Generalizirani trag i normalne forme za logiku interpretabilnosti

Vedran Čačić
PMF – Matematički odsjek
Sveučilište u Zagrebu

Dubrovnik
radiona „Sustavi dokazivanja"
28. lipnja 2012.

Zašto logika interpretabilnosti?

Generalizacija Gödelovih teorema nepotpunosti (Löbov teorem):
Modalni pristup predikatu dokazivosti ($\Box$) $\longrightarrow$ GL:=K4+Löb
Tek 1976. Solovay dokazuje da je GL doista logika dokazivosti PA.

Što je sa uspoređivanjem teorija po snazi (interpretabilnost)?
Modalni pristup ($\rhd$) $\longrightarrow$ IL:=GL+„principi interpretabilnosti"
Razna proširenja su logike interpretabilnosti raznih poznatih teorija.
Berarducci, 1990: PA $\mapsto$ ILM.   Visser, 1988: NBG $\mapsto$ ILP.
GIL (zajednička svim „razumnim" teorijama) se još uvijek traži.

GL (zatvoreni fragment)

sintaksa

Formula: $Fo^{GL} ::= \bot \mid Fo^{GL} \to Fo^{GL} \mid \Box Fo^{GL}$

Ostalo su pokrate:

$$\neg A :\Longleftrightarrow A \to \bot \qquad A \land B :\Longleftrightarrow \neg(A \to \neg B)$$

$$\top :\Longleftrightarrow \neg\bot \qquad A \lor B :\Longleftrightarrow \neg A \to B$$

$$\Diamond A :\Longleftrightarrow \neg\Box\neg A \qquad A \leftrightarrow B :\Longleftrightarrow (A \to B) \land (B \to A)$$

Aksiomi: ($\to$ desno asocijativan)

(PT) sve propozicionalne tautologije

(K) $\Box(A \to B) \to \Box A \to \Box B$

(4) $\Box A \to \Box\Box A$

(Löb) $\Box(\Box A \to A) \to \Box A$

Pravila izvoda: (MP) $\dfrac{A \qquad A \to B}{B}$ (Gen) $\dfrac{A}{\Box A}$

Logika interpretabilnosti (zatvoreni fragment)

sintaksa

Formula: $Fo ::= \bot \mid Fo \to Fo \mid Fo \rhd Fo$
U IL, $\Box$ je pokrata: $\Box A :\Longleftrightarrow \neg A \rhd \bot$; sve ostalo kao u GL

Novi aksiomi: ($\rhd$ veže jače od logičkih veznika)

(J1) $\Box(A \to B) \to A \rhd B$

(J2) $A \rhd B \to B \rhd C \to A \rhd C$

(J3) $A \rhd C \to B \rhd C \to (A \vee B) \rhd C$

(J4) $A \rhd B \to \Diamond A \to \Diamond B$

(J5) $\Diamond A \rhd A$

Veltmanovi okviri

semantika

$\mathfrak{M} = (W, R, S)$, gdje je $S = (S_w)_{w \in W}$:

W: neprazan skup, čije elemente zovemo *svjetovi*

R: tranzitivna relacija na W, takva da je R^{-1} dobro utemeljena
$W[w] := \{u \in W \ : \ w \, R \, u\}$ — skup R-sljedbenika od w

S_w: tranzitivna refleksivna relacija na $W[w]$, proširuje restrikciju od R na $W[w]$ (odnosno, $w \, R \, u \, R \, v$ povlači $u \, S_w \, v$)

$\Vdash$: $\subseteq W \times Fo$, definirana induktivno:

- $w \nVdash \bot$
- $w \Vdash F \rightarrow G$ znači: $w \nVdash F$ ili $w \Vdash G$
- $w \Vdash F \rhd G$ znači: kad god je $W[w] \ni v \Vdash F$, tada postoji $u \Vdash G$ takav da je $v \, S_w \, u$.

GL okvir: (W, R)
Kao što se $\Box$ može izraziti pomoću $\rhd$ u IL, tako se i R može izraziti pomoću S u Veltmanovim okvirima:

$$u \, R \, v \Longleftrightarrow v \, S_u \, v$$

Dubina svijeta i trag formule

Funkcija koja svjetovima pridružuje ordinale, induktivno po R^{-1}:

$$\rho(w) := \sup_{wRv} \rho(v)^+ = \rho\big[W[w]\big]$$

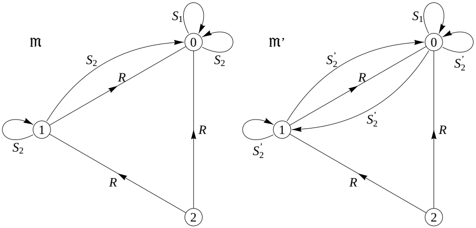
Trag u (zatvorenom fragmentu) GL: podskup od $\mathbb{N}$, takav da je (konačna) dubina od w element traga od φ akko $w \not\Vdash \varphi$.
Direktna posljedica: *svjetovi iste dubine su modalno ekvivalentni.*

To u IL ne vrijedi:

$$2 \not\Vdash \top \rhd \Diamond\top$$

$$2 \Vdash' \top \rhd \Diamond\top$$

Kako se spasiti?



IL sheme

Jedno rješenje: trovaljana logika. $\rho(w) \in tr(\varphi)$ može biti istina, laž, ili „ne znamo”.

Kad već imamo „trovaljanu logiku”, možemo generalizirati IL formule.

$$Sh ::= \bot \mid \star \mid Sh \to Sh \mid Sh \rhd Sh$$

($\star$ predstavlja nepoznate dijelove formule)

Instanca sheme: svaka pojava simbola $\star$ zamijeni se nekom (ne nužno istom) IL formulom.

Generalizirani trag

Preslikavanje $\mathrm{Tr} : Sh \to \{-1, 0, 1\}^{\mathbb{N}}$, definirano induktivno:

- $\mathrm{Tr}(\perp) := (-1, -1, -1, \dots)$
- $\mathrm{Tr}(\star) := (0, 0, 0, \dots)$
- $\mathrm{Tr}(F \to G)_n := \max\{-\mathrm{Tr}(F)_n, \mathrm{Tr}(G)_n\}$
- $\mathrm{Tr}(F \rhd G)_n := \begin{cases} 1, & m(F) \geq \min\{n, M(G)\} \\ -1, & M(F) < n \leq m(G) \\ 0, & \text{inače} \end{cases}$

  uz oznake za F (i analogno za G):

$$m(F) := \min_{\mathrm{Tr}(F)_n \neq -1} n \qquad M(F) := \min_{\mathrm{Tr}(F)_n = 1} n$$

Preciznost generaliziranog traga

„Biti precizniji": refleksivni parcijalni uređaj na $\{-1, 0, 1\}^{\mathbb{N}}$:

$$(a_n)_n \succeq (b_n)_n :\Longleftrightarrow \forall n(b_n \in \{a_n, 0\})$$

Najmanji element: $\mathrm{Tr}(\star) = (0, 0, \ldots)$.
Maksimalne elemente zovemo *determinirani* nizovi
— to su oni koji ne poprimaju vrijednost 0.

„Preciznija shema": shema s preciznijim tragom.
„Determinirana shema": shema s determiniranim tragom.
$m = M$ na determiniranim shemama.
Ako $F \succeq G$, tada $0 \le m(G) \le m(F) \le M(F) \le M(G) \le \infty$.

Korolar

Shema je manje precizna od svoje instance.

Rezultati o generaliziranom tragu

Lema (o slici traga $\longrightarrow$ M. Doko)

Svaki generalizirani trag je konstantan počevši od nekog indeksa. Sve praslike svakog traga su konačni ili kofinitni podskupovi od $\mathbb{N}$.

Lema (o istinitosti)

Neka je $\mathfrak{M} = (W, R, S)$ proizvoljan Veltmanov okvir, $w \in W$ svijet u njemu dubine $\rho(w) = n \in \mathbb{N}$, te F proizvoljna IL shema, i φ njena instanca. Tada:

- *ako je* $\mathrm{Tr}(F)_n = 1$, *tada* $\mathfrak{M}, w \Vdash \varphi$
- *ako je* $\mathrm{Tr}(F)_n = -1$, *tada* $\mathfrak{M}, w \nVdash \varphi$

Direktnim računanjem lako se dokaže:

$$\mathrm{Tr}(\neg F)_n := -\,\mathrm{Tr}(F)_n$$

$$\mathrm{Tr}(F \mathbin{\substack{\vee \\ \wedge}} G)_n := \substack{\max \\ \min}\{\mathrm{Tr}(F)_n, \mathrm{Tr}(G)_n\}$$

Eliminacija $\rhd$ pomoću generaliziranog traga

Korolar (o tragu)

Ako su F i G dvije IL formule s istim determiniranim *tragom, tada $IL \vdash F \leftrightarrow G$.*

GL je podteorija od IL: sve njene formule su formule od IL (uz $\Box$ kao pokratu), a svi njeni aksiomi su teoremi od IL.

Ako je φ GL formula, $\mathrm{Tr}(\varphi) = 1 - 2 \cdot \chi_{tr(\varphi)}$.

Teorem (o eliminaciji $\rhd$)

Ako je F determinirana IL shema, tada su sve njene instance IL-dokazivo ekvivalentne GL formuli čiji je GL-trag $\mathrm{Tr}(F)^{-1}[\{-1\}]$.

