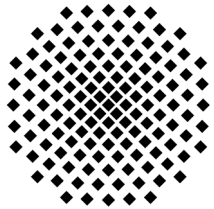


Automatische Verb-Klassifikation unter Einbezug linguistischer Kriterien und semantischer Assoziationsdaten



Sabine Schulte im Walde
Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart


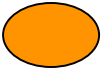

Heidelberg, 23. November 2007

Überblick

1. **Klassifikation**, allgemein und in der CL
2. **Semantische Verbklassen** als Fallbeispiel
3. **Linguistische Kriterien** bei der Evaluierung
4. **Assoziationsdaten** zu Verb-Eigenschaften

Klassifikation

Klassifikation

- Objekte → Klassen (Prozess und Ergebnis)
- Objekte in **gemeinsamen** Klassen:
so **ähnlich** wie möglich
- Objekte in **unterschiedlichen** Klassen:
so **verschieden** wie möglich
- Ziele:
 - » Übersicht über Objekte, Erkennen von Datenstrukturen
 - » Vergleichbarkeit von Objekten
 - » Einordnung von einzelnen Objekten (Klassierung)
 - » Generalisierung über Objekte {    }

Klassifikationsbeispiele

- topographische Karten
- Größe/Haarfarbe von Menschen
- Beurteilung eines Kinofilms

- Biologie: Systematik von Pflanzen/Tieren
- Medizin: Symptomklassifikation → Krankheiten erkennen
- Soziologie: Rollenverhalten in Gruppen

- Wortartenzuweisung (Part-of-Speech Tagging)
- PP-Anbindung (PP-Attachment)
- semantische (Verb-/Nomen-/Adjektiv-/etc.) Klassen

Klassifikation in der Computerlinguistik

- Vielzahl von Klassifikationsproblemen
- Algorithmen der Künstlichen Intelligenz
- Schwierigkeit: **Auswahl eines Klassifikationsalgorithmus**
und **Bestimmung der Parameter**
- Schwierigkeitsgrad abhängig von theoretischer Definition der Aufgabe (Bsp.: Wortarten vs. semantische Klassen)
- Orientierung: **linguistische und kognitive Plausibilität**

Semantische Verbklassen

Semantische Verbklassen

- Verb-Klassifikation auf Basis semantischer Eigenschaften
- Unterschiedliche Arten von semantischen Klassen gemäß der semantischen Eigenschaften
- Beispiele:
 - » Synonymie (WordNet/GermaNet; Kunze 2000):
bekommen, empfangen, erhalten, erlangen, kriegen
 - » Alternationsverhalten (Levin 1993):
buy, catch, earn, find, steal, ...

Semantische Verbklassen als Fallstudie

- Warum das Interesse an semantischen Verbklassen?
- Vielfältige theoretische Definitionen:
 - » Ziel-orientierte Gültigkeit
 - » Theorie vs. Anwendung
- Manuelle vs. automatische Klassifikation
- Ehrgeizige Aufgabe: Lernen von Semantik
- Übertragbarkeit von Entscheidungen auf ähnliche Fälle, z.B. katalanische Adjektive (Boleda 2007)

Semantische Verbklassen: Anwendung

- **Organisation von Verben**, z.B. für Lexikographie
- **Generalisierung über Verben**

→ Ungenügende Daten (Data Sparseness)

→ Anwendungen (Beispiele):

Word Sense Disambiguation

(Dorr & Jones 1996; Kohomban & Lee 2005)

Maschinelle Übersetzung

(Prescher et al. 2000; Koehn & Hoang 2007)

Dokumentenklassifikation (Klavans & Kan 1998)

Akquisition von Subkategorisierung (Korhonen 2002)

Automatische Klassifikation

1. Subkategorisierung als Verb-Eigenschaften
2. Klassifikation der Verben mit k -Means
3. Evaluierung gegen einen Gold Standard

Eigenschaften

- Semantische Eigenschaften für semantische Klassen
- Mögliche Quellen: manuelle Ressourcen vs. **Korpora**
- Potentiell saliente Merkmale für Verb-Modellierung, **Syntax-Semantik-Schnittstelle (Bedeutung ↔ Verhalten)**:
 - » syntaktische Subkategorisierung
 - » Präpositionalphrasen
 - » Argumente: Lexeme oder Generalisierungen
 - » Adverbien etc.

Eigenschaften

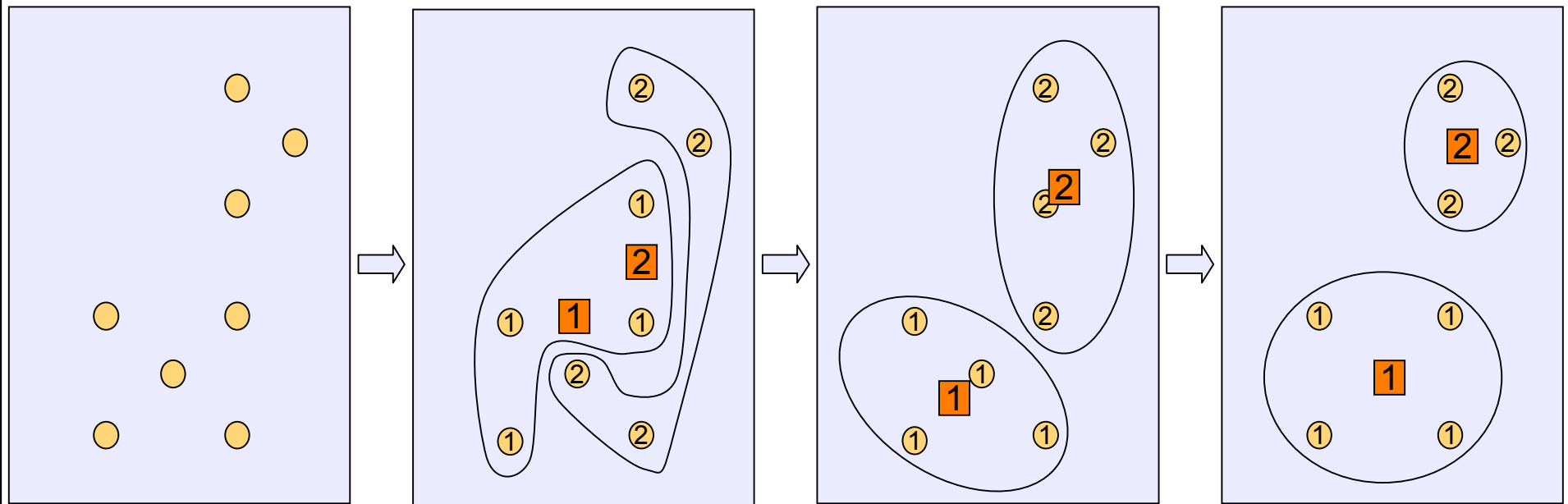
- Quelle: trainierte statistische deutsche Grammatik
- Subkategorisierungs-Rahmen
- GermaNet Selektionspräferenzen
- Beispiel: *glauben*
- Freq → Prob

Subkat-Rahmen	Frequenz
NPnom S-dass <i>Er glaubt, dass ...</i>	1,929
NPnom S-2 <i>Sie glaubt, sie sei ...</i>	1,888
NPnom PP-an <i>Er glaubt an ...</i>	687
NPnom <i>Sie glauben.</i>	608
NPnom NPakk <i>Sie glaubt die Geschichte.</i>	555
...	

Algorithmus: k -Means

- k -Means Algorithmus (Forgy, 1965)
- Unsupervised Hard Clustering (Clustering vs. Klassifikation)
- Iterative Reorganisation von Cluster-Zugehörigkeit:
 1. Anfangs-Klassierung jedes Objekts
 2. Berechnung der Cluster-Schwerpunkte
 3. Ermitteln des nächsten Clusters (Schwerpunkt)
 4. Reorganisation der Cluster-Zugehörigkeit

Algorithmus: k -Means



Ähnlichkeitsmaße

- **Minkowski-Metrik / L_q -Norm:** $L_q(x, y) = \sqrt[q]{\sum_{i=1}^n (x_i - y_i)^q}$

- **Kosinus:** $\cos(x, y) = \frac{\sum_{i=1}^n x_i * y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$

- **Kullback-Leibler-Distanz:** $D(x \parallel y) = \sum_{i=1}^n x_i * \log \frac{x_i}{y_i}$

(geglättete Variante: **Skew Divergence** (Lee 2001))

Klassifikationsergebnis

bilden, erhöhen, festlegen, senken,
steigern, vergrößern, verkleinern

eilen, gleiten, kriechen,
rennen, starren

- Passen Algorithmus und seine Parameter zur Aufgabe?
(Initialisierung, Anzahl der Cluster, Ähnlichkeit) *Ja.*
- **Wie können die Ergebnisse evaluiert werden?**
→ linguistische Kriterien
- **Entsprechen die Verb-Eigenschaften der Aufgabe?**
→ ja, aber ... teilweise widersprüchlich
→ menschliche Urteile zur Verbesserung

Evaluierung

Evaluierung

Wie gut ist das Klassifikationsergebnis?

