Understanding, Predicting, & Influencing Human Decision Making

Krishna P. Gummadi
Max Planck Institute for Software Systems

Outline for my lectures

- Lecture 1:
  - Overview of (my) Social Computing Research
- Lecture 2:
  - On the Temporality of Trust and Privacy
- Lecture 3:
  - On Biases in Search & Recommendations in Crowdsourcing Systems

What is social computing?

- An emerging inter-disciplinary area
- Hard to define precisely
- Lets change the question
- Why do we do Social Computing?

Why we do social computing

- It allows us to flourish
  - Human flourishing: That towards which all activities aim

- "Mathematics for Human Flourishing" by Francis Yu
  - 5 basic desires for human activity
    1. For play or fun
    2. For seeking truth
    3. Pursuing beauty
    4. Fighting for justice
    5. Love for other human beings

1. For Play or Fun

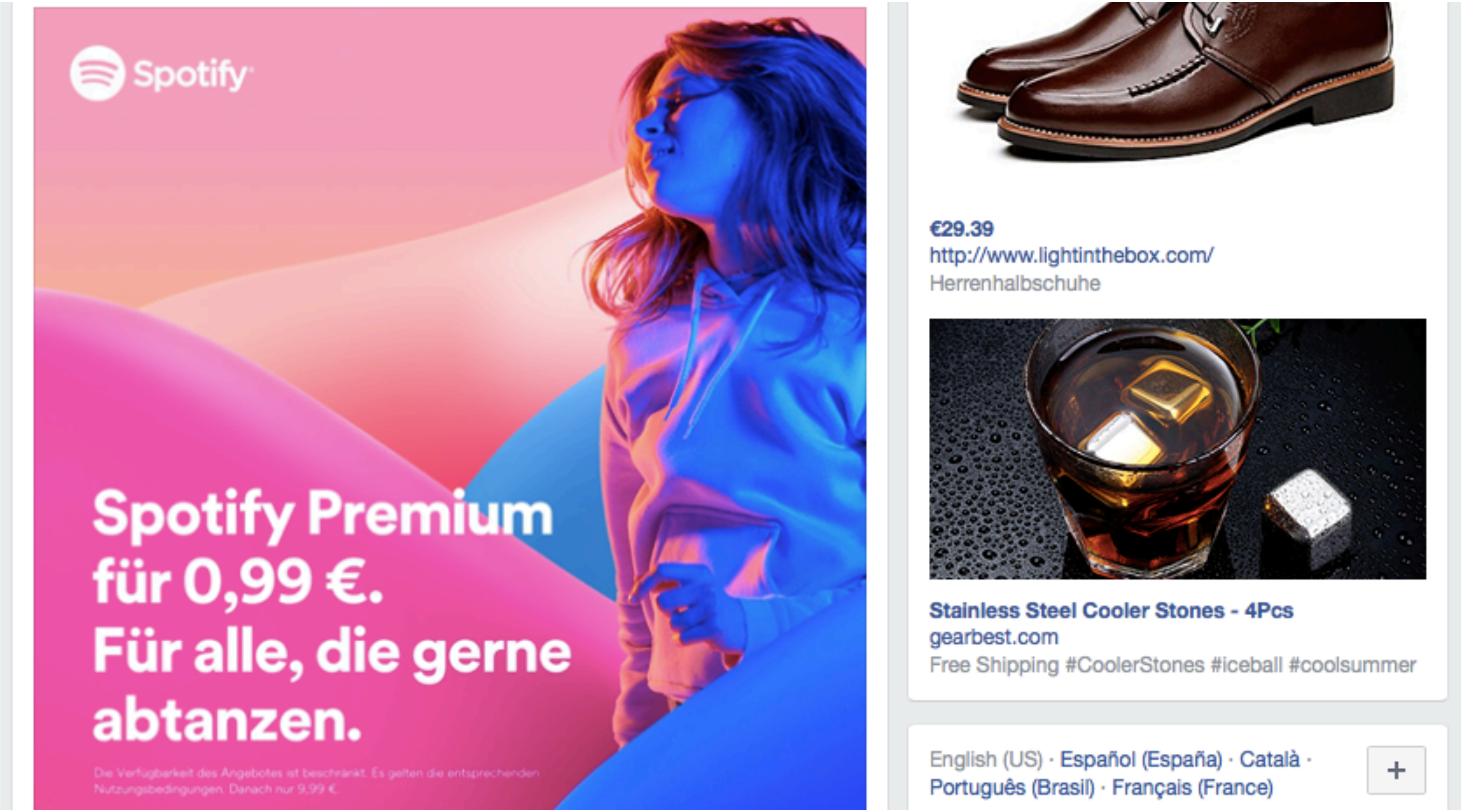
Fun with societal-scale datasets

Facebook ad platform

- By far, the largest social media platform
  - In terms of number of users
  - In terms of data aggregated on users
  - In terms of advertisers & ad revenues
  - In terms of introducing novel & provocative targeting practices

- However, many issues discussed generalize to other social media platforms
  - Like LinkedIn, Twitter, YouTube, Instagram etc.,

Ads on Facebook



Data used for targeting ads

- Facebook gathers lots of data (features) on users

Category	Behaviors															
	Anniversary	*Consumer Class.*	*Digital activities*	*Expats*	*Mobile Device User*	*Multicultural Affinity*	*Seasonal and Events*	*Travel*	*Automotive*	*Charitable donations*	*Financial*	*Job role*	*Media*	*Purchase behavior*	*Residential profiles*	*Business-to-business*
Facebook	1	2	39	74	81	6	2	5								
Acxiom									1	5	25	2	35	23	2	
Epsilon										4	1			5	2	
Experian														3	1	
Datalogix								11	151			1		144		29
Total	1	2	39	74	81	6	2	16	152	9	26	3	35	175	5	29

Category	Demographics								
	Education level	*Generation*	*Home*	*Life Events*	*Parents*	*Politics (US)*	*Relationship status*	*Work*	*Financial*
Facebook	13	3	2	36	9	8	16	26	
Acxiom			19						16
Epsilon						2			
Experian			1						
Datalogix			2		11			1	
Total	13	3	24	36	20	10	16	27	16

Category	Interests										
	Business and industry	*Entertainment*	*Family/relationships*	*Fitness and wellness*	*Food and drink*	*Hobbies and activities*	*Shopping and fashion*	*Sports and outdoors*	*Technology*	*Other*	**Total**
Facebook	39	70	8	11	37	60	21	22	21	2	614
Acxiom											128
Epsilon											14
Experian											5
Datalogix											350
Total	39	70	8	11	37	60	21	22	21	2	1,111

Closer look at features: Examples

- Demographical (gathered by Facebook)
  - Relationship:
    - Interested In: Men and Women, Men, Unspecified, Women
    - Status: Separated, Widowed, Open Relationship, Divorced, In a relationship, Married, Engaged, Unspecified, Single, Complicated Civil Union, Domestic Partnership
- Each user feature is a boolean variable

