Counterfactual Reasoning and Learning Systems

LÉON BOTTOU

MSR

Joint work with

Jonas Peters

Denis Xavier Charles

Elon Portugaly

Patrice Simard

Joaquin Quiñonero Candela

D. Max Chickering

Dipankar Ray

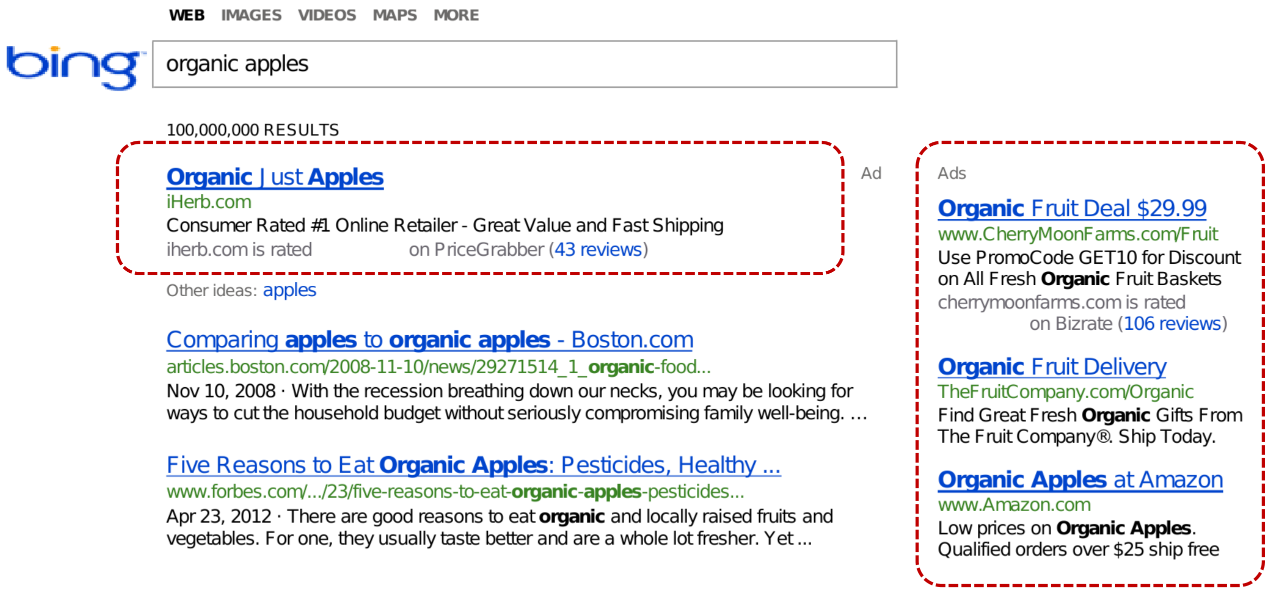
Ed Snelson

Summary

1. Background
2. Counterfactuals
3. Illustration
4. Toolbox
5. Equilibrium

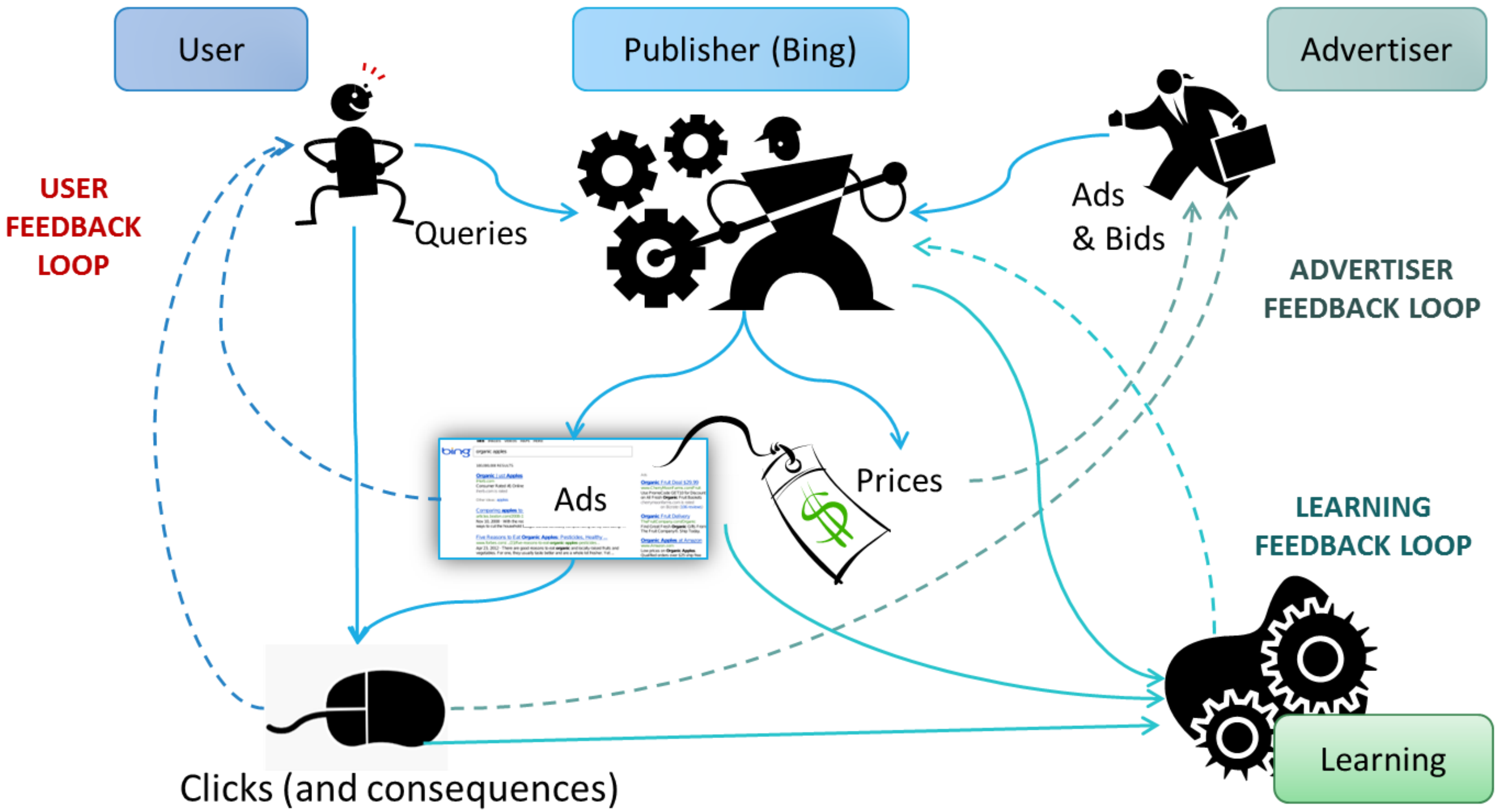
1. Background

The pesky little ads



- An example of real-life machine learning “system”.

Why is it difficult?



Feedback loops

Shifting distributions

- Data is collected when the system operates in a certain way.
  The observed data follows a first distribution.
- Collected data is used to justify actions that change the operating point.
  Newly observed data then follows a second distribution.
- **Correlations observed on data following the first distribution do not necessarily exist in the second distribution.**

Often lead to vicious circles..

Previous work

Auction theory and mechanism design

(Varian, 2007)
(Edelman et al., 2007)
...

- Motivate the design of ad auctions.
- Address the advertiser feedback loop.
- Assume single auction instead of repeated auctions.
- Assume click probabilities are known rather than estimated.
- Ignore impact of ads on future user engagement.
- Ignore how advertisers place a single bid valid for multiple auctions

No clear way to describe the full system
as auctions amenable to theoretical analysis.

Previous work

Multi-armed bandits and extensions

(Robbins, 1952)
(Lai & Robbins, 1985)
(Auer et al., 2002)
(Langford & Zhang, 2007)
(Li et al., 2010)
...

- Regret bounds for “multi armed bandits” (MAB)
- Regret bounds for “contextual bandits” (CB)
- Address the learning feedback loop
- Address the explore/exploit tradeoff
- Ignore impact of ads on future user engagement
- Ignore impact of ad placement on future advertiser bids

No clear way to reduce the full system into MAB or CB problems amenable to theoretical analysis

This work

Causal inference viewpoint

- Changing the ad placement algorithms is an *intervention* on the system.
- We track the consequences of such interventions along the paths of the causal graph.

✓ Much more flexible framework
✓ Leads to powerful estimation methods
✓ Leads to novel learning algorithms
✓ Now in daily use

Overkill?

Pervasive causation paradoxes in ad data

Example:

- Logged data shows a positive correlation between event A *"First mainline ad receives a high score q_1"* and event B *"Second mainline ad receives a click"*.
- Controlling for event C *"Query categorized as commercial"* reverses the correlation for both commercial and noncommercial queries.

A	$P(B\|A)$	$P(B\|A,\neg C)$	$P(B\|A,C)$
q_1 low	124/2000 (6.2%)	92/1823 (5.1%)	32/176 (18.1%)
q_1 high	149/2000 (7.5%)	71/1500 (4.8%)	78/500 (15.6%)

Randomized experiments

How to compare two ad placement systems?

1. Randomly split traffic or users into buckets
2. Apply alternative placement algorithms to distinct buckets.
3. Wait a couple months and compare performance.

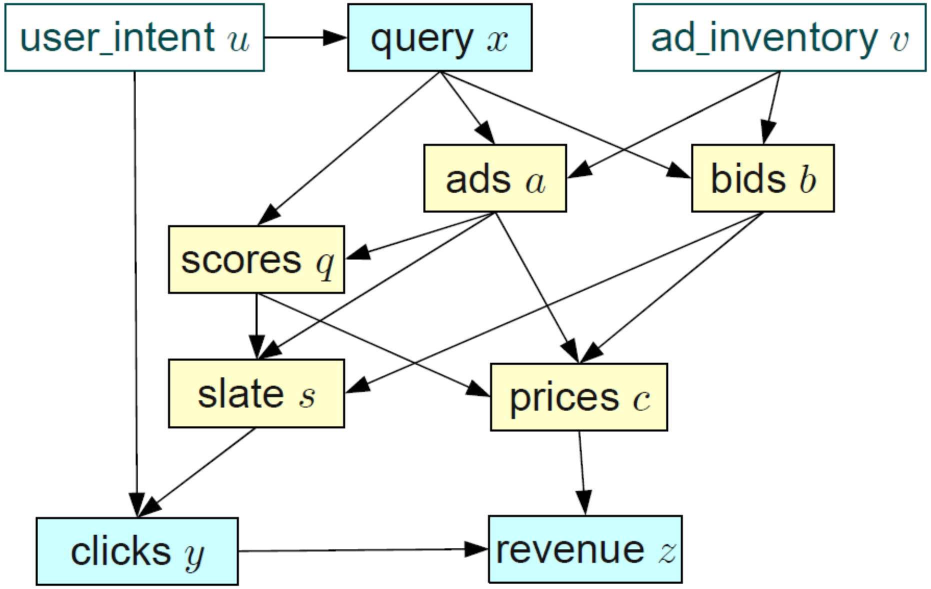
Issues

- Hard to control for advertiser effects
- Need full implementation and several weeks.
- Progress speed limited by available traffic.

We need an alternative.

