Petr Hájek; Ivan Havel; Metoděj Chytil

GUHA - metoda systematického vyhledávání hypotéz

Kybernetika, Vol. 2 (1966), No. 1, (31)--47

Persistent URL: http://dml.cz/dmlcz/125332

KYBERNETIKA ČÍSLO 1, ROČNÍK 2/1966

GUHA - metoda systematického vyhledávání hypotéz*

PETR HÁJEK, IVAN HAVEL, METODĚJ CHYTIL

V tomto článku popisujeme metodu GUHA (automatické hledání generálních unárních hypotéz). Práce sama nepřináší žádné principiálně nové matematické poznatky. Domníváme se však, že její přínos spočívá právě ve skloubení tří věcí: formálního aparátu matematické logiky, možností, které poskytuje technika samočinných počítačů, a vyjadřovacích a metodologických prostředků exaktních věd. Metoda GUHA je aplikací uvedených tří oborů na problémy konkrétního vědeckého bádání.

Článek se dělí na čtyři části. V části I se informativně popisuje charakter metody, možnosti jejího použití a způsob interpretace výsledků, které počítač poskytuje. Matematická formulace úlohy a zdůvodnění zvoleného postupu řešení jsou obsaženy v části II. Ve třetí části popisujeme algoritmus pro řešení úlohy. Konečně v části IV se zmiňujeme o možnostech jeho realizace na počítacím stroji, uvádíme údaje o jedné z prvních praktických aplikací metody a naznačujeme problémy, které v souvislosti s dalším rozvojem metody bude třeba řešit.

Autoři děkují profesoru O. Zichovi za cennou diskusi o některých partiích práce**. Dále děkují kolektivu výpočtového střediska MINSK při SPK v Praze za nevšední ochotu a pomoc při provozním zajištění zkušebních výpočtů.

I.

Metoda GUHA byla vypracována a prvně zkušebně použita v oblasti fyziologického výzkumu, ale má všechny předpoklady, aby vnesla nové formy práce do libovolné oblasti, kde předmětem zkoumání je nějaký (konečný) systém *objektů* a nějaký (konečný) systém *vlastností*, přičemž pro každý objekt je o každé vlastnosti známo, zda ji objekt má či nemá. Objekty mohou být např. pacienti, vlastnostmi různé choroby nebo fakt, zda jim byly podány určité léky, atd. V jiném případě mohou být objekty příslušníci určité skupiny lidí a „vlastnostmi" sociální zařazení, zájmy atd.

* Předneseno na *druhé konferenci o kybernetice*, která se konala v Praze ve dnech 16.—19. listopadu 1965.

** Prof. Zich se společně s Dr. Zelenkou zabýval aplikací matematické logiky na problémy lékařských věd v práci [3]; upozorňujeme čtenáře na tuto práci.

Systém objektů a systém vlastností spolu s určením, který objekt má které vlastnosti, nazýváme *modelem*. Výzkumník má za úkol zjistit obecně platné vztahy mezi zkoumanými vlastnostmi. Přitom obvykle nemá k dispozici údaje o všech existujících objektech (např. o všech osobách trpících určitou chorobou), nýbrž jen o některých, které tvoří, řekněme, podmodel modelu všech existujících objektů. Zjistí-li, že mezi zkoumanými vlastnostmi je pro všechny pozorované objekty jistý vztah, může vyslovit (a dále ověřovat) hypotézu, že tento vztah platí pro vůbec všechny existující objekty. Cílem metody GUHA je pomoci při tvorbě těchto hypotéz.

Její podstata spočívá, zhruba řečeno, v tom, že výzkumný pracovník před začátkem zkoumání nějaké užší oblasti (např. epilepsie) si na nějakém konkrétním modelu (pokud možno reprezentativním) zjistí její pomocí *právě všechny* vztahy, které v tomto modelu platí resp. neplatí. Teprve potom vypracuje jejich individuální logickou analýzou, opřenou samozřejmě již o odborné vědomosti, jednu nebo více pracovních hypotéz, které ověřuje experimentálně. Rozdíl proti dřívějšímu stavu je hlavně ten, že výzkumný pracovník ve fázi formulace pracovní hypotézy nemá k dispozici jen ojedinělé a nepřehledné informace, ale totální přehled všech hypotéz platných ve zkoumaném modelu. Důsledkem toho je jednak skutečnost, že k experimentu přistupuje se znalostí pravděpodobného výsledku, jednak to, že obdobných, pravděpodobně úspěšných, pracovních hypotéz může na základě metody GUHA stanovit velký počet.

Nebude na újmu serióznosti výkladu, přirovnáme-li použití metody GUHA k výlovu rybníka vypuštěním. Teprve po vypuštění rybníka máme jistotu, že známe právě všechny ryby, které rybník v sobě skrývá. Dřívější stav připomíná naproti tomu rybáře s udicí, který se musí spoléhat jen na své umění podmíněné náhodností, než učiní uspokojující úlovek, aniž by přitom věděl, které další ryby a kolik jich rybník ještě v sobě skrývá. Obdobou operace "vypuštění" je v případě metody GUHA provedení mnoha milionů logických operací, které se v rámci zkoumaného modelu pomocí jistého algoritmu uskuteční na samočinném počítači.

Popíšeme nyní zhruba charakter výsledků, které počítač poskytuje, a naznačíme možnosti jejich interpretace.

Výsledky vystupují ze samočinného počítače ve formě tzv. *prostých implicent**, což jsou v jistém smyslu nejjednodušší logické formule, platné pro všechny objekty modelu, a přitom v jistém smyslu nejsilnější s touto vlastností. Neberou se tedy v úvahu vztahy triviální (např. „Jestliže má pacient vlastnost č. 1, pak má vlastnost č. 1.“) anebo vztahy, které lze ještě zesílit. Na tyto prosté implicenty lze celou „pravdu“, kterou v sobě model skrývá, beze zbytku rozložit; prosté implicenty jsou zákony platné v modelu a mohou být považovány za pracovní hypotézy pro celý zkoumaný obor. (Tím ovšem není řečeno, že každá z nich bude vždy zajímavá nebo užitečná.) Podotýkáme, že prosté implicenty mají tvar elementárních disjunkcí a mohou být podle sémantického významu jednotlivých proměnných různým způsobem transfor-

* Přesně řečeno, prostých implicent charakteristické formule modelu; viz část II.

movány např. na tvar implikace. Jednotlivé prosté implicenty lze dále skládat pomocí syntaktických pravidel matematické logiky do složených formulí (složených hypotéz). Popsané transformace je možno provádět různými způsoby podle toho, zda výzkumníka zajímají určité vlastnosti jako příčiny jiných, nebo jako důsledky jiných, či jinak.

