The social impact of algorithmic decision making: Economic perspectives

Maximilian Kasy

Fall 2020

In the news.

There's software used across the country to predict future criminals. And it's biased against blacks.

Paperclip-making robots 'wipe out humanity' in killer AI Doomsday experiment

Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry

Introduction

- Algorithmic decision making in consequential settings is spreading:
  Hiring, consumer credit, bail setting, news feed selection, pricing, ...
- $\Rightarrow$ Public concerns:
  - Are algorithms discriminating?
  - Can algorithmic decisions be explained?
  - Does AI create unemployment?
  - What about privacy?
- $\Rightarrow$ Taken up in computer science:
  "Fairness, Accountability, and Transparency,"
  "Value Alignment," etc.
- What are the normative foundations for these concerns?
  How shall we evaluate decision making systems empirically?
- Economists (among others) have debated related questions
  in non-automated settings for a long time!

Related debates in economics

- **Social Choice** theory.
  How to aggregate individual welfare rankings into a social welfare function?
- **Optimal taxation**.
  How to choose optimal policies
  subject to informational constraints and distributional considerations?
- The economics of **discrimination.**
  What are the mechanisms driving inter-group inequality,
  and how can we disentangle them?
- **Labor economics**, wage inequality, and distributional decompositions.
  What are the mechanisms driving rising wage inequality?

Related debates in economics, continued

- **Causal inference**.
  How can we make plausible predictions
  about the impact of counterfactual policies?
- **Contract theory**, mechanism design, and multi-tasking.
  What are the dangers of incentives based on quantitative performance measures?
- Experimental design and **surrogate** outcomes.
  How can we identify causal effects if the outcome of interest is unobserved?
- Market design, **matching** and optimal transport.
  How can two-side matching markets be organized without a price mechanism?

Some references

- **Social Choice** theory.
  Sen (1995),
  Roemer (1998)
- **Optimal taxation**.
  Mirrlees (1971),
  Saez (2001)
- The economics of **discrimination**.
  Becker (1957),
  Knowles et al. (2001)
- **Labor economics**.
  Fortin and Lemieux (1997),
  Autor and Dorn (2013)
- **Causal inference**.
  Imbens and Rubin (2015)
- **Contract theory**, multi-tasking.
  Holmstrom and Milgrom (1991)
- Experimental design and **surrogates**.
  Athey et al. (2019)
- **Matching** and optimal transport.
  Galichon (2018)

Projects in this agenda that I will discuss

- Kasy, M. and Abebe, R. (2020).
  Fairness, equality, and power in algorithmic decision making.
- Kasy, M. and Abebe, R. (2020).
  Multitasking, Surrogate Outcomes, and the Alignment Problem.
- Kasy, M. and Teytelboym, A. (2020).
  Adaptive combinatorial allocation.

This is all work in progress!

Introduction

Fairness, equality, and power in algorithmic decision making

- Fairness
- Inequality

Multi-tasking, surrogates, and the alignment problem

- Multi-tasking, surrogates
- Markov Decision Problems, Reinforcement learning

Adaptive combinatorial allocation

- Motivation: Refugee resettlement
- Performance guarantee

Conclusion

Introduction

- Public debate and the computer science literature:
  Fairness of algorithms, understood as the absence of **discrimination**.
- We argue: Leading definitions of fairness have three limitations:
  1. They legitimize inequalities justified by "merit."
  2. They are narrowly bracketed; only consider differences of treatment within the algorithm.
  3. They only consider between-group differences.
- Two alternative perspectives:
  1. What is the causal impact of the introduction of an algorithm on **inequality**?
  2. Who has the **power** to pick the objective function of an algorithm?

Fairness in algorithmic decision making – Setup

- Treatment W, treatment return M (heterogeneous), treatment cost c.
  Decision maker's objective
  $$\mu = E[W \cdot (M - c)].$$
- All expectations denote averages across individuals (not uncertainty).
- M is unobserved, but predictable based on features X.
  For $m(x) = E[M|X = x]$, the optimal policy is
  $$w^*(x) = \mathbf{1}(m(X) > c).$$

Examples

- Bail setting for defendants based on predicted recidivism.
- Screening of job candidates based on predicted performance.
- Consumer credit based on predicted repayment.
- Screening of tenants for housing based on predicted payment risk.
- Admission to schools based on standardized tests.

