University of Tsukuba,
LGS'09, 26 - 29 August 2009

White Manipulation in Judgment Aggregation

Davide Grossi
ILLC Amsterdam

UNIVERSITÉ DU LUXEMBOURG

Gabriella Pigozzi
Marija Slavkovik

What is this all about

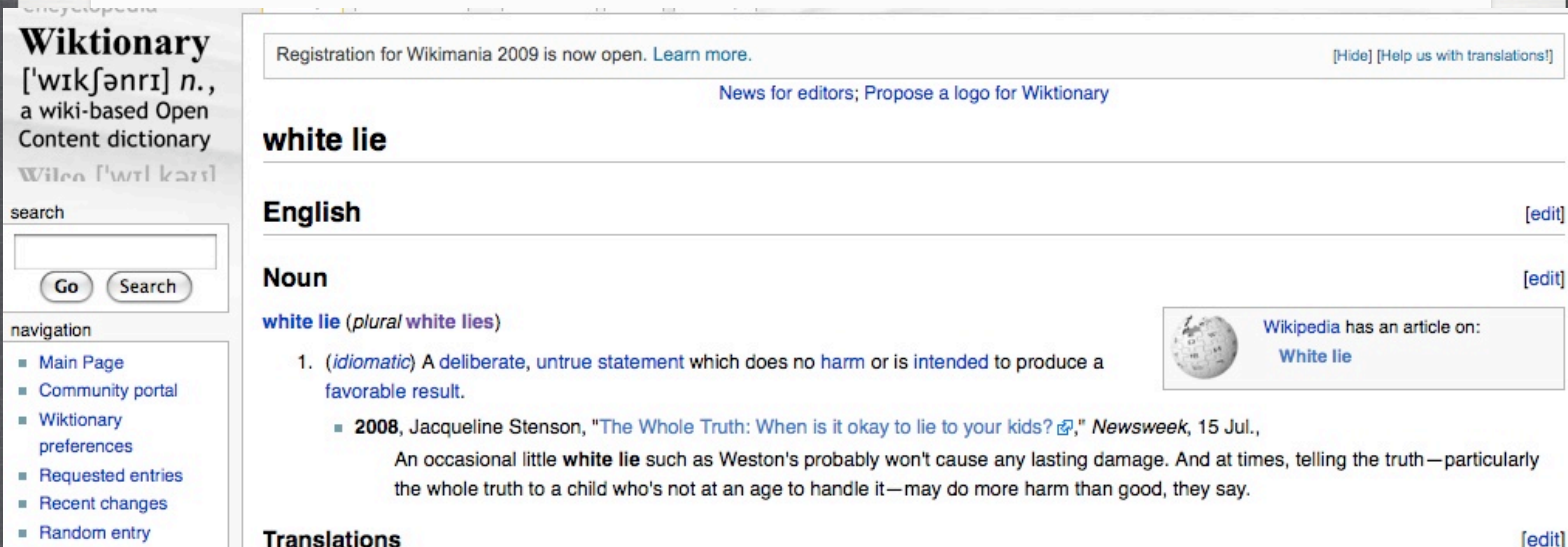
- judgment aggregation (JA) has two problems:
  - aggregation functions that satisfy a desirable set of properties do not exist
  - aggregation operators that exist are manipulable
- the question is: is lying, cheating and manipulation really that bad ?

White Manipulability

- the colloquial term “white lies”

White Manipulability

- the colloquial term “white lies”



White Manipulability

- the colloquial term “white lies”

White Manipulability

- the colloquial term “white lies”

- manipulation - lying with the intent to improve the outcome for the agent who lies

White Manipulability

- the colloquial term “white lies”

- manipulation - lying with the intent to improve the outcome for the agent who lies
- white manipulation - lying with the intent to improve the outcome for all the agents involved

In the rest of the talk

- introduce the basic concepts of judgment aggregation
- redefine the judgment aggregation function
- introduce in JA: scoring functions, social welfare notions
- define white manipulation
- initial results
- conclusions

