

SEMANTISCHE DIALOGMODELLIERUNG MIT GEWICHTETEN MERKMAL-WERTE-RELATIONEN

Markus Huber¹, Christian Kölbl¹, Robert Lorenz¹, Ronald Römer², Günther Wirsching³

Universität Augsburg,
Harman/Becker Automotive Systems GmbH Ulm,
Katholische Universität Eichstätt-Ingolstadt

guenther.wirsching@ku-eichstaett.de

Kurzfassung: Das Hauptergebnis dieser Arbeit ist die Definition von Datenstrukturen für die Schnittstelle zwischen Erkennung und Dialogsteuerung, namentlich den Verstehenshorizont und den Ergebnishorizont als Beispiele stochastischer Merkmal-Werte-Relationen. Außerdem wird ein Algorithmus beschrieben, mit dem sich derart strukturierte Daten automatisch aus einem Entity-Relationship-Modell und einer dazu passenden Relationalen Datenbank generieren lassen. Darüber hinaus wird ein Verfahren vorgestellt, mit dem man die bei der Erkennung berechneten Wahrscheinlichkeiten in Form von Konfidenzen und Wahrscheinlichkeitsmaßen für die Dialogsteuerung verfügbar machen kann.

1 Einleitung und Hintergrund

Dialogsysteme sind erforderlich, um Information zwischen Mensch und Maschine zu übertragen. Prinzipiell kann diese Informationsübertragung in verschiedenen Modi erfolgen: haptisch (z.B. vom Menschen zur Maschine über Tastatur oder Mausclicks), visuell (z.B. von der Maschine zum Menschen via Bildschirm oder umgekehrt via Kamera) oder akustisch über Töne und Mikrofone. Bei einer visuellen oder akustischen Informationsübertragung vom Menschen zur Maschine muss die Maschine Algorithmen zur Mustererkennung enthalten und können nur das erkennen, was schon beim Programmieren für möglich gehalten wurde – und auch das in der Regel nicht mit 100%-iger Sicherheit. In diesem Aufsatz konzentrieren wir uns auf Sprachdialogsysteme, was jedoch nicht ausschließt, dass die hier entwickelte Informationsmodellierung nicht auch für multimodale Dialogsysteme anwendbar wäre.

Sprachdialogsysteme werden z.B. in folgenden Bereichen eingesetzt:

- zum Verteilen ankommender Anrufe in einem Call-Center,
- im KFZ zur Bedienung des Infotainment-Systems sowie des Navigationsgeräts.

Die in diesen Beispielen eingesetzten Dialogsysteme folgen im Wesentlichen einer starren, hierarchisch aufgebauten Menüstruktur. Der automatische Spracherkennung muss in den meisten einzelnen Dialogzuständen nur eine eng begrenzte Anzahl möglicher Äußerungen unterscheiden können, was die Erkenntrate einigermaßen hochhält. In anderen Dialogzuständen kann durchaus eine große Anzahl von Worten unterschieden werden, z.B. zur sprachlichen Eingabe eines Navigationsziels, wobei prinzipiell einige zigtausend Ortsnamen verstanden werden können müssen. Aber beim Versuch, auch in einer weiteren sprachlichen Umgebung eingebettete Ortsnamen verstehen zu können, stoßen marktübliche Dialogsysteme schnell an ihre Grenzen.

Die hier vorgeschlagene Vorgehensweise zum Aufbau wesentlich flexiblerer und intuitiverer Sprachdialogsysteme beruht auf einer Modellierung unsicherer Information durch gewichtete Merkmal-Werte-Relationen, die sich mathematisch als zyklonfreie gerichtete Graphen mit gewissen Gewichtsfunktionen darstellen lassen. Die graphentheoretische Struktur ist dabei durch Untersuchungen zu Semantik und Pragmatik natürlicher Sprache motiviert, während die Gewichtsfunktionen die Unsicherheit modellieren.

