

KONSTRUKTION VON UMP-TRANSDUKTOREN AUS WIZARD-OF-OZ DATEN

Markus Huber¹, Christian Kölbl¹, Robert Lorenz¹, Günther Wirsching²

¹ Universität Augsburg

² Katholische Universität Eichstätt-Ingolstadt

¹ vorname.nachname@informatik.uni-augsburg.de

² guenther.wirsching@ku-eichstaett.de

Zusammenfassung: Wir beschreiben in dieser Arbeit die Konstruktion eines Transduktors zur Übersetzung von Erkennungsergebnissen zwischen der syntaktischen und der semantischen Ebene eines hierarchischen Systems zur Sprachsignalverarbeitung aus Wizard-of-Oz Daten.

Dazu werden in einem ersten Schritt Äußerungs-Bedeutungs-Paare gebildet und dann ein Transduktor konstruiert, der jede Äußerung in die zugehörige Bedeutung übersetzt. Der Transduktor wird nicht durch einen Automaten, sondern durch ein Petrinetz realisiert, da Bedeutungen durch partielle Wörter repräsentiert werden.

1 Einleitung

Wir betrachten in diesem Artikel die Schnittstelle zwischen der syntaktischen und der semantischen Ebene eines auf dem UASR-Ansatz [1, 9, 4] basierenden hierarchischen kognitiven dynamischen Systems zur Sprachsignalverarbeitung.

Die Abbildung 1 gibt einen abstrakten Überblick über den Analyse- und Synthesezweig des Systems (siehe [9] für eine detailliertere Darstellung). Ziel des Systems ist die Steuerung eines natürlichsprachlichen Dialogs, der der Eingabe und Identifikation einer vom Benutzer gewünschten Aktion auf ausgewählten Datenobjekten dient. Benutzeranfragen sollen dabei frei formuliert werden können. Das System integriert sukzessive die Erkennungsergebnisse zu den Äußerungen des Benutzers in einen Informationsstatus (Analyse des akustischen Signals über mehrere Abstraktionsebenen). Falls noch Informationen zur Ausführung der Aktion fehlen, generiert es daraus automatisch eine Nachfrage (Synthese eines akustischen Signals über mehrere Abstraktionsebenen; nicht in der Abbildung dargestellt).

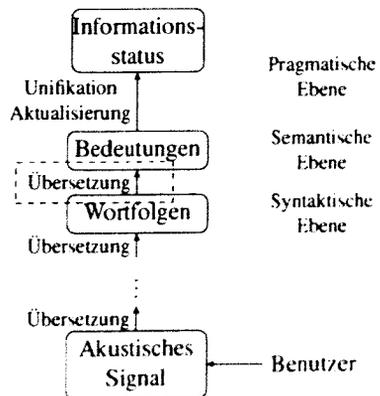


Abbildung 1 - Analyse-Zweig des Systems.

Bis einschließlich zur syntaktischen Ebene werden die Erkennungsergebnisse durch gewichtete Worte über geeigneten Alphabeten dargestellt, wobei die Gewichte Wahrscheinlichkeiten repräsentieren. Die Übersetzung zwischen den Ebenen er-

folgt durch gewichtete endliche Transduktoren (finite state transducers, FSTs). Durch Komposition können diese Transduktoren über mehrere Ebenen zu einem Transduktor zusammengesetzt werden. Die semantische Ebene dient der Interpretation der auf syntaktischer Ebene erkannten Äußerungen. Dazu werden diesen Äußerungen sogenannte *Bedeutungen* zugeordnet. Diese Zuordnung beschreiben wir durch *UMPs* (*utterance-meaning-pairs*) [8]. Die Komponente, welche die Übersetzung realisiert, nennen wir *UMP-Transduktor*. Da wir zur Darstellung von Bedeutungen *partielle Wörter* verwenden, ist eine Realisierung von UMP-Transduktoren durch automatenbasierte FSTs nicht mehr intuitiv [6]. Stattdessen benutzen wir eine neue und allgemeinere Theorie der Petrinetz-Transduktoren (PNTs) [4]. PNTs ermöglichen in natürlicher Weise die Übersetzung zwischen *partiellen Sprachen*, welche aus *partiellen Wörtern* bestehen. Ein partielles Wort besteht dabei nicht aus einer totalen, sondern aus einer partiellen Ordnung auf seinen Symbolen. Da jeder FST ein spezieller PNT ist, ist es möglich, die ganze in der Abbildung dargestellte Hierarchie durch Kombination von FSTs und PNTs zu realisieren.

