KYBERNETIKA ČÍSLO 5, ROČNÍK 3/1967

GUHA - metoda systematického vyhledávání hypotéz II

PETR HÁJEK, IVAN HAVEL, METODĚJ CHYTIL

Práce je pokračováním článku [1], otištěného v Kybernetice v r. 1966, na nějž přirozeně navazuje. Původní algoritmus metody GUHA je zobecněn — hledají se a hodnotí i hypotézy „skoro" pravdivé. Zobecnění a úpravy metody jsou prováděny s cílem jejího přiblížení konkrétním přírodovědeckým úvahám.

V práci [1] jsme podali informativní popis a matematickologické zdůvodnění výzkumné metody GUHA (General Unary Hypotheses Automaton), jakož i návod k realizaci na samočinném počítači a několik prvních zkušeností. Připomeňme čtenáři, že metoda předpokládá, že je dán model (výsledek experimentu) sestávající z konečného počtu m objektů a n vlastností a z údajů, které objekty které vlastnosti mají a které nemají. Cílem metody je na základě teoretických výsledků matematické logiky a praktických možností samočinných počítačů získat automaticky systematický výčet všech informací jistého druhu o daném modelu, heslem řečeno, všechno zajímavé. Zásadně je nutno podotknout, že za zajímavé se nepovažují pouze vztahy dvou ze sledovaných vlastností, ale vztahy libovolného počtu vlastností. Ukazuje se však, že druhému slovu uvedeného hesla („zajímavé") lze dát různé významy. V [1] byly za zajímavé označeny jen takové vztahy sledovaných vlastností, které (kromě dalších podmínek) odpovídají formulím výrokového počtu pravdivým v daném modelu. Přitom však ve čtvrté části práce byly naznačeny tři problémy a), b), c) jakožto možnosti dalšího rozvíjení metody. V práci [2] jsou uvedeny bez důkazů kromě výsledků obsažených v [1] některé další poznatky, odpovídající částečně na zmíněné problémy b) a c), tj. problém rozdílu platnosti dvou syntakticky stejných hypotéz pravdivých v modelu, ale lišících se četnostmi kombinací nul a jedniček, a problém významu hypotéz v modelu sice nepravdivých, ale splňovaných jistým „dostatečným" množstvím $p\%$ objektů (např. 90% nebo jinak dle volby výzkumníka).

Cílem článku, který má čtenář v rukou, je vyložit doplňující poznatky i s důkazy a také se zmínit o realizaci zobecněného projektu na samočinném počítači. Čtenáři,

který se zajímá pouze o možnost aplikace metody a nikoli o její teoretické zdůvodnění, postačí, obeznámí-li se s definicemi 1, 2, větou 1 (bez důkazu), poznámkou za ní a s dikusí o problému implikace (až k větě 2). Připomínáme takovému čtenáři, že je-li p proměnná označující nějakou vlastnost, označuje $\bar{p}$ (negace této proměnné) vlastnost, kterou má nějaký objekt právě tehdy, když nemá vlastnost označenou p. Disjunkce proměnných $p_1, \ldots, p_s$ (označená $p_1 \vee \ldots \vee p_s$) označuje vlastnost, kterou má nějaký objekt právě tehdy, když má alespoň jednu z vlastností označených proměnnými $p_1, \ldots, p_s$. Podobně i v případě, že některé proměnné jsou s negacemi. Vyšetřujeme-li proměnné $p_1, \ldots, p_n$, vyšetřujeme současně s nimi i $\bar{p}_1, \ldots, \bar{p}_n$; tvoříme disjunkce, v nichž se každé z písmen $p_1, \ldots, p_n$ vyskytne nejvýše jednou (s negací nebo bez negace). Takové disjunkce se nazývají elementární. Značka disjunkce je $\vee$; např. $p_1 \vee \bar{p}_2 \vee p_4$ je elementární disjunkce označující vlastnost, kterou má nějaký objekt právě tehdy, když je pro něj pravda alespoň jedna z těchto tří věcí: a) má vlastnost označenou p_1, b) nemá vlastnost označenou p_2, c) má vlastnost označenou p_4. Libovolnou vlastnost, která je (v jistém smyslu) kombinací původních vlastností, můžeme zachytit nějakou formulí výrokového počtu; formule jsou tedy zápisy složených vlastností.

