

1. Übungsblatt (Ersatz) - Abgabe: 31.10.2011

Dies ist das Ersatzübungsblatt, falls Sie nicht in der Lage sind, den Aufzug auszutesten. Es unterscheidet sich nur in Aufgabe 1.2 von dem am Dienstag herausgegebenen Übungsblatt.

Aufgabe 1.1 Maschinelle Übersetzung

- (a) Auf den Webseiten der Computerlinguistik findet sich unter <http://www.coli.uni-saarland.de/page.php?id=whatis> ("What is computational linguistics?") ein Absatz mit der Überschrift "Machines can also help people communicate with each other". Kopieren Sie diesen Absatz (bis zu "... large amounts of text in foreign languages") in das Textfenster von Babelfish (z.B. unter <http://de.babelfish.yahoo.com/>), und lassen sie ihn ins Deutsche übersetzen. Wie würden Sie Qualität und Stil der Übersetzung generell kommentieren?
- (b) Korrigieren Sie den Text: Markieren Sie fehlerhafte oder stilistisch schlechte Übersetzungen, geben Sie an, welche Art von Fehler gemacht wurde (ähnlich, wie Sie das aus der Schule noch kennen: „Ausdruck“, „Satzbau“ etc.), und geben Sie bessere Übersetzungsalternativen an (für die beanstandeten Stellen oder für ganze Sätze).
- (c) Probieren Sie das System auch mit ein paar (mindestens 2) Texten von anderen Textsorten aus: News, Gedicht, persönliche Email. Wie ist das Ergebnis hier im Vergleich? Berichten Sie (mit Beispielen).

Aufgabe 1.2 Dialogsysteme

- (a) Rufen Sie das ANDI-Dialogsystem (0681 302 58121 <http://andiplus.coli.uni-saarland.de/>) an und notieren Sie sich Ihre Anrufer-Nummer, die Ihnen zu Beginn des Telefonats angesagt wird. Versuchen Sie, eine beliebige Verbindung zwischen zwei der unten genannten Haltestellen herauszufinden. (Natürlich können Sie auch beliebige andere Haltestellen im Kreis Saarbrücken erfragen, sofern sie Ihnen bekannt sind).

Mögliche Haltestellen: (Saarbrücken) Hauptbahnhof, (Saarbrücken) Ostbahnhof, Römerkastell, Saarcenter, Burbacher Markt, Rentrischer Straße, Ilsestraße, Kantstraße (Dudweiler), Dudoplatz (Dudweiler), Landessportschule, Universität (Saarbücken), Universität Mensa, Gehlenbergstraße (Dudweiler), Stadtbad, Hansahauss, Johanneskirche, Ludwigskirche, An der Trift, Haus der Zukunft

Rufen Sie den Fragebogen unter <http://andiplus.coli.uni-saarland.de/fragebogen.html> auf und füllen Sie ihn vollständig aus. Geben Sie bitte auch auf dem Übungsblatt ihre Anrufernummer(n) an.

- (b) Versuchen Sie, das Muster, das der Dialogführung des Systems zu Grunde liegt, möglichst übersichtlich aufzuschreiben (nicht unbedingt vollständig): Wie reagiert es in bestimmten Situationen auf bestimmte Benutzeräußerungen? Wie sieht dementsprechend der „Dialogplan“ des Systems aus?

Aufgabe 1.3 Informationssuche

- (a) Benutzen Sie Google (www.google.com), um die Antwort auf die Frage „Wer war US-Präsident als Barack Obama geboren wurde?“ bzw. “Who was American president when Barack Obama was born?” zu erhalten. Beschreiben Sie, wie Sie vorgegangen sind, und die Schwierigkeiten.
- (b) Finden Sie drei weitere Fragen, die aus möglichst unterschiedlichen Gründen nur schwer oder gar nicht mit Google zu finden sind, deren Antworten aber grundsätzlich im Internet vorkommen. Erläutern Sie, warum diese Fragen schwierig sind.

Aufgabe 1.4. Ambiguität

Lesen Sie aus Jurafsky/Martin (Neu-Ausgabe von 2009) Kapitel 1, S. 35-43, Abschnitte 1.1, 1.2, 1.4 und 1.5. Die in 1.3 verwendeten Begriffe sind nicht ohne weiteres verständlich, sie werden im Lauf der Vorlesung eingeführt; 1.3 kann übersprungen werden. Der Text steht in Kopie in der Institutsbibliothek (Vorlesungsordner).

- (a) Versuchen Sie, für die 5 Lesarten des Satzes “I made her duck” aus 1.2 möglichst eindeutige und gut verständliche deutsche Formulierungen zu finden.
- (b) Versuchen Sie, einen deutschen Satz mit maximal 4 Wörtern zu finden, der möglichst viele, aber mindestens 2 Lesarten besitzt. Paraphrasieren Sie die einzelnen Lesarten und erklären Sie, wodurch die Ambiguität zustande kommt.

Abgabe in Gruppen von bis zu drei Studierenden bis **31.10.2011** 18 Uhr entweder als Email im pdf-Format an i2cl@coli.uni-sb.de oder auf Papier im Briefkasten an der Tür von Raum 1.04 in C7.2.