

# **Zusammenschrift wichtiger Fakten – Mustererkennung, Prof. Schuster**

Autor: Marmel ([Marmel.vienna@gmx.at](mailto:Marmel.vienna@gmx.at))

Anm: Diese Fragen sind keine „offiziellen“ Prüfungsfragen von Prof. Schuster, sie gehen aber genauer auf Dinge ein, welche er im Laufe der Vorlesung besonders unterstrichen hat oder welche besonders wichtig erscheinen.

## **1) Was ist ein Muster?**

Ein Muster ist eine Anordnung aus einfachen Bestandteilen, die in ihrer Anordnung und in ihrer Beziehung einen bestimmten, anwendungsabhängigen Sinn ergeben. Muster folgen also gewissen Gesetzmäßigkeiten, haben Struktur.

## **2) Wie unterscheiden sich die Erkenntnisleistungen zwischen Computer und Mensch?**

Der Mensch assoziiert gut, der Computer kombiniert gut.

## **3) Was passiert bei der Clusteranalyse? Was ist hierbei besonders wichtig? Wie kann die Klassenbildung erfolgen?**

Eine Stichprobe aus  $M$  Objekten soll in Klassen unterteilt werden. Wichtig ist, dass ein Maß für die Distanz oder Ähnlichkeit zwischen den Objekten definiert wird. Man benötigt also eine Distanzfunktion für die Mengen und eine für die Objekte.

Die Klassenbildung erfolgt hierarchisch ( $M$  1elementige Klassen vereinigen sich langsam zu größeren Klassen, wobei immer die aneinander näher liegendsten vereint werden) oder nicht-hierarchisch (es wird von einem Element ausgegangen und alle Elemente in der Nähe Epsilon seiner Klasse hinzugefügt. Dasselbe wird mit jedem anderen Element in- und ausserhalb der Klasse durchgeführt).

## **4) Wie kann man bei komplexen Mustern, welche wie zB bei der Texterkennung aus Subelementen bestehen, die Erkennungsrate verbessern?**

Man setzt Rückkopplung ein, d.h. man fällt Entscheidungen erst so spät wie möglich und verwendet gegebenenfalls Information aus der vorigen Erkennungsstufe (zB bei Erkennung, um welches Wort es sich handelt, wird beim Scannen auch darauf eingegangen, zwischen welchen Buchstaben das Programm davor bei der Klassifikation geschwankt hat).

## **5) Welche 3 Ebenen der Bildbeschreibung gibt es und wie kann man deren Modelle definieren?**

Man unterscheidet Bildpunkte, Regionen und Beziehungen (der Bildpunkte oder Regionen). Das deklarative Modell besteht aus Bedingungen, wenn das Bild sie erfüllt wird es in die Klasse aufgenommen. Das prozedurale Modell generiert aus Produktionsregeln Bilder, die mit dem Bild verglichen werden. Ähnelt es einem oder mehreren dieser Bilder, wird es in die Klasse aufgenommen.

## **6) Welche Arten von Merkmalen differenziert man?**

Man unterscheidet photometrische Merkmale (mittlerer Grauwert), geometrische (Rundheit), topologische (Anzahl Löcher in Region), globale (Koeffizient der DFT) oder Texturmerkmale.

## **7) Wie lässt sich die Diskriminationsfähigkeit eines Merkmals darstellen und beurteilen?**

Die Diskriminationsfähigkeit bestimmt, wie gut ein Merkmal zwischen den Klassen unterscheiden lässt. Stellt man das Merkmal auf einem Histogramm dar (Abszisse: Merkmalsausprägung, Ordinate: Häufigkeit getrennt für jede Klasse) kann man sehen, wie

diskriminativ das Merkmal ist. Im Idealfall sind die beiden Klassen disjunkt (überlappen nicht). Je mehr sie überlappen, desto weniger diskriminativ ist das Merkmal. Oft führt die Kombination aus mehreren Merkmalen zu einer besseren Klassifikation.

