MEASURING AND REDUCING DISPARITIES

SURESH VENKATASUBRAMANIAN
UNIVERSITY OF UTAH

DISPARITY IN REPRESENTATION

"Harms of allocation"

How the decisions made by a system affect different groups differently

Demographic parity
Equal odds
Disparate impact

VS

Kate Crawford, NeurIPS 2017

"Harms of representation"

How the way we **represent** groups might cause harm

Stereotyping
Ex-nomination
Under-representation
Denigration

How can we identify and measure disparities in representation?

DISPARITY IN REPRESENTATION VIA GEOMETRY

Can we identify bias in existing (learned) representations and correct it?

DISPARITY IN REPRESENTATION VIA GEOMETRY

Can we identify bias in existing (learned) representations and correct it?

man

doctor

woman

nurse

DISPARITY IN REPRESENTATION VIA GEOMETRY

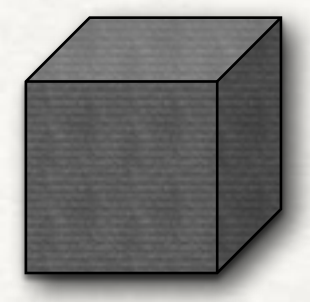
Can we identify bias in existing (learned) representations and correct it?

man

doctor

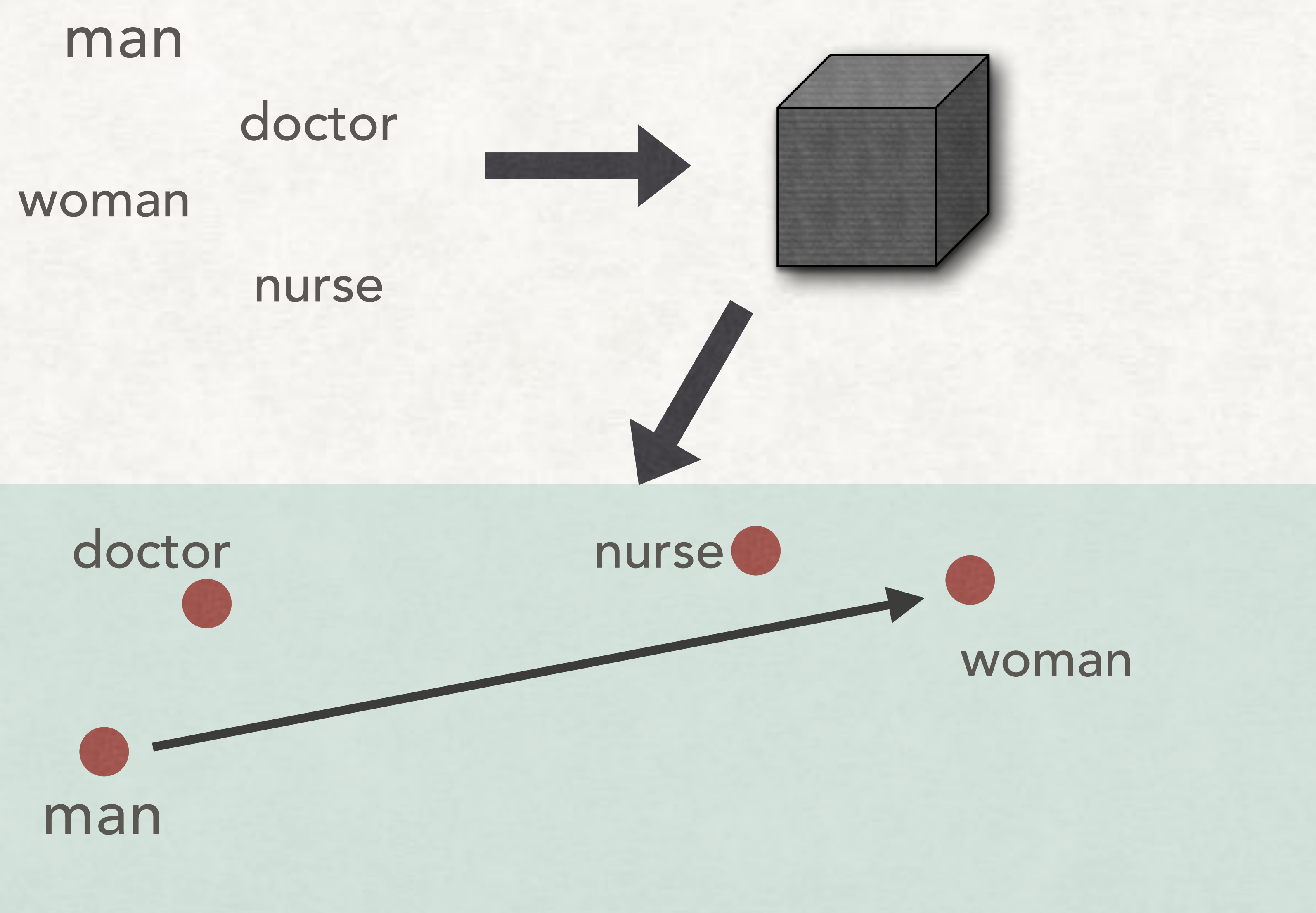
woman

nurse



DISPARITY IN REPRESENTATION VIA GEOMETRY

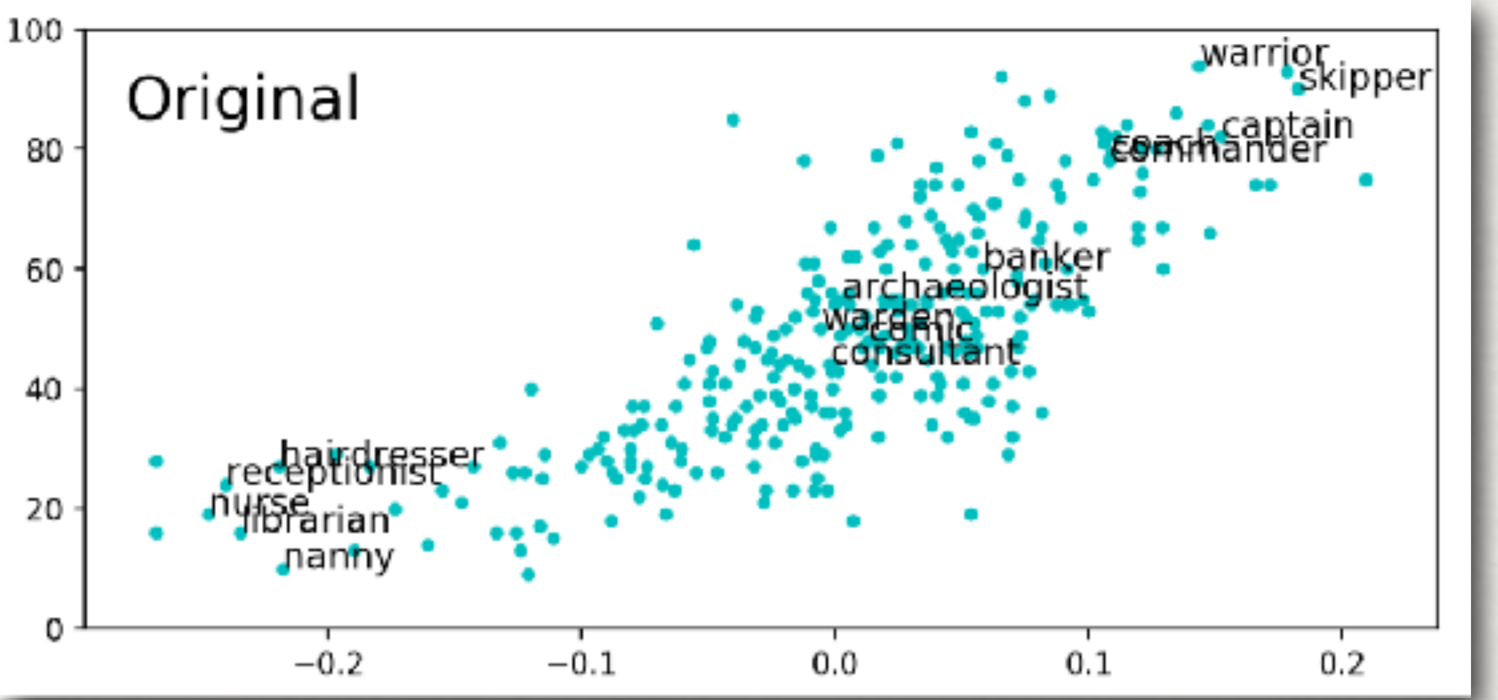
Can we identify bias in existing (learned) representations and correct it?

