TrueSkill and AdPredictor: Machine Learning in the Wild

Thore Graepel

Online Services and Advertising Research

Microsoft Research Cambridge, UK

Joint work with Thomas Borchert, Joaquin Quiñonero Candela, Ralf Herbrich, and Tom Minka

WAPA 2010 – Cumberland Lodge

Outline

The Role of Applications in Machine Learning Research

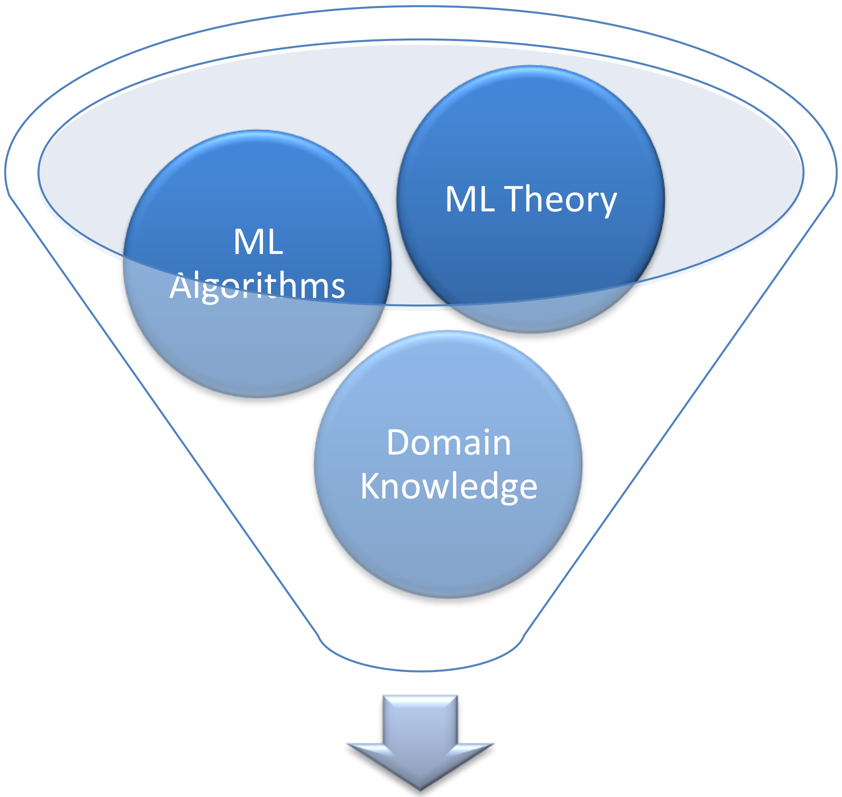
TrueSkill™: Ranking and Matchmaking in Online Games

- Ranking and Matchmaking Task
- TrueSkill Model and Inference
- Empirical Results from Xbox Live and Halo 3
- Specific Challenges in Online Gaming

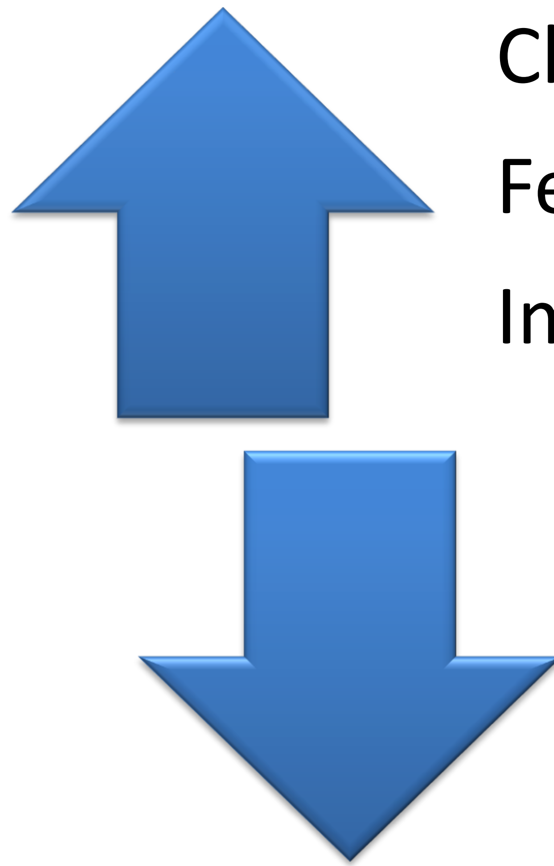
AdPredictor™: Click-Through Rate Prediction for Sponsored Search

- Paid Search Advertising and Click-Through Rate prediction
- AdPredictor Model and Inference
- Results from the Bing Search Engine
- Specific Challenges in Paid Search Advertising

Applications in Machine Learning (ML)



ML Applications



Challenges

Feedback

Impact

Insights

Methods

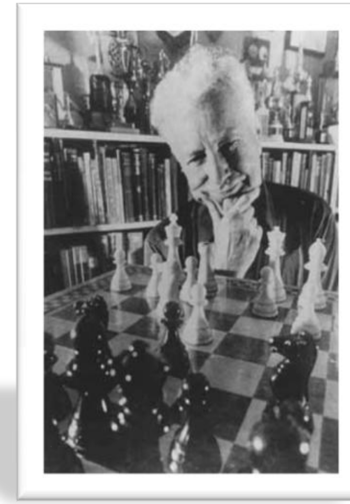
Experience

Only real-world impact of ML work is through applications!

Thore Graepel, Ralf Herbrich, Tom Minka, and our friends from Xbox Live

TRUESKILL

Ranking and Matchmaking



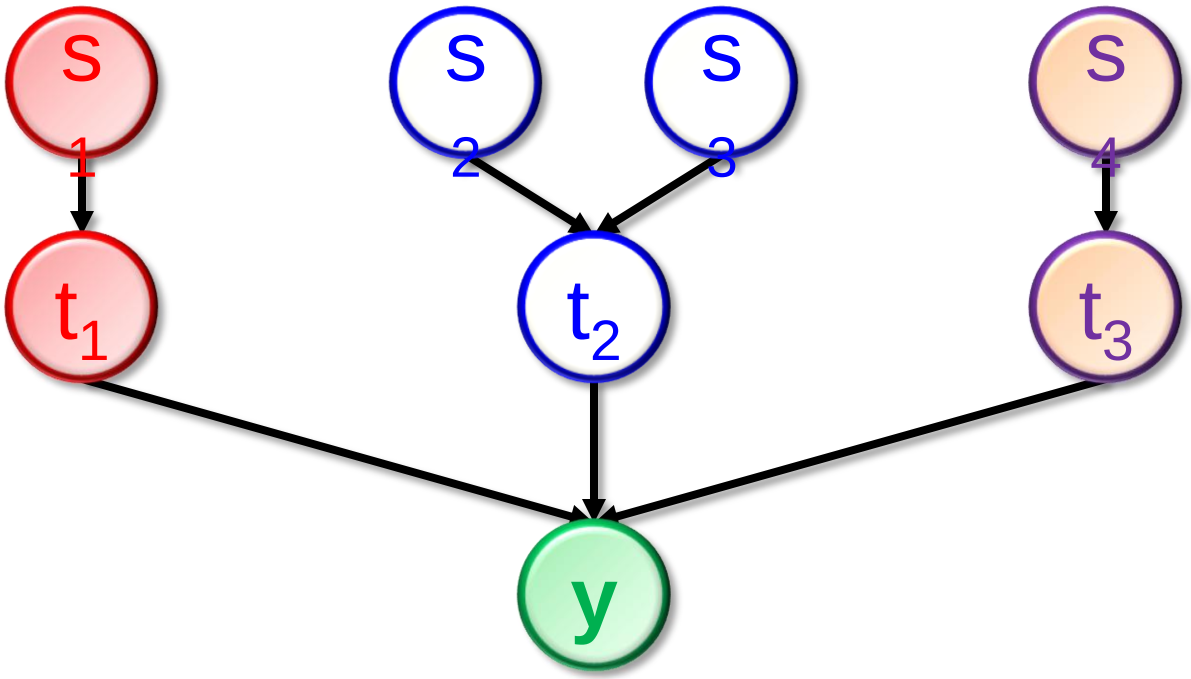
- Competition is central to our lives
  - Evolutionary principle
  - Driving principle of many sports
- Chess Rating for fair competition
  - ELO: Developed in 1960 by Árpád Imre Élő
  - Matchmaking system for tournaments
- Challenges of online gaming
  - Learn from few match outcomes efficiently
  - Support multiple teams and multiple players per team

