The ML Group at CWI +
Monte Carlo Tree Search

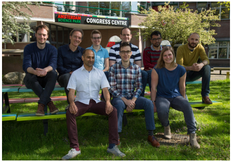
Wouter M. Koolen

CWI
Centrum Wiskunde & Informatica

CWI Scientific Meeting, Friday 16^{th} June, 2017

The Machine Learning Group

The Machine Learning Group



Misspecification
Hypothesis Testing
Statistical Learning Theory

Spiking Neural Networks
Deep Reinforcement Learning

Online Learning and Optimisation
Monte Carlo Tree Search

Overview

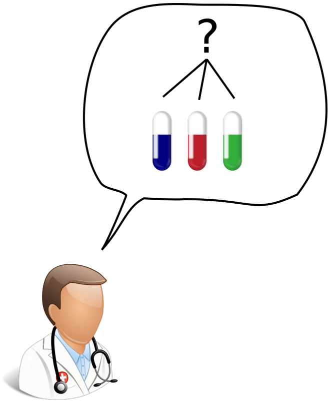
Questions:

- Which data points to collect?
- How many data points do I need?
- How to draw provable conclusions?

Today:

- Best Arm Identification
- Monte Carlo Tree Search

Clinical Trials Example



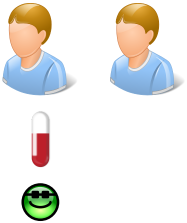
Clinical Trials Example

Clinical Trials Example

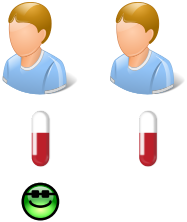
Clinical Trials Example

Clinical Trials Example

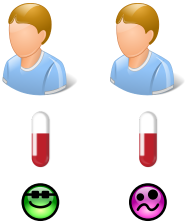
Clinical Trials Example



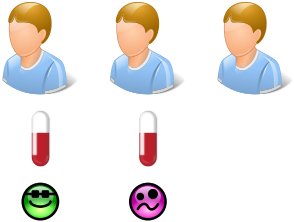
Clinical Trials Example



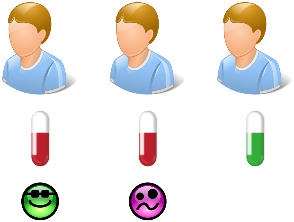
Clinical Trials Example



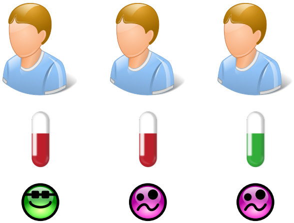
Clinical Trials Example



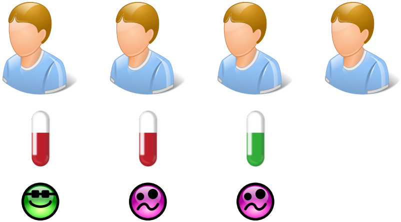
Clinical Trials Example



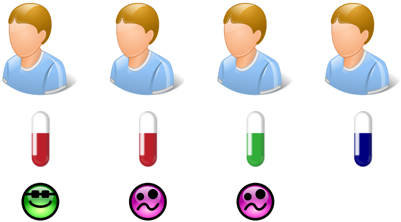
Clinical Trials Example



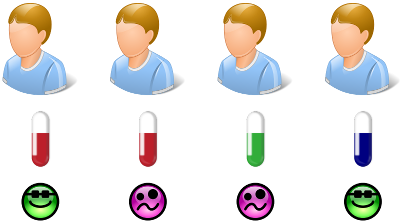
Clinical Trials Example



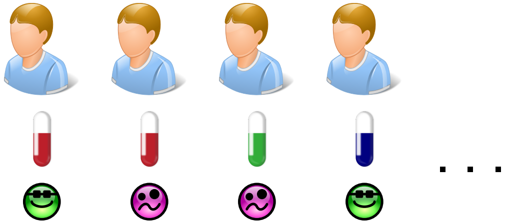
Clinical Trials Example



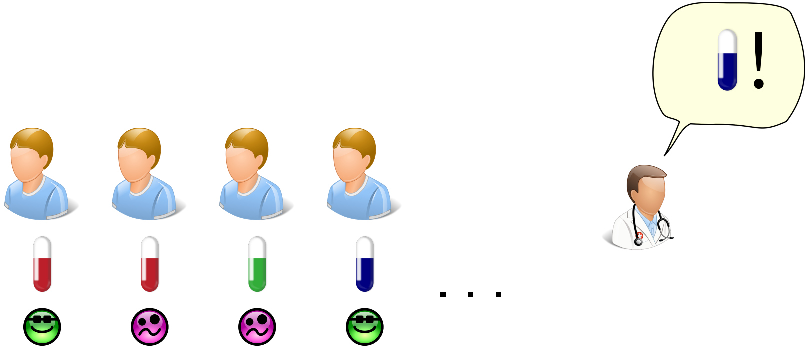
Clinical Trials Example



Clinical Trials Example



Clinical Trials Example



Probably Approximately Correct (PAC) Learning

World: drug success rates $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$

Strategy:

- Adaptive sampling rule I_t
- Stopping rule τ
- Recommendation rule $\hat{\imath}$

Probably Approximately Correct (PAC) Learning

World: drug success rates $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$

Strategy:

- Adaptive sampling rule I_t
- Stopping rule τ
- Recommendation rule $\hat{\imath}$

Definition

A strategy is δ-**PAC** if for all $\boldsymbol{\mu}$