Forma předávaných výsledků dovoluje výzkumníkovi také zodpovědět otázku, jaký vztah platí mezi určitými vybranými vlastnostmi, případně zda vůbec nějaký netriviální platí. Dále lze ve výsledcích snadno vyčíst odpověď na formálně nenáročnou otázku, zda určitá uvažovaná hypotéza (konstruovaná výzkumníkem) platí v modelu, a v případě, že neplatí, dovoluje algoritmus určit všechny možné doplňky k ní, aby se platnou stala. Zdůrazněme tedy: GUHA není metoda na zjištění pouhé platnosti či neplatnosti jednotlivých hypotéz, které už někoho napadly, nýbrž slouží buďto (a především) jako náhrada nápadů (náhrada intuice), čili „nabízení hypotéz" výzkumníkovi („výlov rybníka"), nebo jako korekce hypotézy sice neplatné nebo neúplně formulované, ale platné po určitých dodatcích, tj. „vylepšování hypotéz".

Pro úsporu času a získání přehlednějších odpovědí na právě uvedené otázky lze vyšetřování modelu rozložit na zkoumání tzv. *sond*. Sondu výzkumník předem určí tím, že označí některé proměnné, jejichž vzájemný vztah ho přednostně zajímá, případně některou označí za předpokládaný důsledek, jehož možné příčiny chce zjistit, apod. (viz část III).

Míra platnosti zjištěných výsledků. Jak bylo výše řečeno, lze prosté implicenty považovat z hlediska obecné platnosti za hypotézy, které je nutno dále dokazovat, a to buďto experimentálně nebo i další fází metody GUHA, v obou případech již za pomoci statistiky. V té souvislosti je třeba upozornit na případ, kdy nějaká formule, např. implikace, představuje hypotézu v důsledku toho, že nějakou kombinaci vlastností má třeba jen jeden pacient. Tak např. hypotéza

$$(p_{16} \& p_{27}) \rightarrow p_{20}\,,$$

která značí:

„Každý pacient, který měl abnormální porod a přitom lze u něho zabránit záchvatu v době aury soustředěním na duševní činnost, měl fokální typ záchvatu,"

byla podle sémantických pravidel matematické logiky v pokusném modelu pravdivá. To znamená, že skutečně nenajdeme v modelu pacienta, který má vlastnost p_{16} i p_{27} a přitom nemá p_{20}. Fakt, že studujeme jen prosté implicenty, zaručuje, že alespoň jeden pacient skutečně má obě vlastnosti p_{16}, p_{27} (jinak by byla prostou implicentou formule $\overline{p_{16}} \vee \overline{p_{27}}$). Přitom však byl v rámci celého modelu jen jediný pacient, který obě vlastnosti měl.

Metoda samotná neodpoví na otázku, jaký je rozdíl mezi hypotézou tvaru implikace, která se opírá o přítomnost jisté vlastnosti např. u 50% všech zkoumaných objektů, od takové, která se opírá o přítomnost této vlastnosti jen na objektu jediném. Je třeba však uvážit, že poslední hypotézu nelze zdaleka považovat za bezcennou, a to z těchto důvodů:

a) z hlediska sémantických pravidel je správná (shora uvedená věta skutečně platí pro *všechny* pacienty);

b) výskyt nějaké kombinace vlastností i v jediném případě může někdy mnoho napovědět pro tvorbu závěrečné pracovní hypotézy a směr další výzkumné práce.

Na druhé straně nutno poznamenat, že jsou myslitelná různá zdokonalení popisované metody, motivovaná snahou vědět o výsledných hypotézách víc než pouze to, že platí pro všechny objekty modelu (viz část IV).

II.

V této části podáme některé matematicko-logické definice a uvedeme některé věty (v podstatě většinou dobře známé), abychom mohli přesně formulovat pojem generální unární hypotézy a zjistit jeho vlastnosti, z čehož již snadno vyplyne obecná metoda hledání všech hypotéz. V dalším se předpokládá, že je čtenář obeznámen s pojmy a technikou výrokového počtu.

1. *Formulemi* rozumíme formule výrokového počtu. Budeme uvažovat pouze formule, které obsahují nejvýše n proměnných $p_1, \ldots, p_n$ (kde n je pevné přirozené číslo).

2. Definice. *Model* (sémantický unární model typu n) je konečná neprázdná množina M, na které je dáno n funkcí $P_i(a)$, zobrazujících množinu M do dvouprvkové množiny $\{0, 1\}$. Budeme psát

$$\mathcal{M} = \langle M, P_1, \ldots, P_n \rangle .$$

Prvky $a \in M$ nazýváme *objekty*; je-li $P_i(a) = 1$, říkáme, že objekt *a má i-tou vlastnost*. Množinu

$$K(a) = \langle P_1(a), \ldots, P_n(a) \rangle$$

nazýváme *kartou* objektu a.

Kanonický model typu n je model

$$\mathcal{M}^{(n)} = \langle M^{(n)}, P_1^{(n)}, \ldots, P_n^{(n)} \rangle ,$$

kde $M^{(n)}$ je množina všech uspořádaných n-tic prvků 0, 1, $P_i^{(n)}(\langle u_1, \ldots, u_n \rangle) = 1$ právě když $u_i = 1$.

3. Definice. Budiž dán model $\mathcal{M}$. Každé formuli $\Phi(p_1, \ldots, p_n)$ přiřadíme funkci zobrazující množinu M do množiny $\{0, 1\}$. Tuto funkci budeme nazývat *funkcí asociovanou* k Φ a $\mathcal{M}$. Nechť $u \oplus v$, $u \,.\, v$, $u \,(\rightarrow)\, v$, $u (\equiv) v$, $\bar{v}$ značí po řadě dvouhodnotové pravdivostní funkce přirazené logickým operacím, $\vee$, &, $\rightarrow$, $\equiv$, $-$ (např. $u \,(\rightarrow)\, v = 0 \equiv (u = 1) \,\&\, v = 0$ atd.). Funkci asociovanou k Φ a $\mathcal{M}$ definujeme indukcí.

a) Je-li Φ formule p_i, je $F^{\mathcal{M}}_{\Phi}(a) = P_i(a)$

b)
$$F^{\mathcal{M}}_{\Phi\vee\Psi}(a) = F^{\mathcal{M}}_{\Phi}(a) \oplus F^{\mathcal{M}}_{\Psi}(a)\,,$$
$$F^{\mathcal{M}}_{\Phi\&\Psi}(a) = F^{\mathcal{M}}_{\Phi}(a)\,.\,F^{\mathcal{M}}_{\Psi}(a)\,;$$
$$F^{\mathcal{M}}_{\Phi\to\Psi}(a) = F^{\mathcal{M}}_{\Phi}(a)\,(\to)\,F^{\mathcal{M}}_{\Phi}(a)\,,$$
$$F^{\mathcal{M}}_{\Phi\equiv\Psi}(a) = F^{\mathcal{M}}_{\Phi}(a)\,(\equiv)\,F^{\mathcal{M}}_{\Phi}(a)\,,$$
$$F^{\mathcal{M}}_{\overline{\Phi}}(a) = \overline{F^{\mathcal{M}}_{\Phi}(a)}\,.$$

Funkci asociovanou k Φ a ke kanonickému modelu $\mathcal{M}^{(n)}$ nazýváme *kanonickou funkcí asociovanou* k Φ a značíme ji F_{Φ}.

