

Prüfungsfragen zur VO "Einführung in die Artificial Intelligence"

1) Einführung und Geschichte der Artificial Intelligence

a) Durch welches Resultat wurde Hilberts Programm zunichte gemacht?

Durch das Gödel'sche Theorem (*Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme*), das 1931 gezeigt hat, dass die Zahlentheorie nicht (rekursiv) axiomatisierbar ist. Jede (rekursive) Axiomatisierung der Zahlentheorie ist unvollständig. Erst durch die Einführung eines mächtigeren Formelsystems ω , das k umfasst, kann k selbst vollständig bewiesen werden.

b) Nennen Sie mindestens 3 Wurzeln der Artificial Intelligence!

Philosophie, Psychologie, Linguistik, Informatik, Mathematik

Leibniz (1676-1711) *Calculus Ratiocinator* zur Automatisierung des Mathematischen Schließens

Hilbert (1862-1943) Programm zur vollständigen und konsistenten Axiomatisierung der Mathematik.

Russel und **Whitehead**: *Principia Mathematica* (1910) Versuch der Realisierung von Hilberts Programm.

Gödel 1931: *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme*. Gödel zeigt, dass die Zahlentheorie nicht (rekursiv) axiomatisierbar ist. Zu jeder

Axiomatisierung A der Zahlentheorie gibt es einen wahren Satz G , sodass weder $A \vdash G$, noch $A \vdash \neg G$. (\vdash : ableitbar). D.h. Jede (rekursive) Axiomatisierung der Zahlentheorie ist unvollständig. Hiermit wurde Hilberts Programm zunichte gemacht. Der Traum einer "mechanisierbaren" Mathematik war zu Ende.

Church (1936): Die Prädikatenlogik ist unentscheidbar. Seit diesen negativen Resultaten: Hoffnung, dass wenigstens Teilgebiete des Denkens, Entscheidens und Handelns automatisierbar sind.

Turing, 1950. siehe untere Frage.

McCarthy, 1956: Führt den Begriff der *Artificial Intelligence* ein.

Postulate von McCarthy:

-) Ein Intelligentes System muss eine *Repräsentation* der Welt besitzen, d.h. auf eine Menge von Fakten zugreifen, sie verarbeiten und erweitern können.

-) Zur Repräsentation dienen Mengen von *logischen* Aussagen, bzw. Formeln, die Problemlösungen als Folgerungen enthalten.

-) Die Problemlösung muss *auffindbar* sein. D.h., es muss eine wirksame Suchstrategie für Beweise existieren.

c) Beschreiben Sie die Funktion des Turing-Tests!

Der Turing-Test soll maschinelle Intelligenz testen.

Funktionsweise: Turing Test für maschinelle Intelligenz: Ein nur über Fernschreiber mit je einem Menschen und einem Rechner verbundener Tester muss herausfinden, wer sein menschlicher und wer sein maschineller Partner ist. Gelingt das nicht, liegt künstliche Intelligenz vor. (Seit ca. 1990: Jährlicher Wettbewerb mit Preis \$100.000.-, bis zum Bestehen des Tests).

d) Wie hieß das in Dartmouth vorgestellte erste AI-Programm und was machte es?

1955 entwickelten Newell und Simon „The Logic Theorist“, das von vielen als erstes Programm der Künstlichen Intelligenz betrachtet wurde. Das Programm stellte das Problem in einem Baum dar und verfolgte dann den richtigen Ast der zur Lösung des Problems führte.

e) Wo fand die erste IJCAI Konferenz statt? Was wurde dort u.a. vorgestellt?

“International Joint Conference on Artificial Intelligence”

1969 in Stanford: Roboter Shakey kann log. Prozesse kombinieren und interagieren. Kennt Unterschied zwischen Soll und Ist Zustand.

f) Wie hieß das erste Expertensystem und was war seine Aufgabe?

1974 Erstes Expertensystem MYCIN entsteht. Programm, das Diagnosen zu ansteckenden Krankheiten des Blutes stellt. Erstes Programm wo die Wissensbasis und die logischen

Verknüpfungen getrennt sind.

g) **Wo findet man heutzutage im Alltag Anwendungen der AI?**

-) Expertensysteme und Entscheidungsunterstützung
-) Konfiguration von elektronischen Geräten und Anlagen
-) Medizinische Diagnose
-) Diagnose und Überprüfung technischer Anlagen
-) Game playing, automatische Spieler
-) Robotersteuerung
-) Sprachübersetzungssysteme, natürlichsprachiger Datenbankzugriff
-) Scheduling, Ablaufsteuerung und Optimierung industrieller Prozesse
-) Informationsgewinnung aus dem Internet
-) Shopping agents, automatisches Verhandeln
-) Computerunterstütztes Lernen
-) Mustererkennung
-) Bildverstehen

h) **Wofür sind McCulloch und Pitts bekannt?**

McCulloch-Pitts-Zelle: ist ein von Warren McCulloch und Walter Pitts im Jahr 1943 vorgeschlagenes Neuronenmodell. Beide wollten ein vereinfachtes Modell realer Vorgänge in neuronalen Strukturen entwerfen, um zu klären, ob das Gehirn die turing-berechenbaren Funktionen wirklich berechnen kann. 1943 McCulloch und Pitts: erstes Neuronenmodell kann beliebige arithmetische und logische Funktionen berechnen - noch ohne Lernen

2) **Komplexität: (Methoden zur Lösung schwieriger Probleme)**

a) **Was bedeutet P, NP, EXPTIME ? (Kurze Beschreibung)**

Es geht um Zeitkomplexität von Algorithmen.

P (polynomiell lösbar): Ein Problem ist durch einen Algorithmus beweisbar polynomiell lösbar. (z.B. n^2)

NP (nichtdeterministisch polynomiell lösbar): Kein Problem in NP konnte bisher als nichtpolynomiell lösbar bewiesen werden. Für viele Probleme in NP kennt man aber keinen polynomiellen Algorithmus.

