

Syllabus *Algorithmen auf Sequenzen* (WS 2017/18)

- 17.10. *Administrativa*; Inhaltsübersicht; *Umfrage*;
Maximal Scoring Subsequence Problem (MSS), Biologische Motivation für MSS Problem, Lösungsmethoden mit Analyse: Naive, Dynamische Programmierung, Divide-and-Conquer, Clever-Algorithmus
- 19.10. Clever-Algorithmus (cont.), Laufzeiten; Erweiterung auf All Maximal Scoring Subsequences (AMSS) Problem, Allgemeine Problembeschreibung: prozedural und strukturell, Elementare Eigenschaften der strukturellen Definition
- 24.10. Elementare Eigenschaften der strukturellen Definition, Äquivalenz-Beweis der prozeduralen und strukturellen Definition; Angabe eines linearen Algorithmus
- 25.10. Angabe eines linearen Algorithmus (cont.) und Korrektheitsbeweis, Laufzeitanalyse AMSS; Überblick über verwandte Problemstellungen;
- 26.10. Überblick über verwandte Problemstellungen (cont.); Σ^+ -Bäume, Suffix-Tries und Suffix-Trees; Repräsentationen von Suffix-Bäumen
- 31.10. *Reformationstag*
- 02.11. Speicherplatzsparende Feld-Darstellung; Write-Only-Top-Down-Algorithmus (WOTD), Beispiel für den WOTD-Algorithmus, Laufzeit-Analyse des WOTD-Algorithmus; Suffix-Links, verschachtelte Suffixe und rechtsverzweigende Teilwörter, Idee und elementare Beziehungen für Ukkonens Algorithmus
- 07.11. Aktive Suffixe, elementare Beziehungen für Ukkonens Algorithmus, abstrakter Version Ukkonens Algorithmus, Lokationen und kanonische Lokationen, Beispiel
- 09.11. Ukkonens Algorithmus, Laufzeitanalyse von Ukkonens Algorithmus; Definition von exakten Repeats, Charakterisierung und Erkennung exakter Repeats
- 14.11. Charakterisierung und Erkennung exakter Repeat-Paare, Linksdiversität, Charakterisierung maximaler Repeats, Erkennung maximaler Repeats
- 16.11. Erkennung revers-komplementärer Repeats; Tandem Repeats, Charakterisierung verzweigender Tandem Repeats, Algorithmus zur Erkennung aller rechtsverzweigenden Tandem-Repeats; Beispiel
- 21.11. *am 25.10. vorgeholt*
- 23.11. Laufzeitanalyse für Algorithmus von Stoye und Gusfield; Speedup für den Algorithmus von Stoye und Gusfield für Tandem-Repeats; Alternative D&C-Methode, Laufzeitanalyse, longest common extensions, lowest common ancestors; Beschreibung Conquer-Step von Main-Lorentz und Laufzeitanalyse
- 28.11. k -mismatch Tandem-Repeats, Algorithmus von Landau-Schmidt, Beschreibung Conquer-Step und Laufzeitanalyse; Vokabular und Überdeckungen von Zeichenreihen, Beispiel Vokabular und Überdeckungen von Zeichenreihen;
- 30.11. Vokabular im Suffix-Baum; Skizze des Algorithmus von Gusfield und Stoye; Lempel-Ziv-Zerlegung, Charakterisierung von Tandem-Repeats, Bestimmung einer linkesten Überdeckung in Linearzeit
- 05.12. Bestimmung einer linkesten Überdeckung in Linearzeit (cont.); Partition der linkesten Zerlegung gemäß der Anfangspositionen, sortiert nach Längen der Tandem-Repeats; Teilweise Markierung des Vokabulars im Suffix-Baum
- 07.12. Suffix-Link-Walks, Vervollständigung der Markierung; Vervollständigung der Markierung, Korrektheitsbeweis und Laufzeitanalyse des Algorithmus von Gusfield und Stoye; Definition LCA und RMQ
- 12.12. Algorithmus zur Lösung des LCA-Problem (nach Bender und Farach-Colton); Euler-Touren, Reduktion auf Range-Minimum-Queries; einfaches DP-Preprocessing für die Lösung, Intelligente DP-Preprocessings für die Lösung, Inblock-Anfragen nach Alstrup et al.;

