

Ein Ansatz zum Verstehen natürlicher, gesprochener Sprache durch hierarchisch strukturierte Hidden-Markov-Modelle

Johannes Müller und Holger Stahl
(in alphabetischer Reihenfolge)

Lehrstuhl für Mensch-Maschine-Kommunikation,
Technische Universität München, Arcisstraße 21, D-80333 München,
email: {mue, sta}@mmk.e-technik.tu-muenchen.de

Zusammenfassung. Es wird ein Ansatz zum Verstehen natürlicher, gesprochener Sprache auf der Basis stochastischer Wissensbasen vorgestellt: Das semantische Modell generiert innerhalb eines vorgegebenen Wirkungsfeldes mögliche semantische Gliederungen, welche wortnahe Darstellungen des Bedeutungsinhalts sind. Das syntaktische Modell erzeugt aus einer gegebenen semantischen Gliederung mittels hierarchisch strukturierten Hidden-Markov-Modellen zutreffende Wortketten. Beide Modelle können in ein sprachverstehendes System integriert werden, das in einem 'Top-Down'-Ansatz auf diese stochastischen Wissensbasen zurückgreift. **Schlüsselwörter.** Spracherkennung, Sprachverstehen, Hidden-Markov-Modell, natürlichsprachlicher Mensch-Maschine-Dialog

Abstract. In this paper, an approach for understanding natural speech by means of stochastic knowledge bases is presented: Within a given domain, the semantic model generates possible semantic structures, which are semantic representations close to the word level. Corresponding to such a semantic structure, the syntactic model generates word chains using hierarchical Hidden-Markov-Models. Integrated into a speech understanding system, these stochastic knowledge bases can be utilized for a 'top-down'-approach.

Keywords. speech recognition, language understanding, Hidden-Markov-Model, spoken man-machine-dialogue

1. 'Top-Down'-Ansatz zum Verstehen natürlicher Sprache

Das Verstehen von Sprache kann aufgefaßt werden als Abbildung einer Folge von Merkmalsvektoren O [4] auf den Bedeutungsinhalt S (in diesem Artikel dargestellt durch die sog. semantische Gliederung) der zugrunde liegenden Äußerung. Für die gegebene Beobachtungsfolge O muß also aus der Menge möglicher S dasjenige gefunden werden, welches am wahrscheinlichsten ist. Die zu maximierende a-posteriori-Wahrscheinlichkeit $P(S|O)$ läßt sich mit dem Satz von Bayes umformen:

$$P(S|O) = \frac{P(S, O)}{P(O)} = \frac{P(O|S) \cdot P(S)}{P(O)} \quad (1.1)$$

Die Wahrscheinlichkeiten $P(O|S)$ sowie $P(S)$ sollen mittels stochastischer Methoden bestimmt werden [2] [5] [6]. Aufgrund der hohen Dimensionalitäten von S und O können diese Wahrscheinlichkeiten jedoch nicht direkt aus einer Lernstichprobe geschätzt werden, so daß zusätzliche Repräsentationsebenen, wie die Wortebene W , eingeführt werden müssen. $P(S|O)$ ergibt sich damit zu:

$$P(S|O) = \sum_W \frac{P(O|W) \cdot P(W|S) \cdot P(S)}{P(O)} \quad (1.2)$$

$P(O)$ muß bei der Maximierung nicht berücksichtigt werden, da es bei gegebener Beobachtungsfolge O konstant ist.

$$S_E = \operatorname{argmax}_S P(S|O) = \operatorname{argmax}_S \sum_W [P(O|W) \cdot P(W|S) \cdot P(S)] \quad (1.3)$$

Da für die Ermittlung von S_E nur der wahrscheinlichste Pfad interessiert (Erkennung mit Viterbi-Algorithmus [9]), entartet die Summe zum Maximumoperator:

$$S_E = \operatorname{argmax}_S \max_W [P(O|W) \cdot P(W|S) \cdot P(S)] = \operatorname{argmax}_S \max_W P(O, W, S) \quad (1.4)$$

Abb. 1.1 zeigt die wichtigsten Module des 'Top-Down'-Erkenners, wobei innerhalb jedes dieser Module noch weitere Repräsentationsebenen (z.B. Phonemebene) eingefügt sein können.

