

AUSSCHREIBUNG DER 1. MORPHOLYMPICS FÜR AUTOMATISCHE WORTFORMERKENNUNG

Roland Hausser

7. und 8. März 1994, Universität Erlangen-Nürnberg

Die *Morpholympics* ist ein neuartiger Typ von Tagung, auf dem sich unterschiedliche Systeme zur automatischen Wortformerkennung in einem öffentlichen Wettlauf messen. Der Wettbewerb wird von einer unparteilichen Jury aus 5 Preisrichtern beobachtet, die die Sieger bestimmt.

Im Folgenden werden Ablauf der Morpholympics und Teilnahme-modalitäten genauer erläutert. Eine analoge Veranstaltung für syntaktische Parser ist von der GLDV unter dem Namen *Parsolympics* für das Jahr 1995 geplant.

1 Organisation und Ablauf

1.1 Allgemeine organisatorische Struktur der Morpholympics

1.1.1 Zielsetzung

Ziel der Morpholympics ist ein objektiver, theorieunabhängiger Vergleich existierender Systeme zur automatischen Wortformerkennung. Dabei sollen linguistische Motivation, technische Konzeption, Datenabdeckung und Geschwindigkeit bewertet werden. Für den praktischen Vergleich der Systeme ist es sinnvoll, Testdaten aus einer natürlichen Sprache in den Mittelpunkt zu

stellen. Dies ist die sogenannte *Haupttest-sprache*.

1.1.2 Turnusmäßige Gestaltung

Die Morpholympics werden, je nach Bedarf, alle zwei bis vier Jahre abgehalten. Dabei sollen die Systeme an einer jeweils neuen natürlichen Sprache (*Haupttestsprache*) getestet werden. Tagungssprachen sind Englisch und die jeweilige Haupttestsprache.

Um die Weiterentwicklung bereits getesteter Systeme und Sprachen zu fördern, werden auf einer gegebenen Morpholympics auch die Sprachen früherer Wettkämpfe als *Nebentestsprachen* zugelassen. Auf diese Weise kann der *defending champion* einer ehemaligen Haupttestsprache herausgefordert werden. Voraussetzung für die Aufnahme einer Sprache als Nebentestsprache ist, daß sich mindestens 3 Teilnehmer für diese Sprache melden.

Wenn nach mehreren Morpholympics das Interesse an einer ehemaligen Haupttestsprache wieder genügend angewachsen ist, kann diese zur neuen Haupttestsprache bestimmt werden.

1.1.3 Koordination

Die Verantwortung für die allgemeine Planung, Koordination und anschließende Publikation einer Morpholympics liegt in den Händen eines KOORDINATORS. Der Koordinator soll ein in Computerlinguistik aus

gewiesener Wissenschaftler sein, der gleichzeitig Repräsentant einer nationalen oder internationalen computerlinguistischen Organisation (Trägerorganisation) ist. Der Koordinator sorgt für die Platzierung von Austragungsort und -zeitpunkt, gewinnt die Zusage von fünf kompetenten und neutralen Preisrichtern und trägt die Gesamtverantwortung für den reibungslosen Ablauf. Der Koordinator kann nicht gleichzeitig als Preisrichter fungieren.

1.1.4 Initiierung der Nachfolge Morpholympics

Der Arbeitskreis *Parsing in Morphologie und Syntax* der GLDV bestimmt in Absprache mit dem Koordinator und der Trägerorganisation der letzten Morpholympics den Koordinator der nächsten Morpholympics.

1.1.5 Gestaltung der Preise

Der Preis für den Sieger in der Haupttestsprache einer Morpholympics besteht in einer Urkunde, einer Trophäe und einem Preisgeld. Die Trophäe wird zu Beginn der nächsten Morpholympics dem dann aktiven Koordinator zugestellt und erneut vergeben (Wanderpreis). Abhängig von der Zahl der Teilnehmer ist es möglich, einen zweiten und dritten Preis in der Haupttestsprache, sowie jeweils einen Preis für jede der Nebentestsprachen in Form von Urkunden zu vergeben. Die Preisgelder werden vom Koordinator von industriellen Sponsoren eingeworben.

1.1.6 Zeitlicher Ablauf

Termin und Ort einer Morpholympics, sowie die Haupt- und Nebentestsprachen, werden in der Regel mindestens drei Monate vor der Veranstaltung bekannt gegeben. Zwei Monate vor der Veranstaltung haben potentielle Teilnehmer Zugang zu einem Fragebogen (siehe 2, 'Standardisierte Darstellung der Systeme'), den vorläufigen

Testdaten (siehe 3.1) und dem Anmeldeformular (siehe 4). Um an der Tagung teilzunehmen, muß das ausgefüllte Anmeldeformular bis zu dem vom Koordinator genannten Anmeldeschluß eingereicht werden.