Definitions of fairness

- Most definitions depend on **three ingredients**.
  1. Treatment W (job, credit, incarceration, school admission).
  2. A notion of merit M (marginal product, credit default, recidivism, test performance).
  3. Protected categories A (ethnicity, gender).

- I will focus, for specificity, on the following **definition of fairness**:

$$\pi = E[M|W = 1, A = 1] - E[M|W = 1, A = 0] = 0$$

"Average merit, among the treated, does not vary across the groups a."

This is called "predictive parity" in machine learning,
the "hit rate test" for "taste based discrimination" in economics.

- "Fairness in machine learning" literature: **Constrained optimization**.

$$w^*(\cdot) = \underset{w(\cdot)}{\operatorname{argmax}}\, E[w(X) \cdot (m(X) - c)] \qquad \text{subject to} \qquad \pi = 0.$$

Fairness and $\mathscr{D}$'s objective

Observation

Suppose that W, M are binary ("classification"), and that

1. $m(X) = M$ *(perfect predictability), and*
2. $w^*(x) = \mathbf{1}(m(X) > c)$ *(unconstrained maximization of $\mathscr{D}$'s objective μ).*

Then $w^*(x)$ satisfies predictive parity, i.e., $\pi = 0$.

In words:

- If $\mathscr{D}$ is a firm that is maximizing profits
- and has perfect surveillance capacity
- then everything is fair by assumption
- no matter how unequal the outcomes within and across groups!
- Only deviations from profit-maximization are "unfair."

Reasons for bias

1. **Preference-based** discrimination.
   The decision maker is maximizing some objective other than μ.

2. **Mis-measurement** and biased beliefs.
   Due to bias of past data, $m(X) \neq E[M|X]$.

3. **Statistical discrimination**.
   Even if $w^*(\cdot) = \text{argmax } \pi$ and $m(X) = E[M|X]$,
   $w^*(\cdot)$ might violate fairness if X does not perfectly predict M.

Three limitations of "fairness" perspectives

1. They legitimize and perpetuate **inequalities justified by "merit."**
   Where does inequality in M come from?

2. They are **narrowly bracketed**.
   Inequality in W in the algorithm,
   instead of some outcomes Y in a wider population.

3. Fairness-based perspectives **focus on categories** (protected groups)
   and ignore within-group inequality.

$\Rightarrow$ We consider the impact on inequality or welfare as an alternative.

The impact on inequality or welfare as an alternative

- Outcomes are determined by the **potential outcome equation**

$$Y = W \cdot Y^1 + (1 - W) \cdot Y^0.$$

- The **realized outcome** distribution is given by

$$p_{Y,X}(y,x) = \int \left[p_{Y^0|X}(y,x) + w(x) \cdot \left(p_{Y^1|X}(y,x) - p_{Y^0|X}(y,x) \right) \right] p_X(x) dx.$$

- What is the impact of $w(\cdot)$ on a **statistic** ν?

$$\nu = \nu(p_{Y,X}).$$

- Examples:
  - Variance $\mathrm{Var}(Y)$,
  - "welfare" $E[Y^\gamma]$,
  - between-group inequality $E[Y|A=1] - E[Y|A=0]$.

The impact on inequality or welfare as an alternative

- Outcomes are determined by the **potential outcome equation**

$$Y = W \cdot Y^1 + (1 - W) \cdot Y^0.$$

- The **realized outcome** distribution is given by

$$p_{Y,X}(y,x) = \int \left[p_{Y^0|X}(y,x) + w(x) \cdot \left(p_{Y^1|X}(y,x) - p_{Y^0|X}(y,x) \right) \right] p_X(x) dx.$$

- What is the impact of $w(\cdot)$ on a **statistic** ν?

$$\nu = \nu(p_{Y,X}).$$

- Examples:
  - Variance $\mathrm{Var}(Y)$,
  - "welfare" $E[Y^\gamma]$,
  - between-group inequality $E[Y|A=1] - E[Y|A=0]$.

Influence function approximation of the statistic ν

$$\nu(p_{Y,X}) - \nu(p^*_{Y,X}) \approx E[IF(Y,X)],$$

- $IF(Y,X)$ is the influence function of $\nu(p_{Y,X})$.
  (Formally, the Riesz representer of the Fréchet derivative of ν.)
- The expectation averages over the distribution $p_{Y,X}$.
- Examples:

$$\begin{aligned}
\nu &= E[Y] & IF &= Y - E[Y] \\
\nu &= \mathrm{Var}(Y) & IF &= (Y - E[Y])^2 - \mathrm{Var}(Y) \\
\nu &= E[Y|A=1] - E[Y|A=0] & IF &= Y \cdot \left(\frac{A}{E[A]} - \frac{1-A}{1-E[A]} \right).
\end{aligned}$$

Influence function approximation of the statistic ν

$$\nu(p_{Y,X}) - \nu(p^*_{Y,X}) \approx E[IF(Y,X)],$$

- $IF(Y,X)$ is the influence function of $\nu(p_{Y,X})$.
  (Formally, the Riesz representer of the Fréchet derivative of ν.)

