

Syllabus Algorithmen auf Sequenzen (WS 2020/21)

- KW 45:** *Administrativa*; Inhaltsübersicht; Maximal Scoring Subsequence Problem (MSS), Biologische Motivation für MSS Problem, Lösungsmethoden mit Analyse: Naive, Dynamische Programmierung, Divide-and-Conquer, Clever-Algorithmus, Laufzeiten; Erweiterung auf All Maximal Scoring Subsequences (AMSS) Problem, Allg. Problembeschreibung: prozedural und strukturell, Elementare Eigenschaften der strukturellen Definition
- KW 46:** Elementare Eigenschaften der strukturellen Definition, Äquivalenz-Beweis der prozeduralen und strukturellen Definition Angabe eines linearen Algorithmus und Korrektheitsbeweis, Laufzeitanalyse AMSS, Überblick über verwandte Problemstellungen;
- KW 47:** Σ^+ -Bäume, Suffix-Tries und Suffix-Trees; Repräsentationen von Suffix-Bäumen 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, Aktive Suffixe, elementare Beziehungen für Ukkonens Algorithmus
- KW 48:** Abstrakte Version von Ukkonens Algorithmus, Lokationen und Beispiel; Ukkonens Algorithmus, Laufzeitanalyse von Ukkonens Algorithmus
- KW 49:** Definition von exakten Repeats, Charakterisierung und Erkennung exakter Repeats, Linksdiversität, Charakterisierung maximaler Repeats, Erkennung maximaler Repeats; Erkennung revers-komplementärer Repeats Tandem Repeats, Charakterisierung verzweigender Tandem Repeats, Algorithmus zur Erkennung aller rechtsverzweigenden Tandem-Repeats
- KW 50:** Laufzeitanalyse für Algorithmus von Stoye und Gusfield; Beispiel; 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; k -mismatch Tandem-Repeats, Algorithmus von Landau-Schmidt, Beschreibung Conquer-Step und Laufzeitanalyse;
- KW 51:** Vokabular und Überdeckungen von Zeichenreihen, Beispiel; Vokabular im Suffix-Baum; Skizze des Algorithmus von Gusfield und Stoye; Lempel-Ziv-Zerlegung, Charakterisierung von Tandem-Repeats, Bestimmung einer linkensten Überdeckung in Linearzeit, Partition der linkensten Zerlegung gemäß der Anfangspositionen, sortiert nach Längen der Tandem-Repeats; Teilweise Markierung des Vokabulars im Suffix-Baum, Suffix-Link-Walks, Vervollständigung der Markierung (+); Vervollständigung der Markierung, Korrektheitsbeweis und Laufzeitanalyse des Algorithmus von Gusfield und Stoye (+);
- KW 52:** Definition LCA und RMQ, Algorithmus zur Lösung des LCA-Problem (nach Bender und Farach-Colton), Euler-Konturen, Reduktion auf Range-Minimum-Queries; einfaches DP-Preprocessing für RMQ und Verbesserung; Intelligentes Preprocessing für RMQ, Inblock-Anfragen nach Alstrup et al.; Ausblick succincte RMQ-Datenstrukturen;

24.12. *Weihnachten*

29.12. *Weihnachtspause*

31.12. *Silvester*

05.01. *Neujahrspause*

- KW 01:** Suffix-Arrays, Definition, Konstruktion aus Suffix-Bäumen, Algorithmus von Manber-Myers mit Beispiel und einigen Implementierungsdetails
- KW 02:** Algorithmus von Kärkkäinen-Sanders für Suffix-Arrays: Divide-Schritt, Rekursion, Conquer-Step, Beispiel, Laufzeitanalyse des Algorithmus von Kärkkäinen-Sanders; Einfache und verbesserte binäre Suche in Suffix-Arrays, longest common prefixes, LCP-Tabelle, Verbesserte binäre Suche in Suffix-Arrays; Linearzeit-Algorithmus zur Bestimmung der LCP-Tabelle
- KW 03:** LCP-Intervalle und Kind-Intervalle, konzeptueller LCP-Intervall-Baum, 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, Simulation von Suffix-Links und LCA-Anfragen auf Extended-Suffix-Arrays, Speicherplatzbedarf von Extended Suffix Arrays;
- KW 04:** Burrows-Wheeler-Transformation, LF-Funktion, Inverse der Burrows-Wheeler-Transformation, Inverse mittels LF, Berechnung der LF-Funktion, FM-Index, Rückwärtssuche, Beispiel zur Rückwärtssuche, Rank-Select-Datenstruktur, effiziente Implementierung, Wavelet-Trees, Speicherplatzverbrauch beim FM-Index;
- KW 05:** Genome Rearrangements, biologischer Hintergrund und Modellierung, Rechnen mit Permutationen, Min-SBR, Komplexität von Min-SBR, Breakpoints und triviale untere Schranke für Reversal-Distanz, Strips und Breakpoint-minimierende Reversals, 4-Approximationsalgorithmus, 2-Approximationsalgorithmus, Korrektheitsbeweis (Skizze), Approximationsgüte und Laufzeit (Skizze) Beispiel, 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;
- KW 06:** Orientierte Permutationen und zugehörige unorientierte Permutationen, orientierte Breakpoints und Adjazenzen, Reality-Desire-Diagramm, orientierte Anzahl von Kreisen im RDD, orientierte und unorientierte Kreise im RDD, Untere Schranke der orientierten Reversal-Distanz, Overlap-Graph, Orientierte Komponenten, Orientierte Komponenten sind Mengen von RDD-Kreisen; Hurdles, Super-Hurdles und Fortress, Bessere untere Schranke der orientierten Reversal-Distanz, Beziehung Reality-Desire-Diagramm zu Overlap-Graph, Score eines Reversals, Auswahl orientierter Reversals mit maximalem Score, orientierte Desire-Edges im Overlap-Graphen, Wirkung einer Reversion auf den Overlap-Graphen; Eliminierung von Hurdles: Hurdle-Merging und -Cutting; Beziehung Reality-Desire-Diagramm zu Overlap-Graph, Score eines Reversals aus dem Overlap-Graphen, 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

16.02. Fragestunde

25.02. Klausur

Legende: (+) Teile nicht vollständig, in der Regel ohne Beweise.