- Demographical (aggregated from data brokers)
  - Financial:
    - Income: Geschätztes monatliches Nettoeinkommen 2.000 bis 2.600, 2.600 bis 3.600, 3.600 bis 5.000, über 5.000 Euro

Data aggregation across countries

Country	Facebook	Epsilon	DLX	Experian	Acxiom	Total
US	598	14	350	5	128	1105
UK	598	0	19	17	103	737
France	598	0	0	0	21	619
Germany	598	0	0	0	60	658
Australia	598	0	0	34	24	656
Brazil	598	0	0	20	0	618
Japan	598	0	0	0	17	615
South Korea	598	0	0	0	0	598
Canada	598	0	0	0	0	598
India	598	0	0	0	0	598

How Facebook advertisers target users

1) Traditional targeting: Advertisers specify a boolean formula over the features

- Typically, in a restricted CNF form
  - (F1 v F2 v F3….) ^ (F’1 v F’2 v F’3….) ^ ….. ^ -FK ^ -F’K
- Users are targeted, when their feature values inferred by Facebook satisfy the targeting formula

How Facebook advertisers target users

2) Custom audience targeting: Advertisers upload PII of users they wish to target

Site	Custom audience matching attributes												Minimum Audience	Available targeting attributes								
	First name	Last name	Email	Phone number	ZIP	City	State/Province	Birthday	Gender	Employer	Site user ID	Mobile advertiser ID		Location	Age	Gender	Language	Interests	Behaviors	Mobile device	Keywords	Search terms
Facebook	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	20	✓	✓	✓	✓	✓	✓	✗	✗	✗
Instagram	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	20	✓	✓	✓	✓	✓	✓	✗	✗	✗
Twitter	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓	500	✓	✗	✓	✓	✓	✓	✓	✓	✗
Google	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	1,000	✓	✗	✗	✓	✗	✗	✓	✓	✓
Pinterest	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	100	✓	✗	✓	✓	✓	✗	✓	✓	✗
LinkedIn	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗	✓	100	✓	✓	✓	✗	✗	✗	✗	✗	✗

- Advertisers love custom audience

DEMO

What can we do with FB ad APIs?

- The Good:
  - Journalism and Media Studies
  - Demographics Research
- The Bad:
  - Potential for discriminatory ads
  - Potential for privacy risks

Can ad targeting be discriminatory?

Online **Ads** for High-Paying Jobs Are Targeting Men More Than ...
Adweek - 7 Jul 2015
"We found small instances where there was **discrimination** and **gender-based discrimination** in job **ads**," said ... The issue of bias and **discrimination** in ad targeting isn't new, ... The **Carnegie Mellon researchers** also were alarmed by **ads** that ... (Google's rules forbid serving **ads** based on health information.).

Carnegie Mellon Study Finds **Gender Discrimination** In **Ads** Shown ...
Marketing Land - 8 Jul 2015

When Algorithms **Discriminate**
New York Times - 9 Jul 2015

Probing the Dark Side of Google's Ad-Targeting System
Highly Cited - MIT Technology Review - 6 Jul 2015

Google's Ad System Has Become Too Big to Control
In-Depth - Wired - 9 Jul 2015

Google's algorithm shows prestigious job **ads** to men, but not to ...
Highly Cited - Washington Post - 6 Jul 2015

Marketing Land | New York Tim... | The Guardian | Washington P... | Chicago Tribu... | TechCrunch

View all

Can ad targeting be discriminatory?

[Facebook To Ban '**Ethnic Affinity**' Targeting For Housing ...]
Forbes - 11 Nov 2016
After weeks of push back from U.S. lawmakers, media and civil rights leaders, Facebook FB +0.18% on Friday announced it will stop allowing ...

Facebook disables '**ethnic affinity**' ads for housing, jobs
Engadget - 11 Nov 2016

Facebook to stop ads that target, exclude races
Highly Cited - USA TODAY - 11 Nov 2016

Facebook to Remove '**Ethnic Affinity**' Targeting From Certain Ad ...
International - AgencySpy - 11 Nov 2016

Facebook has discriminated against you, and it's not going to stop
In-Depth - Mashable - 12 Nov 2016

Facebook Bans Targeting Based on Race and **Ethnicity** for Housing ...
Blog - Slate Magazine (blog) - 11 Nov 2016

Engadget TechCrunch Deadline Daily Mail SlashGear USA TODAY

View all

Discrimination via correlated features

- FB's early defense: Ethnic affinity is not ethnicity
- Used voter records from NC to check correlations
  - Voter records have race information
  - Created separate customer lists for different races
  - Checked correlations between their race & ethnic affinity

	Voter Records		Facebook Users					
Attribute	Number	Percent	Uploaded	Matched	Reachable	Reachable %	Corresponding	Corresponding %
White	5,303,383	70.1%	10,000	8,000	6,800	85.0%	5,700	83.8%
Black	1,694,220	22.4%	10,000	7,800	6,300	80.8%	5,200	82.5%
Asian	79,250	1.0%	10,000	7,700	6,600	85.7%	1,900	28.8%
Hispanic	163,236	2.2%	10,000	7,000	5,900	84.3%	3,000	50.8%

Does banning “ethnic affinity” help?

- What about pre-filtered custom lists:
  - using offline info like voter records?

- What of other correlated features?

no	feature name	selectivity	Blacks percentage	rest percentage	ratio
1	Demographic > Ethnic Affinity > African American (US)	17.0%	77.0%	10.9%	7.06
2	Demographic > Politics (US) > US Politics (Very Liberal)	11.8%	49.8%	7.7%	6.44
3	Interests > Entertainment > Music > Gospel music	14.4%	48.3%	14.6%	3.32
4	Interests > Shopping and fashion > Beauty > Hair products	12.2%	40.8%	12.9%	3.15

no	feature name	selectivity	Blacks percentage	rest percentage	ratio
1	Demographic > Politics (US) > US Politics (Very Conservative)	14.4%	4.8%	26.5%	0.18
2	Demographic > Politics (US) > US Politics (Conservative)	16.6%	6.5%	29.7%	0.22
3	Interests > Sports and outdoors > Outdoor recreation > Hiking	11.0%	8.0%	21.6%	0.37
4	Interests > Sports and outdoors > Outdoor recreation > Camping	11.4%	11.5%	22.8%	0.50

Open challenges

- How to detect discriminatory targeting in ads?
  - Particularly, with customer lists?

- How to avoid discriminatory targeting in ads?
  - Detecting & avoiding algo. discrimination is a hot topic
  - But, even here ads pose unique challenges

- Fair targeting might result in unfair ad impressions!
  - Targeting 100 men & 100 women might result in unequal impressions, when costs of their impressions are different!