Charles W. Morris [5] folgend unterscheidet man in der Semiotik zwischen Syntaktik, Semantik und Pragmatik. In Bezug auf gesprochene Sprache beinhaltet dabei die Syntaktik die Regeln zur Aneinanderreihung sprachlicher Zeichen, die Semantik die Verbindung der Zeichen zum Gemeinten und die Pragmatik des Kommunikationsprozesses zeigt sich in der Wirkung auf den Interpretanten. Bei Dialogsystemen besteht das Ziel im Austausch von Informationen zwischen den Dialogteilnehmern, die sich vorab darüber verständigt haben, welchen Ausschnitt von Welt man mit welchen Begriffen¹ dargestellt hat. In Computern verarbeitete Information wird mathematisch greifbar durch die grundlegenden Arbeiten von Codd [2] zur relationalen Algebra und Chen [1] zum Entity-Relationship-Modell (ER-Modell). Vor diesem Hintergrund bestehen die Ziele eines Dialogsystems einerseits darin, Einträge einer Datenbank zu identifizieren, und andererseits darin, die Art der Manipulation wie Lesen, Modifizieren, neu Anlegen, oder Löschen, festzulegen und auszuführen. Die Verbindung zwischen sprachlichen Zeichen und Gemeintem (Semantik) wird durch drei Komponenten realisiert:

1. Im ER-Modell sind *Entitätstypen* durch *Beziehungstypen* verknüpft, wobei jeweils optional *Attribute* sowie *Rollenamen* für an Beziehungstypen beteiligten Entitätstypen vorhanden sein können.
2. Ein im folgenden zu beschreibender Algorithmus erlaubt es, aus einem ER-Modell eine Merkmal-Werte-Relation zu generieren.
3. Lokale Grammatiken assoziieren gewisse sprachliche Ausdrücke mit Merkmalen und Werten aus dieser Merkmal-Werte-Relation.

Schließlich wird die syntaktische Struktur möglicher Äußerungen durch auf lokalen Grammatiken basierende Sprachmodelle beschrieben.

Der Schwerpunkt in diesem Artikel liegt auf der Beschreibung eines Algorithmus zum Übersetzen eines als gegeben angenommenen ER-Modells in eine Merkmal-Werte-Relation. Darauf aufbauend wird der tatsächliche Informationszustand durch eine sich im Verlauf des Dialogs ändernde gewichtete Merkmal-Werte-Relation beschrieben.

2 Dialogdreieck

Die zugrunde liegende Vorstellung ist, dass die Partizipation an einem Dialog in drei Schritte zerlegt werden kann und ein Dialog eine Abfolge dieser Schritte zwischen zwei Partnern ist, die diese abwechselnd vollziehen. Abbildung 1 beinhaltet die Sicht auf einen Teilnehmer eines Dialogs.

Ausgehend von einem bestimmten Informationsstatus (*InfoS*), der den aktuellen Informationsstand repräsentiert, werden nach einer gewissen, flexibel gestaltbaren, Strategie eine *Nachfrage* und gleichzeitig ein Verstehenshorizont (*VerH*) generiert. Die Nachfrage wird an den Dialogpartner übermittelt. Anschließend fungiert der Verstehenshorizont szsg. als Filter und begrenzt das, was von der Antwort des Partners *erkannt* werden kann.

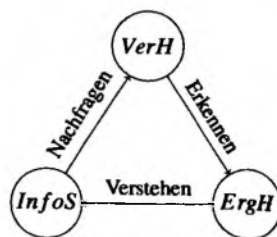


Abbildung 1 - Darstellung eines Dialogteilnehmers

¹Wir unterscheiden hier zwischen Begriffen und sprachlichen Formulierungen. Ein Begriff ist hierbei eine Äquivalenzklasse sprachlicher Formulierungen.

Die Ausgabe dieses Erkennens, der Ergebnishorizont (*ErgH*), wird dann in einem *Verstehensprozess* interpretiert und mit dem bisherigen Informationsstatus zu einem neuen Informationsstatus verarbeitet. Von diesem höheren Stand aus kann der Kreislauf erneut beginnen.

