



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Machine Learning and Statistical Physics of Macromolecular Systems

2526-1-F1703Q025

Obiettivi

Le macromolecole organiche e biologiche sono esempi paradigmatici di sistemi complessi: Da un lato, la loro modellizzazione fisica è complicata dalla presenza di interazione su diverse scale di lunghezza. Dall'altro, frustrazione e barriere sia energetiche che entropiche rendono computazionalmente estremamente costosa l'esplorazione dello spazio conformazionale e lo studio delle proprietà sia dinamiche che di equilibrio di questi sistemi.

Questo corso si propone di condurre lo studente dai principi di base fino alla frontiera dello studio teorico e la simulazione dei sistemi macromolecolari complessi. In particolare, si mostrerà come l'integrazione di tecniche di fisica statistica sviluppate negli ultimi 20 anni con algoritmi di machine learning renda oggi possibile affrontare i principali ostacoli computazionali, offrendo un nuovo strumento di indagine chimico-fisico per lo studio processi fondamentali in biofisica, biologia, scienze dei materiali e ricerca farmacologica.

Il corso è concepito per adattarsi a due tipologie diverse di studenti:

1. studenti provenienti da un percorso teorico in fisica, chimica o scienze dei materiali interessati ad approfondire maggiormente aspetti teorici e concettuali della fisica statistica dei sistemi complessi, della dinamica stocastica e i fondamenti teorici di algoritmi di machine learning che sono utilizzati in questo ambito.
2. studenti provenienti da un percorso sperimentale (in biofisica, biochimica, o biotecnologie) interessati ad approfondire gli aspetti pratici, ovvero imparare ad eseguire ed interpretare di tecniche di simulazione molecolare.

Gli studenti potranno scegliere tra due diverse modalità di esame che riflettono questi due diversi obiettivi formativi.

Contenuti sintetici

Dalla dinamica stocastica alle proprietà di equilibrio di sistemi complessi. Tempi di rilassamento. Approcci multi-scala, campi statistici e stocastici, meccanica statistica degli eventi rari e problemi computazionali ad essa collegati. Cenni di fenomenologia delle macromolecole e in particolare di biopolimeri. Algoritmi AI per l'esplorazione, il campionamento e la analisi di dinamica molecolare. Algoritmi per la predizione di strutture molecolari.

Programma esteso

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 ensemble
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

Prerequisiti

Per studenti di indirizzo teorico:

Meccanica analitica Lagrangiana e Hamiltoniana, meccanica statistica (ergodicità e teoria degli ensemble microcanonici e canonici), meccanica quantistica, metodi matematici per la fisica (spazi di Hilbert, distribuzioni, operatori lineari).

Per studenti di indirizzo biofisico e biochimico:

Concetti rudimentali di meccanica statistica (equilibrio termodinamico e sue proprietà, distribuzione di Boltzmann). Equazione di Newton, leggi della termodinamica, elettrostatica e aspetti fenomenologici di meccanica quantistica (quantizzazione dell'energia, concetto di funzione d'onda e di orbitale)

Modalità didattica

Gli argomenti centrali del corso saranno trattati in lezioni frontali alla lavagna.

Gli studenti interessati al percorso teorico svolgeranno un esame orale sugli argomenti teorici e presenteranno un algoritmo AI a scelta.

Gli studenti interessati al percorso teorico svolgeranno un esame orale su rudimenti di meccanica statistica connessi al corso e porteranno un progetto di simulazione di sistema macromolecolare.

Materiale didattico

- Dispense fornite dal docente.
- M. Tuckermann: "Statistical mechanics: Theory and Molecular Simulations", Oxford Graduate Texts
- T. Schlick: "Molecular Modeling and Simulation: An Interdisciplinary Guide (Interdisciplinary Applied Mathematics)" Springer.

Periodo di erogazione dell'insegnamento

Secondo semestre, secondo anno

Modalità di verifica del profitto e valutazione

Il voto verrà assegnato sulla base di una prova orale finale, tenuto anche conto del contributo dello studente alle sessioni tematiche e relative discussioni.

Orario di ricevimento

In qualsiasi momento, previo accordo con il docente via email

Sustainable Development Goals

IMPRESE, INNOVAZIONE E INFRASTRUTTURE
