



# New Assignment Methods for Chemical Similarity Measures

Der Optimal Assignment Kernel (OAK) ist ein erfolgreiches Ähnlichkeitsmaß, das in einem breitem Spektrum an Screening und Machine Learning Anwendungen eingesetzt wird. Das Assignment auf Atomebene wird dabei durch den Munkres-Algorithmus gelöst, welcher die maximale Summe eines bipartiten Graphmatchings berechnet. Dieser Schritt kann zu Assignments führen, die die Topologie und Geometrie der Moleküle nicht korrekt darstellen. Der OAK bestraft diese Abbildungen nicht, da die physikochemischen Deskriptoren der Atome und der Nachbarschaft nahezu identisch sind. Aus chemischer Sicht sind solche Assignments jedoch nicht korrekt, da die topologische Anordnung der funktionellen Gruppen innerhalb eines Moleküls nicht berücksichtigt wird. Diese Anordnungen können aber wichtige Informationen zur Modellierung der Aktivität beitragen.

Damit diese topologischen Defekte nicht auftreten oder stärker bestraft werden, soll im Rahmen dieser Arbeit mindestens eine neue Methode entwickelt werden, die den Munkres-Algorithmus ersetzt. Hierbei sind verschiedene Ansätze, wie rekursives Sub-Graph- oder Sub-Treematching, als auch Alignmentmethoden auf Atom- bzw. struktureller Merkmalsebene denkbar.

Voraussetzungen:  
Programmiererfahrung in Java

## Kontakt

Andreas Jahn  
Sand 1, Raum A316  
Tel. (07071) 29-77175  
andreas.jahn@uni-tuebingen.de