4. Definice. Budiž dán model $\mathcal{M}$; definujeme indukcí predikát „Objekt a splňuje formuli Φ".

a) Je-li Φ formule p_i, definujeme:
   a splňuje Φ, právě když $P_i(a) = 1$;

b) a splňuje formuli Φ & Ψ, právě když a splňuje Φ a a splňuje Ψ;
   a splňuje $\Phi \vee \Psi$ právě když a splňuje Φ nebo a splňuje Ψ;
   a splňuje $\Phi \to \Psi$ právě když jestliže a splňuje Φ, pak a splňuje Ψ;
   a splňuje $\overline{\Phi}$, právě když a nesplňuje Φ.

Řekněme, že formule Φ je pravdivá v modelu $\mathcal{M}$, jestliže každý objekt $a \in M$ splňuje formuli Φ.

5. Poznámka. Můžeme zavést n unárních predikátů $\mathcal{P}_i(x)$ ($\mathcal{P}_i(x)$ čteme „x má i-tou vlastnost"); naše definice pravdivosti formule je shodná s Tarského definicí pravdivosti vzhledem k formulím predikátového počtu sestrojeným z libovolné formule Φ výrokového počtu takto: každou proměnnou p_i nahradíme predikátem $\mathcal{P}_i(x)$; k takto získané formuli připojíme zpředu generální kvantifikátor $(\forall x)$. Takto sestrojené formule můžeme nazývat *generálními unárními hypotézami*. Je-li generální unární hypotéza sestrojená k formuli Φ pravdivá v modelu $\mathcal{M}$, znamená to, že pro všechny objekty modelu je mezi vlastnostmi $\mathcal{P}_i(x)$ souvislost Φ.

6. Věta. *Budiž $\mathcal{M}$ model, Φ formule, $F^{\mathcal{M}}_{\Phi}$ funkce asociovaná k Φ a $\mathcal{M}$. Formule Φ je pravdivá v modelu $\mathcal{M}$, právě když pro každý objekt $a \in M$ platí $F^{\mathcal{M}}_{\Phi}(a) = 1$.*

Důkaz. Indukcí se ověří následující tvrzení: objekt a splňuje formuli Φ, právě když $F^{\mathcal{M}}_{\Phi}(a) = 1$. Je-li Φ formule p_i, plyne to z definice 4a); např. pro formuli Φ & Ψ platí: (a splňuje (Φ & Ψ)) $\equiv$ (a splňuje Φ a a splňuje Ψ) $\equiv$ ($F^{\mathcal{M}}_{\Phi}(a) = 1$ a $F^{\mathcal{M}}_{\Psi}(a) = 1$) $\equiv F^{\mathcal{M}}_{\Phi\&\Psi}(a) = 1$.

7. Z matematické logiky (viz např. [2]) je známa následující věta („dokazatelná" znamená „dokazatelná ve výrokovém počtu"):

Věta. *Následující tři tvrzení jsou ekvivalentní:*

a) *formule Φ je dokazatelná;*
b) *formule Φ je pravdivá v libovolném modelu;*
c) *formule Φ je pravdivá v kanonickém modelu.*

8. Věta. *Formule $\Phi \to \Psi$ je pravdivá v modelu $\mathcal{M}$, právě když pro každý objekt $a \in M$ platí:*

$$F^{\mathcal{M}}_{\Phi}(a) = 1 \to F^{\mathcal{M}}_{\Psi}(a) = 1\,.$$

Formule $\Phi \equiv \Psi$ je pravdivá v modelu $\mathcal{M}$, právě když pro každý objekt $a \in M$ platí:

$$F_{\Phi}(a) = 1 \equiv F_{\Psi}(a) = 1\,.$$

(To je důsledek věty 6.)

9. Věta. *Jestliže formule Φ je pravdivá v modelu $\mathcal{M}$ a formule $\Phi \to \Psi$ je pravdivá v modelu $\mathcal{M}$, pak je i formule Ψ pravdivá v modelu $\mathcal{M}$.*

(To je důsledek věty 8.)

10. Věta.

$$F^{\mathcal{M}}_{\Phi}(a) = F_{\Phi}(K(a))\,.$$

Důkaz. $F^{\mathcal{M}}_{p_i}(a) = F_{p_i}(K(a)) = p_i(a)$; pro složené formule se tvrzení dokáže indukcí podle definice 3.

11. Definice. Pro každý model $\mathcal{M}$ definujeme charakteristickou formuli modelu $\Phi_{\mathcal{M}}$. K libovolnému objektu $a \in M$ utvořme elementární konjunkci $\tilde{p}_1 \,\&\, \ldots \,\&\, \tilde{p}_n$ tak, že místo i-té vlnovky dáme znak negace, je-li $P_i(a) = 0$, a nedáme nic (dáme prázdný symbol), je-li $P_i(a) = 1$. Disjunkci všech elementárních konjukcí $\varkappa$, k nimž existuje $a \in M$ tak, že $\varkappa$ je konjunkce přiřazená objektu a právě uvedeným způsobem, nazýváme *charakteristickou formulí modelu.*

Je-li u nějaká j-tice nul a jedniček (tj. $u \in M^{(j)}$) a p j-tice proměnných $p_{k_1}, \ldots, p_{k_j}$, označíme $u \circ p$ elementární disjunkci vzniklou tak, že p_{k_i} opatříme negací, právě když $u_i = 0$, a ponecháme bez negace, právě když $u_i = 1$.

Znamená-li nyní p n-tici proměnných $p_1, p_2, \ldots, p_n$, můžeme charakteristickou formuli modelu $\mathcal{M}$ zapsat symbolicky takto:

$$\Phi_{\mathcal{M}} = \bigvee\{K(a) \circ p\}_{a \in M}$$

($\bigvee X$, kde X je konečný systém formulí, značí disjunkci těchto formulí v libovolném pořádku.)

Poznámka. $\Phi_{\mathcal{M}}$ je formule v tzv. normální disjunktivní formě. Pro $x \in M^{(n)}$ je zřejmě $F_{\Phi_{\mathcal{M}}}(x) = 1$ právě když existuje $a \in M$ tak, že $K(a) = x$.

12. Věta. a) *Charakteristická formule $\Phi_{\mathcal{M}}$ modelu $\mathcal{M}$ je pravdivá v modelu $\mathcal{M}$.*

b) *Libovolná formule Ψ je pravdivá v modelu $\mathcal{M}$ právě když je dokozatelné $\Phi_{\mathcal{M}} \to \Psi$.*

Důkaz. Buď $a \in M$; dokážeme, že a splňuje $\Phi_{\mathcal{M}}$. $\Phi_{\mathcal{M}}$ je disjunkce; objekt a ji splňuje (podle definice), právě když splňuje jeden její člen. Mezi jejími členy je formule $K(a) \circ p$. Objekt a zřejmě splňuje tento člen. Ad b): Je-li dokazatelné $\Phi_{\mathcal{M}} \to \Psi$, je $\Phi_{\mathcal{M}} \to \Psi$ pravdivá v modelu $\mathcal{M}$, podle odst. 9 je Ψ pravdivá v modelu $\mathcal{M}$. Nechť $\Phi_{\mathcal{M}} \to \Psi$ není dokazatelná. Pak podle odst. 7 formule $\Phi_{\mathcal{M}} \to \Psi$ není pravdivá v kanonickém modelu $\mathscr{M}^{(n)}$, tedy podle odst. 8 existuje $u \in M^{(n)}$ takové, že $F_{\Phi\mathscr{M}}^{\mathscr{M}(n)}(u) = 1$, $F_{\Psi}^{\mathscr{M}(n)}(u) = 0$. Existuje $a \in M$ tak, že $K(a) = u$. Podle odst. 10 je $F_{\Psi}^{\mathscr{M}}(a) = F_{\Psi}(K(a)) = 0$, tj. objekt a nesplňuje Ψ v $\mathscr{M}$, Ψ není pravdivá v $\mathscr{M}$.