The Skill Rating Problem

- **Given:**
  – Match outcomes: Orderings among k teams consisting o… …tively
- **Questions:**
  – Skill s_i for e…
  – Gl…
  – Fai…

Multiple Team Match Outcome Model

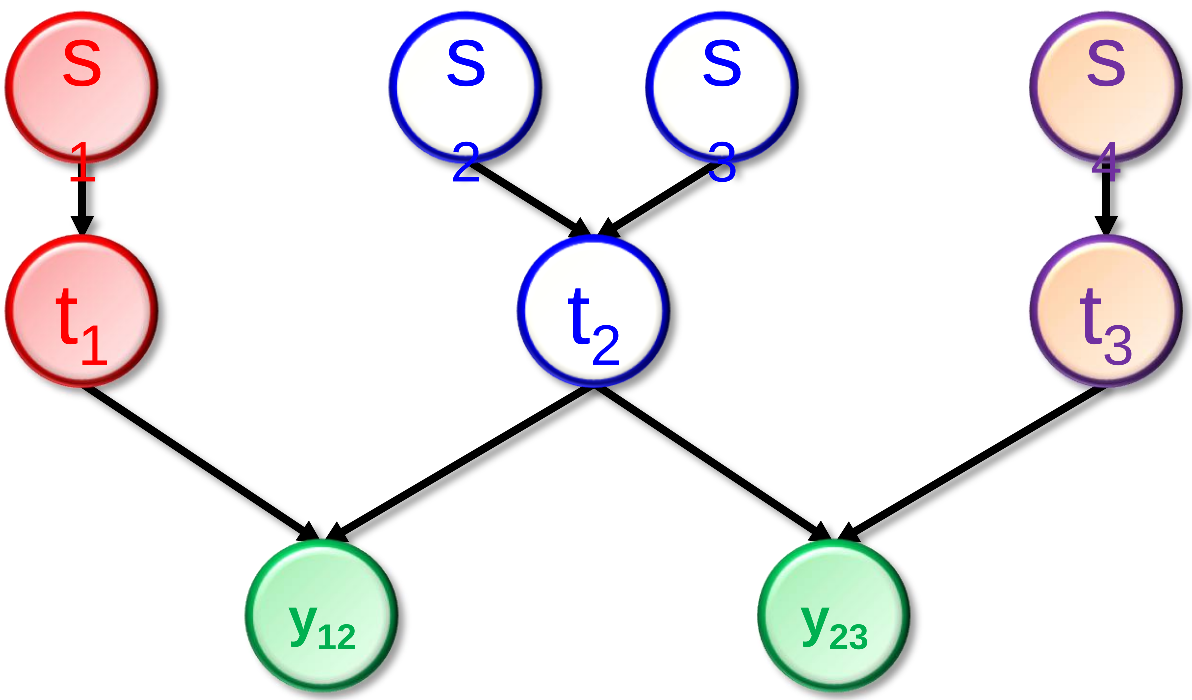
- Possible outcomes: Permutations of the teams



$$P(y|t_1, t_2, t_3) = \mathbb{I}(y = (i, j, k)) \text{ s.t. } t_i > t_j > t_k$$

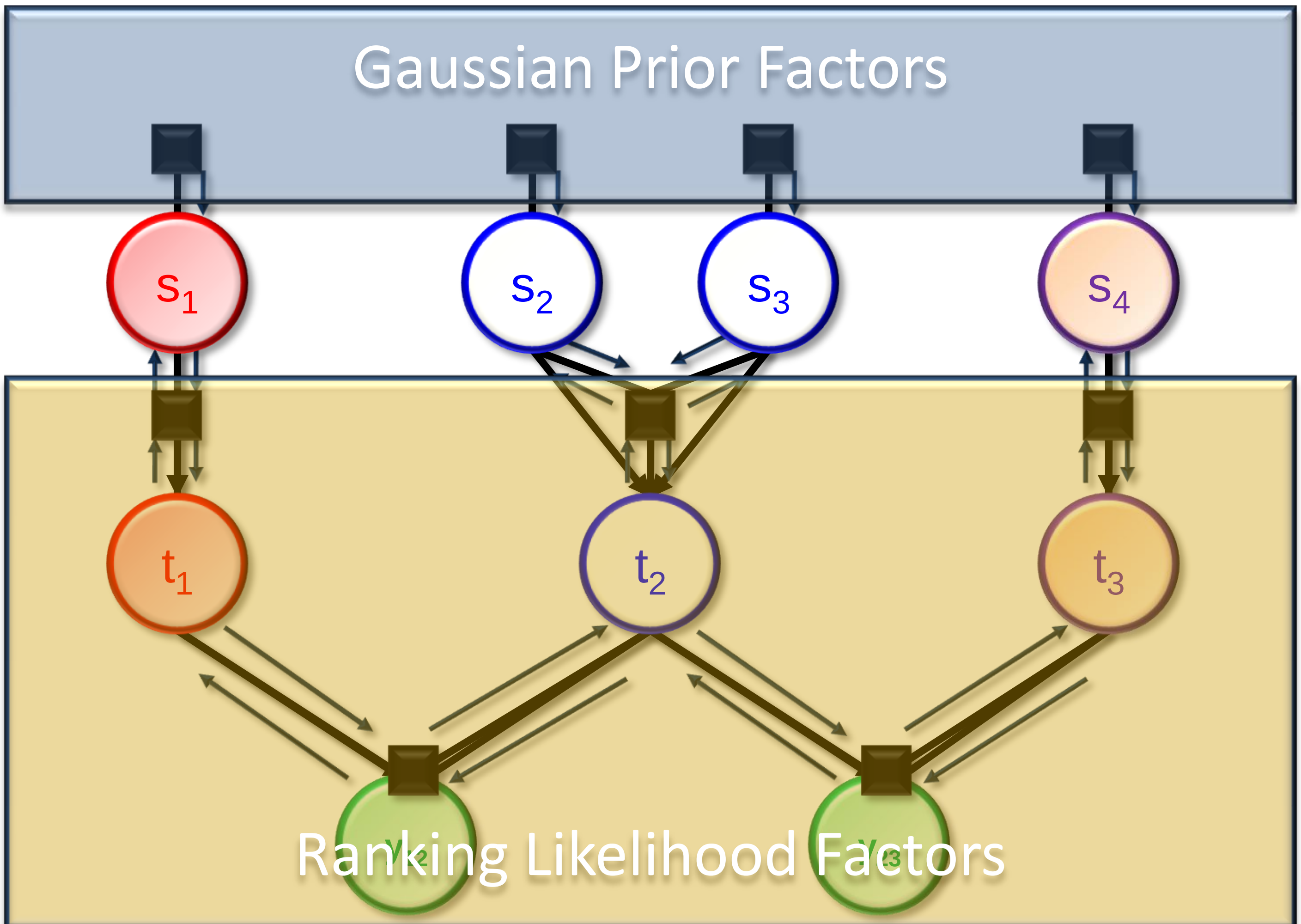
Multiple Team Match Outcome Model

- Observed outcome (1,2,3), use transitivity!
- Skill posterior $P(s_i | y = (1,2,3)) \approx N(\mu_i, \sigma_i^2)$