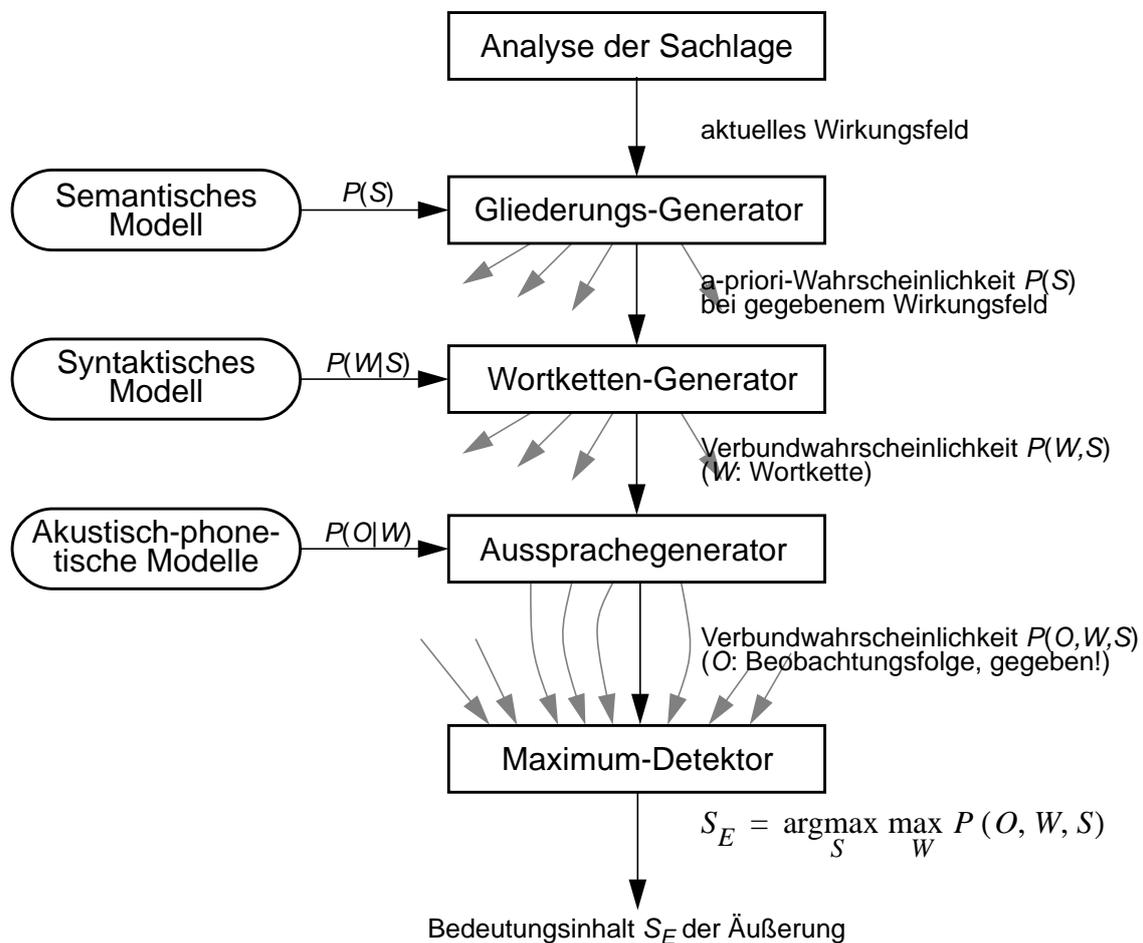


Abb. 1.1: 'Top-Down'-Hierarchie eines sprachverstehenden Systems

- **Semantisches Modell.** Wissensbasis zur Bestimmung von $P(S)$, der a-priori-Wahrscheinlichkeit für den Bedeutungsinhalt S innerhalb eines bestimmten Wirkungsfeldes [3].
- **Syntaktisches Modell.** Wissensbasis zur Bestimmung von $P(W|S)$, der bedingten Wahrscheinlichkeit für eine Wortfolge W bei gegebenem Bedeutungsinhalt S der Äußerung.
- **Akustisch-phonetische Modelle.** Wissensbasen zur Bestimmung von $P(O|W)$, der bedingten Wahrscheinlichkeit für eine Beobachtungsfolge O bei gegebener Wortfolge W [8].

2. Semantische Gliederung

Da es für den Bedeutungsinhalt möglicher Äußerungen eine unbegrenzte Anzahl denkbarer Fälle gibt, erscheint es zweckmäßig, zur Darstellung des Bedeutungsinhaltes die semantische Gliederung S einzuführen. Diese setzt sich aus kleineren bedeutungstragenden Einheiten, welche im folgenden semantische Untereinheiten (oder kurz *Semune*) genannt werden, zusammen.

2.1 Definition der semantischen Gliederung

Vorgeschlagen wird die Verknüpfung mehrerer Semune zu einer baumartigen und somit hierarchischen Struktur, bei der jeweils Semune höherer Hierarchiestufen durch Semune tieferer Hierarchiestufen näher spezifiziert werden. Abb. 2.1 zeigt als Beispiel den Baum einer solchen semantischen Gliederung S , bestehend aus $N=8$ Semunen. Dabei sind die Knoten die Semune s_n . Das Semun der höchsten Hierarchiestufe wird mit s_1 bezeichnet, die übrigen sind wahlfrei durchnummeriert. Die Kanten bilden die Verknüpfungen zwischen den Semunen.

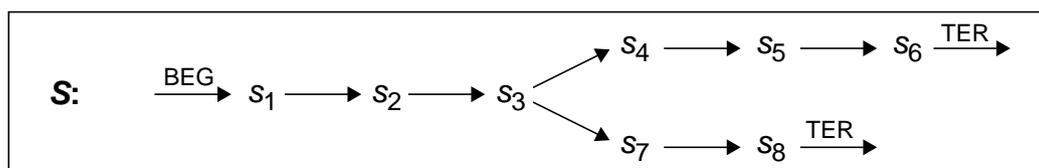
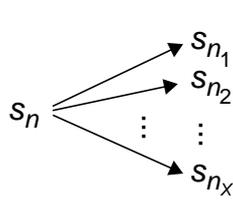


Abb. 2.1: Verknüpfung mehrerer Semune zum Baum einer semantischen Gliederung S

Ein bestimmtes Semun kann X direkte Nachfolger haben, verknüpft durch die Kante \longrightarrow . $\xrightarrow{\text{BEG}}$ bezeichnet die Kante zum Semun s_1 der höchsten Hierarchiestufe, $\xrightarrow{\text{TER}}$ eine terminale Kante. Ein terminaler Knoten ist hierbei nicht zulässig. In obigem Fall ist $X=2$ für s_3 , sowie $X=1$ für alle anderen Semune (auch terminale Kanten zählen). Die Verknüpfungen sind folgendermaßen zu deuten:



Die Semune $s_{n_1}, s_{n_2}, \dots, s_{n_x}$ und deren eventuelle Nachfolger stehen in einer durch das Semun s_n festgelegten Beziehung zueinander und beschreiben miteinander den Teil der semantischen Gliederung, welcher aus dem Semun s_n mit dessen eventuellen Vorgängern besteht.

Ein einzelnes Semun kann im weitesten Sinne als X -stellige prädikatenlogische Relationskonstante [1] angesehen werden. Die Verknüpfung einzelner Semune zur semantischen Gliederung unterscheidet sich jedoch wesentlich von der Darstellung des Bedeutungsinhaltes mittels Prädikatenlogik. Wenn auch mathematisch nicht so exakt, bietet die semantische Gliederung folgende wichtige Vorteile:

- Die semantische Gliederung ist eine wortnahe Darstellung des Bedeutungsinhaltes. Da jedem Semun s_n aus S genau ein bedeutungstragendes Wort w aus der Wortkette W zugeordnet werden kann, ist der Entwurf des syntaktischen Modells (Bestimmung von $P(W|S)$) ohne weitere Repräsentationsebenen möglich.
- Zur Verknüpfung semantischer Untereinheiten gibt es nur einen Mechanismus, nämlich die Kennzeichnung weiterer Semune als sog. Nachfolger. Beim Entwurf der Modelle zur Berechnung der Wahrscheinlichkeiten $P(S)$ und $P(W|S)$ muß daher nur diese eine mögliche Art der Verknüpfung betrachtet werden.

2.2 Definition eines einzelnen Semuns

Ein Semun trägt die semantische Information eines bedeutungstragenden Wortes. Um die eigentliche Bedeutung eines Semuns getrennt von der Menge möglicher Nachfolger modellieren zu können, wird jedes Semun s_n dargestellt durch seinen Typ t_n und seinen Wert v_n :

- Der Typ $t_n \in \{\tau_1, \dots, \tau_p, \dots\}$ gibt die Anzahl X der Nachfolger fest vor und schränkt die Menge möglicher Typen t_{n_1}, \dots, t_{n_X} dieser Nachfolge-Semune ein. Außerdem trifft er eine sinnvolle Auswahl möglicher ihm zuzuordnender Werte. (Für erste Untersuchungen werden nur Semune mit $X=1$ oder $X=2$ betrachtet.)
- Der Wert $v_n \in \{\varphi_1, \dots, \varphi_j, \dots\}$ gibt die eigentliche Bedeutung des Wortes aus W an, welches dem Semun s_n zugeordnet ist.

2.3 Beispiele zum Verständnis

Zur Verdeutlichung der Ausführungen in Kap. 2.1 und Kap. 2.2 soll im folgenden zu zwei verschiedenen Äußerungen, gegeben als Wortkette W , die jeweilige semantische Gliederung S gebildet werden. Die zu verstehenden Äußerungen sollen dabei Befehle darstellen, mit denen ein einfacher Grafikeditor zum Erzeugen dreidimensionaler Objekte auf dem Bildschirm bedient werden kann.

