

Einführung in die Grammatikentwicklung: Übung 8 (17. April 2020)

Ziele:

1. Reduktion der Übergenerierung der Grammatik
2. Einführung eines neuen Valenzmerkmals für Nomina
3. Einteilung der Lexeme in kopffinale und kopfinitiale und entsprechende Beschränkung der Regeln

Übungen:

In Übung 4 haben Sie gesehen, dass es für den Satz *Hilft der Hund dem Hund neben der Katze?* vier statt der erwarteten zwei Lesarten gibt. Das liegt daran, dass das Auftreten der Präposition *neben* rechts der regierten Nominalphrase durch die Grammatik nicht ausgeschlossen wird. In der letzten Übung haben Sie gesehen, dass für die Bedeutungsrepräsentation des Satzes (1a) u. a. die Wortfolge in (1b) erzeugt wird.

- (1) a. Jagt die Katze den Hund?
b. *Jagt Hund der Katze die?

1. Führen Sie analog zu dem Valenzmerkmal SUBCAT ein neues Valenzmerkmal SPR (für Spezifikator) ein. Das SPR-Merkmal wird in Lexikoneintrag von Nomina benutzt, um Determinierer zu selgeieren. Der Lexikoneintrag für *Hund* hat dann die Form:

HEAD	<i>noun</i>
SPR	⟨ DET ⟩
SUBCAT	⟨ ⟩
...	

Damit die Grammatik funktioniert, müssen Sie nun die Spezifikator-Kopf-Regel zur Grammatik hinzufügen. Diese Regel ist analog zur Komplement-Kopf-Regel und kombiniert einen Kopf mit dem Element, das der Kopf in seiner SPR-Liste enthält.

2. Mit dieser Spezialregel ist es jetzt möglich, Köpfe in solche mit Initialstellung und solche mit Finalstellung zu unterteilen: Präpositionen und Nomina gehen ihren Komplementen voran, Adjektive folgen ihren Komplementen und Verben können in Initial- oder Finalstellung stehen.

Führen sie das boolsche Merkmal INITIAL beim Typ *pos* ein. Spezifizieren Sie die Werte für die Untertypen von *pos* entsprechend.

3. Schränken Sie die Verwendung der Grammatikregel so ein, dass in den Kopf-Argument-Regeln mit Kopf-finalstellung auch nur entsprechend spezifizierte Köpfe als Kopftochter zugelassen werden, d. h., Präpositionen wie *an* sollen als Kopftochter in der *argument-head-rule* ausgeschlossen und nur in der *head-argument-rule* zugelassen werden. Wenn Sie das richtig gemacht haben, werden Sätze wie *Die Katze jagt den Hund* nicht mehr verarbeitbar sein. Das ist erwünscht, da solche Sätze mittels Fernabhängigkeiten zu analysieren sind. Fernabhängigkeiten werden in der nächsten Übung eingeführt.

4. Wenn Sie die Sätze mit der Semantik von *Jagt der kluge Hund die Katze?* generieren, werden Sie feststellen, dass auch die Wortfolge *Jagt der Hund kluge die Katze* erzeugt wird. Führen Sie analog zum Merkmal INITIAL ein boolsches Merkmal PRE_MODIFIER ein, das bei Modifikatoren, die vor ihrem Kopf stehen müssen, den Wert + und bei Modifikatoren, die nach dem Kopf angeordnet werden müssen, den Wert – hat. Passen Sie die Grammatikregeln für Adjunktion entsprechend an.

Überlegen Sie, warum man für den Ausschluß der falschen Modifikatorenstellung nicht das INITIAL-Merkmal verwenden kann.

5. Wenn Sie Sätze mit der Bedeutung von (1a) generieren, werden Sie unter anderem den folgenden Satz bekommen:

- (2) Jagen die Katzen die Hunde?

Dieser Satz ist nicht ungrammatisch, bedeutet aber etwas anderes als der Ursprungssatz. (2) wird generiert, weil die Morphologiekomponente neben den Singularformen auch Pluralformen liefert, ohne dass zwischen Singular und Plural semantisch unterschieden würde. Dieser Fehler kann behoben werden, indem in jedem referentiellen Index (der Merkmalstruktur unter SEM|INDEX) Numerus-Information repräsentiert wird. Information über Person, Numerus und Genus ist wichtig für die Bindung von Pronomina.

- (3) a. Der Mann hat ein Buch gekauft. Er liest es.
b. Die Männer kamen in den Schuppen. Sie lachten laut.
c. Ich habe einen Stuhl gekauft. Er gefällt mir.

In den Sätzen in (3) ist klar, worauf sich die Pronomina beziehen, da Genus-, Numerus- und Person-Information bei der Auflösung der Referenz helfen. Da die Indexinformation der Nominalphrasen mit den Argumentstellen des Verbs *jagt* identifiziert wird, ist diese Information in der semantischen Repräsentation für (1a) enthalten, Sätze mit anderen Index-Werten wie z. B. (2) werden deshalb nicht mehr generiert.

- (a) Die Signatur der Grammatik 7 ist bereits so organisiert, dass Merkmalstrukturen vom Typ *object* die Merkmale PER, NUM und GEN haben. Sorgen Sie dafür, dass in den Flexionslexikonregeln für Adjektive und Nomina die entsprechende Information nicht nur mit der Kongruenzinformation der entsprechenden Elemente sondern auch mit der Indexinformation identifiziert wird.

6. Ihre Grammatik sollte jetzt für (1a) nur noch die vier Sätze in (4) erzeugen?

- (4) a. Jagte der Hund die Katze?
b. Jagt der Hund die Katze?
c. die Katze der hund jagte
d. die Katze der hund jagt

In (4) sind Sätze mit Verberst- und Verbletzstellung enthalten, und wir werden in der nächsten Übung lernen, wie sich diese Sätze semantisch unterscheiden. Jetzt soll der Satz (4b) ausgeschlossen werden, der in seiner Zeitform nicht der Eingabe entspricht. Die falsche Lesart kann man parallel zum Ausschluß der Pluralformen unterdrücken: Indizes, die für Ereignisse stehen, bekommen ein Merkmal TEMO, das für Tense/Mode steht und Tempus- und Modus-Information enthält. Diese Information wurde in einer gemeinsamen Typhierarchie repräsentiert, um in der Morphologie Disjunktionen zu vermeiden.

- (a) Bisher wird das Merkmal TEMO beim Typ *fin_verb_i_suffix* eingeführt. Führen Sie einen neuen Typ *temo_intro* ein, der das Merkmal TEMO einführt, und lassen Sie *event* und *fin_verb_i_suffix* von diesem Typ erben. (Ein Merkmal darf nie an zwei verschiedenen Typen eingeführt werden. Braucht man das Merkmal doch in verschiedenen Bereichen einer Struktur, behilft man sich mit einem Hilfstyp wie *temo_intro*. Eine sinnvolle Konvention für die Benennung solcher Typen ist, ihnen das Suffix ‘_intro’ anzuhängen. Bei der Ausgabe von Typhierarchien kann man diese Hilfstypen dann ausblenden. Trale|Draw|Show Intro Types)
- (b) Schauen Sie sich einen Lexikoneintrag für ein finites Verb an und erkunden Sie, wo die Tense/Mode-Information enthalten ist. Sorgen Sie dafür, dass in der Lexikonregel *verb_inf1_1r* die Tense/Mode-Information mit dem Merkmal im Index des Verbs identifiziert wird. In der Analyse von (1a) sollte jetzt ein TEMO-Wert *pres_ind* enthalten sein. Generiert man von der entsprechenden Kante, sollten nur noch zwei Wortfolgen ausgegeben werden.