In diesem Modell lassen sich Morris' semiotische Begriffe wiederfinden. Basierend auf dem Verstehenshorizont wird unter Verwendung einer vorgegebenen Menge lokaler Grammatiken ein Sprachmodell gebildet, welches die Menge an Mustern, die erkannt werden können, auf der syntaktischen Ebene spezifiziert. Durch die Überführung in den Ergebnishorizont werden die erkannten Muster auf die semantische Ebene gehoben, auf der die (variablen) Wortfolgen durch (feste) Merkmal- und Wertennamen repräsentiert sind, was wir als *Verstehen 1. Ordnung* bezeichnen. Anschließend wird die erkannte, im Ergebnishorizont semantisch interpretierte, Information in den bisherigen Kontext eingebettet. Damit entsteht ein neuer Informationsstatus, mit dem der nächste Verstehenshorizont gebildet wird. Dieser Vorgang gehört zur pragmatischen Ebene, wir nennen ihn *Verstehen 2. Ordnung*.

Im Zusammenhang mit diesem Modell bedeutet Dialogstrategie zum einen eine Nachfrage zu stellen und zum anderen den nächsten Verstehenshorizont aufzubauen.

3 Einführende Definitionen

Wir beginnen mit einigen grundlegenden mathematischen Notationen. Mit \mathbb{N} bezeichnen wir die Menge der *nicht-negativen ganzen Zahlen*, mit \mathbb{R} die Menge der *reellen Zahlen* und mit \mathbb{R}_+ die Menge der *nicht-negativen reellen Zahlen*. Ein *Wahrscheinlichkeitsmaß* ω auf einer endlichen Menge V ist eine Abbildung $\omega : V \rightarrow [0, 1]$ mit $\sum_{v \in V} \omega(v) = 1$. Für eine endliche Menge V bezeichnet V^+ wie üblich die Menge aller endlichen Worte über V . Eine binäre Relation R über V heißt *linkstotal*, falls für jedes $v \in V$ ein $w \in V$ existiert mit $(v, w) \in R$. Ein *gerichteter Graph* G ist ein Paar $G = (V, \rightarrow)$, wobei V eine endliche Menge von *Knoten* und die Relation $\rightarrow \subseteq V \times V$ die Menge der *Kanten* ist. Wie üblich schreiben wir auch $v \rightarrow w$ für $(v, w) \in \rightarrow$. Eine endliche Folge von Knoten $v_0 \dots v_n$ ($n \in \mathbb{N}$) mit $v_{i-1} \rightarrow v_i$ ist ein *Pfad* von v_0 nach v_n . Die transitive Hülle einer Relation R ist definiert als die Relation

$$R^+ := \{(v_0, v_n) \mid \text{es gibt einen Pfad im Graphen } (V, R) \text{ von } v_0 \text{ nach } v_n\}.$$

Ein Pfad $v_0 \dots v_n$ mit $v_0 = v_n$ ist ein *Zyklus*. Eine *partielle Ordnung* ist ein gerichteter Graph $(V, <)$, wobei $<$ irreflexiv und transitiv ist. Eine Relation R über einer Menge V lässt sich als gerichteter Graph (V, R) auffassen und umgekehrt. Bekanntlich ist eine Relation R genau dann zyklensfrei, wenn R^+ irreflexiv, also eine partielle Ordnung, ist.

Merkmal-Werte-Relationen

Zur Darstellung semantischer Inhalte während des Informationsflusses verwenden wir den Begriff *gewichtete Merkmal-Werte-Relation*, der wie folgt definiert ist (vgl. Huber et al. [4]):

Definition 1. Gegeben seien zwei disjunkte endliche Mengen F (Merkmalmenge, feature set) und V (Menge atomarer Werte, value set). Eine Merkmal-Werte-Relation (MWR) über F und V ist eine linkstotale, zyklensfreie Relation $mwr \subseteq F \times (F \cup V)$. Zu jedem Merkmal $f \in F$ erhält man die Wertemenge $W(f)$ durch $W(f) := \{v \in F \cup V \mid (f, v) \in mwr\}$. Die MWR bzgl. eines Merkmals f notieren wir als $mwr_f := mwr|_{\{f\} \times (F \cup V)}$.

Eine *stochastische MWR* ist ein Tupel (mwr, ω) bestehend aus einer MWR mwr und einer Menge $\omega = \{\omega_f \mid f \in F\}$ von *Wahrscheinlichkeitsmaßen* $\omega_f : W(f) \rightarrow [0, 1]$ auf den nicht-leeren Wertemengen.

