Volby a volební systémy (NI-VOL), Přednáška č. 11

Judgement Aggregation II

Dušan Knop

Fakulta informačních technologií
České vysoké učení technické v Praze
https://courses.fit.cvut.cz/NI-VOL/

LS 2024/2025, 24. duben 2024

(Verze dokumentu: 22. 4. 2024 17:55)

Plán na dnešek

Naším hlavním cílem dnes bude ilustrovat sílu JA jako nástroje modelování kolektivního rozhodování tím, že ukážeme, jak zakomponovat agregaci preferencí do JA.
Než se však dostaneme k tomu, potřebujeme nějaké další základní nástroje:

- Více příkladů pravidel agregace
- Změna modelu: integritní omezení
- Obohacení modelu: racionální a reálná omezení

Dále:

- Preference jako agenda
- Modelování běžných volebních pravidel

Hlasovací Pravidla

Pro (ordináľní) preference bylo navrženo mnoho hlasovacích pravidel. Viděli jsme:

- Skórovací pravidla: Borda, majorita, veto, k-approval
- Vícekolová: majorita+runoff, STV
- Dodgson: zvolte alternativu nejbližší k tomu, aby byla Condorcetovým vítězem
- Copeland: maximalizujte počet výher head-to-head mínus ty prohrané
- Slater: sestavte head-to-head graf; najděte nejbližší lineární pořadí; zvolte nejlepší
- Kemeny: stejné, ale nyní s použitím váženého grafu
- ...

Na rozdíl od toho v JA bylo navrženo mnohem méně pravidel agregace.
Chceme pochopit proč.

Kemeny pro JA

- Pro profil $J = (J_1, \dots, J_n)$ a množinu stanovisek J bude $J \cap J_i$ shoda mezi J a agentem i.
- Tedy $|J \cap J_i|$ je spokojenost agenta i s J.
- Kemenyho pravidlo maximalizuje celkovou spokojenost:

$$F_{\text{kem}}(J) \in \arg \max_{J \in \mathcal{J}(\Phi)} \sum_{i \in N} |J \cap J_i| = \arg \max_{J \in \mathcal{J}(\Phi)} \sum_{\varphi \in J} |N_{\varphi}^{J}|$$

- Poznámka: F_{kem} je irresolutní.

Alternativní interpretace výše uvedené definice:

- maximalizace celkové shody s vybranými formulemi
- maximalizace průměrné spokojenosti (nebo shody)
- minimalizace kumulativní (Hammingovy) vzdálenosti k profilu

Zamyslete se: Jak byste implementovali toto pravidlo?

Slater pro JA

Slaterovo pravidlo maximalizuje počet formulí podporovaných většinou:

$$F_{\mathsf{sla}}(J) \in \arg\max_{J \in \mathcal{J}(\Phi)} \sum_{\varphi \in J} \mathbf{1}_{|N_\varphi^J| > \frac{n}{2}}$$

Toto pravidlo je také irresolutní.
V praxi je tedy potřeba rozhodovat o remízách.

Binární Agregace s Integritními Omezeními

Vezměme variantu modelu JA založeného na formulích: Agenda Φ obsahuje pouze literály, ale je nyní spárována s tzv. integritním omezením Γ (formule).

- Pojem konzistence rozšíříme na Γ-konzistenci: soubor úsudků $J \subseteq \Phi$ je Γ-konzistentní, pokud $J \cup \{\Gamma\}$ je konzistentní.
- Nechť $J(\Phi, \Gamma)$ je soubor všech úplných a Γ-konzistentních souborů úsudků.
- Nyní nás zajímají agregovací pravidla $F\colon J(\Phi, \Gamma)^n \to 2^{\Phi}$.

Poznámka: Tento model blíže odpovídá tomu, jak jsme poprvé představili doktrinální paradox (právní doktrína $= \Gamma$).
Zamyslete se:

- Jak definovat pravidla Slater a Kemeny pro tento model?
- A co majorita a další kvótní pravidla?