Structural equation model (SEM)

$$
\begin{aligned}
x &= f_1(u, \varepsilon_1) \\
a &= f_2(x, v, \varepsilon_2) \\
b &= f_3(x, v, \varepsilon_3) \\
q &= f_4(x, a, \varepsilon_4) \\
s &= f_5(a, q, b, \varepsilon_5) \\
c &= f_6(a, q, b, \varepsilon_6) \\
y &= f_7(s, u, \varepsilon_7) \\
z &= f_8(y, c, \varepsilon_8)
\end{aligned}
$$

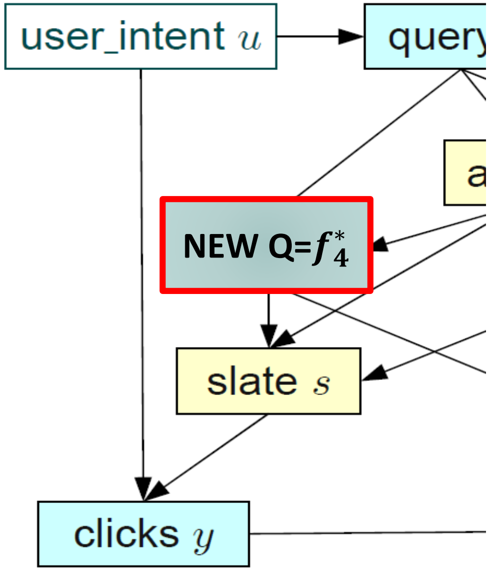


Direct causes / Known and unknown functions

Noise variables / Exogenous variables

Interventions

$$
\begin{aligned}
x &= f_1(u, \varepsilon_1) \\
a &= f_2(x, v, \varepsilon_2) \\
b &= f_3(x, v, \varepsilon_3) \\
q &= \sout{f_4(x, a, \varepsilon_4)} \quad f_4^*(x, a, \varepsilon_4) \\
s &= f_5(a, q, b, \varepsilon_5) \\
c &= f_6(a, q, b, \varepsilon_6) \\
y &= f_7(s, u, \varepsilon_7) \\
z &= f_8(y, c, \varepsilon_8)
\end{aligned}
$$



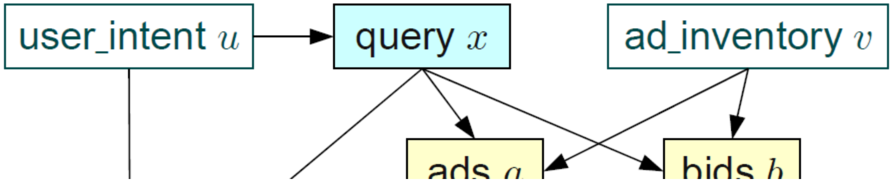
Interventions as algebraic manipulation of the SEM.
Causal graph must remain acyclic.

Isolation

What to do with unknown functions?

- Replace knowledge by statistics.
- Statistics need repeated isolated experiments.
- Isolate experiments by assuming an unknown but invariant joint distribution for the exogenous variables.

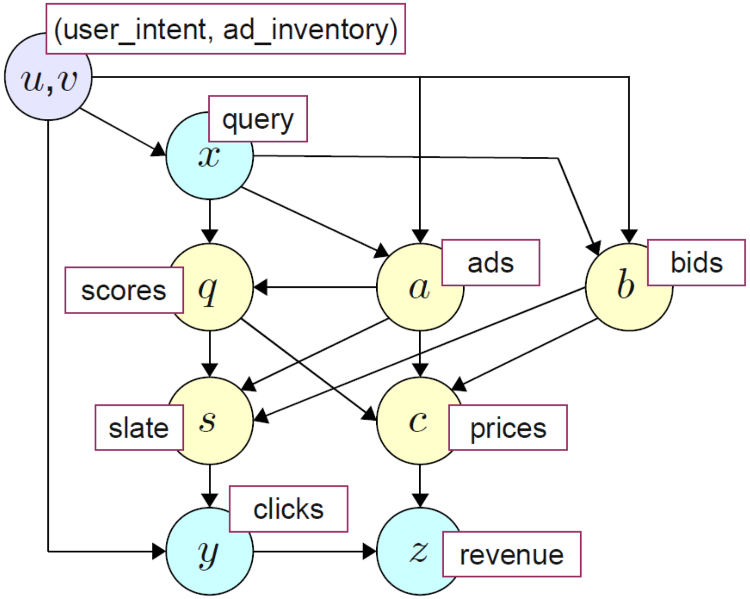
$$P(u, v)$$



⇒ No feedback loops (…yet…)

Markov factorization

$$
\begin{aligned}
\mathbf{P}(\omega) = \quad & \mathbf{P}(u, v) \\
& \times \mathbf{P}(x \mid u) \\
& \times \mathbf{P}(a \mid x, v) \\
& \times \mathbf{P}(b \mid x, v) \\
& \times \mathbf{P}(q \mid x, a) \\
& \times \mathbf{P}(s \mid a, q, b) \\
& \times \mathbf{P}(c \mid a, q, b) \\
& \times \mathbf{P}(y \mid s, u) \\
& \times \mathbf{P}(z \mid y, c)
\end{aligned}
$$



This is a “Bayes network” (Pearl, 1988)

a.k.a. “directed acyclic probabilistic graphical model.”

Markov interventions

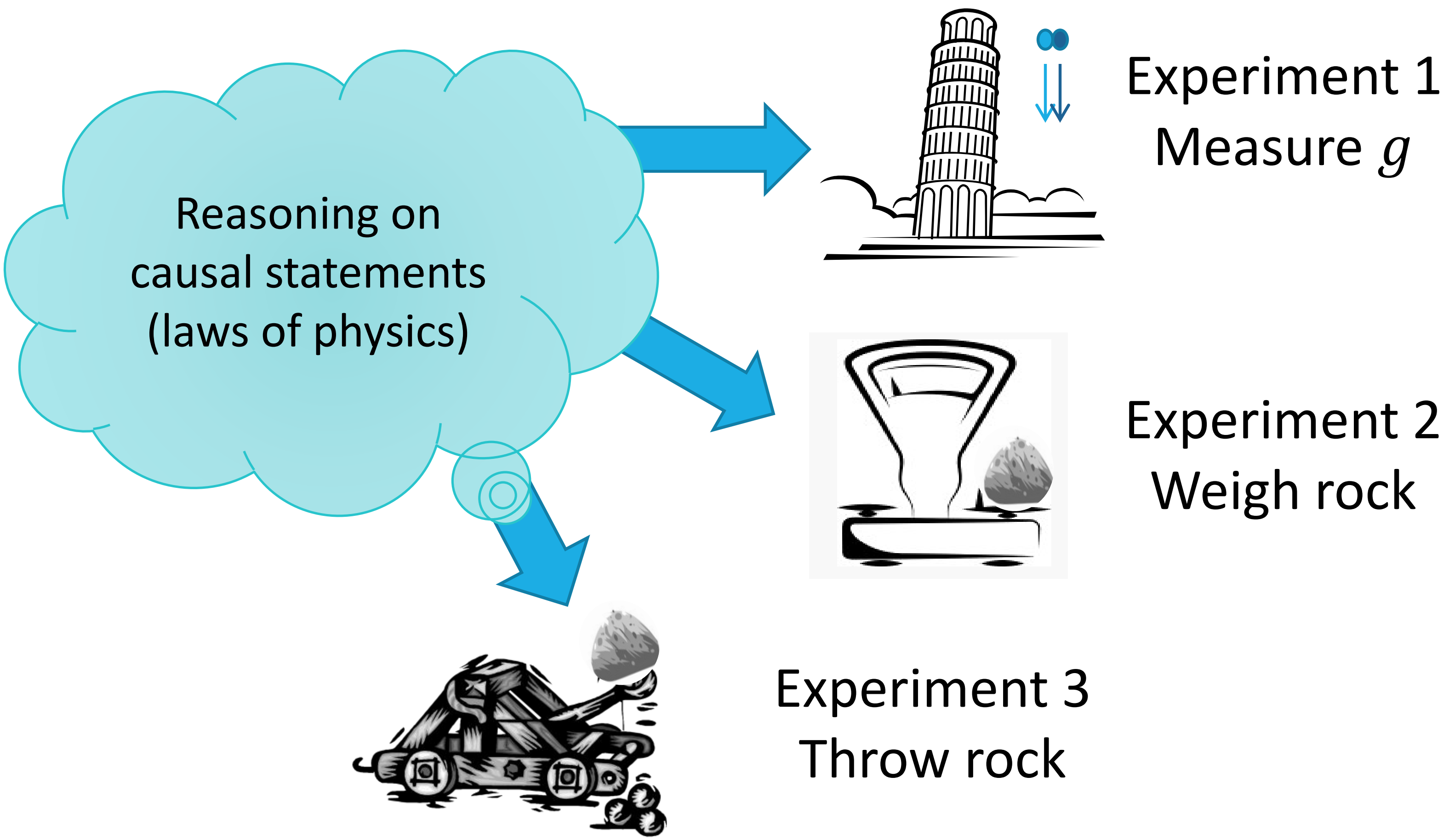
$$
\begin{aligned}
\mathrm{P}^{*}(\omega) = \quad & \mathrm{P}(u, v) \\
& \times \mathrm{P}(x \mid u) \\
& \times \cancel{\mathrm{P}(a \mid x, v)} \quad \mathrm{P}^{*}(q \mid x, a) \\
& \times \mathrm{P}(b \mid x, v) \\
& \times \mathrm{P}(q \mid x, a) \\
& \times \mathrm{P}(s \mid a, q, b) \\
& \times \mathrm{P}(c \mid a, q, b) \\
& \times \mathrm{P}(y \mid s, u) \\
& \times \mathrm{P}(z \mid y, c)
\end{aligned}
$$

Distribution under intervention

Many related Bayes networks are born (Pearl, 2000)

- They are related because they share some factors.
- More complex algebraic interventions are of course possible.

Transfer learning on steroids



2. Counterfactuals

Counterfactuals

Measuring something that did not happen

"How would have the system performed if, when the data was collected, we had used scoring model M' instead of model M?"

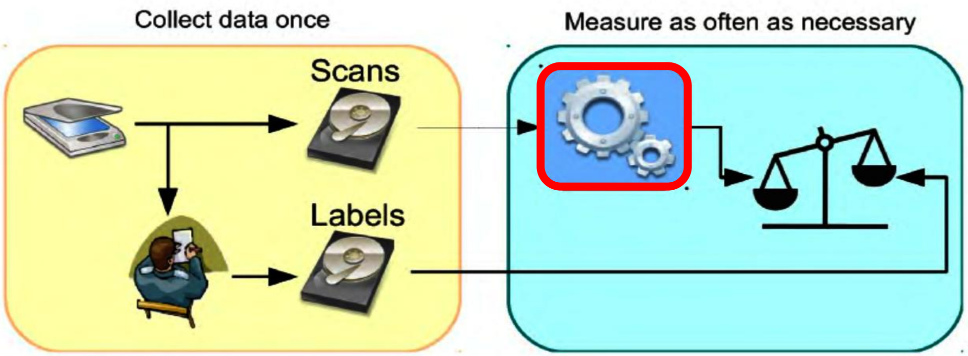
Learning procedure

- Collect data that describes the operation of the system during a past time period.
- Find changes that would have increased the performance of the system if they had been applied during the data collection period.
- Implement and verify…

Replaying past data

Classification example

- Collect labeled data in existing setup
- Replay the past data to evaluate what the performance would have been if we had used classifier θ.

