

Vergleich mehrerer RNA Sekundärstrukturen durch Abbildungen zwischen geordneten Bäumen

Christian Neuenroth

29.05.2012

In der Bioinformatik ist es von Interesse zu wissen, wie ähnlich sich verschiedene RNA Sekundärstrukturen sind. Bei der Betrachtung von diesen Strukturen fällt auf, dass diese Strukturen der eines Baumes „entsprechen“. In der Präsentation wird gezeigt wie man mit den drei einfachen Operationen, Löschen, Einfügen und Ersetzen, Abbildungen zwischen geordneten Bäumen definiert und man definiert den Abstand zweier Bäume als das Minimum der Kosten von allen Abbildungen zwischen diesen Bäumen. Somit können wir einen Vergleich von mehreren RNA Sekundärstrukturen anstellen.

Ich werde auf einen einfachen Algorithmus zur Abstandsberechnung zwischen zwei Bäumen eingehen der bereits seit den 1980er bekannt ist. Anschließend werde ich eine Erweiterung auf beliebig viele Bäume präsentieren.

Zur Reduzierung der Rechenzeit wird eine Methode angegeben mit der wir den wahren Abstand um maximal Faktor 2 überschätzen. Diese Methode ist auch schon für die Approximation des Abstandes mehrerer Lsitien bekannt.