

Programmierkonzepte und Evolutionary Linguistics

Neue Ansätze für die maschinelle Übersetzung

ein Vortrag von
Martin Pyka
und
Johannes Schwall

Inhalte

- Sprache als Konstante
- Language Evolution als neues Paradigma
- Multi-agent Systeme und Language Games
 - Genereller Aufbau
 - Bsp: Talking Heads
 - Forschungsergebnisse
- Schlussfolgerungen der EL

Eigenschaften bisheriger Programme

- Regeln einer Sprache (Grammatik) sind konstant
- Algorithmen (generative Grammatik) sollen alle Sätze nachbilden bzw. verstehen können
- nur bedingt lernfähig

Beispiele

Verbmobil, SysTran

Erkenntnisse

- grosser Wartungsaufwand
- bisher nur sehr spezialisierte oder sehr grobe Simulation möglich
- die Frage nach dem korrekten Weg

Language Evolution als neues Paradigma

"A more robust technology can only come from understanding how language users construct and reconstruct their language as they adapt to the language spoken in their environment and try to keep up with the ever changing communicative challenges arising in their community."

Luc Steel

Aufgaben

- Betrachtung der historischen Evolution der Sprache
- Wechselwirkung zwischen Sprecher und Sprache untersuchen

Sprache ist kein definierbares Regelsystem

"Observation of real verbal behavior shows that language users all the time deviate from what could be considered ideal performance, expressable as a set of rules."

Luc Steel

- verschiedene Sprecher haben verschiedene Grammatiken
- die Performanz des Sprechers hängt von seiner Grammatik und von seiner sozialen, geographischen, historischen und interaktiven Rolle in der Kommunikation mit anderen ab

Wechselwirkung zwischen Sprache und Sprecher

- das individuelle Sprachverhalten bestimmt „die“ Sprache
- die Sprache beeinflusst das Verhalten des Individuums

Wechselwirkungen in der Natur

- Pheromonspuren von Ameisen zur Pfadfindung
- Schwärme

Grundgedanken der Language Evolution

- Sprache verändert sich
- natürliche Selektion im Sprachgebrauch bestimmen Konventionen
- Sprachevolution findet auf einer kulturellen (nicht genetischen) Ebene statt

⇒ eine Universalgrammatik ist dem Menschen nicht angeboren

Evolutionary Linguistics

- die Lehre von der Geschichte und den Ursprüngen von Sprache und ihrer Entwicklung
- Simulationen und künstliche Systeme sind ein wissenschaftliches Instrument zur Analyse sprachlicher Phänomene
- Erhofft sich Antworten auf die Frage nach dem Ursprung von Sprache, wie sich Sprache verändert, wie Kinder Sprache lernen

Agent

- Einheit mit internen Zuständen, die die Interaktion mit der Umgebung bestimmen (Grammatik, Sprachwissen im Allg.)
- Autonom (Lernmechanismen)
- Distributed (kein Einfluss auf die internen Zustände anderer Agenten)

Population

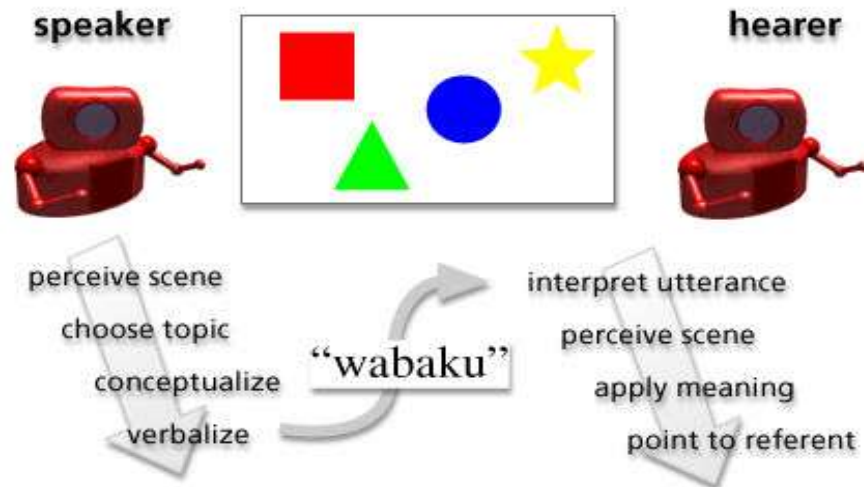
- Anzahl (fest oder flexibel) von Agenten
- Regeln für Nachkommenschaft
- eine vordefinierte Umgebung die optional Interaktivität zulässt (also veränderbar ist)

Simulation/Sprachspiele

- auf reiner Softwareebene (stark abstrahiert) oder durch den Einsatz von Robotik (physische Ebene)
- Entstehung von Konventionen bezügl. des Lexikons, der Grammatik, der Aussprache, der Bedeutung kann untersucht werden
- bisher werden nur einzelne Aspekte der Sprache untersucht
- Führt zu neuer Sprachtheorie, die die Population, Adaption und Evolution als Teil der Linguistik sieht

