Approximate Judgement Aggregation (for the case of the doctrinal paradox)

llan Nehama

Center for the Study of Rationality The Selim and Rachel Benin School of Computer Science and Engineering The Hebrew University of Jerusalem, Israel

Third International Workshop on Computational Social Choice Düsseldorf, Germany, September 14, 2010



- Doctrinal Paradox
- Research Question : Approximate Aggregation
- Approximate Aggregation Results
 - for The Doctrinal Paradox
 - for Other Agendas
 - for a Class of Agendas
- Conclusion



Suppose a defendant is accused in court of murder. In order to prove his guiltiness, one should convince the judge of two independent issues:

- (A) The defendant killed the victim
- (B) The defendant is sane

Conviction is defined to be the conjunction of the first two issues

 $(A \land B)$ The defendant is guilty.



A	B	$A \wedge B$
(Killed)	(Sane)	(Guilty)
0	1	0
1	0	0
1	1	1
0	0	0



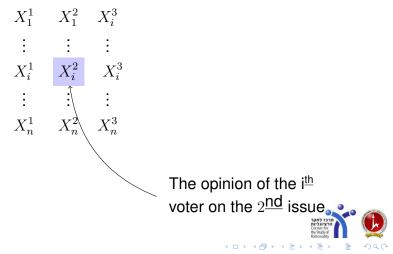
	A	В	$A \wedge B$	
	(Killed)	(Sane)	(Guilty)	
(0	1	0	—
Agonda	1	0	0	
Agenda {	1	1	1	
l	0	0	0	
	0	1	1	\leftarrow inconsistent
	1	0	1	\leftarrow inconsistent
	1	1	0	\leftarrow inconsistent
	0	0	1	\leftarrow inconsistent



	A	В	$A \wedge B$
	(Killed)	(Sane)	(Guilty)
Judge 1	1	0	0
Judge 2	1	1	1
Judge 3	0	1	0
Majority	1	1	0



A profile
$$X \in \{0,1\}^{n \times m}$$
 $\begin{pmatrix} n & : \text{Number of voters} \\ m = 3 & : \text{Number of issues} \end{pmatrix}$



$$X_{1}^{1} \qquad X_{1}^{2} \qquad X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$X_{i}^{1} \qquad X_{i}^{2} \qquad X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$X_{n}^{1} \qquad X_{n}^{2} \qquad X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2}$$

The ith row X_i represents the **consistent** opinion of the ith voter



$$\begin{array}{cccc} X_{1}^{1} & X_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & X_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & X_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{array}$$

The jth column X^j represents the opinions of all voters on the jth issue

$$F\begin{pmatrix} X_{1}^{1} & X_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & X_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & X_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

An **aggregation mechanism** returns for every profile an **aggregated opinion**

$$F: \{\{0,1\}^m\}^n \to \{0,1\}^m$$

$$F\begin{pmatrix} X_{1}^{1} & X_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & X_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & X_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

Definition (Consistency)

F is **consistent** if it returns a consistent result whenever all voters voted consistently

$$a_3 = a_1 \wedge a_2$$



$$F\begin{pmatrix} X_{1}^{1} & X_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & X_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & X_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

Definition (Independence)

F is **independent** if the aggregated opinion of the j^{th} issue depends solely on the votes for the j^{th} issue

the Study of Rationality

$$F\begin{pmatrix} \mathbf{X_{1}^{1}} & X_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ \mathbf{X_{i}^{1}} & X_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ \mathbf{X_{n}^{1}} & X_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

Definition (Independence)

F is **independent** if the aggregated opinion of the j^{th} issue depends solely on the votes for the j^{th} issue

the Study of Rationality

$$F\begin{pmatrix} X_{1}^{1} & \mathbf{X}_{1}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & \mathbf{X}_{i}^{2} & X_{i}^{3} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & \mathbf{X}_{n}^{2} & X_{n}^{3} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

Definition (Independence)

F is **independent** if the aggregated opinion of the j^{th} issue depends solely on the votes for the j^{th} issue

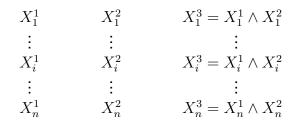
the Study of Rationality

$$F\begin{pmatrix} X_{1}^{1} & X_{1}^{2} & \mathbf{X_{i}^{3}} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{i}^{1} & X_{i}^{2} & \mathbf{X_{i}^{3}} = X_{i}^{1} \wedge X_{i}^{2} \\ \vdots & \vdots & \vdots \\ X_{n}^{1} & X_{n}^{2} & \mathbf{X_{n}^{3}} = X_{n}^{1} \wedge X_{n}^{2} \end{pmatrix} = (\mathbf{a_{1}}, \mathbf{a_{2}}, \mathbf{a_{3}})$$