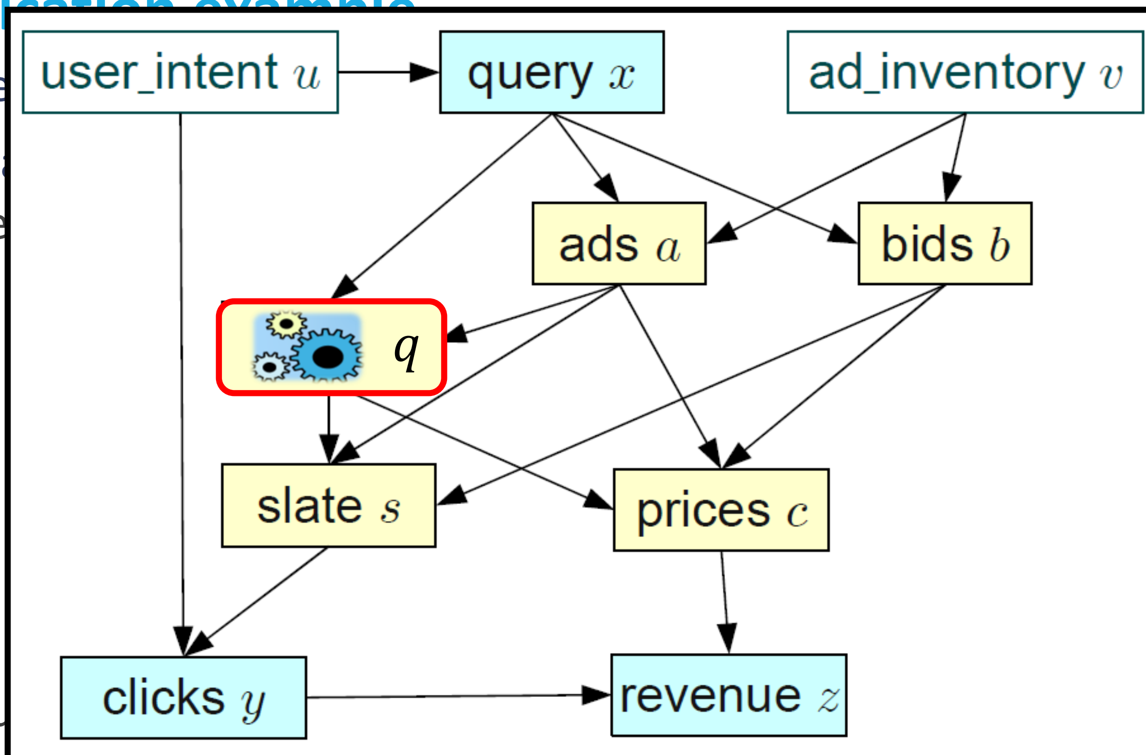


- Requires knowledge of all functions connecting the point of intervention to the point of measurement.

Replaying past data

Classification example

- Colle
- Repl… d have been
  if we



- Requ… ntervention
  to the point of measurement.

Importance sampling

Distribution under intervention

$$
\begin{aligned}
\mathbf{P}^{*}(\omega) = \; & \mathbf{P}(u, v) \\
& \times \mathbf{P}(x \mid u) \\
& \times \mathbf{P}(a \mid x, v) \\
& \times \mathbf{P}(b \mid x, v) \\
& \times \cancel{\mathbf{P}(q \mid x, a)} \qquad \mathbf{P}^{*}(q \mid x, a) \\
& \times \mathbf{P}(s \mid a, q, b) \\
& \times \mathbf{P}(c \mid a, q, b) \\
& \times \mathbf{P}(y \mid s, u) \\
& \times \mathbf{P}(z \mid y, c)
\end{aligned}
$$

Importance sampling

Actual expectation

$$Y = \int_{\omega} \ell(\omega)\, P(\omega)$$

Counterfactual expectation

$$Y^* = \int_{\omega} \ell(\omega)\, P^*(\omega) = \int_{\omega} \ell(\omega) \frac{P^*(\omega)}{P(\omega)} P(\omega) \approx \frac{1}{n} \sum_{i=1}^{n} \frac{P^*(\omega_i)}{P(\omega_i)}\, \ell(\omega_i)$$

Importance sampling

Principle

Reweight past examples to emulate the probability they would have had under the counterfactual distribution.

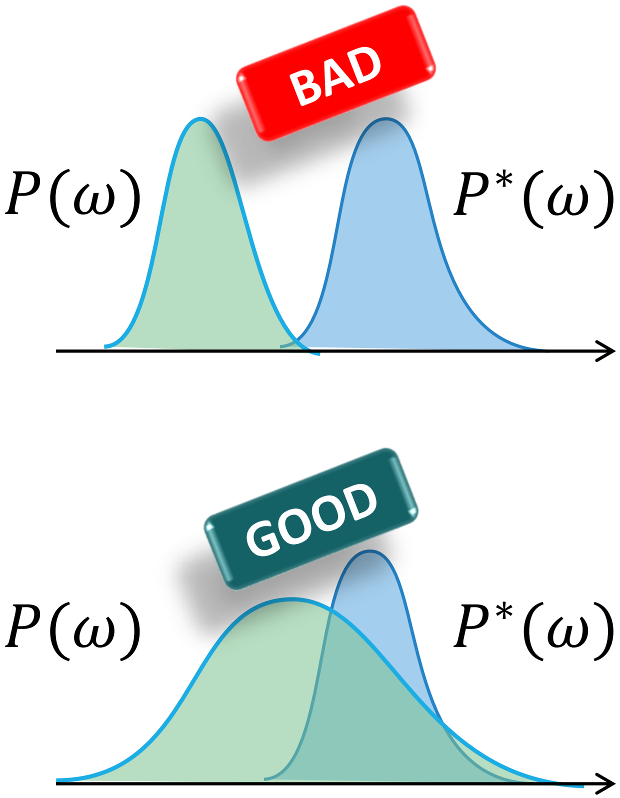
$$w(\omega_i) = \frac{P^*(\omega_i)}{P(\omega_i)} = \frac{P^*(q|x,a)}{P(q|x,a)}$$

Factors in P* not in P

Factors in P not in P*

Only requires the knowledge of the function under intervention (before and after)

Exploration



Quality of the estimation

- Large ratios undermine estimation quality.
- Confidence intervals reveal whether the data collection distribution $P(\omega)$ performs sufficient exploration to answer the counterfactual question of interest.

Confidence intervals

$$Y^* = \int_\omega \ell(\omega)\, w(\omega) P(\omega) \approx \frac{1}{n} \sum_{i=1}^{n} \ell(\omega_i)\, w(\omega_i)$$

Using the central limit theorem?

- $w(\omega_i)$ very large when $P(\omega_i)$ small.
- A few samples in poorly explored regions dominate the sum with their noisy contributions.
- Solution: ignore them.

Confidence intervals (ii)

Well explored area

$$\Omega_R = \{ \omega : P^*(\omega) < R \, P(\omega) \}$$

Easier estimate

$$\bar{Y}^* = \int_{\Omega_R} \ell(\omega) \, P^*(\omega) = \int_{\omega} \ell(\omega) \, \bar{w}(\omega) P(\omega) \approx \frac{1}{n} \sum_{i=1}^{n} \ell(\omega_i) \bar{w}(\omega_i)$$

$$\text{with } \bar{w}(\omega) = \begin{cases} w(\omega) & \text{if } \omega \in \Omega_R \\ 0 & \text{otherwise} \end{cases}$$

This works because $0 \leq \bar{w}(\omega) \leq R$.

Confidence intervals (iii)

Bounding the bias

Assuming $0 \leq \ell(\omega) \leq M$ we have

$$0 \leq Y^* - \bar{Y}^* \leq \int_{\Omega \setminus \Omega_R} \ell(\omega) P^*(\omega) \leq M \; P^*\{\Omega \setminus \Omega_R\} = M[1 - P^*(\Omega_R)]$$

$$= M\left[1 - \int_\omega \bar{w}(\omega) P(\omega)\right] \approx M\left[1 - \frac{1}{n}\sum_{i=1}^{n} \bar{w}(\omega_i)\right]$$

- This is easy to estimate because $\bar{w}(\omega)$ is bounded.
- This represents the cost of insufficient exploration.
- Bonus: this remains true if $P(\omega)$ is zero in some places

Two-parts confidence interval

$$\mathbf{Y}^* - \overline{\mathbf{Y}}_\mathbf{n}^* = (\boldsymbol{Y}^* - \overline{\boldsymbol{Y}}^*) + (\overline{\mathbf{Y}}^* - \overline{\mathbf{Y}}_{\boldsymbol{n}}^*)$$

Outer confidence interval

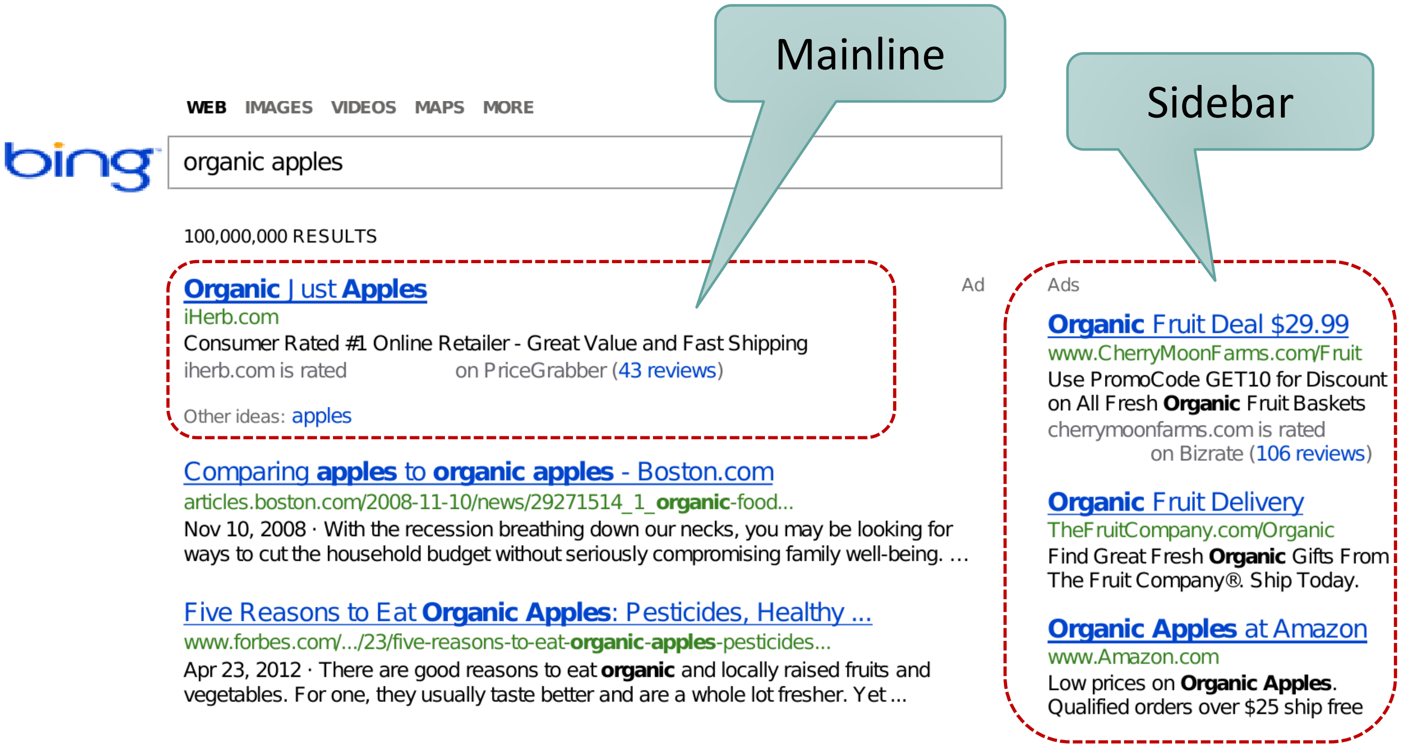
- Bounds $\overline{\mathrm{Y}}^* - \overline{\mathrm{Y}}_n^*$
- When this is too large, we must sample more.

Inner confidence interval

- Bounds $Y^* - \overline{Y}^*$
- When this is too large, we must explore more.

3. Illustration

Mainline ads



Playing with mainline reserves

Mainline reserves (MLRs)

- Rank score thresholds that control whether ads are displayed above the search results.