13. Poznámka. Zřejmě nikdy není pravdivá v žádném modelu formule Φ taková, že $F_{\Phi}(a) = 0$ pro všechna $a \in M^{(n)}$. Je však možné, že jako $\Phi_{\mathscr{M}}$ vyjde tautologie, tj. formule, jejíž kanonická asociovaná funkce je identicky rovná jedné. To nastane právě tehdy, když ke každé n-tici nul a jedniček existuje objekt $a \in M$ takový, že $u = K(a)$. Znamená to, že mezi $p_1, \ldots, p_n$ nelze určit žádné netriviální vztahy; všechny formule, které jsou pravdivé v $\mathscr{M}$, jsou tautologie výrokového počtu.

14. Definice. *Částečný model* modelu $\mathscr{M}$ je model

$$\mathscr{M}' = \langle M, P_{i_1}, \ldots, P_{i_k} \rangle ,$$

kde $1 \leqq i_1 < \ldots < i_k \leqq n$ (zřejmě úloha popsat formule obsahující jen proměnné $p_{i_1}, \ldots, p_{i_k}$ pravdivé v modelu $\mathscr{M}$ je ekvivalentní úloze popsat formule platné v částečném modelu $\mathscr{M}'$).

Podmodel modelu $\mathscr{M}$ je model

$$\mathscr{M}_1 = \langle N, P_1 \upharpoonright N, \ldots, P_n \upharpoonright N \rangle ,$$

kde $N \subseteq M$ a $P_i \upharpoonright N$ značí parcializaci funkce P_i na N.

15. Věta. *K libovolnému modelu $\mathscr{M}$ existuje podmodel $\mathscr{M}_1$ kanonického modelu $\mathscr{M}^{(n)}$ takový, že libovolná formule Φ je pravdivá v $\mathscr{M}$, právě když je pravdivá v $\mathscr{M}_1$.*

Důkaz. Definujme pro $u \in M^{(n)}$ $u \in N \equiv (\exists a \in M)(u = K(a))$ a položme $\mathscr{M}_2 = \langle N, P_1 \upharpoonright N, \ldots, P_n \upharpoonright N \rangle$. Označme $P_i \upharpoonright N = P'_1$. V $\mathscr{M}_1$ je zřejmě $K(u) = u$ pro libovolné $u \in N$ (to platí totiž v $M^{(n)}$), tedy $\Phi_{\mathscr{M}} = \Phi_{\mathscr{M}_1}$. Vzhledem k větě z odst. 12 je tím věta dokázána.

16. Podle věty 12 můžeme algoritmicky nalézt nejsilnější formuli pravdivou v $\mathscr{M}$. Tato formule však může být velmi nepřehledná. Budeme proto hledat formule slabší, než je $\Phi_{\mathscr{M}}$ (ty jsou podle věty 12 pravdivé v $\mathscr{M}$), které jsou v jistém smyslu jednoduché.

Za jednoduché budeme považovat elementární disjunkce o počtu proměnných nepřesahujícím dané číslo. Ukazujeme, že takové formule jednak umožňují systematický postup od nejjednodušších k složitějším v jistém smyslu bez zbytečného prověřování, jednak jsou jako hypotézy dobře čitelné a interpretovatelné (viz odst. 26, 27).

17. Přejmeme z Gluškovovy knihy ([1] str. 308) následující definice:

Formule (výrokového počtu) ψ je *implicentou* formule φ, je-li $\varphi \to \psi$ tautologie výrokového počtu, tj. platí-li pro každé $u \in M^{(n)}$ $F_{\psi}(u) = 0 \Rightarrow F_{\varphi}(u) = 0$. ψ je *prostou implicentou* φ, je-li její implicentou a přitom je to elementární disjunkce taková, že žádná elementární disjunkce, která vznikne z ψ vynecháním některých proměnných není implicentou formule φ.

18. Věta. *Každá formule je ekvivalentní konjunkci všech svých prostých implicent* ([1] str. 309).

Konjunkci všech prostých implicent formule φ nazýváme zkrácenou konjunktivní normální formou formule φ. Někdy ji je ještě možno zjednodušit (vynechat některé prosté implicenty, to je důležité např. při konstrukci automatů, srv. Gluškov), ale pro naše účely se zdá hledání prostých implicent vhodným.

19. Věta. 1) *Všechny prosté implicenty formule $\Phi_{\mathcal{M}}$ jsou pravdivé v modelu $\mathcal{M}$.*

2) *Buď Ψ nějaká formule obsahující jen proměnné $p_{k_1}, \ldots, p_{k_j}$. Formule Ψ je pravdivá v $\mathcal{M}$ tehdy a jen tehdy, je-li důsledkem konjunkce všech prostých implicent formule $\Phi_{\mathcal{M}}$ obsahujících jen proměnné $p_{k_1}, \ldots, p_{k_j}$.*

Důkaz. 1) plyne z věty 12, věty 9 a definice implicenty. 2) Konjunkce prostých implicent je podle věty 22 důsledkem formule $\Phi_{\mathcal{M}}$ a tedy podle věty 12b) pravdivá v $\mathcal{M}$. Je-li tedy Ψ důsledkem konjunkce některých implicent, je pravdivá v $\mathcal{M}$. Nechť obráceně je Ψ pravdivá $\mathcal{M}$. Vyjádřeme Ψ ve zkrácené konjunktivní normální formě, $\Psi \equiv \& \{\Theta; \Theta$ prostá implicenta $\Psi\}$. Každé takové Θ je implicentou formule $\Phi_{\mathcal{M}}$; nějaká část Θ' formule Θ je prostou implicentou $\Phi_{\mathcal{M}}$, přičemž platí $\Theta' \rightarrow \Theta$. Je-li tedy $\Psi = (\Theta_1 \& \ldots \& \Theta_q)$, kde všechna Θ_1 jsou prosté implicenty Ψ, je $(\Theta'_1 \& \ldots \& \Theta'_q) \rightarrow \Psi$, kde každé Θ' je částí Θ a přitom je prostou implicentou formule $\Phi_{\mathcal{M}}$ a neobsahuje jiné proměnné než $p_{k1}, \ldots, p_{k_j}$.