In diesem Artikel untersuchen wir eine strukturierte Vorgehensweise zur Konstruktion von UMP-Transduktoren für konkrete Anwendungen des beschriebenen Systems. Jede solche Anwendung basiert auf einem sogenannten *Weltmodell*. Das Weltmodell legt alle Bedeutungen von Äußerungen, mit denen das System umgehen kann, fest. Mit einem Weltmodell sind folglich indirekt auch alle Äußerungen, die prinzipiell vom System erkannt werden können, festgelegt - das sind alle Äußerungen mit einer Bedeutung aus dem Weltmodell. Wir gehen von einer Anwendung aus, für die schon ein Weltmodell konstruiert wurde.¹ Für eine solche Anwendung kann nun ein Wizard-of-Oz (WOZ) Experiment durchgeführt werden, in dem Versuchspersonen ein funktionierendes System vorgespiegelt wird und deren Dialoge mit dem System aufgezeichnet werden. Aus diesen Aufzeichnungen lässt sich dann semi-automatisch eine Menge von Äußerungen mit zugehörigen Bedeutungen ableiten. Aus den abgeleiteten UMPs kann dann ein UMP-Transduktor konstruiert werden.

Der Artikel ist wie folgt aufgebaut. Im Abschnitt 2 wird ein WOZ-Experiment aus dem Home-Entertainment-Anwendungsbereich mit zugehörigem Weltmodell beschrieben, welches im Rahmen einer Masterarbeit [3] an der BTU Cottbus durchgeführt wurde. In Abschnitt 3 wird anhand der in diesem Experiment gewonnenen Daten die Ableitung von UMPs vorgeführt. Danach wird in Abschnitt 4 ein zugehöriger UMP-Transduktor konstruiert (inklusive einer kurzen Einführung in die Theorie der PNTs).

2 Experiment und Weltmodell

In diesem Abschnitt beschreiben wir kurz das WOZ-Experiment, aus dem die in dieser Arbeit benutzten Daten gewonnen wurden, und das dem Experiment zugrundeliegende Weltmodell.

Das Experiment wurde 2012 an der BTU Cottbus im Rahmen einer Masterarbeit [3] durchgeführt. Insgesamt nahmen zwölf Probanden an der Studie teil, sieben davon männlich und fünf weiblich. Bei der Auswertung wurden die Daten nicht nach Geschlechtern getrennt, daher ist das Wort "Proband" im folgenden geschlechtsneutral zu verstehen.

Zur Durchführung des Experiments saß der Proband in einer Sprecherkabine vor einem Tisch, auf dem sich ein Computermonitor, ein Tablet-PC, ein Mikrofon und Lautsprecherboxen befanden. Die sprachlichen Eingaben des Probanden wurden zum Experimentator geleitet, der außerhalb der Sprecherkabine saß und die Antworten der Maschine zusammenstellte, die dann über eine LAN-Verbindung an den Computer in der Sprechkabine gesendet wurden und dort in

¹Ein Weltmodell kann automatisch aus einem ER-Diagramm mit zugehöriger Datenbank generiert werden [2].

einem Pop-up Fenster auf dem Monitor des Probanden erschienen. Aufgabe der Probanden war es, das System dazu zu bringen, einen bestimmten Film auf Platz 1 einer Favoritenliste zu setzen. Dabei kannten die Probanden zwar nicht den Titel des Films, konnten aber mit Hilfe eines speziell für das Experiment zusammengestellten "Electronic Program Guide" Informationen über Schauspieler und Plot des Films erhalten. Bei der Durchführung des Experiments gelang es acht der zwölf Probanden, das simulierte System dazu zu bringen, die gewünschte Aktion auszuführen. Die Dialoge zwischen Probanden und (simuliertem) System wurden aufgezeichnet und danach verschriftet. Später dienen diese Verschriftungen der Erstellung von UMPs.