1.1.7 Form der 'Wettläufe'

Vor Beginn der Tagung können Teilnehmer auf einem geeigneten Unix-Rechner am Veranstaltungsort einen 10gin erhalten und haben damit Gelegenheit, ihre Systeme per rlogin über das Netz zu installieren. Für die Installation von MAC- und PC- Versionen haben Teilnehmer zwei Tage vor der Veranstaltung Zugang zu den Rechnern am Veranstaltungsort (können aber auch ihre eigenen Geräte mitbringen).

Die offiziell angemeldeten Teilnehmer treffen sich zwischen 9:00 und 9:50 am Morgen des ersten Konferenztages. Bei dieser Gelegenheit wird der ausgefüllte Fragebogen in 6-facher Ausführung bei den Veranstaltern abgegeben (fünf für die Preisrichter, einer für die abschließende Publikation) und die Reihenfolge der 'Wettläufe' durch das Los bestimmt. Jedes System wird durch maximal zwei offizielle Teilnehmer repräsentiert.

Der 'Wettlauf' eines Systems besteht aus

I> einer PRÄSENTATION MIT PRIMÄREM TESTLAUF und

I> zwei SEKUNDÄREN TESTLÄUFEN

Die Präsentation eines Systems dauert 50 Minuten. Zu Beginn der Präsentation erhalten die Teilnehmer die endgültigen Testdaten (siehe 3.3) *on line* und starten ihr System auf einem Unix-Rechner. Dann stellen sie ihr System auf der Grundlage des eingereichten Fragebogens in Form eines Vortrags mit Folien dar. Die Ergebnisse der Tests auf dem Unix-Rechner (primärer Testlauf mit Zahl der erkannten Wortformen, Geschwindigkeit, Speicherbedarf und Dauer des *turn around*, siehe 3.2) werden per LCD-Display während der Präsentation bekannt gegeben und von den Veranstaltern als *hardcopy an* die Preisrichter verteilt.

Nach der Präsentation eines Systems, und parallel zur fortlaufenden Hauptveranstaltung, werden in einem separaten Raum sekundäre Tests unter Aufsicht unabhängiger Protokollführer durchgeführt. In den sekundären Tests werden die endgültigen Testdaten noch einmal auf PC und MAC analysiert. Der Sinn der sekundären Tests ist, die Portierbarkeit des jeweiligen Systems und die Verwendbarkeit unter kommerziellen Bedingungen zu demonstrieren. Die Ergebnisse der sekundären Tests werden den Preisrichtern von den Protokollführern als *hardcopy* nachgereicht.

1.1.8 Grundlagen der Bewertung

Die Entscheidung der Preisrichter basiert auf

1. der standardisierten schriftlichen Darstellung des Systems (2) und - falls vorhanden - Exemplaren der Dokumentation,
2. dem mündlichen Vortrag und
3. den Ergebnissen der endgültigen (primären und sekundären) Tests.

1.1.9 Bestimmung des Siegers, Preisvergabe und Publikation

Nach Abschluß der letzten Präsentation und des letzten sekundären Tests ziehen sich die Preisrichter mit ihren Unterlagen einzeln für ca. 90 Minuten zurück, um ihre Entscheidungen zu treffen, bzw. noch einmal zu überdenken. Anschließend treffen sich die Preisrichter, der Koordinator, Vorstandsmitglieder seiner Organisation und die Vorstandsmitglieder der GLDV für etwa 30 Minuten. Bei dieser Sitzung teilen die Preisrichter ihre Entscheidungen mit und bestimmen die Sieger.

Anschließend werden die Ergebnisse den Teilnehmern und dem Publikum mitgeteilt und die Preise vergeben. Mit ihrer Anmeldung zur Morpholympics haben die Teilnehmer ihr Einverständnis für eine Publika-

tion ihrer Systembeschreibung und der aktuellen Testergebnisse gegeben (siehe 4.2). Es ist die Aufgabe des Koordinators, daß diese baldmöglichst als Sammelband (oder Sonderheft einer Zeitschrift) einem breiteren Publikum bekannt gemacht werden.

