

Dissecting the complex structure of the phenome through machine learning

Initiative: Modellierung und Simulation komplexer Systeme (beendet)

Ausschreibung: Fellowships "Computational Sciences"

Bewilligung: 16.03.2010

Laufzeit: 2 Jahre 6 Monate

Understanding phenotypic variation, and in particular identifying the causal genetic or environmental regulators, is a major aim in biological investigations. The goal of this fellowship is to develop machine learning techniques to model the structure of the underlying complex system based on modern, high-dimensional phenotype datasets. First, the temporal structure of phenotypes that are recorded over time is addressed. By statistical modelling the smoothness of time series is exploited for identifying change points. Second, the structure of images, arising when digital pictures are used as phenotypic variables, is considered. Using machine learning techniques interpretable image features that can be used as quantitative traits, are extracted complementing classical measurements. Finally, the network structure of the phenome is addressed. Different phenotype variables influence each other, resulting in a chain of effects that needs to be modelled to reveal the true causal relationships. The developed algorithms are aimed at and will be tested to understand phenotypic variation in *Arabidopsis thaliana* in direct collaboration with plant biologists at the Max Planck Institute for Developmental Biology.

Projektbeteiligte

Oliver Stegle

Max-Planck-Institut für biologische
Kybernetik
Abt. Empirische Inferenz
Tübingen