Vnoření původního modelu

Jakýkoli scénář JA s agendou Φ lze přeložit do nového modelu (použitím dvojnásobného počtu položek agendy):

- Pro každé $\varphi \in \Phi$ vytvořte kladný literál p_φ a jeho negaci.
- Sestavte integritní omezení jako konjunkci těchto formulí:
  - kódování úplnosti: $p_\varphi \vee p_{\neg\varphi}$ pro všechny kladné $\varphi \in \Phi$
  - kódování konzistence: $\neg \bigvee_{\varphi \in X} p_\varphi$ pro všechny mi-sady $X \subseteq \Phi$

Poznámky:

- Získané integritní omezení může být velmi rozsáhlé (exponenciální ve velikosti Φ).
- Takže tento překlad není vždy v praxi vhodný.
- Překlad v opačném směru je také možný (ale méně zjevný).

Příklad: Racionální a přípustná omezení

- Dosud jsme vždy ukládali stejné podmínky jak na jednotlivé úsudky, tak na kolektivní úsudky.
- Není to náhodou příliš restriktivní?
- Představte si pět členů místní samosprávy, kteří musí rozhodnout, zda schválit financování tří komunitních iniciativ:

	Škola	Divadlo	Parkování
Alice	Ne	Ne	Ano
Bob	Ano	Ano	Ano
Chris	Ano	Ne	Ano
Dana	Ano	Ano	Ne
Eve	Ne	Ano	Ne
Majorita	Ano	Ano	Ano

Racionální „Měl bych podpořit alespoň jednu iniciativu“

Přípustné „Nemůžeme si dovolit financovat všechny iniciativy“

Obohacení Modelu

Obohacujeme model binární agregace s integritními omezeními použitím dvou (možná odlišných) omezení:

Racionální omezení Γ_{in}: předpokládáme, že $J_i \in J(\Phi, \Gamma_{\text{in}})$ pro všechna $i \in N$.

Přípustné omezení Γ_{out}: doufáme, že $F(J) \in J(\Phi, \Gamma_{\text{out}})$.

Cvičení: Jak to ovlivňuje definici našich pravidel?
Diskuse: Γ_{out} by mohlo v porovnání s Γ_{in} být

- náročnější,
- méně náročné,
- stejné,
- nekomparabilní.

Vnoření agregace ordinálních preferencí

V agregaci preferencí agenti vyjadřují preference (lineární pořadí) nad souborem alternativ X a potřebujeme najít kolektivní preference.

Sestavíme preferenční agendu $\Phi_{\succeq}^{X}$ následovně:

$$\Phi_{\succeq}^{X} = \{p_{x \succeq y}, \neg p_{x \succeq y} \mid x, y \in X\}\,.$$

Poté sestavíme integritní omezení Γ jako konjunkci:

Úplnost $p_{x \succeq y} \vee p_{y \succeq x}$ pro všechna $x, y \in X$,

Antisymetrie $\neg(p_{x \succeq y} \wedge p_{y \succeq x})$ pro všechna $x, y \in X$ s $x \neq y$,

Tranzitivita $p_{x \succeq y} \wedge p_{y \succeq z} \rightarrow p_{x \succeq z}$ pro všechna $x, y, z \in X$.

Nyní můžeme simulovat Condorceho paradox:

	$p_{x \succeq y}$	$p_{x \succeq z}$	$p_{y \succeq z}$	Korespondující pořadí
Agent 1	Ano	Ano	Ano	$x \succ y \succ z$
Agent 2	Ne	Ne	Ano	$y \succ z \succ x$
Agent 3	Ano	Ne	Ne	$z \succ x \succ y$
Majorita	Ano	Ne	Ano	Není lineární pořadí

Výzva: Simulace volebních pravidel

- Při použití preferenční agendy je simulace pravidla agregace preferencí, které vrací společné preferenční pořadí v podstatě přímočará.
- Důvodem je, že vstup a výstup jsou stejného typu (pořadí).

Výzva

Simulovat volební pravidla (kde výstupy jsou alternativy).

Nápad: Použijte integritní a přípustná omezení.