Eine *gewichtete MWR* ist ein Tupel (mwr, κ) bestehend aus einer stochastischen MWR mwr und einer Funktion $\kappa : F \rightarrow \mathbb{R}_+$ von *Konfidenzen* zu den Merkmalen.

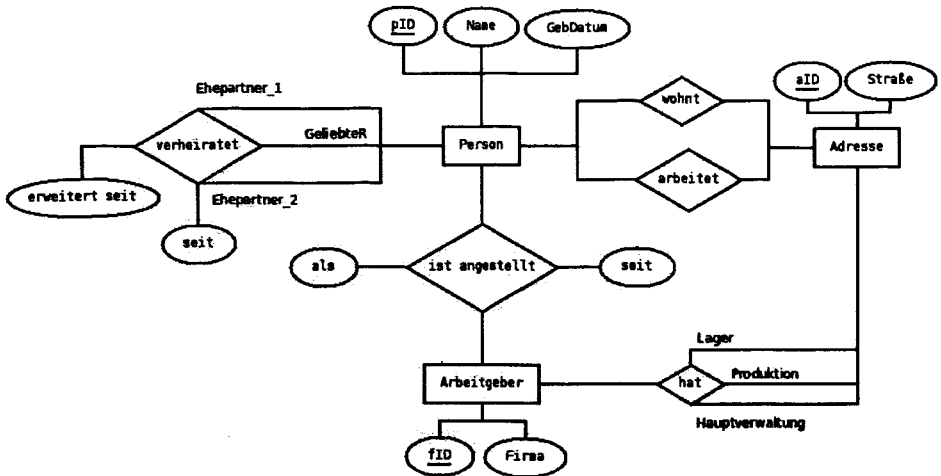


Abbildung 2 - Beispiel ER-Diagramm. (Der Beziehungstyp *verheiratet* modelliert hier eine "typische" Dreiecksbeziehung.)

Die Wahrscheinlichkeitsmaße und Konfidenzen werden uns zur Modellierung von Erkennungsmöglichkeiten und Erkennungsgenauigkeiten sowie Zuverlässigkeiten von Interpretationsmöglichkeiten dienen.

ER-Modell und relationale Datenbank

Um eine Datenbank zu beschreiben benutzen wir das Entity-Relationship-Modell in der Notation von Chen [1], wobei wir aus Platzgründen keine zusammengesetzten Attribute betrachten und zur Identifizierung von Entitäten IDs benutzen. Da unsere Übersetzung in MWRN für beliebige Kardinalitäten möglich ist, gehen wir davon aus, dass es sich immer um $m : n$ -Beziehungen handelt.

Außerdem setzen wir voraus, dass die Datenbank als relationales Datenmodell nach Codd [2] vorliegt und durch das ER-Modell beschrieben wird, wobei wir fordern, dass jeder Beziehungstyp zwischen Entitätstypen in der Datenbank als eigene Relation realisiert ist.

Betrachten wir nun im Folgenden die Übersetzung eines ER-Diagramms in eine Mengennotation, gestützt von unserem Beispiel aus Abbildung 2.

Aus dem ER-Diagramm entnehmen wir nun die Menge $\mathcal{E} = \{E_1, \dots, E_n\}$ aller Entitätstypen, in unserem Beispiel $E_1 = \text{Person}$, $E_2 = \text{Adresse}$, $E_3 = \text{Arbeitgeber}$. Des Weiteren beschreibt $\mathcal{A} = \{A_1, \dots, A_n\}$ die Menge aller Attributmengen, wobei jeweils A_i die Attributmenge zum Entitätstyp E_i bezeichnet, im Beispiel ist $A_1 = \{\text{Name, Geburstdatum, pID}\}$, $A_2 = \{\text{Straße, aID}\}$ und $A_3 = \{\text{Firma, fID}\}$. Weiter sei $\mathcal{R} = \{R_1, \dots, R_m\}$ die Menge aller Beziehungstypen zwischen den Entitätstypen, im Beispiel $R_1 = \text{'verheiratet'}$, $R_2 = \text{'wohnt'}$, $R_3 = \text{'arbeitet'}$, $R_4 = \text{'ist angestellt'}$, $R_5 = \text{'hat'}$. Zusätzlich ordnen wir jedem R_i eine mindestens zweielementige Zeichenkette über \mathcal{E} zu, also $f : \mathcal{R} \rightarrow \mathcal{E}^+ \setminus \mathcal{E}$. Für das Beispiel erhalten wir $f(R_2) = E_1E_2$, $f(R_3) = E_1E_2$, $f(R_4) = E_1E_3$. Zusätzlich können zu einem Beziehungstyp Rollennamen definiert sein, zum Beispiel gibt es beim Beziehungstyp $R_1 = \text{'verheiratet'}$ mit der Zeichenkette $f(R_1) = E_1E_1E_1$ die Rollennamen 'Ehepartner_1', 'Ehepartner_2' und 'GeliebterR'; ebenso hat der Beziehungstyp $R_5 = \text{'hat'}$ die Zeichenkette $f(R_5) = E_3E_2E_2E_2$ mit den Rollennamen 'Hauptverwaltung', 'Produktion' und 'Lager'. Analog zu den Entitätstypattributen sei $\mathcal{B} = \{B_1, \dots, B_m\}$ die Menge