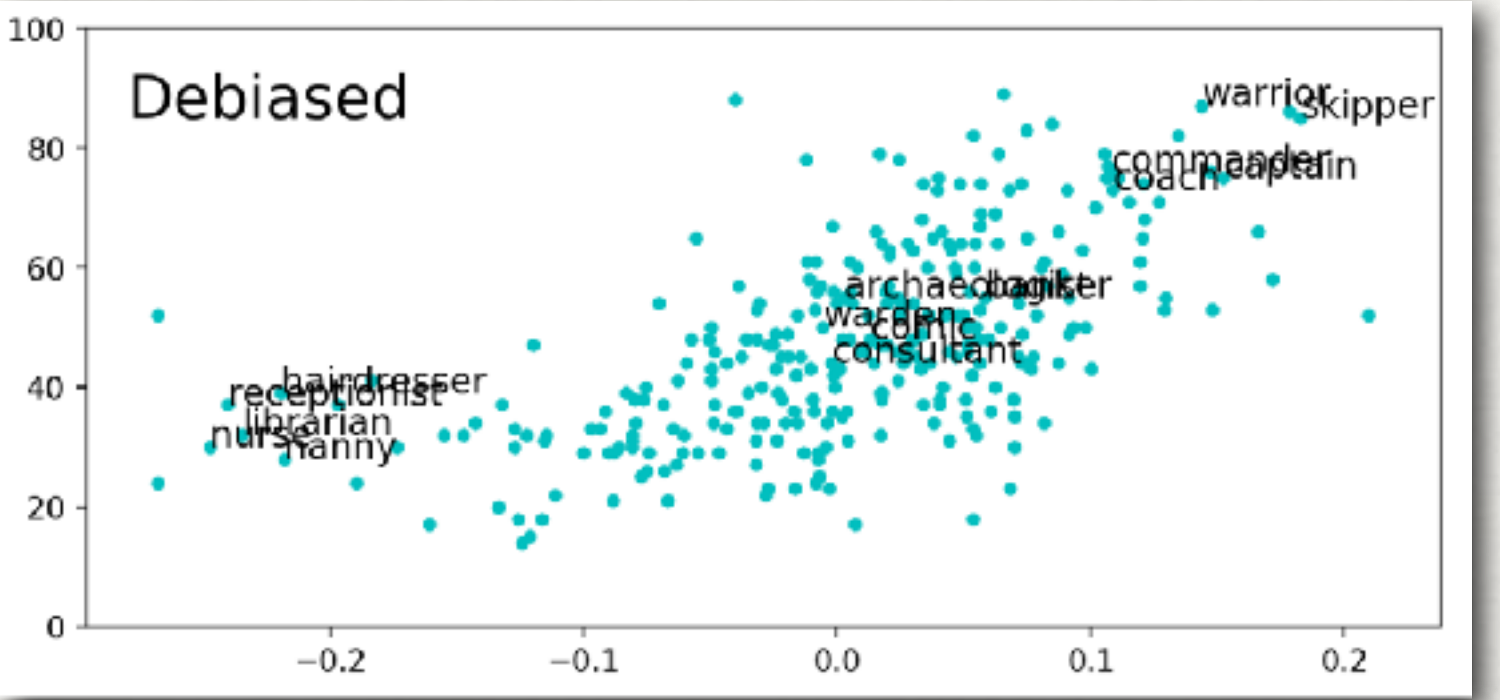


PROBLEM WITH GEOMETRY AS BIAS

Lipstick on a Pig: Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them

- Gonen and Goldberg, ACL2019





STEREOTYPING

"associations and beliefs about the characteristics and attributes of a group and its members that shape how people think about and respond to the group"

— SAGE handbook of prejudice, stereotyping and discrimination.

A specific **mechanism** for stereotyping:

…the tendency to assign characteristics to all members of a group based on stereotypical features shared by a few…

REPRESENTATIVENESS [Bordalo et al 16]

"distorted perception of the relative frequency of a type in the stereotyped group compared to the base group"