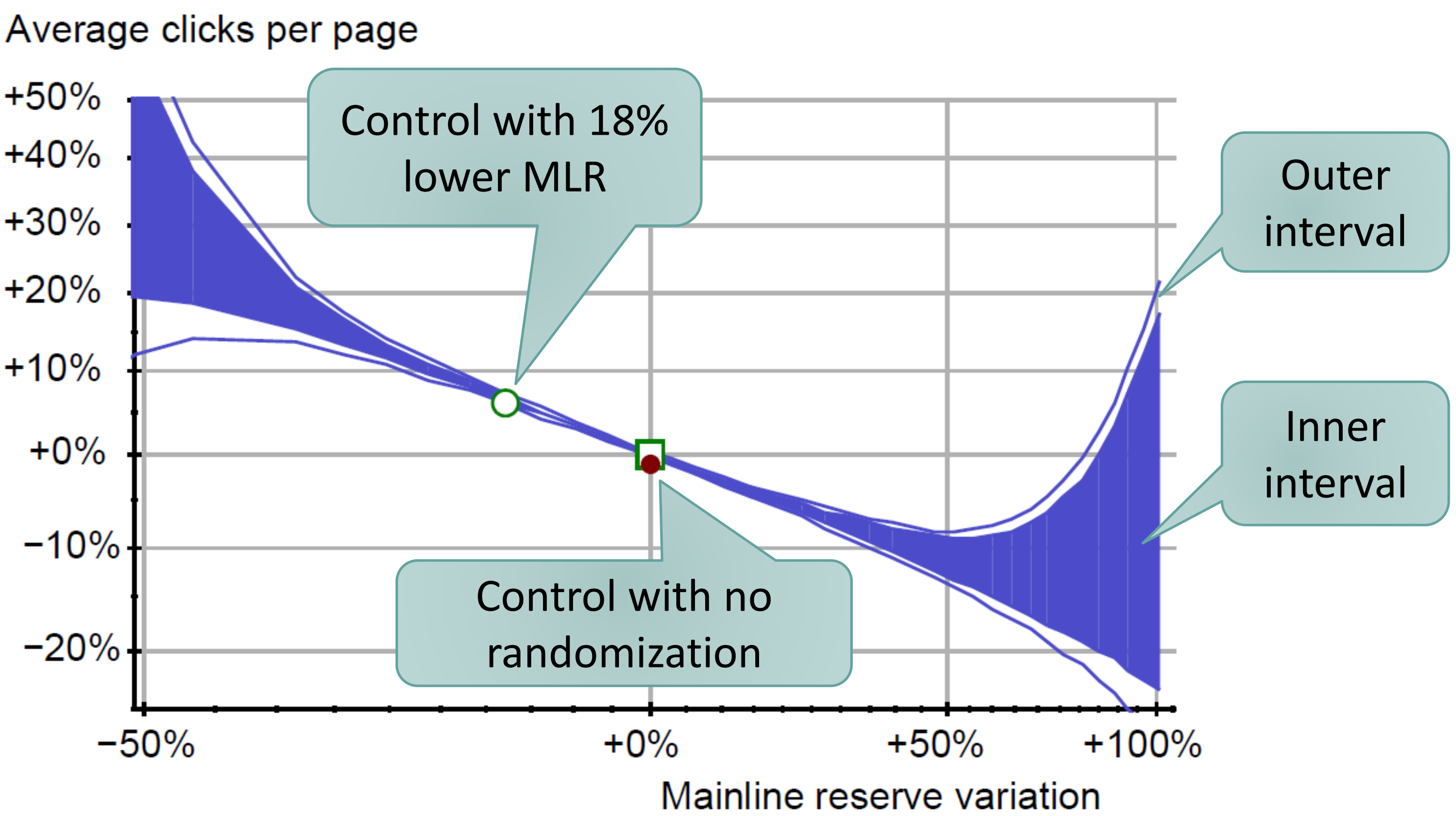
Data collection bucket

- Random log-normal multiplier applied to MLRs.
- 22M auctions over five weeks (summer 2010)

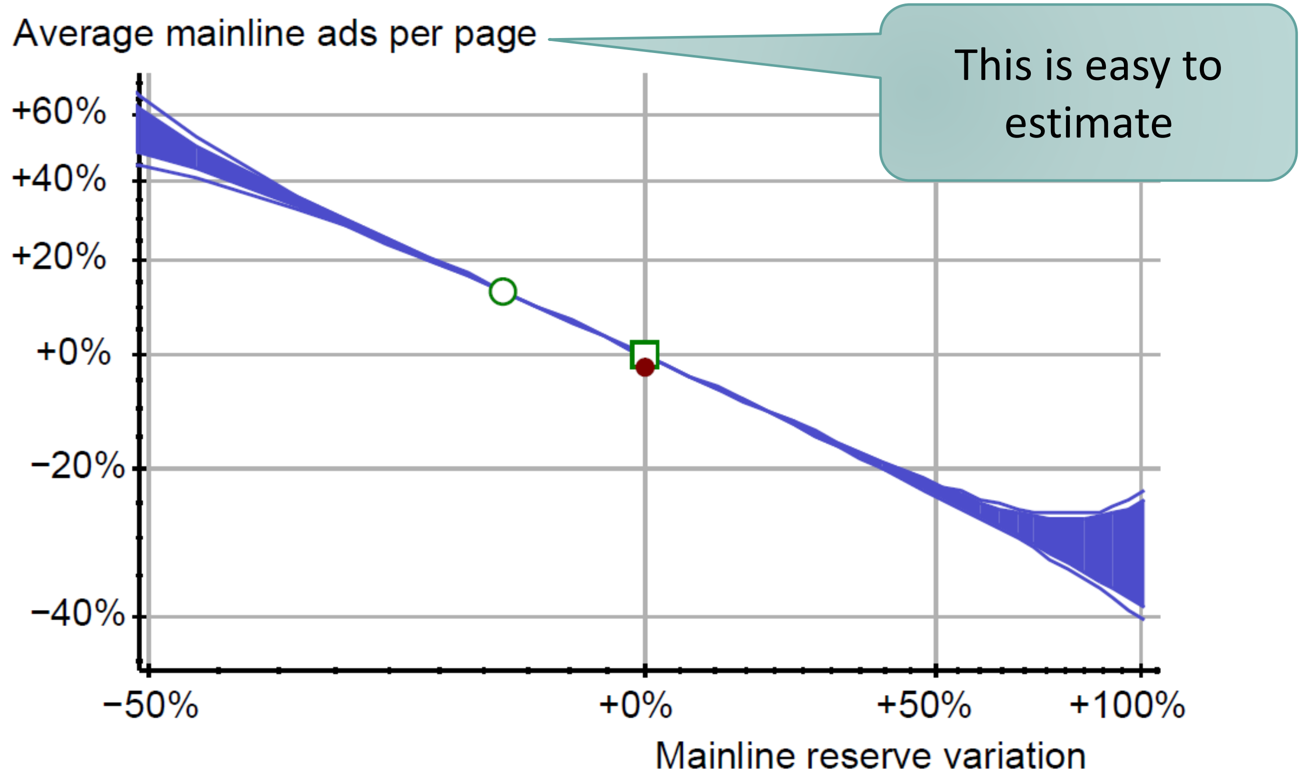
Control buckets

- Same setup with 18% lower mainline reserves
- Same setup without randomization

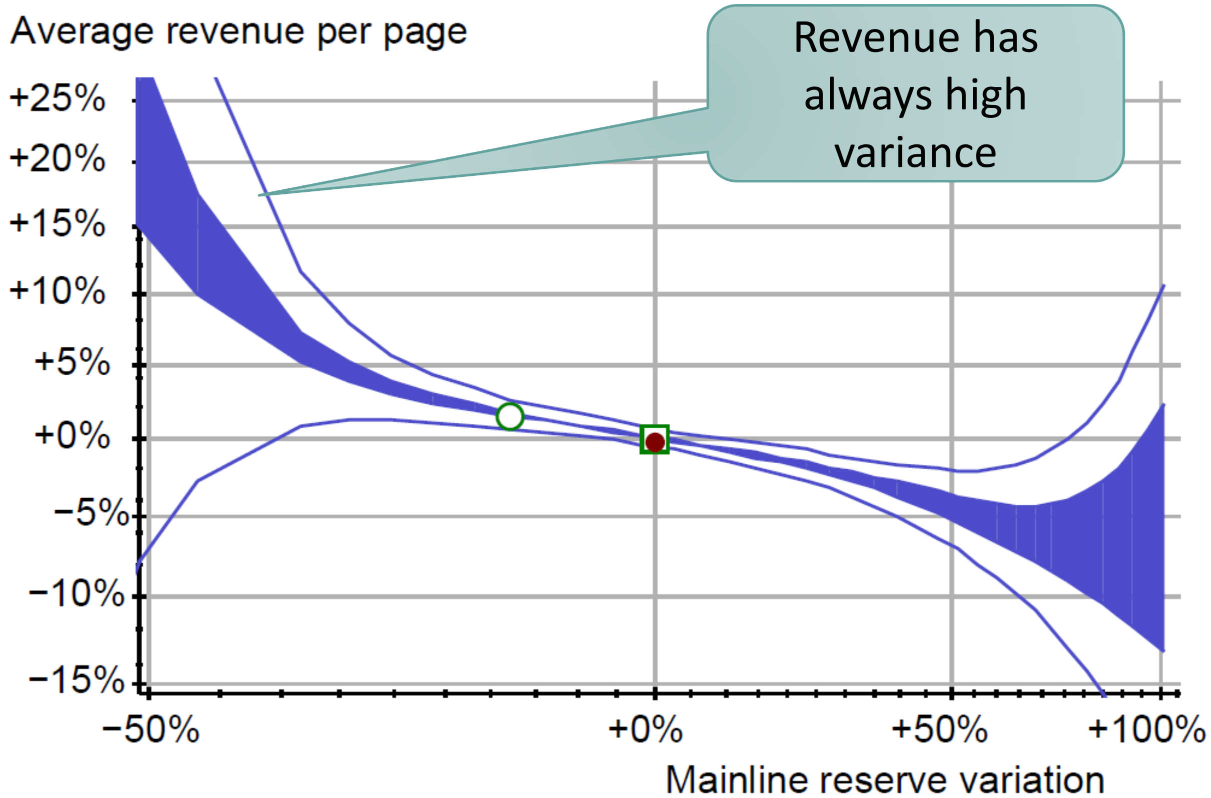
Playing with mainline reserves



Playing with mainline reserves



Playing with mainline reserves



More uses for the same data

Examples

Estimates for different randomization variance

→ Good to determine how much to explore.

Query-dependent reserves

→ Just another counterfactual distribution!

This is the big advantage

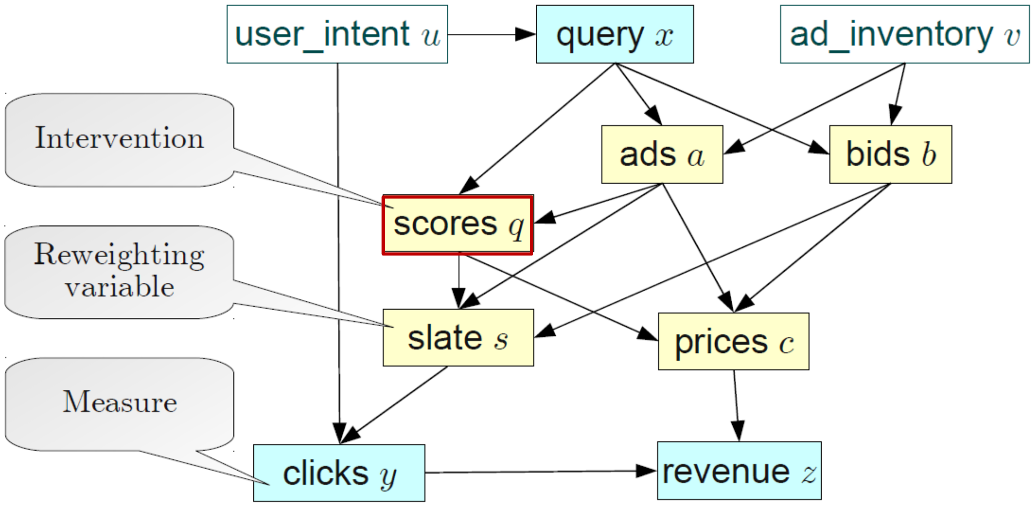
- Collect data first, choose questions later.
- Randomizing more stuff increases opportunities.
- New challenge: making sure that do not leave information on the table.