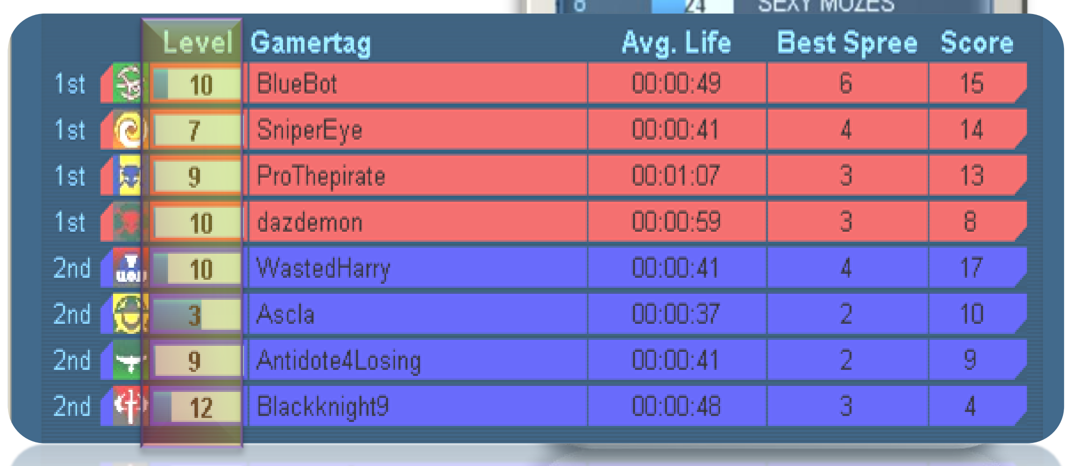
$y_{12} = (1,2)$ $y_{23} = (2,3)$

Efficient Approximate Inference

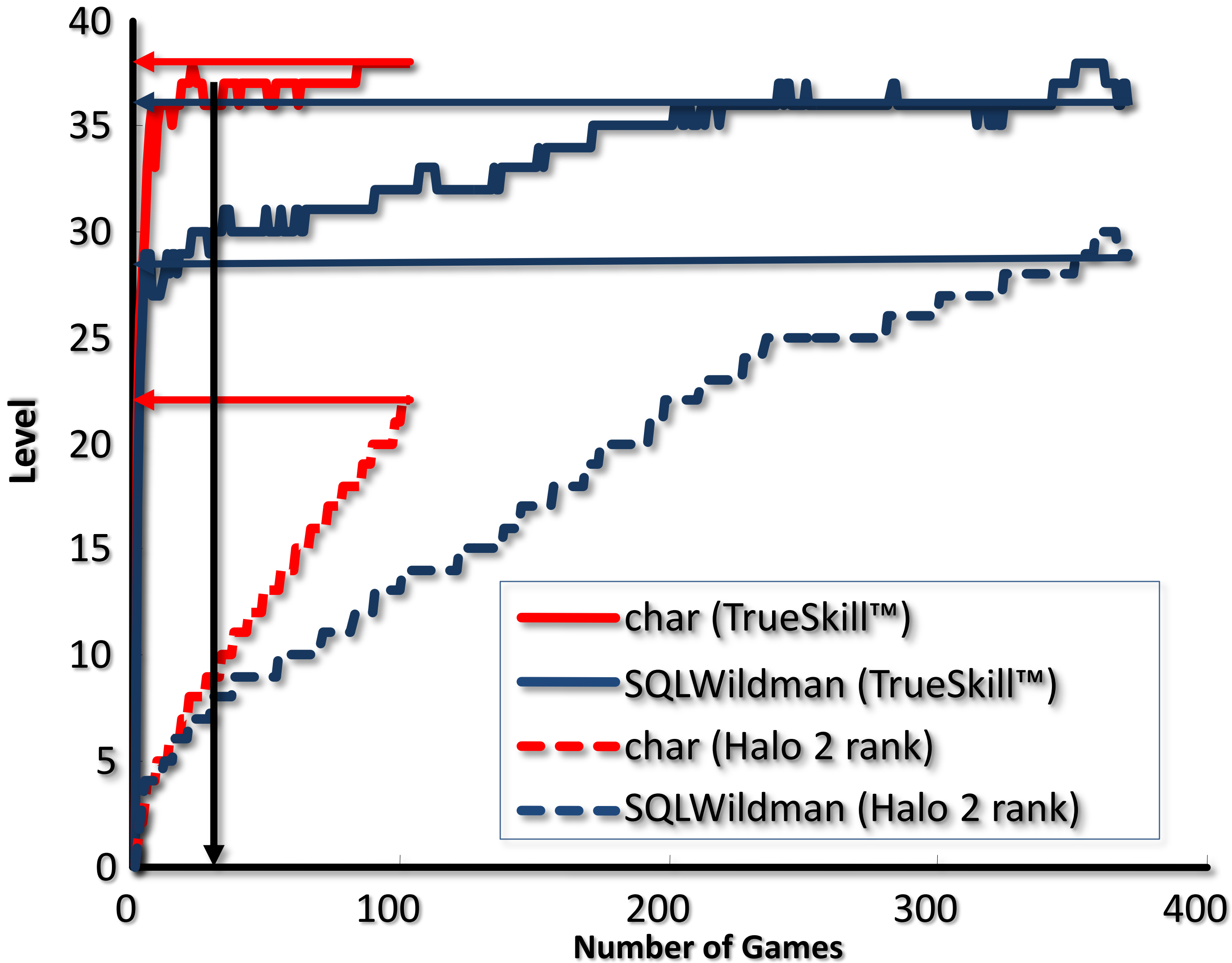


Applications to Online Gaming

- **Leaderboard**
  - Global ranking of all players
  - Rank conservatively according to $\mu_i - 3 \cdot \sigma_i$
- **Matchmaking**
  - For gamers: Most uncertain outcome
  - For inference: Most informative
  - Equivalent!