- **Precision, Recall, F-Score etc.** ← Gold Standard
(Schulte im Walde 2006)
- **Menschliche Urteile** und **Annotatoren-Übereinstimmung**
(Boleda, Schulte im Walde, Badia; to appear)
- Integration in eine **Anwendung**

GS-Evaluierungsmaße

- **Theoretische Statistik**
 - » Rand Index (Rand 1971)
 - » Matching Index (Fowlkes & Mallows 1983)
 - » Rand Index, chance-adjusted (Hubert & Arabie 1985)
- **Clustering von Webseiten**
 - » Mutual Information (Strehl et al. 2000)
- **Koreferenz-Auflösung**
 - » Class-based Precision/Recall (Vilain et al. 1995)
- **Semantische Klassen**
 - » Pair-wise Precision/Recall → Pair-wise F-Score
(Hatzivassiloglou & McKeown 1993; Schulte im Walde & Brew 2002)

Linguistische Kriterien

13 linguistische Kriterien für Wahl des Evaluierungsmaßes:

- Kein Bias (Vorliebe) bei der Anzahl der Cluster.
- Größe und (Un-)Korrektheit der Cluster:

$C_a = \{\text{ahnen, vermuten, wissen}\}$

😊 $C_b = \{\text{ahnen, vermuten, wissen, glauben, denken}\}$

😊 $C_a = \{\text{ahnen, vermuten, wissen}\}$

$C_c = \{\text{ahnen, vermuten, wissen, laufen, lachen}\} \quad [\dots]$

→ Auswahl von drei Evaluierungsmaßen

**Assoziationen →
Verb-Eigenschaften**

Assoziationen → Verb-Eigenschaften

- **Basis:** semantische Assoziationen, also spontane Äußerungen, zu deutschen Verben
- **Annahmen:**
 - » Assoziationen spiegeln linguistische und konzeptuelle Eigenschaften wider, modellieren Bedeutungsaspekte
 - » Leitfaden für Auswahl von Eigenschaften

Web Experiment: Durchführung

klagen

jammern

laut

Trauer

schlecht

Richter

Web Experiment: Datenbeispiel

- Verb: *klagen*
- Die 10 häufigsten Assoziationen:
- Freq → Prob

Assoziation	Freq
<i>Gericht</i>	19
<i>jammern</i>	18
<i>weinen</i>	13
<i>Anwalt</i>	11
<i>Richter</i>	9
<i>Klage</i>	7
<i>Leid</i>	6
<i>Trauer</i>	6
<i>Klagemauer</i>	5
<i>laut</i>	5

Assoziations-basierte Verbklassen

Klasse	Eigenschaften: Assoziationen
bedauern, heulen, jammern, klagen, verzweifeln, weinen	Trauer, weinen, traurig, Tränen, jammern, Angst, Mitleid, Schmerz, Träne, Leid etc.
abnehmen, abspecken, zunehmen	Diät, Gewicht, dick, abnehmen, Waage, Essen, essen, Sport, dünn, Fett etc.

- Assoziationen sind geeignete Verb-Eigenschaften.
- Wie gut können Assoziationen durch Korpus-basierte Eigenschaften (synt. Funktionen, Kookkurrenz etc.) modelliert werden?

Ergebnisse

- **verbesserte Auswahl von Verb-Eigenschaften:**
 - » Ausgesuchte Grammatikfunktionen überragen frühere Rahmen-basierte Verb-Eigenschaften.
 - » Fenster-basierte Eigenschaften (Kookkurrenz) sind ähnlich geeignet wie Grammatikfunktionen.
 - » Adverbien eignen sich als Verb-Eigenschaften (schlechter als Nomen, aber besser als Rahmen).

Zusammenfassung

Zusammenfassung

- Klassifikationen in der Computerlinguistik:
 - » vielfältige Anwendungsmöglichkeiten
 - » Schwierigkeit bei Auswahl von Algorithmus und Definition von Parametern
- Fallstudie als Szenario für **Einsatz linguistischen Wissens** (→ Evaluierung) und **semantischer Assoziationen** (→ Verb-Eigenschaften)
- Status von Klassifikationen in der CL (Maschinelles Lernen & Linguistik, Evaluierung)