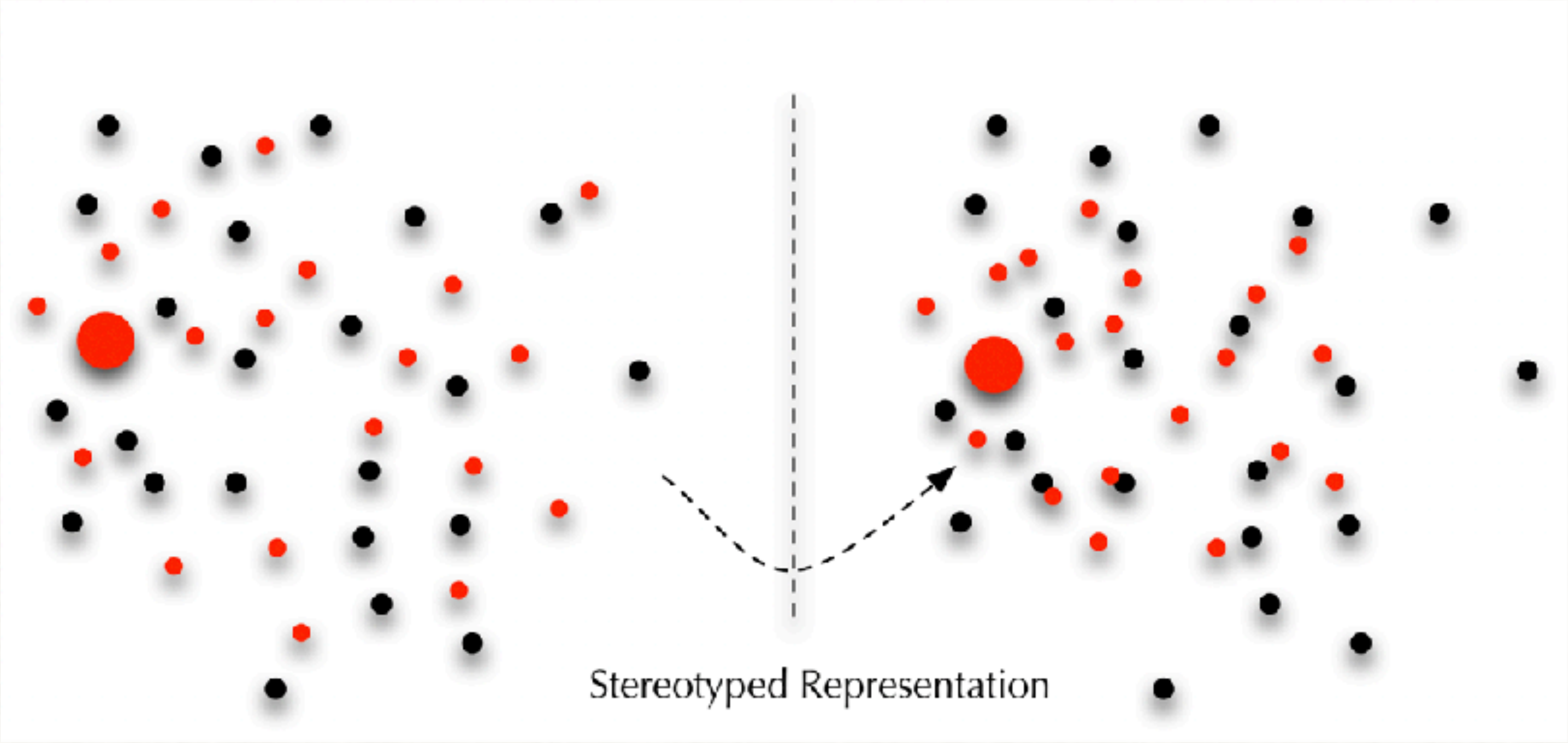
$$R(t, G) = \frac{\Pr(t \mid G)}{\Pr(t \mid \overline{G})}$$ "representativeness"

$$\Pr'(t \mid G) \propto \Pr(t \mid G) R(t, G)^{\rho}$$

ρ is the **degree** of stereotyping

EXEMPLARS

We can model stereotyping as a process by which groups experience non-uniform variance reduction [AFSV19]



Points regress towards an exemplar

$$\mathbf{p}_\alpha = (1 - \alpha)\mathbf{p} + \alpha\mathbf{c}$$

EFFECTS OF STEREOTYPING

- Linear regression (via exemplar stereotyping):

$$y = X\beta + \epsilon$$
$$\beta = (X^\top X)^{-1} X^\top y$$

  - If we set $X' \leftarrow X + \Delta_\alpha$, then β' depends **quadratically** on α

- Can show similar behavior for other tasks (clustering, classification)

Distortion in representation has measurable effect on decision.

BROADER GOALS

DISPARITY IN ACCESS

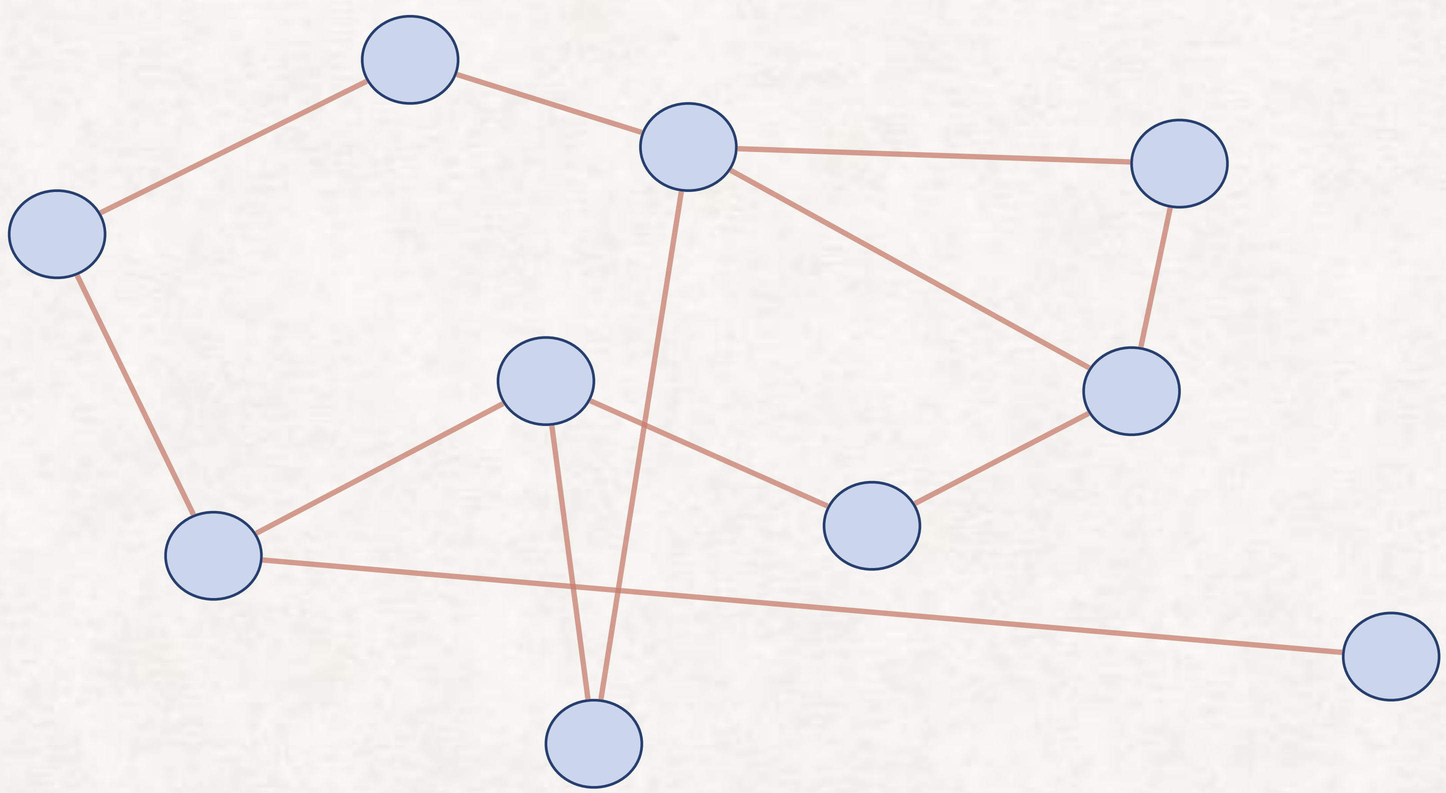
- Social standing within a network confers utility on an individual.
- Social "position" in a network is a class marker defined by the *network*, not the individual. [Coleman]
- Should we be considered about discrimination based on social position? [boyd, Marwick and Levy]

INFORMATION ACCESS

- Social networks grow through recommendations as well as organically
- Network position confers advantage ([Granovetter])
- Access to information that improves network position relies on …. network position
- “edges in social network” == “biased input data”

