Chore: Brute Force Bounties

Elliot Greenlee

April 25, 2018

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Outline

- **Perspectives**
- Motivation
- Requirements
- Domain Test Case
- State of the Art
- Bounty System
- Pathogen Box Workflow
- Future Work

Perspectives

- Introduction to Data Mining (COSC 526)
  - Preprocessing/Prediction focus

- Bioinformatics (COSC 594)
  - Biological usefulness

- Security
  - Integrity

Outline

- Perspectives
- **Motivation**
- Requirements
- Domain Test Case
- State of the Art
- Bounty System
- Pathogen Box Workflow
- Future Work

Motivation

- Data/Computation imbalance
- Brute force simulations and calculations
- Expense barrier
  - Companies
    - Amazon Web Services
    - Dedicated clusters
  - Researchers
    - ?
- Target idle time

Outline

- Perspectives
- Motivation
- **Requirements**
- Domain Test Case
- State of the Art
- Bounty System
- Pathogen Box Workflow
- Future Work

Requirements

- Price
  - Lower for researchers
  - Attract computational power
 - Estimate value

Requirements

- Security
  - Proof of work
  - Result integrity
  - Work sensitivity

Outline

- Perspectives
- Motivation
- Requirements
- **Domain Test Case**
- State of the Art
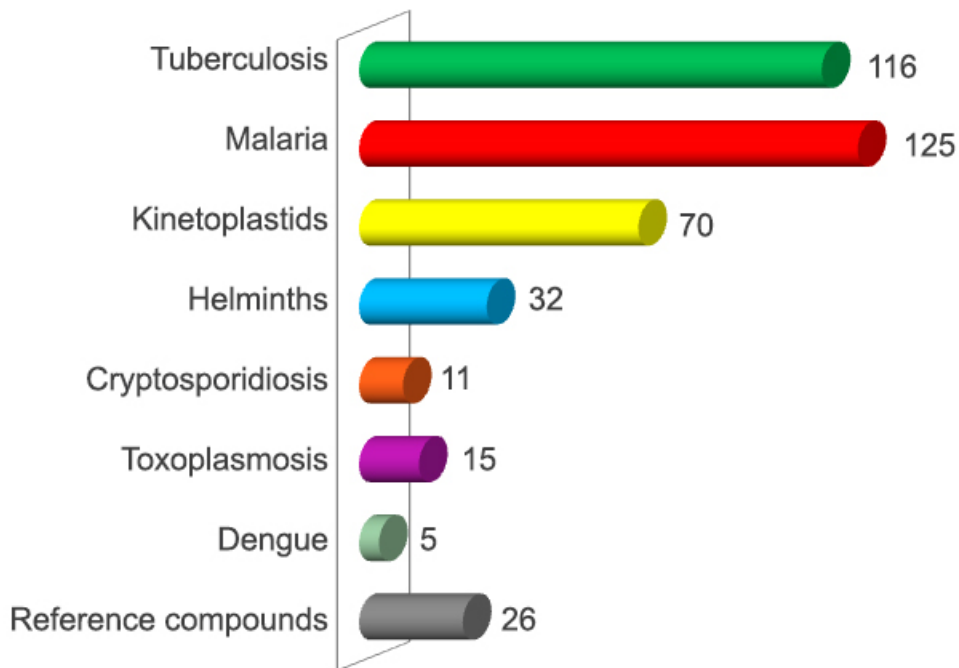- Bounty System
- Pathogen Box Workflow
- Future Work

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Domain Test Case

- Molecular Dynamics
- Atom and molecule interactions
- Important for medicine
- Iterative brute force calculations
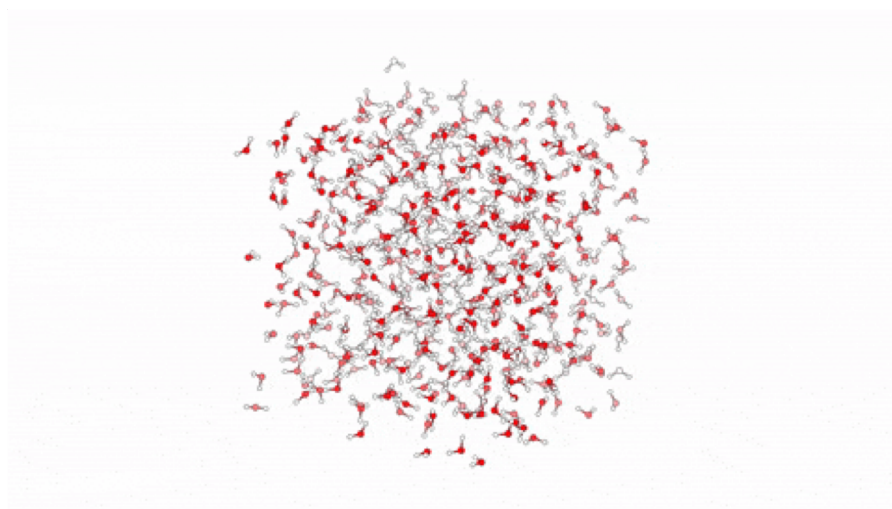
Domain Test Case