Koja je to GL formula? $\longrightarrow$ M. Doko

Primjene teorema o eliminaciji $\rhd$

afirmativne, negativne i prijelazne sheme

Afirmativna shema: $A \succeq A_0 := \star \rhd \star$.
Negativna shema: $N \succeq N_0 := \neg A_0$.
Prijelazna shema: $X \succeq X_0 := \neg(A_0 \rhd N_0)$.

$$\mathrm{Tr}(A) \succeq \mathrm{Tr}(A_0) = (1, 0, 0, 0, \ldots)$$

$$\mathrm{Tr}(N) \succeq \mathrm{Tr}(N_0) = (-1, 0, 0, 0, \ldots)$$

$$\mathrm{Tr}(X) \succeq \mathrm{Tr}(X_0) = (-1, 1, 0, 0, 0, \ldots)$$

Svaka IL formula je ili afirmativna ili negativna.
Negacija mijenja klasu formule: afirmativna$\leftrightarrow$negativna.

Za IL shemu F i GL formulu φ, $F \Longleftrightarrow \varphi$ znači da je svaka instanca od F (lokalno) ekvivalentna φ (shvaćena kao IL formula).

Teorem

Neka je A afirmativna, N negativna, te X prijelazna shema.

$$
\begin{array}{rcr}
\bot \rhd \star & \Longleftrightarrow & \top \\
\star \rhd A & \Longleftrightarrow & \top \\
A \rhd \bot & \Longleftrightarrow & \Box\bot \\
N \rhd X & \Longleftrightarrow & \top \\
X \rhd \bot & \Longleftrightarrow & \Box\Box\bot
\end{array}
$$

Dokaz.

Račun u C++u: $\longrightarrow$ M. Doko $\Box$

Lokalna i globalna ekvivalencija

Dosad smo radili s lokalnom ekvivalencijom, ali postoji i pojam globalne ekvivalencije.

Definicija

Neka su φ i ψ dvije zatvorene IL formule.

- Kažemo da su φ i ψ *lokalno ekvivalentne*, i pišemo $\varphi \Leftrightarrow \psi$, ako za svaki Veltmanov okvir $\mathfrak{M}$ i za svaki svijet w u $\mathfrak{M}$, vrijedi $\mathfrak{M}, w \Vdash \varphi$ ako i samo ako $\mathfrak{M}, w \Vdash \psi$.
- Kažemo da su φ i ψ *globalno ekvivalentne*, i pišemo $\varphi \stackrel{g}{\Longleftrightarrow} \psi$, ako za svaki Veltmanov okvir $\mathfrak{M}$ vrijedi $\mathfrak{M} \Vdash \varphi$ ako i samo ako $\mathfrak{M} \Vdash \psi$.

Globalna ekvivalencija je slabija.
Lokalna ekvivalencija ima svojstvo supstitucije, globalna nema.
Normalne forme za GL su *globalno* ekvivalentne svojim formulama.

Teorem o generalizaciji

još 3 normalne forme

Teorem

$\varphi \overset{g}{\Longleftrightarrow} \psi$ *ako i samo ako* $\Box\varphi \Leftrightarrow \Box\psi$.

Ideja dokaza.

- $\Rightarrow$ Fiksiramo w i konstruiramo novi model nad $W[w]$.
- $\Leftarrow$ Dodamo novi svijet w „ispod” čitavog modela. $\Box$

Teorem

Neka je A afirmativna, N negativna, te X prijelazna IL shema. Vrijede sljedeće globalne ekvivalencije:

$$N \overset{g}{\Longleftrightarrow} \bot \qquad A \rhd N \Leftrightarrow \neg X \overset{g}{\Longleftrightarrow} \Box\bot$$

Dokaz.

Primjenom teorema o generalizaciji, i računom $\longrightarrow$ M. Doko $\Box$

Najjednostavnija formula iz koje $\rhd$ nije eliminabilan

kako smo je našli? $\longrightarrow$ M. Doko

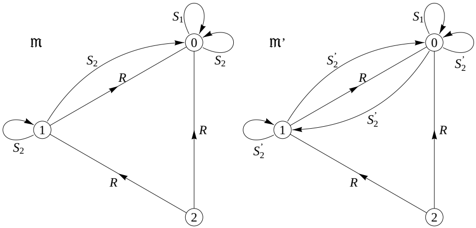
$$\chi := \top \rhd \Diamond\top \to \Box\bot \quad =(\bot \underset{(\rhd)}{\overset{\longrightarrow}{}} \bot)\rhd((\bot \underset{(\rhd)}{\overset{\longrightarrow}{}} \bot)\rhd\bot\to\bot)\to(\bot \underset{(\rhd)}{\overset{\longrightarrow}{}} \bot)\rhd\bot$$

Intuitivno, $\mathfrak{M} \Vdash \chi$ kaže da svaki neterminalni svijet w u $\mathfrak{M}$ ima R-sljedbenika čiji su svi S_w-sljedbenici terminalni.