20. Úkol hledat všechny „zajímavé" hypotézy platné v $\mathcal{M}$ interpretujeme tedy takto: Nalézt všechny prosté implicenty formule $\Phi_{\mathcal{M}}$ kratší než n_0 (s případným dodatkem: a obsahující dané „zajímavé" proměnné v daném „zajímavém" tvaru; viz část III). Algoritmus řešící tuto úlohu bude popsán v části III.

Poznámka. Buď $u \in M^{(j)}$, $p = \langle p_{k_1}, \ldots, p_{k_j} \rangle$. Formule $u \circ p$ je pravdivá v $\mathcal{M}$ (podle definice 4), právě když pro každou kartu $K(a) = \langle P_1(a), \ldots, P_n(a) \rangle$ existuje q, $1 \leqq q \leqq j$, takové, že $P_{k_q}(a) = u_q$.

21. Uveďme nyní na několika příkladech, jak lze pracovat s prostými implicentami charakteristické formule modelu. Předpokládejme, že nás zajímá vztah vlastností $p_2, \ldots, p_j$ k vlastnosti p_1 a omezme se jen na implicenty obsahující p_1 a neobsahující žádné jiné proměnné než $p_1, \ldots, p_j$. Uvědomme si, že následující ekvivalence je tautologie: $(\varphi \vee \overline{\psi}) \equiv (\psi \rightarrow \varphi)$.

Hledáme-li tedy *důsledky* p_1, vybereme všechny prosté implicenty obsahující $\overline{p_1}$, označme je $\overline{p_1} \vee \Phi_1, \ldots, \overline{p_1} \vee \Phi_q$ a po snadné úpravě dostáváme $p_1 \rightarrow (\Phi_1 \& \ldots$ $\ldots \& \Phi_q)$.

Hledáme-li příčiny p_1, vezměme všechny implicenty obsahující p_1 bez negace, řekněme $p_1 \vee \Psi_1, \ldots, p_1 \vee \Psi_n$ a dostáváme $(\overline{\Psi}_1 \vee \ldots \overline{\Psi}_n) \rightarrow p_1$ (po úpravě výrazů Ψ_i bude příčina vyjádřena v normální (zkrácené) disjunktivní formě).

22. V tomto příkladu předpokládejme, že proměnné $p_1, \ldots, p_j$ jsou seřazeny podle stupně obtížnosti ověření, zda daný objekt splňuje p; je-li $i < j$, rozhodneme snáze, zda platí p_i než p_j. Hypotézu o proměnných $p_1, \ldots, p_j$ můžeme (za označení z minulého odstavce) zapsat takto: $(p_1 \vee (\Psi_1 \& \ldots \& \Psi_n)) \& ((\overline{p_1} \vee (\Phi_1 \& \ldots \& \Phi_q))$.

Všechna Φ_i, Ψ_j jsou formule obsahující jen proměnné $p_2, \ldots, p_j$. Koneckonců můžeme naší hypotéze přiřadit graf, jehož vrcholy jsou ohodnoceny proměnnými buďto bez negace nebo s negací. Místo obecné definice uveďme příklad. Mějme proměnné p_1, p_2, p_3, nechť

$$p_1 \vee p_2 \vee p_3\,, \quad p_1 \vee \overline{p_2} \vee \overline{p_3}\,, \quad \overline{p_1} \vee p_3$$

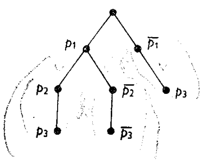
jsou právě všechny prosté implicenty charakteristické formule obsahující jen proměnné p_1, p_2, p_3. V $\mathscr{M}$ je tedy pravdivá formule

$$(p_1 \vee p_2 \vee p_3) \,\&\, (p_1 \vee \overline{p_2} \vee \overline{p_3}) \,\&\, (\overline{p_1} \vee p_3)\,.$$

Upravme podle návodu:

$$(p_1 \vee ((p_2 \vee p_3) \,\&\, (\overline{p_2} \vee \overline{p_3}))) \,\&\, (\overline{p_1} \vee p_3)\,.$$

Naší hypotéze přiřadíme následující ohodnocený graf:



Nyní si představme, že máme nějaký nový objekt $a \in M$, o kterém umíme rozhodnout, zda má vlastnosti p_1, p_2, ale nevíme, zda má p_3. Předpokládáme, že naše hypotéza platí i pro objekt a. Má-li a vlastnost p_1, sledujeme vrchol grafu, označený $\overline{p_1}$. Podle grafu platí $\overline{p_1} \vee p_3$, tedy náš objekt má vlastnost p_3. Nemá-li vlastnost p_1, sledujeme vrchol p_1. Vezměme v úvahu p_2. Nemá-li vlastnost p_2, vyčítáme z vrcholu p_2, že má vlastnost p_3. Má-li vlastnost p_2, vyčítáme z vrcholu $\overline{p_2}$, že nemá vlastnost p_3. (Čtenář si už povšiml, že řetězy v grafu udávají prosté implicenty.) Tato metoda ovšem nemusí dát odpověď: kdybychom usuzovali z vlastností p_1, p_3 na p_2, lze snadno zjistit, že v případě, že objekt a má vlastnost p_1, nerozhodneme, zda má p_2.

III.

Třetí část práce obsahuje popis algoritmu pro řešení úlohy, zformulované v části II.

Navrhovaný a vyzkoušený algoritmus pracuje takto: Pro každé $k = 1, \ldots, n_0$ vytváří všechny možné k-tice z proměnných $p_1, \ldots, p_n$ a jejich negací tak, aby žádná k-tice neobsahovala současně stejnou proměnnou bez negace a s negací.

Každou takovouto k-tici

$$(1) \qquad \tilde{p}_{i_1}, \ldots, \tilde{p}_{i_k}$$

(zde $1 \leqq i_1 < i_2 < \ldots < i_k \leqq n$, $\tilde{p}_{i_j}$ znamená buď p_{i_j} nebo $\overline{p_{i_j}}$) považuje algoritmus za elementární disjunkci

$$(2) \qquad \tilde{p}_{i_j} \vee \tilde{p}_{i_2} \vee \ldots \vee \tilde{p}_{i_k} .$$

Zjišťuje o ní, zda platí v modelu $\mathcal{M}$ (tj. zda je implicentou charakteristické formule Φ), v kladném případě prověřuje dále, zda (2) je prostou implicentou Φ. Na výstupu žádáme od algoritmu právě prosté implicenty Φ.

Práci algoritmu nejlépe objasní hrubé blokové schéma (obr. 1).

Pro optimalizaci celého procesu z hlediska strojového času je důležité vhodně vybrat uspořádání k-tic proměnných. Volili jsme uspořádání lexikografické; přitom základní pořadí proměnných a jejich negací je takovéto: $p_1, \overline{p_1}, p_2, \overline{p_2}, \ldots, \overline{p_{n-1}}, p_n, \overline{p_n}$. Např. první k-ticí při takto zvoleném uspořádání pro každé k je $p_1, p_2, \ldots, p_k$, poslední k-ticí je $\overline{p}_{n-k+1}, \overline{p}_{n-k+2}, \ldots, \overline{p}_n$.

