Univerzalni okviri za neke modalne logike

Vedran Čačić[1]

PMF – Matematički odsjek

14. rujna 2021.
@HLU (Zoom)

[1]Ovaj je rad sufinancirala Hrvatska zaklada za znanost projektom #7459 (CompStruct).

Logika — motivacija

U najširem smislu:
želimo formalizirati matematičko razmišljanje.

Želimo *pitomu* logiku:
da možemo nešto matematički reći o njoj.
Primjeri: svojstvo konačnih modela, odlučivost

Želimo *izražajnu* logiku:
da možemo formalizirati razna razmišljanja.
Primjeri: relacije između objekata, kvantifikacija

Nažalost, to dvoje je često u koliziji.

Ekstremni primjeri

Logika sudova

Vrlo jednostavna logika, dva objekta: $\top$ i $\bot$.
Možemo izraziti samo istinosne funkcije (tablice istinitosti).
Ali imamo brojna lijepa svojstva:
odlučivost, kompaktnost, ...

Logika drugog reda

Možemo izraziti komplicirane tvrdnje,
kao što su aksiom matematičke indukcije na $\mathbb{N}$
ili aksiom potpunosti skupa $\mathbb{R}$.
Ali nemamo čak ni potpun dokazni sustav!

Logika prvog reda

U dosta smislova kompromis
između krajnosti s prošlog slajda.

Možemo modelirati „dovoljno dobro" razne matematičke strukture (PA za $\mathbb{N}$, ZF za kumulativnu hijerarhiju, ...)
Ipak, ne možemo izraziti dobru utemeljenost, niti fiksirati kardinalnost modela (Löwenheim–Skolemov teorem).

Također, imamo svojstva kao što su kompaktnost,
čak i potpunost dokaznog sustava (Gödel),
ali nemamo odlučivost (Churchov teorem)!

Složenost: *model checking* je PSPACE-potpun!
(Ali složenost sama za sebe baš i nije neki argument:
čak i logika sudova ima NP-potpun SAT!)

Modalna logika

Drugačiji (iz perspektive računarstva, bolji) kompromis.

Dovoljno izražajna za neke fenomene:
na primjer, za inverznu dobru utemeljenost.
Svojstvo konačnih modela $\Longrightarrow$ odlučivost.
Štoviše, polinomni *model checking*!

Veza s logikom prvog reda:

- standardna translacija / van Benthemov teorem
- kvantifikatori kao modalni operatori
  (valuacije kao svjetovi)

Nastala iz potrebe opisa modalnosti (nužnost, mogućnost) logičkih tvrdnji, kasnije generalizirana na razne druge neistinitosne operatore (znanje, vjerovanje, prošlost, budućnost, dokazivost, konzistentnost, interpretabilnost).

Modalni operatori — ugrubo

Opisujemo veze između *svjetova* (mogućih stanja stvari) od kojih se sastoji naš model.

Modalne varijable: unarne relacije (skupovi) svjetova.
Unarni modalni operatori: binarne relacije među svjetovima.
Binarni modalni operatori: ternarne relacije na svjetovima.
. . .

$$w \Vdash p :\Longleftrightarrow P_p(w) \tag{1}$$

$$w \Vdash \Box\varphi :\Longleftrightarrow (\forall v : w\,R\,v)(v \Vdash \varphi) \tag{2}$$

$$w \Vdash \varphi \rhd \psi :\Longleftrightarrow (\forall v : w\,R\,v \Vdash \varphi)(\exists u : v\,S_w\,u \Vdash \psi) \tag{3}$$

Koja su minimalna svojstva predikata dokazivosti potrebna da bi Gödelov dokaz teorema nepotpunosti „prošao"?

Osnovna ideja: dokazivost kao modalni operator!

$$\Box\varphi :\Longleftrightarrow \varphi \text{ je dokaziva}$$

Hilbert–Bernaysovi uvjeti $\rightsquigarrow$ modalni aksiomi:

K	$\Box(\varphi \to \psi) \land \Box\varphi \to \Box\psi$
4	$\Box\varphi \to \Box\Box\varphi$
Löb	$\Box(\Box\varphi \to \varphi) \to \Box\varphi$

Formule u zatvorenom fragmentu logike dokazivosti (GL_0):

$$\varphi ::= \bot \mid (\varphi \to \varphi) \mid \Box\varphi$$

Definicija

GL-okvir je struktura $\mathfrak{M} = (W, R)$, gdje je W neprazni skup **svjetova**, a R je tranzitivna relacija na W takva da je R^{-1} dobro utemeljena: ne postoji niz $w_0\ R\ w_1\ R\ w_2\ R \cdots$.