Mögliche Äußerungen (geschrieben jeweils als Wortkette) wären beispielsweise:

$W_1 =$ 'lösche alle grünen quader'

$W_2 =$ 'zeichne bitte zwei kugeln und mache doch den roten kegel klein'

Es muß ein Satz möglicher Semuntypen und -werte existieren, mit dem der Bedeutungsinhalt einer beliebigen Äußerung innerhalb des Wirkungsfeldes als semantische Gliederung dargestellt werden kann. Nachstehende Tabelle listet mögliche Semune für das Beispiel des Grafikeditors auf. (Die Tabelle erhebt keinen Anspruch auf Vollständigkeit der Semune, die zur Darstellung aller möglichen Äußerungen nötig wären. Sie soll lediglich verdeutlichen, wie bereits mit wenigen Typen und Werten der Bedeutungsinhalt sehr vieler Äußerungen abgedeckt wird.)

Tabelle 2.1: Beispiele für mögliche Semuntypen und -werte, mit denen Äußerungen zur Bedienung eines Grafikeditors als semantische Gliederungen dargestellt werden können.

Typ τ_i	Werte φ_i	Typen möglicher Nachfolger	X	Erläuterung
befehl1	erzeugen, löschen	form, logik	1	Aktion mit 1 Nachfolger
befehl2	ändern	1.: form, logik 2.: farbe, gröÙe, lage, (terminal)	2	Aktion mit 2 Nachfolgern
form	kugel, zylinder, kegel, quader	anzahl, logik, gröÙe, lage, farbe, (terminal)	1	Form der Objekte
anzahl	1, 2, ..., viele, alle	gröÙe, lage, farbe, (terminal)	1	Anzahl der Objekte
gröÙe	groÙ, klein, normal	lage, farbe, (terminal)	1	GröÙe der Objekte
lage	mitte, unten, oben, rechts, links	farbe, (terminal)	1	Position der Objekte
farbe	rot, grün, blau, gelb	(terminal)	1	Farbe der Objekte
logik	und, oder	(Beide Nachfolge-Typen müssen gleich oder 'befehl1' und 'befehl2' sein. Ein Nachfolge-Typ kann 'logik' sein.)	2	logische Verknüpfung

Zur Vereinheitlichung der Wahrscheinlichkeitsrechnung wird in die Menge der existierenden Typen $\{\tau_1, \dots, \tau_i, \dots\}$ auch der Nachfolge-Typ (terminal) aufgenommen. Dieser Typ definiert kein Semun, ihm wird kein Wert zugeordnet. Mit den in Tab. 2.1 vorgeschlagenen Semunen lassen sich die semantischen Gliederungen S_1 und S_2 zu den Wortketten W_1 und W_2 angeben. Man beachte, daß die in der Wortkette W_2 vorhandenen, bedeutungslosen Füllwörter wie 'bitte' und 'doch' nicht in der semantischen Gliederung S_2 repräsentiert sind.

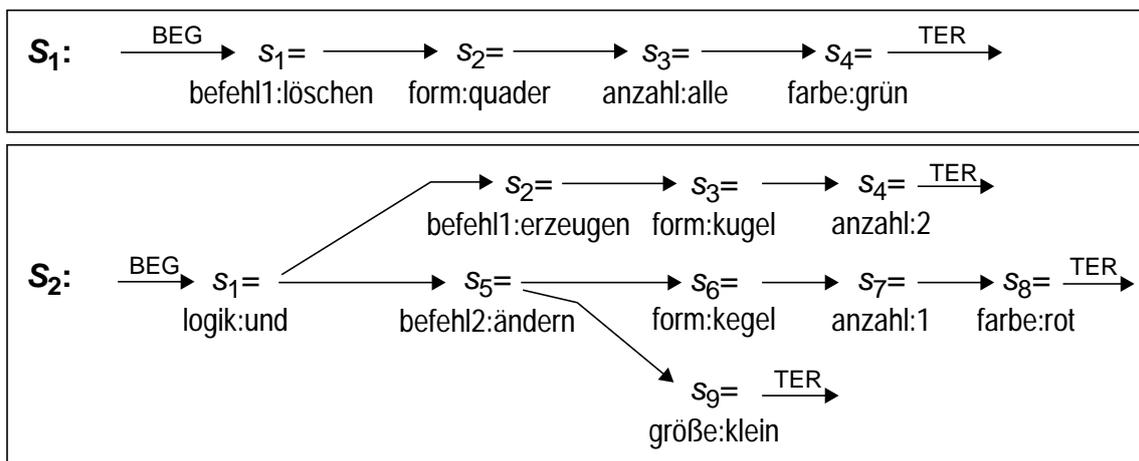


Abb. 2.2: Semantische Gliederungen S_1 und S_2 zu den Wortketten W_1 und W_2 , wobei jedes Semun s_n durch ein konkretes Typ-Wert-Paar $t_n : v_n$ dargestellt ist.

3. Semantisches Modell und Gliederungs-Generator

Der Gliederungs-Generator (vgl. Abb. 1.1) hat die Aufgabe, alle möglichen semantischen Gliederungen S mit den zugehörigen a-priori-Wahrscheinlichkeiten $P(S)$ zu erzeugen, die zu einem bestimmten Wirkungsfeld [3] denkbar und sinnvoll sind.