1.2 Spezifika der 1. Morpholympics

1.2.1 Koordinator

Koordinator der 1. Morpholympics ist Roland Hausser als Vorstandsmitglied der GLDV und Urheber dieses Veranstaltungstyps.

1.2.2 Termine

Anmeldeschluß: 20. Februar 1994

Veranstaltungsdatum der 1. Morpholympics (1994):

Montag, den 7. März und Dienstag, den 8. März, 1994

1.2.3 Veranstaltungsort:

Abteilung Computerlinguistik,
Universität Erlangen-Nürnberg,
Bismarckstr. 12, 91054 Erlangen.

Kommentare, Fragen, etc. bitte E-Mail an:
rrh@linguistik.uni-erlangen.de (Roland Hausser)

Für remote logins und technische Hilfe bitte E-Mail an: rolf@linguistik.uni-erlangen.de (Rolf Haberrecker)

1.2.4 Testsprache(n)

Die Haupttestsprache der 1. Morpholympics ist Deutsch. Nebentestsprachen können beim Koordinator bis zum 6. Januar 1994 beantragt werden.

1.2.5 Preisrichter der 1. Morpholympics 1994

Prof. Dr. I. S. Batori
Universität Koblenz-Landau

Prof. Dr. Grzegorz Dogil

Universität Stuttgart

Prof. Dr. Günther Görz

Universität Erlangen-Nürnberg

Prof. Dr. Winfried Lenders

Universität Bonn

Prof. Dr. Wolfgang Wahlster

Universität Saarbrücken

2 Standardisierte Selbstdarstellung

Die folgenden Fragen zur linguistischen Motivation, technischen Konzeption, Datenabdeckung und Geschwindigkeit der teilnehmenden Systeme soll in einer standardisierten Selbstbeschreibung der zu testenden Systeme resultieren. Sie dient den Preisrichtern als schriftliche Unterlage, die bei der Bewertung mit in Betracht gezogen wird, und ist Grundlage der mündlichen 'Präsentation' (siehe 1.1. 7).

2.1 Name und Herkunft des angemeldeten Systems:

2.2 Konzeptuelle Kriterien:

2.2.1 Deklarative Spezifikation lexikalischer Einträge und Regeln

Stellen Sie bitte die Struktur der lexikalischen Einträge und der Regel(type)n der Morphologie-Grammatik schematisch dar. Geben Sie bitte zu jedem Schema ein realistisches (unediertes) Beispiel.

2.2.2 Bezug zwischen lexikalischen Einträgen und Wortformen

Zeigen Sie bitte anhand einer schematischen Ableitungsskizze, wie in Ihrem System das Verhältnis von Allomorphie und Konkatenation behandelt wird. Skizzieren Sie bitte den Bezug zwischen lexikalischem Eintrag und den abgeleiteten Wortformen *anlächle/anlächeln* in Ihrem System.

2.2.3 Verständlichkeit und linguistische Motivation der Regeln

Beschreiben Sie bitte, wie in Ihrem System die morphologischen Prozesse der Flexion, Derivation und Komposition behandelt werden. Illustrieren Sie die Handhabung dieser Prozesse an der Ableitung der Formen

Tisch, Tisches, Tischen
vorbeischwammst,
vorbeischwämme,
vorbeigeschwommenen
Hausdächern, Häusermeers
Unabhängigkeitserklärung
unlesbares
durchdachte
gut, besser, besten

aus den Lexikoneinträgen Ihres Systems. Zitieren Sie die hierbei verwendeten Regeln und erläutern Sie ihre Funktionsweise.

2.2.4 Morpho-syntaktische Analyse (Kategorisierung)

Erläutern Sie bitte anhand der oben analysierten Beispiele, daß Ihr System Wortformen nach traditionellen Kriterien wie Genus, Numerus, Kasus, (Verb-)Modus, Tempus, Person, Komparation etc. charakterisiert. Werden Morphemgrenzen dargestellt (Segmentierung) und/oder syntaktisch relevante Eigenschaften wie Valenzrahmen berücksichtigt? Wie schätzen Sie die Möglichkeit ein, Ihr System im Rahmen unterschiedlicher Syntaxsysteme zu verwenden (mithilfe einer automatischen Transduktion Ihrer morpho-syntaktischen Analysen in die entsprechenden Formate)?

2.2.5 Behandlung der Generierung

Werden in Ihrem System dieselben Regeln für die Generierung verwendet wie für die Analyse? Illustrieren Sie Ihr System bitte an der Generierung der Paradigmen von *geben* und *Küchentisch*.