***

Na rozdíl od práce [2], kde je logická teorie podána v rámci predikátového počtu, přidržujeme se tedy zde výkladu v rámci počtu výrokového soudíce, že obě možné verze mají své přednosti i nevýhody, nehledě na to, že rozdíl obou je nepatrný. Vlastnosti modelu jsou označeny výrokovými proměnnými $p_1, \ldots, p_n$; v [1] nás zajímaly tzv. prosté implicenty formule $\Phi_{\mathcal{M}}$ ($\Phi_{\mathcal{M}}$ je charakteristická formule modelu), neboli – jak říkáme v [2] – prosté disjunkce modelu $\mathcal{M}$. Přidržíme se názvu „prosté disjunkce modelu $\mathcal{M}$". V [1] je ukázáno, že elementární disjunkce D je prostou disjunkcí modelu $\mathcal{M}$ právě tehdy, když (a) je pravdivá v $\mathcal{M}$, přičemž (b) žádná disjunkce, vzniklá z D vynecháním některých členů, není pravdivá v $\mathcal{M}$. Zobecníme nyní pojem pravdivosti formule a prosté disjunkce modelu a dokážeme větu analogickou větě 19 z [1].

Definice 1. Buď p racionální číslo (v dalším pevně zvolené), buď $0 < p < 1$. (Např. $p = 0{,}9$ nebo $p = 0{,}95$ atd.) Buď A formule, buď $\mathcal{M}$ model o počtu prvků m. Buď m_A počet těch prvků modelu $\mathcal{M}$, které splňují A. Řekneme, že formule A je *skoro pravdivá* v $\mathcal{M}$, jestliže $p \leqq m_A/m < 1$ (tj. bylo-li např. $p = 0{,}95$, je A skoro pravdivá v $\mathcal{M}$, jestliže alespoň 95% objektů splňuje A a přitom ovšem A není úplně pravdivá v $\mathcal{M}$, tj. aspoň jeden objekt nesplňuje A).

Definice 2. Elementární disjunkce A se nazývá *skoro prostou disjunkcí* modelu $\mathcal{M}$, jestliže

1. A je skoro pravdivá v $\mathcal{M}$, přičemž
2. žádná disjunkce vzniklá z A vyškrtnutím některého členu už v $\mathcal{M}$ skoro pravdivá není.

Věta 1. *Buď $\mathscr{M}$ model, A formule. Je-li A skoro pravdivá v $\mathscr{M}$, pak je důsledkem jistých prostých a skoro prostých disjunkcí modelu $\mathscr{M}$. Detailněji: Jestliže A obsahuje pouze proměnné $p_{k_1}, \ldots, p_{k_h}$, a je skoro pravdivá v $\mathscr{M}$, pak je důsledkem jistých prostých a skoro prostých disjunkcí modelu $\mathscr{M}$ obsahujících pouze proměnné $p_{k_1}, \ldots, p_{k_h}$.*