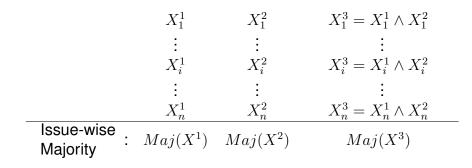
Definition (Independence)

F is **independent** if the aggregated opinion of the j^{th} issue depends solely on the votes for the j^{th} issue

the Study of Rationality

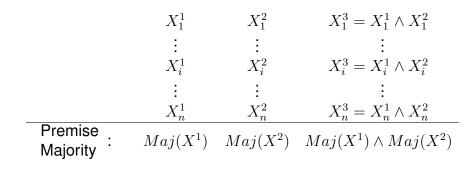






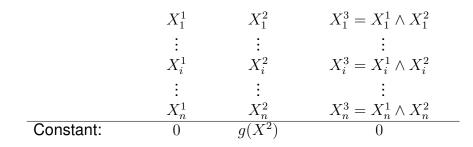
Independence: \checkmark Consistency: X





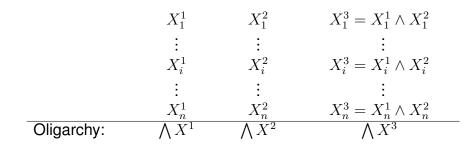
Independence: X Consistency: \checkmark





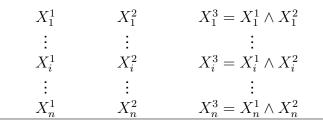
Independence: \checkmark Consistency: \checkmark





Independence: \checkmark Consistency: \checkmark





Independence: Consistency:

Are there any other consistent and independent aggregation mechanisms?



$\langle A, B, A \wedge B \rangle$ - Oligarchy

Definition (Oligarchy)

An oligarchy of S returns 1 iff all the members of S voted 1.

$$u_S(\bar{x}) = \bigwedge_{i \in S} x_i$$



Theorem

Let *F* be an independent and consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$. Then there exists three boolean functions $f, g, h : \{0, 1\}^n \rightarrow \{0, 1\}$ s.t. $F(X) = \langle f(X^1), g(X^2), h(X^3) \rangle$ and $f = h \equiv 0$ or $g = h \equiv 0$ or f = g = h and it is an oligarchy.

This theorem is a direct corollary from a series of works in the more general framework of aggregation. (E.g., Nehring&Puppe 2007, Holzman&Dokow 2008)



Theorem

Let F be a δ -independent and δ -consistent aggregation mechanism for $\langle A,B,A\wedge B\rangle$.





Theorem

Let F be a $\,\delta$ -independent and δ -consistent aggregation mechanism for $\langle A,B,A\wedge B\rangle$.

Then

Definition (δ -consistent)

F is δ -consistent if the following test fails with probability at most δ :

Choose a consistent profile X uniformly at random. Check whether F(X) is a consistent opinion.

4 B b

< ∃→

Theorem

Let F be a δ -independent and δ -consistent aggregation mechanism for $\langle A,B,A\wedge B\rangle$.

Then

Definition (δ -independent)

F is δ -independent if the following test fails with probability at most δ :

Choose a consistent profile X uniformly at random. Choose an issue j uniformly at random.

Choose a random consistent profile Y s.t. $X^j = Y^j$. Check whether $(F(X))^j$ equals $(F(Y))^j$



Theorem

Let *F* be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$. Then

Notice that

0-consistency≡Consistency 0-independence≡Independence



Theorem

Let *F* be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$.

Then

Notice that

0-consistency≡Consistency 0-independence≡Independence

Moreover, for $\delta < C \cdot 4^{-n} \approx \frac{1}{\text{Number of profiles}}$

 δ -consistency \equiv Consistency δ -independence \equiv Independence



Theorem

Let $\delta > exp(n, \epsilon)$ Let F be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$. Then there exists an independent and consistent aggregation mechanism G that agrees with F on at least $1 - \epsilon$ of the profiles.

Notice that

0-consistency≡Consistency 0-independence≡Independence

Moreover, for $\delta < C \cdot 4^{-n} \approx \frac{1}{\text{Number of profiles}}$

 δ -consistency \equiv Consistency δ -independence \equiv Independence



Theorem

Let *F* be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$. Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.

The other direction is trivial

Theorem

Let *F* and *G* be two aggregation mechanisms for $\langle A, B, A \wedge B \rangle$ such that

- G is independent and consistent
- F and G agree on at least 1ϵ of the profiles

then F is ϵ -independent and 6ϵ -consistent.



Main result for $\langle A, B, A \wedge B \rangle$

Theorem

For any $\epsilon > 0$ and $\delta = poly(\epsilon, n)$: $(\delta \approx C \cdot n^{-2} \epsilon^5)$ Let *F* be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \wedge B \rangle$. Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.