Tabelle 1 - Relationen zu den Entitätstypen

Person			Adresse		Arbeitgeber	
Name	Geburtsdatum	pID	Straße	aID	Firma	fID
Meyer	01.01.1970	1	Lindenstraße	1	Quelle	1
Mayr	03.03.1973	2	Ahornstraße	2	Karstadt	2
Müller	02.02.1972	3	Schillerstraße	3	GM	3

Tabelle 2 - Relationen zu den Beziehungstypen

verheiratet					wohnt		arbeitet	
Ehepartner_1	Ehepartner_2	GeliebteR	seit	erweitert seit	pID	aID	pID	aID
1	3	2	1996	1998	1	1	1	2
					3	1	3	2
					2	3	2	3

ist angestellt				hat			
pID	fID	als	seit	fID	Lager	Produktion	Hauptverwaltung
1	1	Insolvenzverwalter	01.07.2009	1	3	1	2
2	2	Buchhalter	01.09.2008	2	3	1	2
3	1	Lageristin	01.09.1988	3	3	1	2

aller Beziehungstypattributmengen, wobei B_i die Menge der Attribute zum Beziehungstyp R_i bezeichnet, im Beispiel $B_1 = \{ 'seit', 'erweitert seit' \}$, $B_2 = B_3 = B_5 = \emptyset$, $B_4 = \{ 'seit', 'als' \}$.

Die Füllung der Datenbank liege uns in Form der Relationen aus Tabelle 1 und Tabelle 2 vor.

Die Beschränkung einer Relation E_i (bspw. Person) auf eine Teilmenge $\beta \subset A_1, \dots, A_n$ aller Attribute (hier z.B. pID) nennt sich *Projektion* $\pi_\beta(E_i)$ ($\pi_{pID}(\text{Person})$) und meint nach Codd [2]: $\pi_\beta(E_i) := \{ t_\beta | t \in E_i \}$ (die o.g. Projektion liefert also die Relation Person beschränkt auf deren Spalte pID). Eine weitere Operation auf einer Relationen E_i ist die *Selektion*, mit der man Tupel (Zeilen) gemäß eines Ausdrucks – der Selektionsbedingung – auswählen kann. Nach Codd [2] lautet die Notation der Selektion $\sigma_{\text{Ausdruck}}(E_i) := \{ t | t \in E_i \wedge t \text{ erfüllt Ausdruck} \}$.

4 Konstruktion von elementaren MWREN

Bevor wir mit der Konstruktion von MWREN beginnen definieren wir folgende Mengen:

Bezeichne zunächst für ein $R \in \mathcal{R}$ und $1 \leq j \leq |f(R)|$: $l_j(f(R)) = 'j\text{-ter Buchstabe von } f(R)'$ die Position des Entitätstyps E innerhalb des Beziehungstypwortes $f(R)$.

Weiter bezeichne $\mathfrak{R}_E := \{ (R, i) \in \mathcal{R} \times \mathbb{N} \mid l_i(f(R)) = E \}$ alle Paare von Beziehungstypen und Positionen, an denen der Entitätstyp E auftritt.

