

---

## Aufgabenblatt 7

### Suffixbäume & Longest Common Substring

---

Abgabe: **14.12.2004** bis **17 Uhr** über Goya

Maximal: 20 Punkte

Namen nicht vergessen!

---

1. Konstruiert den Suffixbaum zum String 'ottomotor\$' nach Ukkonen. Stellt die Schritte 4-8 (Anhängen von 'omoto') und den fertigen Baum graphisch dar und benennt die jeweils angewendete Extensionsregel! **4 P**
2. Beschreibt und begründet kurz die hauptsächlichen Probleme bei der Implementierung von Suffixbäumen. **4 P**
3. Um eine Konsensussequenz in unterschiedlichen DNA-Abschnitten zu finden, kann man nach den längsten gemeinsamen Substrings suchen. Aus diesen Konsensussequenzen kann man schließen, welche Abschnitte in verwandten Genen über die Zeit und in verschiedenen Organismen erhalten geblieben sind.  
Skizziert den Ablauf, um in den Worten 'ottomotor' und 'protomotif' den längsten gemeinsamen Substrings (*longest common substring*) zu suchen. **6 P**
4. Beschreibt kurz, wie sich der Longest-Common-Substring-Algorithmus auf  $n$  Strings erweitern lässt! **3 P**
5. Es besteht ein Zusammenhang zwischen dem Suffixbaum zu einem String und dem Baum zu seiner rückwärts geschriebenen Fassung. Beschreibt ihn!  
*Hinweis:* Suffixlinks helfen hier weiter. **3 P**