- Restricting ourself to independent mechanisms.
- Applying an (agenda independent) technique to extend the result to δ-independence and δ-consistency.



Given an independent $\delta\text{-consistent}$ aggregation mechanism $F=\langle f,g,h\rangle$



Given an independent $\delta\text{-consistent}$ aggregation mechanism $F=\langle f,g,h\rangle$

Definition (Influence (Banzhaf Power Index))

The **influence** of the ith voter on f is the probability he can change the result by changing his vote.

$$I_i(f) = \Pr[f(x) \neq f(x \oplus e_i)]$$

Definition (Ignorability)

The **ignorability** of the ith voter on f is the probability f returns 1 although i voted 0.

$$P_i(f) = \Pr[f(x) = 1 | x_i = 0]$$

10/17

Given an independent δ -consistent aggregation mechanism $F=\langle f,g,h\rangle$ We show that

• f is an oligarchy iff

$$\forall i : I_i(f)P_i(f) = 0$$



Given an independent $\delta\text{-consistent}$ aggregation mechanism $F=\langle f,g,h\rangle$ We show that

• f is an oligarchy iff

$$\forall i : I_i(f)P_i(f) = 0$$

$$\forall i : I_i(f)P_i(g) \leqslant 4\delta$$



Techniques - How did we get this result?

Given an independent $\delta\text{-consistent}$ aggregation mechanism $F=\langle f,g,h\rangle$ We show that

• f is an oligarchy iff

$$\forall i : I_i(f)P_i(f) = 0$$

۰

$$\forall i : I_i(f)P_i(g) \leqslant 4\delta$$

• Let u be the oligarchy of the voters with small ignorability (either $P_i(f)$ or $P_i(g)$) Then,

f and g are close to uF is close to $\langle u, u, u \rangle$.



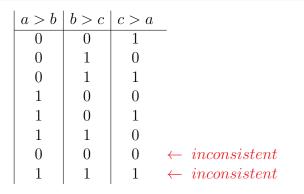
10/17

- Doctrinal Paradox
- Research Question : Approximate Aggregation
- Approximate Aggregation Results
 - for The Doctrinal Paradox
 - for Other Agendas
 - for a Class of Agendas
- Conclusion



- Doctrinal Paradox
- Research Question : Approximate Aggregation
- Approximate Aggregation Results
 - for The Doctrinal Paradox
 - for Other Agendas
 - Preference Agenda
 - XOR Agenda $\langle A, B, A \oplus B \rangle$
 - for a Class of Agendas
- Conclusion







Theorem (Condorcet Paradox)

Pair-wise majority might lead to inconsistent outcome.



Theorem (Condorcet Paradox)

Pair-wise majority might lead to inconsistent outcome.

Theorem (Arrow's Theorem 1950)

So is any other non-dictatorial aggregation mechanism that satisfies independence and Pareto.



Theorem (Condorcet Paradox)

Pair-wise majority might lead to inconsistent outcome.

Theorem (Arrow's Theorem 1950)

So is any other non-dictatorial aggregation mechanism that satisfies independence and Pareto.

Theorem (Kalai 2002, Mossel 2009)

For any $\epsilon > 0$:

Let F be an independent, $K\epsilon$ -consistent (and balanced) preference aggregation mechanism.

Then there exists an independent and consistent aggregation mechanism G(i.e., dictatorship) that agrees with F on at least $1 - \epsilon$ of the profiles.



Other Agendas - $\langle A, B, A \oplus B \rangle$

A	B	$A \oplus B$	
0	1	1	
1	0	1	
1	1	0	
0	0	0	
0	1	0	\leftarrow inconsistent
1	0	0	\leftarrow inconsistent
1	1	1	\leftarrow inconsistent
0	0	1	\leftarrow inconsistent



Other Agendas - $\langle A, B, A \oplus B \rangle$

Theorem

For any $\epsilon > 0$ and $\delta = poly(\epsilon, n)$: $(\delta = C \cdot \epsilon)$

Let *F* be a δ -independent and δ -consistent aggregation mechanism for $\langle A, B, A \oplus B \rangle$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.



Techniques - How did we get this result?

- Restricting ourself to independent mechanisms.
- Applying an (agenda independent) technique to extend the result to δ-independence and δ-consistency.



Techniques - How did we get this result?

Given an independent δ -consistent aggregation mechanism $F = \langle f, g, h \rangle$ We describe f,g, and h using Fourier representation and prove that

$$1 - 2\delta = \sum_{\chi} \widehat{f}(\chi)\widehat{g}(\chi)\widehat{h}(\chi)$$

when

- The summation is over all functions χ s.t. (χ, χ, χ) is consistent
- $\left| \widehat{f}(\chi) \right|$ equals 1 2d for d being the distance between f and χ .

in order to get that *F* is 'close to' $\langle \chi, \chi, \chi \rangle$.