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =X$ is good at teaching	$b = X$ is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a = X$ is good at teaching	$b = X$ is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a = X$ is good at teaching	$b = X$ is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

$$f : \text{profiles} \longrightarrow \text{judgment sets}$$

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

$$f : \text{profiles} \longrightarrow \text{judgment sets}$$

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

$$f : \text{profiles} \longrightarrow \text{judgment sets}$$

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $\boxed{x \leftrightarrow (a \wedge b)}$

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

$$f : \text{profiles} \longrightarrow \text{judgment sets}$$

- impasse

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- judgment aggregation functions are not manipulable if they satisfy independence and (weak) monotonicity[1]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- judgment aggregation functions are not manipulable if they satisfy independence and (weak) monotonicity[1]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- judgment aggregation functions are not manipulable if they satisfy independence and (weak) monotonicity[1]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a = X$ is good at teaching	$b = X$ is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- judgment aggregation functions are not manipulable if they satisfy independence and (weak) monotonicity[1]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the premise based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the premise based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the premise based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the premise based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the premise based procedure is manipulable [2]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the conclusion based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the conclusion based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the conclusion based procedure

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- the conclusion based procedure is manipulable [3]

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	a = X is good at teaching	b = X is good at research	x = hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- distance based merging

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- distance based merging

Judgment Aggregation

- how individual judgments on logically connected issues can be aggregated into a collective judgment on the same issues
- hiring committee example with rule $x \leftrightarrow (a \wedge b)$:

	$a =$ X is good at teaching	$b =$ X is good at research	$x =$ hire X
prof. A	*yes*	*no*	*no*
prof. B	*yes*	*yes*	*yes*
prof. C	*no*	*yes*	*no*
Majority	*yes*	*yes*	*no*

- distance based merging is manipulable [4]

The ideas are ...

- impasse $\downarrow$ is in the set of possible outputs of the aggregation function, but not part of any profile
- assume that agents have preferences over outputs and neither of the agents prefers the output $\downarrow$
- scoring functions determine preference ordering over the elements of the set of possible outputs of the aggregation function

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$
- example we work with - quota rule f^q
- score function we define as a function that, given a judgment set, scores all other possible outcomes based on that judgment set
- we work with examples of distance based scoring functions:

$$HS_e : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$
- example we work with - quota rule f^q
- score function we define as a function that, given a judgment set, scores all other possible outcomes based on that judgment set
- we work with examples of distance based scoring functions:

$$HS_e : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$
- example we work with - quota rule f^q
- score function we define as a function that, given a judgment set, scores all other possible outcomes based on that judgment set
- we work with examples of distance based scoring functions:

$$HS_e : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$
- example we work with - quota rule f^q
- score function we define as a function that, given a judgment set, scores all other possible outcomes based on that judgment set
- we work with examples of distance based scoring functions:

$$HS_e : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$
- example we work with - quota rule f^q
- score function we define as a function that, given a judgment set, scores all other possible outcomes based on that judgment set
- we work with examples of distance based scoring functions:

$$HS_e : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

JA function & scoring functions

- JA function we defined as $\quad f : \Omega \longrightarrow \Phi \cup \{\downarrow\}$

Definition. [Manipulation] Let s be a scoring function. An aggregation function f is manipulable if and only if there exists a judgment profile $\omega \in \Omega$ and an agent i such that $f(\omega) \prec_i^s f(\omega')$, where $\omega' \in \Omega$ is some i-variant of ω.

$$VS_k : \Phi \longrightarrow (\Phi^{\downarrow} \longrightarrow \mathbb{N})$$

Social welfare notions in JA

- using a scoring function, a preference profile can be built from a judgment profile
- having a preference profile, social welfare notions can be applied to JA
- Utilitarian social welfare

$$USW_{s(\omega)}(\varphi) = \sum_{i=1}^{n} s(\varphi)(\varphi_i)$$

- Egalitarian social welfare

$$ESW_{s(\omega)}(\varphi) = \mathtt{max}\{s(\varphi)(\varphi_i) \mid \varphi_i \in \Phi\}$$

Social welfare notions in JA

- using a scoring function, a preference profile can be built from a judgment profile

Social welfare notions in JA

- using a scoring function, a preference profile can be built from a judgment profile

Definition. [White manipulability] Let $\mathcal{SW}$ be a social welfare function and s a scoring function. An aggregation function f is white manipulable if and only if there exists an agent i and a judgment profile $\omega \in \Omega$ such that $f(\omega) \prec_i^s f(\omega')$ and $\mathcal{SW}(f(\omega)) < \mathcal{SW}(f(\omega'))$, where $\omega' \in \Omega$ is some i-variant of ω.