Důkaz. Skoropravdivost je definována racionálním číslem p, $0 < p < 1$. Definovali jsme, že formule je skoro pravdivá, když ji splňuje alespoň $p' = p \,.\, 100$ procent objektů (ale ne všechny). Odstraňme z modelu $\mathscr{M}$ všechny objekty, které nesplňují A (těch není víc než $(100 - p')\%$), tím vznikne model $\mathscr{M}'$, v němž je A pravdivé a tudíž dle věty 19 z [1] důsledkem konjunkce jistých prostých disjunkcí modelu $\mathscr{M}'$, což jsou rozhodně skoro pravdivé, nebo dokonce pravdivé disjunkce modelu $\mathscr{M}$. Nechť to jsou disjunkce $D_1, \ldots, D_j$. Z každé elementární disjunkce pravdivé (skoro pravdivé) v $\mathscr{M}$ však lze vyškrtáváním členů dospět k prosté (skoro prosté) disjunkci modelu $\mathscr{M}$ (nejméně o jednom členu), nechť to jsou disjunkce $D_1^*, \ldots, D_j^*$. Zřejmě každá ohvězdičkovaná disjunkce implikuje příslušnou neohvězdičkovanou, tudíž konjunkce všech ohvězdičkovaných implikuje konjunkci $D_1 \,\&\, \ldots \,\&\, D_j$, která je však ekvivalentní formuli A.

Poznámka. Tvrzení obrácené k větě 1 neplatí (na rozdíl od věty 19 z [1]); např. každá z formulí p_1, p_2 může být splněna 90% objektů, ale konjunkce $p_1 \,\&\, p_2$ může být splněna méně než 90% objektů a tudíž (pro $p = 0{,}9$) nebýt skoro pravdivá v $\mathscr{M}$.

V [1] (část III) je definován pojem *sondy*; elementární disjunkci můžeme nazývat zajímavou, jsou-li (v závislosti na sondě) splněny podmínky 1–4 z [1] str. 43.

Na základě věty 1 je účelné definovat obecněji úkol počítače takto:

(I) generovat všechny zajímavé elementární disjunkce (kratší dříve než delší) a tisknout ty z nich, které jsou prosté nebo skoro prosté.

Problém implikace. Jak již bylo zmíněno, elementární disjunkce jsou ekvivalentní jistým implikacím. Použitím mj. toho, že

$$A_1 \vee A_2 \text{ je ekvivalentní } \bar{A}_1 \to A_2 \,,$$

$$\overline{(A_1 \vee A_2)} \text{ je ekvivalentní } \bar{A}_1 \,\&\, \bar{A}_2 \,,$$

$$\bar{\bar{A}}_1 \text{ je ekvivalentní } A_1 \,,$$

(kde A_1, A_2 jsou libovolné formule), můžeme ke každé elementární disjunkci s nejméně dvěma komponentami nalézt různé ekvivalentní implikace tvaru $K \to D$, kde K je elementární konjukce (definovaná zřejmým způsobem) a D je elementární disjunkce. Takové implikace mohou být kombinovány užitím toho, že

$$(K_1 \to D) \,\&\, (K_2 \to D) \text{ je ekvivalentní } (K_1 \vee K_2) \to D \,,$$

$$(K \to D_1) \,\&\, (K \to D_2) \text{ je ekvivalentní } K \to (D_1 \,\&\, D_2) \,.$$

Tímto způsobem může být mechanicky odvozena řada důsledků (skoro) prostých disjunkcí. Podle věty 19 z [1] je každý důsledek prostých disjunkcí pravdivý v modelu. Srovnejme však příklad uvedený v tabulce 1.

Tabulka 1.

p_1	p_2	I	II	III	IV
1	1	90	1	87	3
1	0	0	0	8	9
0	1	1	90	2	8
0	0	9	9	3	80

Předpokládejme, že model má 100 objektů; v případě I 90 z nich (má p_1 a má p_2), pro žádný neplatí, že (má p_1 a nemá p_2) atd. V případech I a II je disjunkce $\bar{p}_1 \vee p_2$ prostá; v případech III a IV je skoro prostá (pro $p = 0{,}9$). Implikace $p_1 \rightarrow p_2$ je tedy pravdivá resp. skoro pravdivá. To znamená, že nejsou žádné objekty (resp. je jen málo objektů), které mají p_1 ale nemají p_2.

Je však přirozené klást si tyto dvě otázky: 1. Kolik je objektů, které mají p_1? Kolik z nich má také p_2? První otázka může být nazvána otázkou signifikance*) nalezené implikace, druhá otázkou relativní pravdivosti (nebo relativní pravděpodobnosti) implikace.