Welche Formen dienen bei der Generierung als Eingabe (z.B. Stamm oder Infinitiv beim Verb)? Wie wird die Generierung spezifischer Zielformen (z.B. 3. Person Singular Präsens von *lernen*) gehandhabt?

2.2.6 Übertragbarkeit auf andere Sprachen

Auf welche anderen Sprachen wurde Ihr System bisher angewendet? Gab es dort spezifische morphologische Phänomene (z.B. Vokalharmonie oder *intercalation* von Morphemen), für die sich Ihr System als besonders geeignet erwies?

2.3 Technische Konzeption und Einsatzfähigkeit

2.3.1 Zielsetzung der Konzeption

Gibt es Anwendungsaspekte, auf die bei Konzeption und Implementation Ihres Systems besonderer Wert gelegt wurde? Mögliche Aspekte sind z.B. Geschwindigkeit, Vollständigkeit der Datenabdeckung, Modellierung einer spezifischen linguistischen Theorie, Verwendbarkeit bei den verschiedensten Sprachen, geringer Bedarf an primärem oder sekundärem Speicher, Lauffähigkeit auf bestimmten Maschinen (z.B. PC) oder unter bestimmten Betriebssystemen, Unterstützung von akustischer Spracherkennung, Verwendung durch einen Syntaxparser und die Indizierung textueller Datenbanken. Wie haben sich Ihre speziellen Einsatzziele auf die Konzeption ausgewirkt, und inwiefern ist Ihr System für diese Einsatzziele besonders geeignet?

2.3.2 Portabilität der Software und der Daten

In welcher Programmiersprache/Version ist Ihr System bisher implementiert worden? Auf welchen Maschinen und mit welchen Betriebssystemen/Version haben Sie Ihr System bisher zum Laufen gebracht? Bei welchen Umgebungen gab es Probleme?

2.3.3 Schnittstellen zur Syntax und zur Semantik

Gibt es in Ihrem System Schnittstellen zur Syntax und zur Semantik? Wie sind sie spezifiziert und werden sie tatsächlich verwendet?

2.3.4 Hilfestellung bei Benutzerfehlern

Demonstrieren Sie bitte die Güte der Fehlermeldungen und des Debuggers in Ihrem System anhand der *trace* dreier aussagekräftiger Beispiele.

2.3.5 Größenbeschränkung des Systems

Welche Größenbeschränkungen gibt es in Ihrem System bzgl. der Speicherbelegung, des Lexikons oder der Länge der Eingabe?

2.3.6 Schnittstelle zu Nicht-ASCII Zeichen

Kann Ihr System Nicht-ASCII Zeichen (z.B. Sonderzeichen wie Umlaute, Akzente etc., oder nicht-lateinische Sprachzeichen wie Hangul) in der Eingabe verarbeiten und in der Ausgabe darstellen? Wie ist dies technisch gelöst?

2.3.7 Benutzerfreundlichkeit des *turn around*

Wieviele Schritte sind erforderlich, um an Ihrem System linguistisch-empirische Modifikationen vorzunehmen, und wie langwierig ist die damit einhergehende Kompilation? Demonstrieren Sie bitte einen solchen Vorgang mit einer *getimeten trace*. Fügen Sie bitte (im Druckbild deutlich markierte) erläuternde Kommentare ein, um dem Leser einen schnellen Überblick zu verschaffen.

Stellen Sie bitte auch dar, wie robust Ihr System auf unbekannte Wortformen reagiert und welche Hilfestellungen es bei der Eingabe neuer Wörter ins Lexikon gibt.

2.3.8 Transparenz und Vollständigkeit der Dokumentation

Welche Dokumentation gibt es zu Ihrem System? Umfaßt sie (a) eine Darstellung der zugrundeliegenden linguistischen Theorie mit algebraischer Definition und Komplexitätsanalyse, (b) ein Bedienungsmanual, daß es neuen Benutzern ermöglicht, Ihr System schrittweise zu erlernen und auf neue Sprachen anzuwenden und/oder (c) den kommentierten Quellcode? Wann wurde die Dokumentation das letzte Mal überarbeitet? Falls aktuelle Teile Ihrer Dokumentation veröffentlicht wurden, geben Sie bitte die bibliographischen Daten an. Bitte legen Sie Exemplare Ihrer Dokumentation zu den Fragebogen bei.