EXPTIME (exponentiell lösbar): Ein Problem ist durch einen Algorithmus beweisbar nur in exponentiell steigender Zeit lösbar.

b) **Nennen Sie zwei NP-vollständige Probleme**

Rucksackproblem, Graphen-Dreifärbbarkeit, Traveling Salesman Problem, Kreuzworträtsel

c) **Nennen Sie drei Lösungsansätze für NP-vollständige Probleme**

- randomisierte lokale Suche
- Approximation:
- Identifikation leicht lösbarer Subklassen

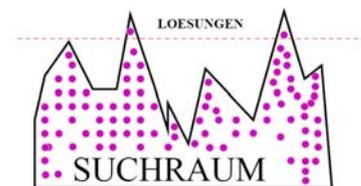
d) **Beschreiben Sie kurz das Verfahren der Randomisierten Suche**

Verfahren:

1. Erzeuge zufälligen Lösungskandidaten.
2. Führe solange wie möglich lokale Verbesserungen durch.
3. Wenn Lösung gefunden: Ausgeben und Programm beenden.
4. Wenn Zeitlimit erreicht, Abbruch: "Keine Lösung gefunden"
5. Gehe zu Schritt 1.

Vorteile der randomisierten lokalen Suche:

- Funktioniert erstaunlich gut für viele Probleme von Praktischer Relevanz.
- Führt dann mit hoher Wahrscheinlichkeit zum Erfolg.
- Liefert vollständige Lösungen.
- Ist ein natürliches Verfahren.



Nachteile:

- Funktioniert schlecht, falls keine Lösung existiert.
- Liefert keinen "Unlösbarkeitsbeweis"
- Funktioniert schlecht bei Instanzen mit wenigen Lösungen.
- Funktioniert nur mit Bewertungsfunktion. (Z.B. nicht geeignet zum Knacken von Codes.)
- Lokales Optimum (Suboptimum) muss nicht unbedingt eine Lösung sein

e) **Beschreiben Sie kurz das Verfahren der NP-Approximation**

Ziel: Finde annähernd optimale Lösung bzgl. Bewertungsfunktion w .

Ein Maximierungsproblem ist ϵ -approximierbar, wenn es einen Algorithmus T mit polynomieller Laufzeit gibt, sodass für alle Instanzen x gilt:

$$\max w(x) - w(T(x)) / \max w(x) \leq \epsilon.$$

Approximationsgrad $\gamma(A)$ eines Maximierungsproblems A : Größte untere Schranke für ϵ -Approximierbarkeit. $\gamma(A) = \inf \{ \epsilon \mid A \text{ ist } \epsilon\text{-approximierbar} \}$.

Beispiele

- $\gamma(\text{RUCKSACK})=0$; beliebig approximierbar.
- $\gamma(\text{TSP})=1$ außer $P=NP$; überhaupt nicht approximierbar.
- $\gamma(\text{TSP/EUCLID}) \leq 1/3$
- $\gamma(A)=0$ bedeutet nicht, dass A polynomiell lösbar!

Vorteil der Approximationsmethode:

- Approximation in der Praxis oft ausreichend (brauche keine exakte Lösung).
- Auch auf manche beweisbar exponentielle Probleme anwendbar.

Nachteile:

- Wenige praktische Probleme approximierbar.
- Nicht auf Entscheidungsprobleme anwendbar.

Approximationstheorie: Eleganter und reichhaltiger Zweig der theoretischen Informatik.

f) **Wodurch ist NP-Härte bei Graphenproblemen oft bedingt?**

Durch Zyklizität der Graphen.

g) **Nennen Sie eine polynomielle Subklasse für das Dreifärbbarkeitsproblem**

azyklische Graphen, Graphen mit konstant beschränkter Baumweite ($\text{bw}(G) \leq \text{fixer Konstante } k$).

3) Wissensrepräsentationa) **Nennen Sie die drei Ebenen eines Wissensbasierten Systems**

Wissensrepräsentation, Meta-Prozessor, Inferenzkomponente (2 und 3 = Wissensverarbeitung)

b) **Vor- und Nachteile der prozeduralen Wissensrepräsentation**

Verfahren zur Konstruktion, Verknüpfung und Anwendung von Wissen

Vorteil: sehr effizient

Nachteile: aufwendig zu ändern, falls sich Wissen ändert; Die Reihenfolge der Problem-lösungsschritte muss detailliert algorithmisch „vorgedacht“ sein; Es besteht die Gefahr unerwünschten Systemverhaltens bei „nicht vorgedachten“ Situationen

c) **Wie ist der Ablauf in einem Regelinterpreter?**

Regelinterpreter: sucht nach anzuwendenden (relevanten) Regeln, wertet sie aus und löst fällige Aktionen aus.

Ein Regelinterpreter ist für die Abarbeitung von Regeln zuständig. Dabei kann der Regelinterpreter zur Abarbeitung zwei Strategien verfolgen:

Vorwärtsverkettung (forward chaining = explorierend):

Bei dieser Strategie wird ausgehend von einer vorhandenen Datenbasis aus der Menge der Regeln, deren Vorbedingung durch die Datenbasis erfüllt ist, eine Regel ausgewählt, ihr Aktionsteil ausgeführt und damit die Datenbasis geändert. Dies wird solange wiederholt, bis keine Regel mehr anwendbar ist.

Rückwärtsverkettung (backward chaining = verifizierend):

Bei dieser Strategie startet der Regelinterpretier mit einem vorgegebenen Ziel. Wenn das Ziel nicht in der Datenbasis bekannt ist, entscheidet der Regelinterpretier zunächst, ob es abgeleitet werden kann oder erfragt werden muss.

Im Fall der Ableitung werden alle Regeln abgearbeitet, in deren Aktionsteil das Ziel enthalten ist. Wenn bei der Überprüfung der Aussage einer Regel ein Parameter unbekannt ist, wird ein Unterziel zur Bestimmung dieses Parameters generiert und versucht, dieses Unterziel durch Ableiten zu erreichen. Das Endergebnis ist die Bestimmung eines Wertes für das vorgegebene Ziel und alle Unterziele, die Evaluation der relevanten Regeln und das Stellen der notwendigen Fragen. Die Rückwärtsverkettung enthält also implizit eine Dialogsteuerung, wobei die Reihenfolge der gestellten Fragen von der Reihenfolge der Regeln zur Ableitung eines Parameters und von der Reihenfolge der Aussagen in der Vorbedingung einer Regel abhängt. Diese Abhängigkeit von der Reihenfolge erschwert eine Modularisierung des Regelsystems.