Teorem

Ne postoji GL formula φ takva da je $\chi \overset{g}{\Longleftrightarrow} \varphi$.

Dokaz slikom.



Vidimo $\mathfrak{M} \Vdash \chi$, ali $\mathfrak{M}' \nVdash \chi$.
Pretpostavimo suprotno. Tada $\mathfrak{M} \Vdash \varphi$, ali $\mathfrak{M}' \nVdash \varphi$.
No $\mathfrak{M}$ i $\mathfrak{M}'$ imaju isti GL okvir, s kojim se moraju slagati na GL formulama: $\mathfrak{N} \Vdash \varphi$ i $\mathfrak{N} \nVdash \varphi$, što je kontradikcija. $\square$

Koliko je dobar generalizirani trag?

brojanje IL formula

IL sheme često nisu determinirane. No što je s IL formulama? Na koliki dio njih se može primijeniti teorem o eliminaciji $\rhd$, eventualno nakon primjene teorema o generalizaciji?

Što uopće znači „koliki dio"? IL formula ima beskonačno mnogo. No možemo ih stratificirati po složenosti (broju simbola $\to$ i $\rhd$).

Označimo sa $\#_n$ broj IL formula složenosti n. Imamo rekurziju

$$\#_0 = 1, \qquad \#_{n+1} = 2\sum_{k=0}^{n} \#_k \#_{n-k}$$

čije je rješenje (eksponencijalno uvećani Catalanovi brojevi)

$$\#_n = \frac{2^n}{n+1}\binom{2n}{n}.$$

Koliko je dobar generalizirani trag?

asimptotski udio

Asimptotski udio formula sa svojstvom P je limes (ako postoji)

$$\mu(P) := \lim_n \frac{\text{broj formula složenosti } n \text{ sa svojstvom } P}{\text{broj svih formula složenosti } n \ (= \#_n)}$$

Dakle, $\mu(P)$ je približna vjerojatnost da slučajno odabrana IL formula dovoljno velike složenosti ima svojstvo P.

Primjer

Asimptotski udio afirmativnih formula je

$$\mu(\mathit{afirmativne}) = \frac{1 + \frac{3}{\sqrt{17}}}{2} \approx 86.38\%,$$

što se može dobiti na sljedeći način:
gramatika $\rightarrow$ rekurzija $\rightarrow$ funkcija izvodnica $\rightarrow$ asimptotski udio

Koliko je dobar generalizirani trag?

determinirane i normalne formule

Prisjetimo se, IL formula je determinirana ako joj je trag determiniran, odnosno ako je lokalno ekvivalentna nekoj GL formuli.

Za IL formulu kažemo da je *normalna* ako joj je generalizacija determinirana, odnosno ako je globalno ekvivalentna nekoj GL formuli (dakle booleovskoj kombinaciji formula oblika $\Box^n\bot$).

Brojanjem formula iz gornjih ekvivalencija (5+3) dobije se:

Teorem (V. Čačić, V. Kovač)

$$\mu(\textit{determinirane}) \geq \frac{\sqrt{4679\sqrt{17}-2519}}{8\sqrt{323}} \approx 90\%$$

$$\mu(\textit{normalne}) \geq \tfrac{55}{126} - \tfrac{45}{14\sqrt{17}} + \tfrac{8}{63}\sqrt{\tfrac{-2519+4679\sqrt{17}}{323}} + $$

$$+\tfrac{32}{63}\sqrt{\tfrac{3877-242\sqrt{17}+8\sqrt{-154217+51073\sqrt{17}}}{8995}} \approx 94.22\%$$

Što dalje?

- Normalne forme za djelić zatvorenog fragmenta IL koji nije u GL (asimptotski, manje od 6%; empirijski, manje od 2%)
- Egzaktne vrijednosti za $\mu(\mathit{determinirane})$ i $\mu(\mathit{normalne})$, ili čak μ(one koje imaju $\substack{\textit{lokalno}\\ \textit{globalno}}$ ekvivalentnu GL formulu)
- Nepostojanje normalnih formi za $\mathit{Prop} \neq \emptyset$ (po analogiji sa Solovayevim rezultatom za GL)
- Jesu li navedene (5 lokalnih i 3 globalne) ekvivalencije esencijalno jedini slučajevi gdje se $\rhd$ može eliminirati u IL?
- Univerzalno rješenje problema „zavisnosti” različitih potformula formule čiji generalizirani trag računamo
- U kojim sve proširenjima od IL je $\rhd$ eliminabilan u zatvorenom fragmentu, i može li se to dobiti generaliziranim tragom?
- Jednostavni strukturalni kriterij za modalnu ekvivalenciju svjetova (ono što je dubina za GL okvire)