- Pathogen Box drugs
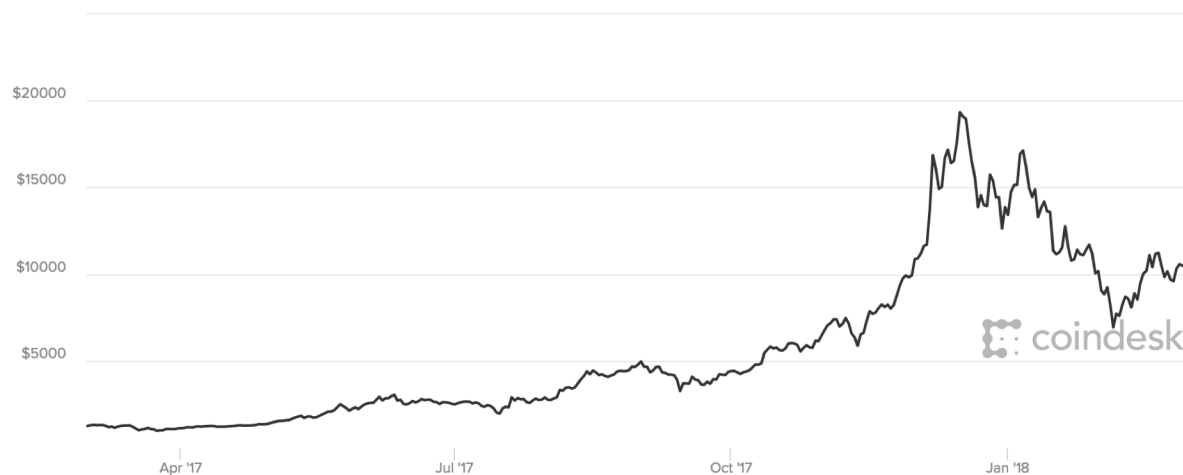- Neglected subtropical diseases

Outline

- Perspectives
- Motivation
- Requirements
- Domain Test Case
- **State of the Art**
- Bounty System
- Pathogen Box Workflow
- Future Work

State of the Art

- Folding@home
- Pande lab at Stanford
- Simulates molecular dynamics
- Distributed to user's idle time
- 18 years, 100 petaflops
- Charity

State of the Art

- CureCoin
- Cryptocurrency value
- Part of the Folding@home ecosystem
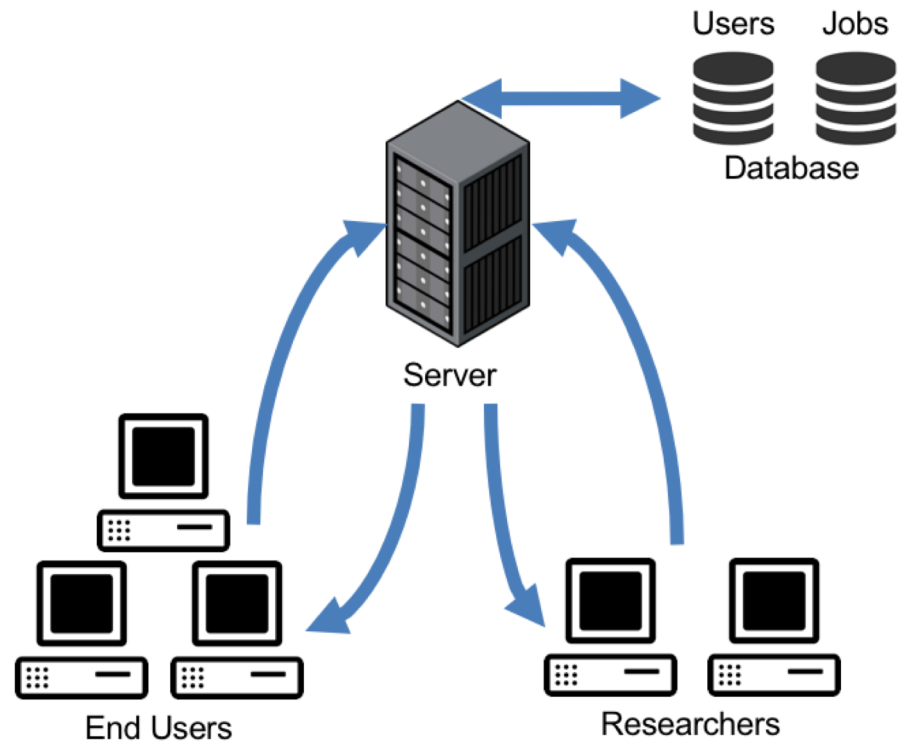- Direct computational power from altcoins to folding

Outline

- Perspectives
- Motivation
- Requirements
- Domain Test Case
- State of the Art
- **Bounty System**
- Pathogen Box Workflow
- Future Work

Bounty System

- Modifications to a traditional job board

- Central authority
- Bounty system

Bounty System

User

| Field	Description
id	To identify the user
money_id	Bank/paypal account
money	Money in the account
completed_jobs	Number of jobs completed by the user
submitted_jobs	Number of jobs submitted by a researcher
computer_rating	User computer performance relative to a standard

Bounty System

Job

Field	Description
id	To identify the job
job_files	Location of the job files
hours	How long in standard units to run the job
money	Total value of completing the job
results_files	Location of the results files
taken	Boolean if the job is currently being run

Bounty System

A new researcher gives $500 to complete 2 lengthy tuberculosis jobs, at 20% tax

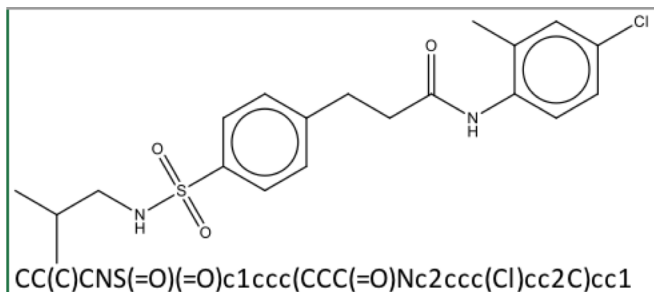| Job ID	Value	Hours
Dengue 1	$10 per hour	10
Dengue 2	$10 per hour	10
Malaria 1	$0 per hour	10
Malaria 2	$0 per hour	10
Malaria 3	$0 per hour	10
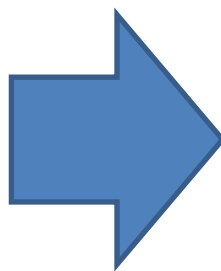
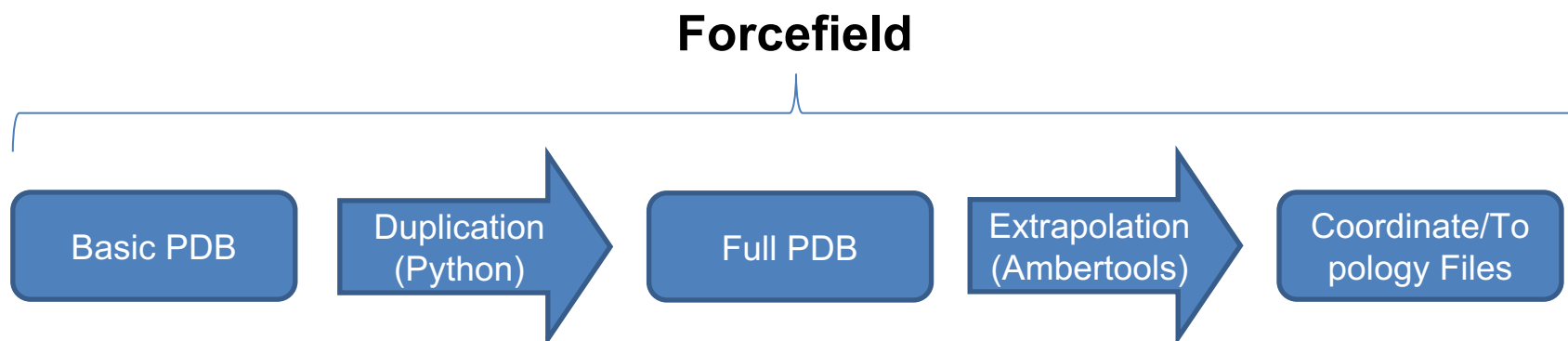| Job ID	Value	Hours
Dengue 1	$12 per hour	10
Dengue 2	$12 per hour	10
Tuberculosis 1	$10 per hour	20
Tuberculosis 2	$10 per hour	20
Malaria 1	$2 per hour	10
Malaria 2	$2 per hour	10
Malaria 3	$2 per hour	10