4. Toolbox

Using the causal structure

Improved confidence intervals

- Example: users click without knowing the ad scores or the click prices.
- Technique: “shifting” the reweighting variables



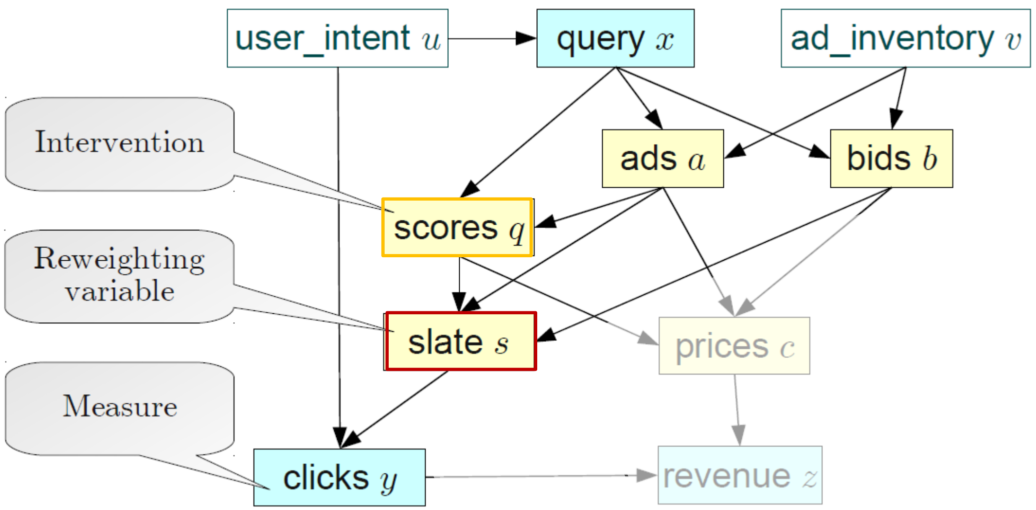
Standard weights

$$w(\omega_i) = \frac{P^*(\omega_i)}{P(\omega_i)} = \frac{P^*(q|x,a)}{P(q|x,a)}$$

Using the causal structure

Improved confidence intervals

- Example: users click without knowing the ad scores or the click prices.
- Technique: “shifting” the reweighting variables



Standard weights

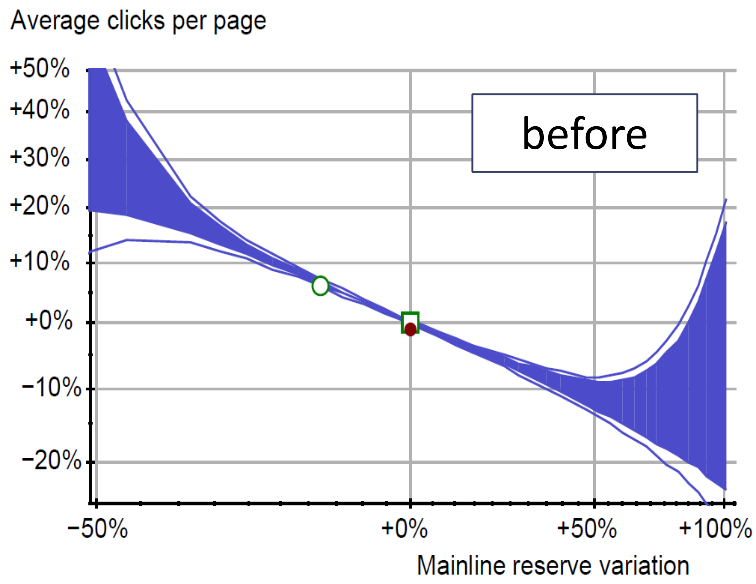
$$w(\omega_i) = \frac{P^*(\omega_i)}{P(\omega_i)} = \frac{P^*(q|x,a)}{P(q|x,a)}$$

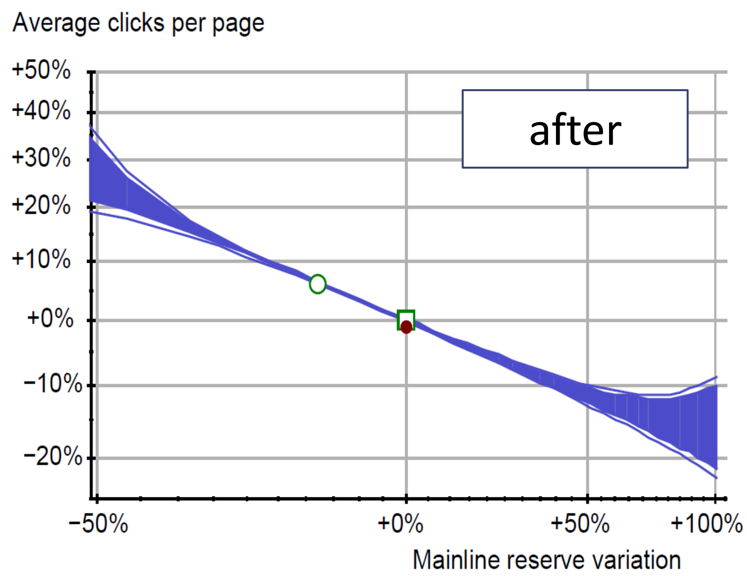
Shifted weights

$$w(\omega_i) = \frac{P^*(\omega_i)}{P(\omega_i)} = \frac{P^*(s|x,a,b)}{P(s|x,a,b)}$$

Using the causal structure

Experimental validation





Estimating differences

Comparing two potential interventions

Is scoring model M_1 better than M_2 ?

$\Delta =$ Click-thru-rate if we had used model M_1 − Click-thru-rate if we had used model M_2

Improved confidence via variance reduction

- Example: since seasonal variations affect both models nearly identically, the variance resulting from these variations cancels in the difference.

Estimating derivatives

Infinitesimal interventions

$$\frac{\partial CTR}{\partial \theta} = \left(\text{Click rate if we had used } M(\theta + d\theta) - \text{Click rate if we had used model } M(\theta) \right) \Big/ \mathrm{d}\theta$$

- Related to "policy gradient" in RL.
- Optimization algorithms learn model parameters θ.

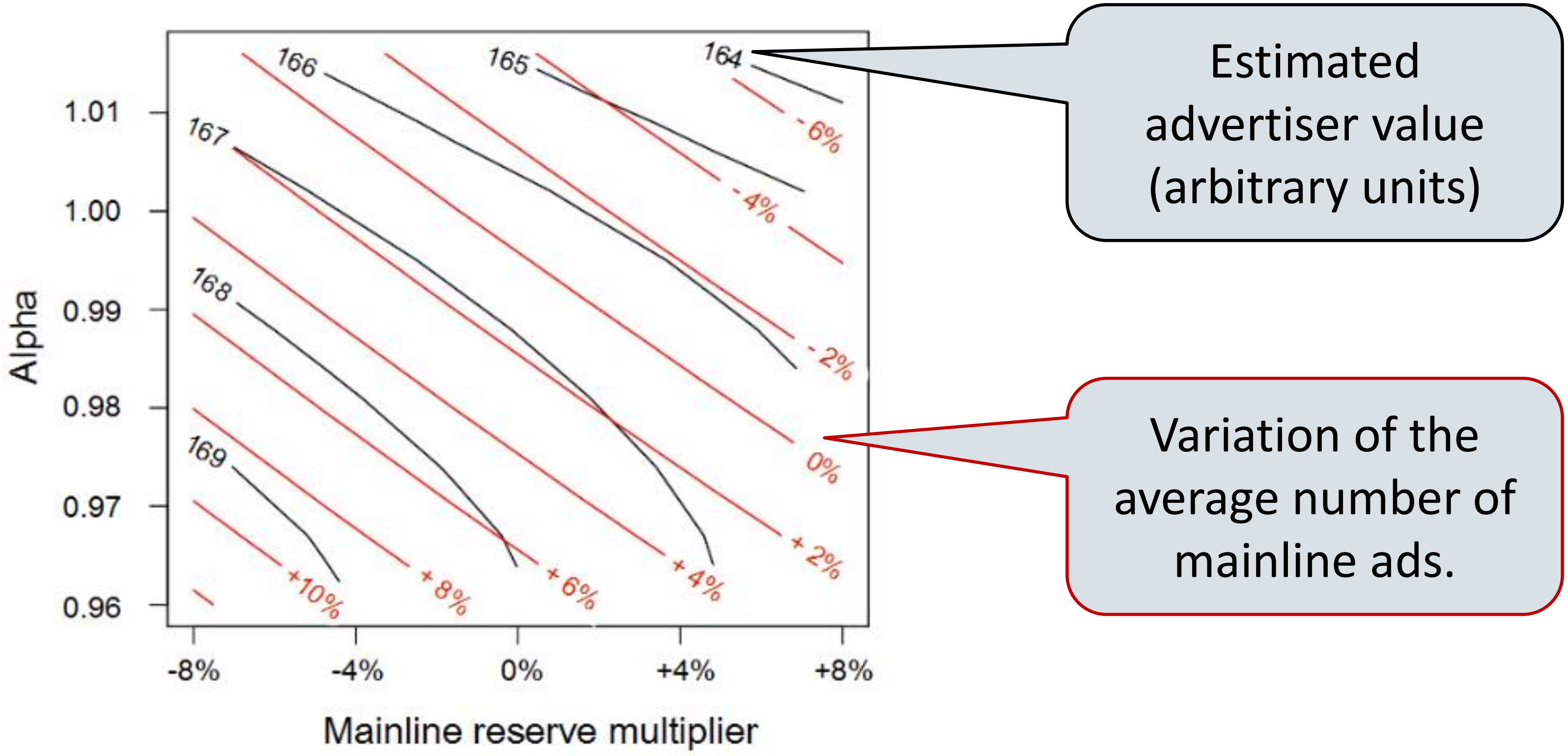
Derivatives and optimization

Example
Tuning squashing exponents and reserves

- Ads ranked by decreasing $bid \times pClick^{\alpha}$
- Lahaie and McAfee (2011) show that $\alpha < 1$ is good when click probability estimation gets less accurate.
- Different α_k and reserves ρ_k for each query cluster k.

Derivatives and optimization

Level curves for one particular query cluster



Learning as counterfactual optimization

- Does it generalize?

  Yes, we can obtain uniform confidence intervals.

- Sequential design?

  Thompson sampling comes naturally in this context.

- Metering exploration wisely?

  Inner confidence interval tells how much exploration we need to answer a counterfactual question.

  But it does not tell which questions we should ask.