Hiring example revisited

Φ	(a,b,x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

Hiring example revisited

Φ	(a,b,x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

20 possible profiles

Hiring example revisited

Φ	(a,b,x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

Hiring example revisited

Φ	(a,b,x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

$$(1,0,0) \succ_2^{HS_e} (0,0,0) \succ_2^{HS_e} (0,1,0) \sim_2^{HS_e} (1,1,1) \succ_2^{HS_e} \downarrow$$
$$(1,1,1) \succ_1^{HS_e} (1,0,0) \sim_1^{HS_e} (0,1,0) \succ_1^{HS_e} (0,0,0) \succ_1^{HS_e} \downarrow$$
$$(0,1,0) \succ_3^{HS_e} (0,0,0) \succ_3^{HS_e} (1,0,0) \sim_3^{HS_e} (1,1,1) \succ_3^{HS_e} \downarrow$$

Hiring example revisited

Φ	(a, b, x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

$$(1,0,0) \succ_2^{HS_e} (0,0,0) \succ_2^{HS_e} (0,1,0) \sim_2^{HS_e} (1,1,1) \succ_2^{HS_e} \downarrow$$
$$(1,1,1) \succ_1^{HS_e} (1,0,0) \sim_1^{HS_e} (0,1,0) \succ_1^{HS_e} (0,0,0) \succ_1^{HS_e} \downarrow$$
$$(0,1,0) \succ_3^{HS_e} (0,0,0) \succ_3^{HS_e} (1,0,0) \sim_3^{HS_e} (1,1,1) \succ_3^{HS_e} \downarrow$$

Hiring example revisited

Φ	(a,b,x)	$x \leftrightarrow (a \wedge b)$
(prof. A) φ_1	(1,0,0)	
(prof. B) φ_2	(1,1,1)	
(prof. C) φ_3	(0,1,0)	
φ_4	(0,0,0)	

$f(\omega)$

$$(1,0,0) \succ_2^{HS_e} (0,0,0) \succ_2^{HS_e} (0,1,0) \sim_2^{HS_e} (1,1,1) \succ_2^{HS_e} \downarrow$$
$$(1,1,1) \succ_1^{HS_e} (1,0,0) \sim_1^{HS_e} (0,1,0) \succ_1^{HS_e} (0,0,0) \succ_1^{HS_e} \downarrow$$
$$(0,1,0) \succ_3^{HS_e} (0,0,0) \succ_3^{HS_e} (1,0,0) \sim_3^{HS_e} (1,1,1) \succ_3^{HS_e} \downarrow$$

Coordinated white manipulation

- one agent can white manipulate alone and improve the social welfare
- the group can agree on how to manipulate and this improve the social welfare
- idea: negotiate on how to lie
- example: fallback bargaining

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers “fallback” on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ d=1 \qquad d=2 \qquad d=3 \quad d=4 \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

consensus for
r = n

- hiring example:

$$M^h = \begin{pmatrix} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \\ d=1 & d=2 & d=3 & d=4 \end{pmatrix}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \left(\begin{array}{cc|cc} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{array}\right) \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

consensus for r = n

- hiring example:

$$M^h = \begin{array}{c} \left(\begin{array}{cccc} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{array}\right) \\ d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} {\scriptstyle d=1} & {\scriptstyle d=2} & {\scriptstyle d=3} & {\scriptstyle d=4} \end{array} \end{array}$$

consensus for r = n

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers “fallback” on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ d=1 \quad\quad d=2 \quad\quad d=3 \quad d=4 \end{array}$$

Fallback **B**argaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

consensus for r = 2

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

Fallback **B**argaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers “fallback” on less and less preferred alternatives

$$M = \begin{array}{c} \left(\begin{array}{c|ccc} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{array}\right) \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

consensus for r = 2

- hiring example:

$$M^h = \begin{array}{c} \left(\begin{array}{cccc} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{array}\right) \\ d=1 \qquad d=2 \qquad d=3 \quad d=4 \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

consensus for r = 2

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

r = n

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \left(\begin{array}{ccc|c} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{array} \right) \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array} \qquad \mathrm{r} = \mathrm{n}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]