In [2] haben wir eine einheitliche Datenstruktur zur Repräsentation aller Komponenten der semantischen und pragmatischen Ebene aus Abbildung 1 entworfen. Diese Datenstruktur nennen wir *Merkmal-Werte-Relation (MWR)*. Wir verzichten im weiteren auf eine ausführliche formale Darstellung und arbeiten stattdessen mit folgender Beschreibung einer MWR: Eine MWR ist ein azyklischer gerichteter Graph. Dieser definiert eine Hierarchie von semantischen Kategorien, genannt *Merkmalen*, wie zum Beispiel *Schauspieler*, *Name* und *Vorname*. Manchen Merkmalen werden Datenwerte zugeordnet, sogenannte *Werte*, wie zum Beispiel *Hugh* und *Kate*. Schließlich werden IDs von Datenobjekten den Datenwerten der Objekte zuordnet.

Insbesondere stellen wir das einer Anwendung zugrundeliegende Weltmodell als MWR dar. Dieses legt alle Datenobjekte mit ihren Werten und deren durch Merkmale beschriebenen semantische Bedeutungen fest. Im folgenden beschreiben wir das Weltmodell, das dem beschriebenen Experiment zugrunde lag. Hierbei gehen wir aus Gründen der Übersichtlichkeit nur auf einen Teil der Merkmale, Datenwerte und Datenobjekte ein, die in [3] beschrieben werden (dort wird das Weltmodell durch ein ER-Diagramm dargestellt).

Als Datenobjekte stehen Medienobjekte und Schauspieler zur Verfügung. Abbildung 2 zeigt einen Ausschnitt des Weltmodells, der das Merkmal *Schauspieler* beschreibt. Ein Schauspieler wird danach durch die Merkmale *Vorname* und *Nachname* beschrieben, welche wiederum noch durch das Merkmal *Name* zusammengefasst werden. Die MWR enthält die zugehörigen Daten von genau zwei Datenobjekten vom Typ *Schauspieler*. Deren IDs schließen die MWR ab. Abbildung 3 zeigt

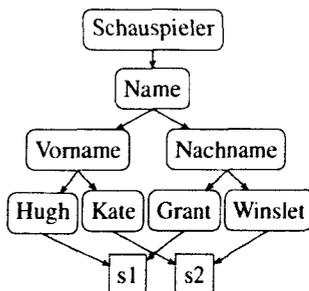


Abbildung 2 - MWR zur Beschreibung des Merkmals *Schauspieler* mit zwei zugehörigen Datenobjekten.

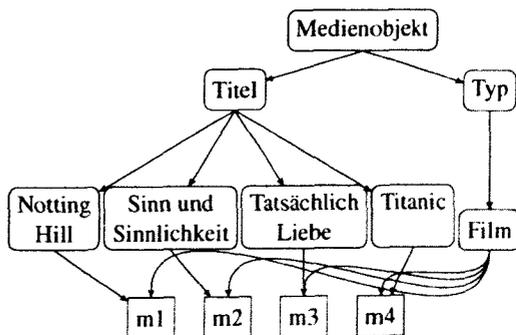


Abbildung 3 - MWR zur Beschreibung des Merkmals *Medienobjekt* mit vier zugehörigen Datenobjekten.

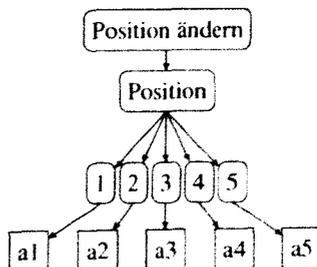


Abbildung 4 - MWR zur Beschreibung der Aktion *Position ändern*.

einen anderen Ausschnitt des Weltmodells, der das Merkmal *Medienobjekt* mit vier zugehörigen Datenobjekten beschreibt (bei denen es sich ausschließlich um Filme handelt - natürlich sind auch andere Typen wie Serien und dergleichen denkbar).