- 14.12. Inblock-Anfragen nach Alstrup et al. (cont.); Ausblick succincte RMQ-Datenstrukturen; Suffix-Arrays, Definition, Konstruktion aus Suffix-Bäumen, Idee des Algorithmus von Manber-Myers;
- 19.12. Algorithmus von Manber-Myers, Algorithmus von Kärkkäinen-Sanders für Suffix-Arrays: Divide-Schritt, Rekursion Conquer-Step,
- 21.12. Kärkkäinen-Sanders: Conquer-Step (cont.), Beispiel, Laufzeitanalyse des Algorithmus von Kärkkäinen-Sanders; Einfache und verbesserte binäre Suche in Suffix-Arrays, longest common prefixes, LCP-Tabelle
- 26.12. *Weihnachten*
- 28.12. *Weihnachtspause*
- 02.01. *Neujahrspause*
- 04.01. *Neujahrspause*
- 09.01. Verbesserte binäre Suche in Suffix-Arrays; longest common prefixes, LCP-Tabelle Linearzeit-Algorithmus zur Bestimmung der LCP-Tabelle; LCP-Intervalle und Kind-Intervalle, konzeptueller LCP-Intervall-Baum
- 11.01. Ermittlung von Kind-Intervallen von LCP-Intervallen mittels ℓ -Indizes, Optimales Suchen in Suffix-Arrays, Simulation von Suffix-Baum Algorithmen auf Extended Suffix-Arrays, Simulation von Parent-Links
- 16.01. Simulation von Suffix-Links und LCA-Anfragen auf Extended-Suffix-Arrays, Speicherplatzbedarf von Extended Suffix Arrays; Burrows-Wheeler-Transformation
- 18.01. Burrows-Wheeler-Transformation und LF-Funktion, Inverse der Burrows-Wheeler-Transformation, Inverse mittels LF, Berechnung der LF-Funktion, FM-Index, Rückwärtssuche
- 23.01. Rückwärtssuche (cont.), Beispiel zur Rückwärtssuche, Rank-Select-Datenstruktur, effiziente Implementierung, Wavelet-Trees, Speicherplatzverbrauch beim FM-Index
- 25.01. Genome Rearrangements, biologischer Hintergrund und Modellierung, Rechnen mit Permutationen, Komplexität von Min-SBR, Breakpoints und triviale untere Schranke für Reversal-Distanz, Strips und Breakpoint-minimierende Reversals, 4-Approximationsalgorithmus
- 30.01. 2-Approximationsalgorithmus, Korrektheitsbeweis (Skizze), Approximationsgüte und Laufzeit (Skizze) Kommentare zur 2-Approximation; (Erweiterte) Breakpoint-Graphen und Zyklenzerlegung Änderung der Zyklenzahl einer maximalen alternierenden Zyklenzerlegung bei Anwendung einer Reversion, eine bessere untere Schranke für die Reversal-Distanz
- 02.02. Eine bessere untere Schranke für die Reversal-Distanz (cont.); Orientierte Permutationen und zugehörige unorientierte Permutationen, Reality-Desire-Diagramm, orientierte und unorientierte Kreise im RDD, Orientierte Komponenten
- 06.02. Kurzzusammenfassung: Orientierte Komponenten sind Mengen von RDD-Kreisen, Hurdles, Super-Hurdles und Fortress, Untere Schranke der orientierten Reversal-Distanz, Beziehung Reality-Desire-Diagramm zu Overlap-Graph, Score eines Reversals, Auswahl orientierter Reversals mit maximalem Score, Eliminierung von Hurdles: Hurdle-Merging und -Cutting; Beziehung Reality-Desire-Diagramm zu Overlap-Graph, Score eines Reversals, Wirkung eines orientierten Reversals auf den Overlap-Graphen, sichere Reversals, orientierte Reversals mit maximalem Score sind sicher, Eliminierung von Hurdles: Hurdle-Merging und -Cutting; Ausblick und Übersicht über verwandte Probleme
- 08.02. Fragestunde
- 15.02. *Klausur*