### **8) Was für Pixelrelationen (Nachbarschaftsbeziehungen) gibt es?**

- $N_4(p)$ : Die Menge der 4 direkten Nachbarpixel heißt 4-Nachbarn von  $p$  und wird  $N_4(p)$  geschrieben

- $N_D(p)$ : Die Menge der 4 indirekten Nachbarpixel (die diagonal anliegenden Pixel) schreibt man  $N_D(p)$

- $N_8(p)$ : Direkte und indirekte Nachbarn zusammen nennt man die 8-Nachbarn von  $p$  und schreibt man  $N_8(p)$ .

Zusammenhängend sind die Pixel aber erst, wenn sie Nachbarn (je nach verwendeter Beziehung) sind und gleichzeitig gleiche Grauwerte besitzen. Im Allgemeinen definiert man zur Modellierung der Ähnlichkeit eine Menge  $V$  von Grauwerten, und nennt die Pixel ähnlich, wenn beide Grauwerte aus  $V$  besitzen. Zwei Pixel  $p$  und  $q$  mit Grauwerten aus  $V$  sind:

- 4-zusammenhängend, wenn  $q \in N_4(p)$

- 8-zusammenhängend, wenn  $q \in N_8(p)$

Man nennt dementsprechend eine Menge von Pixeln 4-Weg, wenn sie alle 4-zusammenhängend sind, und 4-Region, wenn es für alle Pixelpaare aus  $R$  einen 4-Weg gibt, der in ganz  $R$  liegt (Analog für 8-zusammenhängend).

Mittels des Connected Components Labeling lassen sich über zusammenhängende Pixel je nach Beziehung zusammenhängende Regionen markieren.

### **9) Was sind die Merkmale von Blobs?**

Die räumliche Verteilung der Bildpunkte in einem Blob wird charakterisiert durch Fläche, Schwerpunkt, Umfang, Euler Zahl (Anzahl Komponenten minus Anzahl Löcher), der zweite zentrale Moment (-> Berechnung der wichtigsten Parameter einer Ellipse), Bounding Box (kleinstmögliches Rechteck, das alle Punkte eines Blobs umschließt), konvexe Hülle (Durchschnitt aller konvexer [jede Gerade zw 2 Punkten liegt vollständig in der Region] Regionen des Blobs).

### **10) Was für Formeldescriptoren unterscheidet man?**

Es gibt:

- Formfaktor
- Rundheit
- Kompaktheit
- Aspect Ratio (Lage-Ratio)

Für konvexe Hülle auch:

- Solidity
- Convexity

### **11) Was ist ein lokales Merkmal?**

Ein Merkmal ist lokal, wenn sein Wert nur von einer kleinen Untermenge von Pixel abhängt. Um lokale Merkmale zu entdecken, kann man das Bild mit Schablonen des gewünschten Merkmals durchsuchen. (ist aber für die Detektion von Merkmalen nicht besonders spezifisch).

### **12) Was versteht man unter Segmentierung?**

Segmentierung ist die Zerlegung eines Bildes in eine Anzahl von (zusammenhängenden) Bildpunkt-Mengen, die zu einer Klasse jeweils ähnlicher Objekte gehören.

Allen Segmentierungsverfahren ist gemeinsam, dass jedem Pixel  $n$  Merkmale zugeordnet werden, aufgrund deren Eigenschaften dann versucht wird, den Pixel in eine Klasse einzuteilen.