Notacija: $R[w] := \{u \in W\ :\ w\ R\ u\}$ — R-sljedbenici od w

Teorem (adekvatnost, potpunost, konačni modeli)

- *Ako je $\varphi \in GL_0$ teorem, istinita je na svakom svijetu svakog GL-okvira.*
- *U suprotnom, postoji* **konačni** *GL-okvir, i svijet u njemu, na kojem je φ lažna.*

Dubina svijeta i okvira

Dubina: funkcija $\rho\colon W \to \mathbf{On}$, definirana rekurzivno po (dobro utemeljenoj relaciji) R^{-1}:

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

Lema (o dubini)

Ako je α ordinal i w svijet takav da je $\rho(w) > \alpha$, tada postoji $v \in R[w]$ takav da je $\rho(v) = \alpha$.

Korolar

Slika dubine, $\operatorname{rng} \rho$, je također ordinal; zovemo ga „dubina okvira" i označavamo $\rho(\mathfrak{M})$.

Propozicijske varijable i modeli

U čitavoj GL (za razliku od GL_0) dozvoljavamo i modalne propozicijske varijable u formulama ($Prop := \{p, q, \ldots\}$).

Model je kao okvir, samo pored W i relacijâ koje modeliraju modalne operatore (poput R) ima i *relaciju forsiranja* zadanu između W i $Prop$ (i proširenu rekurzivno na sve formule). Dakle, GL-okvir je zapravo model za GL_0 (i slično za IL).

Ovo o čemu govorimo dalje uglavnom se odnosi na $Prop = \emptyset$, pa je svejedno govorimo li o modelima ili okvirima.

Modalna ekvivalencija

Definicija

Kažemo da su svjetovi w i w' (možda u različitim modelima) **modalno ekvivalentni** ako na njima vrijede iste formule. (Na okvirima, naravno, gledamo samo *zatvorene* formule.)

Modalna ekvivalencija je slabija od izomorfizma:
svaka modalna formula „vidi" okvir samo do fiksne **konačne** dubine (broj ugniježđenih modalnih operatora u formuli).

Dakle, na primjer, svjetovi na dubinama $\omega + 1$ i $\omega + 2$ jesu modalno ekvivalentni, ali nisu izomorfni.

(Također, okviri dubinā $\omega + 2$ i $\omega + 3$ nisu izomorfni.)

Napomena

Za svaki ordinal α postoji okvir $(\alpha, \ni)$ dubine točno α.

Univerzalni okvir za *GL*

postojanje

Okvir $\bar{\mathfrak{N}} := (\omega + 1, \ni)$ je **univerzalan**:
za svaki *GL*-okvir $\mathfrak{N}$ i za svaki svijet w u njemu, svijet $\bar{w} := \min\{\rho(w), \omega\}$ u $\bar{\mathfrak{N}}$ je modalno ekvivalentan w.

Štoviše, $\bar{w}$ je **jedini** takav svijet u $\bar{\mathfrak{N}}$: očito je jedini na svojoj dubini, a svjetovi na različitim konačnim dubinama nikada nisu modalno ekvivalentni (kao ni svijet na konačnoj i svijet na beskonačnoj dubini — ω je jedini takav u $\bar{\mathfrak{N}}$).

Zašto? Jer za svaki $n \in \omega$ postoji formula koja **karakterizira** svjetove dubine točno n:

$$\chi_0 := \Box\bot$$
$$\chi_{n+1} := \left(\Box^{n+2}\bot \land \Diamond\chi_n\right) \qquad (\Diamond\varphi \text{ je pokrata za } \neg\Box\neg\varphi).$$

Svijet ω je karakteriziran eliminacijom:
nijedna formula χ_n ne vrijedi na njemu.