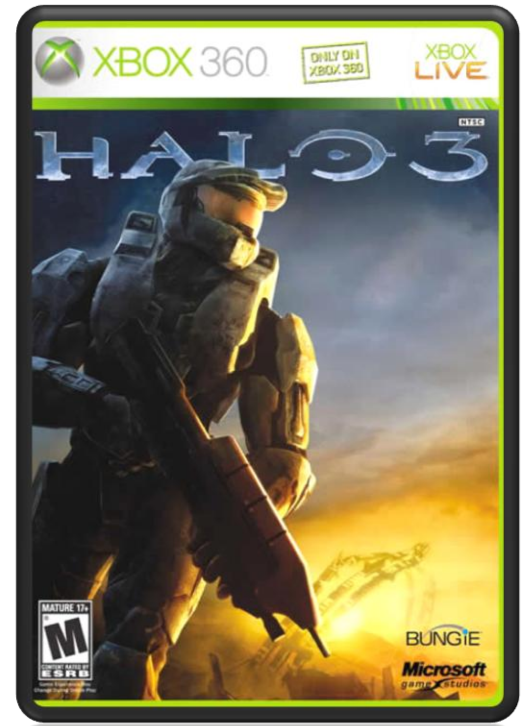


Convergence Speed

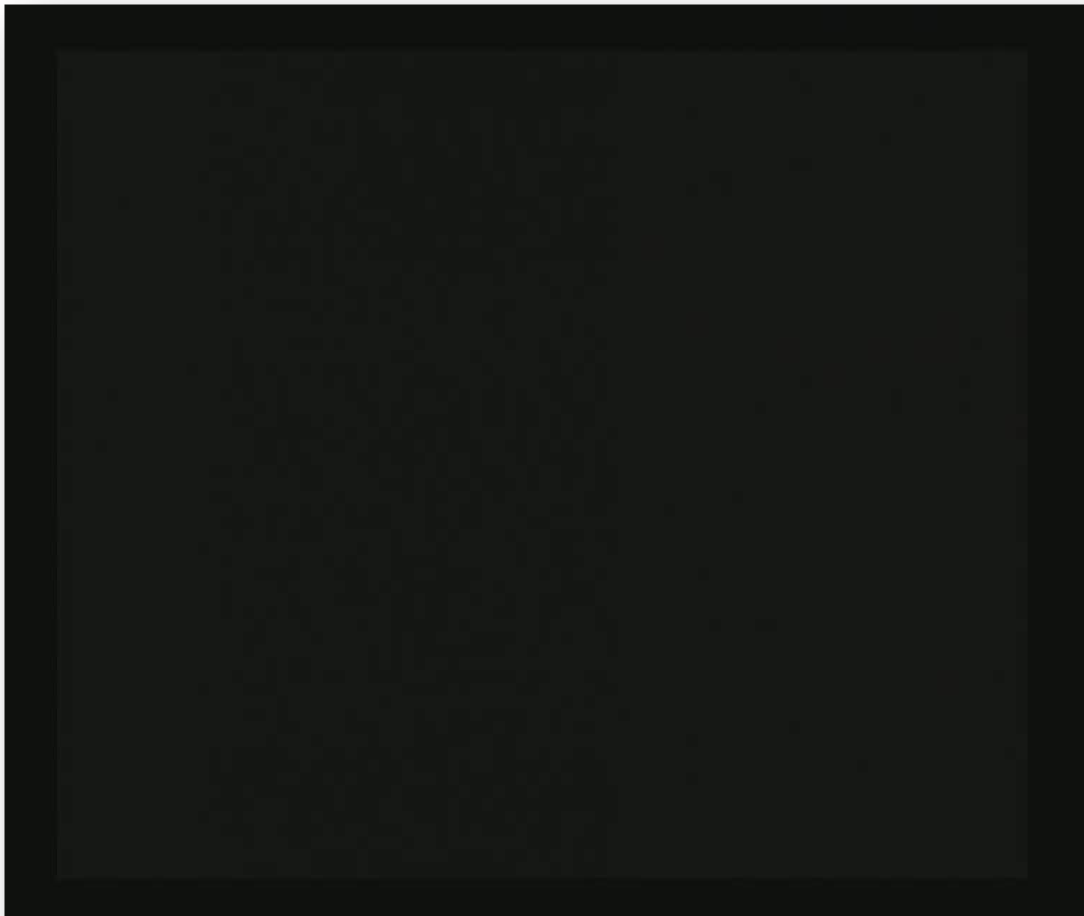


Xbox 360 & Halo 3

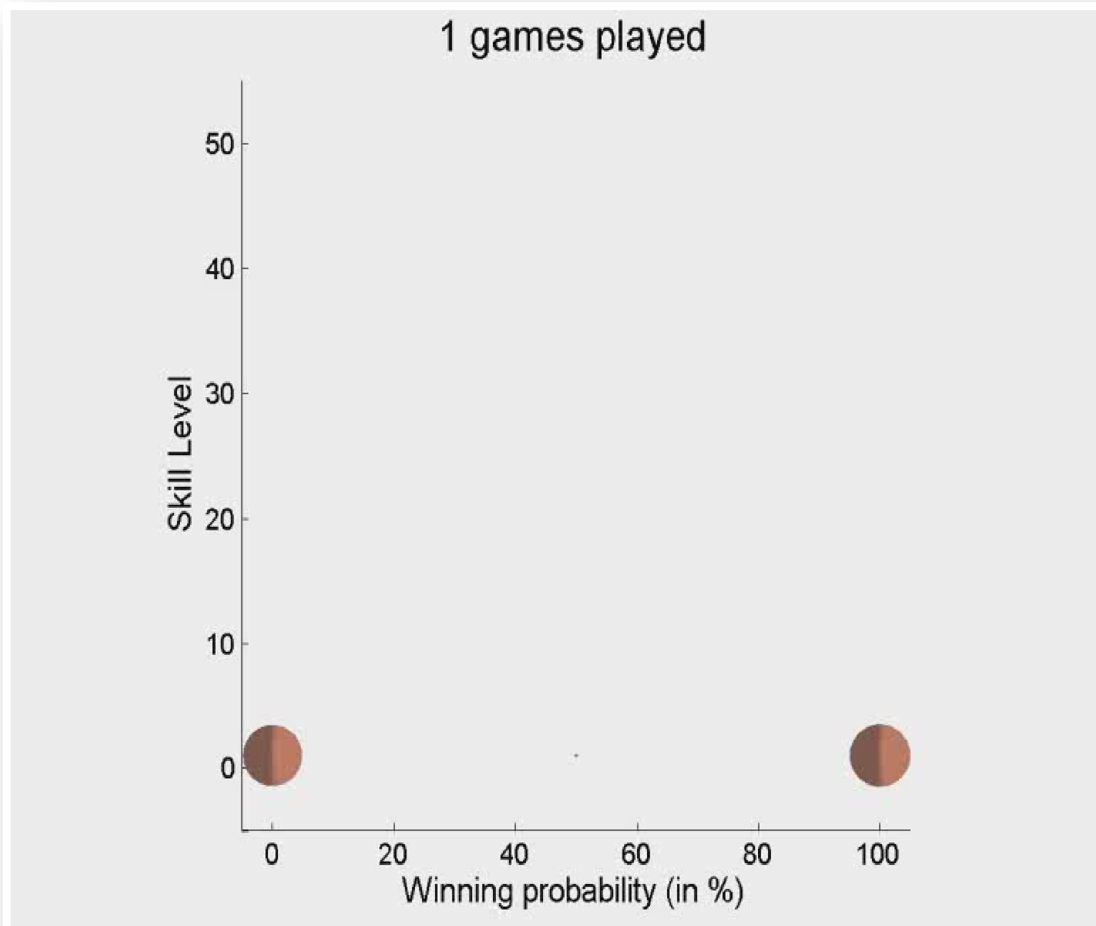
- **Xbox 360 Live**
  - Launched in September 2005
  - Every game uses TrueSkill™ to match players
  - > 20 million players
  - > 2 million matches per day
  - > 2 billion hours of game play
- **Halo 3**
  - Launched on 25th September 2007
  - Largest entertainment launch in history
  - > 200,000 player concurrently (peak: 1,000,000)



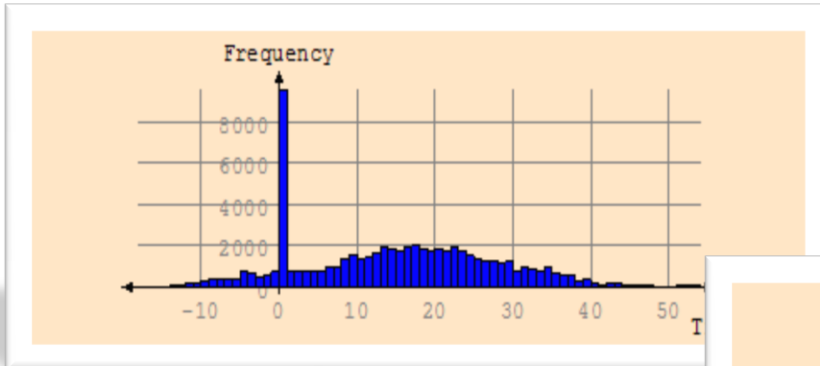
Halo 3 in Action



Halo 3 Analysis: Fair Matches?

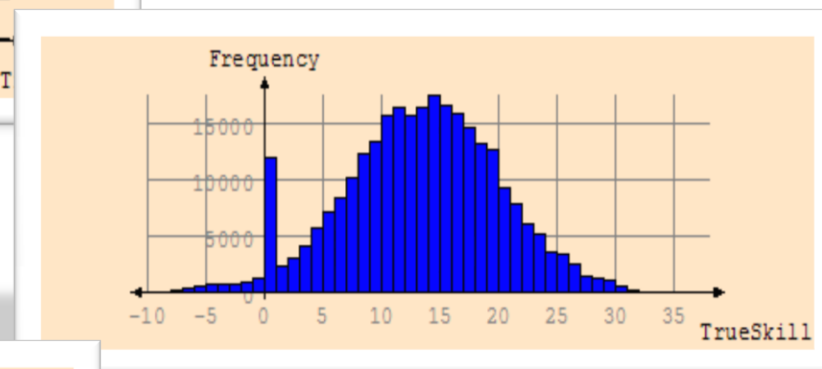


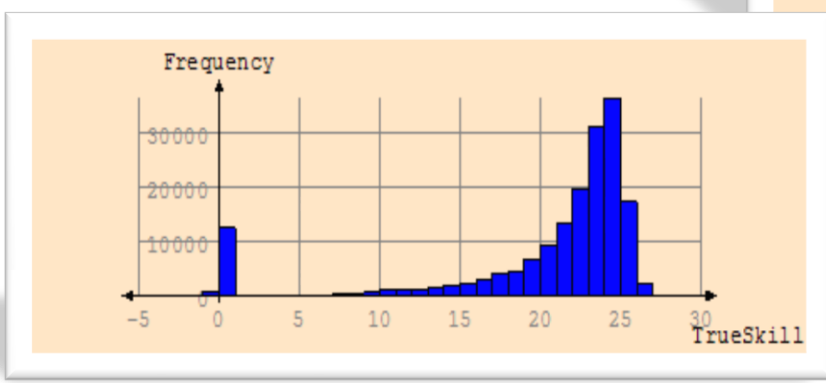
Skill Distributions of Online Games



Golf (18 holes): 60 levels

Car racing (3-4 laps): 40 levels





UNO (chance game): 10 levels

Special Requests from Bungie

- Great, but could you please
  - create a flat skill distribution for more plausible matchmaking!
  - make the skill convergence slower for hill-climbing experience!
  - reduce the updates if low quality matches or bad connections are detected!
  - factor in the time a player actually participates in a match?