- The expectation averages over the distribution $p_{Y,X}$.

- Examples:

$$\begin{aligned}
\nu &= E[Y] & IF &= Y - E[Y] \\
\nu &= \mathrm{Var}(Y) & IF &= (Y - E[Y])^2 - \mathrm{Var}(Y) \\
\nu &= E[Y|A=1] - E[Y|A=0] & IF &= Y \cdot \left(\frac{A}{E[A]} - \frac{1-A}{1-E[A]} \right).
\end{aligned}$$

The impact of marginal policy changes on profits, fairness, and inequality

Proposition

Consider a family of assignment policies $w(x) = w^*(x) + \epsilon \cdot dw(x)$*. Then*

$$\partial_\epsilon \mu = E[dw(X) \cdot l(X)], \quad \partial_\epsilon \pi = E\left[dw(X) \cdot p(X)\right], \quad \partial_\epsilon \nu = E[dw(X) \cdot n(X)],$$

where

$$\begin{aligned} l(X) &= E[M|X = x] - c, \\ p(X) &= E\left[(M - E[M|W = 1, A = 1]) \cdot \frac{A}{E[WA]} \right. \\ &- \left. (M - E[M|W = 1, A = 0]) \cdot \frac{(1 - A)}{E[W(1 - A)]} \right| X = x \right], \\ n(x) &= E\left[IF(Y^1, x) - IF(Y^0, x) | X = x\right]. \end{aligned}$$

The impact of marginal policy changes on profits, fairness, and inequality

Proposition

Consider a family of assignment policies $w(x) = w^*(x) + \epsilon \cdot dw(x)$. *Then*

$$\partial_\epsilon \mu = E[dw(X) \cdot l(X)], \quad \partial_\epsilon \pi = E\left[dw(X) \cdot p(X)\right], \quad \partial_\epsilon \nu = E[dw(X) \cdot n(X)],$$

where

$$\begin{aligned} l(X) &= E[M|X = x] - c, \\ p(X) &= E\left[(M - E[M|W = 1, A = 1]) \cdot \frac{A}{E[WA]}\right. \\ &- \qquad \left.(M - E[M|W = 1, A = 0]) \cdot \frac{(1 - A)}{E[W(1 - A)]}\right| X = x \Bigg], \\ n(x) &= E\left[IF(Y^1, x) - IF(Y^0, x)|X = x\right]. \end{aligned}$$

Uses of the proposition 1

1. Elucidate the **tension** between objectives.
   - Profits vs. fairness vs. equality vs. welfare?
   - Suppose $\pi < 0$, $n(x) > 0$ is positive, while $p(x) < 0$.
     Then increasing $w(x)$ is good for welfare and bad for fairness.
   - $\Rightarrow$ Characterizes which parts of the feature space drive
     the tension between alternative objectives.
2. Solve for **optimal assignment** subject to constraints.
   - E.g. maximize μ subject to $\pi = 0$.
   - Then $w(x) = \mathbf{1}(l(x) > \lambda p(x))$.

Uses of the proposition 1

1. Elucidate the **tension** between objectives.
    - Profits vs. fairness vs. equality vs. welfare?
    - Suppose $\pi < 0$, $n(x) > 0$ is positive, while $p(x) < 0$.
      Then increasing $w(x)$ is good for welfare and bad for fairness.
    - $\Rightarrow$ Characterizes which parts of the feature space drive the tension between alternative objectives.
2. Solve for **optimal assignment** subject to constraints.
    - E.g. maximize μ subject to $\pi = 0$.
    - Then $w(x) = \mathbf{1}(l(x) > \lambda p(x))$.

Uses of the proposition 1, continued

3. **Power and inverse welfare weights**
   - For a given $w(\cdot)$, what objective is implicitly maximized?
   - What are the weights for different individuals that rationalize $w(\cdot)$?
4. **Algorithmic auditing**.
   - Similar to distributional decompositions in labor economics.
   - Cf. Fortin and Lemieux (1997); Firpo et al. (2009).