Popsané uspořádání k-tic poskytuje určitou možnost zkrácení procesu. Nechť např. disjunkce

$$(3) \qquad \tilde{p}_{i_1} \vee \tilde{p}_{i_2} \vee \ldots \vee \tilde{p}_{i_j} \vee \ldots \vee \tilde{p}_{i_k} \quad (j < k)$$

platí v $\mathcal{M}$, ale není prostou implicentou Φ. Zjistíme-li, že disjunkce $\tilde{p}_{i_1} \vee \tilde{p}_{i_2} \vee \ldots \vee \tilde{p}_{i_j}$, která je vlastní částí (3), také platí v $\mathcal{M}$, můžeme ihned „posunout proměnnou $\tilde{p}_{i_j}$", tj. přejít ke zkoumání disjunkce (4)

$$(4) \qquad \tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{j-1}} \vee \overline{p_{i_j}} \vee p_{i_j+1} \vee \ldots \vee p_{i_j+k-j}$$

v případě, že $\tilde{p}_{i_j} = p_{i_j}$, a ke zkoumání disjunkce

$$(5) \qquad \tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{j-1}} \vee p_{i_j+1} \vee \ldots \vee p_{i_j+k-j+1}$$

v případě, že $p_{i_j} = \overline{p_{i_j}}$.

Všechny elementární disjunkce, které v našem uspořádání budou mezi (3) a (4), resp. mezi (3) a (5), budou sice platit v $\mathcal{M}$, žádná z nich však nebude prostou implicentou Φ a tudíž jsou pro nás nezajímavé.

Další možnost úspory strojového času poskytují různé varianty té části algoritmu, která zkoumá, zda vyšetřovaná disjunkce je prostou implicentou Φ.

Algoritmus může pracovat takto:

Počínaje vůbec první zjištěnou prostou implicentou ukládat každou další do paměti počítače a rozhodování o tom, zda zkoumaná disjunkce, platná v $\mathcal{M}$, je prostou implicentou Φ, lze provádět následovně. Porovnat ji s každou dříve utvořenou (a tudíž uloženou do paměti) prostou implicentou Φ a zjišťovat, zda některou z nich neobsahuje jako vlastní část. Jestliže žádnou neobsahuje, pak vzhledem k pořadí, ve kterém jsou disjunkce vytvářeny, je možné tvrdit, že zkoumaná disjunkce je prostou implicentou Φ. Tento způsob je výhodný hlavně z počátku, pokud počet nalezených prostých implicent není příliš velký.

Další vyzkoušená varianta, kterou je možno do algoritmu zapojit, pracuje takto: Ze zkoumané disjunkce (2), platné v $\mathscr{M}$, po řadě vytvoří disjunkce

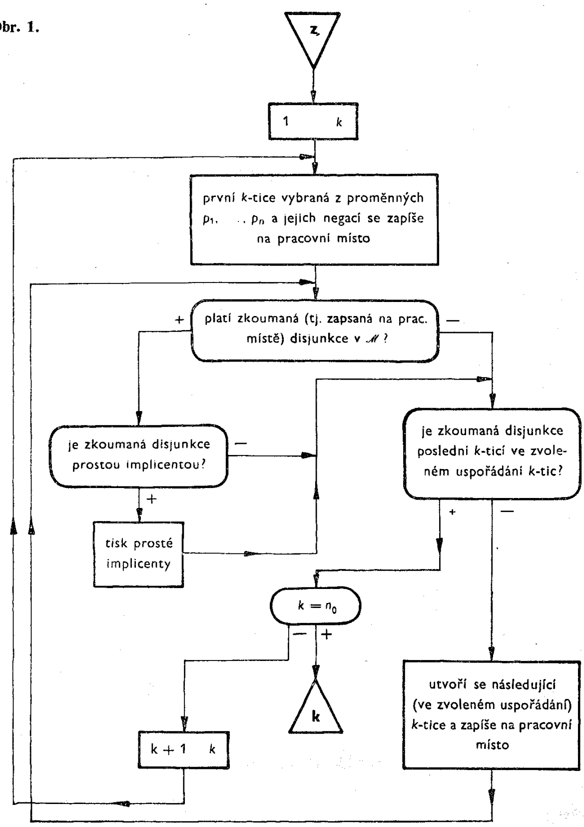
$$(2^{(k)}) \qquad \tilde{p}_{i_1} \vee \tilde{p}_{i_2} \vee \ldots \tilde{p}_{i_{k-1}} ,$$

$$(2^{(k-1)}) \qquad \tilde{p}_{i_1} \vee \tilde{p}_{i_2} \vee \ldots \vee \tilde{p}_{i_{k-2}} \vee \tilde{p}_{i_k} ,$$

$$\vdots \qquad \qquad \vdots$$

$$(2^{(1)}) \qquad \tilde{p}_{i_2} \vee \ldots \vee \tilde{p}_{i_k} .$$

Obr. 1.



Disjunkce $(2^{(j)})$ se dostane z (2) vynecháním proměnné $\tilde{p}_{i_j}$.

Pro každou z formulí $(2^{(k)})$, ..., $(2^{(1)})$ algoritmus prověří, zda platí ještě v $\mathscr{M}$. (Vynechání kterékoliv proměnné v (2) tuto formuli zesiluje, jelikož je to disjunkce, a je možné, že přestane pak platit v $\mathscr{M}$.)

Platí-li některé $(2^{(j)})$ $1 \leqq j \leqq k$ v $\mathscr{M}$, znamená to, že v (2) můžeme vynechat proměnnou $\tilde{p}_{i_j}$, aniž bychom narušili platnost této formule v $\mathscr{M}$. Tudíž (2) nebyla prostou implicentou Φ. V takovém případě algoritmus přechází ze zkoumání další disjunkce.

Jestliže při zkoumání pravdivosti formulí $2^{(k)}, \ldots, 2^{(1)}$ zjistíme, že hned první z nich, totiž $(2^{(k)})$ v $\mathscr{M}$ platí, postupujeme dál speciálním způsobem, který opět šetří strojový čas. Pokoušíme se totiž v $(2^{(k)})$ vynechat ještě další proměnné zprava, tj. zkoumáme po řadě formule

$$\tilde{p}_{i_j} \vee \ldots \vee \tilde{p}_{i_{k-2}},$$

$$\tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{k-3}} \text{ atd.}$$

Nechť $\tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{k-l}}$ $(l \geqq 1)$ je poslední disjunkce v této posloupnosti, která ještě platí v $\mathscr{M}$; tj.

$$\tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{k-l-1}} \text{ již v } \mathscr{M} \text{ neplatí.}$$

Můžeme pak ihned „posunout proměnnou $\tilde{p}_{i_{k-l}}$", to znamená přejít ke zkoumání formule

$$\tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{k-l-1}} \vee \overline{p_{i_{k-l}}} \vee \tilde{p}_{i_{k-l}+1} \vee \ldots \vee p_{i_{k-l}+l}$$

v případě, že $\tilde{p}_{i_{k-l}} = p_{i_{k-l}}$ a ke zkoumání formule

$$\tilde{p}_{i_1} \vee \ldots \vee \tilde{p}_{i_{k-l-1}} \vee p_{i_{k-l}+1} \vee \ldots \vee p_{i_{k-l}+l+1}$$

v případě, že $\tilde{p}_{i_{k-l}} = \overline{p_{i_{k-l}}}$.

Postupujeme-li právě popsaným způsobem, nepotřebujeme pro rozhodování o tom, zda zkoumaná disjunkce je prostou implicentou Φ, vůbec žádnou informaci o dříve nalezených prostých implicentách.

V paměti počítače není třeba rezervovat místo pro ukládání prostých implicent a kromě toho poskytuje tento způsob ještě další podstatnou výhodu: je možné kdykoliv bez komplikací výpočet přerušit. Při opětovném zahájení je třeba jenom zapsat na pracovní místo formuli, u které byla práce algoritmu zastavena.

Konečně, zlepšuje se poněkud i stabilita výpočtu oproti náhodným chybám počítače. Zatímco v první variantě špatně určená prostá implicenta zapsaná do paměti má vliv na celou řadu dalších, v druhé variantě k podobnému jevu nedochází.

V praxi se nejlépe osvědčil kombinovaný způsob – začít první variantou a po určité době pokračovat druhou.

Popíšeme ještě určité zobecnění algoritmu. Často dochází k situacím, ve kterých popsaný algoritmus nemůže poskytnout odpověď, protože tato odpověď je skryta mezi velkým množstvím výsledků (hypotéz), které nás v danou chvíli nezajímají.

Vyhledání všech hypotéz by trvalo dlouhou dobu, anebo není z důvodu nedostačující operační rychlosti počítače vůbec proveditelné. Navíc by bylo nutné ze záplavy hypotéz oddělit dodatečně zajímavé a to v podstatě manuelně.

Uvedeme tři příklady, ukazující motivy zobecnění algoritmu.

1. Již v části I byla naznačena situace, kdy máme v souvislosti s určitým modelem $\mathcal{M}$ utvořenu jistou hypotézu, např. ve tvaru elementární disjunkce. Víme o ní, že v celém $\mathcal{M}$ neplatí (tj. v $\mathcal{M}$ jsou objekty, které ji nesplňují). Očekáváme ale, že po přidání určitých „doplňků“ bude tato hypotéza v $\mathcal{M}$ již platit. Chtěli bychom znát právě tyto doplňky.

2. Potřebujeme vědět, co všechno v daném modelu vyplývá z jedné nebo několika příčin (jejich konjunkce), nebo naopak, které jevy mohou způsobit určitý důsledek.

3. Zajímá nás souvislost určitých vytčených proměnných, přitom je u některých z nich lhostejné, zda do hypotézy vstoupí jako příčiny nebo důsledky, u jiných to lhostejné není.

Každá z těchto úloh může být řešena pomocí zobecněného algoritmu, který pracuje takto: Po zavedení údajů (tj. karet objektů modelu) do paměti počítače vkládáme do ní ještě tzv. charakteristická slova, která vymezují z množiny P všech proměnných $p_1, \ldots, p_n$ tři disjunktní podmnožiny A, B, C. Proměnné typu A (tj. prvky množiny A) nazýváme „*absolutně zajímavými*“ proměnnými, proměnné typu B jsou „*zajímavé*“ proměnné a konečně proměnné typu C budeme nazývat „*ostatními*“ proměnnými. Poznamenejme, že $A \cup B \cup C$ nemusí být celé P, a dále, že každá z množin A, B, C může být prázdná. V množině $A \cup B \cup C$ zadáme ještě podmnožinu M (případně prázdnou) proměnných, u kterých předepíšeme tvar (tj. stanovíme, zda se proměnná má vyskytovat v nalezených implicentách s negací nebo bez negace). Proměnné, u kterých je vytčen negativní tvar, tvoří množinu $N \subseteq M$; proměnné z $M - N$ jsou vytčeny v pozitivním tvaru. Budeme říkat, že pětice A, B, C, M, N určuje *sondu* v modelu $\mathcal{M}$ (viz část I). Úlohu formulujeme takto: nalézt všechny prosté implicenty Φ modelu $\mathcal{M}$ vyhovující následujícím požadavkům:

1. Každá implicenta obsahuje všechny proměnné typu A.
2. Každá implicenta obsahuje alespoň jednu proměnnou typu B.
3. Všechny proměnné, vyskytující se v každé z implicent, jsou buď typu A, nebo typu B, nebo typu C.
4. Každá proměnná s předepsaným tvarem se může v libovolné implicentě vyskytovat pouze v tomto předepsaném tvaru.

Nyní ukážeme, jak je možné pomocí zobecněného algoritmu získat řešení tří výše formulovaných typických úloh.

K první úloze: Všechny proměnné původní hypotézy prohlásíme za absolutně zajímavé a u všech předepíšeme příslušný tvar, totiž právě ten, ve kterém se vyskytují v původní hypotéze. Máme-li určité představy o charakteru očekávaných doplňků, zavedeme některé proměnné do procesu jako proměnné typu B, v opačném případě volíme $B = \emptyset$.

Druhou úlohu vyřešíme tak, že při hledání důsledků ze známých příčin prohlásíme příčiny za „absolutně zajímavé" (nebo „zajímavé") a předepíšeme jim negaci. Při hledání příčin určitého jevu bude proměnná odpovídající důsledku absolutně zajímavá ve tvaru bez negace.

Třetí úlohu řešíme tak, že všechny vytčené proměnné prohlásíme za absolutně zajímavé. U některých předepíšeme tvar, u jiných nikoliv.

Povšimněme si ještě, jak se mění charakter úlohy v závislosti na tom, které z množin A, B, C volíme prázdnými.

Vyloučíme-li triviální případ $A = B = C = \emptyset$, zbývá 7 možností. Odhlédneme-li od možnosti předepsat tvar proměnných, je případ $A = B = \emptyset$ i případ $A = C = \emptyset$ totožný s původním algoritmem. $B = C = \emptyset$ dá jen částečný výsledek původního algoritmu, totiž hypotézy největší délky. Případ $A = \emptyset$, $B \neq \emptyset \neq C$ je velmi zajímavý, stejně jako případ $B = \emptyset$, $A \neq \emptyset \neq C$. Ten je téměř totožný s případem $C = \emptyset$, $A \neq \emptyset \neq B$.

Poslední možnost je, že žádná z množin A, B, C nebude prázdná.

IV.

Pro ilustraci uvedeme některé údaje o modelech, vyšetřených samočinným počítačem MINSK 2. Objekty těchto modelů bylo 230 epileptiků, evidovaných v rámci Protizáchvatové poradny polikliniky v Praze 1. O každém z nich byly k dispozici údaje o 36 „vlastnostech".* V připojené tabulce jsou uvedeny počty zjištěných implicent a to jednak pro úplný model, obsahující všech 36 proměnných, jednak pro dva částečné modely, obsahující jen některé vybrané proměnné.

Počet proměnných v modelu	Počet proměnných v jedné implicentě									
	1	2	3	4	5	6	7	8	9	10 ...
36	0	154	2050	...						
9	0	15	13	10	2	0	0	0	0	
12	0	4	99	107	55	29	9	...		