Adversarial Behaviour

- Skill Boosting
  - Players collude by choosing a common, but rare language that ensures they are matched and can arrange the game outcome.
  - Players “de-rank” accounts and play in teams with the de-ranked account. When their team wins, their other account’s rank is boosted!
- Now paid services available for boosting players ranks!

Incentive Alignment

- TrueSkill only takes into account the final game outcome at team level
  - No individual performance is counted
  - No proxies for skill (e.g., number of kills, flags etc) are taken into account.
- Why?
  - TrueSkill is universal and can be applied to almost any known game
  - Ratings are aligned with the objective of the game. In team games, players have no incentives to promote their own ranking at the cost of the team.

Closed Training Loop

- TrueSkill was developed and tested based on real-world training data from Halo 2.
- Since TrueSkill's estimates are used for deciding which players to match next:
  - The composition of the training sample is changed by the earlier estimates of the system
  - A loop is created in which TrueSkill feeds on data from matches it has created
  - The prediction performance of TrueSkill may go down over time because it creates harder and harder prediction tasks for itself!

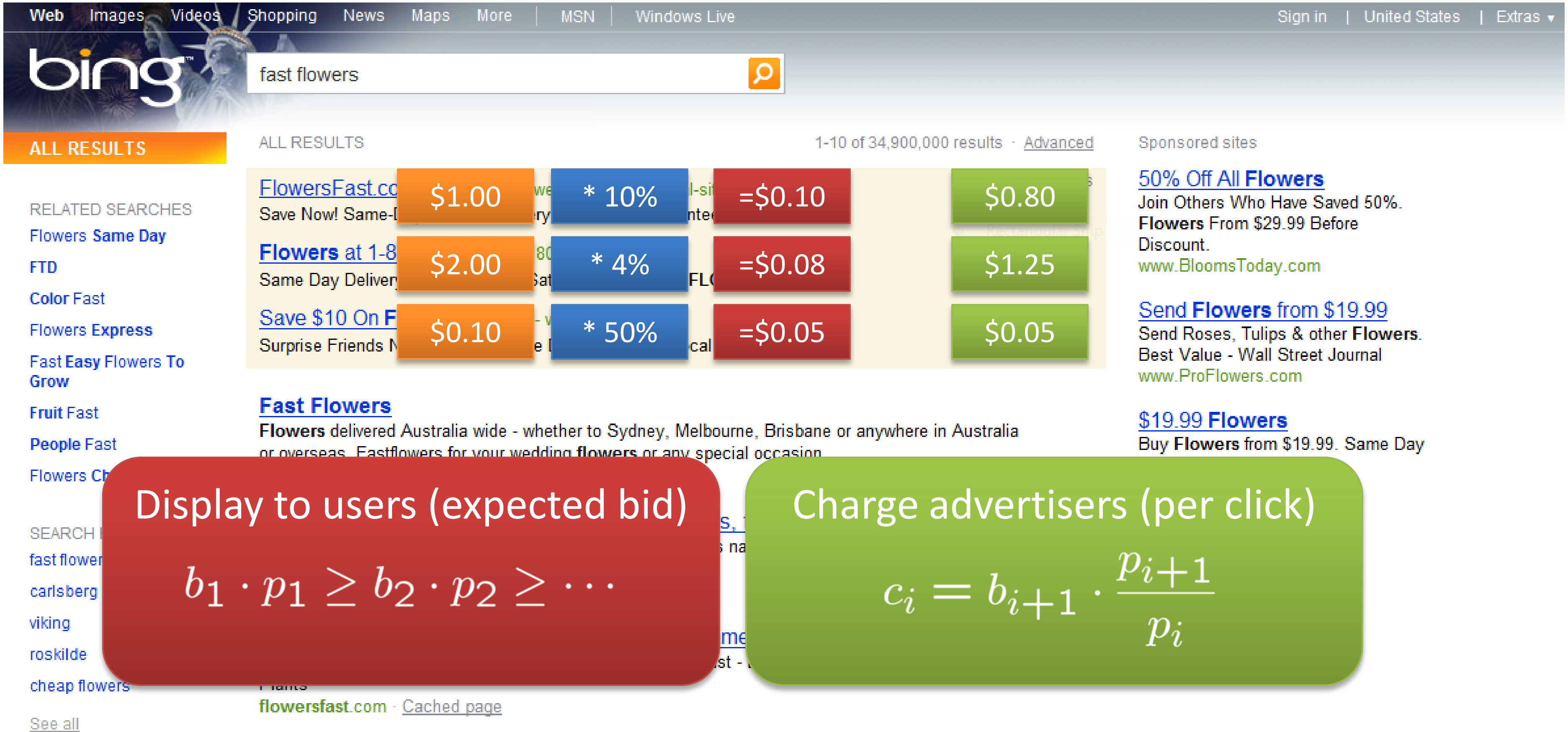
My former boss Bill…

Why are they wasting their time on computer games?

Great stuff, guys, but how about online advertising? Are the ads not competing for the attention of the users just like players in a game?!

Thore Graepel, Joaquin Quiñonero Candela, Thomas Borchert, Ralf Herbrich & our friends from Microsoft adCenter

ADPREDICTOR



Display to users (expected bid)

$$b_1 \cdot p_1 \geq b_2 \cdot p_2 \geq \cdots$$

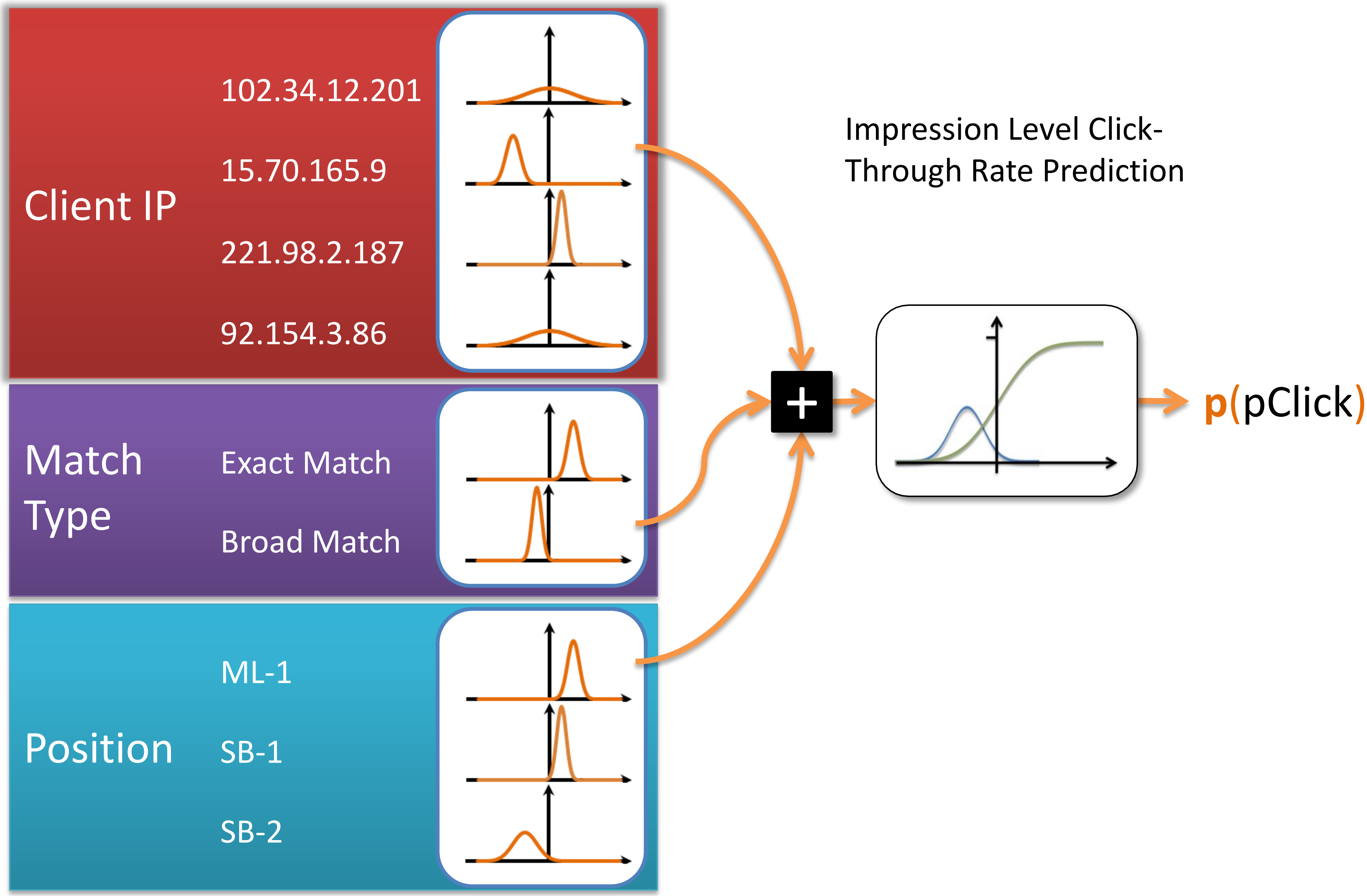
Charge advertisers (per click)