Outline

- Perspectives
- Motivation
- Requirements
- Domain Test Case
- State of the Art
- Bounty System
- **Pathogen Box Workflow**
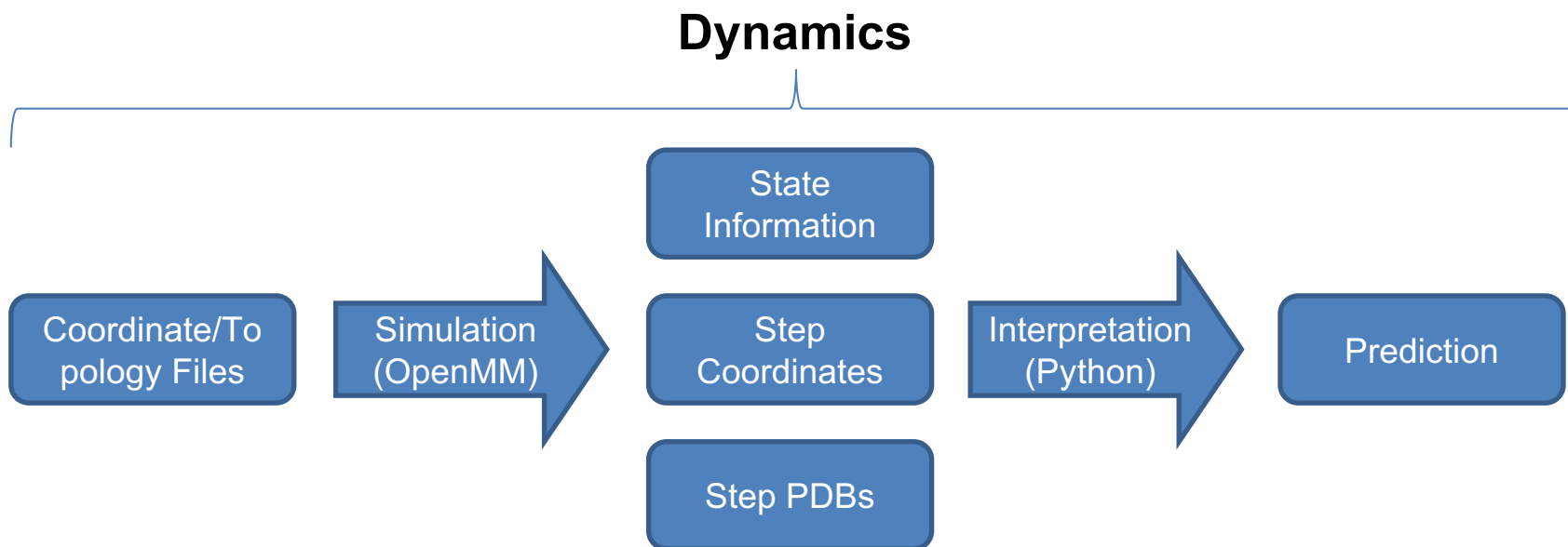- Future Work

Pathogen Box Workflow

Molecule

Pathogen Box Metadata → Extraction (Python) → SMILES → Conversion (Cactus) → Basic PDB

CC(C)CNS(=O)(=O)c1ccc(CCC(=O)Nc2ccc(Cl)cc2C)cc1

Pathogen Box Workflow

Forcefield

Basic PDB → Duplication (Python) → Full PDB → Extrapolation (Ambertools) → Coordinate/Topology Files

Pathogen Box Workflow

Dynamics

Coordinate/Topology Files → Simulation (OpenMM) → State Information / Step Coordinates / Step PDBs → Interpretation (Python) → Prediction
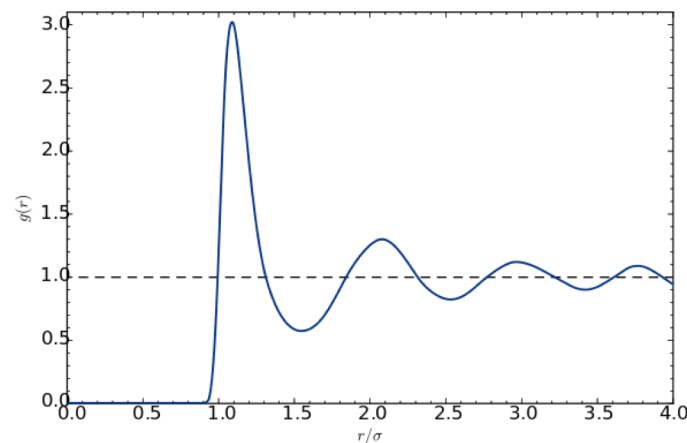
Pathogen Box Workflow

- Radial distribution function
- Particle pair distances
- Bin and normalize to ideal gas
- Compare water against drug molecule
- View over time

Outline

- Perspectives
- Motivation
- Requirements
- Domain Test Case
- State of the Art
- Bounty System
- Pathogen Box Workflow
- **Future Work**

Future Work

- Deliver predictions for drug molecules
- Use more appropriate standards/settings
- Improve the molecule descriptions

- Move prototype to an accessible server
- Improve security of the prototype

- Evaluate new job types

Chore: Brute Force Bounties

Elliot Greenlee

April 25, 2018