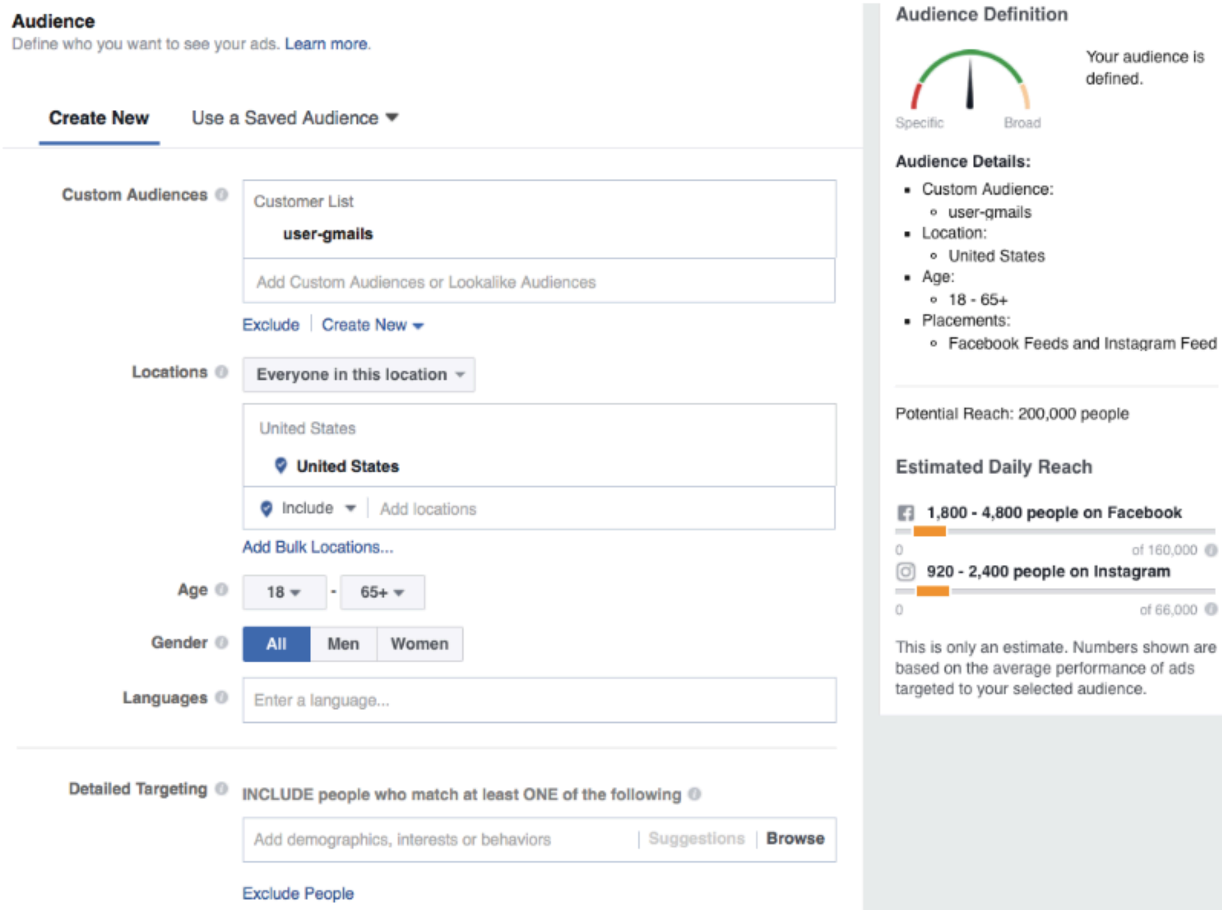
What can we do with FB ad APIs?

- The Good:
  - Journalism and Media Studies
  - Demographics Research

- The Bad:
  - Potential for discriminatory ads
  - Potential for privacy risks

Potential audience reach estimate

- A feature of Facebook's advertiser interface



Privacy risks from exact estimates

- Assume exact estimates of audience size
- Then, given a user's PII
  - Like phone-num. or email-id. or name-address
- Any advertiser can check if the user is on Facebook
- And retrieve all the user's info Facebook inferred
  - Including the financial info provided by data brokers!

Precision of audience reach estimates

- Reverse-engineered how the estimates work
- No estimates given when the audience reach < 20
- Estimates are rounded
  - Audience reach < 1000, rounded down to closest 10
  - Audience reach < 10000, rounded down to closest 100
  - Audience reach < 100000, rounded down to closest 1000
  - .....
- Are these noisy estimates sufficient for privacy?

Privacy risks from audience estimates

- Given any customer list S with and a user U’s PII
  - Like phone-num. or email-id. or name-address
- Create a new customer list with S + U
- Is audience reach for S + U is more than S?
  - If it does, user U is on FB
    - One can similarly retrieve all the info FB has on the user
  - If not, either U is not on FB OR it’s a rounding error

The probability of rounding errors

- For a list S with audience reach > 20 & < 1000
  - Rounding error probability is 0.9

- For K-lists with audience reach > 20 & < 1000
  - Chance of every try suffering a rounding error is 0.9^K
  - Chance of at-least one try not being rounded is 1-0.9^K
    - For K = 100, this is chance is 99.999974%

- So by creating 100 lists with reach > 20 & < 1000
  - One can w.h.p. retrieve all data FB has on any user

How to create such customer lists

- Use public voter records in the US!
  - Randomly sample names/addresses from records
  - Till you get a customer list of size > 20 & < 1000!
  - Repeat the process 100 times!

- Validation: Used it to retrieve all data FB has on us!
  - User transparency tool?
  - But, could be used to retrieve data on others as well!

Open challenges

- Audience estimates are very useful for advertisers
- How to preserve the estimates without data leaks
  - Differential privacy? Other noisy estimates?
    - But, how can one enforce privacy budgets?

Why we do social computing

1. ~~For play or fun~~
2. For seeking truth
3. Pursuing beauty
4. Fighting for justice
5. Love for other human beings

