



ALGEBARSKI PRISTUP PROBLEMU ZADOVOLJENJA OGRANIČENJA AN ALGEBRAIC APPROACH TO THE CONSTRAINT SATISFACTION PROBLEM

Kristina Asimi, *Fakultet Tehničkih nauka, Novi Sad*

Oblast - MATEMATIKA U TEHNICI

Kratak sadržaj - *U ovom radu bavićemo se jednim NP-kompletnim problemom - problemom zadovoljenja ograničenja (CSP), koji se određenim restrikcijama može svesti na P problem.*

Ključne reči: Problem zadovoljenja ograničenja, Klonovi operacija

Abstract - *In this paper, we consider an NP-complete problem - Constraint Satisfaction Problem (CSP), which can be restricted to a P problem.*

Keywords: Constraint satisfaction problem, Clones

1. UVOD

Jedan od najpoznatijih nerešenih problema u matematici i računarstvu jeste takozvani "P versus NP problem". Problem traži odgovor na pitanje da li svaki problem čije rešenje može biti provereno u polinomnom vremenu može biti rešen u polinomnom vremenu. Naime, klasa problema za koje postoji algoritam koji u polinomnom vremenu dovodi do rešenja naziva se *klasa P*, a klasa problema za koje odgovor može biti proveren u polinomnom vremenu zove se *klasa NP*. Pitanje je, dakle, da li je $P = NP$. Potkласa klase *NP* je klasa *NP-kompletnih problema*. *NP*-kompletan problem je *NP* problem na koji se može svesti svaki drugi *NP* problem. Uzmimo za primer sudoku - igru u kojoj je data delimično popunjena mreža brojeva i čiji je cilj da se mreža popuni poštujući određena pravila. Ako nam neko ponudi rešenje sudoku problema, mi ćemo ga lako proveriti, ali će nam biti potrebno mnogo više vremena da dođemo do rešenja. Dakle, sudoku je *NP* problem, ali čini se da ne spada u klasu *P* problema.

Problemom zadovoljenja ograničenja mogu se izraziti mnogi kombinatorni problemi. Grubo rečeno, to je problem u kojem je data kolekcija ograničenja (uslova) definisanih nad skupom promenljivih i cilj je dodeliti vrednosti tim promenljivim tako da uslovi budu zadovoljeni.

Sudoku je jedan primer problema zadovoljenja ograničenja. Data je kvadratna mreža 9x9, što je ukupno 81 polje. Ona je podeljena na manje kvadratne mreže 3x3. Svakom polju treba dodeliti broj iz skupa $A = \{1, 2, \dots, 9\}$ (s tim da su neka polja već popunjena) tako da se u svakoj koloni svaki broj iz skupa A pojavljuje tačno jednom, u svakom redu se svaki broj iz skupa A pojavljuje tačno jednom i u svakom manjem kvadratu se svaki broj iz A pojavljuje tačno jednom.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bila dr Jovanka Pantović, red. prof.

				9			
8				7			3
2				4	8	9	7 1
						1	8
				6	1	7	9
						3	5
4	5	3	6				
9			5		2		
1	6	2		4	7	5	

Više o igri sudoku i sličnim problemima može se naći u [1]. Paradigma problema zadovoljenja ograničenja je veoma opšta i uključuje mnoge poznate probleme iz računarstva i matematike. U poslednjih 20 godina problem zadovoljenja ograničenja je privukao ogromnu pažnju u teorijskom računarstvu. Glavni razlog za to jeste taj što problem zadovoljenja ograničenja balansira između opštosti i strukture: predstavlja odgovarajući formalni model za brojne računarske pojave, a opet svaki pojedinačni model sadrži dovoljno podataka za detaljnu analizu i dubok uvid u pojavu koju opisuje.

2. PROBLEM ZADOVOLJENJA OGRANIČENJA

Sada dajemo standardnu formalnu definiciju instance problema zadovoljenja ograničenja nad konačnim domenom [2].

Definicija 1 *Instanca problema zadovoljenja ograničenja je uređena trojka $P = (V, D, \mathcal{C})$, gde je*

- V konačan skup promenljivih,
- D konačan domen,
- \mathcal{C} konačan skup ograničenja (uslova) oblika $C = (x, R)$ pri čemu je za neki prirodan broj n
 - $x = (x_1, \dots, x_n) \subseteq V^n$ (polje ograničenja),
 - $R \subseteq D^n$ (relacija ograničenja).