Uses of the proposition 1, continued

3. **Power and inverse welfare weights**
   - For a given $w(\cdot)$, what objective is implicitly maximized?
   - What are the weights for different individuals that rationalize $w(\cdot)$?
4. **Algorithmic auditing**.
   - Similar to distributional decompositions in labor economics.
   - Cf. Fortin and Lemieux (1997); Firpo et al. (2009).

A recipe for algorithmic auditing in 5 steps

1. **Normative choices**
   - Relevant outcomes Y for individuals' welfare?
   - Measures of welfare or inequality ν? Quantiles are a good default!
2. **Calculation of influence functions**
   - At the appropriate baseline distribution,
   - evaluated for each (Y_i, X_i), and stored in a new variable.
3. **Causal effect estimation**
   - Estimate and impute $n(x) = E\left[IF(Y^1, x) - IF(Y^0, x) | X = x\right]$.
   - E.g., assume $W \perp (Y^0, Y^1)|X$,
     estimate $n(\cdot)$ using causal forest approach of Wager and Athey (2018).

A recipe for algorithmic auditing in 5 steps, continued

4. **Counterfactual assignment probabilities**
   - Impute $\Delta w(x_i) = w(x_i) - w^*(x_i)$ for all i in the sample.
5. **Evaluation of distributional impact**
   - Calculate $\hat{\Delta}\nu = \alpha \cdot \frac{1}{n} \sum_i \Delta w(x_i) \cdot n(x_i)$.

The value alignment problem

- Much recent attention; e.g. Russell (2019):

  > *[...] we may suffer from a failure of value alignment—we may, perhaps inadvertently, imbue machines with objectives that are imperfectly aligned with our own*

- The debate in tech & CS focuses on robotics and the grand, e.g. Bostrom (2003):

  > *Suppose we have an AI whose only goal is to make as many paper clips as possible. The AI will realize quickly that it would be much better if there were no humans because humans might decide to switch it off. [...]*

Value alignment and observability

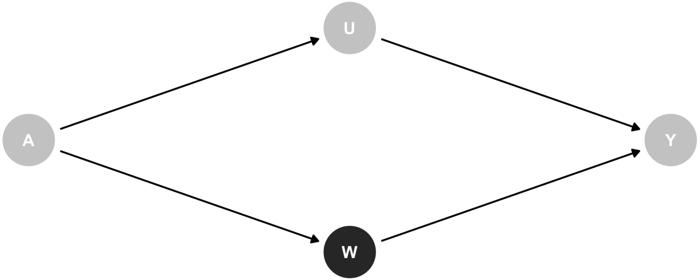
- There are many **examples** outside robotics in the **social world**, right now:
  - Social media feeds maximizing clicks.
  - Teachers promoted based on student test scores.
  - Doctors paid per patient.
  - ...
- One unifying theme: Lack of **observability** of welfare.
- How to design
  reward functions / incentive systems / adaptive treatment assignment algorithms,
  when our true objective is not observed?

Static setup

- Action $A \in \mathscr{A} \subset \mathbb{R}^n$,
- observed mediators (surrogates) $W \in R^k$,
- unobserved mediators $U \in R^l$,
- unobserved outcome (welfare) $Y \in R$.

$$
\begin{aligned}
W &= g_w(A, \epsilon_w), \\
U &= g_u(A, \epsilon_u), \\
Y &= g_y(W, U, \epsilon_y) \\
&= h(A, \epsilon_w, \epsilon_u, \epsilon_y).
\end{aligned}
$$

Both U and W are mediators

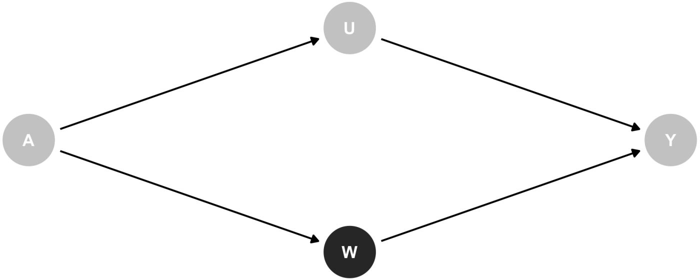


Static setup

- Action $A \in \mathscr{A} \subset \mathbb{R}^n$,
- observed mediators (surrogates) $W \in R^k$,
- unobserved mediators $U \in R^l$,
- unobserved outcome (welfare) $Y \in R$.