Užitečná omezení

Využijte omezení k popisu vlastností binárních relací:

Úplnost $\bigwedge_{x,y}(p_{x\succeq y} \vee p_{y\succeq x})$

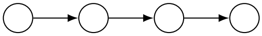
Anti-symetrie $\bigwedge_{x\neq y}\neg(p_{x\succeq y} \wedge p_{y\succeq x})$

Tranzitivita $\bigwedge_{x,y,z}(p_{x\succeq y} \wedge p_{y\succeq z} \rightarrow p_{x\succeq z})$

Slabé uspořádání `Úplnost` $\wedge$ `Tranzitivita`

Uspořádání `Slabé uspořádání` $\wedge$ `Anti-symetrie`

Soubory hodnocení splňující omezení `Uspořádání` korespondují s relacemi jako tato (uspořádání):



Užitečná omezení

Použijme $p_{x \succ y}$ jako zkratku pro $p_{x \succeq y} \wedge \neg p_{y \succeq x}$.

Další omezení:

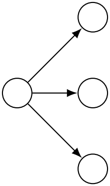
NoChain $\bigwedge_{x,y,z} \neg(p_{x \succ y} \wedge p_{y \succ z})$

Zakořeněné $\bigwedge_x \bigwedge_{y \neq x} p_{x \succ y}$

Dichotomous: `Slabé uspořádání` $\wedge$ `NoChain`

Winner: `Dichotomous` $\wedge$ `Zakořeněné`

Soubory s úsudky splňující `Winner` odpovídají vztahům jako je tento:



Platí:

- `Winner` $\models$ `Zakořeněný` (triviálně)
- `Uspořádání` $\models$ `Zakořeněný`

Simulace volebních pravidel

- Buď X množina alternativ a F pravidlo agregace pro $\Phi_{\succeq}^{X}$.
- Racionální omezení $\Gamma_{\text{in}} = \texttt{Ranking}$.
- Tedy: sada úsudků odpovídá preferenčním profilům.
- Omezení přípustnosti Γ_{out} požadujeme $\Gamma_{\text{out}} \models \texttt{Rooted}$.
- Tedy: z libovolného Γ_{out}-konzistentního výstupu můžeme vybrat vítěznou alternativu.
- Pro F, které zaručuje Γ_{out}-konzistentní výstupy na Γ_{in}-konzistentních profilech, řekneme, že F **simuluje volební pravidlo** F', pokud vždy funguje náseldující postup:
  1. Přeložíme preferenční profil do profilu sady úsudků
  2. použijte (možná irrezolutní) F na tomto profilu sady úsudků
  3. vyberte „kořenovou alternativu“ z každého výstupního souboru
  4. množina extrahovaných alternativ jsou vítězi podle F'

Výsledky simulace

Detaily důkazů jsou složité (a doposud byla analýza provedena pouze pro lichá n), ale pochopení intuica není příliš obtížné:

Racionalita	Přípustnost	SlaterJA	KemenyJA
$\Gamma_{\text{in}} = \texttt{Ranking}$	$\Gamma_{\text{out}} = \texttt{Ranking}$	Slater	Kemeny
$\Gamma_{\text{in}} = \texttt{Ranking}$	$\Gamma_{\text{out}} = \texttt{Winner}$	Copeland	Borda

Diskuze: Možná to vysvětluje řídkost pravidel pro JA?

Výpočetní úvahy

Pravidla agregace úsudků, která jsme diskutovali v předchozích přednáškách, byla všechna snadno vypočitatelná (např. počítání do dané kvóty). Dnešní „optimalizační“ pravidla jsou však odlišná:

- Je třeba prohledávat všechny možné modely, abychom našli ten „nejlepší“. Může jich být exponenciálně mnoho.
- Případ je podobný i pro model založený na formulích: musíme ověřit (možná) exponenciálně mnoho konzistentních souborů úsudků.
- A samo ověřování konzistence je obecně obtížný problém.

Kontrola Racionality

- Předtím, než se začneme naplno věnovat složitosti, pojďme zvážit mnohem základnější problém.
- Možná nejzákladnější problém v agregaci je určit, zda informace poskytnutá jednotlivým agentem je správně utvořená.

Racionalita

Vstup:	Agenda Φ, integritní omezení Γ, soubor úsudků $J \in 2^{\Phi}$.
Otázka:	Je J prvkem $J(\Phi, \Gamma)$?