3.1 Wahrscheinlichkeiten im semantischen Modell

Die dem Gliederungs-Generator zugrundeliegende Wissensbasis ist das semantische Modell, welches aus Trainingsdaten (das sind viele vorliegende semantische Gliederungen aus einem bestimmten Wirkungsfeld) erstellt werden muß. Die Vielfalt von S ist jedoch zu groß, um $P(S)$ direkt aus einem Trainingskorpus zu schätzen, denn es ist nicht anzunehmen, daß alle denkbaren und sinnvollen semantischen Gliederungen auch im Training gesehen werden. Das semantische Modell muß eine endliche Anzahl Parameter enthalten, die sich einerseits zuverlässig aus begrenztem Trainingsmaterial bestimmen lassen, die andererseits jedoch auf die jeweilige a-priori-Wahrscheinlichkeit $P(S)$ einer unbegrenzten Menge semantischer Gliederungen S schließen lassen. Vorgeschlagen wird daher die Abschätzung folgender bedingter Wahrscheinlichkeiten 1. Ordnung:

- Die *Beginnwahrscheinlichkeit* $p_{\text{BEG}}(\tau_i)$ gibt an, mit welcher Wahrscheinlichkeit sich in der höchsten Hierarchiestufe ein Semun des Typs τ_i befindet:

$$p_{\text{BEG}}(\tau_i) = P(\text{Typ} = \tau_i \mid \text{in höchster Hierarchiestufe}) \quad (3.1)$$

- Die *Wertwahrscheinlichkeit* $p_{\tau_i}(\varphi_j)$ gibt an, mit welcher Wahrscheinlichkeit ein Semun des Typs τ_i den Wert φ_j annimmt:

$$p_{\tau_i}(\varphi_j) = P(\text{Wert} = \varphi_j \mid \text{Typ} = \tau_i) \quad (3.2)$$

- Die *Folgewahrscheinlichkeit* gibt an, mit welcher Wahrscheinlichkeit ein Semun des Typs τ_i die Typen $\tau_{i_1}, \dots, \tau_{i_X}$ als Nachfolger hat:

$$p_{\tau_i}(\tau_{i_1}, \dots, \tau_{i_X}) = P(\text{Nachfolge-Typen} = \tau_{i_1}, \dots, \tau_{i_X} \mid \text{Typ} = \tau_i) \quad (3.3)$$

3.2 Bestimmung von $P(S)$

Die Bestimmung der a-priori-Wahrscheinlichkeit $P(S)$ sei anhand des unten dargestellten Ausschnitts des Baumes von S erläutert.

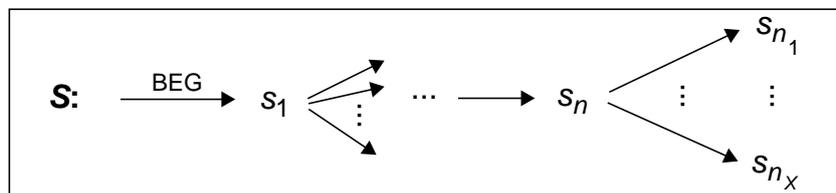


Abb. 3.1: Ausschnitt einer semantischen Gliederung S

- Die *Beginnkante* zum Semun s_1 wird mit der Beginnwahrscheinlichkeit

$$f_{\text{BEG}} = p_{\text{BEG}}(t_1), \quad (3.4)$$

- jeder *Knoten*, d.h. jedes Semun s_n , wird mit der Wertwahrscheinlichkeit

$$e_n = p_{t_n}(v_n) \quad (3.5)$$

- und *alle Kanten* (auch terminale Kanten), welche von einem Semun s_n ausgehen, werden mit der Folgewahrscheinlichkeit f_n beaufschlagt:

$$f_n = p_{t_n} \left(t_{n_1}, \dots, t_{n_X} \right) \quad (3.6)$$

Unter der Annahme der statistischen Unabhängigkeit aller Terme in Gl. (3.4)-(3.6) ergibt sich die a-priori-Wahrscheinlichkeit $P(S)$ als gemeinsames Produkt der Beginnwahrscheinlichkeit f_{BEG} mit allen Folge- und Wertwahrscheinlichkeiten der N Semune, die die semantische Gliederung S enthält:

$$P(S) = f_{\text{BEG}} \cdot \prod_{n=1}^N (e_n \cdot f_n) \quad (3.7)$$

Obwohl die Annahme der statistischen Unabhängigkeit in der Praxis nicht gegeben ist, rechtfertigt die einfache Modellierbarkeit die Anwendung von Gl. (3.7). Beliebiger genau könnte $P(S)$ bestimmt werden, wenn nicht bedingte Wahrscheinlichkeiten 1. Ordnung, sondern k -ter Ordnung mit $k \rightarrow \infty$ verwendet würden. Die Anzahl der zu trainierenden Parameter nimmt allerdings mit steigender Ordnung k exponentiell zu, was einerseits die Trainierbarkeit erschweren würde, andererseits bei einer späteren 'Top-Down'-Decodierung den zu bearbeitenden Hypothesen-Suchraum viel zu stark anwachsen ließe.