$$
\begin{aligned}
W &= g_w(A, \epsilon_w), \\
U &= g_u(A, \epsilon_u), \\
Y &= g_y(W, U, \epsilon_y) \\
&= h(A, \epsilon_w, \epsilon_u, \epsilon_y).
\end{aligned}
$$

Both U and W are mediators



Mis-specified reward and regret

- Reward function $R = r(W)$.
- We want to analyze algorithms that maximize R (rather than Y).
- Optimal and pseudo-optimal action, regret

$$
\begin{aligned}
a^* &= \operatorname*{argmax}_{a \in \mathscr{A}} \; E[Y|do(A=a)] \\
a^+ &= \operatorname*{argmax}_{a \in \mathscr{A}} \; E[R|do(A=a)] \\
\Delta &= E[Y|do(A=a^*)] - E[Y|do(A=a^+)].
\end{aligned}
$$

(Using the do-calculus notation of Pearl 2000).

Related Literature 1: Multi-tasking

- Holmstrom and Milgrom (1991):
  Why are high-powered economic incentives rarely observed?
- Agent chooses effort $A \in \mathbb{R}^2$,
  to maximize their expectation of utility R,
  which depends on a monetary reward
  that in turn is a linear function of noisy measures of effort, (W_1, W_2).

Multi-tasking, continued

- In their model, the certainty equivalent of agent rewards equals

$$\alpha + \beta_1 a_1 + \beta_2 a_2 - C(a_1 + a_2) - \frac{1}{2} r \left(\beta_1^2 \sigma_1^2 + \beta_2^2 \sigma_2^2\right),$$

  where $C(\cdot)$ is the increasing and concave cost of effort.
- $\Rightarrow$ Positive effort a_j on both dimension only if $\beta_1 = \beta_2$.
- If $\sigma_j = \infty$ for one j (unobservable component),
  $\beta_1 = \beta_2 = 0$ for the optimal contract.

Related Literature 2: Surrogate oucomes

- Athey et al. (2019): Suppose W satisfies the surrogacy condition

$$A \perp Y | W.$$

- For exogenous A, this holds if all causal pathways from A to Y go through W.
- Let $\hat{y}(W) = E[Y|W]$, estimated from auxiliary data. Then

$$E[Y|A] = E[\hat{y}(W)|A].$$

- Implication: For $r(W) = \hat{y}(W)$, $a^+ = a^*$!

Related Literature 2: Surrogate oucomes

- Athey et al. (2019): Suppose W satisfies the surrogacy condition

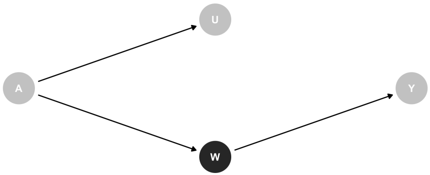
$$A \perp Y | W.$$

- For exogenous A, this holds if all causal pathways from A to Y go through W.
- Let $\hat{y}(W) = E[Y|W]$, estimated from auxiliary data. Then

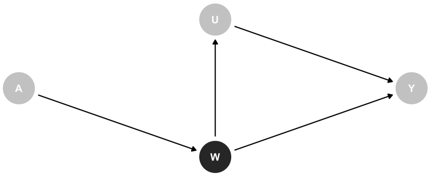
$$E[Y|A] = E[\hat{y}(W)|A].$$

- Implication: For $r(W) = \hat{y}(W)$, $a^+ = a^*$!

No effect from U to Y



W is a mediator for U



Related Literature 2: Surrogate oucomes

- Athey et al. (2019): Suppose W satisfies the surrogacy condition

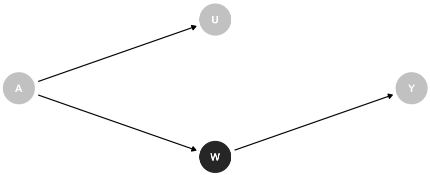
$$A \perp Y | W.$$

- For exogenous A, this holds if all causal pathways from A to Y go through W.
- Let $\hat{y}(W) = E[Y|W]$, estimated from auxiliary data. Then

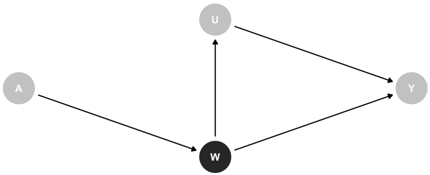
$$E[Y|A] = E[\hat{y}(W)|A].$$

- Implication: For $r(W) = \hat{y}(W)$, $a^+ = a^*$!

