UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Machine Learning and Statistical Physics of Macromolecular Systems

2526-1-F1703Q025

Aims

Organic and biological macromolecules are paradigmatic examples of complex systems:
On the one hand, their physical modeling is complicated by the presence of interactions at multiple length scales. On the other hand, frustration and both energetic and entropic barriers make the exploration of conformational space, as well as the study of dynamic and equilibrium properties, computationally extremely expensive.

This course aims to guide students from the basic principles to the cutting edge of theoretical study and simulation of complex macromolecular systems. In particular, it will demonstrate how the integration of statistical physics techniques developed over the last 20 years with machine learning algorithms now makes it possible to tackle major computational challenges, offering a new physico-chemical tool for investigating fundamental processes in biophysics, biology, materials science, and pharmaceutical research.

The course is designed to accommodate two different types of students:

1. Students with a theoretical background in physics, chemistry, or materials science who are interested in deepening their understanding of the theoretical and conceptual aspects of statistical physics of complex systems, stochastic dynamics, and the theoretical foundations of machine learning algorithms used in this field.

2. Students with an experimental background (in biophysics, biochemistry, or biotechnology) who are interested in the practical aspects, namely learning how to perform and interpret molecular simulation techniques.

The students will be able to choose between two different exam formats, reflecting these two distinct learning objectives.

Contents

Relaxation times. Multiscale approaches, statistical and stochastic fields, statistical mechanics of rare events and related computational challenges.
Overview of the phenomenology of macromolecules, with a focus on biopolymers.
AI algorithms for the exploration, sampling, and analysis of molecular dynamics.
Algorithms for molecular structure prediction.

Detailed program

Chapter 1: Biomolecules as complex open systems

The theoretical foundation of physics-based simulation of biomolecules
The concept of multi-scale description of soft and biological matter

Chapter 2: Stochastic dynamics

Dynamics in open macromolecular systems: characteristic scales and fluctuation-dissipation processes
Microscopic derivation of the generalized Langevin equation. Ohmic and overdamped limit.
Rudiments of Stochastic Calculus and stochastic measure.
Fokker-Plank equation and stochastic path integrals

Chapter 3: Statistical mechanics of thermally activated processes

Reactive events and transition path ensamble
Stochastic descriptors of reactive dynamics: committor function, transition path density and transition path current
Kinetics of Mean-first passage times and transition path time
Markov state models

Chapter 4: Molecular Dynamics (MD)

Ergodic Integrators
MD in the NPT and NVT ensembles.
Practical Simulations using GROMACS: set-up, execution and analysis of MD simulations (6 ore)

Chapter 5: Statistical Computing for molecular simulations

Collective variables and potential of mean-force
Diffusive distance and diffusion maps
Configuration clustering and Markov State Model construction
Intrinsic manifold and intrinsic dimensionality

Chapter 5: Enhanced sampling methods and machine learning for molecular simulations

Open challenges in simulating complex biological matter
Enhanced sampling algorithms
Machine Learning for Molecular Simulations

Prerequisites

Analytical mechanics in Lagrangian and Hamiltonian formulation, statistical mechanics (ergodicity and the theory of microcanonical and canonical ensembles), quantum mechanics, and mathematical methods for physics (Hilbert spaces, distributions, linear operators).

For students with a biophysical or biochemical background:

Basic concepts of statistical mechanics (thermodynamic equilibrium and its properties, Boltzmann distribution). Newton's equation, the laws of thermodynamics and the notion of free-energy, electrostatics, and phenomenological aspects of quantum mechanics (energy quantization, the concept of wavefunction and orbital).

Teaching form

The core topics of the course will be covered through traditional lectures at the blackboard.

Students following the theoretical track will take an oral exam on the theoretical topics and present an AI algorithm of their choice.

Students following the applied/experimental track will take an oral exam on basic concepts of statistical mechanics related to the course and present a simulation project of a macromolecular system.

Textbook and teaching resource

- Typeset notes provided by the lecturer.
- M. Tuckermann: "Statistical mechanics: Theory and Molecular Simulations", Oxford Graduate Texts
- T. Schlick: "Molecular Modeling and Simulation: An Interdisciplinary Guide (Interdisciplinary Applied Mathematics)" Springer.

Semester

Second semester, second year

Assessment method

The grade will be assigned on the basis of a final oral examination, also taking in consideration the student's contribution to the special topic session (or computational project) and related discussion.

Office hours

The lecturer will be available any time, previous arranging the time and date by email.

Sustainable Development Goals

INDUSTRY, INNOVATION AND INFRASTRUCTURE