

Liste von Vorschlägen fällt leichter als die buchstabenweise Generierung eines Textes. In gewisser Weise findet eine automatische Rechtschreibprüfung des Textes statt. Es wird berichtet, dass behinderte Personen mittels Textvorhersage qualitativ bessere Texte herstellen konnten, als bei herkömmlicher Eingabe.

## 7 Berechnung der maximal möglichen KSR

Zur Abschätzung der erzielbaren KSR führen wir als weitere Größe die durchschnittliche Wortlänge  $l_{ow}$  ein. Wenn jedes Wort sofort – das heißt noch bevor der erste Buchstabe eingegeben wurde – in der Vorschlagsliste aufscheint, kann es durch einen einzigen Tastendruck pro Wort ausgewählt werden. Dieses Idealmaß wird oft als (ideale) Benchmark für Textvorhersagesysteme angesehen. Es soll hier mit  $e_1$  bezeichnet werden (Index  $I$  für einen Anschlag / eine Selektion).

$$e_1 = 1 - \frac{1}{l_{ow} + 1}$$

Der Term  $(l_{ow} + 1)$  ist die durchschnittliche Wortlänge plus nachfolgendes Leerzeichen, das ja bei der Textvorhersage bei jedem ausgewählten Wort automatisch generiert wird.

Mit einer für die deutsche Sprache üblichen  $l_{ow}$  von 5 ergibt sich der Benchmarkwert zu:

$$e_1 = 1 - \frac{1}{5 + 1} = 0,83$$

bzw. die maximal theoretisch erreichbare KSR mit 83%.

Nach eigenen Abschätzungen sind aber 2-3 Anschläge der realistische Wert, womit sich die Leistung eines realen Textvorhersageprogramms zwischen den Werten für die  $e_2$  und  $e_3$  einstellen wird:

$$e_2 = 1 - \frac{2}{5 + 1} = 0,67$$

$$e_3 = 1 - \frac{3}{5 + 1} = 0,50$$

Mit anderen Worten werden reale Textvorhersageprogramme eine KSR zwischen 55% und 60% kaum überschreiten können. Mit diesen Werten ist jedoch das Fernziel unter Einbeziehung syntaktischer Zusammenhänge gesteckt. Eigene Messungen, die noch ohne Verwendung syntaktischer Methoden erfolgten, zeigten, dass durchschnittlich mit einem Anschlag mehr gerechnet werden muss. Die zu erwartende Leistung wird demnach zwischen:

$$e_3 = 1 - \frac{3}{5 + 1} = 0,50$$

$$e_4 = 1 - \frac{4}{5 + 1} = 0,33$$

also die KSR 33% bis 50% liegen.

## 8 Schlussfolgerung und Herausforderung

Die hier durchgeführte Abschätzung einer maximal möglichen bzw. realistisch zu erwartenden KSR und jenes Grenz-Einsparungsfaktors, ab dem kein Zeitgewinn mehr vorhanden ist, macht deutlich, dass für wirklich nützliche Textvorhersageprogramme alle Register gezogen werden müssen und dass neben einer effektive Programmierung der grundlegenden Algorithmen selbst großes Gewicht auf die Einbeziehung linguistischer Zusammenhänge und die Optimierung der Mensch-Maschine-Schnittstelle gelegt werden muss.

Um für die deutsche Sprache den Nachholbedarf geringer werden zu lassen, wurde unter Förderung der Oesterreichischen Nationalbank das Projekt PROPHET durchgeführt, in dem grundlegende Arbeiten zur Textvorhersage in der deutschen Sprache durchgeführt wurden. In Simulationen an verschiedenen Texten wurde der Einfluß zahlreicher Parameter (wie Größe und Verwaltung des Wörterbuches, Länge der Vorschlagsliste etc.) gemessen. Ohne grammatikalische oder syntaktische Beziehungen auszunützen, konnten KSR-Werte in der Größenordnung von 40 bis 45% erreicht werden. Bei Texten, deren Wortschatz durch das Wörterbuch vollständig oder zumindest fast vollständig abgedeckt wurde, stieg die KSR sogar bis weit über 50%.

Es konnte somit durch Simulation gezeigt werden, dass brauchbare KSR-Werte auch in der deutschen Sprache erreicht werden können, dass aber das Ziel, Werte um die 55% bis 60% zu erhalten, erst durch Auswertung syntaktischer Beziehungen und die zusätzliche Verwendung von „*Abbreviation Expansion*“ erreicht werden kann.

## 9 Danksagung und Kontakte

Das Projekt PROPHET – Deutschsprachiges Textvorhersageprogramm für kommunikationsbehinderte Menschen wurde vom Jubiläumsfond der OeNB unter der Nummer 6273 gefördert. Ab Oktober 1998 wird der Abschlussbericht vorliegen.

### fortec -

Forschungsgruppe für Rehabilitationstechnik  
Institut für Allgemeine Elektrotechnik und Elektronik  
Technische Universität Wien  
1040 Wien, Gußhausstraße 27/359-3B  
Tel: +43 (1) 504-1830,  
Fax: +43 (1) 504-1830 /12  
E-Mail: [zw@fortec.tuwien.ac.at](mailto:zw@fortec.tuwien.ac.at)  
<http://fortec.iaee.tuwien.ac.at/fortec>

# Filmcomputer

gesammelt von Teemu Schabl

Dinge, die Computer in Filmen können:

- 1 Textverarbeitungsprogramme zeigen niemals einen Cursor
- 2 Man muss niemals die Leertaste benutzen, um lange Sätze einzugeben.
- 3 Filmcharaktere machen niemals Tippfehler.
- 4 Alle Bildschirme zeigen 5cm-Buchstaben
- 5 High-Tech-Computer, wie die von der NASA, dem CIA oder solchen Regierungsinstitutionen, haben einfach zu bedienende grafische Oberflächen.
- 6 Diejenigen, die keine grafische Oberfläche haben, haben unglaublich gute textbasierte Eingabeaufforderungen, die englische Sätze korrekt verstehen und ausführen.
- 7 Kommandozeileninterpreter geben einem Zugriff zu jeglicher Information, die man benötigt, wenn man einfach "ACCESS THE SECRET FILES" auf der nächsten Tastatur eingibt.
- 8 Man kann ebenfalls einen Computer mit einem zerstörerischen Virus infizieren, indem man einfach "UPLOAD VIRUS" eingibt (->"Fortress").
- 9 Alle Computer sind verbunden. Man kann auf Informationen auf dem Desktop-Computer des Bösewichts zugreifen, selbst wenn dieser ausgeschaltet ist.
- 10 Leistungsstarke Computer piepen jedesmal, wenn man eine Taste drückt oder der Bildschirm sich ändert. Einige Computer verlangsamen sogar die Ausgabe auf dem Bildschirm, damit es nicht schneller geht als man lesen kann (wirklich hervorragende Computer emulieren dazu noch das Geräusch eines Nadeldruckers).
- 11 Alle Computerpanels arbeiten mit Tausenden Volt und haben explosive Geräte direkt unter ihrer Oberfläche. Fehlfunktionen werden von einem hellen Lichtblitz, einer Rauchwolke, einem Funkenschauer, und einer Explosion, die dich wegschleudert, angezeigt.
- 12 Leute, die etwas auf einem Computer eingeben, können ihn ohne Datenverlust sofort ausschalten.
- 13 Ein Hacker kann jederzeit in die gesicherten Computer der Welt einbrechen, indem sie das geheime Passwort mit zwei Versuchen erraten. *weiter auf Seite 43* ➤