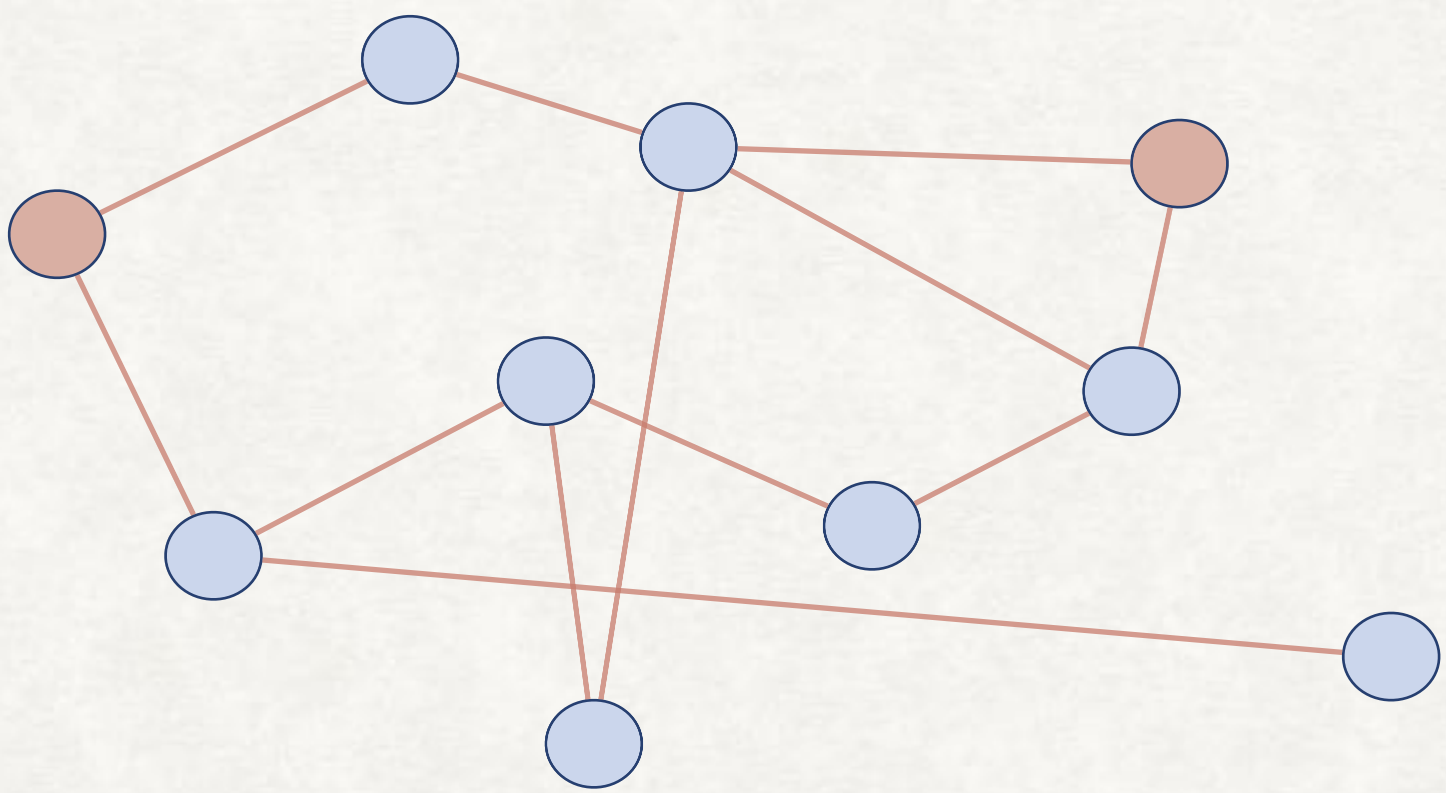
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and k seeds, how many nodes can be be *influenced?*



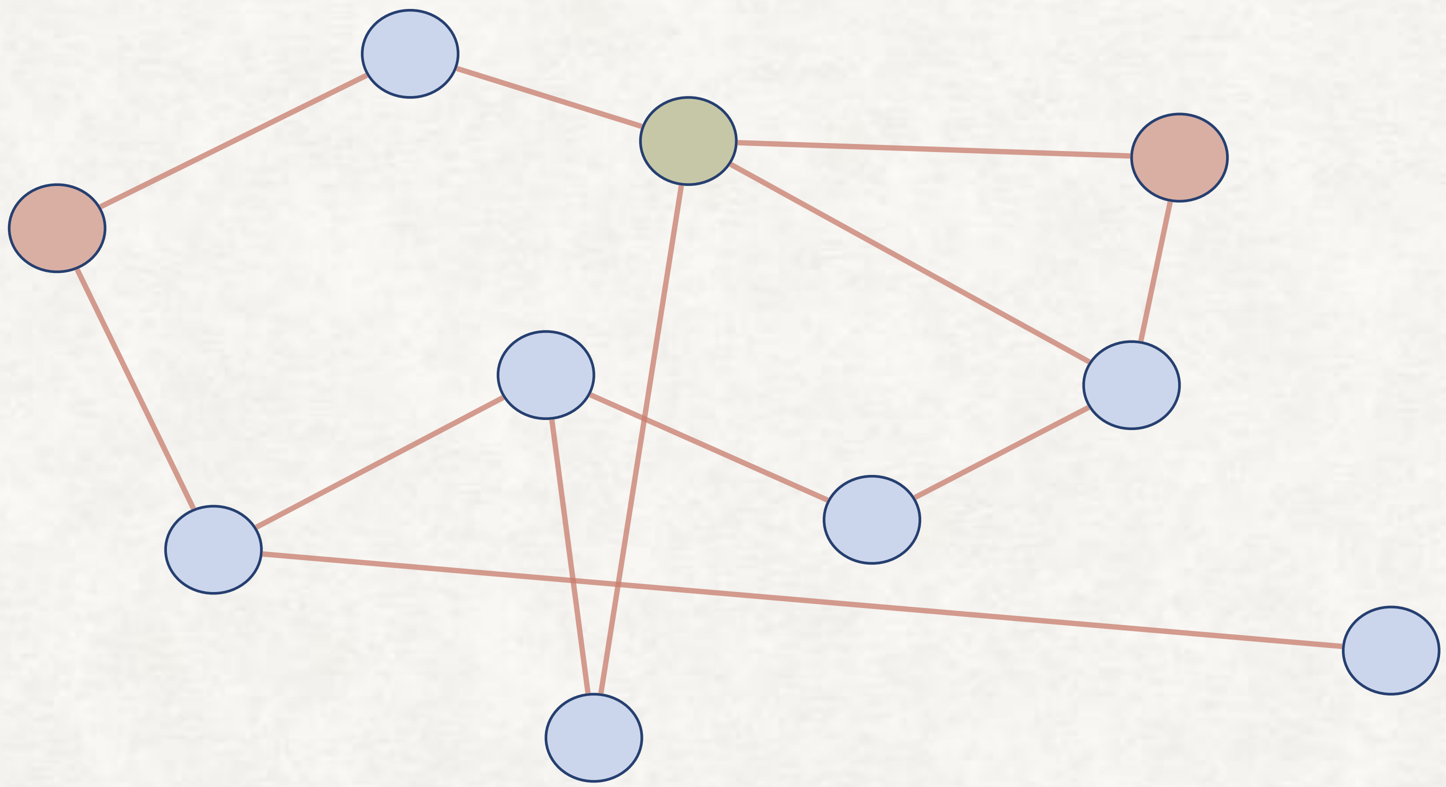
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and k seeds, how many nodes can be be *influenced?*