d) **Was ist eine Beschreibungslogik (description logic)?**

Beschreibungslogiken (engl. description logics) sind eine Familie von Sprachen zur Wissensrepräsentation. Die meisten Beschreibungslogiken sind eine Untermenge der Prädikatenlogik erster Stufe, im Gegensatz zu dieser aber entscheidbar. Dies ermöglicht über eine Beschreibungslogik zu schließen, d.h. aus vorhandenem Wissen neues Wissen zu gewinnen.

e) **Was wird mit Skripten beschrieben?**

Formalismus, um Vorgänge zu repräsentieren (im Prinzip Prozesse im Sinne des Prozessengineering) bestehend aus:

-) Ereignisse, die stattfinden
-) Reihenfolgeinformationen
-) Eintrittsbedingungen und Resultate
-) beteiligte Rollen

Repräsentation ähnlich zu Frames, aber zusätzliche Interpretation, die im Metaprozessor realisiert werden kann

f) **Was bedeutet "die klassische Logik ist monoton"?**

Für alle Wissensbase KB1 und KB2 gilt:

Alles was aus KB1 logisch folgt, folgt auch aus der erweiterten Wissensbasis KB1 KB2.

Klassische Logik (wie Aussagen- und Prädikatenlogik) verfügen über die Monotonie-Eigenschaft. Diese besagt im Wesentlichen, dass durch Schlussfolgerungen lediglich neues Wissen gewonnen, nicht aber bereits vorhandenes Wissen revidiert werden kann.

4) **Konfiguration**

a) **Was ist eine key component?**

Schlüsselkomponente: Teil der in der Konfiguration enthalten sein muss

b) **Erläutern Sie die strukturbasierte Konfiguration.**

-) Zerlegungsbäume können auch rein hierarchisch sein(Computer-Gehäuse-Bus)
-) Entwickelt aus Stücklistenverarbeitung, einfacher Ansatz aber beschränkt
-) Vorgaben durch key Components und globale Anforderungen (Platz, Kosten, Strom)

c) **Beschreiben Sie Horizont- und Schwelleneffekte.**

Horizont-Effekt: spätes Feststellen von getroffenen Fehlentscheidungen. In der Regel sind keine korrekten Abschränkungen für Heuristiken möglich, da immer ein noch nicht überprüfter Constraint auftauchen kann.z.B.: Hätte man geahnt, dass eine zusätzliche Festplatte notwendig ist, hätte man ein anderes Netzteil vorgesehen.

Schwellen-Effekt: • kleine Unterschiede in der Spezifikation können große Effekte auslösen.

-) Z.B. {A, D, D} statt {A, D, B}: ein Gehäuse reicht nicht mehr - größere Kosten, Gewicht, oder größerer Stromverbrauch benötigt größeres Netzteil, ...
-) Problematik: globale Ressourcen (Gewicht, Kosten) abhängig von lokalen Entscheidungen

d) **Was ist Partial Commitment?**

~tws. Bindung

Es wird Komponente A oder B eingebaut, beide brauchen aber Komponente C. C wird daher festgelegt und A, B wird erst am Schluss entschieden. (BSP: CD-Player & Verstärker)

e) **Was ist Propose-Revise?**

Stelle eine Lösung zusammen, benutze Prüfungswissen, um Schwachstellen zu finden, benutze heuristisches Reparaturwissen, um zu reparieren (z.B. VT: System für Entwurf von Aufzügen)

f) **Erläutern Sie den Unterschied in der Vorgehensweise zwischen regelbasierter und modellbasierter Konfiguration.**

-) 1 Constraint => mehrere Regeln. Bestimmte Möglichkeiten werden ausgelassen, ungültige Varianten müssen explizit deklariert werden.

-) Regeln werden durch ausdrucksstärkeres Konzept der Constraints ersetzt

-) Variablen (Package, Roof) entsprechen Plätzen in der Konfiguration, Werte den Komponenten $\text{dom}(\text{Package}) = \{\text{lux}, \text{std}\}$

-) Lösung: alle Variablen zugewiesen; alle Constraints erfüllt

Bereichsabhängige Vorgehensweise ersetzt durch allgemeines Suchverfahren (korrekt bzgl. Constraints, vollständig, steuerbar durch Heuristiken)

g) **Welcher Unterschied ergibt sich zwischen constraintbasierter und regelbasierter Konfiguration bei der Beschreibung der Wissensbasis?**

-) deklarativ - Was ist eine Lösung, statt wie wird sie gebaut

-) Schnellere Beschreibung --> leichtere Wartung

-) 1 Constraint-> mehrere Regeln. Bestimmte Möglichkeiten werden ausgelassen, ungültige Varianten müssen explizit deklariert werden.

-) ODER Verknüpfungen müssen in Regeln umgesetzt werden

h) **Geben Sie die Vorteile des constraintbasierten Konfigurationsansatzes gegenüber dem regelbasierten an.**

-) deklarativ - Was ist eine Lsg, statt wie wird sie gebaut;

-) Schnellere Beschreibung, leichtere Wartung

-) 1 Constraint-> mehrere Regeln. Bestimmte Möglichkeiten werden ausgelassen, ungültige Varianten müssen explizit deklariert werden.

-) Lsg. : alle Variablen zugewiesen; alle Constraints erfüllt

Bereichsabhängige Vorgehensweise ersetzt durch allgemeines Suchverfahren(korrekt bzgl. Constraints, vollständig, steuerbar durch Heuristiken)

-) Generative Constraint

Weiterentwicklung des Constraintgraphen; Komponenten mit Ports und Attributen

Verfahren:

- Auswahl der Komponenten der Verbindung
- Entnahme der Komp. aus Katalog

5) Diagnosea) **Was versteht man unter Diagnose? Was ist die Idee der Modellbasierten Diagnose?**

Diagnose ist die Identifikation von Komponenten eines (physikalischen) Systems, die für ein fehlerhaftes Verhalten verantwortlich sind. Die Modellbasierte Diagnose vergleicht den tatsächlichen Zustand eines Systems mit dem Sollzustand des Modells, auf dem das System basiert.

b) **Warum heißt die Modellbasierte Diagnose auch "Diagnosis from First Principles"?**

Weil man das Verhalten eines physikalischen Systems mit einem theoretischen Modell vergleicht und aus dem Unterschied zwischen diesen die Diagnose ableitet.

c) **Welche Verfahren zur Diagnose gibt es? Was sind die Nachteile? Was sind die Vorteile?**

Decision Trees, Neuronale Netze, Regelbasierte Systeme, Bayes'sche Netze

Nachteile dieser Systeme: spezifisch für ein System, schlecht an andere Systeme adaptierbar,

Erstellung eines Modells i.A. schwer, eingeschränkte Diagnose, wenig Bezug zu physikalischen Modellen

Vorteile: schnelle Diagnoseberechnung, begrenzter Speicherplatz notwendig, Einfachfehler meistens ausreichend

d) **Definieren Sie Diagnosesystem, Diagnoseproblem, und Diagnosen.**

Diagnosesystem: besteht aus einer Systembeschreibung und einer Menge von Komponenten.

