

2. Aufgabenblatt vom Donnerstag, den 31. Oktober 2002 zur Vorlesung

Algorithmen und Datenstrukturen (für Bioinformatik) (Frank Hoffmann)

Abgabe am Donnerstag, den 07. November 2002 bis 16⁰⁰

1. **Korrektheit von Aho–Corasick** (5 Punkte) Im Gusfield–Buch ist der Korrektheitsbeweis für den Knuth–Morris–Pratt Matching–Algorithmus dargestellt. Verallgemeinern Sie diesen Beweis, um die Korrektheit des Aho–Corasick–Algorithmus nachzuweisen.
2. **Vorbereitung der Implementierung des Aho–Corasick–Algorithmus** (8 Punkte) Ziel dieses mittelfristigen Projektes wird es sein, den in der Vorlesung angegebenen Aho–Corasick–Algorithmus zu implementieren und zu testen.
Sie sollen sich zunächst über die Gesamtstruktur Ihres Programmes, zu definierende Klassen und Methoden, Kapselungen etc. klar werden und dies gut strukturiert aufschreiben und kommentieren. Die Mustermenge wird von einer Textdatei (pro Zeile ein Muster) eingelesen.
3. **Endliche Automaten und reguläre Sprachen** (6 Punkte) Geben Sie dfa's (endliche deterministische Automaten) durch ihre Zustandsdiagramme an, die die folgenden formalen Sprachen L über dem Alphabet $\Sigma = \{0, 1\}$ entscheiden. Geben Sie ebenso reguläre Ausdrücke an, die diese Sprachen beschreiben.

- $L = \{w \in \Sigma^* \mid w \neq 11, w \neq 111\}$
- $L = \{w \in \Sigma^* \mid \text{jede ungerade Stelle ist } 0\}$
- $L = \{w \in \Sigma^* \mid w \text{ ist durch } 8 \text{ teilbare Zahl in Binärdarstellung}\}$