### **13) Welche Segmentierungsverfahren gibt es?**

- Parallele Segmentierung: Die Unterteilung wird einzeln für jedes Pixel gemacht (-> Pixelbetrachtungsreihenfolge spielt keine Rolle)
- Sequentielle Segmentierung: Jede Entscheidung zur Einteilung hängt von zuvor getroffenen Entscheidungen ab.
- Homogene Subpopulation: Man trennt das Bild in  $k$  Gebiete, die in einer Kombination von Merkmalen homogen sind. -> Cluster-Analyse
- Morphologie: Eine Maske mit einer bestimmten Form wird über das Bild gelegt. Stimmen Maske und Teil des Bildes überein, wurde das Objekt gefunden.
- Segmentierung aufgrund von Ähnlichkeiten:
  - Bottom up: man geht von einem Punkt aus und verschmilzt solange Punkte, wie es die Homogenitätsbedingung zulässt (=Region growing)
  - Top-down: Man zerlegt das Bild solange, bis homogene Bereiche entstanden sind (split and merge)
- Schwelldwertverfahren: betrachtet nur ein Merkmal (zB Grauwert). Über das Histogramm kann man den Pixel dann nach Wert der Klasse mit nächstem Gipfel zuordnen. Dabei soll die Abweichung der Werte innerhalb der Klassen minimiert und die der Werte zwischen den Klassen maximiert werden. Jedem Grauwert werden Wahrscheinlichkeiten für die Zugehörigkeit zu einem Objekt zugeordnet. Hierbei muss man jeweils Eigenarten der Klassen berücksichtigen:
  - Große Flächen mit weichen Grenzen: Grenze so wählen, dass Pixel mit Werten ober- und unterhalb der Grenze minimiert
  - Wertebereiche von Objekt und Hintergrund überlappen: 2 Grenzen verwenden: Pixel, die nicht zwischen den Grenzen liegen, eindeutig klassifizieren, die anderen aufgrund ihrer Nachbarn.
  - Ungleiche Beleuchtung oder Sensorempfindlichkeit: Blöcke mit lokalen Schwellwerten bilden, dann räumlich über Schwellwerte interpolieren.
- Kantenfinder: versuchen Grenzen zwischen Objekten mittels Grauwertänderung zu finden, ebenfalls mit Grenzwert.

### **14) Was ist eine Textur?**

Textur ist die nicht-zufällige (regelmäßige) Anordnung von lokalen elementalen Grundmustern, die sich über eine Region wiederholen, die im Verhältnis zu den Grundmustern groß ist.

Man hat darin Texturprimitiva (kleinste Einheiten, welche aufgrund bestimmter Eigenschaften eine Einheit bilden) welche eine bestimmte räumliche Verteilung einnehmen (Anordnung, kann zufällig, funktionell, strukturell oder probabilistisch sein).

### **15) Was besagt die Hypothese von Julesz?**

Unterschiede in Bildern werden vom Menschen nur in der Statistik erster (=Grauwertverteilung) oder zweiter (=Verteilung der Grauwertpaare in gegebener räumlicher Entfernung) erkannt, Unterschiede höherer Ordnung werden nicht mehr wahrgenommen. Neuere Arbeiten weisen darauf hin, dass man das ganze auf die erste Ordnung + Textone (lokal gut sichtbare Merkmale) beschränken kann.

### **16) Welche Eigenschaften haben Primitiva?**

- Grobheit: räumliche Wiederholungsrate, je größer die Periode, desto größer die Struktur

- Kontrast: je größer die Grauwertskala, desto kontrastreicher
  - Richtung der Primitiva: Form und Platzierungsregel (gerichtet  $\Leftrightarrow$  ungerichtet, ausgerichtet  $\Leftrightarrow$  durcheinander)
  - Form der Primitiva: linienhaft oder fleckig
  - Gleichmäßigkeit: regelmäßig oder unregelmäßig
  - Rauheit: geschmeidig (rund) oder ungeschmeidig (zackig)
- Diese Charakteristika können auch im Leistungsspektrum erkannt werden.

### **17) Was ist der Unterschied zwischen schwacher und starker Textur?**

Schwache Texturen haben zwischen den Primitiva geringe räumliche Interaktion (man verwendet statistische Verfahren), bei starken Texturen ist die räumliche Interaktion zwischen Primitiven gleichmäßig (man verwendet strukturelle Verfahren).

### **18) Was sind die Aufgaben der Texturanalyse?**

- Texturerkennung: feststellen, ob eine Textur einem Bezugsmuster entspricht
- Texturklassifikation: Textur wird einer von K Texturklassen zugeordnet.
- Textursegmentierung: Grenzen zwischen Regionen unterschiedlicher Textur finden.