Pro naše dva modely agregace to znamená:

- Pro formulově založenou agregaci úsudků (s $\Gamma = \top$):
  Je daný soubor úsudků J úplný a konzistentní?
- Pro binární agregaci s integritními omezeními (jen literály v Φ):
  Je daný soubor úsudků J modelem integritního omezení Γ?

Jaká je složitost tohoto problémů?

Složitost kontroly racionality

Složitost se liší pro oba modely agregace:

Věta

Racionalita je v P pro binární agregaci s integritními omezeními, ale NP-úplná pro formulově založenou JA.

Důkaz.

BA Jedná se o kontrolu modelu pro výrokovou logiku, což je snadné (v pravdivostní tabulce pro Γ zkontrolujte řádek odpovídající J).

JA Kontrola konzistence je SAT (takže úplnost je snadná). $\square$

Problém určení výsledku

- Dnes se budeme zabývat pouze formulově založenou JA.
- Nejprve se zaměříme na rozhodovací pravidla F (= jediný vítěz).
- „Hledací/Funkční“ problém, který nás zajímá, problém nalezení $F(J)$ pro zadaný profil J.
- Zde je odpovídající rozhodovací problém:

$\textsc{Rozhodování}(F)$

Vstup:	Agenda Φ, profil $\mathbf{J} \in J(\Phi)^n$, formule $\varphi \in \Phi$.
Otázka:	Je φ prvkem $F(\mathbf{J})$?

- Proč je F součástí názvu problému, ale Φ součástí vstupu?
- Pokud můžete efektivně vyřešit $\textsc{Rozhodování}(F)$, pak můžete také efektivně vyřešit hledací problém.

Proč by následující formulace nebyla dobrá?

Zadaná Φ, $\mathbf{J}$, a $J \subseteq \Phi$, rozhodněte, zda $F(\mathbf{J}) = J$.

Kvótní pravidla

Kvótní pravidlo F_q je definováno funkcí $q \colon \Phi \to \{0, 1, \ldots, n+1\}$:

$$F_q(J) = \{\varphi \in \Phi | \#N_J\varphi > q(\varphi)\}$$

Pro libovolné kvótní pravidlo je určení výsledku řešitelné:

Lemma

$\textsc{Rozhodování}(F_q)$ je v P pro každé q.

Důkaz.

Triviální. Stačí jen spočítat, zda $|N_J|\varphi > q(\varphi)$. □

Agregace z předpokladů

Premisové pravidlo F_{pre} pro předpoklady Φ_p a závěry Φ_c:

$$F_{\text{pre}}(\mathbf{J}) = \Delta \cup \{\varphi \in \Phi \mid \Delta \models \varphi\},$$

kde $\Delta = \left\{\varphi \in \Phi \mid |N_\varphi^{\mathbf{J}}| > \frac{n}{2}\right\}$.
Předpokládejme, že

- předpoklady $=$ literály a
- Φ uzavřené vzhledem k výrokovým proměnným

(zaručuje konzistenci a úplnost, alespoň pro liché n)

Lemma

Za těchto předpokladů je $\textsc{Rozhodování}(F_{\text{pre}})$ v P.

Důkaz.

- Pro předpoklady jde jen o počítání.
- Pro závěry jde o kontrolu modelu pro výrokovou logiku. $\square$

Určení výsledku pro irresolutní pravidla

- Většina prakticky užitečných pravidel agregace ve skutečnosti není resolutní a může vrátit (neprázdnou) množinu výherních souborů:
$$F\colon J(\Phi)^n \to 2^{(2^\Phi)} \setminus \{\emptyset\}$$
- Předpokládejme, že pro problém hledání jsme spokojeni s výpočtem jednoho z výsledných souborů.

$\textsc{Rozhodování}^\star(F)$

Vstup: Agenda Φ, profil $\mathbf{J} \in J(\Phi)^n$, podmnožina $L \subseteq \Phi$.
Otázka: Existuje J', které patří do $F(\mathbf{J})$ takové, že $L \subseteq J'$?

Pozorování

- Hledání lze řešit opakovaným řešením $\textsc{Rozhodování}^\star(F)$.
- Práce s formulemi φ místo L by nefungovala.