$$c_i = b_{i+1} \cdot \frac{p_{i+1}}{p_i}$$

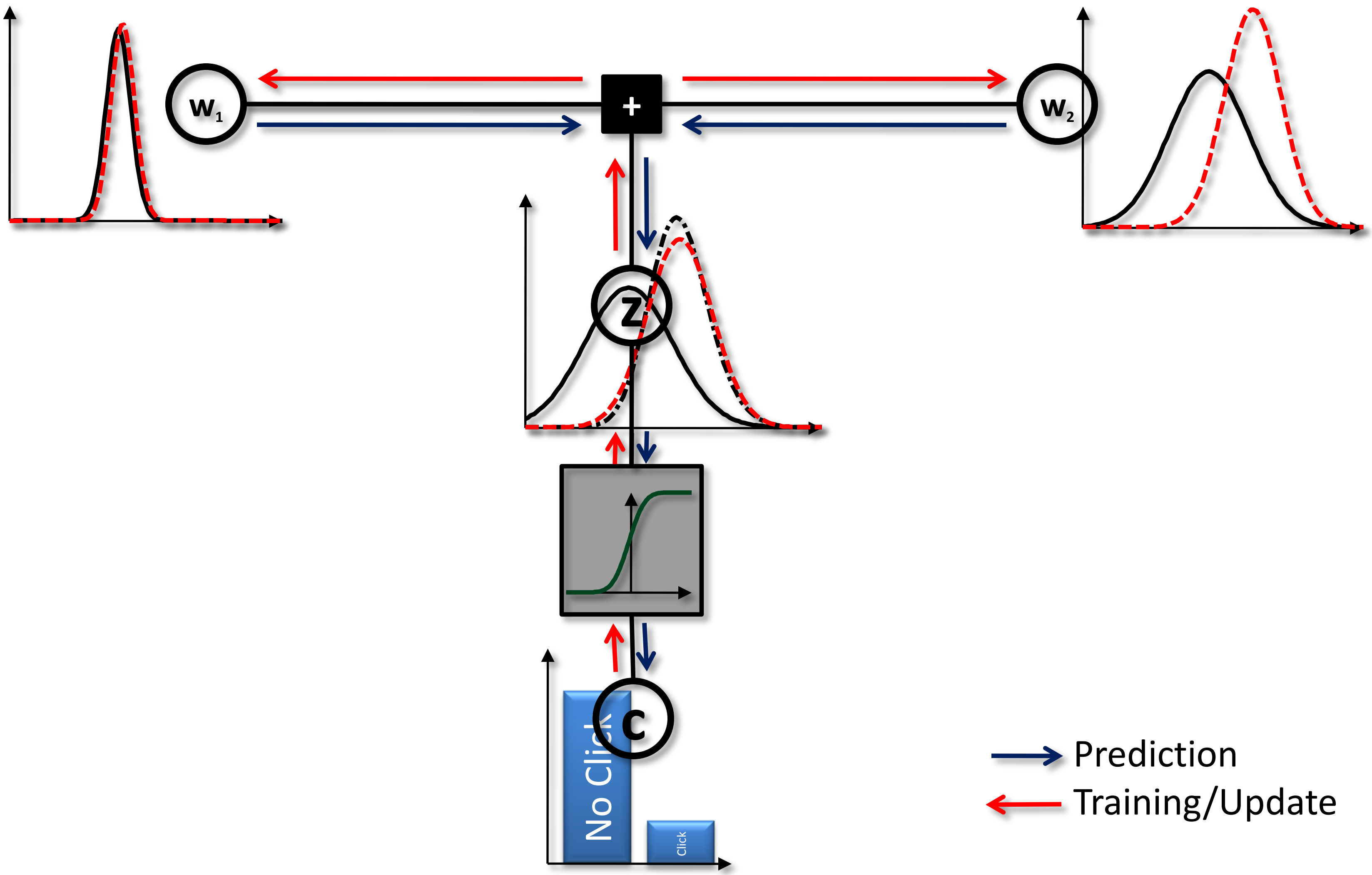
Importance of accurate probability estimates:

- Increase user satisfaction by better targeting
- Provide better deal to advertisers
- Increase revenue by showing ads with high click-thru rate

AdPredictor: Bayesian Probit Regression



Training Algorithm in Action



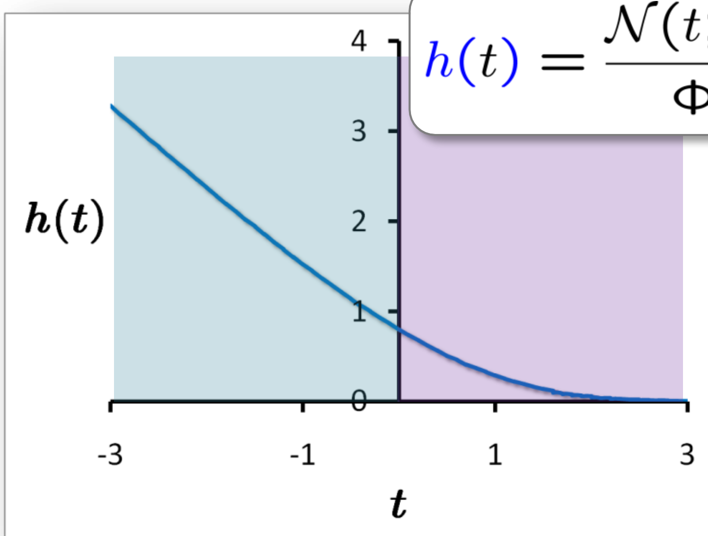
Closed Form Updates for Click

$$\mu_i \leftarrow \mu_i + \frac{\sigma_i^2}{s} \cdot h\left[\frac{\sum_{j=1}^{d} \mu_j}{s}\right]$$

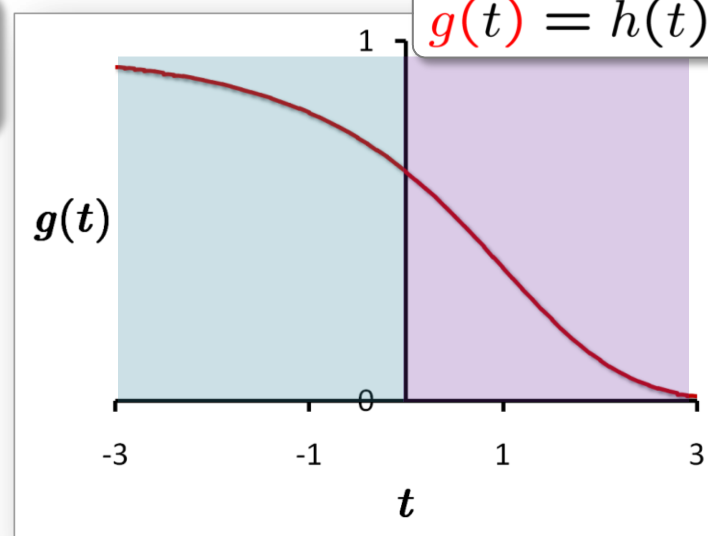
$$\sigma_i^2 \leftarrow \sigma_i^2 \left(1 - \frac{\sigma_i^2}{s^2} \cdot g\left[\frac{\sum_{j=1}^{d} \mu_j}{s}\right]\right)$$