2. For Seeking Truth

Feasibility of fair decision making

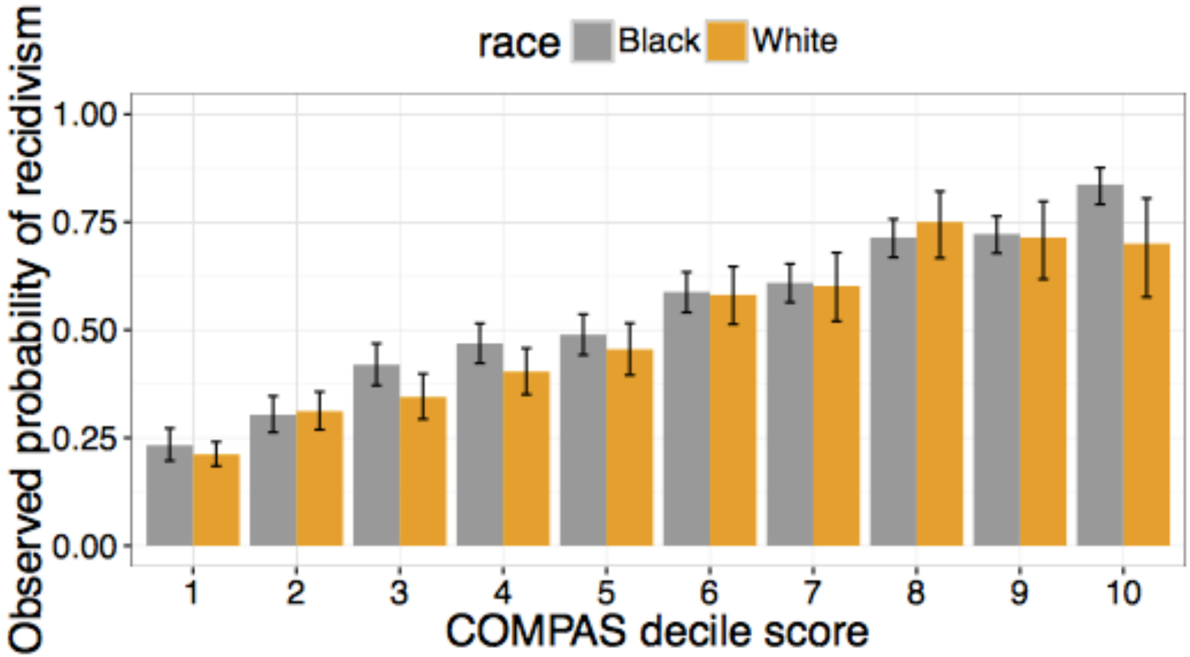
A decision making scenario: Recidivism risk prediction

- Estimate likelihood of a criminal re-offending in the near future
- Used by judges to determine sentencing terms
- Systemic bias against certain demographic groups
- Can machine judgements do better?

COMPAS recidivism risk estimator

- Built by a commercial company, Northpointe, Inc.
- Inputs: Based on a long questionnaire
- Outputs: Used across US by judges, probation, and parole officers
- Are COMPAS' estimates fair to salient social groups?

Is COMPAS fair to all groups?



- Northpointe: In each estimated risk level, blacks & whites have similar probabilities of recidivating
  - **YES!**

Is COMPAS fair to all groups?

Black Defendants	Low	High
Survived	990	805
Recidivated	532	1369
FP rate: 44.85		
FN rate: 27.99		

White Defendants	Low	High
Survived	1139	349
Recidivated	461	505
FP rate: 23.45		
FN rate: 47.72		

- ProPublica: False positive & false negative rates are considerably worse for blacks than whites
  - **NO!**

Who is right about COMPAS?

- **Both!** Depends on how you measure fairness!
- How many fairness measures can one define?
  - How many different error rate measures can one define?

		Predicted Label		
		$\hat{y} = 1$	$\hat{y} = -1$	
True Label	$y = 1$	True positive	False negative	$P(\hat{y} \neq y \mid y = 1)$ False Negative Rate
	$y = -1$	False positive	True negative	$P(\hat{y} \neq y \mid y = -1)$ False Positive Rate
		$P(\hat{y} \neq y \mid \hat{y} = 1)$ False Discovery Rate	$P(\hat{y} \neq y \mid \hat{y} = -1)$ False Omission Rate	$P(\hat{y} \neq y)$ Overall Misclass. Rate

But, aren't the measures similar?

- Doesn't satisfying one satisfy another?
  - **NO!**
- They present inherent trade-offs!
  - When base recidivism rates for blacks and whites differ, it is impossible to achieve similar FPR, FNR, & FDR!
- No solution can be simultaneously fair according to both ProPublica and Northpointe analysis!

Learning fair decision making

- To learn, we define & optimize a risk (loss) function
  - Over all examples in training data

$$L(\mathbf{w}) = \sum_{i=1}^{N} (y_i - \mathbf{w}^T \mathbf{x}_i)^2 \qquad L(\mathbf{w}) = \sum_{i=1}^{N} -\log p(y_i | \mathbf{x}_i, \mathbf{w})$$

  - Risk function captures inaccuracy in prediction
  - So learning is cast as an optimization problem

$$minimize\ L(\mathbf{w})$$

Learning fair decision making

- Idea: Cast fairness notions as constraints on learning

- Optimize for accuracy under those constraints

$$
\begin{aligned}
&\textit{minimize } L(\mathbf{w}) \\
&\textit{subject to} \quad P(\hat{y}|\mathbf{x}, z) = P(\hat{y}|\mathbf{x}) \\
&\qquad\qquad\quad\;\; P(\hat{y}=1|z=0) = P(\hat{y}=1|z=1) \\
&\qquad\qquad\quad\;\; P(\hat{y}\neq y|z=0) = P(\hat{y}\neq y|z=1)
\end{aligned}
$$

- Technical challenge: Computational methods to solve the constrained optimization efficiently

Learning fair classifiers

- Previous formulation: Non-convex, hard-to-learn

$$\begin{aligned} &\textit{minimize } L(\mathbf{w}) \\ &\textit{subject to } \; P(\hat{y} \neq y | z = 0) = P(\hat{y} \neq y | z = 1) \end{aligned}$$

Learning fair classifiers

- New formulation: Convex-concave, can learn efficiently using convex-concave programming

$$
\begin{array}{r|l}
\textit{minimize} & L(\mathbf{w}) \\
\textit{subject to} & \frac{-N_1}{N}\sum_{i=1}^{N_0} g_{\mathbf{w}}(y_i, \mathbf{x}_i) + \frac{N_0}{N}\sum_{i=1}^{N_1} g_{\mathbf{w}}(y_i, \mathbf{x}_i) \leq \mathbf{c} \\
 & \frac{-N_1}{N}\sum_{i=1}^{N_0} g_{\mathbf{w}}(y_i, \mathbf{x}_i) + \frac{N_0}{N}\sum_{i=1}^{N_1} g_{\mathbf{w}}(y_i, \mathbf{x}_i) \geq -\mathbf{c}
\end{array}
$$

All misclassifications $g_{\mathbf{w}}(y, \mathbf{x}) = min(0, y d_{\mathbf{w}}(\mathbf{x})),$

False positives $g_{\mathbf{w}}(y, \mathbf{x}) = min\left(0, \frac{1+y}{2} y d_{\mathbf{w}}(\mathbf{x})\right),$ or

False negatives $g_{\mathbf{w}}(y, \mathbf{x}) = min\left(0, \frac{1-y}{2} y d_{\mathbf{w}}(\mathbf{x})\right),$

The case for machine judgements

- Formalizes decision making
  - Requires goals & constraints to be defined clearly
  - Biases in decisions can be detected and constrained

- Reveals the inherent trade-offs between multiple learning objectives and their utility

- Can find optimal solution points that are beyond the reach of human computational abilities