$$\mathbb{P}_{\boldsymbol{\mu}}\left(\hat{\imath} \neq \arg\max_i \mu_i\right) \;\leq\; \delta.$$

Probably Approximately Correct (PAC) Learning

World: drug success rates $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$

Strategy:

- Adaptive sampling rule I_t
- Stopping rule τ
- Recommendation rule $\hat{\imath}$

Definition

A strategy is δ-**PAC** if for all $\boldsymbol{\mu}$

$$\mathbb{P}_{\boldsymbol{\mu}}\left(\hat{\imath} \neq \arg\max_i \mu_i\right) \leq \delta.$$

Want: δ-PAC strategy with low **sample complexity** $\mathbb{E}[\tau]$.

Thompson Sampling

- Assume prior distributions on success rate μ_i of each drug i.
- Each round t
  - Draw a world $\tilde{\boldsymbol{\mu}}_t$ from posteriors
  - Try the best drug for it $I_t = \arg\max_i \tilde{\mu}_{t,i}$
  - Update the posterior

Video.

Information Theoretic Lower Bounds

How many samples are **really necessary** in a world μ?

Information Theoretic Lower Bounds

How many samples are **really necessary** in a world $\boldsymbol{\mu}$?

Change of measure idea

Consider a world $\boldsymbol{\mu}'$ that is close to $\boldsymbol{\mu}$ but has a different best drug

$$\arg\max_i \mu'_i \neq \arg\max_i \mu_i$$

Few observations might as well have come from $\boldsymbol{\mu}'$ ...
... but then your answer is wrong.
Need **many** observations.

Information Theoretic Lower Bounds

How many samples are **really necessary** in a world $\boldsymbol{\mu}$?

Change of measure idea

Consider a world $\boldsymbol{\mu}'$ that is close to $\boldsymbol{\mu}$ but has a different best drug

$$\arg\max_i \mu'_i \neq \arg\max_i \mu_i$$

Few observations might as well have come from $\boldsymbol{\mu}'$ ...
... but then your answer is wrong.
Need **many** observations.

Theorem (KCG'15)

Any δ-PAC algorithm needs

$$\mathbb{E}[\tau] \geq T^*(\boldsymbol{\mu}) \ln \frac{1}{\delta} \quad \textit{where} \quad \frac{1}{T^*(\boldsymbol{\mu})} = \max_{\boldsymbol{w} \in \triangle} \min_{\boldsymbol{\mu}'} \sum_i w_i \, \mathrm{KL}(\mu_i \| \mu'_i)$$

