

7. Übungsblatt - Abgabe: 17.12.2013

Aufgabe 7.1 - WordNet

Lesen Sie Kapitel 19 aus Jurafsky & Martin von 19.1 bis einschließlich 19.3. (Der Text steht als Kopiervorlage im Vorlesungsordner, wenn Sie direkt aus dem Buch kopieren: benutzen Sie bitte die 2. Auflage.) Beantworten Sie mit Hilfe des Texts die folgenden Fragen:

- (a) Was versteht man unter einem Synset? Erläutern Sie an einem deutschen Beispiel.
- (b) Tabelle 19.2 gibt die WordNet-Relationen für Nomen an. Geben Sie für jede dieser Relationen jeweils ein Beispiel, bei dem der Begriff „Hund“ in seiner „Hauptlesart“ auf der linken Seite der Relation steht oder begründen Sie kurz, warum es für die Relation kein geeignetes Beispiel mit diesem Wort gibt. (Ignorieren Sie den letzten Eintrag „Derivationally Related Form“.)
- (c) Überlegen Sie sich drei Beispiele für unterschiedliche Arten von Relationen (und benennen Sie diese), die in WordNet nicht vorkommen, aber prinzipiell wünschenswert oder denkbar wären und geben Sie jeweils ein deutsches Beispiel für eine Instanz dieser Relation.

Aufgabe 7.2 - Semantische Relationen

Die folgenden Daten stammen aus einem Assoziationsexperiment, das in Saarbrücken durchgeführt wurde: Versuchspersonen wurde ein Wort präsentiert (Triggerwort), und sie sollten semantisch ähnliche Wörter dazu aufzählen. Im folgenden sind drei Triggerwörter, jeweils mit 5 häufig genannten assoziierten Wörtern, angegeben:

- **Fahrrad:** Fahrzeug, Sport, Rad, Pedal, Moped
 - **Erdbeere:** Kuchen, Frucht, Sahne, Garten, süß
 - **Buch:** Seite, Roman, Autor, Bibliothek, lesen
- (a) Beurteilen Sie für alle Paare von Triggerwort und Assoziation, ob die zwischen ihnen bestehende semantische Relation durch eine der WordNet-Relationen (Hyponymie/Hyperonymie, Meronymie/Holonymie, Antonymie) plausibel beschrieben wird.
 - (b) Geben Sie in den Fällen, in denen dies Ihrer Ansicht nach der Fall ist, die Relation an (z.B.: „Fahrrad – Fahrzeug: Hyponymie“) und versuchen Sie die Beziehung in der WordNet-Datenbank aufzufinden. Problem ist, dass es kein deutsches WordNet mit einer frei zugänglichen Web-Schnittstelle gibt. Nehmen Sie deswegen den Umweg über das Englische: Passende englische Übersetzung wählen und angeben (ggf.

mithilfe von Wörterbuch oder LEO) und im englischen WordNet (<http://wordnetweb.princeton.edu/perl/webwn>) recherchieren. Geben Sie die Resultate der Recherche an.

- (c) Versuchen Sie in den Fällen, in denen Ihrer Ansicht nach keine WordNet-Relation vorliegt, die semantische Beziehung mit eigenen Worten zu charakterisieren.

Aufgabe 7.3 - WordNet

Carnivores are a group of mammals that includes dogs, bears and cats. Cats have specialized teeth, sharp claws and a variety of coat patterns that enable them to blend in with their surroundings. They include tigers, lions and many other species. Some carnivores form social groups that facilitate hunting and the rearing of young. Lions, e.g. form prides which consist of several related females, their young and a few males.

Das WordNet 3.1-Webinterface (<http://wordnetweb.princeton.edu/perl/webwn>) erlaubt es, die komplette Hypernym-Kette (d.h. alle Oberbegriffe, unter „inherited hypernym“) für einen gegebenen Begriff abzufragen. Geben Sie jedes Substantiv aus dem obigen Text ein, wählen Sie jeweils die im Kontext des Dokuments angemessene Lesart aus (und geben Sie diese weils explizit an), und fragen Sie die Oberbegriffe (Hypernyme) ab. Konstruieren Sie den (möglicherweise nicht zusammenhängenden) Graphen, der alle markierten Wörter enthält. Sie müssen nicht jeden einzelnen WordNet- Oberbegriff in Ihre Darstellung übernehmen: Sie können sich auf die – strukturell oder inhaltlich – wichtigeren beschränken.

Feedback

Bitte beantworten Sie kurz folgende Fragen:

- (a) Wie schwer fanden Sie den Stoff der aktuellen Vorlesung? (sehr einfach, einfach, durchschnittlich, schwer, sehr schwer)
- (b) Wie viel Zeit haben Sie mit der Bearbeitung des aktuellen Übungsblattes verbracht?
- (c) Haben Sie sonstige Fragen oder Anmerkungen zur aktuellen Vorlesung? Wenn ja, welche?
-

Abgabe in Gruppen von bis zu drei Studierenden bis **17.12.2013** 10 Uhr entweder als Email- Anhang im pdf-Format an i2cl@coli.uni-sb.de mit dem Betreff **i2cl homework 8** und Ihrer Matrikelnummer oder auf Papier im Briefkasten an der Tür von Raum 1.04 in C7.2.