### **19) Was ist die Statistik erster Ordnung und was sind ihre Merkmale?**

Die Statistik erster Ordnung betrachtet die einzelnen Grauwerte und ihre Verteilung (Histogramm). Unterschiede in der Struktur können so jedoch nicht beschrieben werden. Dennoch gibt es Bilder, die sich sehr gut durch die Statistik erster Ordnung alleine charakterisieren lassen. Merkmale sind:

- Mittlerer Grauwert
- Variation der Grauwerte (Spannwerte, mittl Abweichung, Varianz)
- Schiefe der Grauwert-Verteilung
- Exzeß
- Relative Extrema Density
- Textural Edgeness (Kantenfrequenz) -> über Kantendetektion
- Grauwert-differenzen (Mittelwert, Kontrast)
- Angular Second Moment
- Entropie

### **20) Was ist die Statistik zweiter Ordnung?**

Die Statistik zweiter Ordnung beschreibt die räumlichen Beziehungen zwischen Grauwerten.

### **21) Was ist die Co-occurrence matrix? Welche Merkmale ergeben sich daraus?**

Die 2-dimensionale Co-Occurrence Matrix ist eine quadratische  $Q \times Q$  Matrix, welche die räumliche Anordnung und die räumliche Abhängigkeit der Grauwerte beschreibt, wobei Q die Zahl der Quantisierungsstufen ist. Da eine Texturbeschreibung mittels der Co-occurrence-Matrix viel zu umständlich und redundant wäre, extrahiert man daraus nur bestimmte Merkmale:

- Energie (= Homogenität, Größe des Grauwertbereichs)
- Kontrast (= lokale Bildvariation)
- inverses Differenzmoment
- Entropie (negatives Informationsgehalt)
- Korrelation (Bildlinearität)
- und viele mehr...

### **22) Textursegmentierung: Was sind Vor- und Nachteile von Flächen-basiertem und Grenzen-basiertem Ansatz?**

Beim Flächenbasierten Ansatz erhält man oft zusammenhängende Flächen und somit geschlossene Grenzen. Allerdings ist es oft notwendig, die Anzahl der Texturen im Voraus zu kennen und Grenzen festzulegen, um die Texturen zu unterscheiden.

Vorteil des grenzenbasieren Ansatzes ist es, dass man diese Anzahl nicht kennen muss, dass Grenzen dafür aber unterbrochen sein können. Das Ganze geschieht kantenorientiert.

Optimal ist eine Kombination beider Ansätze.

### **23) Wovon hängt die Güte eines Merkmals ab?**

Seine Güte hängt davon ab, wie schnell/einfach es berechnet werden kann und wie diskriminativ es ist (wie gut es zwischen interessierenden Klassen unterscheidet). Wenn es nicht in erster Dimension diskriminativ ist, dann möglichst in zweiter zB gemeinsam mit einem anderen Merkmal.

### **24) Was macht man, wenn die Merkmale statistisch unabhängig sind?**

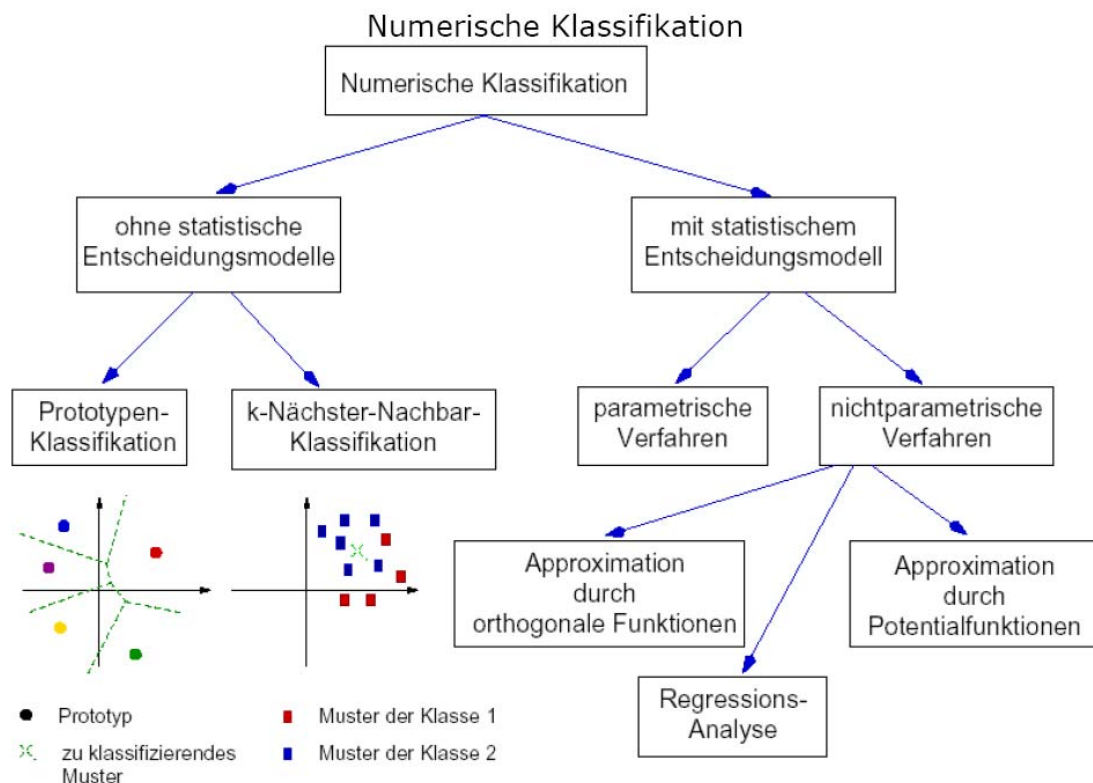
In diesem Idealfall genügt es, die Merkmale nach ihrer Güte zu ordnen und nur die besten zu nehmen.