Kažemo da funkcija dodeljivanja $f : V \rightarrow D$ zadovoljava ograničenje $C = ((x_1, \dots, x_n), R)$ ako $(f(x_1), \dots, f(x_n)) \in R$. Funkcija $f : V \rightarrow D$ je rešenje instance problema zadovoljenja ograničenja ako zadovoljava sva ograničenja.

Cilj je, dakle, naći preslikavanje skupa promenljivih u dati domen tako da budu zadovoljeni svi postavljeni uslovi. Instanca problema zadovoljenja ograničenja razmatra se sa tri zanimljiva aspekta:

- (i) Odlučivanje. Da li data instanca ima rešenje? Problem koji je usko povezan sa ovim problemom jeste *problem traženja rešenja*, tj. problem koji se sastoji u konstruisanju nekog rešenja ako postoji bar jedno rešenje.
- (ii) Optimizacija. Ako instanca nema rešenje, naći optimalno dodeljivanje, tj. ono koje zadovoljava maksimalan broj uslova. Proučavaju se i *algoritmi optimizacije*, čiji je cilj, na primer, naći dodeljivanje takvo da zadovoljava bar 80% od broja uslova zadovoljenih optimalnom dodelom.
- (iii) Prebrojavanje. Koliko rešenja ima data instanca? Ovaj problem takođe ima aproksimativnu verziju: *aproksimativno prebrojavanje*.

Problem zadovoljenja ograničenja možemo definisati i nad fiksiranim skupom relacija.

Definicija 2 Jezik ograničenja \mathcal{D} je konačan skup relacija nad zajedničkim konačnim domenom D . Sa $CSP(\mathcal{D})$ označavamo restrikciju generalnog CSP problema na instance gde je domen D i sve relacije ograničenja su iz \mathcal{D} .

Godine 1993. Feder i Vardi¹ su formulisali Hipotezu o dihotomiji problema zadovoljenja ograničenja.

Hipoteza dihotomije. Za svaki konačan jezik ograničenja \mathcal{D} problem $CSP(\mathcal{D})$ je u P ili je NP -kompletan. U ono vreme dve stvari su podržavale hipotezu: Šeferova teorema o dihotomiji za sve jezike nad dvoelementnom domenu i Helova i Nešetrilova² teorema o dihotomiji za jezike koji se sastoje od jedne binarne simetrične relacije.

Feder i Vardi su uvideli da su poznati polinomni slučaji povezani sa svojstvima algebarskog zatvaranja i postavili pitanje da li polinomna rešivost za CSP može uvek biti objašnjena na taj način. To su potvrdili Dževons, Koen i Hissens, i ovi i sledeći radovi bazirani na ovoj povezanosti sa algebrrom odneli su oblast na drugi nivo, koji verovatno ne bi bio dostignut samo sa kombinatoričkim alatom.

Ipak, Hipoteza se još uvek ne smatra teoremom. Objavljeno je nekoliko dokaza, ali je za neke od njih ustavljeno da su netačni. Međutim, dva dokaza (jedan je objavio Bulatov, a drugi Žuk³ [4]) su u fazi provere od strane eksperata iz uže naučne oblasti i za sada nisu opovrgnuti.

3. PRIMERI

Ilustrovaćemo definiciju na nekim opštim i konkretnim primerima instanci problema zadovoljenja ograničenja.

3.1. SAT problemi

Prvi problem za koji je dokazano da je NP -kompletan jeste problem zadovoljenja iskazne formule ili skraćeno *SAT* (od engl. satisfiability). Problem zadovoljenja iskazne formule je pitanje da li za datu iskaznu formulu postoje vrednosti promenljivih za koje je formula tačna.

Neka je data iskazna formula ϕ . Svaka iskazna formula može se zapisati u konjunktivnoj normalnoj formi

$$\phi = X_1 \wedge \cdots \wedge X_n$$

¹Moshe Vardi (1954-) izraelski matematičar

²Jaroslav Nešetřil (1946-) češki matematičar

³Dmitriy Zhuk (1984-) ruski matematičar

gde je $n \geq 1$, a $X_i, i = 1, \dots, n$ su klauze tj. disjunkcije literalu [3].