Diagnoseproblem: besteht aus dem Diagnosesystem und einer Menge von Beobachtungen

Diagnose: Teilmenge der Komponentenmenge mit folgender Eigenschaft: Unter der Annahme, dass in der Diagnose alle Komponenten nicht funktionieren und alle anderen Komponenten funktionieren, erhält man keinen Widerspruch zu den Beobachtungen.

e) **Zählen Sie die Vor- und Nachteile der Modellbasierten Diagnose auf.**

Es werden physikalische Modelle verwendet, nur das korrekte Verhalten muss modelliert werden, es handelt sich um komponentenorientierte Modelle, Wiederverwendung der Modelle ist möglich, die Erstellung von Modellen komplexer Systeme gestaltet sich einfach, Änderungen sind nur lokal, es existieren schnelle Algorithmen und Berechnungszeit sowie Speicherplatz sind begrenzt.

f) **Was ist der Unterschied zwischen Quantitativen Modellen und Qualitativen Modellen.**

Quantitative Modelle haben einen unendlichen Wertebereich, verwenden Methoden der Analysis, insbesondere Differentialgleichungen.

Qualitative Modelle haben einen finiten Wertebereich und verwenden Methoden der Algebra sowie qualitative Differentialgleichungen.

g) **Was sind Fehlermodelle? Wozu werden sie benötigt?**

Fehlermodelle dienen der Einschränkung von Diagnosen; es sind dies Beschreibungen des Fehlverhaltens von Komponenten bzw. des Systems.

6) **Constraint Satisfaction Systeme**

a) **Aus welchen Komponenten besteht ein CSP? Wie lassen sich diese darstellen?**

Ein Constraint Satisfaction Problem besteht (formal) aus einer Menge von Variablen, einer Menge von Wertebereichen dieser Variablen und einer Menge von Constraints (Bedingungen), welche die Variablen erfüllen müssen. Die Darstellung erfolgt mit Hilfe dieser Regeln.

b) **Geben Sie ein praktisches Beispiel (nicht aus der Vorlesung und nicht das Kino-Beispiel) für ein CSP an. Welcher Aspekt des Beispiels entspricht welcher Komponente des CSP?**

Man möchte mit mehreren Personen eine Reise machen. Die Variablen sind dabei das Urlaubsziel, die Anfangs- und Endtag der Reise, der Umfang etc., Wertebereiche werden durch in Frage kommende Urlaubsziele und entsprechende Angebote festgelegt. Constraints sind Wünsche und Ansprüche der einzelnen Teilnehmer, die zur Verfügung stehende Zeit etc.

c) **Beschreiben Sie chronologisches Backtracking. Wodurch kann man die Effizienz des Verfahrens beeinflussen?**

Partielle Lösungen werden so lange erweitert, bis eine Lösung gefunden wird. Look-back prüft, welche bereits belegte Variable für eine Inkonsistenz verantwortlich ist und führt Backtracking bis zu dieser Variable durch. Look-ahead prüft, welche zukünftigen Variablenbelegungen auf Grund der bisherigen Belegungen nicht mehr möglich sind; diese Belegungen müssen nicht mehr geprüft werden.

d) **Was ist der Unterschied von Backjumping und Conflict-Directed Backjumping?**

Läuft die Suche bei der momentan belegten Variable x_i in eine Sackgasse, so springt Backjumping zur tiefsten Variable x_j , die mit x_i verglichen wurde. Beim Conflict-Directed Backjumping gibt es zu jeder Variablen x_i gibt es eine Konfliktmenge $A(x_i)$, die alle Variablen enthält, die inkonsistent mit der aktuellen Belegung von x_i sind. Bei Inkonsistenz von x_i mit der Belegung einer Variablen x_j , so wird x_j zu $A(x_i)$ hinzugenommen. Läuft x_i in eine Sackgasse, wird zur tiefsten Variablen x_j in $A(x_i)$ zurückgesprungen, $A(x_i) \setminus x_j$ wird zu $A(x_j)$ hinzugenommen.

e) **Nennen Sie zwei unterschiedliche Darstellungsarten von Constraints.**

Binäre Constraints: Constraint Graph, Erweiterter Constraint Graph

Nicht-binäre Constraints: Hypergraph

f) **Wie lassen sich unterschiedliche Lösungsverfahren für CSPs vergleichen?**

Ein derartiger Vergleich erfolgt unter Zuhilfenahme von Anzahl der besuchten Knoten im Suchbaum und Anzahl der Konsistenztests.

g) **Welche Vorteile hat es, ein gegebenes Problem als CSP zu lösen?**

CSP ermöglichen einheitliche Lösungsverfahren für verschiedenste Problemstellungen.

7) Wissensbasierte Suche

a) **Warum braucht man in der künstlichen Intelligenz im Speziellen effiziente Suchverfahren?**

Die Laufzeit nicht-effizienter Suchverfahren wäre schlichtweg zu lang.

b) **Was ist der Unterschied zwischen einer informierten und einer uninformierten Suchstrategie?**

Die informierte Suche verwendet zusätzliches Wissen über die Struktur, in der die Suche erfolgt.

c) **Was ist der Vorteil einer Suche mit iterativer Vertiefung?**

Die Vorteile von Breitensuche und Backtracking werden kombiniert.

d) **Was versteht man unter einer heuristischen Suche?**

Die Suche verwendet zusätzliches Wissen über die Struktur, in der die Suche erfolgt, um die Effizienz zu erhöhen.

e) **Aus welchen beiden Komponenten setzt sich die Schätzfunktion eines A* Algorithmus zusammen?**

Die Schätzfunktion ist die Summe zweier Schätzfunktionen, die jeweils die Kosten eines Knotens im Graphen zum Ausgangsknoten bzw. zum Zielknoten angeben.

f) **Welchen Vorteil besitzt der IDA* Algorithmus gegenüber dem A* Algorithmus?**

Man erhält linearen Speicheraufwand im Gegensatz zu exponentiellem Speicheraufwand beim A* Algorithmus.

g) **Was versteht man unter "alpha/beta pruning"?**

Alpha/Beta-Pruning ist ein Verfahren, um die Anzahl der besuchten Knoten in einem vom Minimax-Algorithmus durchlaufenen Baums zu reduzieren, der bei Nullsummenspielen eingesetzt wird.