(日)

For any $\epsilon > 0$, $m, n \ge 1$, and $\delta = poly\left(\frac{1}{n}, \epsilon, m\right)$:

Let X be a premise-conclusion agenda over m issues in which each issue is either a premise, or a conclusion of at most two premises.

Let F be a δ -independent and δ -consistent aggregation mechanism for $\mathbb X$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.



For any $\epsilon > 0$, $m, n \ge 1$, and $\delta = poly\left(\frac{1}{n}, \epsilon, m\right)$:

Let X be a premise-conclusion agenda over m issues in which each issue is either a premise, or a conclusion of at most two premises.

Let F be a δ -independent and δ -consistent aggregation mechanism for $\mathbb X$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.

For instance:
$$\langle A, B, A \oplus B \rangle$$

 $\langle A, B, A \wedge B, A \vee B \rangle$
 $\langle A, B, C, A \wedge B \vee C \rangle$
 $\langle A, B, C, A \wedge B, B \oplus C, A \wedge C \rangle$
 $\langle A \wedge B, B \wedge C, C \wedge A \rangle$

16/17

For any $\epsilon > 0$, $m, n \ge 1$, and $\delta = poly\left(\frac{1}{n}, \epsilon, m\right)$:

Let X be a premise-conclusion agenda over m issues in which each issue is either a premise, or a conclusion of at most two premises.

Let F be a δ -independent and δ -consistent aggregation mechanism for $\mathbb X$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.

For instance:
$$\langle A, B, A \oplus B \rangle$$

 $\langle A, B, A \wedge B, A \vee B \rangle$
 $\langle A, B, C, A \wedge B \vee C \rangle$
 $\langle A, B, C, A \wedge B, B \oplus C, A \wedge C \rangle$
 $\langle A \wedge B, B \wedge C, C \wedge A \rangle$

16/17

For any $\epsilon > 0$, $m, n \ge 1$, and $\delta = poly\left(\frac{1}{n}, \epsilon, m\right)$:

Let X be a premise-conclusion agenda over m issues in which each issue is either a premise, or a conclusion of at most two premises.

Let F be a δ -independent and δ -consistent aggregation mechanism for $\mathbb X$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.



For any $\epsilon > 0$, $m, n \ge 1$, and $\delta = poly\left(\frac{1}{n}, \epsilon, m\right)$:

Let X be a premise-conclusion agenda over m issues in which each issue is either a premise, or a conclusion of at most two premises.

Let F be a δ -independent and δ -consistent aggregation mechanism for $\mathbb X$.

Then there exists an independent and consistent aggregation mechanism *G* that agrees with *F* on at least $1 - \epsilon$ of the profiles.

Technique: $\bullet \land$ and \oplus represent all boolean functions of two arguments.

• Induction over the number of issues.





• We defined the question of approximate aggregation.



- We defined the question of approximate aggregation.
- We proved approximate aggregation theorems for $\langle A,B,A\wedge B\rangle$ and $\langle A,B,A\oplus B\rangle$.



- We defined the question of approximate aggregation.
- We proved approximate aggregation theorems for $\langle A,B,A\wedge B\rangle\,$ and $\langle A,B,A\oplus B\rangle$.
- We proved approximate aggregation theorems for a class of premise conclusion agendas.



- We defined the question of approximate aggregation.
- We proved approximate aggregation theorems for $\langle A,B,A\wedge B\rangle\,$ and $\langle A,B,A\oplus B\rangle$.
- We proved approximate aggregation theorems for a class of premise conclusion agendas.
- Open question:



- We defined the question of approximate aggregation.
- We proved approximate aggregation theorems for $\langle A,B,A\wedge B\rangle\,$ and $\langle A,B,A\oplus B\rangle$.
- We proved approximate aggregation theorems for a class of premise conclusion agendas.
- Open question:
 - Find an agenda and an aggregation mechanism that is δ-independent and δ-consistent but is far from any independent consistent aggregation mechanism.



- We defined the question of approximate aggregation.
- We proved approximate aggregation theorems for $\langle A,B,A\wedge B\rangle\,$ and $\langle A,B,A\oplus B\rangle$.
- We proved approximate aggregation theorems for a class of premise conclusion agendas.
- Open question:
 - Find an agenda and an aggregation mechanism that is δ-independent and δ-consistent but is far from any independent consistent aggregation mechanism.

Thank You



17/17

More information

email: ilan.nehama@mail.huji.ac.il
Homepage: www.cs.huji.ac.il/~ilan_n
Paper: http://arxiv.org/abs/1008.3829

Please write me any comments/questions/suggestions you have.