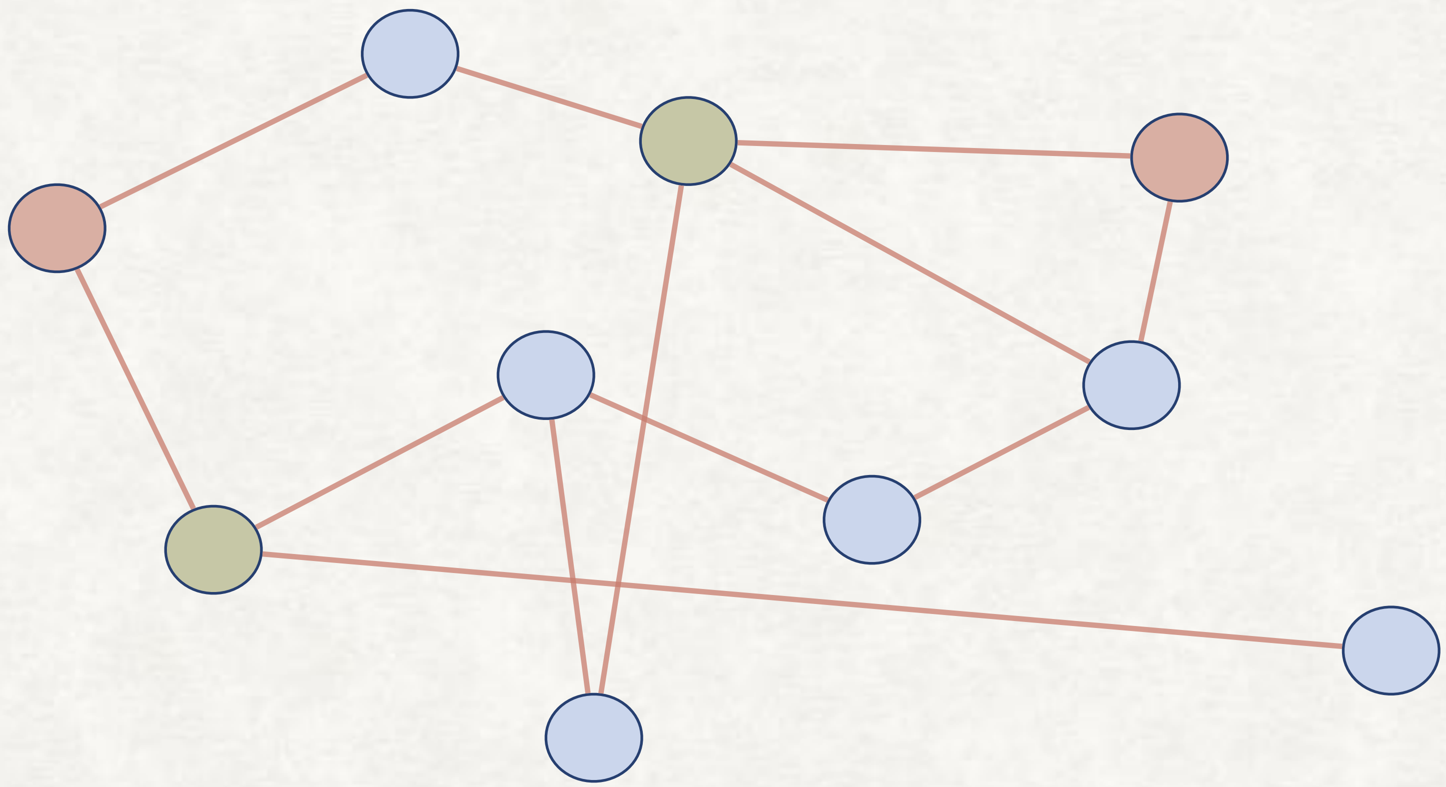
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and k seeds, how many nodes can be be *influenced?*



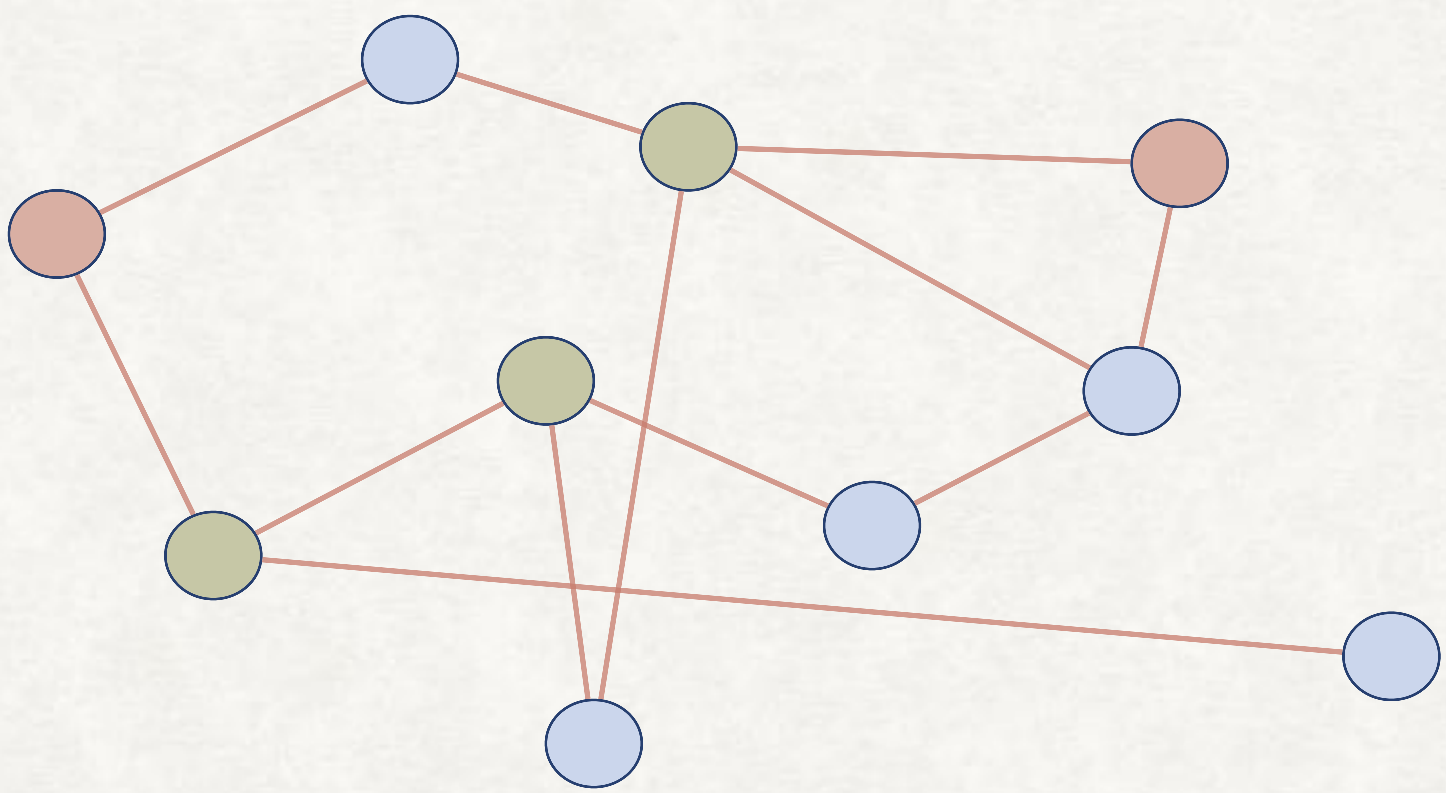
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and k seeds, how many nodes can be be *influenced?*