8) Wissensbasiertes Planen

a) **Aus welchen Teilen besteht die Beschreibung eines Planungsproblems?**

???

b) **Aus welchen Teilen besteht ein Planungsoperator?**

Ein Planungsoperator besteht aus vier Listen: Einschränkungen, Vorbedingungen, Hinzufügungen und Aufhebungen.

c) **Was versteht man unter dem Situationskalkül?**

Der Situationskalkül ist eine auf der Prädikatenlogik basierende Beschreibung von Auswirkungen von Aktionen auf Situationen in einer Modellwelt.

d) **Was ist der Unterschied zwischen einer Eintrittsbedingung und einer Einschränkung bei einem Planungsoperator?**

Einschränkungen sind Bedingungen für die Instanziierung von Variablen, Eintritts- oder Vorbedingungen sind Bedingungen, die erfüllt sein müssen, damit der Operator ausgeführt werden kann.

e) **Spezifizieren Sie einen Planungsoperator, der beschreibt, wie ein Besucher in ein Kino geht!**

???

f) **Was ist eine nicht-lineare Planung?**

Bei nichtlinearer Planung werden einzelne Pläne ineinander kombiniert, es erfolgt regressive Planung (Einschub von Operatoren an geeignete Stellen des Plans).

g) **Was wird bei einer Situationsabstraktion gemacht?**

Es erfolgt die Bewertung von Aussagen von Situationen und Vorbedingungen.

9) *Maschinelles Lernen*

a) **Schreiben Sie einen abduktiven Syllogismus auf.**

Resultat: Hugo ist verdächtig.

Regel: Verdächtig ist man, wenn man kein Alibi hat.

Fall: Hugo hat kein Alibi.

oder:

Resultat: Sokrates ist sterblich

Regel: Alle Menschen sind sterblich

Fall: daher ist Sokrates ein Mensch

b) **Welche Arten maschinellen Lernens gibt es?**

Neurale Netzwerke (cf. Gehirn)

Genetische Algorithmen (cf. Evolution)

Induktive Logische Programmierung (Verallgemeinern)

Abduktion auf logischen Wissensbasen (Theorien erfinden)

Decision tree learning (Bsp: Tiere: Ratten)

Computationales/Statistisches Lernen (PACLearning)

Probabilistic reinforcement learning

c) **Was ist der Unterschied zwischen Unsupervised und Supervised Learning?**

Supervised Learning:

-) Trainingsdaten werden klassifiziert

-) Dem lernenden System wird mitgeteilt, ob ein Fall ein positives oder ein negatives Beispiel des Zielkonzepts darstellt

Unsupervised Learning:

-) Kategorienbildung ist das Grundproblem des unüberwachten Lernens

-) Wie kann ein Agent die Objekte in sinnvolle Kategorien einteilen, wenn eine Menge von Objekten gegeben ist, die unterschiedliche Eigenschaften aufweisen

Das lernende System bildet und bewertet die Konzepte selbständig:

z.B Naturwissenschaften:

-) Hypothese wird formuliert, um Beobachtungen zu erklären

-) Hypothesen werden bewertet anhand Kriterien wie Einfachheit, Allgemeingültigkeit,...

-) Hypothesen werden durch selbst entworfene Experimente getestet

d) **Beschreiben Sie den ID3 Algorithmus.**

-) Repräsentiert Konzepte als Entscheidungsbäume

-) Objekte werden klassifiziert, indem ihre Eigenschaften auf bestimmte Werte überprüft werden

e) **Geben Sie ein Beispiel für den ID3 Algorithmus an.**

Verschuldungsrisiko Einschätzung:

Eigenschaften: Kreditwürdigkeit, aktuelle Verschuldung, Sicherheiten, Einkommen

f) **Was ist Clustering?**

Konzeptuelles Clustering:

-) Sammlung nicht klassifizierter Objekte

-) Methode, um die Ähnlichkeiten von Objekten zu messen

-) Das Ziel besteht darin, die Objekte in eine Hierarchie von Klassen einzuordnen, die einem bestimmten Qualitätsstandard genügen, wie z.B die Ähnlichkeiten von Objekten zu maximieren, die der gleichen Klasse angehören

g) **Beschreiben Sie ein Beispiel für Reinforcement-Learning.**

Reinforcement-Learning:

-) Ein Agent wird hier einer Umgebung ausgesetzt und empfängt in diesem Kontext Rückmeldungen

-) Der Agent muss handeln und die Rückmeldungen zu diesen Handlungen interpretieren

-) Es gibt kein Lehrer, der direkt auf jede Handlung reagiert

-) Der Agent muss stattdessen selbst eine Strategie zur Interpretation der Rückmeldungen entwickeln
z.B: Radfahren lernen

10) Moderne Heuristiken

a) Was sind Heuristiken und warum werden sie verwendet?

Definitionen: „das Studium der Methoden und Regeln von Entdeckung und Erfindung“, „a process that may solve a given problem, but offers no guarantees of doing so“, Als Heuristik bezeichnet man eine Strategie, die das Streben nach Erkenntnis und

b) das Finden von Wegen zum Ziel planvoll gestaltet.

Verwendung: Ein Problem besitzt eine exakte Lösung, doch die Rechenkosten für die Suche nach dieser Lösung sind unerschwinglich (Schach TSP); Ein Problem hat keine Lösung, da die Problembeschreibung oder die verfügbaren Daten zweideutig sind (med. Diagnose, Sehvermögen)

c) Was ist der Unterschied zwischen Simulated Annealing und Stochastic Hill-Climbing?

Unterscheiden sich hauptsächlich in der Art und Weise wie sie: die Nachbarschaft erkunden & die nachfolgende Lösung der aktuellen Lösung akzeptieren.

