Shlukování založené na fuzzy ekvivalenci

Vilém Vychodil

21. května 2002

Abstrakt

Tento text obsahuje jemný úvod do jedné základní shlukovací metody – shlukování založeném na vlastnostech fuzzy ekvivalence. Všechny potřebné základní pojmy jsou shrnuty v úvodním paragrafu textu, u čtenáře jsou předpokládání základní znalosti pojmů relace, zobrazení a podobně. Tento text vznikl z důvodu podpory vývoje analyzátoru podobností zdrojových kódů popsaných gramatikou. Autor může být kontaktován prostřednictvím elektronické pošty na adrese <vilem.vychodil@upol.cz>.

Residuovaný svaz

V celém textu pracujeme s úplným residuovaným svazem $\mathcal{L} = \langle [0,1], \vee, \wedge, \otimes, \rightarrow, 0, 1 \rangle$, to jest v úplném residuovaném svazu s nosičem jímž je reálný interval $[0,1]$, spojení a průseku odpovídají $\vee = \max$ a $\wedge = \min$. Dvojice $\langle \otimes, \rightarrow \rangle$ tvoří adjungovaný pár, to jest pro libovolné $a, b, c \in [0,1]$ platí

$$a \otimes b \leq c \quad \text{právě když} \quad a \leq b \rightarrow c. \tag{1}$$

Podmínka (1) se nazývá adjunkce. Operace $\otimes$ se nazývá součin nebo též t-norma a platí, že $\langle L, \otimes, 1 \rangle$ je komutativní monoid. Každá t-norma určuje právě jedno residuum, pro které je podmínka (1) splněna. Pokud by totiž $\langle \otimes, \rightarrow_1 \rangle$, $\langle \otimes, \rightarrow_2 \rangle$ byly dva adjungované páry, potom by ze vztahu (1) pro libovolná $a, b, c \in [0,1]$ platilo

$$a \leq b \rightarrow_1 c \quad \text{právě když} \quad a \otimes b \leq c \quad \text{právě když} \quad a \leq b \rightarrow_2 c, \tag{2}$$

Jelikož ale platí $b \rightarrow_1 c \leq b \rightarrow_1 c$, potok ze vztahu (2) plyne $b \rightarrow_1 c \leq b \rightarrow_2 c$. Stejně tak ze vztahu $b \rightarrow_2 c \leq b \rightarrow_2 c$ plyne $b \rightarrow_2 c \leq b \rightarrow_1 c$, to jest platí $\rightarrow_1 = \rightarrow_2$. T-norma a residuum jsou v úplném residuovaném svazu jinými slovy vzájemně jednoznačně určeny.

Definice 1. Na $[0,1]$ definujeme tři základní residuované svazy následovně.

1. Residuovaný svaz $\langle [0,1], \vee, \wedge, \otimes, \rightarrow, 0, 1 \rangle$, kde

$$a \otimes b = \max(a + b - 1, 0), \tag{3}$$

$$a \rightarrow b = \min(1 - a + b, 1), \tag{4}$$

se nazývá *ȷ ukasiewiczova algebra.*

2. Residuovaný svaz $\langle [0,1], \vee, \wedge, \otimes, \rightarrow, 0, 1 \rangle$, kde

$$a \otimes b = a \wedge b, \tag{5}$$

$$a \rightarrow b = \begin{cases} 1 & \text{když } a \leq b, \\ b & \text{jinak,} \end{cases} \tag{6}$$

se nazývá *Gödelova algebra.*

3. Residuovaný svaz $\langle [0,1], \vee, \wedge, \otimes, \rightarrow, 0, 1\rangle$, kde

$$a \otimes b = a \cdot b, \tag{7}$$

$$a \rightarrow b = \begin{cases} 1 & \text{když } a \leq b, \\ \frac{b}{a} & \text{jinak,} \end{cases} \tag{8}$$

se nazývá *produktová algebra*.

Všechny tři výše uvedené t-normy jsou navíc spojité a libovolnou jinou spojitou t-normu lze definovat právě pomocí nich. Důkaz tohoto tvrzení je však netriviální a je zcela mimo rozsah tohoto textu. Kromě základních operací residuovaného svazu bývá dobré definovat i unární operaci negace $\neg a = a \rightarrow 0$ a binární operaci biresiduum $a \leftrightarrow b = (a \rightarrow b) \wedge (b \rightarrow a)$. V dalším textu jsou shrnuty základní fuzzy množinové pojmy.