Im Experiments wurde genau eine ausführbare Aktion betrachtet, nämlich das Ändern der Platzierung von Medienobjekten auf Position 1 in einer Favoritenliste. Da generell mehrere Aktionen zur Verfügung stehen können, müssen auch diese im Weltmodell repräsentiert werden. Dazu werden ebenfalls Merkmale benutzt. Abbildung 4 zeigt einen Ausschnitt des Weltmodells, der obige Aktion beschreibt. Hierbei wird angenommen, dass die Favoritenliste aus fünf Positionen besteht und jede dieser Positionen geändert werden kann. Für jede Position ergibt sich eine Aktion mit eigener ID. Die bisherigen Weltmodell-Ausschnitte beschreiben noch nicht, dass die Positionen der Aktion *Position ändern* mit Medienobjekten besetzt werden sollen (und nicht z.B. mit Schauspielern). Eine solchen Beziehung zwischen Aktions- und Daten-

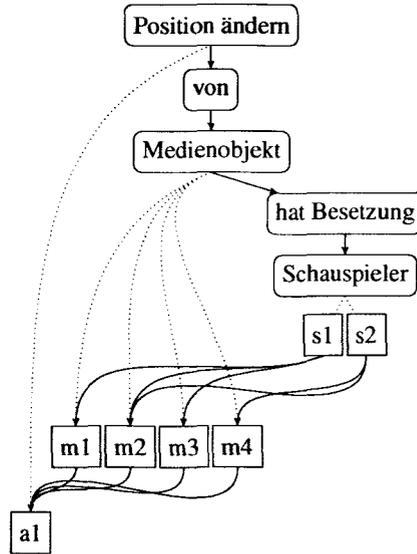


Abbildung 5 - Übersicht über das Weltmodell.

Merkmalen wird durch ein weiteres Merkmal, einen sogenannten *Anker*, und Verbindungen zwischen Objekt- und Aktions-IDs beschrieben. Ein Anker verbindet in Beziehung stehenden Merkmale in eine Richtung über zwei gerichtete Kanten. Abbildung 5 illustriert die Verbindung zwischen der Aktion *Position ändern* und dem Daten-Merkmal *Medienobjekt* über den Anker *von*, wobei diesmal aus Platzgründen nur die Aktion mit ID 1 gezeigt wird und die Struktur der Merkmale nicht im Detail dargestellt (sondern durch gestrichelte Linien abstrahiert) wird. Das vollständige Weltmodell ergibt sich aus Abbildung 5 durch Einsetzen der detaillierten Strukturen. Da es prinzipiell möglich sein soll jedes Medienobjekt auf Position 1 zu setzen, ist jede *Medienobjekt*-ID mit der betrachteten *Aktion*-ID verbunden.

Schließlich können Daten-Merkmale zueinander in Beziehung stehen, zum Beispiel hat ein Medienobjekt eine Besetzung aus Schauspielern. Eine solche Beziehung wird ebenfalls durch Anker beschrieben und ist in Abbildung 5 illustriert. Dabei sind die Schauspieler genau mit den Medienobjekten über IDs verbunden, in denen sie mitgewirkt haben. Bei Beziehungen zwischen Daten-Merkmalen modellieren wir immer nur explizit eine Richtung der Beziehung, obwohl wir auch die andere Richtung (*Schauspieler* gehört zur Besetzung von *Medienobjekt*) in Äußerungen zulassen (bei Beziehungen zwischen Aktions- und Datenmerkmalen ist allerdings immer nur die Richtung vom Aktions-Merkmal zum Daten-Merkmal zugelassen). Zusammengefasst können somit Datenobjekte in Äußerungen nicht nur über ihre Eigenschaften, sondern auch über ihre Beziehungen zu anderen Datenobjekten beschrieben werden.

3 Vorgehensmodell

Hier zeigen wir exemplarisch, wie sich aus gegebenen WOZ-Daten und einer durch ein Weltmodell gegebenen Merkmalstruktur eine Menge von Äußerungen mit zugehöriger Bedeutung

gewinnen lässt, welche vom System erkannt werden können. Die angesprochene *Merkmalsstruktur* ist dabei durch den Teil des Weltmodells ohne Datenwerte und IDs gegeben. Die Semantikschnittstelle im Analysezeit des Systems ist dazu da, möglichst jeder im Kontext sinnvollen Äußerung eines Benutzers ihre *Bedeutung* zuzuordnen. Eine Bedeutung repräsentieren wir wie schon das Weltmodell durch eine MWR. Dabei wird die benötigte Kontextinformationen durch einen wie in [7] beschriebenen, ebenfalls durch eine MWR gegebenen, Verstehenshorizont dargestellt (auf den Verstehenshorizont wird im weiteren Verlauf nicht weiter eingegangen). Die Aufgaben des Transduktors sind also die Beschreibung der Menge der im Kontext sinnvollen Wortfolgen, und die Übersetzung jeder Wortfolge in eine MWR, die deren Bedeutung darstellt.