### **25) Was macht man, wenn die Merkmale nicht statistisch unabhängig sind?**

In diesem Fall kann man (bei sehr kleinen Werten) alle Merkmalskombinationen bewerten, oder über Vorwärts- (man geht vom Merkmal mit max Gütewert aus und nimmt nach und nach die Werte, die in Kombination den besten Gütewert haben) oder Rückwärtsselektion (man lässt nach und nach die schlechtesten Werte wegfallen).

Eine weitere Möglichkeit ist, die Karhunen-Loeve-Transformation anzuwenden. Hierbei transformiert man den Merkmalsraum so, dass die einzelnen Merkmale unkorreliert sind und wählt dann die wichtigsten Merkmale aus. Die Kovarianzmatrix beschreibt hierbei die Korrelationen zwischen den Merkmalen.

### **26) Was sind die Arten von Verfahren der numerischen Klassifikation?**



**27) Was sind die Unterschiede mit oder ohne statistischem Entscheidungsmodell?**

Große Ausnahmen werden vom Wahrscheinlichkeitsmodell kaum erkannt, wenn die Wahrscheinlichkeiten stimmen ist es aber besser.

Bei Änderungen der Wahrscheinlichkeiten ist das statistische Entscheidungsmodell schlechter.

**28) Was ist das Problem der Kurvenapproximation?**

Wenn man zu genau interpoliert, ist dies schlecht, denn der Algorithmus verliert seine Verallgemeinerungsfähigkeit. Deshalb muss man zwischen Lern- und Teststichprobe differenzieren.

**29) Was sind die Ziele überwachten/unüberwachten Lernens?**

Das überwachte Lernen hat als Ziel, eine Zuordnungsvorschrift zu finden, die unbekannte Muster möglichst oft auf die richtige Klasse abbildet.

Das unüberwachte Lernen soll die Muster geeignet aufteilen und damit ähnliche Muster zusammenfassen (Clustering).

**30) Welche Methoden des überwachten Lernens gibt es?**

Bei der Reklassifikation ist die Lernstichprobe mit der Teststichprobe identisch. Beim Jackknife-Test werden in mehreren Versuchen bis zu 20% der Stichprobe als Teststichprobe verwendet und der Rest als Lernstichprobe.

**31) Was macht man dagegen, dass die numerischen Merkmale unterschiedliche Einheiten haben können?**

Man skaliert den Merkmalsraum. Dies kann mittels der Spannweite oder der Standardabweichung geschehen. Vorteil der Skalierung mit der Standardabweichung ist, dass alles so robuster gegenüber Ausreißern ist.

**32) Wie groß muss die Lernstichprobe sein, wenn man n Merkmale hat?**

Sie muss eine Größe von  $2^n$  haben

**33) Wieviele Grenzflächen braucht man, um K Klassen zu trennen? Was ist die Alternative dazu?**

Man benötigt  $K*(K-1)/2$  Grenzflächen. Als Alternative kann man K Diskriminanzfunktionen verwenden.