$$s^2 = \beta^2 + \sum_{j=1}^{d} \sigma_j^2$$

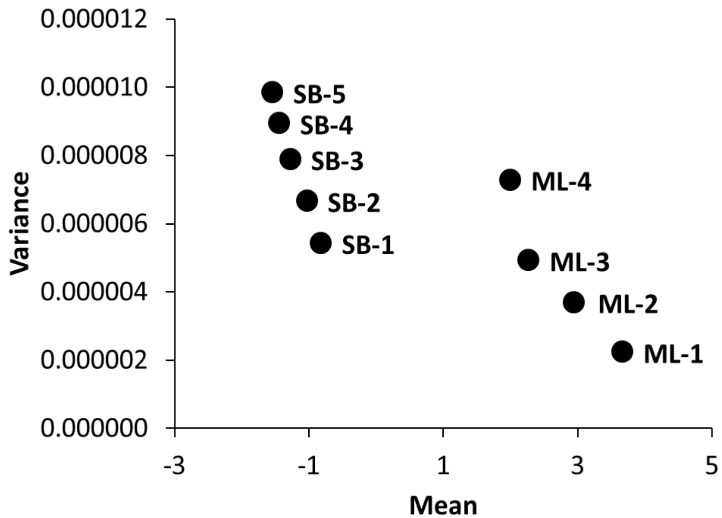
$$h(t) = \frac{\mathcal{N}(t;\, 0, 1)}{\Phi(t)}$$



$$g(t) = h(t) \cdot [h(t) + t]$$

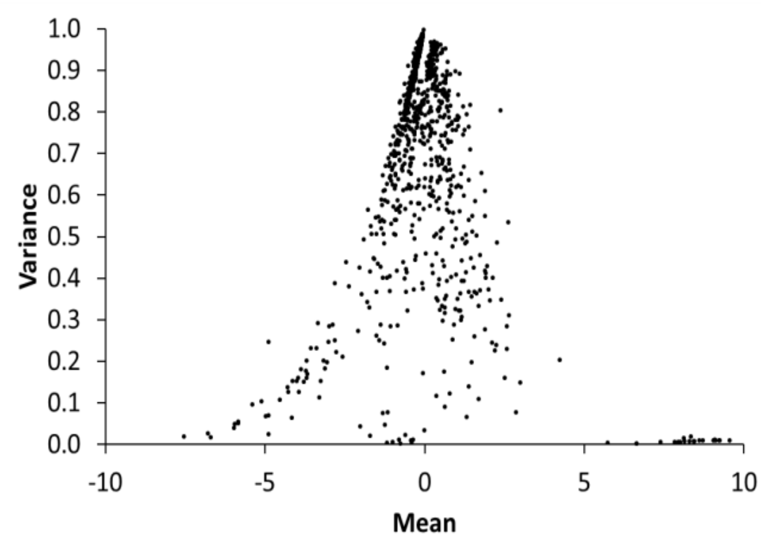


Weight Parameters from Production

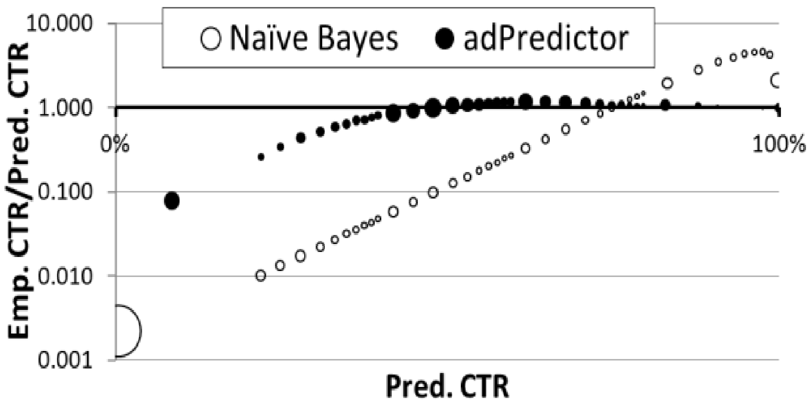


- Means and variances of the display position feature.
- Mainline ML, sidebar SB

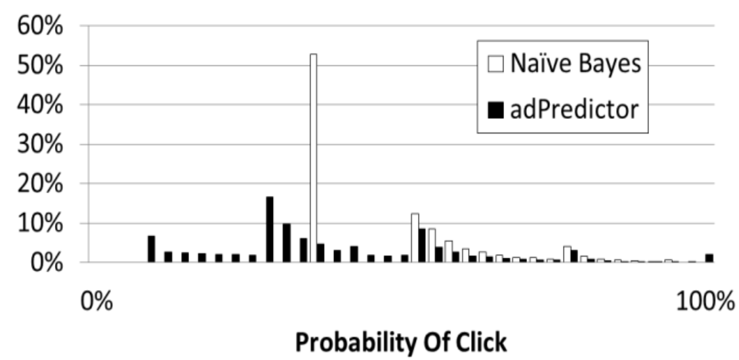
- Means and variances of weights for user ID feature.
- Very high mean may indicate fraud



Comparison with Naïve Bayes



Calibration: AdPredictor vs. Naïve Bayes



Courage: AdPredictor vs. Naïve Bayes

$$RIG := \frac{CE + H(\bar{p})}{H(\bar{p})}$$

$$CE := \frac{1}{T}\sum_{t=1}^{T} y_t \log \hat{p}_t + (1 - y_t) \log(1 - \hat{p}_t)$$

$$H(\bar{p}) := -(\bar{p} \log \bar{p} + (1 - \bar{p}) \log(1 - \bar{p})).$$

Algorithm	RIG	AUC
adPredictor	**61.42%**	**95.6%**
adPredictor (calibrated)	61.35%	**95.6%**
Naïve Bayes	-41.54%	89.4%
Naïve Bayes (calibrated)	33.86%	89.3%

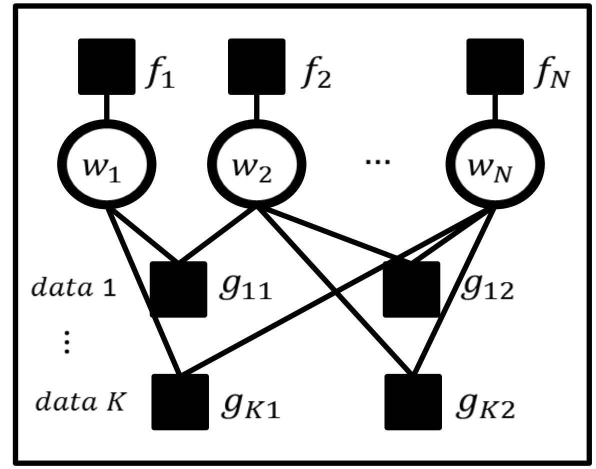
Web Scale Implementation

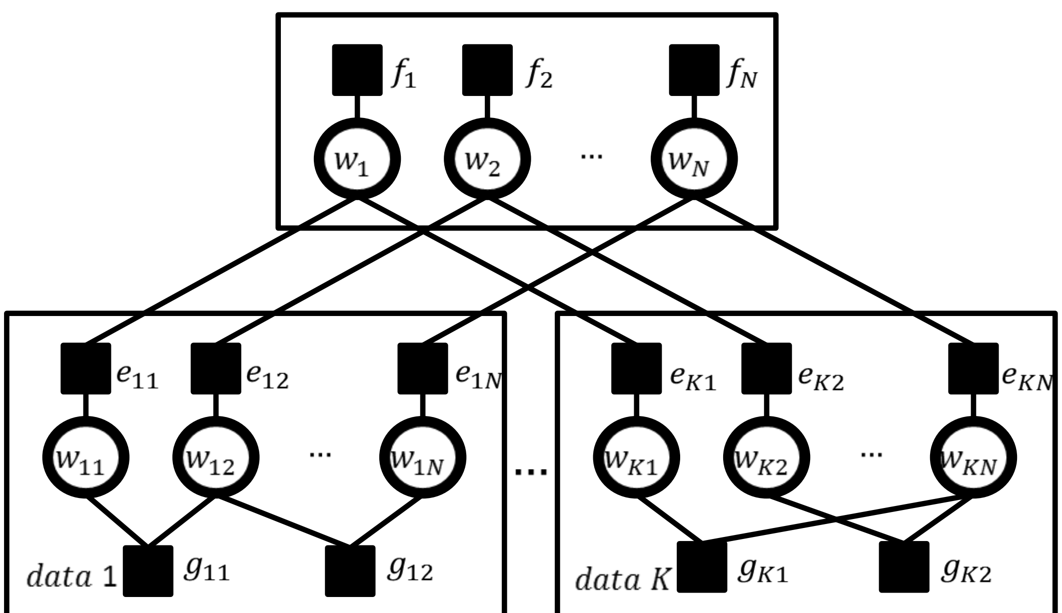
AdPredictor drives almost 100% of Bing's sponsored search traffic