Hill-Climbing: Anfangslösung s --> erzeuge Nachbarschaft $N(s)$ --> selektiere Lösung mit bester Fitness --> besser als aktuelle Lösung?

besser: selektierte Lösung = aktuelle & erzeuge wieder neue Nachbarschaft

schlechter: Ende

Stochastic Hill-Climbing: Nur eine Lösung aus der Nachbarschaft wird selektiert --> Diese Lösung wird für die nächste Iteration mit einer Wahrscheinlichkeit akzeptiert, die von der Evaluierungsdifferenz zwischen dieser Lösung und der aktuellen Lösung abhängig ist

Simulated Annealing: Während der Suche wird auch eine schlechtere Lösung mit einer bestimmten Wahrscheinlichkeit akzeptiert. Unterscheidet sich von der Stochastic-Hill-Climbing-Methode dadurch, dass der Parameter T während der Suche verändert wird (anfangs hoch, zu Ende niedrig).

Was sind die wichtigen Fragen bei Simulated Annealing?

Wichtige Fragen für Simulated Annealing:

-) Wie bestimme ich die Anfangs-„Temperatur“ T ?

-) Wie bestimme ich die Funktion $g(T,t)$?

-) Wie bestimme ich die „termination condition“?

-) Wie bestimme ich das „halting criterion“?

d) Welche Lokalen Suchmethoden sind deterministisch?

Tabu-Suche

e) Was ist eine Tabu Liste? Warum wird sie verwendet?

-) In der Tabu-Liste werden für eine fixe Anzahl von durchlaufenen Iterationen bestimmte Informationen zur Suche gespeichert (z.B. bereits besuchte Lösungen)

-) Zyklen (Besuchen von gleichen Lösungen) werden vermieden

-) Die Akzeptanz der Lösung für die nächste Iteration hängt ab von: Fitness der Lösung, Tabu-Liste

f) Welche Operatoren werden bei Genetischen Algorithmen verwendet? Was tun Sie?

Kreuzung: Der Kreuzungsmechanismus teilt zwei Lösungskandidaten auf und tauscht Komponenten gegeneinander aus, um neue Kandidaten zu produzieren.

Mutation: Ein Aspekt eines Kandidaten wird in zufälliger Weise verändert.

g) Geben Sie ein Beispiel für einen Kreuzung-Operator an.

Lösung 1: 111111, Lösung 2: 000000 Kreuzung zwischen 1 und 2: 111000 (oder 000111)

11) Unsicheres Schließen

a) Nennen Sie zwei verschiedene Arten von Unsicherheit und geben Sie Beispiele an.

Bayes'sche Interferenzregel: Monty's Hall-Rätsel

Fuzzy-Interferenz: Diagnosesysteme in der Medizin, Qualitätskontrolle, Automatikgetriebe

b) Was besagt die und wann gilt für Ereignisse die Additionsregel in der Wahrscheinlichkeitstheorie?

$p(A+B)=p(A)+p(B)-p(A \cap B)$

- c) **Drei Todeskandidaten warten auf ihre Hinrichtung. Allerdings ist bekannt, dass einer der drei begnadigt werden soll. Der Kandidat 1 fragt den Gefängniswärter: "Hör mal, kannst Du mir verraten, wer von den beiden anderen dran glauben muss? Einer ist auf jeden Fall an der Reihe, also verrätst Du mir kein großes Geheimnis." Der Wärter überlegt und sagt "Irgendwie hast Du recht. Also, Kandidat 2 ist fällig." Jetzt ist der erste Todeskandidat erleichtert, denn er denkt: "Bleiben zwei übrig, einer davon überlebt, also ist meine Überlebenschance von 1/3 auf 1/2 gestiegen." Was sind die tatsächlichen Überlebenschancen der drei Kandidaten wenn sonst nichts bekannt ist?**
- d) **Wie lautet die Bayes'sche Inferenzregel?**
- e) **Der HIV-Test liefert mit hoher Wahrscheinlichkeit die richtige Antwort (die "Sensitivität" des Tests, d.h. die Wahrscheinlichkeit, dass das Testergebnis positiv ist wenn die Testperson infiziert ist, ist 0.998; die "Spezifität" des Tests, d.h. die Wahrscheinlichkeit, dass das Testergebnis negativ ist wenn die Testperson nicht infiziert ist, ist 0.99). Die von Experten am häufigsten genannte Zahlen für Westeuropa besagen, dass rund ein Promille der sexuell aktiven Bevölkerung infiziert sind. Wäre es dann sinnvoll, einen Zwangs-HIV-Test für alle in Österreich lebenden Erwachsenen einzuführen? Nehmen wir an, ein zufällig ausgewählter Erwachsener testet positiv. Wie gross ist dann die Wahrscheinlichkeit, dass er oder sie trotzdem nicht mit HIV infiziert ist (schreiben Sie auch die verwendete Formel auf; falls kein Taschenrechner zur Hand, so schätzen Sie die Größenordnung des Ergebnisses ab)?**

$$p(A|B) = p(A) \cdot p(B|A) / p(B) = p(B|A) \cdot p(A) / (p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A))$$

- d) **Wie lautet die Bayes'sche Inferenzregel?**

$$p(A|B) = p(A) \cdot p(B|A) / p(B) = p(B|A) \cdot p(A) / (p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A))$$

- e) **Der HIV-Test liefert mit hoher Wahrscheinlichkeit die richtige Antwort (die "Sensitivität" des Tests, d.h. die Wahrscheinlichkeit, dass das Testergebnis positiv ist wenn die Testperson infiziert ist, ist 0.998; die "Spezifität" des Tests, d.h. die Wahrscheinlichkeit, dass das Testergebnis negativ ist wenn die Testperson nicht infiziert ist, ist 0.99). Die von Experten am häufigsten genannte Zahlen für Westeuropa besagen, dass rund ein Promille der sexuell aktiven Bevölkerung infiziert sind. Wäre es dann sinnvoll, einen Zwangs-HIV-Test für alle in Österreich lebenden Erwachsenen einzuführen? Nehmen wir an, ein zufällig ausgewählter Erwachsener testet positiv. Wie gross ist dann die Wahrscheinlichkeit, dass er oder sie trotzdem nicht mit HIV infiziert ist (schreiben Sie auch die verwendete Formel auf; falls kein Taschenrechner zur Hand, so schätzen Sie die Größenordnung des Ergebnisses ab)?**