Fuzzy množina a α-řez

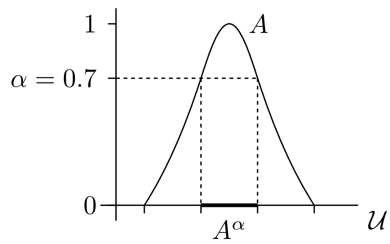
Nechť $\mathcal{L} = \langle [0,1], \vee, \wedge, \otimes, \rightarrow, 0, 1\rangle$ je úplný residuovaný svaz. $\mathcal{L}$-množinou neboli fuzzy množinou je libovolné zobrazení $A\colon U \rightarrow [0,1]$, kde U je universum prvků. Zobrazení A se nazývá $\mathcal{L}$-množinou v universu U. Operace s fuzzy množinami lze definovat přirozeně pomocí operací residuovaného svazu $\mathcal{L}$. Například sjednocení, průnik a rozdíl fuzzy množin $A\colon U \rightarrow [0,1]$, $B\colon U \rightarrow [0,1]$ lze definovat předpisy

$$(A \cup B)(x) = A(x) \vee B(x), \tag{9}$$

$$(A \cap B)(x) = A(x) \wedge B(x), \tag{10}$$

$$(A \setminus B)(x) = A(x) \otimes \neg B(x). \tag{11}$$

Pro fuzzy množinu $A\colon U \rightarrow [0,1]$ lze pro libovolné $\alpha \in [0,1]$ definovat α-řez, to jest podmnožinu $A^\alpha \subseteq U$, $A^\alpha = \{x;\, x \in U, \alpha \leq A(x)\}$. Řez tedy není $\mathcal{L}$-množinou, nýbrž bivalentní podmnožinou universa prvků, které mají v A příslušnost alespoň α. Graficky je α-řez znázorněn na následujícím obrysu fuzzy množiny.



Pro α-řez evidentně platí $\alpha \leq \beta$, právě když $A^\beta \subseteq A^\alpha$.

Fuzzy tolerance a ekvivalence

Pokud je universum tvořeno kartézským součinem či kartézskou mocninou, pak hovoříme o $\mathcal{L}$-relaci respektive o fuzzy relaci. Je-li například U množina, pak je $A\colon U \times U \rightarrow [0,1]$ binární fuzzy relace. Každé dvojici $\langle x, y\rangle \in U \times U$ je zobrazením A přidělena míra příslušnosti z intervalu $[0,1]$. U binárních fuzzy relací můžeme přirozeně zavádět vlastnosti jako v případě bivalentních binárních relací.

Relace tolerance a ekvivalence je možné vymezit dvěma, respektive třemi axiomy,

$$(\forall x)(\mathfrak{E}(x,x)), \tag{12}$$
$$(\forall x)(\forall y)(\mathfrak{E}(x,y) \Rightarrow \mathfrak{E}(y,x)), \tag{13}$$
$$(\forall x)(\forall y)(\forall z)(\mathfrak{E}(x,y) \otimes \mathfrak{E}(y,z) \Rightarrow \mathfrak{E}(x,z), \tag{14}$$

kde $\mathfrak{E}$ označuje binární relační symbol. Axiomy (12)–(14) mají samozřějmě i svoji fuzzy sémantickou interpretaci. Předpokládejme nyní, že relační symbol $\mathfrak{E}$ je realisován binární fuzzy relací $E\colon U \times U \to [0,1]$ a spojky $\otimes$, $\Rightarrow$ jsou realisovány operacemi residuovaného svazu $\otimes$ a $\to$. Navíc $x \to y = 1$ právě když $1 \leq x \to y$, což je z adjunkce právě když $1 \otimes x \leq y$, tedy právě když $x \leq y$. To jinými slovy znamená, že $\|\varphi \Rightarrow \psi\| = 1$ právě když $\|\varphi\| \leq \|\psi\|$. Předchozí axiomy lze tedy interpretovat následovně. Pro libovolné $x, y, z \in U$ musí platit

$$E(x,x) = 1, \tag{15}$$
$$E(x,y) = E(y,x), \tag{16}$$
$$E(x,y) \otimes E(y,z) \leq E(x,z), \tag{17}$$

kde $E(x,y)$ je zjednodušení značení místo $E(\langle x,y\rangle)$. Každá fuzzy relace splňující (15), (16) se nazývá *fuzzy tolerance*, každá relace splňující (15)–(17) se nazývá *fuzzy ekvivalence.* Následující tvrzení charakterisují vztah fuzzy tolerancí a některých fuzzy ekvivalencí k jejich α-řezům.