No effect from U to Y



W is a mediator for U



Discrete actions, linear rewards, and regret

- Suppose the set of feasible actions is finite.
- Allow for randomized actions.

$$A \in \left\{ a : \sum_i a_i = 1, \; a_i \geq 0 \right\}.$$

- Vector / matrix of linear regression coefficients $\beta^{Y|A}$, $\beta^{W|A}$, $\beta^{Y|W}$.
- Linear rewards $r(W) = W \cdot \rho$.
- $\Rightarrow$ Regret bound:

$$\Delta \leq 2 \cdot \left\| \beta^{Y|A} - (\beta^{W|A} \cdot \rho) \right\|_\infty$$

- Surrogate rewards: $\rho = \beta^{Y|W}$.
- Optimal rewards: $\rho^* \in \text{argmin}_{\rho} \max_j \beta_j^{Y|A} - \beta_{j^+}^{Y|A}$.

$$\Rightarrow \quad \Delta \leq 2 \cdot \min_\rho \left\| \beta^{Y|A} - (\beta^{W|A} \cdot \rho) \right\|_\infty .$$

Discrete actions, linear rewards, and regret

- Suppose the set of feasible actions is finite.
- Allow for randomized actions.

$$A \in \left\{ a : \sum_i a_i = 1, \; a_i \geq 0 \right\}.$$

- Vector / matrix of linear regression coefficients $\beta^{Y|A}$, $\beta^{W|A}$, $\beta^{Y|W}$.
- Linear rewards $r(W) = W \cdot \rho$.
- $\Rightarrow$ Regret bound:

$$\Delta \leq 2 \cdot \left\| \beta^{Y|A} - (\beta^{W|A} \cdot \rho) \right\|_\infty$$

- Surrogate rewards: $\rho = \beta^{Y|W}$.
- Optimal rewards: $\rho^* \in \text{argmin}_\rho \max_j \beta_j^{Y|A} - \beta_{j^+}^{Y|A}$.

$$\Rightarrow \quad \Delta \leq 2 \cdot \min_\rho \left\| \beta^{Y|A} - (\beta^{W|A} \cdot \rho) \right\|_\infty .$$

Dynamic version:
Markov Decision Problems and Reinforcement Learning

- Sutton and Barto (2018); François-Lavet et al. (2018)).
- States s, actions a, transition probabilities $P(s, a, s')$, and conditionally expected true rewards $Y(s, a)$.
- Probability $x(a|s)$ of choosing action a.
  $\Rightarrow$ Markov process with stationary distribution $\pi(s, a)$, average welfare

$$V_x = \operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} Y(S_T, A_T) = \sum_{s,a} Y(s, a) \pi(s, a).$$

- Two equivalent problems:
  Optimal policy $x(s, a) \Leftrightarrow$ Optimal stationary distribution $\pi(s, a)$.

Mis-specified rewards in the dynamic setting

- $\Rightarrow$ Linear programming problem

$$\pi^*_Y(\cdot) = \operatorname{argmax} \sum_{s,a} Y(s,a)\pi(s,a), \text{ subject to}$$

$$\sum_{s,a} \pi(s,a) = 1, \qquad \pi(s,a) > 0 \text{ for all } s, a.$$

$$\sum_{a} \pi(s',a) = \sum_{s,a} \pi(s,a) \cdot P(s,a,s') \text{ for all } s' \tag{1}$$

- Replacing $Y(\cdot)$ with potentially misspecified rewards $R(\cdot)$
  $\Rightarrow \pi^*_R$, with expected regret

$$\Delta = \sum_{s,a} Y(s,a)\left(\pi^*_Y(s,a) - \pi^*_R(s,a)\right).$$

- Formally this is like the static setup, with the added constraint (1)!

Mis-specified rewards in the dynamic setting

- $\Rightarrow$ Linear programming problem

$$\pi^*_Y(\cdot) = \operatorname{argmax} \sum_{s,a} Y(s,a)\pi(s,a), \text{ subject to}$$

$$\sum_{s,a} \pi(s,a) = 1, \qquad \pi(s,a) > 0 \text{ for all } s, a.$$

$$\sum_{a} \pi(s',a) = \sum_{s,a} \pi(s,a) \cdot P(s,a,s') \text{ for all } s' \tag{1}$$

- Replacing $Y(\cdot)$ with potentially misspecified rewards $R(\cdot)$
  $\Rightarrow \pi^*_R$, with expected regret

$$\Delta = \sum_{s,a} Y(s,a) \left(\pi^*_Y(s,a) - \pi^*_R(s,a)\right).$$

- Formally this is like the static setup, with the added constraint (1)!