Instanca problema zadovoljenja formule ϕ jeste $(V, \{0, 1\}, \mathcal{C})$, gde je V skup svih promenljivih koje se pojavljuju u klauzama $X_i, i = 1, \dots, n$, a $\mathcal{C} = \{(x_1, R_1), \dots, (x_n, R_n)\}$. Uslovi (x_k, R_k) , $k = 1, \dots, n$, su dati na sledeći način. Za svaku klauzu X_k , gde su x_1^k, \dots, x_j^k promenljive koje se pojavljuju u X_k , neka je $x_k = (x_1^k, \dots, x_j^k)$ i $R_k = \{0, 1\}^{j \setminus \{(a_1, \dots, a_j)\}}$ gde je $a_i = 1$ ako je promenljiva x_i^k negirana u X_k , a u suprotnom $a_i = 0$ (to jest, R_k je skup svih j -torki koje čine klauzu X_k tačnom). Rešenja ovog problema zadovoljenja ograničenja su baš valvacije u kojima je formula ϕ tačna.

Specijalan slučaj ovog problema jeste *3-SAT* problem, kod kojeg svaka klauza sadrži tačno 3 literala. Na primer, neka je

$$A = (x_1 \vee x_2 \vee \neg x_3) \wedge (x_3 \vee x_4 \vee x_5).$$

Pitanje je da li možemo promenljivama x_1, x_2, x_3, x_4, x_5 dodeliti vrednosti iz skupa $\{0, 1\}$ tako da formula A bude tačna. Drugim rečima, tražimo valvaciju

$$v : \{x_1, x_2, x_3, x_4, x_5\} \rightarrow \{0, 1\}$$

sa osobinom

$$\begin{aligned} (v(x_1), v(x_2), v(x_3)) &\in \{0, 1\}^3 \setminus \{(0, 0, 1)\} \text{ i} \\ (v(x_3), v(x_4), v(x_5)) &\in \{0, 1\}^3 \setminus \{(0, 0, 0)\} \end{aligned}$$

Možemo proveriti da za valvaciju v za koju je $v(x_1) = v(x_2) = v(x_3) = v(x_4) = v(x_5) = 1$ formula A ima vrednost 1, dakle zadovoljiva je.

3-SAT je ekvivalentan problemu $CSP(\mathcal{D}_{3SAT})$, gde je $D_{3SAT} = \{0, 1\}$ i

$$\mathcal{D}_{3SAT} = \{S_{ijk} : i, j, k \in \{0, 1\}\},$$

pri čemu je $S_{ijk} = \{0, 1\}^3 \setminus \{(i, j, k)\}$.

Uopšteno, za svaki prirodan broj k *k-SAT* označava sličan problem gde je svaka klauza disjunkcija k literalu. Kako je *3-SAT* NP -kompletan, sledi da je *k-SAT* NP -kompletan za svako $k \geq 3$. S druge strane, *2-SAT* je rešiv u polinomnom vremenu.

3.2. Sistemi linearnih jednačina

Neka je domen skup $\mathbb{Z}_3 = \{0, 1, 2\}$ i neka je dat sledeći sistem linearnih jednačina

$$\begin{aligned} x + z &= 1 \\ x + y - z &= 0 \\ 2x - y + z &= 0 \end{aligned}$$

gde su sabiranje i množenje po modulu 3. Instanca ovog problema je

$$(\{x, y, z\}, \mathbb{Z}_3, \{x + z = 1, x + y - z = 0, 2x - y + z = 0\})$$

Treba svakoj od promenljivih x, y, z dodeliti neki broj iz skupa $\{0, 1, 2\}$ tako da budu zadovoljene sve tri jednakosti. Ovaj problem ima rešenje. Naime, $(x, y, z) = (0, 1, 1)$ zadovoljava zadati sistem linearnih jednačina.