Lemma 1. (O isotonii součinu) *Nechť $\mathcal{L} = \langle [0,1], \vee, \wedge, \otimes, \to, 0, 1\rangle$ je libovolný residuovaný svaz. Potom je operace $\otimes$ isotonní, to jest pro libovolné $a, b, c, d \in [0,1]$, $a \leq b$, $c \leq d$ platí $a \otimes c \leq b \otimes d$.*

Důkaz. Mějme $a \leq b$, $c \leq d$. Z reflexivity $\leq$ plyne $b \otimes d \leq b \otimes d$, to jest $b \leq d \to (b \otimes d)$. Jelikož $a \leq b$, platí i $a \leq d \to (b \otimes d)$, což platí právě když $a \otimes d \leq b \otimes d$ to jest právě když $d \leq a \to (b \otimes d)$. Z $c \leq d$ plyne $c \leq a \to (b \otimes d)$, to jest $a \otimes c \leq b \otimes d$. Což bylo dokázat. □

Věta 1. (O řezech fuzzy tolerance) *Binární fuzzy relace $E : U \times U \to [0,1]$ je fuzzy tolerance, právě když je každý její α-řez relace tolerance, to jest právě když je E^α bivalentní tolerance pro libovolné $\alpha \in [0,1]$.*

Důkaz. „⇒“ Jelikož pro fuzzy ekvivalenci platí $E(x,x) = 1$ pro libovolné x, pak rovněž $\langle x,x\rangle \in E^\alpha$ pro libovolné $\alpha \in [0,1]$. Navíc jelikož platí $E(x,y) = E(y,x)$ pro libovolná $x, y \in [0,1]$, relace E^α je evidentně symetrická. Každá E^α je tedy relace tolerance, což bylo dokázat.

„⇐“ Nyní předpokládejme, že E^α je bivalentní ekvivalence pro libovolné $\alpha \in [0,1]$. Postupně dokážeme platnost vztahů (15)–(17). Jelikož je dle předpokladu každá E^α reflexivní, $\langle x,x\rangle \in E^\alpha$ pro každé $\alpha \in [0,1]$, tedy i pro $\alpha = 1$. To jest pro $x \in U$ je $E(x,x) = 1$. Fuzzy relace je reflexivní. Předpokládejme $\langle x,y\rangle \in E^\alpha$ implikuje $\langle y,x\rangle \in E^\alpha$. To jest když $E(x,y) \leq \beta$, pak $E(y,x) \leq \beta$. Z $E(x,y) \leq E(x,y)$ potom plyne $E(y,x) \leq E(x,y)$. Stejnou úvahou lze dojít k opačné nerovnosti. Dohromady $E(x,y) = E(y,x)$ pro libovolné $x, y \in U$. Relace je symetrická a v důsledku fuzzy tolerance, což bylo dokázat. □

Věta 2. (O řezech fuzzy ekvivalence) *Uvažujme Gödelovu strukturu, to jest residuovaný svaz $\mathcal{L} = \langle [0,1], \vee, \wedge, \otimes, \to, 0, 1\rangle$, kde $\otimes = \wedge$. Pak binární fuzzy relace $E\colon U \times U \to [0,1]$ je $\mathcal{L}$-ekvivalence, právě když je každý její α-řez bivalentní relací ekvivalence na $U \times U$.*

Důkaz. Vzhledem k faktům presentovaným ve větě 1. stačí dokázat pouze transitivitu. Důležitým poznatkem je, že $\otimes$ je v případě Gödelovy struktury idempotentní.

„$\Rightarrow$“ Nechť je $\alpha \in [0,1]$ libovolné a platí $\langle x,y\rangle \in E^\alpha$, $\langle y,z\rangle \in E^\alpha$. Jelikož $E(x,y) \geq \alpha$ a zároveň $E(y,z) \geq \alpha$, z isotonie plyne $E(x,y) \otimes E(y,z) \geq \alpha \otimes \alpha$. Z idempotence $\otimes$ dále plyne $E(x,y) \otimes E(y,z) \geq \alpha$. Vzhledem k tomu, že platí (17), musí být i $E(x,z) \geq \alpha$. Odtud již přímo dostáváme $\langle x,z\rangle \in E^\alpha$, což bylo dokázat.