Teorem

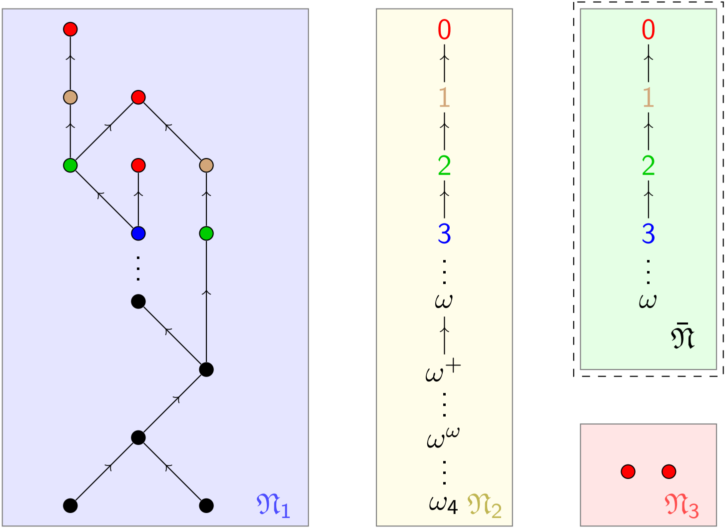
Štoviše, sâm $\bar{\mathfrak{N}}$ je jedinstveni (do na izomorfizam) GL-okvir s tim svojstvom (da svaki svijet svakog GL-okvira u njemu ima jedinstvenog modalnog ekvivalenta).

Doista, mora imati svijet na svakoj konačnoj dubini, također i na beskonačnoj dubini, i u svakoj od tih „hrpa" mora biti točno jedan svijet (kako su svjetovi iste dubine modalno ekvivalentni, inače ne bi vrijedilo svojstvo jedinstvenosti). Taj jedini svijet na beskonačnoj dubini mora biti točno na dubini ω, jer inače bi po lemi o dubini imao sljedbenika na dubini ω, pa opet ne bismo imali svojstvo jedinstvenosti.

R je isto u potpunosti fiksirana: za svjetove w i v takve da je $\rho(v) < \rho(w)$, mora biti $w \, R \, v$ po lemi o dubini i svojstvu jedinstvenosti, a onda ne smije biti $v \, R \, w$ (niti $w \, R \, w$) zbog inverzne dobre utemeljenosti ($v \, R \, w \, R \, v \, R \, w \, R \, v \, R \cdots$).

Univerzalni okvir za *GL*

dijagram



GL dobro modelira dokazivost u mnogim razumnim aritmetičkim teorijama.
1976. Solovay dokazuje: *GL* je upravo logika dokazivosti *PA*.

No što je s uspoređivanjem (proširenja) aritmetičkih teorija s obzirom na relativnu snagu dokazivanja?
Precizno, koja proširenja neke osnovne teorije mogu interpretirati neka druga?

Također primjenjujemo modalni pristup:
nad nekom osnovnom teorijom T,

$$\begin{array}{ccc} \Box\varphi & \leftrightsquigarrow & T \text{ dokazuje } \varphi \\ (\varphi \rhd \psi) & \leftrightsquigarrow & T + \varphi \text{ interpretira } T + \psi \end{array}$$

Interpretacija — općenita definicija

Teorija T **interpretira** teoriju S ako postoji preslikavanje j koje „pretvara" formule jezika od S u formule jezika od T, koje komutira s logičkim veznicima i čuva dokazivost:

$$j(\varphi \land \psi) = \big(j(\varphi) \land j(\psi)\big) \text{ (i slično za } \neg,\ \lor,\ \to,\ \leftrightarrow)$$
$$S \vdash \varphi \text{ ako i samo ako } T \vdash j(\varphi)$$

te postoji formula $s(x)$ nad jezikom od T za koju vrijedi

$$j(\forall x\, \varphi) = \forall x \big(s(x) \to j(\varphi)\big) \text{ (i slično za } \exists)$$
$$T \vdash \exists x\, s(x)$$

Primjer

- ZF interpretira PA (von Neumannovi ordinali)
- algebra interpretira geometriju (analitička geometrija)

Jedna dokazivost — mnogo interpretabilnosti

$IL := GL +$ „principi interpretabilnosti"

- **različiti** principi za različite osnovne teorije T

Proširenja IL su logike interpretabilnosti raznih teorija:

kad	tko	T	princip	proširenje
1988	Visser	NBG	permanencije	ILP
1990	Berarducci	PA	Montagnin	ILM

GIL (zajednička svim razumnim teorijama) se još uvijek traži.
Goris–Joosten–Mikec 2010–2020: ILP_0, ILW, ILR, ILR^*, ...