4. Syntaktisches Modell und Wortketten-Generator

Mit Hilfe des syntaktischen Modelles werden alle zu einer semantischen Gliederung S möglichen Wortketten W mit den bedingten Wahrscheinlichkeiten $P(W|S)$ erzeugt. Im wesentlichen werden folgende Anforderungen an das Modell gestellt:

1. Erzeugen genau eines bedeutungstragenden Wortes w pro Semun s_n .
2. Einstreuen zusätzlicher, bedeutungsloser Füllwörter.
3. Zeitliche Anordnung dieser Wörter zu Wortketten.

Es wird ein stochastischer Ansatz vorgeschlagen, bei dem zu jeder semantischen Gliederung S ein spezielles syntaktisches Modell λ_s vorliegt, so daß gilt:

$$P(W|\lambda_s) = P(W|S) \quad (4.1)$$

Das syntaktische Modell stellt ein Hidden-Markov-Modell dar [7], wobei zur Wortkettengenerierung die stochastischen Prozesse Zustandswechsel und Emission von Worten aus bestimmten Zuständen des Modelles überlagert werden.

Die Wahrscheinlichkeit für das Durchlaufen der Zustandsfolge Q im Modell λ_s ergibt sich als Produkt über alle Übergangswahrscheinlichkeiten entlang Q :

$$P(Q|\lambda_s) = \prod_{\substack{\text{alle Zustands-} \\ \text{übergänge} \\ \text{des Pfades } Q}} (\text{jeweilige Übergangswahrscheinlichkeit}) \quad (4.2)$$

Die Wahrscheinlichkeit, daß entlang der Zustandsfolge Q die Wortkette W emittiert wurde, ergibt sich als Produkt aller Emissionswahrscheinlichkeiten entlang des Pfades Q :

$$P(W|Q, \lambda_S) = \prod_{\substack{\text{alle Zustände} \\ \text{entlang } Q \text{ mit} \\ \text{Wortemissionen}}} (\text{jeweilige Emissionswahrscheinlichkeit}) \quad (4.3)$$

Unter der Annahme, daß die beiden Prozesse statistisch unabhängig voneinander sind, ergibt sich die Verbundwahrscheinlichkeit für das Durchlaufen eines ganz bestimmten Pfades Q bei gleichzeitiger Emission der Wortfolge W zu

$$P(W, Q|\lambda_S) = P(W|Q, \lambda_S) \cdot P(Q|\lambda_S) . \quad (4.4)$$

Die gesuchte Wahrscheinlichkeit $P(W|\lambda_S)$ ist die Summe obiger Verbundwahrscheinlichkeiten über alle Pfade Q :

$$P(W|\lambda_S) = \sum_{\text{alle } Q} P(W, Q|\lambda_S) \quad (4.5)$$

Für die unbegrenzte Anzahl möglicher semantischer Gliederungen gibt es jedoch eine ebenfalls unbegrenzte Anzahl von syntaktischen Modellen. Daher ist es nötig, das syntaktische Modell in kleinere Einheiten zu unterteilen, deren Parameter aus begrenztem Trainingsmaterial (das sind mögliche Wortketten zu vielen semantischen Gliederungen) geschätzt werden können. Diese kleineren Einheiten werden im folgenden als elementare Hidden-Markov-Modelle (EHMME) bezeichnet. Ausgehend von der Struktur einer semantischen Gliederung S ist es naheliegend, für jedes Semun s_n ein solches EHMM in das syntaktische Modell einzugliedern, wobei die einzelnen EHMME hierarchisch verknüpft werden.

4.1 Parameter eines elementaren Hidden-Markov-Modells

Abhängig von der Anzahl X der Nachfolger des zugeordneten Semuns werden EHMME mit $X+2$ Zuständen Z_1, \dots, Z_{X+2} gebildet.