Why we do social computing

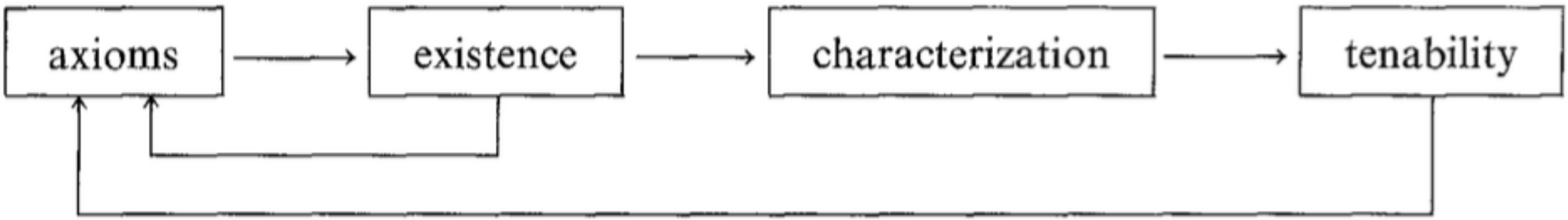
1. ~~For play or fun~~
2. ~~For seeking truth~~
3. Pursuing beauty
4. Fighting for justice
5. Love for other human beings

3. Pursuing Beauty

Rethinking fair decision making

How to reason about fairness?

- Reflective Equilibrium
  - From A Theory of Justice by John Rawls



- Is there a role for human moral instincts?

Our fairness axioms, so far

	A_1	A_2	…	A_m
$User_1$	$x_{1,1}$	$x_{1,2}$	…	$x_{1,m}$
$User_2$	$x_{2,1}$			$x_{2,m}$
$User_3$	$x_{3,1}$			$x_{3,m}$
…	…			…
$User_n$	$x_{n,1}$	$x_{n,2}$	…	$x_{n,m}$

Decision
Accept
Reject
Reject
…
Accept

- Based only on what the decision outcomes are
  - Inputs / outputs of the black-box
- Not on the process of decision making
  - How the input data was used to arrive at decisions

Revisiting recidivism risk prediction

- Is it fair to use the data in COMPAS questionnaire?

Feature
prior offenses arrest charge description arrest charge degree # juvenile felony offenses # juvenile misdemeanor offenses # juvenile other offenses age gender race

- What do our human moral instincts tell us?

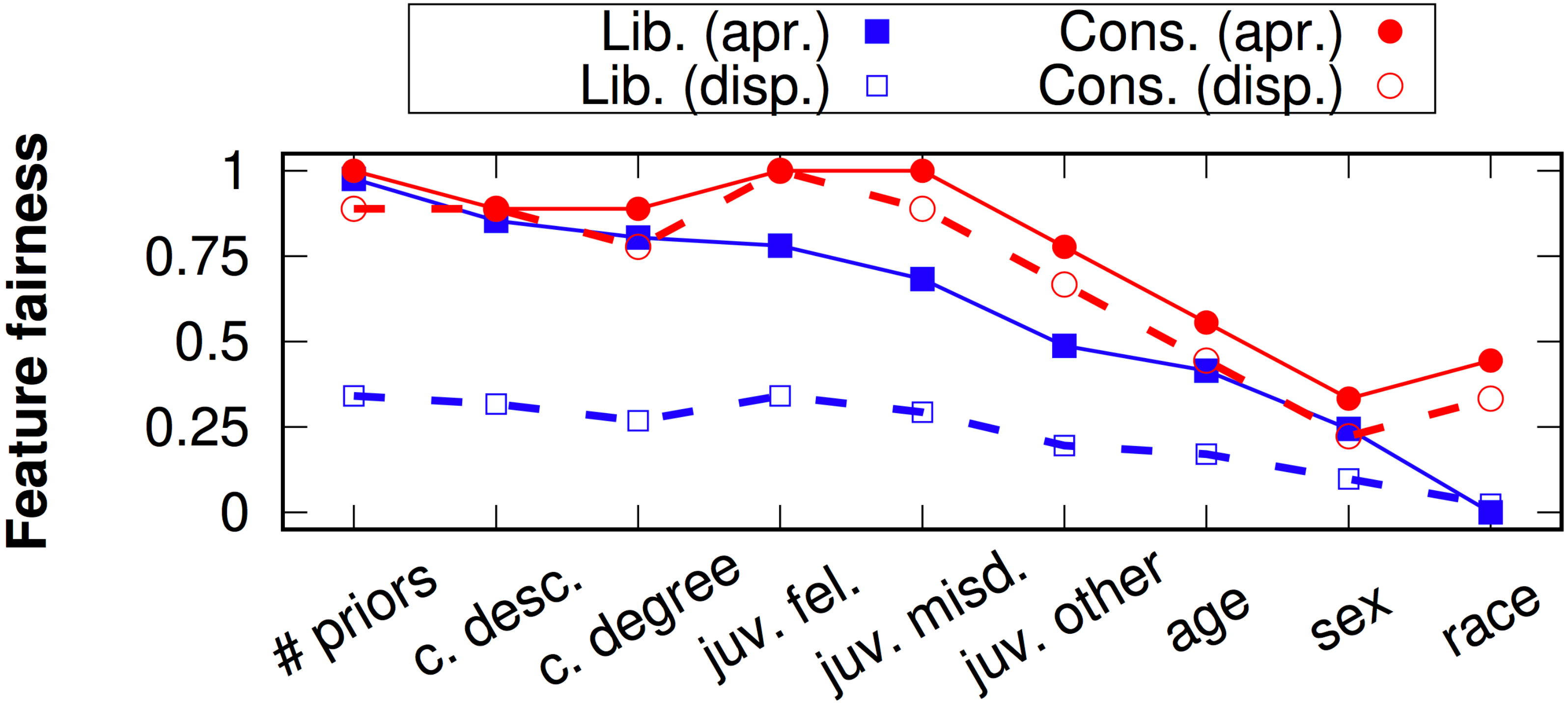
Revisiting recidivism risk prediction

- Is it fair to use the data in COMPAS questionnaire?

Feature	Q. 1 (a priori)	Q. 2 (if more accurate)	Q. 3 (if increases disparity)
# prior offenses	95%	93%	83%
arrest charge description	86%	92%	71%
arrest charge degree	85%	91%	69%
# juvenile felony offenses	74%	80%	61%
# juvenile misdemeanor offenses	65%	71%	53%
# juvenile other offenses	63%	69%	52%
age	44%	61%	32%
gender	26%	55%	24%
race	21%	42%	17%

- What do our human moral instincts tell us?

Do we share same moral instincts?