„$\Leftarrow$“ Označme $E(x,y) = \alpha$, $E(y,z) = \beta$. Stačí ukázat, že $\alpha \otimes \beta \leq E(x,z)$. Ale z lemmy o isotonii a z neutrality 1 plyne, že $\alpha \otimes \beta \leq \alpha$, $\alpha \otimes \beta \leq \beta$. To ale znamená, že $\langle x,y\rangle \in E^{\alpha\otimes\beta}$, $\langle y,z\rangle \in E^{\alpha\otimes\beta}$. Z transitivity $E^{\alpha\otimes\beta}$ dostáváme $\langle x,z\rangle \in E^{\alpha\otimes\beta}$. Potom ale $\alpha \otimes \beta \leq E(x,z)$, což bylo dokázat. □

Transitivní uzávěr fuzzy relace

Binární fuzzy relace lze přirozeným způsobem skládat. Máme-li fuzzy relace $A\colon U \times U \to [0,1]$, $B\colon U \times U \to [0,1]$, složením fuzzy relací A, B vznikne fuzzy relace $A \circ B\colon U \times U \to [0,1]$, kde

$$(A \circ B)(x,z) = \bigvee_{\forall y \in U} A(x,y) \otimes B(y,z). \tag{18}$$

Označíme nyní symbolem $\subseteq$ fuzzy množinovou inklusi. Pro fuzzy množiny A, B platí $A \subseteq B$, pokud pro každé $x \in U$ platí $A(x) \leq B(b)$. Inkluse má zajímavý vztah vzhledem ke skládání reflexivních fuzzy relací. Pokud je A reflexivní fuzzy relace, pak pro $B(x,y) = \alpha$ platí $B(x,y) \otimes B(y,y) = \alpha \otimes 1 = \alpha$. To jest $\langle x,y\rangle$ má v $A \circ B$ míru příslušnost alespoň α, lze tudíž konstatovat $B \subseteq A \circ B$.

Pomocí skládání lze definovat i mocniny binární relace. Položíme-li $R^1 = R$, $R^n = R^{n-1} \circ R$, pak můžeme R^n, pro $n \in \mathbb{N}$ nazvat n-tou mocninou fuzzy relace. Je-li R reflexivní, pak $R \subseteq R^n$ pro libovolné $n \in \mathbb{N}$. Navíc v případě reflexivní relace evidentně nedochází ke snožování žádného stupně příslušnosti. Algoritmus mocnění relace lze implementovat efektivně, jelikož $R^n = R^{\frac{n}{2}} \circ R^{\frac{n}{2}}$ pro n sudé a $R^n = R \circ R^{n-1}$ pro n liché.

Při aplikaci fuzzy ekvivalencí při shlukové analýze se lze velmi často dostat do situace, kdy máme k disposici pouze relaci fuzzy tolerance. Pokud není relace transitivní, budeme ji v našem případě muset nahradit příbuznou relací, která transitivní je. Nabízí se možnost reflexivní a symetrickou relaci transitivně uzavřít. Transitivním uzávěrem fuzzy relace R vznikne nejmenší relace T taková, že $R \subseteq T$ a T je transitivní.

Věta 3. (O transitivním uzávěru) *Nechť $R\colon U \times U \to [0,1]$ je reflexivní fuzzy relace definovaná na konečném universu prvků $|U| = n$. Pak relace $Q = R^{n-1}$ je fuzzy kvaziuspořádání a platí $R \subseteq Q$.*

Důkaz. Jelikož je relace reflexivní, podle předchozího výkladu platí $R^i \subseteq R^{i+1}$, pro libovolné $i \in \mathbb{N}$ speciálně pak $R \subseteq R^{n-1}$. Stačí tedy ověřit, že relace R^{n-1} je transitivní.