2.3.9 Verfügbarkeit und Wartung

Unter welchen Voraussetzungen und in welcher Form kann ein potentieller Benutzer Ihr System als laufendes Softwarepaket bekommen? An welchen Institutionen wurde Ihr System von wann bis wann für welche Sprachen benutzt? Wird das System von Ihnen z. Zt. aktiv gepflegt, bzw. weiterentwickelt? In welchen Anwendungen wird das System z. Zt. von Ihnen genutzt? An welche Personen (email-adresse) in Ihrem Team kann sich ein Benutzer im Fall von *bugs* und anderen Problemen wenden?

3 Testdaten und Fragebogen

3.1 Vorläufige Tests des Systems an bekannten Daten

Zwei Monate vor den Morpholympics stellen die Preisrichter zwei Sorten von Testdaten *on line* zur Verfügung:

I> eine Wortliste mit besonders interessanten morphologischen Phänomenen (inzwischen als 'Schmankerlnliste' bekannt)

I> einen repräsentativen Text ausreichen der Länge.

Diese Testdaten sollen den teilnehmenden Systemen zur Vorbereitung auf die Morpholympics und zur Berechnung vorläufiger *bench marks* dienen.

3.2 Portabilität

Um die Portabilität Ihres Systems zu zeigen, testen Sie bitte Ihr System an den in 3.1 genannten Daten, und zwar auf einer Unix-Workstation, einem PC und einem Apple-Macintosh. Für jeden der drei Tests beantworten Sie bitte im Rahmen dieses Fragebogens folgende Fragen:

Typ der Maschine, des Operating Systems und der Programmiersprache, sowie

I> Geschwindigkeit (Wortformen pro Sekunde)

I> Speicherbelegung

I> Zeit, die für einen *turn around benötigt* wird (siehe 2.2.7)

Bitte fügen Sie die Liste analysierter und nicht-analyzierter Wortformen als Anlage bei, wobei die Zahl der analysierten und der nicht-analyzierten Wortformen sowohl hier als auch am Anfang der Anlage genannt wird. Geben Sie hier bitte auch Geschwindigkeit, benötigten Speicherplatz und 'turn around'-Zeit für jeden der drei Maschinentests an.

3.3 Endgültige Tests des Systems an neuen Daten

Die in 3.2 beschriebenen Testläufe werden während der Morpholympics an neuen, aber ähnlichen Testdaten wiederholt. Dabei wird der Testlauf auf der Workstation während der Präsentation vorgeführt (primärer Testlauf). Die Testläufe auf dem PC und dem Mac (sekundärer Testläufe) werden nach der Präsentation durchgeführt (siehe 1.1. 7).

3.4 Auswahl und Verteilung der vorläufigen Testdaten

Potentielle Teilnehmer und Zuschauer bei der Morpholympics werden hiermit höflich gebeten, Wortformen oder Texte, die sie für relevant halten, an einen der Preisrichter, Herrn Prof. Lenders, unter lenders@uni-bonn.de zu schicken. Die Preisrichter werden diese Daten bei der Zusammenstellung der vorläufigen und endgültigen Testdaten in Betracht ziehen.

Die vorläufigen Testdaten können ab 6. Januar 1994 via ftp von

server name:
sol.linguistik.uni-erlangen.de
login name: anonymous
password: your user name
directory: morpholympics

abgerufen werden. Die Auswahl und Weitergabe der vorläufigen und endgültigen Testdaten liegt in der Verantwortlichkeit der Preisrichter.

3.5 *on line*-Abruf von Fragebogen und Anmeldeformular

Die vorliegende Darstellung von Ablauf und Teilnahmemodalitäten kann bereits jetzt in den Tagungssprachen Englisch und Deutsch von dem in 3.4 genannten Server abgerufen werden.

4 Anmeldeformular für die 1. MORPHOLYMPICS

4.1 Anmeldung des Systems

Das folgende System zur automatischen Wortformererkennung für *on line* Texte im Deutschen nimmt an dem Wettlaufen der 1. Morpholympics teil, das am 7. und 8. März 1994 in Erlangen von der Gesellschaft für Linguistische Datenverarbeitung (GLDV) veranstaltet wird.

Name und Herkunft des Systems:

Name und Adresse der Personen, die das System bei den Morpholympics vorstellen:

4.2 Einverständniserklärung für die Publikation

Wir werden 6 Kopien der standardisierten Beschreibung unseres Systems zu Beginn der Morpholympics beim Veranstalter einreichen. Wir sind damit einverstanden, daß diese Beschreibung vom Koordinator in der offiziellen Publikation über die Ergebnisse der Morpholympics veröffentlicht wird.

Unterschriften, Ort und Datum