Mis-specified rewards in the dynamic setting

- $\Rightarrow$ Linear programming problem

$$\pi_Y^*(\cdot) = \text{argmax} \sum_{s,a} Y(s,a)\pi(s,a), \text{ subject to}$$

$$\sum_{s,a} \pi(s,a) = 1, \qquad \pi(s,a) > 0 \text{ for all } s,a.$$

$$\sum_{a} \pi(s',a) = \sum_{s,a} \pi(s,a) \cdot P(s,a,s') \text{ for all } s' \tag{1}$$

- Replacing $Y(\cdot)$ with potentially misspecified rewards $R(\cdot)$
  $\Rightarrow \pi_R^*$, with expected regret

$$\Delta = \sum_{s,a} Y(s,a) \left(\pi_Y^*(s,a) - \pi_R^*(s,a)\right).$$

- Formally this is like the static setup, with the added constraint (1)!

Next steps

- Compare regret
  1. For arbitrary $R = r(W)$,
  2. for surrogate rewards $R = E[Y|W]$,
  3. for the oracle optimal (regret minimizing) R.
- Worst-case bounds on regret,
  as a function of the characteristics of the setting.
- Empirical examples.
- Questions:
  1. ML references relating value alignment to observability?
  2. What results would get CS audience interested
     in the connection to multi-tasking / surrogates?

Motivation: Ethics of experimentation

- "Active learning" often involves experimentation in social contexts.
- This is especially sensitive when vulnerable populations are affected.
- (When) is it ethical to experiment on humans?

Kant (1791):

> *Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end.*

Combinatorial allocation problems

- Our motivating example:
  - **Refugee resettlement** in the US: Done by resettlement agencies like HIAS.
  - A fairly small number of slots in different locations is available.
  - Refugees without prior ties have been distributed pretty much randomly.
  - Alex Teytelboym's prior work with HIAS, Annie MOORE:
    Estimate refugee-location match effects on employment, using past data,
    find optimal matching, implement.
  - This project: **Learning while matching.**

- Many policy problems have a similar form:
  - Resources, agents, or locations need to be allocated to each other.
  - There are various feasibility constraints.
  - The returns of different options (combinations) are unknown.
  - The decision has to be made repeatedly.

Sketch of setup

- There are J **options** (e.g., matches) available to the policymaker.
- Every period, the policymaker's **action** is to choose at most M options.
- Before the next period, the policymaker observes the **outcomes** of every chosen option (combinatorial semi-bandit setting).
- The policymaker's **reward** is the sum of the outcomes of the chosen options.
- The policymaker's **objective** is to maximize the cumulative expected rewards.
- Equivalently, the policymaker's objective is to minimize **expected regret**—the shortfall of cumulative expected rewards relative to the oracle optimum.

Notation:

- Actions

$$a \in \mathcal{A} \subseteq \{a \in \{0,1\}^J : \|a\|_1 = M\}.$$

- Expected reward:

$$R(a) = \mathbf{E}[\langle a, Y_t \rangle | \Theta] = \langle a, \Theta \rangle.$$

Sketch of setup

- There are J **options** (e.g., matches) available to the policymaker.
- Every period, the policymaker's **action** is to choose at most M options.
- Before the next period, the policymaker observes the **outcomes** of every chosen option (combinatorial semi-bandit setting).
- The policymaker's **reward** is the sum of the outcomes of the chosen options.
- The policymaker's **objective** is to maximize the cumulative expected rewards.
- Equivalently, the policymaker's objective is to minimize **expected regret**—the shortfall of cumulative expected rewards relative to the oracle optimum.

Notation:

- Actions
$$a \in \mathcal{A} \subseteq \{a \in \{0,1\}^J : \|a\|_1 = M\}.$$
- Expected reward:
$$R(a) = \mathbf{E}[\langle a, Y_t \rangle | \Theta] = \langle a, \Theta \rangle.$$

Thompson sampling

- Take a random action $a \in \mathcal{A}$, sampled according to the distribution

$$\mathbf{P}_t(A_t = a) = \mathbf{P}_t(A_t^* = a),$$

  where $\mathbf{P}_t$ is the posterior at the beginning of period t.
- This assumption implies in particular that

$$\mathbf{E}_t[A_t] = \mathbf{E}_t[A^*].$$

- Introduced by Thompson (1933) for treatment assignment in adaptive experiments.