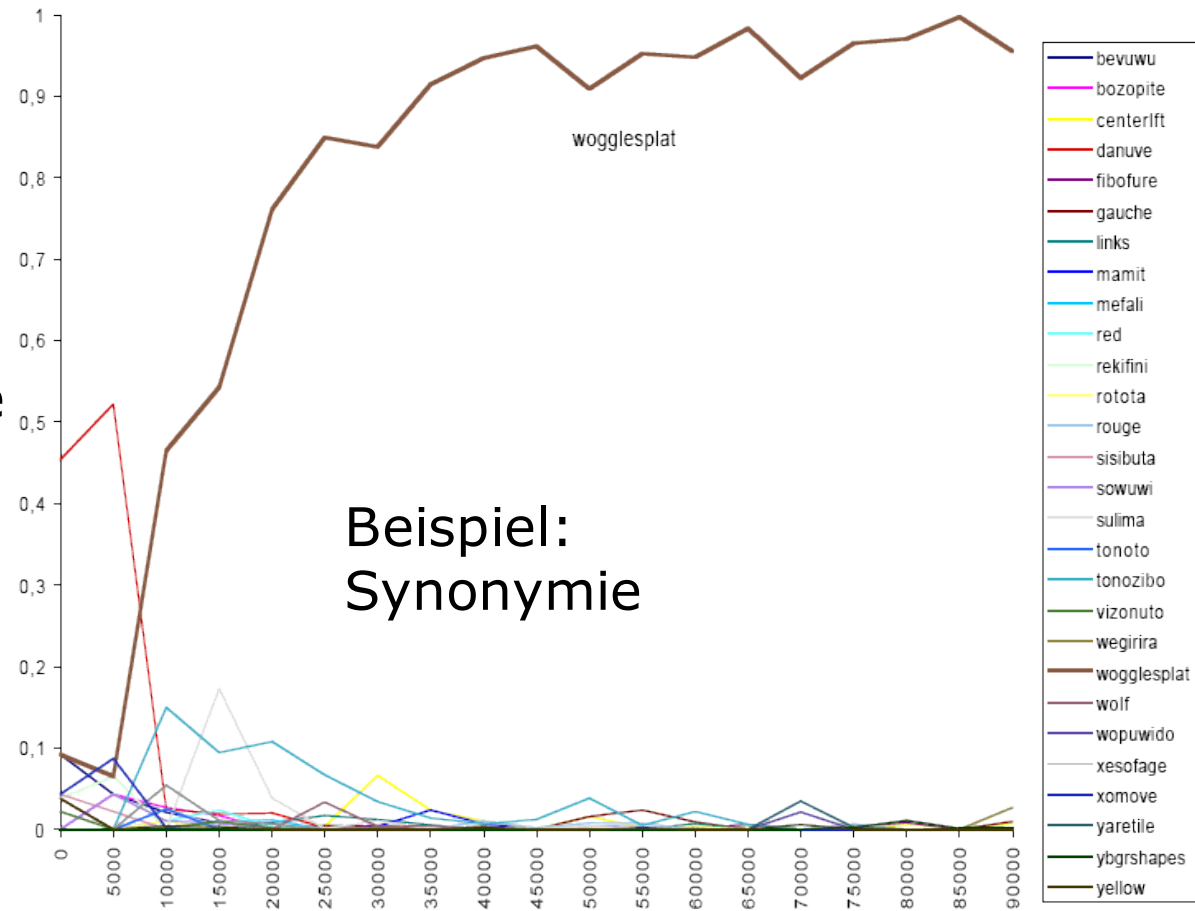
Beschreibung

- Agenten, die Rollen annehmen (Speaker, Hearer)
- Magnetsteine auf Whiteboard bilden den Kontext
- Agent kann kategorisieren
- Speaker benennt ein Objekt, Hearer muss es erraten

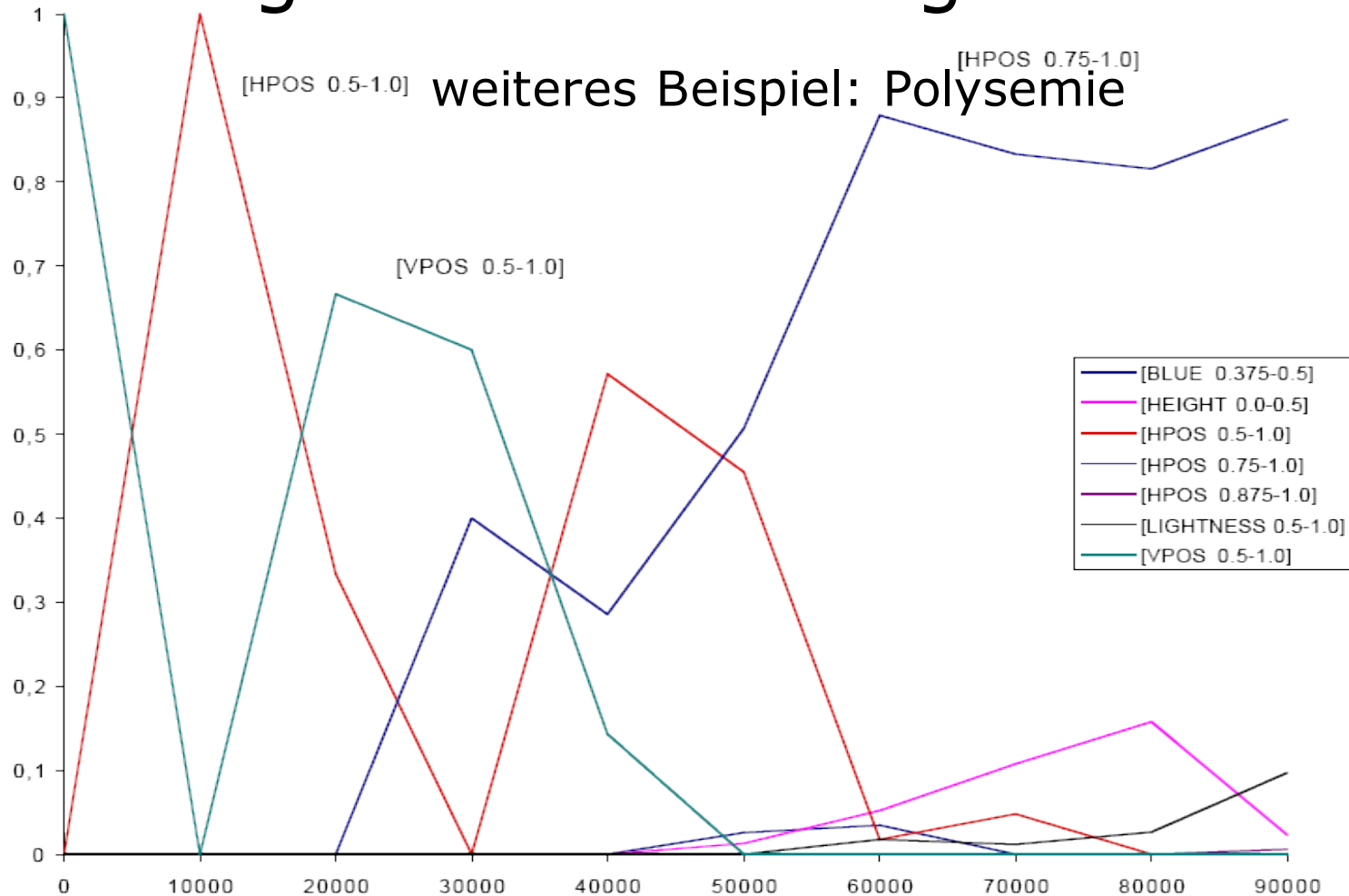


Reaching and maintaining coherence

- notwendig für eine erfolgreiche Kommunikation
- Auftreten typischer Phänomene der natürlichen Sprache

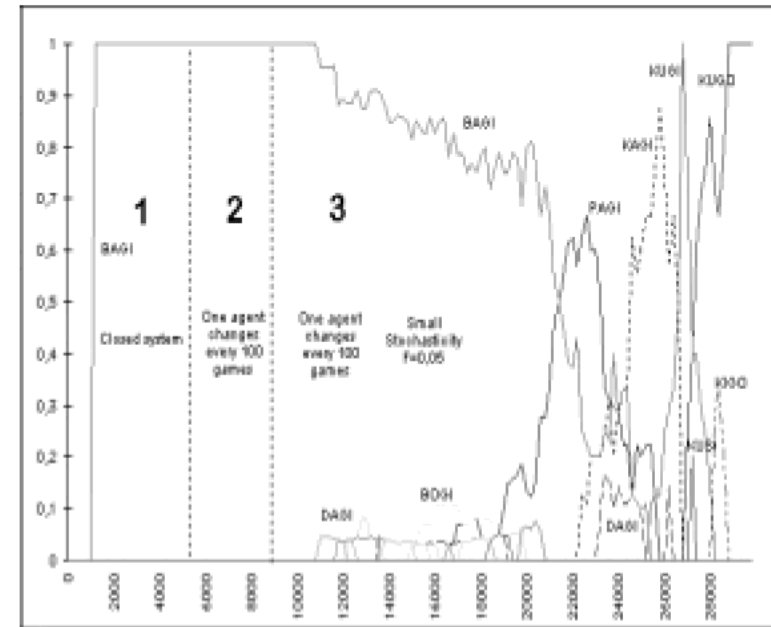


Reaching and maintaining coherence



Reaching and maintaining coherence

- Aufgrund von „Übertragungsfehlern“ verschwinden alte Wortformen und entstehen neue
- „*Evolution due to performance deviation*“



Faktoren der Language Evolution

- Der physiologische und neuronale Apparatus für die Sprache
- die Umwelt (Umgebung)
- Minimalvoraussetzungen für eine „erfolgreiche“ Kommunikation
 - Robust gegen Fehler
 - Lernfähigkeit
 - ausreichende Intelligenz um mit komplexen Bedeutungen umgehen zu können

"So we arrive at a whole new research program for linguistics, one that tries to explain the characteristics and universal tendencies in language rather than merely describe them and then declare them to be innate."

Luc Steel