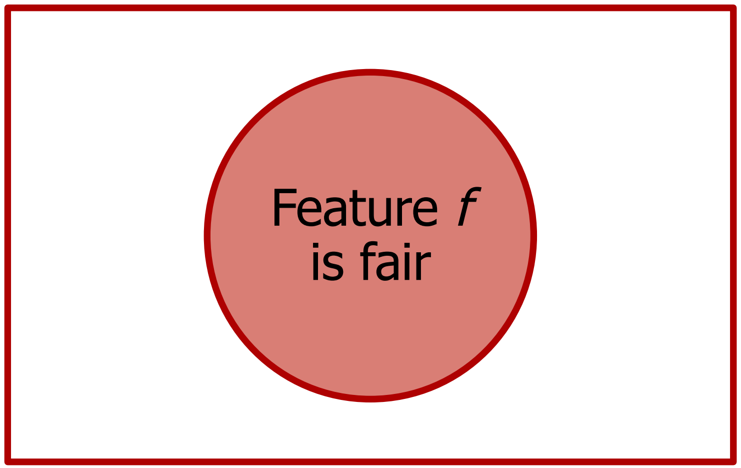


- There exist interesting differences
- But, our instincts are more similar than different

How to account for these instincts in fair decision making?

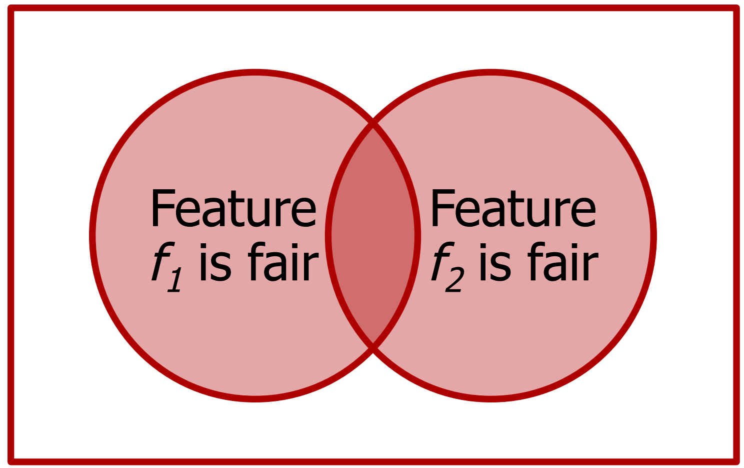
Individual feature

- We have fraction of people that consider using any single feature fair



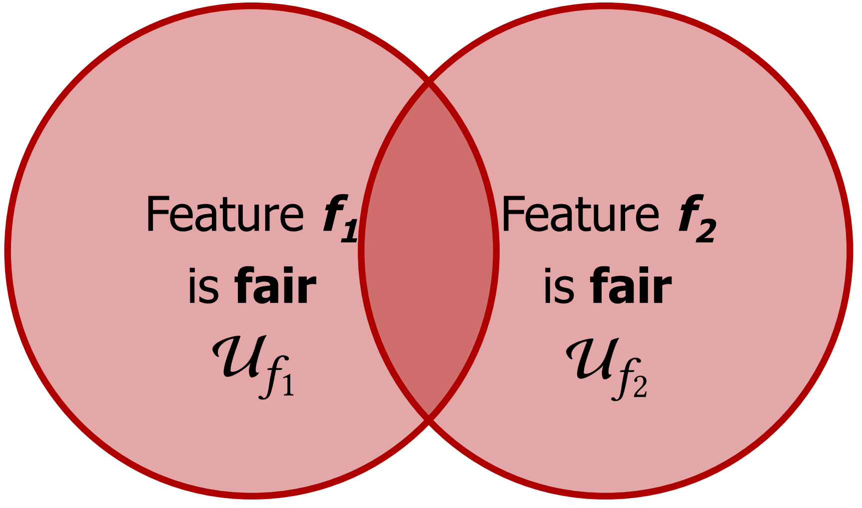
Set of features

- We can compute fraction of people that consider using all of the feature set fair