3.2. Bojenje grafa

Dat je graf (V, E) . Neka je $V = \{v_1, \dots, v_m\}$ skup čvorova i svaku granu iz E koja spaja čvorove v_i i v_j obeležimo sa e_{ij} . Za fiksiran prirodan broj k , k -bojenje je dodeljivanje boja iz skupa $D_k = \{0, 1, \dots, k-1\}$ čvorovima grafa tako da susedni čvorovi budu obojeni različitim bojama. Postavlja se pitanje da li je takvo bojenje moguće. Instanca ovog problema je

$$(V, D_k, \{C_{ij} : e_{ij} \in E\})$$

$$C_{ij} = ((v_i, v_j), \rho_k), \quad \rho_k = \{(a, b) \in D_k^2 : a \neq b\}$$

Ako je moguće izvršiti k -bojenje na grafu, kažemo da je graf k -obojiv. Specijalno, ako je graf 2-obojiv, to znači da je bipartitan (skup čvorova V se može podeliti na dva disjunktna skupa V_1 i V_2 tako da je $V = V_1 \cup V_2$ pri čemu unutar grafa V_i , $i = 1, 2$, nema grana). Ovaj problem je NP -kompletan za $k \geq 3$. Za $k = 2$ (tj. odlučivanje da li je dati graf bipartitan) ovaj problem je rešiv u polinomnom vremenu.

3. KLONOVI I PP-FORMULE

Neka je A konačan skup sa bar dva elementa. Za svako $n \geq 1$ označimo sa $O_A^{(n)}$ skup svih n -arnih operacija na A , to jest, $O_A^{(n)} = A^{A^n}$, a sa O_A označimo skup svih finitarnih operacija na A , to jest, $O_A = \bigcup_{n \geq 1} O_A^{(n)}$. Arnost operacije f označavamo sa $ar(f)$. Za skup nekih finitarnih operacija $F \subseteq O_A$ uvodimo oznaku $F^{(n)} = F \cap O_A^{(n)}$. Označimo sa π_i^n i -tu n -arnu projekciju:

$$\pi_i^n(x_1, \dots, x_i, \dots, x_n) = x_i,$$

a sa Π_A skup svih projekcija svih konačnih arnosti skupa A .

Neka su $f \in O_A^{(n)}$ i $g_1, g_2, \dots, g_n \in O_A^{(m)}$. Kompozicija operacija f i g_1, \dots, g_n je operacija $h = f(g_1, g_2, \dots, g_n) \in O_A^{(m)}$ definisana na sledeći način:

$$h(x_1, x_2, \dots, x_m) = f(g_1(x_1, x_2, \dots, x_m), g_2(x_1, x_2, \dots, x_m), \dots, g_n(x_1, x_2, \dots, x_m)).$$

Definicija 3 Skup $C \subseteq O_A$ se naziva klon operacija na A ako važi

$$(i) \quad \Pi_A \subseteq C, \quad i$$

$$(ii) \quad \text{ako } f \in C^{(n)} \text{ i } g_1, g_2, \dots, g_n \in C^{(m)}, \text{ onda i } h = f(g_1, g_2, \dots, g_n) \in C.$$

Više o klonovima može se naći u [5].

Definicija 4 Neka su \mathcal{D} i \mathcal{E} jezici ograničenja na istom domenu $D = E$. Kažemo da \mathcal{D} pp-definiše \mathcal{E} ako svaka relacija iz \mathcal{E} može biti definisana formulom prvog reda koja koristi samo relacije iz \mathcal{D} , relaciju jednakosti, konjunkciju i egzistencijalnu kvantifikaciju.

Ova terminologija potiče iz teorije modela, gde se formula prvog reda naziva *primitive* ako se može zapisati u obliku $\exists y \bigwedge_{i < n} \alpha_i(x, y)$ gde je za sve $i < n$ $\alpha_i(x, y)$ literal (atomička formula ili negacija atomičke formule). Primitive formula koja ne sadrži negaciju naziva se *primitive pozitivna* (pp-) formula.

Bez korišćenja matematičke logike, gornju definiciju možemo preformulisati na sledeći način. \mathcal{D} pp-definiše \mathcal{E} ako \mathcal{D} i \mathcal{E} imaju isti domen i svaka relacija $R_i \in \mathcal{E}$ se može predstaviti konstrukcijom koja koristi relacije iz \mathcal{D} kao što sledi. Postoji instanca P_i od $CSP(\mathcal{D} \cup \{=\})$ i podskup X_i skupa promenljivih iz P_i takvi da skup svih rešenja od P_i , kada se projektuje na X_i , daje baš relaciju R_i .