V případech I, II je druhá otázka zodpovězena uspokojivě (a platí to obecně pro pravdivé implikace): všechny objekty, které mají p_1, mají i p_2. Ale první otázka je zodpovězena různě: V případě I je „dost" objektů majících p_1, v případě II je jich „málo". To samé platí v případech III, IV, ale otázka relativní pravdivosti je nyní zodpovězena různě: V případě III je 95 objektů majících p_1, 87 z nich má p_2, tj. více než 91%. V případě IV je 12 objektů majících p_1 a jen tři z nich mají p_2, tj. 25%. Tedy ne všechny implikace ekvivalentní prostým a skoro prostým disjunkcím jsou „dobré" vzhledem k našim dvěma otázkám. Zavedeme proto tuto definici:

Definice 3. Implikace $A \rightarrow B$ je *relativně (skoro) pravdivá* v modelu $\mathcal{M}$, jestliže B je (skoro) pravdivá v modelu složeném ze všech objektů modelu $\mathcal{M}$ splňujících A. (tj. jestliže relativní četnost objektů splňujících B vůči objektům splňujícím A je 1 (nebo alespoň p)).

Věta 2. *Každá relativně (skoro) pravdivá implikace je důsledkem jistých prostých (prostých a skoro prostých) disjunkcí.*

Důkaz. Ukážeme, že každá relativně pravdivá implikace je také pravdivá a že

* Upozorňujeme, že tento pojem signifikance není totožný s pojmem signifikance běžným ve statistice.

každá relativně skoro pravdivá implikace je také skoro pravdivá v modelu. Mějme implikaci $A \to B$, označme m_{11} počet objektů splňujících A i B, m_{01} počet objektů nesplňujících A a splňujících B, podobně m_{10}, m_{00}. Je-li $A \to B$ relativně pravdivá, znamená to, že $m_{10} = 0$. To ale též znamená, že $A \to B$ je pravdivá v celém původním modelu. Je-li $A \to B$ relativně skoro pravdivá, znamená to, že $p \leqq m_{11}/(m_{11} + m_{10}) < 1$ ($m_{11} + m_{10}$ je počet objektů modelu tvořeného všemi objekty splňujícími A, m_{11} je počet těch z nich, které splňují $A \to B$). Počet všech objektů celého modelu je $m_{11} + m_{10} + m_{01} + m_{00}$, počet těch z nich, které splňují $A \to B$, je $m_{11} + m_{01} + m_{00}$. Zřejmě

$$\frac{m_{11}}{m_{11} + m_{10}} \leqq \frac{m_{11} + m_{01} + m_{00}}{m_{11} + m_{10} + m_{01} + m_{00}} < 1$$

a tudíž je $A \to B$ skoro pravdivá. Tvrzení naší věty tedy plyne z věty 1 a z věty 19 v [1].

Tedy prosté a skoro prosté disjunkce obsahují „celou pravdu" i v tomto modifikovaném pojetí, co se týče relativní pravdivosti implikace. Musíme však zkoušet, které implikace ekvivalentní (skoro) prostým disjunkcím jsou relativně (skoro) pravdivé. To je další úkol počítače. Zvolíme pevné číslo s (menší než počet všech objektů v modelu) a definujeme:

Definice 4. Budiž $K \to D$ implikace (K elementární konjunkce a D elementární disjunkce) ekvivalentní jisté prosté nebo skoro prosté disjunkci (značme ji A) modelu $\mathcal{M}$. Antecedent K je dobrý vzhledem k A, jestliže 1. alespoň s objektů modelu $\mathcal{M}$ splňuje K (tj. K je signifikantní), 2. v případě, že A je skoro pravdivá, je implikace $K \to D$ relativně skoro pravdivá.

(Je-li A pravdivá, pak 1. je jediný požadavek; v tomto případě už plyne fakt, že $K \to D$ je relativně pravdivá, z faktu, že A je pravdivá.)