**34) Wie berechnet man in der statistischen Klassifikation den optimalen Klassifikator?**

Der optimale Klassifikator berechnet zu jeder Klasse  $\omega_k$  die **Prüfgröße**

$$u_k(\vec{x}) = \sum_{j=1}^K r_{kj} \times p(\omega_j) \times p(\vec{x} | \omega_j) \quad \forall k=1, \dots, K$$

K.....Anzahl der Klassen

$r_{kj}$  ..... Kosten

$p(\omega_j)$ .....(a-priori) Wahrscheinlichkeit der Klasse  $\omega_j$

$p(\vec{x} | \omega_j)$ ..... bedingte Wahrscheinlichkeit des Merkmalsvektors  $\vec{x}$

**35) Wie wählt man die Kostenfunktion beim Bayes'schen Klassifikator?**

$R(k_j)$  ist 0 wenn die Klassen ident sind und 1 wenn die Klassen verschieden sind. Damit wird das Muster der Klasse zugeordnet, bei der die a-priori Wahrscheinlichkeit multipliziert mit der a-posteriori Wahrscheinlichkeit den größten Wert hat.

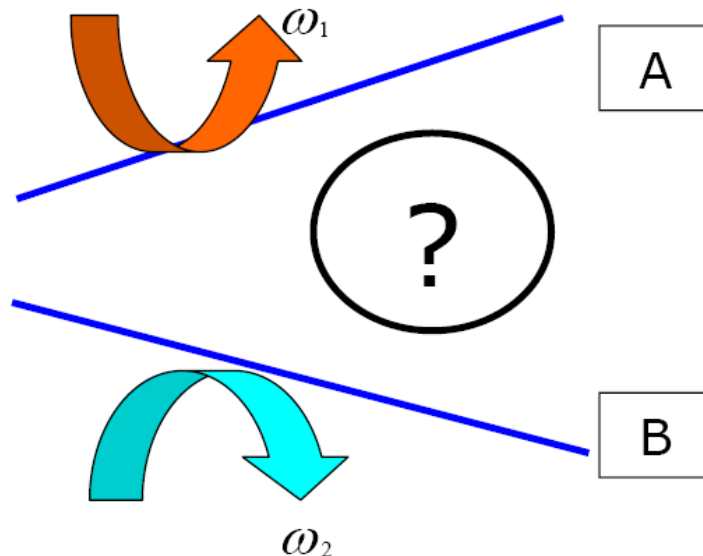
### **36) Wie verläuft der sequentielle Wahrscheinlichkeitstest?**

Erhebe Merkmal n

Berechnung der **Quotienten**  $\tau_n = \frac{p(\vec{x}_n | \omega_1)}{p(\vec{x}_n | \omega_2)}$

n.....Anzahl der aktuell erhobenen Merkmale

**Vergleich** dieses Quotienten  $\tau_n$  mit 2 Intervallgrenzen A und B.



- Wenn  $\tau_n \geq A \Rightarrow$  Muster  $\rightarrow$  Klasse  $\omega_1$
- Wenn  $\tau_n \leq B \Rightarrow$  Muster  $\rightarrow$  Klasse  $\omega_2$
- Wenn  $B < \tau_n < A$  muss ein weiteres Merkmal hinzugenommen werden

### **37) Was ist der Unterschied zwischen parametrischer und nichtparametrischer Klassifikation?**

• Parametrische Klassifikation:

Der Typ der Wahrscheinlichkeitsdichteverteilungen ist vorgegeben und es müssen lediglich die Parameter dieser Verteilung aus der Lernstichprobe geschätzt werden.

• Nichtparametrische Klassifikation:

Der Typ der Wahrscheinlichkeitsdichteverteilungen ist nicht vorgegeben, sondern muss ebenfalls geschätzt werden.

### **38) Wieso nimmt man bei parametrischer Klassifikation die Normalverteilung an?**

Nach dem Razor-Prinzip (Prinzip des unzureichenden Grundes) nimmt man immer die Methode, die einfach ist und ein Phänomen am ehesten erklären kann, solange nichts massiv gegen diese Methode spricht.