Wir konstruieren nun zu jedem Entitätstyp $E_i \in \mathcal{E}$ eine MWR (ohne Gewichte) wie folgt (die Algorithmusbeschreibung wird von dem Entitätstyp Person unseres laufenden Beispiels gestützt):

Wir starten bei einem Entitätstyp E_i ($E_1 = \text{Person}$) und nehmen nun E_i in die Merkmalmenge F_i (F_1) der elementaren MWR zu E_i (mwr_{E_i}) auf. Zum Merkmal E_i erhalten wir eine Wertemenge $W(E_i)$, die sich aus zwei disjunkten Teilmengen zusammensetzt: Der Menge A_i aller Attribute zu E_i (bei Person sind das Name, Geburtsdatum, pID) und der Menge aller möglichen Repräsentationen² von Beziehungstypen an denen E_i beteiligt ist, also \mathfrak{R}_{E_i} (im Beispiel ist $\mathfrak{R}_{E_1} =$

²Aufgrund von n-ären, reflexiven und rollenbehafteten Beziehungstypen kann ein Entitätstyp in verschiedenen Kombinationen eines Beziehungstyps auftreten. Somit wird für jede der Rollen zu 'verheiratet' bei Person eine

$\{(verheiratet, 1), (verheiratet, 2), (verheiratet, 3), (wohnt, 1), (arbeitet, 1), (ist\ angestellt, 1)\}$.
 Folglich erhalten wir also $W(E_i) = A_i \cup \mathfrak{R}_{E_i}$. Jedes Element aus dieser Wertemenge $W(E_i)$ hat ebenfalls wieder Werte, aus diesem Grund sind auch \mathfrak{R}_{E_i} und A_i wiederum Merkmalmengen und werden F_i hinzugefügt.

Nun müssen wir die Wertemengen für die gerade hinzugefügten Merkmale definieren. Die Wertemenge eines jeden $a \in A_i$ ist dann die Projektion der Relation E_i auf das Attribut a , versehen mit dem Attribut a . Wir erhalten $\forall a \in A_i : W(a) = (\pi_a(E_i), a)$ (im Beispiel: $\pi_{pID}(Person) = \{(1, pID), (2, pID), (3, pID)\}$, $\pi_{Name}(Person) = \{(Meyer, Name), (Mayr, Name), \dots\}$). Weiter suchen wir noch $W((R_j, k))$ mit $(R_j, k) \in \mathfrak{R}_{E_i}$. Die Wertemenge zu einem (R_j, k) besteht nun aus den Beziehungstypattributen B_j versehen mit dem Beziehungstyp und der Rolle. Wir erhalten also $W((R_j, k)) = B_j \times \{(R_j, k)\}$. (Für R_4 erhalten wir

$$W(ist\ angestellt, 1) = \{(seit, ist\ angestellt, 1), (als, ist\ angestellt, 1)\},$$

für R_1 entsprechend

$$W(verheiratet, 1) = \{(seit, verheiratet, 1), (erweitert\ seit, verheiratet, 1)\}.$$

Schließlich fehlen uns noch die Wertemengen zu den Elementen aus $B_j \times \{(R_j, k)\}$, welche analog zu A_i konstruiert werden, nämlich durch

$$W((b, R_j, k)) = \pi_b(R_j) \times \{(R_j, k)\} \forall b \in B_j$$

$$(W(als, ist\ angestellt, 1) = \{(Insolvenzverw., ist\ angestellt, 1), (Buchhalter, ist\ angestellt, 1), (Lagerist, ist\ angestellt, 1)\}.$$

Verstehenshorizont

Aus den elementaren MWRen können wir den Verstehenshorizont (vgl. Abschnitt 2) recht einfach aufbauen. Zuerst hängen wir die elementaren MWRen unter eine künstliche Wurzel und erweitern die so entstandene MWR zu einer stochastischen MWR (vgl. Definition 1). Jedes Merkmal m sowie die künstliche Wurzel erhalten also ein Wahrscheinlichkeitsmaß ω_m auf ihren entsprechenden Wertemengen $W(m)$.

Die Wahrscheinlichkeitsmaße sagen aus, auf welchen Werten die Präferenz im Verstehen liegen soll, also welche Muster aktiviert werden sollen.