Kemeny

Kemeny maximalizuje shodu s přijatými formulemi:

$$F_{\mathsf{kem}}(\mathbf{J}) = \arg \max_{J \in J(\Phi)} \sum_{\varphi \in J} |N_{\varphi}^{\mathbf{J}}|$$

- Zjevně je určení výsledku pro Kemenyho poměrně obtížné.
- Naivní algoritmus by postupoval následovně:
  - Projdi všechny (bezdoplňkové a úplné) sady posudků (existuje jich exponenciálně mnoho).
  - Pro každou z nich zkontroluj, zda je konzistentní (NP-těžký problém).
  - Pro konzistentní sady zjisti celkovou shodu (snadné).
  - Vrať maximum (nebo jedno z maxim, v případě shod).
- Ale možná existuje chytřejší způsob?

Kemeny

Lépe než použití naivního algoritmu, ale náš problém je stále velmi výpočetně složitý.
Dále ukážeme:

Věta (Endriss, Grandi, Porello, 2012)

Problém určení výsledku pro Kemeny je $\mathsf{P}_{||}^{\mathsf{NP}}$-úplný.

- (Připomenutí) $\mathsf{P}_{||}^{\mathsf{NP}} = \mathsf{P}^{\mathsf{NP}}[\log]$ je třída problémů, které lze řešit v polynomiálním čase s logaritmickým počtem dotazů na NP-orákulum (adaptivní).
- Těžkost lze dokázat pomocí redukce z problému určení vítěze pro Kemenyho pravidlo ve volbách (Hemaspaandra et al., 2005).
- Tuto část důkazu přeskočíme a ukážeme pouze, jak dokázat náležení do $\mathsf{P}_{||}^{\mathsf{NP}}$.

Kemenyho skóre

Připomeňme si, že Kemeny hledá konzistentní soubor úsudků J, který maximalizuje skóre Kemenyho:

$$\mathsf{ks}^{\mathbf{J}}(J) = \sum_{\varphi \in J} |N_{\varphi}^{\mathbf{J}}| \,.$$

Uvažujme nejprve tento problém:

KemenySkore

Vstup:	Agenda Φ, profil $\mathbf{J} \in J(\Phi)^n$, podmnožina $L \subseteq \Phi$, $K \in \mathbb{N}$.
Otázka:	Existuje $J' \in J(\Phi)$ takové, že $L \subseteq J'$ a $\mathsf{ks}^{\mathbf{J}}(J') \geq K$?

Pozorování

KemenySkore je v NP.

Algoritmus

Lemma

Problém určení výsledku Kemenyho pravidlo je v Θ_p^2.

Důkaz.

- Použijeme NP-orákulum, které umí řešit $\textsc{KemenySkore}$.
- Určíme výsledek vyzkoušením všech možných hodnot pro K.
- Ale: počet dotazů na orákulum by byl nad-logaritmický, protože maximální K může být jakékoliv číslo mezi 1 a $K^* = \frac{|\Phi|}{2} \cdot |N|$.
- Ale můžeme provést chytřejší prohledávání prostoru všech K (binární vyhledávání):
  - Dotaz na $\textsc{KemenySkore}$ s $K := \frac{1}{2} \cdot K^*$
  - Pokud ANO, pokračujeme s $K := \frac{3}{2} \cdot K$ $(= \frac{3}{4} \cdot K^*)$
  - Pokud NE, pokračujeme s $K := \frac{1}{2} \cdot K$ $(= \frac{1}{4} \cdot K^*)$
  - A tak dále. . .
- Počet (adaptivních) dotazů je logaritmický: $O(\log_2 K^*)$. □

Něco pozitivního

- Zda je použití Kemenyho pravidla v praxi skutečně obtížné, závisí z velké části na agendě Φ.
- Existují agendy, pro které je určení výsledku zpracovatelné? Ano!

Věta

Pro agendy, které se skládají pouze z literálů, je problém určení výsledku pro Kemenyho pravidlo v P.

Úkol: Dokažte to!

- Výše uvedený výsledek je příliš zjednodušený a v podstatě nemá praktické využití, ale existují další výsledky pro zajímavější třídy agend:
- R. de Haan. Hunting for Tractable Languages for Judgment Aggregation. KR2018.

Co bychom po dnešku měli znát

- TODO