

Algorithm Engineering-Projekt

String Alignment mit Neeldeman-Wunsch

Tom Wegener, 18INM/TZ

2019-02-10

Inhaltsverzeichnis

1	Einleitung	1
1.1	generelle Aufgabenstellung	1
1.2	spezielle Problembeschreibung	1
1.3	Charakteristika der Eingabedaten	1
1.4	Messumgebung	1
2	Algorithmen und Optimierung	1
2.1	Needleman-Wunsch	1
2.2	Paralleler Needleman-Wunsch	1
3	Laufzeitmessungen	1
4	Ausblick	1

1 Einleitung

1.1 generelle Aufgabenstellung

Am Anfang des Semesters soll sich für ein Projekt entschieden werden, dieses dann in vier Schritten über den Zeitraum des Semesters ausprogrammiert werden. Zu diesen vier Schritten gehört jeweils eine Abgabe.

Zuerst sollte ein Parser entwickelt werden, der die ausgewählten Daten ausliest und abspeichert. Anschließend eine erste Version des Algorithmus ausprogrammiert werden und erste Laufzeiten gemessen werden, die anschließend auch graphisch dargestellt werden. Aus den Laufzeiten sollten dann als dritte Aufgabe Optimierungsmöglichkeiten aufzeigen. Außerdem wurde die Verwendung eines Profilers empfohlen. In der vierten Aufgabe werden dann zwei Algorithmen auf unterschiedliche Art und Weise verglichen.

1.2 spezielle Problembeschreibung

In dieser Arbeit wird das Problem String-Alignment behandelt. Die Daten werden aus der Datenbank des "National Institute of biotechnical Information"(ncbi) in Form von Fasta-Dateien entnommen und durch einen Parser eingelesen.

Anschließend wird ein Wert ausgegeben, der der Ähnlichkeit von den zwei Strings entspricht. Dafür gibt es verschiedene Algorithmen.

Das Projekt wurde zuerst teilweise in Python umgesetzt und anschließend in Go (bzw golang) übersetzt und fertig gestellt, um die Laufzeiten niedrig zu halten.

1.3 Charakteristika der Eingabedaten

Die

1.4 Messumgebung

2 Algorithmen und Optimierung

2.1 Needleman-Wunsch

2.2 Paralleler Needleman-Wunsch

3 Laufzeitmessungen

4 Ausblick

Das Projekt verlief im Blick auf den Fortschritt relativ konstant, es konnte jedoch gegen Ende der Projektzeitraums ein deutlicher Anstieg der Produktivität und Anstrengung festgestellt werden. Das hat vermutlich damit zu tun, dass wir selbst etwas mehr den Druck, dass das Projekt bald fertig sein muss, gespürt haben, aber auch damit, dass Bachelor- sowie Master-Studenten die Arbeitspaket-Formulierung deutlich verbessert haben und eher viele, aber kleinere Pakete erstellt worden sind, wodurch eine höhere Flexibilität und Produktivität erreicht werden konnte.

Außerdem erschweren Prüfungsphasen auch eine konstante Arbeit am Projekt, da sich in diesen die Aufmerksamkeit auf die Prüfungen legt und somit die Arbeit am Projekt deutlich zurückgeht.

Das Projekt hat mir sehr gut gefallen, da wir keinen bereits vorhandenen Code refactorn mussten oder uns in solchen einlesen mussten, sondern frisch starten konnten. Auch das Thema war interessant und hat uns dadurch, dass es uns eine Auseinandersetzung mit neuen Gegebenheiten abverlangt hat, auch ein Erlernen neuer sowie ein Verbessern und Anwenden bereits vorhandener Fähigkeiten ermöglicht.

Der Wechsel des Projektes und Kundens war eine Hürde, die wir erfolgreich gemeistert haben und mit der ich auch sehr zufrieden bin, da unser Produkt schon viel Konkurrenz hatte, bevor wir angefangen haben und ich keinen Unique Selling Point ausmachen konnte, der uns von der Konkurrenz abgegrenzt hätte. Auch der Anspruch des ersten Kundens, den Quellcode des Projektes nicht zu veröffentlichen, wäre meiner Meinung nach für ein Projekt von Studierenden nicht angemessen gewesen.