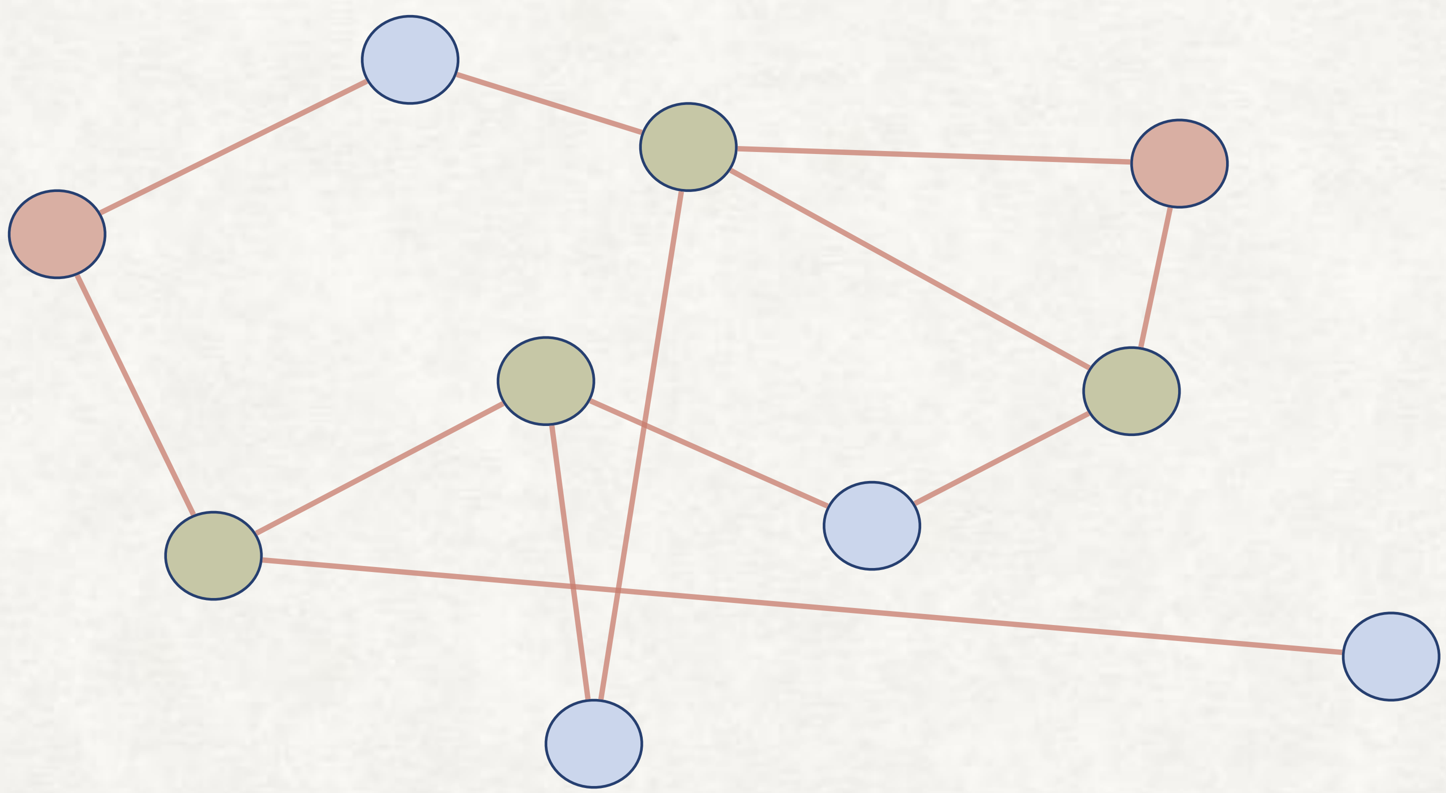
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and *k* seeds, how many nodes can be be *influenced?*



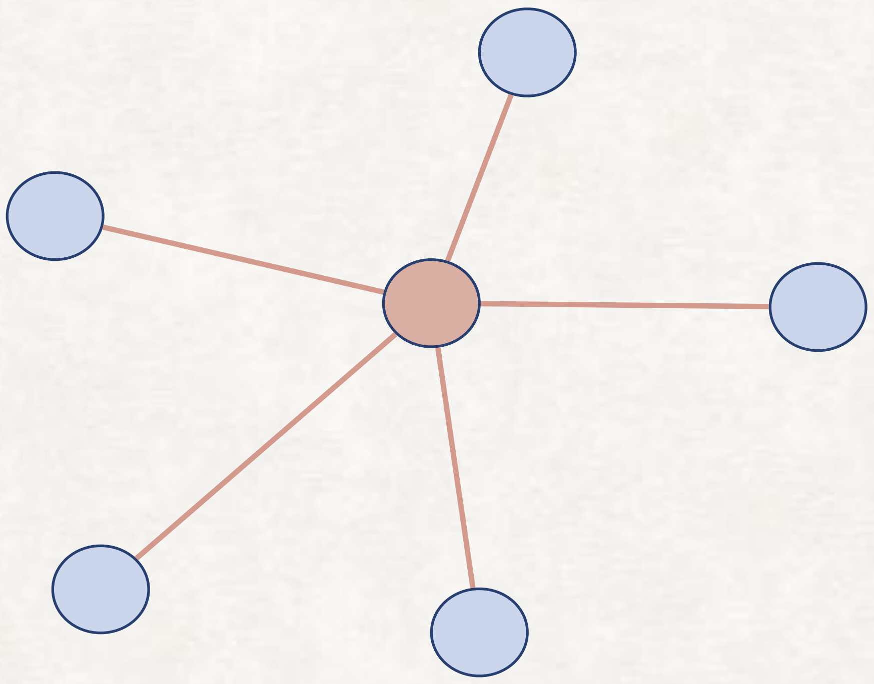
INFLUENCE MAXIMIZATION

Given a graph, a mechanism for spreading information and k seeds, how many nodes can be be *influenced?*



