

# HABILITATION

Fachbereich Mathematik und Informatik

## EINLADUNG

zum Habilitationsvortrag

Im Rahmen seines Habilitationsverfahrens wird

**Herr Dr. Hugues Richard**

am **Mittwoch, d. 03. Juli 2024**

um **16:00 Uhr** im **Seminarraum 005, Takustr. 9, 14195 Berlin**

einen Vortrag über das Thema:

**Differentiable sequence alignment and associated learning strategies**

halten.

Der Vortrag wird ca. 45 Minuten dauern (Zusammenfassung s.u.).

Die Universitätsöffentlichkeit ist dazu herzlich eingeladen.

gez. Prof. Dr. G. Rote

Dekan des FB Mathematik und Informatik

Abstract: Biological sequence alignment is one of the most foundational methods for genome sequence studies. As a result, it is at the core of most bioinformatics analysis: it precludes the construction of an overlap graph for sequence assembly, it is the first step of transcriptomics or ChIP-Seq analysis, and it can detect homologous relationships between proteins. A large corpus of work has been done on efficiently finding the set of optimal or suboptimal alignments given a predefined scoring function. However, despite first efforts aiming at optimizing score and alignment jointly, scoring functions are usually designed independently (e.g. PAM and BLOSUM matrices) and then fixed for the alignment step. Some of the work used the connection between gap-affine alignment scores and Hidden Markov Models (HMM) to provide strategies that estimate the score function together with the alignment. Recently, it was proposed to model the scores as differentiable distributions and integrate those into an end-to-end machine learning framework. In such a setup, the scoring function is optimized together with the alignment for a given task. This talk will present in an historical context how we arrived at this solution of differentiable scoring functions. After summarizing the properties of sequence alignment and highlighting its link to pair HMM and the preliminary work that has been done to estimate score functions from sequence data, I will present the methods that were proposed recently with end-to-end differentiable learning.