$$p(TI|PI)=0,998, p(TN|PI)=0,002$$

$$p(TN|PN)=0,99, p(TI|PN)=0,01$$

$$p(PI)=0,001$$

$$p(PI|TI)=p(TI|PI) \cdot p(PI) / (p(TI|PI) \cdot p(PI) + p(TI|PN) \cdot p(PN))=0,091$$

$$p(\neg PI|TI)=1-p(PI|TI)=0,909$$

Eine flächendeckende Durchführung eines solchen Tests wäre daher nicht sinnvoll

- f) **Entwerfen Sie eine kleine Fuzzy-Regelbasis (ohne Angabe der Fuzzy Mengen; nur die linguistischen Regeln sind gefragt) für die Auftau-Automatik eines Mikrowellen-Herdes. Nehmen Sie dabei an, dass geeignete Sensoren zur Verfügung stehen.**
- gefriergut kalt+schwer => hohe leistung
 gefriergut mittel-kalt+schwer => hohe leistung
 gefriergut nicht-kalt+schwer => mittlere leistung
 gefriergut kalt+halbschwer => hohe leistung
 gefriergut mittel-kalt+halbschwer => mittlere leistung
 gefriergut nicht-kalt+halbschwer => mittlere leistung
 gefriergut kalt+leicht => mittlere leistung
 gefriergut mittel-kalt+leicht => niedrige leistung
 gefriergut nicht-kalt+leicht => niedrige leistung

12) Intelligente Agentensysteme

- a) **Nennen Sie die zwei Teilgebiete und zwei großen Fragestellungen der DAI!**

Distributed Problem Solving, Concurrent Multi-Agent Systems; die Frage, wie und von wem vorhandene Ressourcen aufgeteilt werden können und die Frage, wie vorhandene Aufgaben auf verschiedene Agenten aufgeteilt werden können.

- b) **Auf welchen drei Beschreibungsebenen kann ein wissensbasierter Agent beschrieben werden?**

Abstrakte Ebene, logische Ebene, Implementationsebene

- c) **Geben Sie drei Techniken an die in der Distributed AI verwendet werden.**

???

- d) **Nennen Sie drei Eigenschaften, die ein Agent aufweisen sollte um sich von einem "gewöhnlichen" Programm abzuheben!**
Autonomie, Sozialfähigkeit, Reaktionsfähigkeit, Aktionsfähigkeit
- e) **Beschreiben Sie den Unterschied zwischen reaktiven und deliberativen Agenten.**
Bei **deliberativen** Agenten ergibt sich ausgehend von einem wissensbasierten System eine deliberative Architektur. Der Reasoning-Prozess findet in einem Modell der Realität statt. **Reaktive** Agenten haben kein zentrales, symbolisches Modell der Realität, in dem agiert wird.
- f) **Erläutern Sie den Unterschied zwischen remote procedure calling und remote programming anhand eines kurzen Beispiels!**
RPC: zum Löschen von n Dateien werden $2n+1$ Nachrichten ausgetauscht; der Client kann Operationen auf dem Server ausführen.
RP: hier wird nur eine Nachricht versendet; der Client kann den Server veranlassen, bestimmte Aktionen durchzuführen.
- g) **Wie sieht das Schichtenlayer von KQML aus? Welche Gestalt hat ein KIF Satz?**
Schichten von KQML: Content Layer, Message Layer, Communication Layer
KIF-Satz ist ein prädikatenlogischer Ausdruck in Präfixform

13) Informationsagenten im Internet

- a) **Erläutern Sie folgende Begriffe: SGML, XML, HTML. Worin unterscheidet sich HTML von XML?**
SGML: Standard Generalized Markup Language: standardisierte Markupsprache für Beschreibung von Struktur und Inhalt maschinenlesbarer Daten
XML: Extended Markup Language: Teilmenge von SGML
HTML: Hypertext Markup Language: SGML-Applikation, Textauszeichnungssprache
HTML enthält Daten und Layoutinformationen, XML beliebige Daten in bestimmter Struktur
- b) **Welche Abfrage- und Transformationssprachen für XML kennen Sie? Führen Sie drei an.**
XPath, XQuery und XSL
- c) **Welche der im Vortrag erwähnten Sprachen eignen sich zur Beschreibung nichtmobiler Agenten? Welche davon sind Skriptsprachen?**
Skriptsprachen: Tcl/Tk, Perl, Telescript, ECMA Script (JavaScript)
Nicht-Skriptsprachen: Java, C/C++ etc.
- d) **Skizzieren Sie das Agency-Intelligence-Mobility-Diagramm und erläutern Sie kurz wo das MindIt-Tool darin anzusiedeln ist, bzw. der ContinualQuery Agent.**
???
- e) **Schreiben Sie drei der genannten Eigenschaften/Vorteile von Java auf.**
objektorientiert, maschinenunabhängig, für viele Plattformen verfügbar
- f) **Schreiben Sie eine XML Darstellung eines Bucheintrages die zu folgendem DTD Fragment konform ist:**
- <!ELEMENT book (author+, title, price)>**
 - <!ELEMENT title (#PCDATA)>**
 - <!ELEMENT author (#PCDATA | lastname | firstname | fullname)*>**
 - <!ELEMENT price (#PCDATA)>**
 - <!ATTLIST price currency CDATA "USD" source (list|regular|sale) list taxed CDATA #FIXED "yes">**
- ```
<?xml version="1.0" ?>
<book>
<author>bla</author>
<title>bla</title>
<price currency="USD" source="list">
```

```
taxed="yes">12345</price>
</book>
```

g) **Erläutern Sie die Begriffe Wrapper und Mediator.**

Ein **Mediator** ist ein Informationssystem, das die Daten mehrerer Informationsquellen durch virtuelle Integration zusammenfasst. Ein **Wrapper** extrahiert aus Webseiten Informationen, um sie einem Mediator zur Verfügung zu stellen.