Outlook

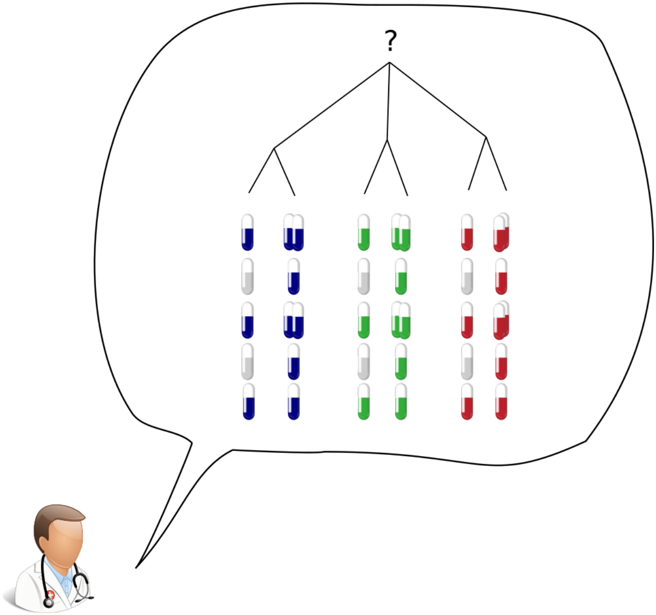
Optimal algorithm now available [GK'16].

- Matching lower bound
- Characterise proportion of draws of each arm

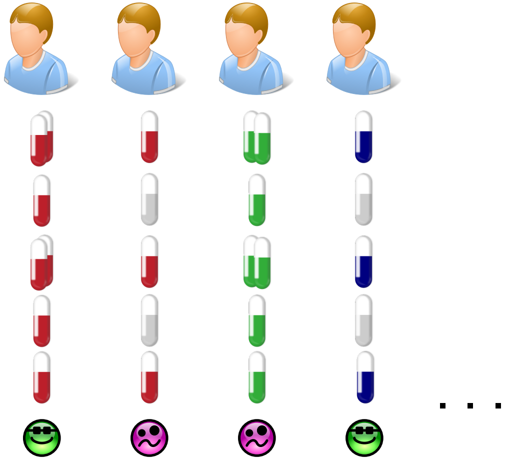
"Top Two" Thompson Sampling gets very close [R'16].

How to answer more challenging questions?

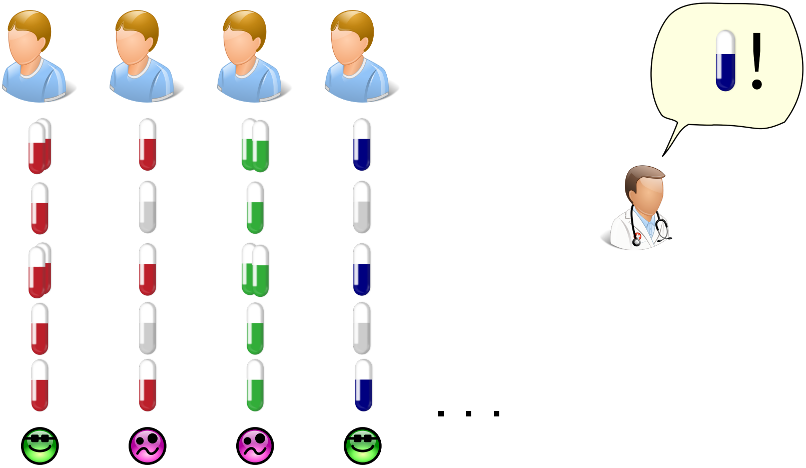
Robust Clinical Trials Example



Robust Clinical Trials Example



Robust Clinical Trials Example



My Project

Goal: develop complete theory of tree search

- Lower bounds
  - Optimal weights are often sparse
  - hints at **pruning**
  - computational challenges
- Well-developed understanding of depth 2 [GKK'15]
- Upgrading efficient algorithms [THT'14, GKK'15, KK'17]

My Project

Goal: develop complete theory of tree search

- Lower bounds
  - Optimal weights are often sparse
  - hints at **pruning**
  - computational challenges
- Well-developed understanding of depth 2 [GKK'15]
- Upgrading efficient algorithms [THT'14, GKK'15, KK'17]

Applications beyond robust statistics: