

Zusammenfassung zur Vorlesung Basismodul Computerlinguistik

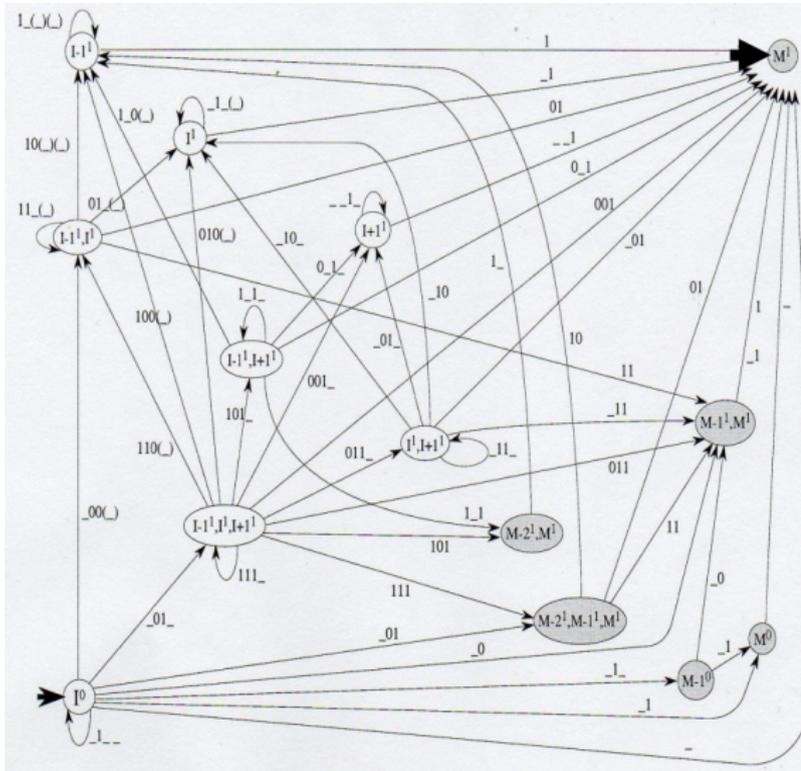
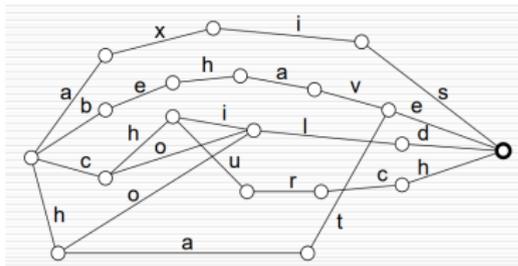
Universelle Levenshtein-Automaten & Themenschwerpunkte in der Klausur

20.01.2022

- 1 Universelle Levenshtein-Automaten: Eine Tafelübung
- 2 Themenschwerpunkte der Klausur

Universelle Levenshtein-Automaten: Eine Tafelübung

Aufgabe: Gegeben seien ein Lexikonautomat \mathcal{L} und Patternwort *chold*. Verwende den Universellen Levenshtein-Automaten mit der Schranke 1 $A^{\vee}(1)$ und finde alle Wörter im \mathcal{L} , deren Levenshtein-Abstand mit dem Patternwort kleiner gleich 1 ist.



Basiswissen

- Grundbegriffe: Sprache, Endlicher Automater
- Determinisierung eines NDEA (Potenzmengenkonstruktion)
- DEA-Minimierung
- Nerode-Konstruktion

Lexikonautomat

- **Daciuk-Algorithmus:** Minimierung eines Lexikonautomaten
- Perfektes Hashing
- Tarjan-Tabellen
- Nachträgliches Hinzufügen / Löschen von Wörtern

Universelle Levenshtein-Automaten

- Generelles Verständnis
- Nachfolgezustände berechnen
- Für Schranke k aus gegebenem Wort Bitvektoren berechnen

Textkorrektur

- Verständnis

- Wenn nicht explizit gefordert: partiellen Automaten entwerfen
- Können auch Textaufgaben drankommen
- Lösung lieber doppelt checken, wenn alles fertig ist