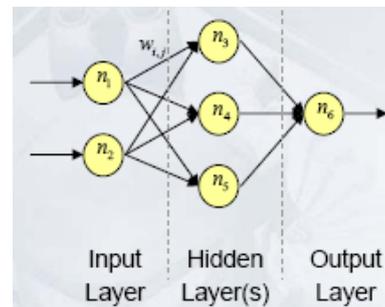
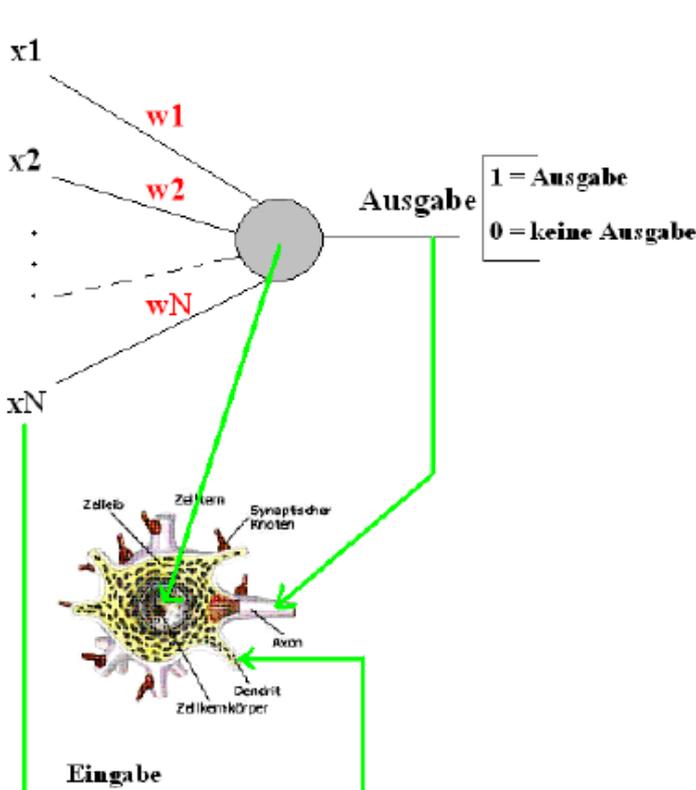
**14) Neuronale Netze**

a) **Welche Lernmethode gab der stagnierenden Forschung im Bereich der kNN ab 1986 neuen Schwung?**

**1986: Rumelhart** et al: Backpropagation als mathematisch elegante Lernmethode in n-Schichten Modellen --> Verbesserung der Lernalgorithmen, Anwendungen, ...

**Backpropagation** (englisch: Rückpropagierung) ist ein verbreitetes Verfahren für das Einlernen von künstlichen neuronalen Netzen. Formuliert wurde es zuerst 1974 durch Paul Werbos. Bekannt wurde es allerdings erst durch die Arbeit von David E. Rumelhart, Geoffrey E. Hinton und Ronald J. Williams ab 1986 und führte zu einer „Renaissance“ der Erforschung künstlicher neuronaler Netze.

b) **Aus welchen 3 Funktionen besteht ein künstliches Neuron? Nennen Sie auch die korrespondierenden 3 Bestandteile eines biologischen Neurons.**



**Inputfunktion:** Dendriten

$$net_j(o) = \sum_{i=1}^n w_{ij} \cdot o_{ij}$$

**Aktivierungsfunktion:** Zellkern (Zellkörper)

$$a_j = act_j(net_j(o))$$

**Outputfunktion:** Axon

$$o_j = f(a_j) = a_j$$

[http://www.intergenomics.de/literaturseminar/Referat\\_Folien\\_NeuronaleNetze.pdf](http://www.intergenomics.de/literaturseminar/Referat_Folien_NeuronaleNetze.pdf)

c) **Aus welchen 3 Schichten (Layers) sind kNNs normalerweise aufgebaut?**

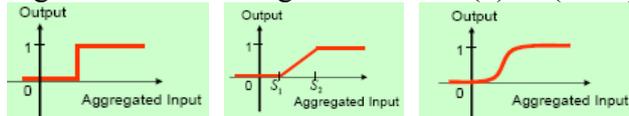
Input Layer, Hidden Layer(s), Output Layer

d) **Nennen Sie mind. 2 verschiedene Arten von Aktivierungsfunktionen.**

Threshold Logic Function:  $act(x) = 1$  (wenn  $net(x) \geq S$ ), sonst 0;

Threshold Linear Function:  $act(x) = 1$  (wenn  $net(x) \geq S$ ),  $-ax$  falls  $S_1 \leq net(x) \leq S_2$ , sonst 0;

Sigmoide Aktivierungsfunktion:  $\text{act}(x)=1/(1+e^{-x})$ ;  $\text{act}(x)=\tanh(x)$ ;  $\text{act}(x)=\sin(x)$



- e) **Beschreiben Sie kurz die wesentlichen Unterschiede zwischen Feedforward-Netzen und Recurrent-Netzen.**

**Feedforward-Netze:** ohne Rückkopplung; Es gibt keinen Pfad, der von einem Neuron direkt oder über andere Neuronen wieder zurückführt. Daten werden nur in eine Richtung weitergegeben.

**Recurrent-Netze:** mit Rückkopplung; direkt, indirekt, lateral, vollständig gebunden

- f) **Neben der Veränderung der Gewichte, welche prinzipiell möglichen Lernprozesse in kNN sind Ihnen noch bekannt? Nennen Sie mind. drei.**

Entwicklung neuer Verbindungen; Löschen existierender Verbindungen; Modifikation der Verbindungsstärke (Veränderung der Gewichte); Modifikation des Schwellenwertes; Modifikation der Aktivierungs- bzw. Ausgabefunktion; Entwicklung neuer Neuronen; Löschen bestehender Neuronen

- g) **Beschreiben Sie kurz (2 Sätze) die prinzipielle Funktionsweise des Backpropagation Algorithmus (Formeln sind nicht notwendig).**

Ein Eingabemuster wird angelegt und vorwärts durch das Netz propagiert. --> Die Ausgabe des Netzes wird mit der gewünschten Ausgabe verglichen. --> Die Differenz der beiden Werte wird als Fehler des Netzes erachtet --> Der Fehler wird nun wieder über die Ausgabe- zur Eingabeschicht zurück propagiert, dabei werden die Gewichtungen der Neuronenverbindungen abhängig von ihrem Einfluss auf den Fehler geändert. Dies garantiert bei einem erneuten Anlegen der Eingabe eine Annäherung an die gewünschte Ausgabe.

Dieser Fragenkatalog wurde von keinem der vortragenden Professoren kontrolliert und wurde rein aus den Vorlesungsfolien, Google-Searches, Wikipedia, usw. erstellt. Wir können also nicht für seine Richtigkeit garantieren. Falls jemand zu den nicht beantworteten Fragen noch Antworten findet, oder bestehende Antworten verbessern möchte, so ist er herzlich dazu eingeladen.

**!!! Viel Glück bei der Prüfung !!!**