3.1 Semantische Analyse der WOZ-Daten

Der erste Schritt ist, jedem der im WOZ-Experiment aufgenommenen Sätze eine Bedeutung in Form einer MWR zuzuordnen. Dieser Schritt muss im wesentlichen von Hand durchgeführt werden, wobei darauf zu achten ist, dass die MWR aus dem Weltmodell generierbar ist. Zum Beispiel passt die in Abbildung 6 gezeigte Bedeutung auf die Äußerung: *„Ich suche einen Film wo die Hauptdarstellerin aus Titanic mitspielt“*.

In dieser Äußerung wird ein Medienobjekt durch seinen Typ und seine Besetzung beschrieben. Die Beziehung des gesuchten Medienobjekts zu einem zur Besetzung gehörenden Schauspieler wird das Anker-Merkmal *hat Besetzung* repräsentiert. Der zur Besetzung gehörende Schauspieler wiederum wird durch seine Zugehörigkeit zur Besetzung des Medienobjekts *m4* mit dem Titel *Titanic* beschrieben. Hierfür wird dasselbe Anker-Merkmal wiederbenutzt, aber diesmal in seiner Rückrichtung *ist in Besetzung von*. Dieses Beispiel zeigt, dass Datenobjekte über die im Weltmodell definierten Beziehungen mehrfach ineinander verschachtelt werden können. Schließlich wird in unserem Beispiel der Schauspieler durch diese Beziehung eindeutig als *s2* identifiziert (die Besetzung des Medienobjekts *m4* besteht nur aus einem Schauspieler). Der Schauspieler *s2* wiederum identifiziert eindeutig das Medienobjekt *m2*, da in der gezeigten Bedeutungs-MWR berücksichtigt ist, dass nicht nach *m4* gesucht wird. Sind alle im WOZ-Experiment aufgenommenen Sätzen mit Bedeutungen versehen – wobei für die Modellierung des Analysezeitweigs die Benutzeräußerungen genügen –, dann ist der nächste Schritt die Auswahl von Sätzen (Wortfolgen), die zu den Merkmalsstrukturen der Bedeutungen passen.

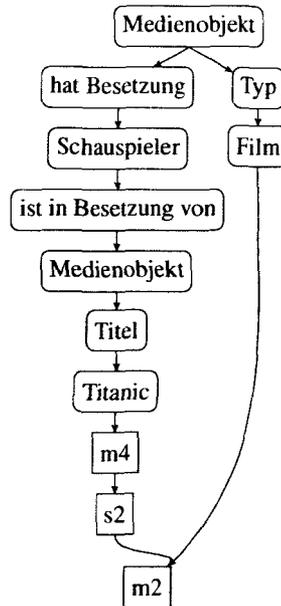


Abbildung 6 - Eine Bedeutung als MWR.

3.2 Generierung von Wortfolgen

Abstraktion von Datenwerten: Natürlich soll das Dialogsystem nicht nur genau die im WOZ-Experiment gesprochenen Sätze verstehen können, sondern auch gewisse Varianten davon. Zum Beispiel können in unserer Beispieläußerung anstelle des Wortes *„Titanic“* auch ein andere Titel von Medienobjekten genannt werden. Deshalb ersetzen wir jeden in Äußerungen vorkommenden Datenwert durch einen eindeutigen Bezeichner des Merkmals, zur dem der Datenwert gehört. Das führt zur Generalisierung

“Ich suche einen <Medienobjekt-Typ> wo die Hauptdarstellerin aus <Medienobjekt-Titel> mitspielt”