Measuring (un)fairness of classifier

- Fairness: Fraction of people that consider all features used fair

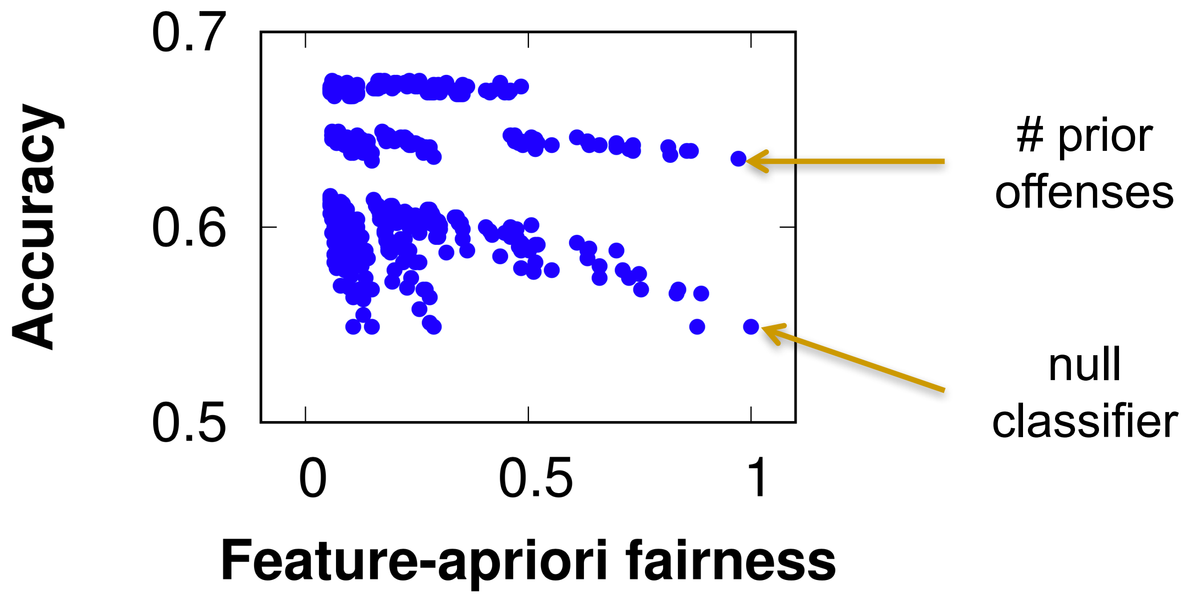


$$\frac{|\bigcap_{f \in \mathcal{F}} \mathcal{U}_f|}{|\mathcal{U}|}$$

- Unfairness: 1 - Fairness

Learning fair classifiers

- Tradeoffs between fairness and prediction accuracy

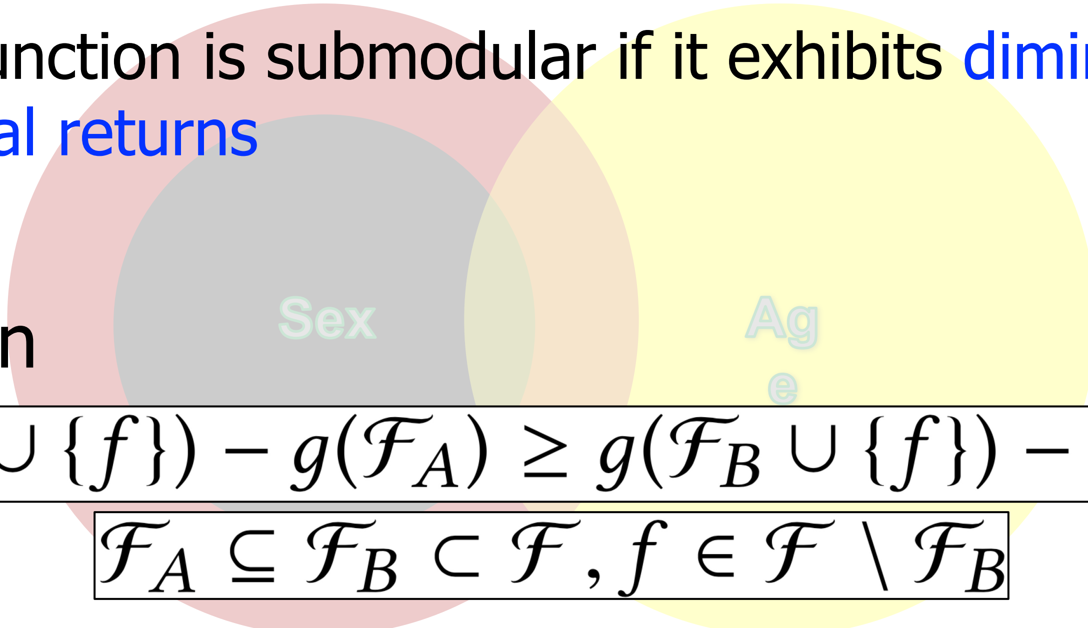


- How to compute the tradeoff efficiently?
  - Not scalable – 2^n classifiers, n = number of features
- Can we do better?

Properties of our unfairness measure

- Unfairness measure is submodular

- Intuition
  - A set function is submodular if it exhibits diminishing marginal returns



- Definition

$$\boxed{g(\mathcal{F}_A \cup \{f\}) - g(\mathcal{F}_A) \geq g(\mathcal{F}_B \cup \{f\}) - g(\mathcal{F}_B),}$$

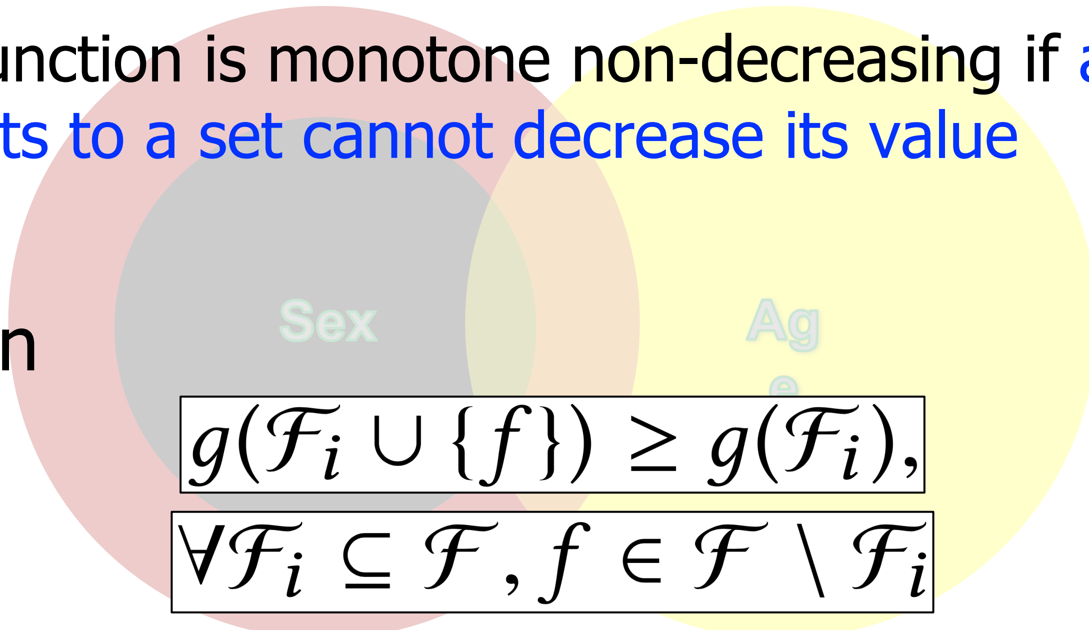
$$\boxed{\mathcal{F}_A \subseteq \mathcal{F}_B \subset \mathcal{F}, f \in \mathcal{F} \setminus \mathcal{F}_B}$$

Properties of our unfairness measure

- Feature unfairness is monotone non-decreasing

- Intuition
  - A set function is monotone non-decreasing if adding elements to a set cannot decrease its value

- Definition



$$\boxed{g(\mathcal{F}_i \cup \{f\}) \geq g(\mathcal{F}_i),}$$
$$\boxed{\forall \mathcal{F}_i \subseteq \mathcal{F}, f \in \mathcal{F} \setminus \mathcal{F}_i}$$

Formulating fair classifier learning

- Maximize accuracy subject to a constraint on unfairness

- Unfairness budget (threshold) ***t***
  - Majority of users must consider all features fair → $t = 1/2$
  - Supermajority → $t \in \{1/3, 2/5, 1/4\}$

- SCSK: Submodular Cost Submodular Knapsack

$$\begin{aligned} \underset{\mathcal{S} \subseteq \mathcal{F}}{\text{maximize}} \quad & accuracy(\mathcal{S}) \\ \text{subject to} \quad & unfairness(\mathcal{S}) \leq t \end{aligned}$$

Alternate formulation

- Minimize unfairness subject to a constraint on accuracy

- Accuracy threshold ***t***
  - ***t*** ≥ 0.9
  - ***t*** ≥ 0.9 * accuracy of classifier that uses all features

- SCSC: Submodular Cost Submodular Cover

$$\begin{aligned} \underset{\mathcal{S} \subseteq \mathcal{F}}{\text{minimize}} \quad & \mathit{unfairness}(\mathcal{S}) \\ \text{subject to} \quad & \mathit{accuracy}(\mathcal{S}) \geq t \end{aligned}$$

Why we do social computing

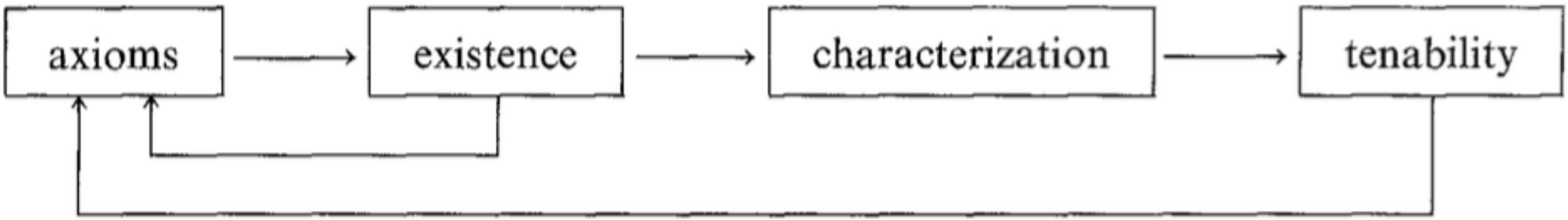
1. ~~For play or fun~~
2. ~~For seeking truth~~
3. ~~Pursuing beauty~~
4. Fighting for justice
5. Love for other human beings

4. Fighting for Justice

Deep diving into our moral instincts

How to reason about fairness?