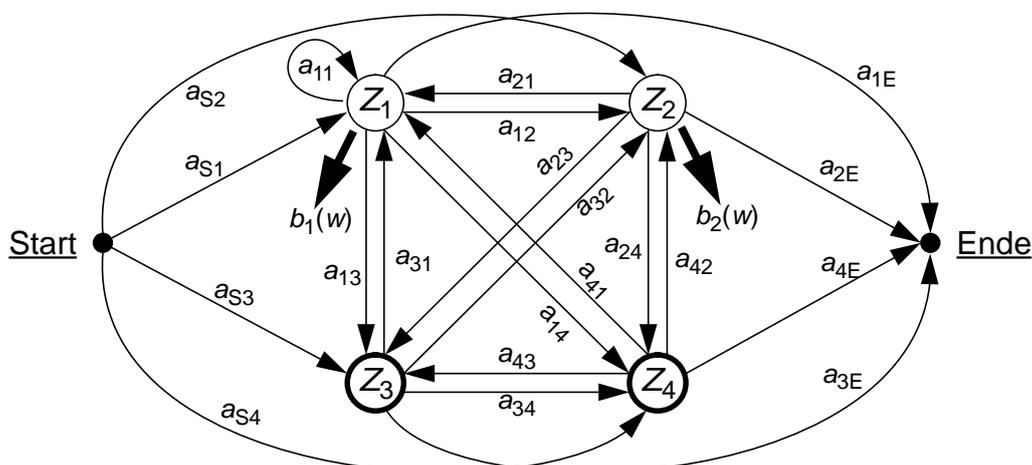


Abb. 4.1: Elementares Hidden-Markov-Modell mit vier Zuständen

Prinzipiell stellt das EHMM ein ergodisches Hidden-Markov-Modell 1. Ordnung dar, allerdings mit bestimmten Auflagen:

- Der Zustand Z_1 emittiert ein bedeutungsloses Wort.
- Der Zustand Z_2 emittiert ein bedeutungstragendes Wort, das den Bedeutungsinhalt des zugeordneten Semuns repräsentiert. Er wird genau einmal durchlaufen.
- Die weiteren Zustände symbolisieren die EHMMe der nachfolgenden Semune. Diese Sub-Modelle werden jeweils genau einmal durchlaufen.

Ein EHMM wird durch folgende Parameter beschrieben:

- Die *Übergangswahrscheinlichkeit* a_{ij} ist diejenige Wahrscheinlichkeit, mit der vom aktuellen Zustand Z_i in den nächsten Zustand Z_j übergegangen wird:

$$a_{ij} = P(\text{nächster Zustand} = Z_j | \text{aktueller Zustand} = Z_i) \quad (4.6)$$

Übergänge in die Zustände Z_3 und Z_4 stellen dabei Übergänge zum Start des jeweiligen Sub-Modelles dar. Übergänge aus den Zuständen Z_3 und Z_4 stellen Übergänge vom Ende des jeweiligen Sub-Modelles dar.

- Die *Emissionswahrscheinlichkeit* $b_i(w)$ ist diejenige Wahrscheinlichkeit, mit der von einem Zustand Z_i das Wort w emittiert wird:

$$b_i(w) = P(\text{emittiertes Wort} = w | \text{aktueller Zustand} = Z_i), \quad i \in \{1, 2\} \quad (4.7)$$

Um die Anzahl der zu schätzenden Modellparameter zu begrenzen, wird angenommen, daß mit Ausnahme der Emissionswahrscheinlichkeit $b_2(w)$ alle Parameter nur vom Typ des zugeordneten Semuns abhängen. Es sind somit nur so viele verschiedene EHMM-Parametersätze nötig, wie es mögliche Semun-Typen τ_i gibt.

4.2 Verknüpfung mehrerer EHMMe zum syntaktischen Modell

Der strukturelle Aufbau des syntaktischen Modelles λ_S aus einzelnen EHMMen ist angelehnt an die Struktur der semantischen Gliederung S , die aus einzelnen Semunen s_n besteht. λ_S enthält also für jeden Knoten (d.h. jedes Semun) im Baum

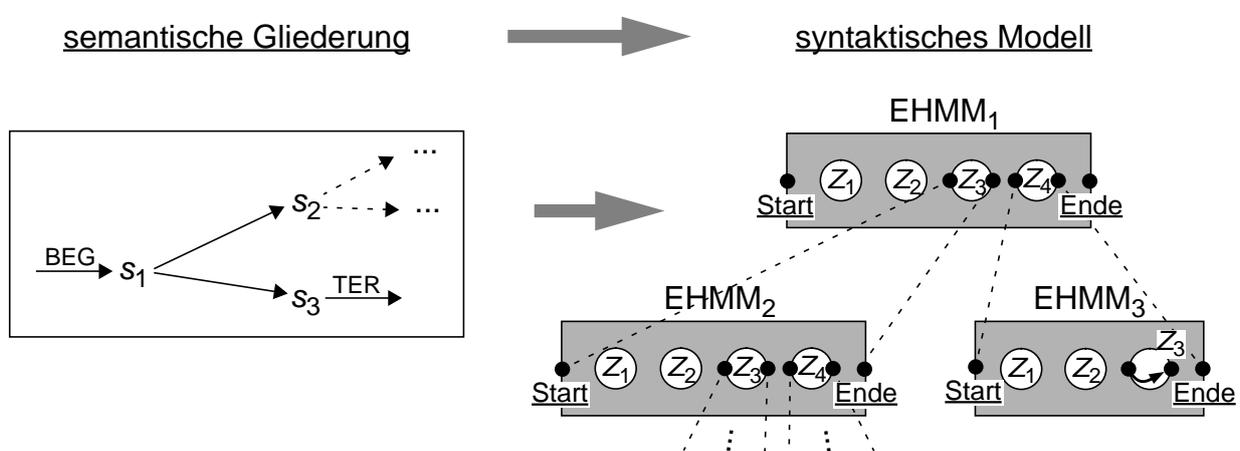


Abb. 4.2: Aufbau des syntaktischen Modelles ausgehend von der semantischen Gliederung