Z definice skládání matic plyne, že $R^{n-1}(x,y)$ je vypočteno jako supremum ze součinů měr příslušnosti hran nacházejících se ve všech $(n-1)$-prvkových sledech mezi vrcholy x a y. Zároveň každý sled $\langle x_0,x_1\rangle, \langle x_1,x_2\rangle, \ldots, \langle x_{k-1},x_k\rangle$ o $k \leq n-1$ prvcích lze díky reflexivitě relace doplnit na sled $\langle x_0, x_1 \rangle, \langle x_1, x_2 \rangle, \ldots, \langle x_{k-1}, x_k \rangle, \langle x_k, x_k \rangle, \langle x_k, x_k \rangle \ldots$ délky $n-1$ a díky neutralitě 1 platí

$$R^{n-1}(x_0, x_1) \otimes R^{n-1}(x_1, x_2) \otimes \cdots \otimes R^{n-1}(x_{k-1}, x_k) =$$
$$R^{n-1}(x_0, x_1) \otimes R^{n-1}(x_1, x_2) \otimes \cdots \otimes R^{n-1}(x_{k-1}, x_k) \otimes R^{n-1}(x_k, x_k) \otimes R^{n-1}(x_k, x_k) \otimes \cdots$$

Uvažujme nyní $R^{n-1}(x, y) = \alpha$, $R^{n-1}(y, z) = \beta$. Hodnota $R^{n-1}(x, z)$ je supremem součinů měr příslušnosti hran nacházejících se ve všech $(n-1)$-prvkových sledech mezi vrcholy x a z. To jest $\alpha \otimes \beta$ vyjadřuje součin hodnot jednoho z možných sledů, tento sled má ale $2n-2$ vrcholů. Jelikož je relace R n-prvková, lze v něm tedy najít opakující se vrchol. Odtud dostáváme, že $\alpha \otimes \beta \leq R^{n-1}(x, z)$. Relace $Q = R^{n-1}$ je kvaziuspořádání. □

Shlukování založené na fuzzy ekvivalenci

Samotné shlukování je založeno na vytváření řezů fuzzy tolerance, nebo fuzzy ekvivalence. Pro konečnou množinu prvků $|U| = n$ je nejprve vytvořena fuzzy relace R, která representuje jejich vzájemnou podobnost. Čím je $R(x, y)$ větší, tím jsou si x a y podobnější. Relace R by měla být přirozeně reflexivní a symetrická. Každý prvek je sám sobě podobný s největší mírou příslušnosti a vzájemná podobnost dvou prvků je stejná. Relace R je tedy obecně fuzzy tolerance.

Relace R však nemusí být obecně transitivní. Při shlukové analýze se buďto pracuje pouze s fuzzy tolerancí, nebo se provádí její transitivní uzávěr, viz větu 3. Transitivním uzavřením sice obdržíme fuzzy ekvivalenci, v některých případech však může dojít k nepříjemnému efektu – výsledná relace je dost nepodobná relaci výchozí. Při vytváření transitivního uzávěru je rovněž potřeba vhodně volit operaci součinu.

Třetím problémem je vytváření řezů. Řezy $\mathcal{L}$-ekvivalence, kde $\mathcal{L}$ je Gödelova struktura jsou ekvivalence. Při použití jiné struktury tomu tak již být nemusí. Východiskem ze situace je používání tolerancí místo ekvivalencí – pomocí řezů lze potom vytvářet normální pokrytí příslušná těmto tolerancím. Normální pokrytí již není obecně disjunktní, ale zachovává některé důležité vlastnosti.

Příklad 1. Mějme fuzzy relaci R, S na množině $U = \{a, b, c, d\}$. Míry příslušnosti jsou zadány následujícími tabulkami,

R	a	b	c	d
a	1	0.9	0.8	0.85
b	0.9	1	0.75	0.9
c	0.8	0.75	1	0.8
d	0.85	0.9	0.8	1

S	a	b	c	d
a	1	0.9	0.8	0.9
b	0.9	1	0.8	0.9
c	0.8	0.8	1	0.8
d	0.9	0.9	0.8	1

$S^{0.9}$	a	b	c	d
a	1	1	0	1
b	1	1	0	1
c	0	0	1	0
d	1	1	0	1

Fuzzy relace R je ekvivalence vzhledem k produktové struktuře. Například ale vzhledem ke Gödelově struktuře není, jelikož například $\min(0.9, 0.9) \not\leq 0.85$. Relace S vznikla transitivním uzavřením relace R vzhledem k $\otimes = \wedge$, to jest $S = R^3$. Poslední matice symbolizuje α-řez matice S pro $\alpha = 0.9$. Řez $S^{0.9}$ je ekvivalence a $U/S^{0.9} = \{\{a, b, d\}, \{c\}\}$.

Reference

[1] Klir, G. J. – Yuan, B. *Fuzzy Sets and Fuzzy Logic. Theory and Applications.* Prentice Hall P T R, Upper Saddle River, NJ 07458, 1995. (pages 357–378).