  This was not a problem in practice…

5. Equilibrium

Revisiting the feedback loops

Tracking the equilibrium

If we increase the ad relevance thresholds :

- We show less ads and lose revenue in the short time.
- Users see more relevant ads, are more likely to click on ads in the future, possibly making up for the lost revenue [eventually].
- Advertisers will [eventually] update their bids.
  It could go both ways because they receive less clicks
  from more engaged users…

Counterfactual equilibrium

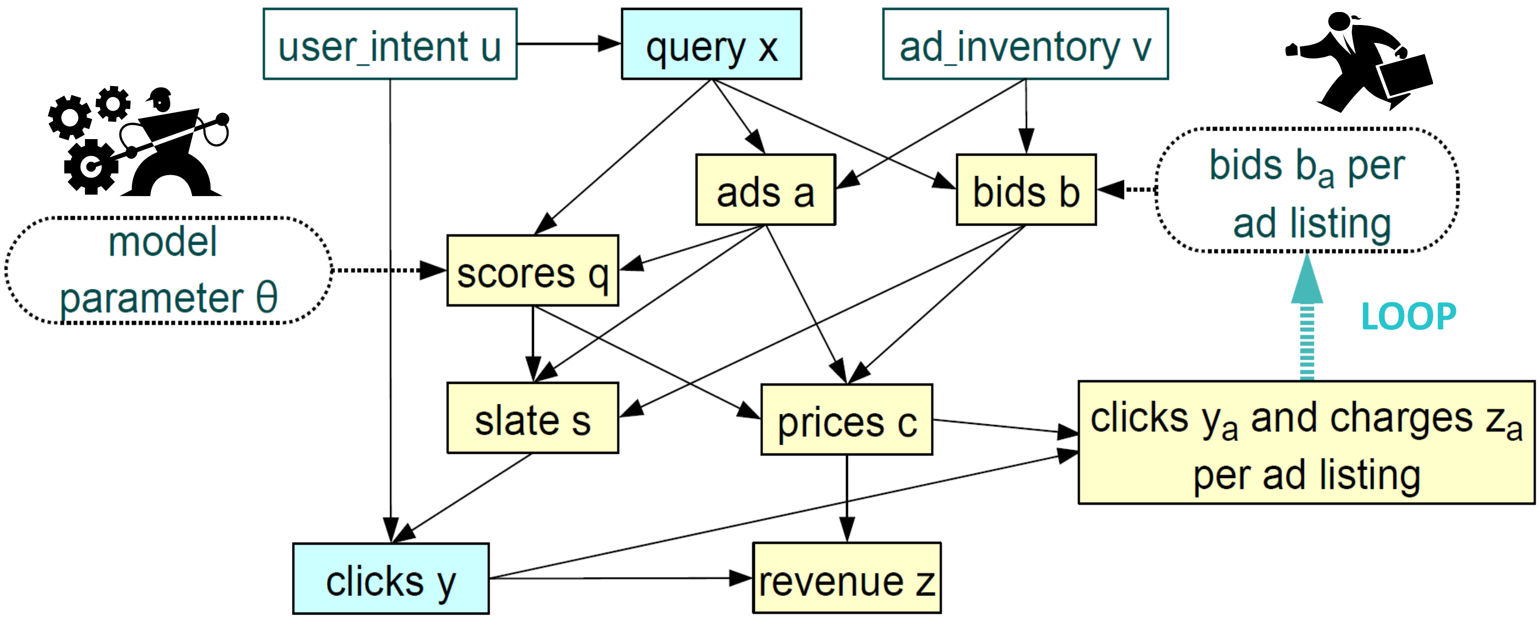
Counterfactual question

> *"What would have been the system performance metrics if we had applied an infinitesimal change to the parameter θ of the scoring model long enough to reach the equilibrium during the data collection period?"*

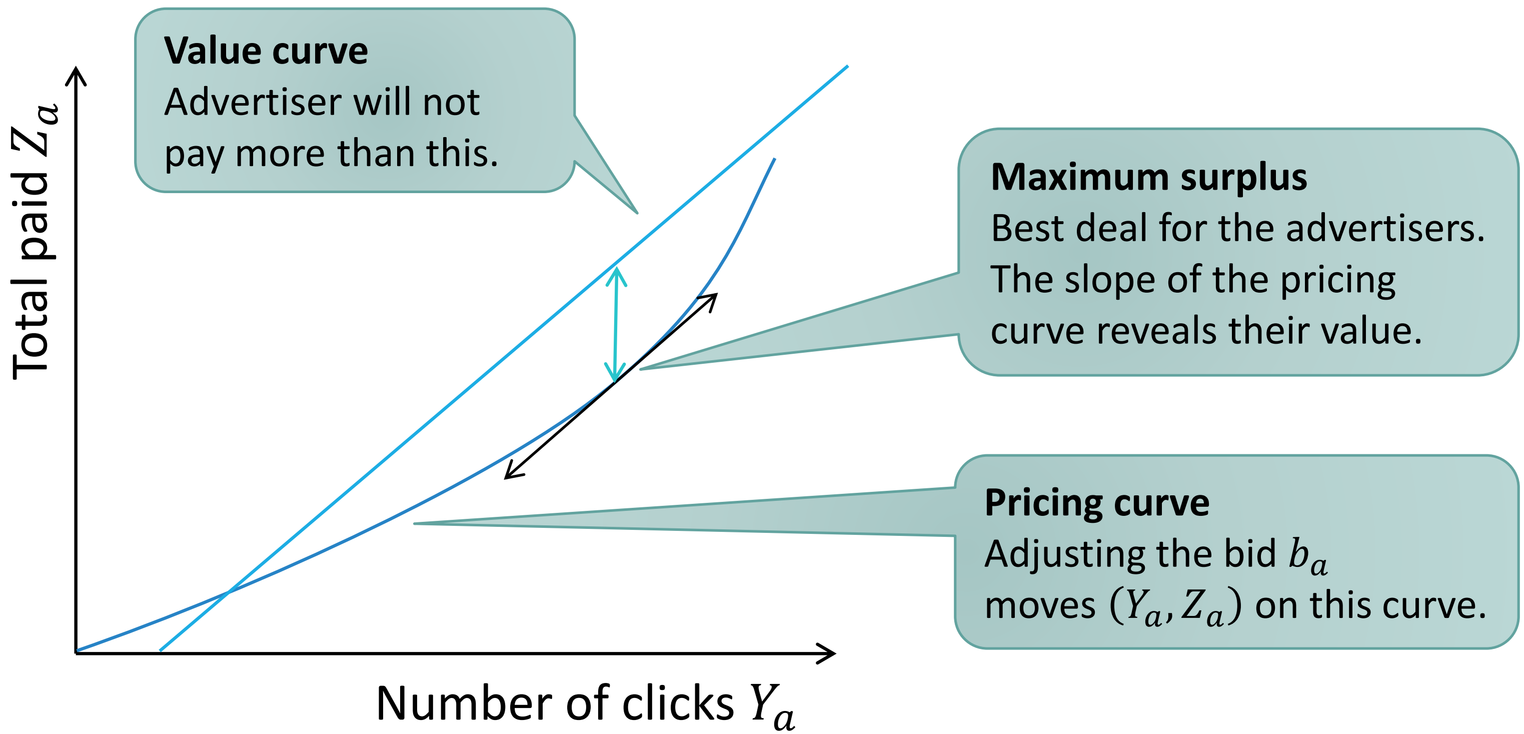
We can answer using *quasi-static analysis*.

(this comes from physics.)

Advertiser feedback loop



Rational advertiser



Rational advertisers keep $V_a = \frac{\partial Z_a}{\partial Y_a} = \frac{\partial Z_a}{\partial b_a} / \frac{\partial Y_a}{\partial b_a}$ constant!

Estimating values

When the system reaches equilibrium, we can compute

$$V_a = \frac{\partial Z_a}{\partial b_a} \Big/ \frac{\partial Y_a}{\partial b_a} = \frac{\partial E_{\boldsymbol{b},\theta}(z_a)}{\partial b_a} \Big/ \frac{\partial E_{\boldsymbol{b},\theta}(y_a)}{\partial b_a}$$

Counterfactual derivatives

- Complication: we cannot randomize the bids. However, since ads are ranked by *bids×scores*, we can interpret a random score multiplier as a random bid multiplier (need to reprice.)

Feedback loop equilibrium

Derivative of surplus vector $\Phi = \left[\ldots \frac{\partial Z_a}{\partial b_a} - V_a \frac{\partial Y_a}{\partial b_a} \ldots \right] = 0.$

$$\mathrm{d}\Phi = \frac{\partial \Phi}{\partial \theta} \mathrm{d}\theta + \sum_a \frac{\partial \Phi}{\partial b_a} \mathrm{d}b_a = 0$$

Equilibrium condition

Solving the linear system yields $\frac{\mathrm{d}b_a}{\mathrm{d}\theta}$.

Then we answer the counterfactual question

$$\mathrm{d}Y = \left(\frac{\partial Y}{\partial \theta} + \sum_a \frac{\partial Y}{\partial b_a} \frac{\mathrm{d}b_a}{\mathrm{d}\theta} \right) \mathrm{d}\theta$$

Multiple feedback loops

Same procedure

1. Write total derivatives.
2. Solve the linear system formed
   by all the equilibrium conditions.
3. Substitute into the total derivative
   of the counterfactual expectation of interest.

Conclusion

Main messages

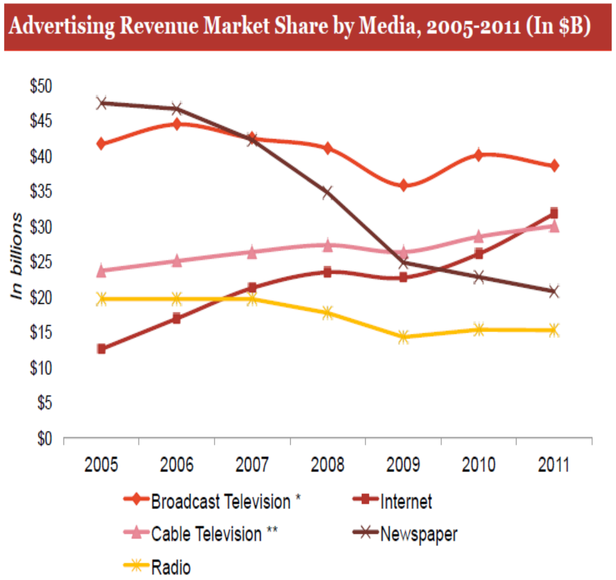
- Relation between explore-exploit and correlation-causation.
- The causation framework provides a rich and modular toolbox.
- The differential equilibrium analysis methods of physics apply.

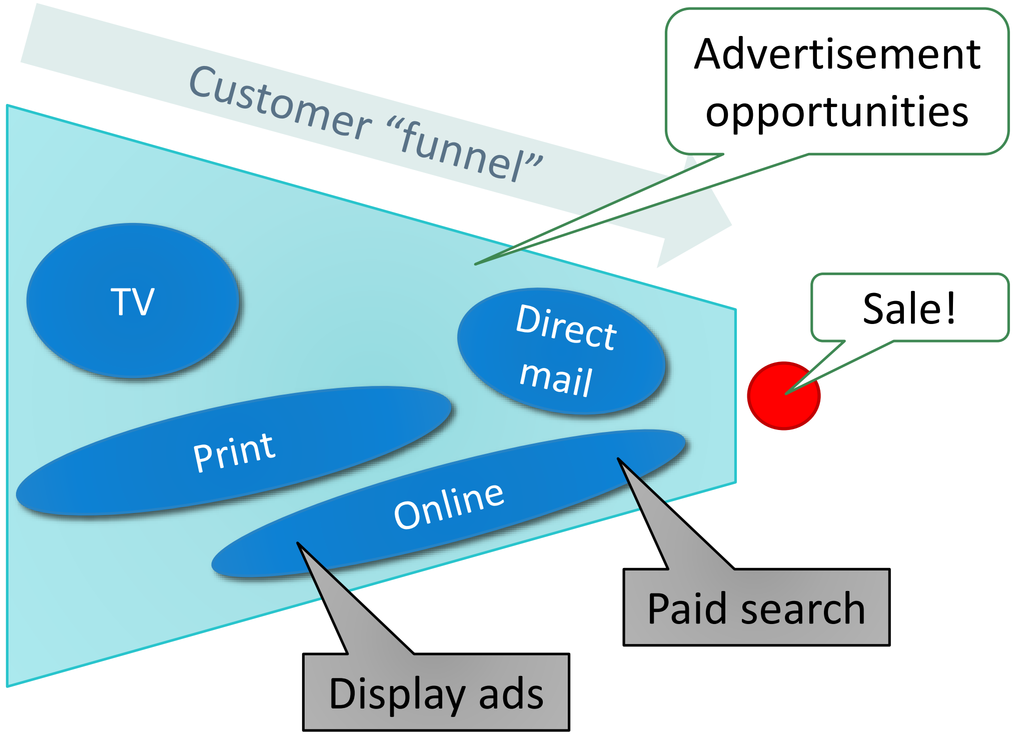
Tech report available at http://leon.bottou.org/papers.