Ergebnishorizont

Der Ergebnishorizont (vgl. Abschnitt 2) besteht ebenfalls aus elementaren MWRen, die zu stochastischen MWRen erweitert wurden. Er unterscheidet sich allerdings strukturell vom Verstehenshorizont. Die elementaren MWRen stehen nicht notwendigerweise lose nebeneinander bevor sie unter einer künstlichen Wurzel zusammengefasst werden, sondern sind, je nach Erkanntem, in Beziehung zueinander gesetzt. Dies wird dadurch ausgedrückt, dass eine elementare MWR an passender Stelle eines Beziehungstyp-Rollen-Merkmals $((R_j, k))$ „angehängt“ wird. So kann man bspw. darstellen, dass eine erkannte Adresse zu einer erkannten Person gehört.

Durch die Wahrscheinlichkeitsmaße können wir ausdrücken, welche Zuordnung von Werten zu einem Merkmal dem Erkannten entspricht.

Beziehung benötigt, die berücksichtigt werden muss. Bildlich gesprochen muss jede Kante eines Beziehungstyps aufgelöst werden.

Informationsstatus

Der Informationsstatus ist eine Menge von gewichteten Merkmal-Werte-Relationen. Jede dieser MWRen – *Welt* genannt – ergibt sich aus dem ER-Modell und der relationalen Datenbank, indem die oben beschriebenen elementaren MWRen durch zusätzliche Knoten erweitert und mit Wahrscheinlichkeitsmaßen und Konfidenzfunktionen versehen werden, wobei sowohl die Wahrscheinlichkeitsmaße als auch die Konfidenzen bei jedem Schritt mit Hilfe des Ergebnishorizonts aktualisiert werden.

Eine Welt entsteht also dadurch, dass wir für jeden Entitätstyp E_i die entsprechende elementare MWR um eine ID-Schicht erweitern. Zu jedem Paar $(R_j, k) \in \mathfrak{R}_{E_i}$ erhalten wir eine Menge³

$$\{(F, m, R_j, k) \in \mathcal{E} \times \{1, \dots, |f(R_j)|\} \setminus \{k\} \times \{(R_j, k)\} \mid l_m(f(R_j)) = E_i\}$$

über deren Elemente wir nun die Merkmale der ID-Schicht durch

$$\pi_{(ID \text{ von } F)}(F) \times \{(F, m, R_j, k)\}$$

erhalten. Die Wertemenge eines solchen Merkmals ergibt sich durch

$$W(v, F, m, R_j, k) = \pi_{(ID \text{ von } E_i \text{ in der Rolle } k)}(\sigma_{(ID \text{ von } F \text{ in der Rolle } m=v)}(R_j)).$$

Zusätzlich definieren wir noch die Wertemengen der $q \in W(a \in A_i)$ durch

$$W(q) = \pi_{(ID \text{ von } E_i)}(\sigma_{(a=q)}(E_i)).$$

5 Steuerung des Dialogs

Wie in Abschnitt 2 beschrieben, besteht ein Dialog aus der zyklischen Abfolge von Nachfragen, Erkennen und Verstehen. Beim Nachfragen wird einerseits eine Eingabeaufforderung formuliert und andererseits ein Verstehenshorizont generiert, der dem Erkennen übermittelt wird. Dieser erzeugt nun zunächst aus dem VerH und einem Vorrat an lokalen Grammatiken ein für die Erkennung geeignetes Sprachmodell. Der Erkennen benutzt dann dieses Sprachmodell, um das akustische Signal anhand der sich aus dem Sprachmodell ergebenden bewerteten Wortfolgen weich zu klassifizieren. Daraus ergeben sich für die einzelnen Erkennalternativen Wahrscheinlichkeiten, die es erlauben, den ErgH als stochastische MWR aufzubauen. Der nächste Schritt der Dialogsteuerung ist das Aktualisieren des Informationsstatus' unter Verwendung des Erkennungsergebnisses in der Form eines Ergebnishorizonts, d.h. das Einbinden der erkannten Information in die bestehende Information.