### **39) Wie funktioniert die Prototypen Klassifikation?**

Es gibt ein „Prototypenmuster“. Die weiteren Muster werden je nach Abstand zu dem Prototypen der entsprechenden Klasse dem nächsten Prototypen zugeordnet.

#### **40) Wie funktioniert die Nächster Nachbar Klassifikation?**

Bei dieser Klassifikation wird ein Objekt der Klasse zugeordnet, der sein nächster Nachbar im Merkmalsraum angehört.

#### **41) Was sind Musterprimitiva?**

Musterprimitiva sind die Grundbausteine eines Musters, also zB Striche aus denen Buchstaben zusammengesetzt sind oder Phoneme der gesprochenen Sprache.

#### **42) Wieso eignen sich besonders reguläre und kontextfreie Sprachen für die syntaktische Klassifikation?**

Reguläre Sprachen können durch endliche Automaten und kontextfreie Sprachen durch Keller-Automaten oder Parsing-Algorithmen erkannt werden.

#### **43) Welche Typen von Grammatiken gibt es und welche Form haben ihre Regeln? Welche davon verwendet man in der syntaktischen Klassifikation?**

- Typ 0 (Chomsky-Grammatik):  $\alpha \rightarrow \beta$
- Typ 1 (kontextsensitive Grammatik):  $\alpha_1 A \alpha_2 \rightarrow \alpha_1 \beta \alpha_2$
- beschränkte Grammatik:  $a \rightarrow b$  (mit Länge  $a \leq \text{Länge } b$ )
- Typ 2 (kontextfreie Grammatik):  $A \rightarrow \beta$
- Typ 3 (reguläre Grammatik):  $A \rightarrow \alpha B$  oder  $A \rightarrow \alpha$

Für die Mustererkennung sind nur die Grammatiken G verwendbar, für die **entscheidbar** ist, ob eine Symbolkette Element der von einer Grammatik erzeugten Sprache ist oder nicht. Für die Sprachen, die von einer Typ-1 (kontextsensitive), Typ-2 (kontextfreie) oder Typ-3 (reguläre) Grammatik erzeugt werden, ist dieses Problem entscheidbar.

Am häufigsten verwendet:

Grammatiken vom Typ-2 (kontextfrei) oder Typ-3 (regulär), weil für diese relativ einfache Analysealgorithmen zur Verfügung stehen.

#### **44) Was ist der Vorteil der Realsierung der syntaktischen Klassifikation durch endliche Automaten?**

Die Laufzeit: Sie beträgt bei einer Eingabelänge n maximal n Zustände, damit ist die Laufzeit nicht von der Anzahl Endzustände abhängig (die in der Medizin bei vielen Diagnosen gewaltig sein kann).

#### **45) Was für Grundtypen von Fehlern gibt es und wie kann man sie behandeln?**

3 Grundtypen von Fehlern:

- Ersetzung eines Symbols durch ein anderes  $\alpha a \beta \rightarrow \alpha b \beta$
- Einfügen eines Symbols  $\alpha \beta \rightarrow \alpha a \beta$
- Auslassen eines Symbols  $\alpha a \beta \rightarrow \alpha \beta$

Methoden zur Fehlerkorrektur:

- Man erweitert die Musterbeschreibungssprache, indem man alle fehlerhaften Symbolketten einbezieht. Die Erkennungsalgorithmen ändern sich nicht.

Nachteil:

Durch die Erweiterung mit allen Fehlermöglichkeiten kann die Sprache so umfangreich werden, dass dies zu einer ineffizienten Analyse führen kann.

- Die Sprache wird für fehlerfreie Grundsymbole konstruiert, die Erkennungsalgorithmen werden aber modifiziert. Man verwendet dabei ein Maß der Übereinstimmung der fehlerhaften Kette mit der korrekten.

zB: Wenn man die einzelnen Substitutionswahrscheinlichkeiten kennt, kann man zu einer gestörten Symbolkette jene finden, die am wahrscheinlichsten in die fehlerhafte Symbolkette transformiert wird.