MORE SLIDES

ADS

Advertisement primer





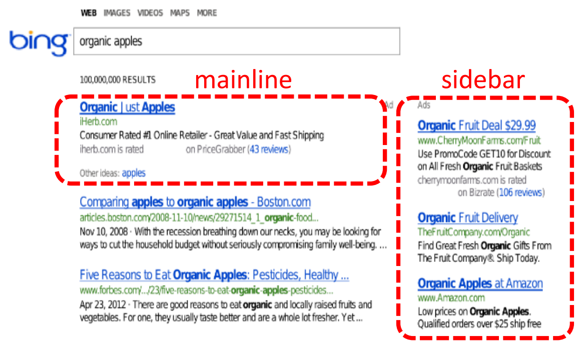
Self interest

User	• <u>Expects</u> results that satisfy her interests • <u>Possibly</u> by initiating business with an advertiser • Future engagement depends on her satisfaction….
Advertiser	• <u>Expects</u> to receive potential customers • Expects to <u>recover</u> clicks costs from resulting business • Return on investment impacts future ads and bids…
Publisher	• <u>Expects</u> click money • <u>Learns</u> which ads work from past data. • In order to <u>preserve</u> future gains, publishers must ensure the continued satisfaction of users and advertisers. (this changes everything!)

Performance metrics

First order performance metrics

- Average number of ads shown per page
- Average number of **mainline ads** per page
- Average number of ad clicks per page
- Average revenue per page (RPM)



Should we just optimize RPM?

Showing lots of mainline ads improves RPM.
Users would quickly go away!

Increasing the reserve prices also improves RPM.
Advertisers would quickly go away!

Performance metrics

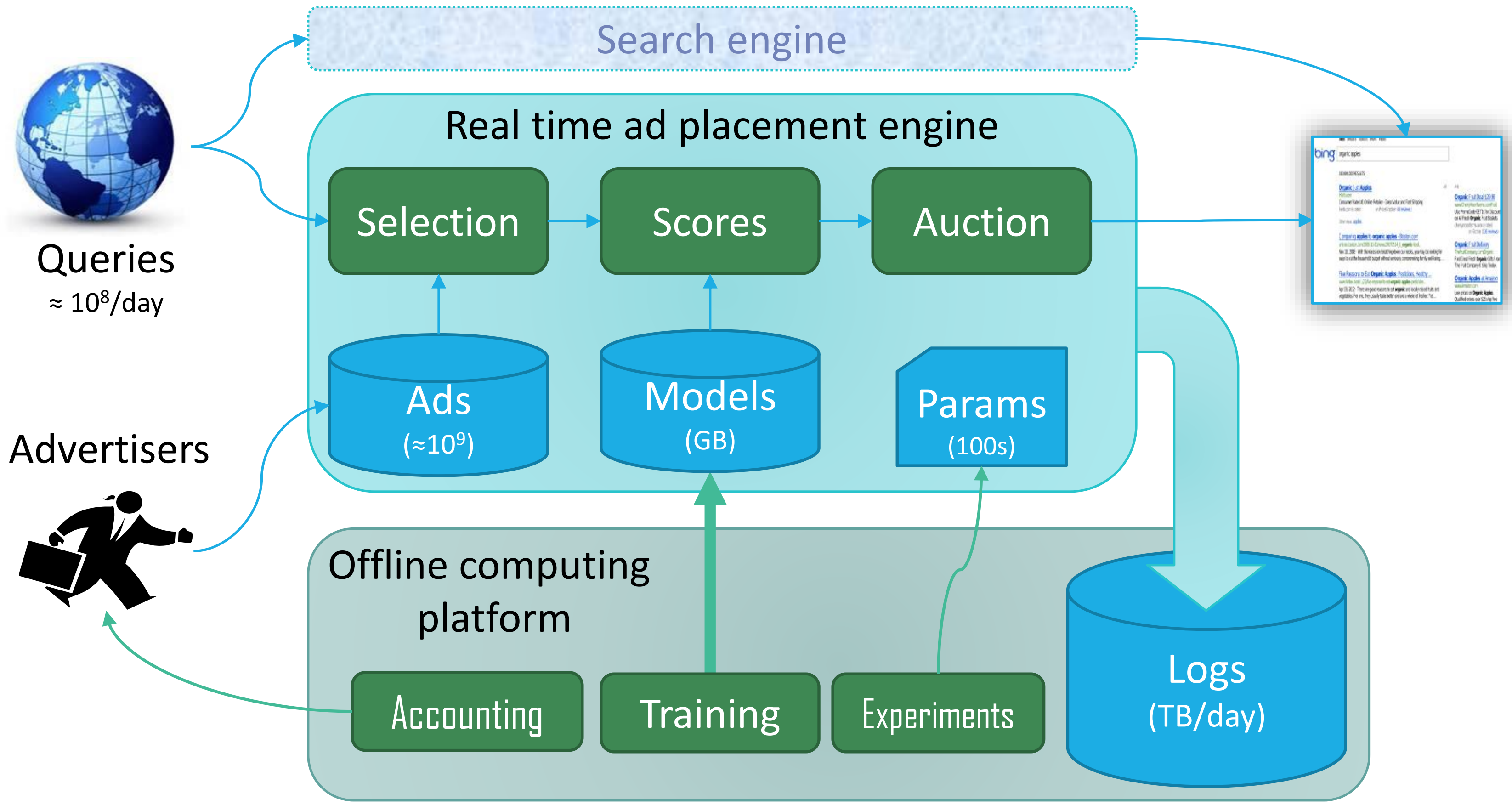
First order performance metrics

- Average number of ads shown per page
- Average number of mainline ads per page
- Average number of ad clicks per page
- Average revenue per page (RPM)
- Average relevance score estimated by human labelers
- Average number of bid-weighted ad clicks per page
- …

Monitor heuristic indicators of user fatigue

Monitor heuristic indicators of advertiser value

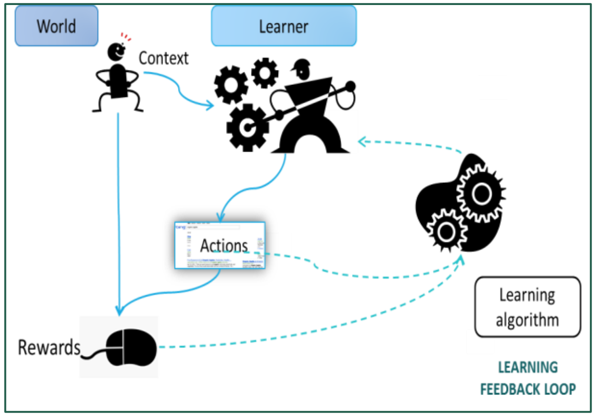
The plumbing



MORE SLIDES

CONTEXTUAL BANDITS

Contextual bandits (CBs)



Framework

- World select context x
- Learner decides action $a = \pi(x)$
- World announces reward $r(x, a)$

Results

- Randomized data collection (i.e., exploration) enables offline unbiased evaluation of an alternate policy π^*.
- Solid analysis of the explore/exploit trade-off, that is, how much exploration is needed at each instant.

Structure

Actions have structure

- What we learn by showing a particular ad for a particular query tells us about showing similar ads for similar queries.

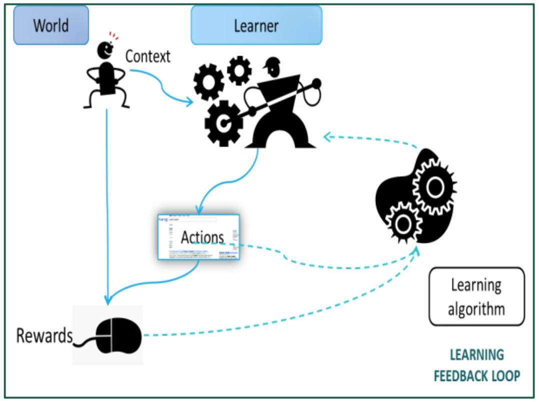
Policies have structure

- One action is a set of ads displayed on a page.
  But computationally feasible policies score each ad individually.

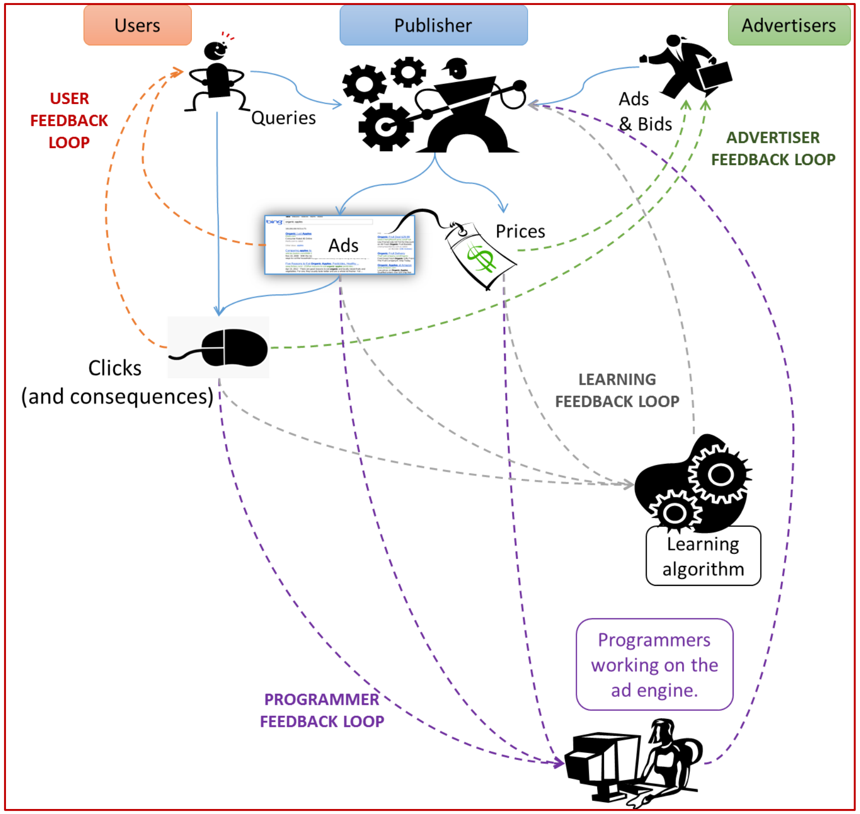
Rewards have structure

- Actions are set of ads with associated click prices.
  Chosen ads impact users, chosen prices impact advertisers.

The causal graph has structure



≠



MORE SLIDES

SIMPSON IN ADS

Causal paradoxes in ad data

A legitimate question

"Does it help to know the estimated click probability of the first mainline ad in order to estimate the click probability of the second mainline ad?"

Naïve approach

- Collect past data for pages showing at least two ads.
- Split them in two groups according to the estimated click probability q_1 computed for the first ad.
- Count clicks on the second ad and compare.

Causal paradoxes in ad data

	CTR_2
q_1 low	124/2000 (6.2%)
q_1 high	149/2000 (7.5%)

Confounding factors…

- Commercial queries get higher q_1.
  They also receive more clicks everywhere…
- Let us split the data according to the estimated click probability q_2 computed for the second ad.

Causal paradoxes in ad data

	CTR_2	if q_2 low	if q_2 high
q_1 low	124/2000 (6.2%)	92/1823 (5.1%)	32/176 (18.1%)
q_1 high	149/2000 (7.5%)	71/1500 (4.8%)	78/500 (15.6%)

Simpson reversal!

- q_1 and CTR_2 have a (confounding) common cause.
- Controlling for a common cause can reverse the conclusion.
- What about the common causes we do not know?
- How to reason about such problems?

MORE SLIDES

LOOPS

Toy example

Two queries

Q1: “cheap diamonds” (50% traffic)
Q2: “google” (50% traffic)

Three ads

A1: “cheap jewelry”
A2: “cheap automobiles”
A3: “engagement rings”

More simplifications

- We show only one ad per query
- All bids are equal to $1.