Regret bound

Theorem

Under the assumptions just stated,

$$\mathbf{E}_1\left[\sum_{t=1}^{T}(R(A^*)-R(A_t))\right] \leq \sqrt{\frac{1}{2}JTM\cdot\left[\log\left(\frac{J}{M}\right)+1\right]}.$$

Features of this bound:

- It holds in finite samples, there is no remainder.
- It does not depend on the prior distribution for Θ.
- It allows for prior distributions with arbitrary statistical dependence across the components of Θ.
- It implies that Thompson sampling achieves the efficient rate of convergence.

Regret bound

Theorem

Under the assumptions just stated,

$$\mathbf{E}_1\left[\sum_{t=1}^{T}\left(R(A^*)-R(A_t)\right)\right]\leq\sqrt{\frac{1}{2}JTM\cdot\left[\log\left(\tfrac{J}{M}\right)+1\right]}.$$

Verbal description of this bound:

- The worst case expected regret (per unit) across all possible priors goes to 0 at a rate of 1 over the square root of the sample size, $T\cdot M$.
- The bound grows, as a function of the number of possible options J, like $\sqrt{J}$ (ignoring the logarithmic term).
- Worst case regret per unit does not grow in the batch size M, despite the fact that action sets can be of size $\binom{J}{M}$!

Key steps of the proof

1. Use Pinsker's inequality to **relate expected regret** to the information about the optimal action A^*. Information is measured by the **KL-distance** of posteriors and priors.

   (This step draws on Russo and Van Roy (2016).)

2. Relate the **KL-distance** to the **entropy reduction** of the events $A_j^* = 1$.

   The combination of these two arguments allows to bound the expected regret for option j in terms of the entropy reduction for the posterior of A_j^*.

   (This step draws on Bubeck and Sellke (2020).)

3. The total **reduction of entropy** across the options j, and across the time periods t, can be no more than the **sum of the prior entropy** for each of the events $A_j^* = 1$, which is bounded by $M \cdot \left[\log\left(\frac{J}{M}\right) + 1\right]$.

Key steps of the proof

1. Use Pinsker's inequality to **relate expected regret** to the information about the optimal action A^*. Information is measured by the **KL-distance** of posteriors and priors.

   (This step draws on Russo and Van Roy (2016).)

2. Relate the **KL-distance** to the **entropy reduction** of the events $A_j^* = 1$.

   The combination of these two arguments allows to bound the expected regret for option j in terms of the entropy reduction for the posterior of A_j^*.

   (This step draws on Bubeck and Sellke (2020).)

3. The total **reduction of entropy** across the options j, and across the time periods t, can be no more than the **sum of the prior entropy** for each of the events $A_j^* = 1$, which is bounded by $M \cdot \left[\log\left(\frac{J}{M}\right) + 1\right]$.

Key steps of the proof

1. Use Pinsker's inequality to **relate expected regret** to the information about the optimal action A^*. Information is measured by the **KL-distance** of posteriors and priors.

   (This step draws on Russo and Van Roy (2016).)

2. Relate the **KL-distance** to the **entropy reduction** of the events $A_j^* = 1$.

   The combination of these two arguments allows to bound the expected regret for option j in terms of the entropy reduction for the posterior of A_j^*.

   (This step draws on Bubeck and Sellke (2020).)

3. The total **reduction of entropy** across the options j, and across the time periods t, can be no more than the **sum of the prior entropy** for each of the events $A_j^* = 1$, which is bounded by $M \cdot \left[\log\left(\frac{J}{M}\right) + 1\right]$.

Conclusion and summary

- Artificial intelligence, machine learning, algorithmic decision making raise many normative questions.
- Especially when applied in consequential settings in social contexts.
- Many of these normative questions echo those encountered by economists in non-automated settings.
  $\Rightarrow$ Much scope for interdisciplinary exchange!
- This talk: Three projects in this agenda, connecting
  1. Algorithmic fairness to empirical tests for discrimination, social welfare analysis, distributional decompositions.
  2. The value alignment problem and reinforcement learning to multi-tasking, surrogate outcomes.
  3. Active learning and ethics of experiments to matching markets, semi-bandits.
- To be continued...

Thank you!