- logička konstanta $\bot$, veznik $\to$, modalni operator $\rhd$
- (zatvorena) formula: $\varphi ::= \bot \mid (\varphi \to \varphi) \mid (\varphi \rhd \varphi)$
- sve ostalo ($\neg$, $\top$, $\wedge$, $\vee$, $\leftrightarrow$, $\Box$, $\Diamond$) su pokrate:

$$\neg A :\Longleftrightarrow A \to \bot \qquad A \vee B :\Longleftrightarrow A \to \neg B$$

$$\Box A :\Longleftrightarrow \neg A \rhd \bot$$

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

$$\top :\Longleftrightarrow \neg\bot \qquad A \leftrightarrow B :\Longleftrightarrow (A \to B) \wedge (B \to A)$$

Definicija (Modalna dubina formule)

$$\begin{aligned} \delta(\bot) &:= 0 \\ \delta(\varphi \to \psi) &:= \max\left\{\delta(\varphi), \delta(\psi)\right\} \\ \delta(\varphi \rhd \psi) &:= \max\left\{\delta(\varphi), \delta(\psi)\right\} + 1 \end{aligned}$$

Aksiomi:

PT	(sve sudovne tautologije)
Löb	$\Box(\Box A \to A) \to \Box A$
J1	$\Box(A \to B) \to (A \rhd B)$
J2	$((A \rhd B) \land (B \rhd C)) \to (A \rhd C)$
J3	$((A \rhd C) \land (B \rhd C)) \to ((A \lor B) \rhd C)$
J4	$(A \rhd B) \to (\Diamond A \to \Diamond B)$
J5	$\Diamond A \rhd A$

Pravila zaključivanja: MP: $\dfrac{A \quad A \to B}{B}$ Gen: $\dfrac{A}{\Box A}$

Propozicija

GL je podteorija od *IL*: K slijedi iz J3, a 4 iz J4 i J5.

Definicija

Veltmanov okvir je struktura $\mathfrak{M} = \big(W, R, (S_w)_{w \in W}\big)$, gdje

- (W, R) je GL-okvir
- za svaki w, S_w je refleksivna tranzitivna relacija na $R[w]$, koja tamo proširuje R $(R \cap R[w]^2 \subseteq S_w)$

$\Vdash$: rekurzivno definirana relacija između svjetova i formula

- $w \nVdash \bot$ za svaki svijet w
- $w \Vdash F \to G$ znači $w \nVdash F$ ili $w \Vdash G$
- $w \Vdash F \rhd G$ znači: za svaki v takav da $w \mathrel{R} v \Vdash F$,
  postoji u takav da $v \mathrel{S_w} u \Vdash G$.

Teorem

Nema univerzalnog Veltmanovog okvira.

Dokaz: Pretpostavimo suprotno, da
postoji Veltmanov okvir $\bar{\mathfrak{M}}$ takav da
za svaki Veltmanov okvir $\mathfrak{M}$,
za svaki svijet w u $\mathfrak{M}$,
postoji jedinstveni svijet $\bar{w}$ u $\bar{\mathfrak{M}}$ takav da
za svaku zatvorenu IL-formulu φ,
$w \Vdash \varphi$ ako i samo ako $\bar{w} \Vdash \varphi$.

Konstruirat ćemo $\mathfrak{M}$ i w tako da $\bar{w}$ ne može postojati.

Prvo moramo definirati **karakteristične formule** za IL.

Konačno mnogo tipova formula
zadane dubine

Za svaki konačni skup IL_0-formula S, definiramo skup svih savršenih disjunktivnih normalnih formi nad formulama iz S:

$$D_0(S) := \left\{ \bigvee_{\psi \in K} \psi \ : \ K \subseteq \Big\{ \bigwedge_{\varphi \in T} \varphi \wedge \bigwedge_{\varphi \in S \setminus T} \neg\varphi \ : \ T \subseteq S \Big\} \right\}$$

(uz $\bigwedge \emptyset = \top$, $\bigvee \emptyset = \bot$), i nastavljamo rekurzivno:

$$D_{n+1}(S) := D_0\Big(D_n(S) \cup \big\{ \varphi \rhd \psi \ : \ \varphi, \psi \in D_n(S) \big\} \Big);$$

tada svaka $\varphi \in IL_0$ ima ekvivalentnu $\bar{\varphi} \in D_{\delta(\varphi)}(\emptyset)$.