- Reflective Equilibrium
  - From A Theory of Justice by John Rawls



- Why do we feel some features are fair & not others?
- What are our axioms for our moral instincts?

Revisiting COMPAS features

Feature
prior offenses
arrest charge description
arrest charge degree
juvenile felony offenses
juvenile misdemeanor offenses
juvenile other offenses
age
gender
race

- Norms
  - Socio-political-legal-economic-privacy norms

1. Volitional or Non-Volitional?
2. Reliably assessed?
3. Relevant?
   - Probity vs. Prejudicial value?
4. Causality?
   - Reverse causality?

Axioms for morality of feature use

- Hypothesis:
  - Volitionality, Reliability, Relevancy, Causality matter
- Conducting a large-scale user survey-based study
  - To check for opinion consensus on these axioms
  - Hopefully, we will find pattern maps between axioms and moral instincts

Why we do social computing

1. ~~For play or fun~~
2. ~~For seeking truth~~
3. ~~Pursuing beauty~~
4. ~~Fighting for justice~~
5. Love for other human beings

5. Love for other human beings

Human-Computer Symbiosis
BEYOND Human-Computer Interfaces
BEYOND Human vs. Computers!

Is fair decision making a ML problem?

- The travails of an Indian Ola cab driver

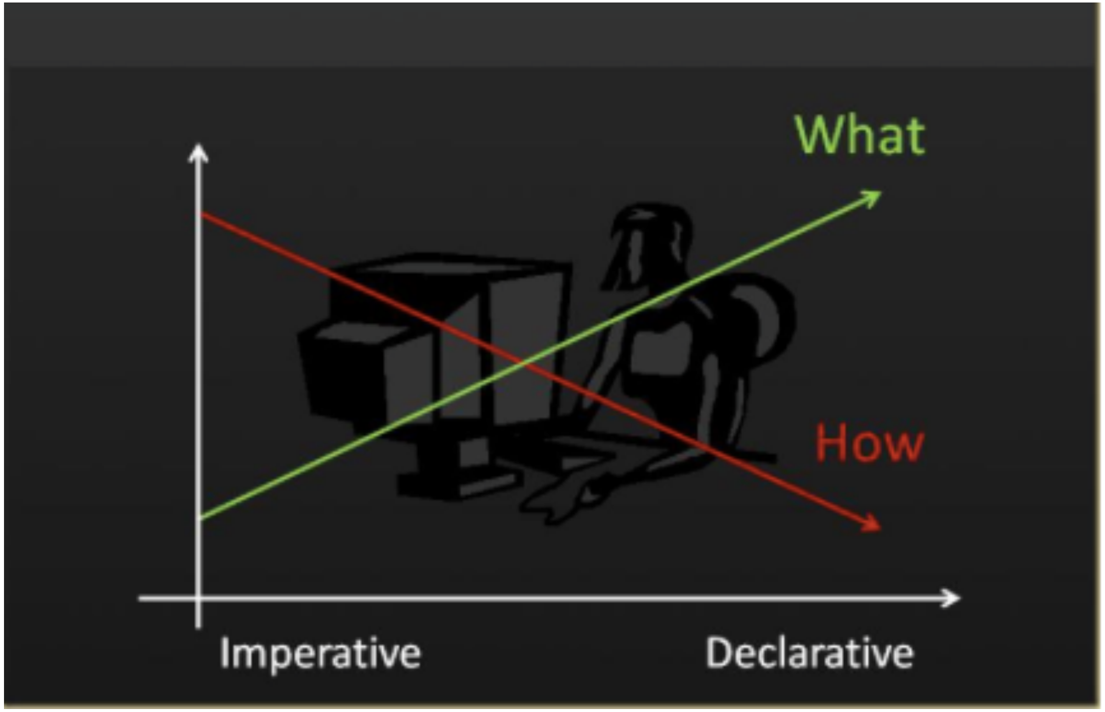
- What he needs is informational justice

- Informational justice is a distributed systems problem
  - How do we build accountable systems?
  - It has information security / privacy aspects too!

Is fair decision making a CS problem?

- How we engineer (program) computers
  - Imperative programming:
    - You describe the procedure for making decisions
      - Not what you want from the decisions
  - Declarative programming:
    - You declare the outcome goals of your decision making
      - Not how you want to make decisions
    - Leveraging machines to find optimal decision procedure

Imperative vs. Declarative Programming



- In practice, any complex software system has elements of both

The excitement about AI/ML

- Can get away with lazy declarative engineering
  - Get some training data – examples of past decisions
  - Declare a default goal – decision prediction accuracy

- Miraculously, lazy engineering appears to work!
  - But, does it really work?

The achilles heels of lazy AI/ML

Even assuming no training data biases, AI/ML decisions

1. Optimize for a single decision outcome goal, ignoring
   - Fairness: Equal prediction accuracy for all salient social groups
   - Worst-cases: Lower bound worst-case prediction accuracy
   - Norms: Should use or not use data in a specific manner

2. Optimal for a static NOT an evolving society, because
   - Training data becomes unrepresentative
   - Feedback loops are not accounted for in the first place
   - Decision outcome goals change over time!

Can we guard the achilles heels?

- Can we account for fairness & other norms in ML decision making?
  - Maybe! Even with declarative engineering
    - Declare multiple decision outcome objectives when training
    - Lots of ongoing research on specifying the objectives to machines

- Can we design ML decision making for an evolving society?
  - Not sure! Need more imperative / procedural engineering

How we engineer human decisions

- In democratic societies, we mix imperative and declarative decision making in ingenious ways

- A continuum from procedural to declarative & back
  - Constitution defines procedures for making laws
    - Laws interpret constitution to reach desired outcomes
  - Laws define procedures for making executive rules
    - Executive rules interpret laws to reach desired outcomes
  - Further, decision validity can be challenged in courts

- Designed to work for an evolving society!

Human-Computer Symbiotic Decision Making

- Rather than human vs. computer decision making
- We use understandable human decision procedures to define objectives for decision outcomes
  - A computational perspective is needed for understanding
- We trust machines to find optimal decision procedures that achieve the well-defined objectives

Why we do social computing

1. ~~For play or fun~~

2. ~~For seeking truth~~

3. ~~Pursuing beauty~~

4. ~~Fighting for justice~~

5. ~~Love for other human beings~~

What is social computing?

- Hopefully, you have a feel

- What do social computists do?
  - Look via a computational lens at the world around them
  - Try to find problems that allow them to flourish
  - They leverage optimization, databases, networking, systems, ML, IR, NLP, and social sciences (economics, political science, psychology, and law)