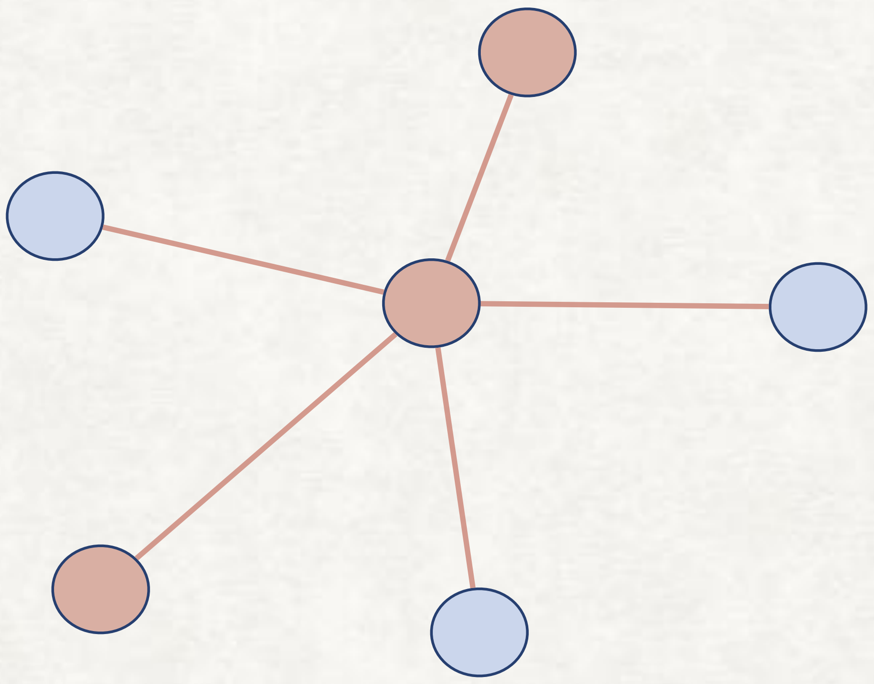
INFORMATION CASCADES

An *active* node sends information on each edge with probability α and then becomes *inactive.*



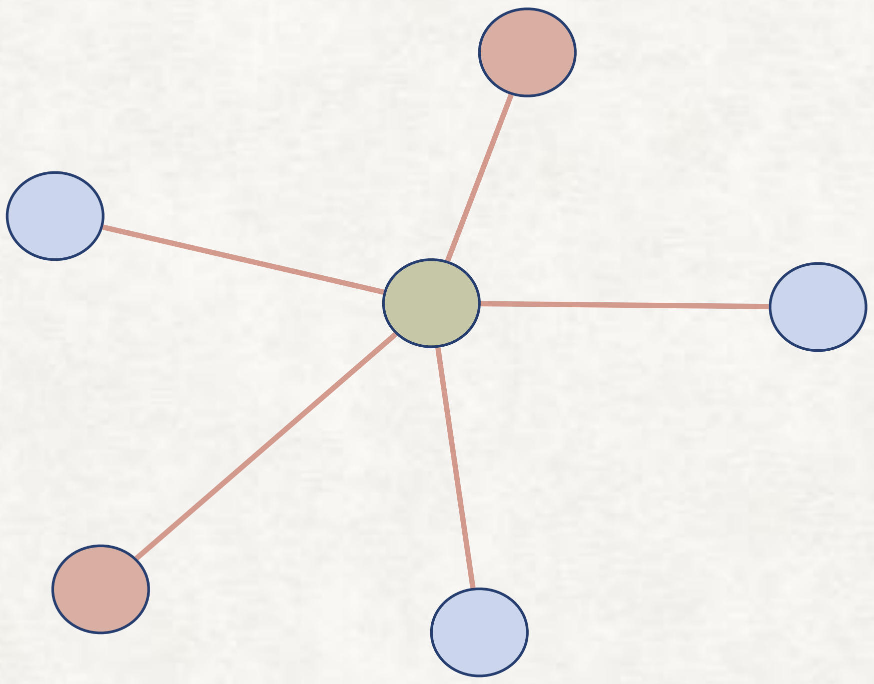
INFORMATION CASCADES

An *active* node sends information on each edge with probability α and then becomes *inactive.*



INFORMATION CASCADES

An *active* node sends information on each edge with probability α and then becomes *inactive.*



WELFARE FUNCTIONS

$p_v =$ Probability that v gets the information.

Welfare function $\mu : [0,1]^n \to \mathbb{R}$

Welfare of vertex set $V = \{v_1, v_2, \ldots, v_s\}$ in graph G with seed set S:

$$\mu_G(S, V) = \mu(v_1, v_2, \ldots, v_s)$$

ϕ-mean

$$\mu_\phi = (1/n \sum_v p_v^\phi)^{1/\phi}$$

$$\mu_r = 1/n \sum_v p_v$$

$$\mu_{-\infty} = \min_v p_v$$

Find $S^+, |S^+| = k$ so that $\mu_G(S^+ \cup S)$ is maximized.

ACCESS GAPS

Definition: *Access gap* of a partition *V, V′* of *G* with seed set *S* is

$$\mu(S, V) - \mu(S, V')$$

Definition: In a graph *G*, the *rich get richer* if there is a partition (*V, V′*) such that the optimal intervention *S** satisfies

$$\mu(S^*, V') - \mu(S^*, V) > \mu(S, V') - \mu(S, V) > 0.$$

The access gap increases after intervention

THE RICH ALWAYS GET RICHER

Proposition

Suppose μ is symmetric, increasing, and for any $x_1, \ldots, x_m$ in $[0, 1]$, there is some $1 \leq \phi < \infty$ such that

$$\min_i x_i \leq \mu(x_1, \ldots, x_m) \leq \left(\frac{1}{m} \sum_{i=1}^{m} x_i^{\phi} \right)^{1/\phi} .$$

Then if $0 < \alpha < \frac{1}{2\phi}$, $\exists\, G, S$ where the rich get richer.

THE RICH ALWAYS GET RICHER

Proposition

Suppose μ is symmetric, increasing, and for any $x_1, \ldots, x_m$ in $[0,1]$, there is some $1 \leq \phi < \infty$ such that

$$\min_i x_i \leq \mu(x_1 \ldots) \leq \left(1 \; m \ldots \right)^{1/\phi}$$

Then if $0 < \alpha < \frac{1}{2\phi}$, $\exists\, G, S$ where th

UNDER ANY WELFARE FUNCTION, THE RICH GET RICHER

A RELAXED CRITERION: BALANCE

μ is k-imbalanced if $\exists G$, seed set S, partition V, V' and optimal interventions S^* and S_V (for V only) with no more than k seeds such that

1. $\mu(S, V) \leq \mu(S_V, V)$ (there's a way to improve welfare of V)
2. $\mu(S_V, V) \leq \mu(S, V')$ (...but V''s welfare is then better)
3. $\mu(S, V') \leq \mu(S^*, V')$ (V''s welfare can be improved by the global optimum)
4. $\mu(S^*, V) \leq \mu(S, V)$ (...but V's welfare does not improve)

A RELAXED CRITERION: BALANCE

μ is k-imbalanced if $\exists G$, seed set S, partition V, V' and optimal interventions S^* and S_V (for V only) with no more than k seeds such that