Konačno mnogo tipova **svjetova**

zadane dubine

Dakle, za svaki $n \in \mathbb{N}$ postoji samo konačno mnogo *IL*-formulā modalne dubine n, do na logičku ekvivalenciju.

Korolar

Za svaki n postoji samo konačno mnogo svjetova dubine n, do na modalnu ekvivalenciju.

Skica dokaza.

Svaka dva modalno neekvivalentna svijeta dubine n moraju se razlikovati na nekoj formuli φ modalne dubine n ili manje, pa se moraju razlikovati i na $\bar{\varphi}$;
a takvih formula ima samo $\sum_{i=0}^{n} \# D_i(\emptyset) =: m$.

Dakle, ne može postojati više od 2^m modalno neekvivalentnih svjetova. □

Kontraprimjer

Promotrimo Veltmanov okvir $\mathfrak{M}_{\lightning} := \big(W, R, (S_w)_{w\in W}\big)$, gdje

$$P := [4 \cdot\cdot\, 7] \times [1 \cdot\cdot\, 3] \cup (>_3) \quad R := \{8\} \times [1 \cdot\cdot\, 7] \cup P$$

$$W := [1 \cdot\cdot\, 8] \quad S_1 := \emptyset \quad S_2 := \{1\}^2 \quad S_3 := \{1,2\}^2$$

$$S_4 := (\geq_3) \cup \{(1,2)\} \quad S_5 := S_6 := (\geq_3)$$

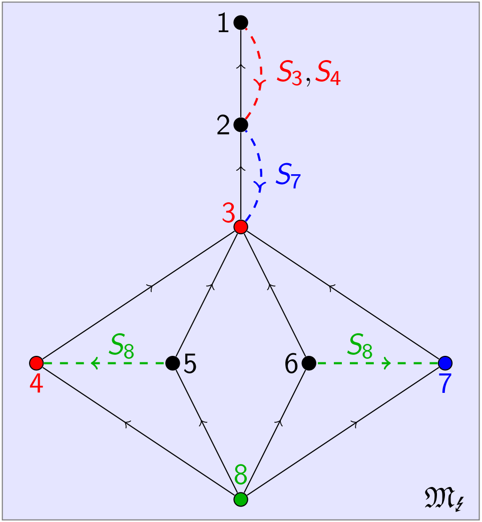
$$S_7 := (\geq_3) \cup \{(2,3)\} \quad S_8 := \{(5,4),(6,7)\} \cup P \cup (=_7)$$

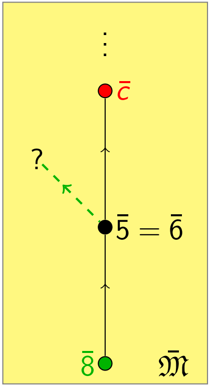
Notacija: $[i \cdot\cdot\, j] := \{i, i+1, \ldots, j\} \quad (R_n) := R \cap [1 \cdot\cdot\, n]^2$

Čini se kompliciranim, i *jest* komplicirano ☺, ali ne toliko: mnoge strelice su forsirane definicijom Veltmanovog okvira.

Na sljedećem slajdu nacrtane su samo neforsirane strelice.

Univerzalni okvir za *IL*
kontraprimjer — dijagram





Potpuna karakterizacija svjetova konačne dubine

Propozicija

Za svaki $w \in W$ postoji formula

$$\xi_w := \chi_{\rho(w)} \wedge \bigwedge_{\substack{\varphi \in S \\ w \Vdash \varphi}} \varphi \wedge \bigwedge_{\substack{\varphi \in S \\ w \nVdash \varphi}} \neg\varphi, \quad \text{gdje je } S := D_{\rho(w)}(\emptyset),$$

koja ga u potpunosti karakterizira: bilo koji svijet v je modalno ekvivalentan s w ako i samo ako $v \Vdash \xi_w$.

Doista, prvi konjunkt odabire samo svjetove na dubini $\rho(w)$, a onda ostali konjunkti određuju koji točno tip svijeta (od konačno mnogo njih na toj dubini) predstavlja w.