Large volume of data $O(10^9)$ examples per week →Massively parallel training	System's predictions determine composition of future training sample →Exploration	Continuous Model Training →Maintain controlled memory footprint by pruning

Distributed Training

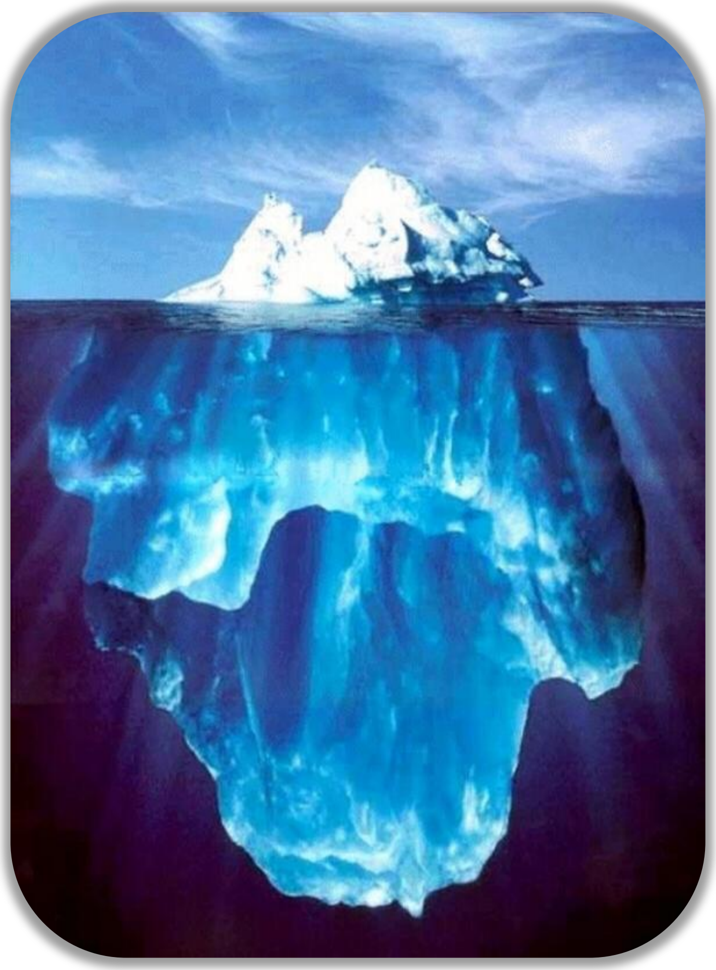
- Single instance of model leads to inefficient blocking

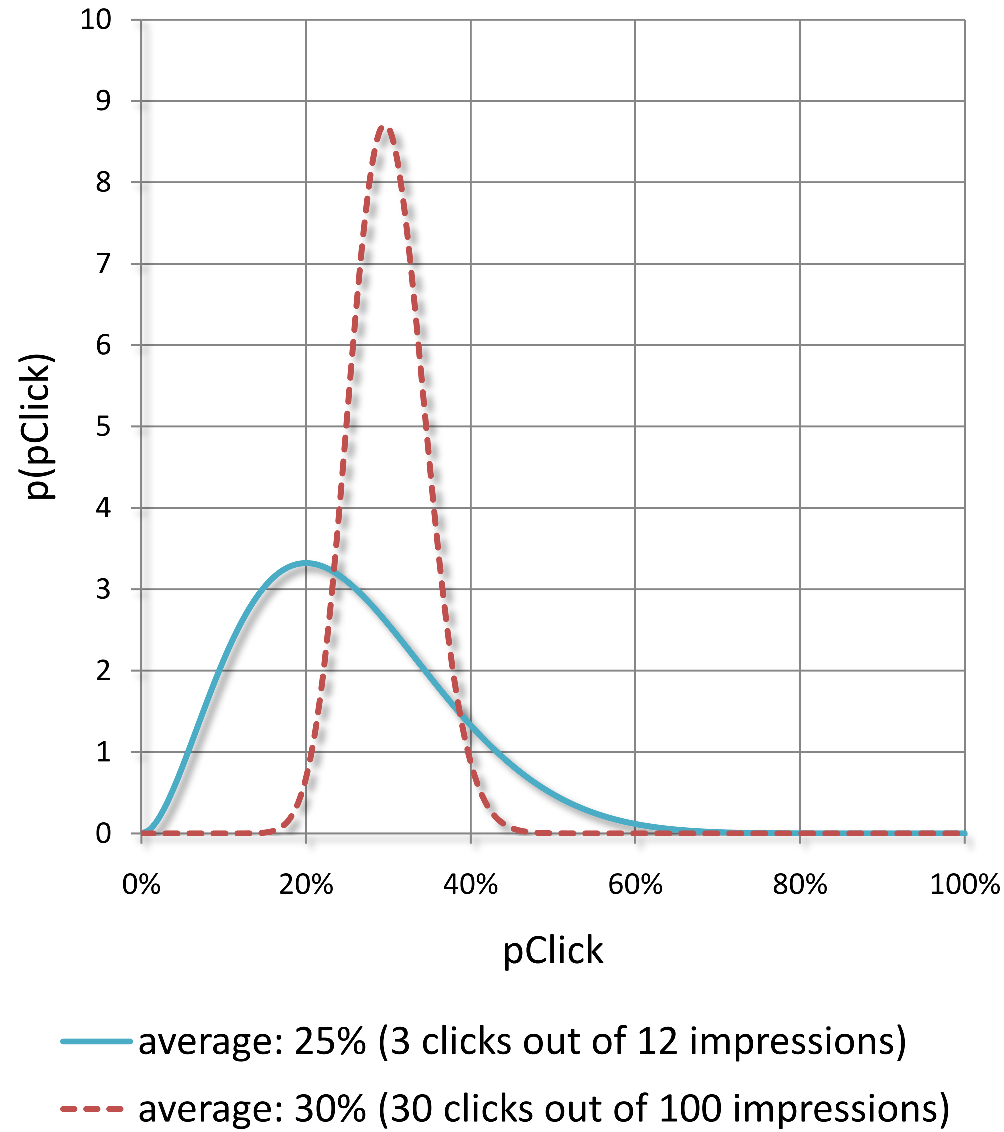




- Model replication with equality factors for consistency → Distributed Training

Principled Exploration





Adversarial Behaviour

- Click-fraud by using bots or sweat-shops to deplete competitor's budget
- Advertisers may benefit from the in-dubio-pro-reo attitude of the exploration system
- Possible collusion among advertisers by sharing market (phase of the moon) or manipulating bids

Closed Training System

- AdPredictor was designed based on static training data from its predecessor
- Since AdPredictor's CTR estimates are used for deciding which ads to show next:
  - The composition of the training sample is changed by the earlier estimates of the system
  - A loop is created in which AdPredictor feeds on data from ads shown based on its estimates
  - Static performance measures are meaningless in the loopy setting!
  - Difference to TrueSkill: Ads are shown based on high CTR, while matches are created based on similar skill! In both cases: Source entropy up → more difficult tasks

Conclusions

Large Scale Bayesian Models Work in Practice!

- Model uncertainty explicitly
- Automatically adapt learning step size
- Make accurate, calibrated predictions

Online Applications

- Large scale requires parallel training and model pruning
- Closed loop requires active learning and exploration
- Watch out for self-interested agents in the system
- The application is always different from the abstraction!

Thank you!

thoreg@microsoft.com