Spezielle sprachliche Generalisierungen: Eine weitere Generalisierung, die sich fast zwangsläufig aus der Generalisierung mit Datenwerten ergibt, ist das Zulassen von Worten mit direkt vergleichbarer Bedeutung. In unserem Beispiel scheint es klar, dass auch eine Äußerung verstanden werden soll, bei der “der Hauptdarsteller” anstelle von “die Hauptdarstellerin” gesagt wird. Dadurch werden aus unserem Beispielsatz die beiden Sätze

“Ich suche einen <Medienobjekt-Typ> wo die Hauptdarstellerin aus <Medienobjekt-Titel> mitspielt”

“Ich suche einen <Medienobjekt-Typ> wo der Hauptdarsteller aus <Medienobjekt-Titel> mitspielt”

Erweiterung durch synonymische Syntagmen: Beim nächsten Generalisierungsschritt ist das Ziel, auch im Detail anderslautende Äußerungen mit der gleichen Bedeutung zuzulassen, z.B. anstelle unserer Beispieläußerung auch

“Ich suche einen Film mit der Schauspielerin, die in Titanic die Hauptrolle gespielt hat”

Dieser Generalisierungsschritt ist mit großer Vorsicht anzuwenden, denn im WOZ-Experiment hat der Proband ja vielleicht nicht ohne Grund gerade eine bestimmte Formulierung gewählt. Auf der anderen Seite ist aber sinnvoll, Formulierungsvarianten, die im gegebenen Kontext wirklich austauschbar erscheinen, zu berücksichtigen. In unserem Beispiel erhalten wir durch diese Generalisierung zusätzlich die Sätze:

“Ich suche einen <Medienobjekt-Typ> mit der Schauspielerin, die in <Medienobjekt-Titel> die Hauptrolle gespielt hat”

“Ich suche einen <Medienobjekt-Typ> mit dem Schauspieler, der in <Medienobjekt-Titel> die Hauptrolle gespielt hat”

Glattziehen der Grammatik: Beim Einsetzen von beliebigen Datenbankeinträgen kann es zu grammatischen Inkongruenzen kommen, zum Beispiel bei den beiden Medienobjekt-Typen “Film” und “Serie” (wobei letzterer zwar nicht in unserem Weltmodell gezeigt wird, aber im Experiment genannt wurde), bei denen der Genus des ersten “maskulin” und der des zweiten “feminin” ist. Einige Beispiele glattgezogener Expansionen sind:

“Ich suche einen Film mit der Schauspielerin, die in Titanic die Hauptrolle gespielt hat”

“Ich suche eine Serie wo der Hauptdarsteller aus Notting Hill mitspielt”

4 Konstruktion eines UMP-Transduktors

In diesem Abschnitt führen wir in die Theorie der Petrinetz-Transduktoren (PNTs) ein und zeigen exemplarisch, wie sich aus den im letzten Abschnitt gewonnenen UMPs ein UMP-Transduktor in PNT-Technologie konstruieren lässt. Wir werden dazu Petrinetze nur informal und mit einer sehr einfachen Syntax einführen. Ein detaillierte formale und vollständige Beschreibung kann in [5] nachgelesen werden. Ein *Petrinetz* besteht aus *Transitionen* (in Grafiken dargestellt durch Quadrate), *Stellen* (Kreise) und einer Flussrelation zwischen Stellen und Transitionen (gerichtete Kanten). Die Transitionen stellen Aktionen dar und die Stellen Vor- und Nachbedingungen dieser Aktionen. Der Zustand eines Netzes wird durch eine Verteilung von Marken auf Stellen, eine sog. *Markierung*, repräsentiert. Ist eine Stelle markiert, so ist die zugehörige Bedingung erfüllt. Ein Transition kann in einer Markierung schalten, wenn alle