Ein möglicher Ausgangspunkt für einen Dialog ist ein „leerer“ Informationsstatus, was bedeutet, dass die Wahrscheinlichkeitsmaße Gleichverteilungen und die Konfidenzen 0 sind. Die Dialogstrategie fordert den Benutzer nun zur Eingabe einer Anfrage auf und generiert gleichzeitig einen genügend breiten VerH. Nehmen wir an, die beiden folgenden Anfragen können mit dem Sprachmodell, das der Erkennen aus dem gegebenen VerH generiert hat, verstanden werden:

1. „Wann hat Meyer Geburtstag?“
2. „Wann hat Meyer aus der Lindenstraße Geburtstag?“

³Die verschiedenen „Personen“ erhalten wir aufgrund der Rollen und die Ziffern geben die Position des Entitätstyps im Wort R_j , hier *PersonPersonPerson*, an.

Nehmen wir weiter an, einer dieser beiden Sätze wurde vom Erkennen verstanden und im ErgH repräsentiert. Im ersten Fall erhalten wir die elementare MWR für den Entitätstyp Person, wobei die Wahrscheinlichkeitsverteilungen auf den Attributen bzgl. Name und auf dessen Werten bzgl. „Meyer“ und „Mayr“ angepasst werden. Im zweiten Fall erhalten wir eine größere MWR, die aus der elementaren MWR des ersten Falles besteht und einer weiteren elementaren MWR für den Entitätstyp Adresse, der am Knoten 'wohnt,1' der ersten hängt.

Anschließend wird der InfoS mit Hilfe dieses ErgH aktualisiert und damit die im ErgH codierte semantische Information eingearbeitet und für den weiteren Dialogablauf verfügbar gemacht. Die Aktualisierung erfolgt durch einen Algorithmus, der einerseits auf dem klassischen prädikatenlogischen Unifikationsalgorithmus⁴ basiert und andererseits auf einer Bayesianisch motivierten Verknüpfung von Konfidenzen und Wahrscheinlichkeitsmaßen einzelner Merkmale (vgl. Huber [3]). Bei der Aktualisierung wird auch klar, dass im zweiten Fall die Entität „Meyer“ eindeutig identifiziert ist, wohingegen das im ersten Fall noch nicht zutrifft. Bei der ersten Alternative muss also die Dialogsteuerung den Dialog mit einer passenden weiteren Nachfrage und einem entsprechenden VerH fortsetzen.

Ist das Erkenergebnis uneindeutig, z.B. weil vom Erkennen zwei sich gegenseitig ausschließende Satzbedeutungen mit genügend hoher Wahrscheinlichkeit bewertet wurden, werden bei der Aktualisierung zwei verschiedene „Parallelwelten“, nämlich Kopien, die sich anhand ihrer Gewichtsfunktionen unterscheiden, erzeugt. Die Dialogsteuerung hat jetzt die zusätzliche Aufgabe, sich im weiteren Verlauf mittels geeigneter Strategien für eine Welt zu entscheiden.

Literatur

- [1] CHEN, P. P.-S.: *The entity-relationship model-toward unified view of data*. Communications of the ACM, 1(1):9–36, 3 1976.
- [2] CODD, E. F.: *A relational model of data for large shared data banks*. Communications of the ACM, 13(6):377–387, 6 1970.
- [3] HUBER, M.: *Semantische Informationsbearbeitung unter Berücksichtigung von Konfidenzzahlen*. Diplomarbeit in Mathematik. Katholische Universität Eichstätt-Ingolstadt, 2009.
- [4] HUBER, M., C. KÖLBL, R. LORENZ und G. WIRSCHING: *Ein Petrinetz-Modell zur Informationsübertragung per Dialog*. In: LOHMANN, N. und K. WOLF (Hrsg.): *Proceedings of the 15th German Workshop on Algorithms and Tools for Petri Nets, AWPN 2008, Rostock, Germany, September 26–27, 2008*, Bd. 380 d. Reihe CEUR Workshop Proceedings, S. 15–24. CEUR-WS.org, Sep. 2008.
- [5] MORRIS, C. W.: *Foundations of the theory of signs*, vol. 1-2 of *Foundations of the unity of science*. University of Chicago Press, Chicago, 12 ed., 1975.

⁴Dieser Algorithmus wird gerade im Rahmen einer Mathematik-Diplomarbeit erarbeitet.