Teorema 5 Ako \mathcal{D} pp-definiše \mathcal{E} , onda se $CSP(\mathcal{E})$ može svesti na $CSP(\mathcal{D})$.

4. POLIMORFIZMI I CSP

Za $m > 0$ m -arna relacija ρ na A je podskup od A^m . Označimo sa $R_A^{(m)}$ skup svih m -arnih relacija na A , tj. $R_A^{(m)} = \mathcal{P}(A^m)$, a sa R_A skup svih finitarnih relacija na A , tj. $R_A = \bigcup_{m \geq 1} R_A^{(m)}$. Za skup $Q \subseteq R_A$ uvodimo oznaku $Q^{(m)} = Q \cap R_A^{(m)}$.

Uvodimo vezu između relacija i operacija preko relacije kompatibilnosti.

Definicija 6 Operacija $f \in O_A^{(n)}$ je kompatibilna sa relacijom $\rho \in R_A^{(m)}$ (f je polimorfizam od ρ ili ρ je invarijantno za f) ako

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \in \rho, \dots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \in \rho \Rightarrow \begin{bmatrix} f(a_{11}, \dots, a_{1n}) \\ f(a_{21}, \dots, a_{2n}) \\ \vdots \\ f(a_{m1}, \dots, a_{mn}) \end{bmatrix} \in \rho.$$

Ekvivalentno, ako je (a_{ij}) matrica dimenzija $n \times m$ čiji su redovi elementi iz ρ , onda f primenjena na kolone daje m -torku koja je takođe u ρ . Drugim rečima, f je polimorfizam od ρ ako f primenjena komponentno na bilo koju n -torku elemenata iz ρ daje element iz ρ .

Primer 1 Ako je \leq relacija porekla na skupu A , onda „ f je polimorfizam od \leq “ znači da je f monotona u odnosu na \leq . Na primer, sabiranje je polimorfizam za standardnu relaciju \leq .

Polimorfizam indukuje Galoaovu vezu [6] između operacija i relacija, a odgovarajući operatori $\overrightarrow{\rho}$ i $\overleftarrow{\rho}$ se obično označavaju sa $PolQ$ i $InvF$ za $F \subseteq O_A$ i $Q \subseteq R_A$, tj.

$$InvF = \{\rho \in R_A | \rho \text{ je invarijantno za svaku operaciju iz } F\}; \\ PolQ = \{f \in O_A | f \text{ je polimorfizam svake relacije iz } Q\}.$$

Teorema 7 (Bondarčuk, Kalužnin⁴, Kotov, Romov 1969) Neka je $C \subseteq O_A$. Tada su sledeća tvrdjenja ekvivalentna.

$$(1) \quad C \text{ je klon.}$$

$$(2) \quad C = PolQ \text{ za neko } Q \subseteq R_A.$$

$$(3) \quad C = PolInvC.$$

Teorema 8 Neka su \mathcal{D} i \mathcal{E} jezici ograničenja nad istim domenom $D = E$. Tada \mathcal{D} pp-definiše \mathcal{E} ako i samo ako je $Pol\mathcal{D} \subseteq Pol\mathcal{E}$.

⁴Lev Arkad'evich Kalužnin (1914-1990)

Kombinovanjem ove teoreme i teoreme 5 dobijamo da kompleksnost problema $CSP(\mathcal{D})$ zavisi samo od klena $Pol\mathcal{D}$. Preciznije, ako je $Pol\mathcal{D} \subseteq Pol\mathcal{E}$, onda se $CSP(\mathcal{E})$ može svesti na $CSP(\mathcal{D})$. Povrh toga, dokaz prethodne teoreme nam daje generičku pp-definiciju jezika \mathcal{E} iz \mathcal{D} , koja nam daje generičku redukciju problema $CSP(\mathcal{E})$ na $CSP(\mathcal{D})$.