Ihre Vorbedingungen erfüllt sind. Das Schalten einer Transition verändert die Markierung so, dass danach ihre Vorbedingungen nicht mehr erfüllt sind, aber dafür ihre Nachbedingungen. Im Petrinetz N_1 aus Abbildung 7 können ausgehend vom Anfangszustand nacheinander die Transitionen a, b, c und so weiter schalten. Der wesentliche Unterschied zu Automaten ist, dass sich der globale Zustand eines Netzes aus mehreren lokalen Zuständen zusammensetzt. Wenn zwei Transition keine Vorbedingungen teilen und beide in einem Zustand aktiviert sind, so können beide unabhängig voneinander in beliebiger Reihenfolge oder auch simultan schalten. Solche Transitionen heißen *nebenläufig* (in einem solchen Zustand). Die Existenz von nebenläufigen Transitionen ermöglicht es, das Schalten von *Schrittsequenzen* zu definieren, wobei ein Schritt eine Menge nebenläufiger Transitionen ist. Im Petrinetz N_4 aus Abbildung 7 können ausgehend vom Anfangszustand nacheinander die Schritte $\{x\}, \{g, k\}, \{h, l\}$ und so weiter schalten. Auf der Basis von Schrittsequenzen definiert man schließlich *partiell geordnete Abläufe* (po-Abläufe) von Netzen. Ein solcher Ablauf ist eine partiell geordnete Menge von Knoten, die mit Transitionsnamen beschriftet sind, eine sogenannte *partielles Wort*. Die Knoten repräsentieren Transitions-Schaltvorgänge und die partielle Ordnung (gerichtete Kanten) eine "früher als"-Beziehung zwischen diesen Schaltvorgängen. Gibt es zwischen zwei Schaltvorgängen keine Kanten, dann sind diese nebenläufig im oben beschriebenen Sinne. Eine BPO ist ein po-Ablauf eines Netzes, wenn er konsistent ist mit den Schrittsequenzen, die im Netz ausgehend vom Anfangszustand schalten können. Das Petrinetz N_1 aus Abbildung 7 hat einen po-Ablauf der Form $a \rightarrow b \rightarrow c \rightarrow \dots \rightarrow e$. Das Petrinetz N_4 hat als po-Ablauf ein partielles Wort in der Form der MWR aus Abbildung 6 (natürlich mit Transitionsbeschriftungen).

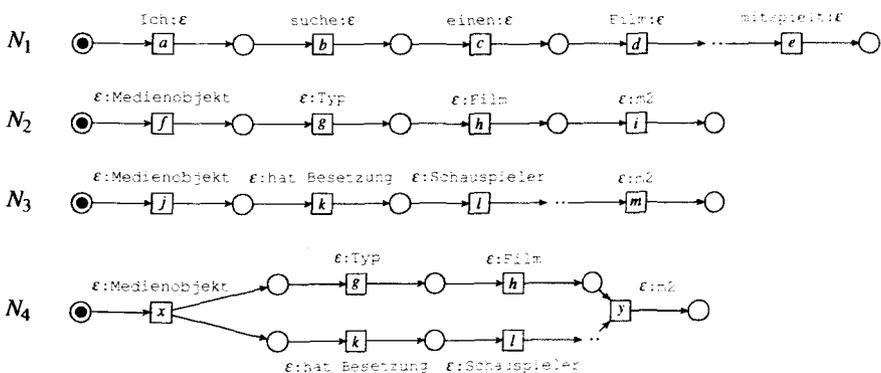


Abbildung 7 - UMP-Transduktoren, realisiert als PNTs.

Ein *Petrinetz-Transduktor* (PNT) ist ein Petrinetz, dessen Transitionen eine Beschriftung mit einem *Eingabe-Symbol* x und einem *Ausgabe-Symbol* y haben (grafisch annotiert in der Form $x : y$). Schaltet eine Transition, so wird das entsprechende Eingabe-Symbol x gelesen und das Ausgabe-Symbol y geschrieben. Soll kein Symbol gelesen oder geschrieben werden, so wird das Symbol ϵ für das leere Wort verwendet. Jeder PNT hat einen ausgezeichneten Startzustand und einen ausgezeichneten Endzustand, beide jeweils repräsentiert durch eine einzelne Stelle. Es werden nur solche po-Abläufe betrachtet, die vom Start- in den Endzustand führen. Die Menge der po-Abläufe eines PNT N bezeichnen wir mit $LPO(N)$. Ein *Eingabewort* von N erhält man, indem man die Knoten eines po-Ablaufs aus $LPO(N)$ statt mit Transitionsnamen mit den zugehörigen Eingabesymbolen beschriftet. Analog wird das zu einem Eingabewort gehörende *Ausgabewort* gebildet. Der PNT N_1 aus Abbildung 7 hat als Eingabe die in Abschnitt 3 betrachtete Äußerung und keine Ausgabe. Die PNTs N_2 und N_3 haben keine Eingabe und generieren als Ausgabe jeweils einen Pfad der MWR aus Abbildung 6. Der PNT N_4 hat keine Eingabe und

