

Übung zu Algorithmen auf Sequenzen

Blatt 1

Ausgabe: 11.10.2018 Besprechung: 18.10.2018

Aufgabe 1.1

Wie viele Teilstrings und wie viele Teilsequenzen hat eine Sequenz der Länge n stets? Wie viele *unterschiedliche* Teilstrings und Teilsequenzen eine Sequenz der Länge n hat, hängt allerdings von der Sequenz selbst ab. Geben Sie hierfür untere und obere Schranken an, sowie Extrembeispiele.

Aufgabe 1.2

Es gibt $2^4 = 16$ binäre Strings der Länge 4. Geben Sie einen möglichst kurzen String an, in dem alle 16 verschiedenen Strings als Substrings vorkommen. Wie ist die Minimallänge eines solchen Strings? Formulieren Sie eine Hypothese über die Minimallänge eines Strings x , der alle Strings Länge q eines σ -elementigen Alphabets enthält.

Aufgabe 1.3

1. Wie viele verschiedene DNA-Sequenzen der Länge q gibt es? (Das ist einfach.)
2. Zeigen Sie, dass auf DNA-Sequenzen der Länge q durch \sim eine Äquivalenzrelation definiert wird (Symmetrie, Reflexivität, Transitivität): $s \sim t$ genau dann wenn $s = t$ oder $s = \text{revcomp}(t)$, wobei revcomp die Funktion ist, die einer DNA-Sequenz ihr reverses Komplement zuordnet.
3. Wie ist Frage 1. zu beantworten unter Berücksichtigung von \sim ? (Achtung: Die Antwort ist komplexer als man denkt. Am besten schreibt man für kleine $q = 1, 2, 3, 4$ alle DNA-Sequenzen zusammen mit ihren reversen Komplementen einmal auf.)

Aufgabe 1.4

1. Berechnen Sie explizit $207 // 16$ und $207 \% 16$ im Binärsystem mit Bitoperationen.
2. Sei x eine positive ganze Zahl (32-bit int). Schreiben Sie eine möglichst einfache Funktion, die testet, ob x eine Zweierpotenz ist.
3. Was tut der Ausdruck $x \& (-x)$? Wofür ist das nützlich?