Má-li elementární disjunkce A j členů ($j \geqq 2$), pak je $2^j - 2$ implikací $K \to D$ ekvivalentních A. Řekneme, že konjunkce K_1 je částí konjunkce K_2 (nebo je obsažena v K_2), jestliže každý člen konjunkce K_1 je i členem K_2.

Věta 3. *Je-li K dobrý antecedent disjunkce A, pak každá elementární konjunkce, která je částí K, je také dobrým antecedentem disjunkce A.*

Důkaz. Nechť K je ekvivalentní elementární konjukci $K_1 \,\&\, K_2$, nechť K_1 vzniká z K vyškrtnutím těch komponent, které tvoří K_2. Nechť A je ekvivalentní implikaci $(K_1 \,\&\, K_2) \to D$. Pak je také ekvivalentní implikaci $K_1 \to (\bar{K}_2 \vee D)$. Je-li antecedent $K_1 \,\&\, K_2$ dobrý, znamená to, že 1. formuli $K_1 \,\&\, K_2$ splňuje alespoň s objektů a 2. v případě, že A je skoro pravdivá, je $(K_1 \,\&\, K_2) \to D$ relativně skoro pravdivá. Pak tím spíše formuli K_1 splňuje alespoň s objektů. Pokud je tedy A pravdivá, víme hned, že K_1 je dobrý antecedent. Je-li A skoro pravdivá, musíme ještě ověřit, že implikace $K_1 \to (\bar{K}_2 \vee D)$ je skoro pravdivá, což se provede podobným odhadem, jako v důkazu věty 2.

Řekneme, že K je maximální dobrý antecedent disjunkce A, je-li to její dobrý antecedent a přitom není obsažen v žádném větším dobrém antecedentu disjunkce A.

Počítač má tedy druhý úkol:

(II) Ke každé prosté a skoro prosté disjunkci nalézt všechny její maximální dobré antecedenty.

Zdá se být užitečné volit $s \leqq 1/(1 - p)$. (Např. pro $p = 0{,}9$, $s \leqq 10$, pro $p = 0{,}95$, $s \leqq 20$ atd.) Pak platí následující věta:

Věta 4. *Buď A prostá disjunkce, K některý odpovídající antecedent. K je dobrý právě tehdy, když je signifikantní. (To je zřejmé.) Buď A skoro pravdivá disjunkce, K některý odpovídající antecedent. K je dobrý tehdy a jen tehdy, když příslušná implikace je relativně skoro pravdivá.*

Důkaz. Nechť A je skoro pravdivá disjunkce ekvivalentní implikaci $K \to D$. Máme se přesvědčit, že splňuje-li K méně než s objektů, pak $K \to D$ není relativně skoro pravdivá. Nechť m_{11} objektů splňuje $K \& D$, m_{10} objektů splňuje $K \& \bar{D}$. Dle předpokladu víme, že $m_{11} + m_{10} < s$, $s \leqq 1/(1 - p)$, $m_{10} \geqq 1$ (neboť A není pravdivá). Označme $m_{11} + m_{10} = t$, pak

$$\frac{m_{11}}{m_{11} + m_{10}} \leqq \frac{t - 1}{t} < \frac{s - 1}{s} \leqq p ;$$

$K \to D$ není relativně skoro pravdivá.

***

Dva vytčené úkoly pro počítač vyžadují nepříliš velkou změnu strojového programu (oproti programu popsanému blokovým schématem v [1]); zejména na generování elementárních disjunkcí se nic nemění. I v nové verzi jsou možné úspory „přeskakováním" některých úseků v uspořádané množině elementárních disjunkcí nebo alespoň tím, že v jistých úsecích stačí verifikovat, zda je pracovní disjunkce prostá, protože není-li pravdivá, nemůže (v těchto úsecích) být skoro prostá (srv. diskusi o disjunkci $(2^{(k)})$, [1] str. 42).