generiert als Ausgabe die komplette MWR aus Abbildung 6. Wie für FSTs kann man auch für PNTs nützliche Kompositions-Operationen definieren und implementieren, mit denen man aus einfachen PNTs Schritt für Schritt komplexere PNTs konstruieren kann. Wie für FSTs stehen Operationen wie Vereinigung, Produkt, iterativer Abschluss und Sprach-Komposition zur Verfügung. Darüber hinaus lassen sich neue Operationen definieren, die nicht auf FSTs anwendbar sind, da sie zu Nebenläufigkeit führen, wie parallele oder synchrone Komposition. Beispielsweise entsteht N_4 durch synchrone Komposition von N_2 und N_3 . Durch das Produkt von N_1 mit N_4 erhält man einen Transduktor zur Übersetzung einer Äußerung in eine Bedeutung. Durch Vereinigung lassen sich mehrere Äußerungen zu einer Bedeutung oder mehrere Bedeutungen zu einer Äußerung integrieren.

Aktuell arbeiten wir an einer Erweiterung von PNTs um Gewichte zur Modellierung von Wahrscheinlichkeiten. Daran anschließende Schritte beinhalten die Entwicklung effizienter Algorithmen zur Optimierung von PNTs und zur Berechnung der N besten Bedeutungen.

Literatur

- [1] R. Hoffmann, M. Eichner, and M. Wolff. Analysis of verbal and nonverbal acoustic signals with the Dresden UASR system. In *Verbal and Nonverbal Communication Behaviours*, volume 4775 of *LNAI*, pages 200–218. Springer, 2007.
- [2] M. Huber, C. Kölbl, R. Lorenz, R. Römer, and G. Wirsching. Semantische Dialogmodellierung mit gewichteten Merkmal-Werte-Relationen. In *Proceedings of "Elektronische Sprachsignalverarbeitung (ESSV)"*, volume 53 of *Studientexte zur Sprachkommunikation*, pages 25–32, 2009.
- [3] K. Karnagel. Semantische Modellierung von System-Benutzer-Interaktionen am Beispiel einer Home-Entertainment-Anwendung. Masterarbeit, 2012.
- [4] R. Lorenz and M. Huber. Petri Net Transducers in Semantic Dialogue Modelling. In *Proceedings of "Elektronische Sprachsignalverarbeitung (ESSV)"*, volume 64 of *Studientexte zur Sprachkommunikation*, pages 286–297, 2012.
- [5] R. Lorenz and M. Huber. Towards a Theory of weighted Petri Net Transducers. In *Applications and Theory of Petri Nets, 34th International Conference, Proceedings*. submitted, 2013.
- [6] G. Wirsching. Nichtsequentialität in der Sprachverarbeitung mit FST. In *Proceedings of "Elektronische Sprachsignalverarbeitung (ESSV)"*, volume 64 of *Studientexte zur Sprachkommunikation*, pages 26–33, 2012.
- [7] G. Wirsching, M. Huber, C. Kölbl, R. Lorenz, and R. Römer. *Semantic Dialogue Modeling*, volume 7403 of *LNCS*, pages 104 – 113. Springer, 2012.
- [8] G. Wirsching and C. Kölbl. Language Modeling with Utterance-Meaning-Pairs. Technical Report 2011-12, Institute of Computer Science, University of Augsburg, 2011.
- [9] M. Wolff, R. Römer, and R. Hoffmann. Hierarchische kognitive dynamische Systeme zur Sprach- und Signalverarbeitung. In *Proceedings of "Elektronische Sprachsignalverarbeitung (ESSV)"*, volume 64 of *Studientexte zur Sprachkommunikation*, pages 96–103, 2012.