Toy example

True conditional click probabilities

	A1 (cheap jewelry)	**A2** (cheap autos)	**A3** (engagement rings)
Q1 (cheap diamonds)	7%	2%	9%
Q2 (google)	2%	2%	2%

Step 1: pick ads randomly.

$$CTR = \frac{1}{2}\left(\frac{7+2+9}{3} + \frac{2+2+2}{3}\right) = 4\%$$

Toy example

Step 2: estimate click probabilities

- Build a model based on a single Boolean feature:
  F : “*query and ad have at least one word in common*”

	A1 (cheap jewelry)	**A2** (cheap autos)	**A3** (engagement rings)
Q1 (cheap diamonds)	7%	2%	9%
Q2 (google)	2%	2%	2%

$$P(Click|F) = \frac{7+2}{2} = 4.5\%$$

$$P(Click|\neg F) = \frac{9+2+2+2}{4} = 3.75\%$$

Toy example

Step 3: place ads according to estimated pclick.

Q1: show A1 or A2. (predicted pclick 4.5% > 3.75%)

Q2: show A1, A2, or A3. (predicted pclick 3.75%)

	A1 (cheap jewelry)	A2 (cheap autos)	A3 (engagement rings)
Q1 (cheap diamonds)	7%	2%	~~9%~~
Q2 (google)	2%	2%	2%

$$CTR = \frac{1}{2}\left(\frac{7+2}{2} + \frac{2+2+2}{3}\right) = 3.25\%$$

Toy example

Step 4: re-estimate click probabilities with new data.

	A1 (cheap jewelry)	**A2** (cheap autos)	**A3** (engagement rings)
Q1 (cheap diamonds)	7%	2%	~~9%~~
Q2 (google)	2%	2%	2%

$$P(Click|F) = \frac{7+2}{2} = 4.5\%$$

$$P(Click|\neg F) = \frac{2+2+2}{3} = 2\%$$

- We keep selecting the same inferior ads. 
- Estimated click probabilities now seem more accurate. 

What is going wrong?

- Estimating Pclick using click data collected by showing random ads.

	A1	A2	A3
Q1	7%	2%	9%
Q2	2%	2%	2%

1. Feature F identifies relevant ads using a narrow criterion.
2. Feature F misses a very good ad for query Q1.
3. P(Click|¬F) is pulled down by queries that do not click.
4. Ads for query Q1 are ranked incorrectly.

Adding a feature that singles out the case (Q1,A3)

- would improve the pclick estimation metric.
- would rank Q1 ads more adequately.

What is going wrong?

- Re-estimating Pclick using click data collected by showing ads suggested by the previous Pclick model.

	A1	A2	A3
Q1	7%	2%	~~9%~~
Q2	2%	2%	2%

In this data, A3 is never shown for query Q1.

$P(Click|\neg F)$ seems more accurate because we have removed the case (Q1,A3)

Adding a (Q1,A3) feature

- would not improve the Pclick estimation on this data.
- would not help ranking (Q1,A3) higher.

Further feature engineering based on this data

- would always result in eliminating more options, e.g. (Q1,A2).
- would never result in recovering lost options, e.g. (Q1,A3).

We have created a black hole!

(Q,A) can be occasionally sucked by the black hole.

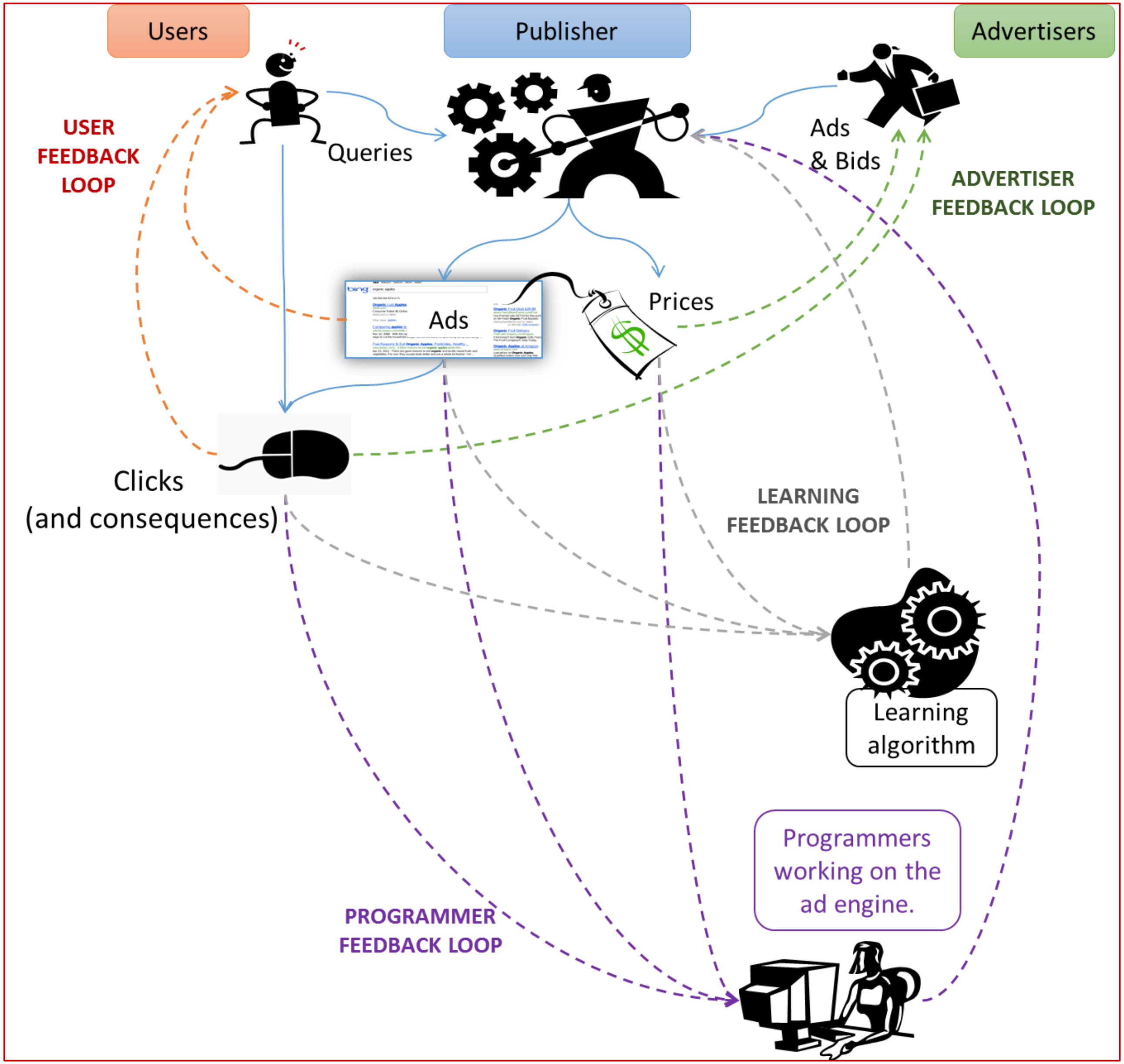
- All kinds of events can cause ads to disappear.
- Sometimes, advertisers spend extra money to displace competitors.

(Q,A) can be born in the black hole.

- Ads newly entered by advertisers
- Ads newly selected as eligible because of algorithmic improvements.

Exploration

- We should sometimes show ads that we would not normally show in order to train the click prediction model.

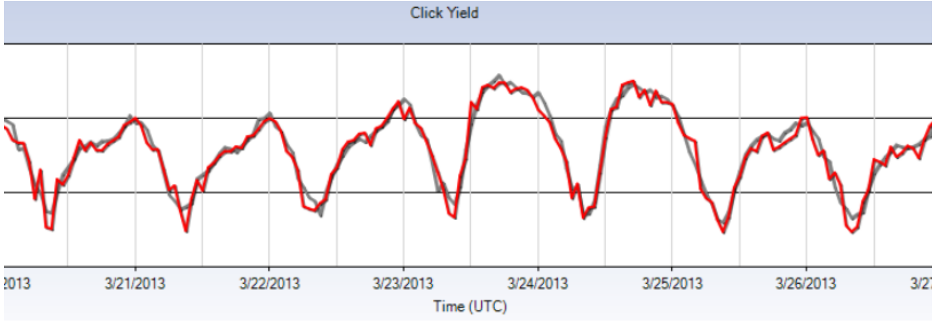

Users
Publisher
Advertisers
USER FEEDBACK LOOP
Queries
Ads & Bids
ADVERTISER FEEDBACK LOOP
Ads
Prices
Clicks (and consequences)
LEARNING FEEDBACK LOOP
Learning algorithm
Programmers working on the ad engine.
PROGRAMMER FEEDBACK LOOP

MORE SLIDES

VARIANCE REDUCTION

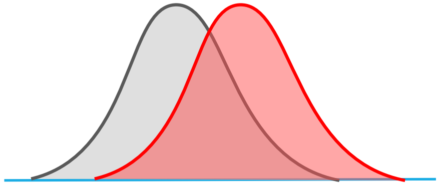
Variance reduction

Hourly average click yield for treatment and control



$$\left(Y - \frac{1}{n}\sum y_i\right) \sim \mathcal{N}\left(0, \frac{\sigma}{\sqrt{n}}\right)$$

Daily effects increases the variance of both treatment and control.



Daily effects affect treatment and control in similar ways!
Can we subtract them?

Variance reduction

- Treatment estimate $\quad Y^* \approx \hat{Y}^* = \frac{1}{|T|} \sum_{i \in T} \; y_i$
- Control estimate $\quad Y \approx \hat{Y} = \frac{1}{|C|} \sum_{i \in C} \; y_i$

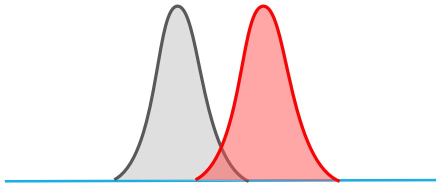
- Predictor $\zeta(X)$ tries to estimate Y on the basis of solely the context X.

- Then $Y^* - Y = \big(Y^* - \zeta(X)\big) - \big(Y - \zeta(X)\big)$

$$\approx \frac{1}{|T|} \sum_{i \in T} \big(y_i - \zeta(x_i) \big) - \frac{1}{|C|} \sum_{i \in C} \big(y_i - \zeta(x_i) \big)$$

This is true regardless of the predictor quality.

But if it is any good, $\mathrm{var}[Y - \zeta(X)] < \mathrm{var}[Y]$, and



Counterfactual differences

Which scoring model works best?

- Comparing expectations under counterfactual distributions $P^+(\omega)$ and $P^*(\omega)$.

$$Y^+ - Y^* = \int_\omega [\ell(\omega) - \zeta(\nu)]\, \Delta w(\omega)\; P(\omega)$$

$$\approx \frac{1}{n} \sum_{i=1}^{n} [\ell(\omega_i) - \zeta(\nu_i)]\, \Delta w(\omega_i)$$

with $\Delta w(\omega) = \frac{P^+(\omega)}{P(\omega)} - \frac{P^*(\omega)}{P(\omega)}$

Variance captured by predictor $\zeta(\nu)$ is gone!

Counterfactual derivatives

Counterfactual distribution $P^{\theta}(\omega)$

$$\frac{\partial Y^{\theta}}{\partial \theta} = \int_{\omega} [\ell(\omega) - \zeta(\nu)] \, w'_{\theta}(\omega) P(\omega) \approx \frac{1}{n} \sum_{i=1}^{n} [\ell(\omega_i) - \zeta(\nu_i)] \, w'_{\theta}(\omega_i)$$

with $w'_{\theta}(\omega) = \frac{\partial w_{\theta}(\omega)}{\partial \theta} = w_{\theta}(\omega) \frac{\partial \log P^{\theta}(\omega)}{\partial \theta}$

$w_{\theta}(\omega)$ can be large but there are ways…