

Aufgabe 4.1: MoSS (T)_Artischocken

In Abbildung 1 sehen Sie eine Darstellung des MoSS Algorithmus zur Vermeidung von Substruktur-Duplikaten.

1. Erweitern Sie das Beispiel aus Abbildung 1, an dem mit einem Stern (*) markierten Ast, um 2 weitere Stufen.
2. Markieren Sie die Muster, die aufgrund einer Regelverletzung nicht in die Lösungsmenge aufgenommen werden und geben Sie die dafür zuständige Regel an.

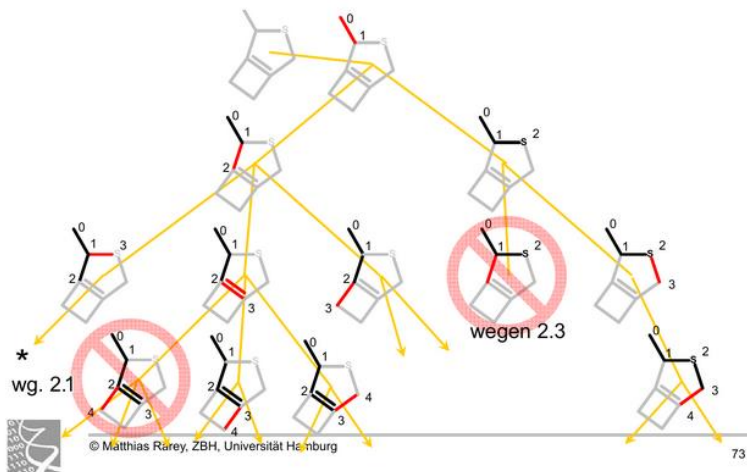


Abbildung 1: Frequent Pattern Mining

1.)

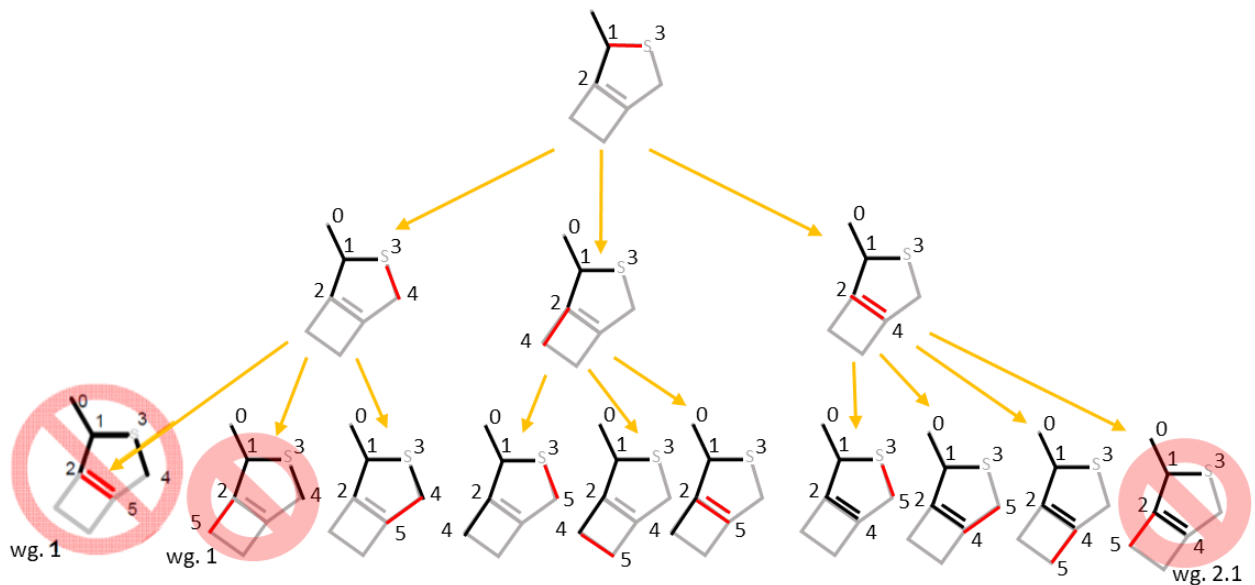


Abbildung 2: Erweiterung des Beispiels aus der Aufgabenstellung um zwei weitere Stufen; rot durchgestrichen markiert: regelunzulässige Erweiterungen nach Frequent Pattern Mining zur Vermeidung von Substruktur-Duplikaten

Algorithmen für Frequent Pattern Mining

- Vermeidung von Substruktur-Duplikaten:

Atome erhalten eine aufsteigende Nummer nach Einfügereihenfolge (Atom-Ordnungsnummer).

1. Erweitere stets nur an Atomen, deren Ordnungsnummer \geq der Ordnungsnummer des Atoms ist, an dem zuletzt hinzugefügt wurde.
2. Definiere eine Ordnung \angle auf ausgehende Bindungen an einem Atom, erweitere nur nach aufsteigender \angle -Ordnung:
 1. Einfach- \angle Aromatische \angle Doppel- \angle Dreifachbindung
 2. Ringschlussbindungen (führen zu Atomen der Substruktur) \angle offene Bindungen (führen zu neuen Atomen)
 3. Ordnungszahl Zielatom
3. Verbiете Ringschlussbindungen von Atomen mit hoher Ordnungsnummer zu Atomen mit kleinerer Ordnungsnummer

Abbildung 3: Ausschnitt aus der Vorlesung CIW, der die Regeln des Frequent Pattern Mining enthält¹

In Abbildung 2 ist die um zwei Stufen am (*) erweiterte Beispiel aus Abbildung 1 zu sehen und in Abbildung 3 sind die Regeln zur Vermeidung von Substruktur-Duplikaten als Ausschnitt aus der Vorlesung zu sehen.

In Abbildung 2 sind genau drei Erweiterungen unzulässig. Zwei davon (links im Bild zu sehen), sind unzulässig, da sie gegen Regel 1 verstoßen. Regel 1 besagt, dass nur an Atomen erweitert werden darf, wenn deren Ordnungsnummer größer/gleich der Ordnungsnummer des Atoms ist, an dem zuletzt eine Erweiterung stattgefunden hat (siehe Abb. 3). Bei den Erweiterungen war die Ordnungsnummer an der zuletzt erweitert wurde 3; die neue Erweiterung sollte jedoch an Atom 2 stattfinden. Da 2 jedoch nicht größer/gleich 3 ist, sind diese Erweiterungen unzulässig.

Die rot durchgestrichene Erweiterung rechts in Abbildung 2 ist unzulässig, da sie gegen Regel 2.1 verstößt. Diese Regel besagt, dass eine Ordnung auf den Bindungen besteht und nur nach aufsteigender Ordnung erweitert werden darf. Da hier einmal eine Bindung mit Doppelbindung an Atom 2 angebaut wurde, darf keine Einfachbindung, die in der Reihenfolge davorliegt, an Atom 2 angebaut werden.

¹ Rarey, Vorlesung CIW 2020, K2_05, Folie 68