Ovdje su nam potrebne samo ξ_1 to ξ_8 — ali zapravo, čitava teorija karakterističnih formula može se primijeniti na bilo koji svijet **konačne dubine** bilo kojeg okvira.

Modalne (ne)ekvivalencije

4 i 7 nisu modalno ekvivalentni: $4 \Vdash \xi_1 \rhd \xi_2$, a $7 \nVdash \xi_1 \rhd \xi_2$.

5 i 6 **jesu** modalno ekvivalentni: imaju iste sljedbenike i S.

- Jedinstvenost reprezentanta u univerzalnom okviru: $\bar{5}$ i $\bar{6}$ je jedan te isti svijet u $\bar{\mathfrak{M}}$; označimo ga s t.
- $\xi_5 \Leftrightarrow \xi_6$, pa imaju zajedničkog $D_3(\emptyset)$-ekvivalenta $\bar{\xi_5} = \bar{\xi_6}$; označimo je s ψ.

Također, ψ ne vrijedi ni na kojem drugom svijetu u $\mathfrak{M}_{\not{t}}$ (osim 5 i 6): jedini koji dolaze u obzir (dubine 3) su 4 i 7, a njih eliminiraju formule $\neg(\xi_1 \rhd \xi_2)$ i $\neg(\xi_2 \rhd \xi_3)$ redom.

Konfiguracija

iz perspektive svijeta 8

Na svijetu 8 vrijede formule

- χ_4 (formula koja kaže da je 8 na dubini 4);
- $\psi \rhd (\xi_4 \vee \xi_7)$, budući da sljedbenik od 8 na kojem vrijedi ψ mora biti ili 5 ili 6, a svaki od njih ima S_8-susjeda (4 odnosno 7) na kojem vrijedi $\xi_4 \vee \xi_7$;

ali **ne** vrijede formule

- $\psi \rhd \xi_4$, jer $8 \mathrel{R} 6 \Vdash \psi$, ali 4 (jedini svijet u W na kojem vrijedi ξ_4) nije S_8-susjed od 6;
- $\psi \rhd \xi_7$ (sasvim analogno, jer $8 \mathrel{R} 5 \not\mathrel{S}_h 7$).

Tvrdimo da je takva konfiguracija nemoguća u $\bar{\mathfrak{M}}$: dakle, $\bar{8}$ ne može postojati.

Zbog $\bar{8} \nVdash \psi \rhd \xi_4$ postoji R-sljedbenik od 8 na kojem vrijedi ψ, ali na čijem nijednom $S_{\bar{8}}$-susjedu ne vrijedi ξ_4. Međutim, jer je $\psi \in D_3(\emptyset)$ karakteristična formula (svjetova 5 i 6), taj sljedbenik mora biti $\bar{5} = \bar{6} = t$.

Zaključujemo: $t \; S_{\bar{8}} \; u$ povlači $u \nVdash \xi_4$, i analogno $u \nVdash \xi_7$.

Ali $\bar{8} \Vdash \psi \rhd (\xi_4 \vee \xi_7)$ znači da
svaki R-sljedbenik od $\bar{8}$ na kojem vrijedi ψ (specijalno t), mora imati $S_{\bar{8}}$-susjeda u na kojem vrijedi $\xi_4 \vee \xi_7$.
Po prethodnom odlomku, takav u ne može postojati.

Prijedlozi za daljnja istraživanja

- Ključni problem je jedinstvenost reprezentanta u univerzalnom okviru. Možemo li oslabiti to svojstvo i još uvijek dobiti nešto korisno?
  - Možda, ali ne možemo ga potpuno maknuti, jer onda nemamo jedinstvenost samog univerzalnog okvira.
- Postoje brojna proširenja logike *IL*.
  Neka od njih, kao što je *ILF*, podudaraju se s *GL* u svom zatvorenom fragmentu [Hájek–Švejdar], i zato sigurno **imaju** univerzalni okvir. Gdje je granica?
- Postoje druge semantike za *IL*:
  među ostalim *generalizirana Veltmanova semantika* [Verbrugge–Vuković–Mikec].
  Postoji li univerzalni **generalizirani** Veltmanov okvir?
- Karakteristične formule su stravično duge (tetracija)!
  Postoje li polinomno duge karakteristične formule?
  - Vjerojatno [Perkov].