1. $\mu(S, V) \leq \mu(S_V, V)$ (there's a way to improve welfare of V)
2. $\mu(S_V, V) \leq \mu(S, V')$ (...but V''s welfare is then better)
3. $\mu(S, V') \leq \mu(S^*, V')$ (V''s welfare can be improved by the global optimum)
4. $\mu(S^*, V) \leq \mu(S, V)$ (...but V's welfare does not improve)

Minimax welfare $\mu_{-\infty}$ is balanced (improve minimum access)

A RELAXED CRITERION: BALANCE

μ is k-imbalanced if $\exists G$, seed set S, partition V, V' and optimal interventions S^* and S_V (for V only) with no more than k seeds such that

1. $\mu(S, V) \leq \mu(S_V, V)$ (there's a way to improve welfare of V)
2. $\mu(S_V, V) \leq \mu(S, V')$ (...but V''s welfare is then better)
3. $\mu(S, V') \leq \mu(S^*, V')$ (V''s welfare can be improved by the global optimum)
4. $\mu(S^*, V) \leq \mu(S, V)$ (...but V's welfare does not improve)

Minimax welfare $\mu_{-\infty}$ is balanced (improve minimum access)

Let $\phi > -\infty, \alpha < 1$

Then μ_ϕ is $\Omega(n)$ -imbalanced

A RELAXED CRITERION: BALANCE

μ is k-imbalanced if $\exists G$, seed set S, partition V, V' and optimal interventions S^* and S_V (for V only) with no more than k seeds s

1. $\mu(S, V) \leq \mu(S_V, V)$ (there's a way to i
2. $\mu(S_V, V) \leq \mu(S, V')$ (...but V''s wel
3. $\mu(S, V') \leq \mu(S^*, V')$ (V''s welfare ca pti- mum)
4. $\mu(S^*, V) \leq \mu(S, V)$ (...but V's welfare not improve)

AND NO OTHER WELFARE FUNCTION IS BALANCED

Minimax welfare $\mu_{-\infty}$ is balanced (improve minimum access)

Let $\phi > -\infty, \alpha < 1$

Then μ_ϕ is $\Omega(n)$ -imbalanced

ON CISE/SBE COLLABORATIONS

- **Questioning the frame**
  - CS does not naturally take a critical stance towards its problems
  - But we are seeing more of it now

What are the **traps** we encounter if we ignore the **social** in a sociotechnical system [SbFVV19]

What are effective ways to harness the power of computer science for social change? [ABKLRR20]

ON CISE/SBE COLLABORATIONS

- **Questioning the frame**
  - CS does not naturally take a critical stance towards its problems
  - But we are seeing more of it now

Fairness and Abstraction in Sociotechnical Systems

What are the **traps** we encounter if we ignore the **social** in a sociotechnical system [SbFVV19]

What are effective ways to harness the power of computer science for social change? [ABKLRR20]

ON CISE/SBE COLLABORATIONS

- **Questioning the frame**
  - CS does not naturally take a critical stance towards its problems
  - But we are seeing more of it now

Fairness and Abstraction in Sociotechnical Systems

What are the **traps** we encounter if we ignore the **social** in a sociotechnical system [SbFVV19]

Roles for computing in social change

What are effective ways to harness the power of computer science for social change? [ABKLRR20]

ON CIS/SBE COLLABORATIONS

- CS is not merely a "have data, will compute" field.
- The computational lens allows for precision, and therefore allows us to articulate limits.

The (im)possibility of fairness: different value systems require different mechanisms for fair decision making*

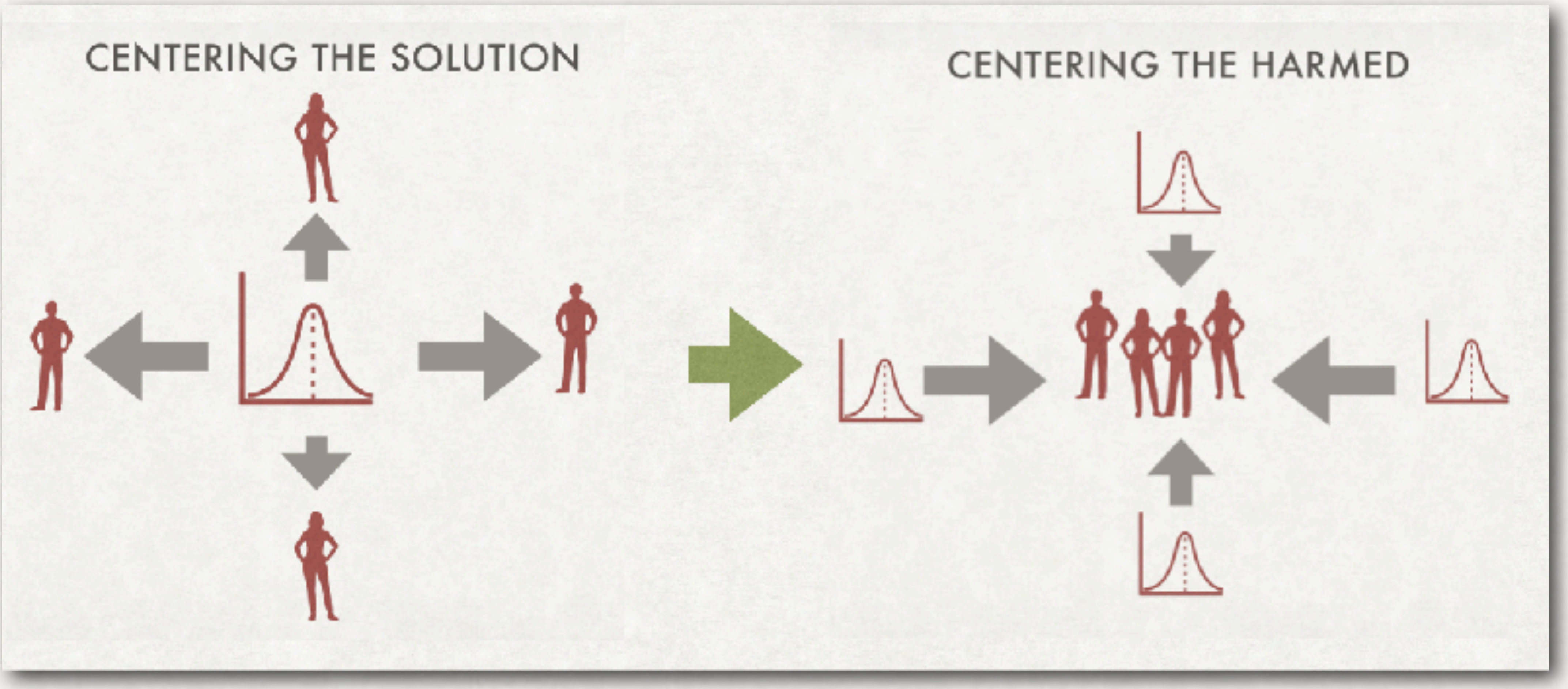
[FSV16]

Inherent Trade-Offs in the Fair Determination of Risk Scores

[KMR16]

ON CIS/SBE COLLABORATIONS

- The "technology is problematic" frame assumes a fixed formulation of technical questions. We can change that!



ON CISE/SBE COLLABORATIONS