- bargainers “fallback” on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{matrix} d=1 & d=2 & d=3 & d=4 \end{matrix} \end{array}$$

  - hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ \begin{matrix} d=1 & d=2 & d=3 & d=4 \end{matrix} \end{array} \qquad \mathrm{r} = \mathrm{n}$$

Fallback **B**argaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \begin{array}{cccc} {\scriptstyle d=1} & {\scriptstyle d=2} & {\scriptstyle d=3} & {\scriptstyle d=4} \end{array} \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array}$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \begin{pmatrix} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{pmatrix} \\ \scriptstyle d=1 \quad d=2 \quad d=3 \quad d=4 \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \begin{pmatrix} 100 & 000 & 111, 010 & \downarrow \\ 111 & 100, 010 & 000 & \downarrow \\ 010 & 000 & 111, 100 & \downarrow \end{pmatrix} \\ d=1 \quad d=2 \quad d=3 \quad d=4 \end{array} \quad r = 2$$

Fallback Bargaining

- introduced by S.J. Brams and D.M. Kilgour (1998)[5]
- bargainers "fallback" on less and less preferred alternatives

$$M = \begin{array}{c} \left(\begin{array}{cccc} a & b & c & d \\ a & c & b & d \\ b & a & d & c \end{array}\right) \\ \begin{array}{cccc} {\scriptstyle d=1} & {\scriptstyle d=2} & {\scriptstyle d=3} & {\scriptstyle d=4} \end{array} \end{array}$$

- hiring example:

$$M^h = \begin{array}{c} \left(\begin{array}{cc|cc} 100 & 000 & 111,010 & \downarrow \\ 111 & 100,010 & 000 & \downarrow \\ 010 & 000 & 111,100 & \downarrow \end{array}\right) \\ \begin{array}{cccc} d=1 & d=2 & d=3 & d=4 \end{array} \end{array} \quad r = 2$$

Fallback Bargaining & WM

- if $\downarrow$ is the least preferred outcome, it will not be the result of the bargaining (for both scoring functions)
- for r=n, the utilitarian social welfare of the bargaining output is the highest

Our contribution

- treat the inconsistency as an impasse and the impasse as a possible outcome
- introduce the idea of manipulability as a positive concept
- extend the judgment aggregation framework with an automatically built preference profile
- introduce social welfare concepts in the judgment aggregation framework

Future work

- analyze further the fallback bargaining for other social welfare functions
- analyze other agreement reaching protocols for the use of white manipulation
- analyze profiles with different preferences regarding the impasse
- redefine manipulation concepts in terms of coalition manipulation concepts
- extend the JA framework towards game theory

References

1. manipulability of JA functions:

F. Dietrich and C. List. Strategy-proof judgment aggregation. STICERD - Political Economy and Public Policy Paper Series 09, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE, Aug 2005.

2. manipulability of premise based JA:

F. Dietrich and P. Mongin. The premiss-based approach to judgment aggregation. Research memoranda, Maastricht : METEOR, Maastricht Research School of Economics of Technology and Organization, 2008.

3. manipulability of conclusion based JA:

A complete conclusion-based procedure for judgment aggregation G.Pigozzi, M. Slavkovik and L.van der Torre. In Proceedings of 1rst Conference on Algorithmic Decision Theory (forthcoming)

4. manipulability of distance based merging:

P. Everaere, S. Konieczny, and P. Marquis. The strategy-proofness landscape of merging. *J. Artif. Intell. Res. (JAIR)*, 28:49–105, 2007.

5. fallback bargaining

S. Brams and D. Kilgour. Fallback bargaining. Working Papers 98-10, C.V. Starr Center for Applied Economics, New York University, 1998.

S. Brams, D. Kilgour, and M. Sanver. A minimax procedure for negotiating multilateral treaties. In R. Avenhaus and I. Zartman, editors, *Diplomacy Games*, pages 265–282. Springer Berlin Heidelberg, May 2007.