Byly vypracovány, odladěny a jsou používány dva programy pro novou verzi, program v jazyce FORTRAN IV vyzkoušený na počítači IBM 7040 ve Vídni a program pro MINSK 22.

Na závěr podotkněme, že podaná koncepce není zdaleka jediným možným zobecněním původní verze; autoři této práce se sami zabývají možností zobecnění v jiném směru. Základním programem zůstává: získat z modelu prostředky matematické logiky a počítacích strojů (vedle případných dalších prostředků) všechno zajímavé. Bude tedy lépe rozumět zkratce GUHA – General Unary Hypotheses Automaton – jako obecný automat na unární hypotézy.

(Došlo dne 23. února 1967.)

[1] P. Hájek, I. Havel, M. Chytil: GUHA — metoda systematického vyhledávání hypotéz. Kybernetika *2* (1966), 1, 31—47.
[2] P. Hájek, I. Havel, M. Chytil: The GUHA method of automatic hypotheses determination. Computing *1* (1966), 4.

SUMMARY

The GUHA Method of Systematical Hypotheses Searching II

PETR HÁJEK, IVAN HAVEL, METODĚJ CHYTIL

This paper is a continuation of [1]. It contains a description of a generalization of the GUHA-method, which arose from the attempts to solve the problems formulated there.

Again, the question of automatic generation and verification of interesting hypotheses is dealt with. The computer „offers" such hypotheses to an experimenter, who, of course, has to supply it with an original data, so called model. A model is a finite nonempty system of objects and a finite system of properties; it is known, for each object and each property, whether or not the object possesses the property.

The necessary logical formalization is carried out on the basis of the propositional calculus. In addition to the formulas true in a model $\mathscr{M}$ also the „almost true" formulas (i.e. the ones fulfilled by a certain sufficient number of objects) are looked for (cf. Definition 1). In this way there is generalized also the notion of a prime implicent of $\Phi_{\mathscr{M}}$ (which is called prime disjunction in this paper):

Definition. An elementary disjunction A is called an almost prime disjunction of a model $\mathscr{M}$ if (1) A is almost true in $\mathscr{M}$ and (2) no elementary disjunction obtained by omitting some components in A is almost true in $\mathscr{M}$.

Theorem. *Let $\mathscr{M}$ be a model, A — formula.* 1. *If A is almost true in $\mathscr{M}$ then it is (logically) implied by a conjunction of some prime and almost prime disjunctions of $\mathscr{M}$.* 2. *Let A consist only of the variables $p_{k_1}, \ldots, p_{k_h}$ and be almost true in $\mathscr{M}$. Then A is the logical consequence of the conjunction of prime and almost prime disjunctions of $\mathscr{M}$ containing only the variables $p_{k_1}, \ldots, p_{k_h}$.*

This theorem and the similar one from [1] make possible to formulate the task of the computer as follows: to generate all the interesting elementary disjunctions and print the prime and almost prime ones (for the notion of an „interesting disjunction" cf. [1]).

Further, there is investigated the problem of implications logically equivalent to prime (almost prime) disjunctions. Two different implications which are equivalent to the same prime disjunction may differ from one another as for the number of objects fulfilling their antecedents, in the case of almost prime disjunction also as for their relative truthfulness (cf. Definition 3).

We define the notion of „a good antecedent" of a prime (almost prime) disjunction. The computer has to find all the maximum good antecedents of a given (almost) prime disjunction.

The generalization was motived by an endeavour to make the method more convenient for the natural sciences experimenters needs. (It was clear, that the aspects of the theory of probability should have been included). Certainly, the generalization described is not the only possible. The authors, however, consider the main program of the method, namely „to obtain from the model by means of mathematical logic and computer technique (and maybe by some other means) everything interesting" to be useful.

Petr Hájek, Ivan Havel, Matematický ústav ČSAV, Žitná 25, Praha 1; Metoděj Chytil, Matematické oddělení Fyziologického ústavu ČSAV, Rudé armády 319, Praha 8 — Kobylisy.