vor deren Umsetzungen simuliert werden und heraus der maximale Motivationswert ermittelt wird.
10. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass
20 interne Sensorikeinheiten einen Motivationswert, einen Effektorzustand, eine Effektoreinstellung, eine Effektorposition, eine Größe, ein Gewicht, eine Abmessung, eine Geschwindigkeit, eine Beschleunigung, ein Abbremsen, einen Batteriestand, eine Reichweite, eine Position, eine Prozessorauslastung, eine Speicherauslastung und/ oder einen Systemparameter der Vorrichtung erfassen.
- 25 11. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass empirisch technische Einflussfaktoren auf den Motivationswert ermittelt werden und diese mit einer abgespeicherten Metrik zur Bestimmung des maximalen Motivationswerts korreliert werden.
12. Vorrichtung eingerichtet zum Ausführen eines Verfahrens nach einem der
30 vorhergehenden Ansprüche, aufweisend:

- eine Initialisierungseinheit eingerichtet zum Initialisieren (100) mittels randomisierten Betätigens (101) von Effektoren und Auslesen (102) einer Vielzahl von internen Sensorikeinheiten eingerichtet zur Messung von internen Vorrichtungseigenschaften und einem Auslesen (103) von einer Vielzahl von externen Sensorikeinheiten eingerichtet zur
- 5 Messung von externen Umgebungseigenschaften;
- eine Speichereinheit eingerichtet zum Abspeichern (104) von Wechselbeziehungen, welche zwischen dem Betätigen (101) der Effektoren und den ausgelesenen Vorrichtungseigenschaften und Umgebungseigenschaften bestehen als Handlungsanweisungen, derart, dass eine Handlungsanweisung ein Ausgangstripel von
- 10 Betätigen (101), Vorrichtungseigenschaften und Umgebungseigenschaften in ein Zieltripel überführt; und
- eine Steuerungseinheit eingerichtet zum Ansteuern (105) der Vorrichtung gemäß mindestens einer vordefinierten Zielvorrichtungseigenschaft und/ oder mindestens einer vordefinierten Zielumgebungseigenschaft unter Verwendung von mindestens einer
- 15 abgespeicherten (104) Handlungsanweisung, welche angibt wie die mindestens eine vordefinierte Zielvorrichtungseigenschaft und/ oder die mindestens eine vordefinierte Zielumgebungseigenschaft anhand des Ausgangstripels und des Zieltripels eingestellt wird, wobei diejenige mindestens eine Handlungsanweisung verwendet wird, welche einen Motivationswert maximal erhöht und der Motivationswert angibt wie nahe das Überführen
- 20 von Ausgangstripel in das Zieltripel die mindestens ein vordefinierte Zielvorrichtungseigenschaft und die mindestens eine vordefinierte Zielumgebungseigenschaft erreicht.
13. Systemanordnung aufweisend mehrere Vorrichtungen nach Anspruch 12 welche kommunikationstechnisch gekoppelt sind und Handlungsanweisungen austauschen.
- 25 14. Computerprogrammprodukt, umfassend Befehle, die bei der Ausführung des Programms durch mindestens einen Computer diesen veranlassen, die Schritte des Verfahrens nach einem der Ansprüche 1 bis 11 auszuführen.
15. Computerlesbares Speichermedium, umfassend Befehle, die bei der Ausführung durch mindestens einen Computer diesen veranlassen, die Schritte des Verfahrens nach einem
- 30 der Ansprüche 1 bis 11 auszuführen.