Primer 2 Podsetimo se 3-SAT problema. Njegov jezik je $\mathcal{D}_{3SAT} = \{S_{ijk} : i, j, k \in \{0, 1\}\}$, gde je $S_{ijk} = \{0, 1\}^3 \setminus \{(i, j, k)\}$ za sve $i, j, k \in \{0, 1\}$. Jasno, sve projekcije su polimorfizmi jezika \mathcal{D}_{3SAT} , međutim, to su i svi polimorfizmi tog jezika. Na osnovu prethodne teoreme, odatle sledi da \mathcal{D}_{3SAT} pp-definiše svaki jezik ograničenja sa domenom $\{0, 1\}$.

Primer 3 I za jezik $\mathcal{D}'_{3COLOR} = \{\neq_3, C_0, C_1, C_2\}$ na domenu $\{0, 1, 2\}$ su jedini polimorfizmi projekcije. Odatle sledi da \mathcal{D}'_{3COLOR} pp-definiše svaku relaciju na $\{0, 1, 2\}$. Pokažimo kako dokaz prethodne teoreme proizvodi pp-definiciju neke relacije, recimo, ternarne relacije

$$R = \{(0, 1, 1), (0, 2, 1)\}.$$

Kako R sadrži 2 elementa, pp-definišemo 3^2 -arnu relaciju

$$S = \{(f(0, 0), f(0, 1), \dots, f(2, 2)) : f \text{ je binarni polimorfizam jezika } \mathcal{D}'_{3COLOR}\}$$

koja odgovara skupu svih binarnih polimorfizama jezika \mathcal{D}'_{3COLOR} :

$$\begin{aligned} S(x_{00}, x_{01}, \dots, x_{22}) \text{ ako i samo ako} \\ \bigwedge_i (x_{ii} = i) \wedge \bigwedge_{i_1 \neq i_2, j_1 \neq j_2} (x_{i_1 j_1} \neq_3 x_{i_2 j_2}). \end{aligned}$$

Proverimo da li formula zaista definiše S . Jasno je da indeksi pokazuju kako interpretirati svaku 3^2 -orku iz S kao binarnu operaciju na $\{0, 1, 2\}$. Prvi deo konjunkcije govori da je interpretirana operacija kompatibilna sa C_0, C_1 i C_2 , dok drugi deo govori da je kompatibilna sa \neq_3 . Sada egzistencijalno kvantifikujemo svaku promenljivu iz S osim x_{00}, x_{12} i x_{11} - izuzeci su one promenljive čiji indeksi odgovaraju prvoj, drugoj i trećoj (respektivno) koordinati trojki iz R . Dobijena ternarna relacija $R'(x_{00}, x_{12}, x_{11})$ sadrži sve

elemente iz R , jer S sadrži dve torke koje odgovaraju slučajevima kad je f projekcija $\pi_i^2 : \{0, 1, 2\}^2 \rightarrow \{0, 1, 2\}$, i svi elementi iz R' su sadržani u R , jer je relacija R kompatibilna sa svakim polimorfizmom jezika \mathcal{D}'_{3COLOR} .

5. ZAKLJUČAK

Zaključak je da kompleksnost problema $CSP(\mathcal{D})$ zavisi samo od polimorfizama jezika \mathcal{D} .

6. LITERATURA

- [1] Yato T., Seta T., *Complexity and Completeness of Finding Another Solution and Its Application to Puzzles*, IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E86-A No.5 pp.1052-1060, May 2003
- [2] Barto L., Krokin A., Willard R., *Polymorphisms, and How to Use Them*, <http://drops.dagstuhl.de/opus/volltexte/2017/6959/pdf/DFU-Vol7-15301-1.pdf>
- [3] Janićić P., *Matematička logika u računarstvu*, Beograd, 2008.
- [4] Zhuk D., *The Proof of CSP Dichotomy Conjecture*, <https://arxiv.org/abs/1704.01914>
- [5] Aichhinger E., *Basics of clone theory*, <http://www.algebra.uni-linz.ac.at/Students/UniversalAlgebra/s11/clonebasics2.pdf>
- [6] Börner F., *Basics of Galois connections*, Complexity of Constraints, volume 5250 of Lecture Notes in Computer Science, Springer, 2008.

Kratka biografija:



Kristina Asimi rođena je u Zrenjaninu 1994. god. Master rad na Fakultetu tehničkih nauka iz oblasti Matematika u tehnici odbranila je 2018. god.