von S genau ein EHMM $_n$. Jede Kante im Baum von S bewirkt eine Verknüpfung der entsprechenden EHMMe. Abb. 4.2 zeigt links einen Ausschnitt der semantischen Gliederung, rechts dargestellt ist der entsprechende Teil des syntaktischen Modelles.

4.3 Wortemissionen

- Die Emissionen im Zustand Z_1 sind bedeutungslose Wörter. Die Emissionswahrscheinlichkeit $b_1(w)$ ist nur abhängig vom Semun-Typ t_n .
- Die Emissionen im Zustand Z_2 sind bedeutungstragende Wörter. Die Emissionswahrscheinlichkeit $b_2(w)$ ist abhängig vom Typ t_n und vom Wert v_n .

Ein Beispiel für die möglichen Emissionen aus den Zuständen Z_1 und Z_2 zum Typ befehl1 (aus Tab. 2.1) zeigt folgende Abbildung:

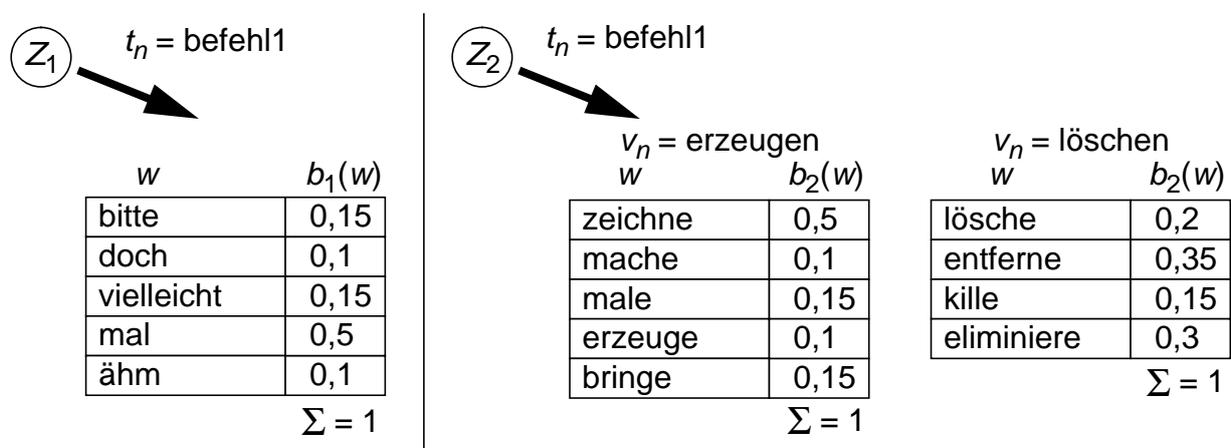


Abb. 4.3: Emissionen bedeutungsloser und bedeutungstragender Wörter mit $b_i(w)$

5. Literatur

- [1] G. Görz: *Einführung in die künstliche Intelligenz*, Addison-Wesley, 1993
- [2] H. Höge: *Statistische Modelle für die Spracherkennung*, Tagungsband Fortschritte der Akustik - DAGA 1993 (Frankfurt am Main, Deutschland), S. 11-30
- [3] M. Lang: *Mensch-Maschine-Kommunikation 1*, Vorlesungsskript, Technische Universität München, 1994
- [4] M. Lang, H. Stahl: *Spracherkennung für einen ergonomischen Mensch-Maschine-Dialog*, mikroelektronik, Bd. 8 (1994), H. 2, S. 79-82
- [5] R. Pieraccini et al.: *A Speech Understanding System Based on Statistical Representation of Semantics*, Proc. ICASSP 1992 (San Francisco, USA), S. I 193-I 196
- [6] R. Pieraccini, E. Levin, E. Vidal: *Learning how to Understand Language*, Proc. Eurospeech 1993 (Berlin, Deutschland), S. 1407-1412
- [7] L. R. Rabiner: *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*, Proc. IEEE, Bd. 77 (1989), H. 2, S. 257-286
- [8] H. Roll, H. Stahl, J. Müller: *Training von Hidden-Markov-Modellen zur Erkennung fließender Sprache*, Diplomarbeit, Technische Universität München, 1994
- [9] A.J. Viterbi: *Error Bounds for Convolutional Codes and an Asymptotical Optimal Decoding Algorithm*, IEEE Trans. Information Theory, Bd. 61 (1973), S. 268-278