Snadnými odhady se dá zjistit (a praxe to také potvrzuje), že se vzrůstajícím počtem proměnných v modelu rychle roste doba potřebná k výpočtu a také (většinou) počet výsledků, takže se oddělení sémanticky zajímavých hypotéz stává téměř nepro-

* Dále byly zkoumány modely sestrojené pro studium problematiky předčasných porodů a pedagogické problematiky (dětská četba).

veditelným.* Druhé potíži lze čelit použitím zobecněného algoritmu s rozlišením významnosti proměnných a vytčením tvaru (s negací či bez negace). První potíž vede ke zvýšeným nárokům na použitý typ počítače. Již z popisu algoritmu je zřejmé, že zpracování úlohy neklade téměř žádné nároky na vstupní jednotku a vnější paměťová zařízení. Také výstupní jednotka je zatížena poměrně velmi málo. Např. na stroji MINSK 2 jsme vystačili s jedním blokem operativní paměti (tj. se 4096 dvojkovými slovy o 36 řádech). Pro zápis jedné karty modelu bylo vyhraženo jedno paměťové místo (počet proměnných $n = 36$), zkoumaná elementární disjunkce byla zapsána na dvou paměťových místech (na prvním byly jednotky v řádech proměnných vyskytujících se v disjunkci, na druhém jednotky v řádech těch proměnných, které se v disjunkci vyskytují s negací). Použité operace jsou převážně logického charakteru. Velikost modelů, dovolujících strojové zpracování (tj. počet objektů a počet proměnných), podstatně závisí na operační rychlosti stroje. Úloha je uspokojivě řešitelná na malých samočinných počítačích (které jsou tč. u nás k dispozici) jen pro modely s menším počtem proměnných.

Další rozvíjení metody GUHA spočívá v tom, že budou respektovány též některé statistické problémy, zejména:

a) problém, do jaké míry je zjištěná hypotéza i obecně platná (jaká je pravděpodobnost, že formule pravdivá v modelu „pokusných objektů“ bude pravdivá i v modelu „všech existujících objektů“);

b) problém rozdílu platnosti dvou syntakticky stejných hypotéz pravdivých v modelu, ale lišících se četnostmi výskytu kombinací nul a jedniček (srv. konec odstavce I);

c) problém významu hypotéz sice v modelu nepravdivých, ale splňovaných jistým „dostatečným“ množstvím objektů. (Algoritmus by šlo zobecnit tak, aby takové hypotézy systematicky vyhledával na předepsané hladině četnosti výskytu.)

Jiná cesta zobecnění spočívá v tom, že by se studovala pravdivost jiných druhů hypotéz než jsou generální unární hypotézy. Např. tak, že by se vzaly v úvahu binární predikáty a jim odpovídající binární relace na množině M. Pak by bylo lépe upustit od formulí výrokového počtu a studovat přímo formule predikátového počtu.

Dalo by se též uvažovat o užití metody GUHA při minimalizaci booleovských výrazů a sice k hledání zkrácené konjunktivní normální formy booleovské funkce např. až 10 proměnných. Metoda by byla výhodná zvláště pro případ, že tato funkce nenabývá příliš často hodnoty 1. (Předpokládá se, že element realizující konjunkci je „dražší“ než element realizující disjunkci.)

* Že skutečně získáme i řadu hypotéz sémanticky nezajímavých, budiž dokumentováno na první implicentě prvního ze zkoumaných modelů $\overline{p_1} \vee \overline{p_5}$, kde p_1 znamená „pacient je muž“ a p_5 „u pacienta byla zjištěna závislost záchvatů na graviditě“, tj. „u žádného muže nebyla zjištěna závislost epileptických záchvatů na graviditě“.

Pro aplikace metody GUHA bude mít velký význam i volba vhodných (ne „jalových“) modelů. Autoři zamýšlejí publikovat na jiném místě článek určený pro uživatele metody GUHA, v němž mají být dána jistá kritéria pro volbu modelů a interpretaci zjištěných hypotéz.

(Došlo dne 19. září 1965)

LITERATURA

[1] В. М. Глушков: Синтез цифровых автоматов. Физматгиз, Москва 1962.

[2] A. Grzegorczyk: Zarys logiki matematycznej, PWN, Warszawa 1961.

[3] J. Zelenka, O. Zich: Matematisch-logisches Modell der Vestibular- und Gehörstörungen. Rozpravy ČSAV *75*, (1965), sešit 9.

SUMMARY

GUHA — the Method of Systematical Hypotheses Searching

PETR HÁJEK, IVAN HAVEL, METODĚJ CHYTIL

The presented method is an application of the mathematical logic and computer technique to the research problems of concrete sciences.

Let us assume a model (precisely a unary semantic model), i.e. a finite non-empty system of objects and a finite system of (unary) properties to be given. It is known, moreover, for each object and each property, whether or not the object possesses the property. (E.g. objects are patients and properties are diseases or facts that some medicines were taken (administrated) etc.). The research worker usually looks for the relations among the properties of the model that hold true for all the objects. E.g., whether for each object, the following holds true: if the object possesses the property p_1, and does not possess the property p_2, then it possesses the property p_3. If he finds this relation hold true for all the objects of his model (the model of objects under consideration), he is able to formulate and verify the hypothesis of validity of such a relation for all existing objects in general.

All possible hypotheses of this kind (the so called general unary hypotheses) can be generated and verified in the model automatically, by means of computer technique. In the output device of the computer there will appear all the hypotheses that hold true in model. Hence the GUHA-method (General Unary Hypotheses Automaton) can be considered as a substitution for an intuition or as "an offering of hypotheses".

There arise two problems to be solved:

1. It is necessary to choose a suitable class of formulae that will be limited to investigating the model.

2. It is necessary to choose an optimal way of generation and verification of hypotheses.

The first problem is solved with help of mathematical logic methods. To each property of the model there corresponds a single propositional variable, the hypotheses are identified with formulae of propositional calculus. The notion of the formula true in the model we define in accordance with well-known definition of Tarski, which (at least in the case of a finite model) coincides with usual conception of validity. It can be proved, that to each model $\mathcal{M}$ there exists a single formula $\Phi_{\mathcal{M}}$ (the so called characteristic formula of the model $\mathcal{M}$) so that:

1. $\Phi_{\mathcal{M}}$ holds true in $\mathcal{M}$,
2. an arbitrary formula Ψ holds true in $\mathcal{M}$ if and only if $\Phi_{\mathcal{M}}$ implies Ψ (in propositional calculus).

The significance of the formula $\Phi_{\mathcal{M}}$ is only theoretical. It is expedient to look for the so called prime implicents of $\Phi_{\mathcal{M}}$.

Definition. An elementary disjunction Ψ (i.e. the disjunction of some variables or variables with negation) is called to be a prime implicent of $\Phi_{\mathcal{M}}$ if

1. Ψ holds true in $\mathcal{M}$,
2. any disjunction arising by omission of some element of Ψ does not hold true in $\mathcal{M}$.

Theorem. 1. $\Phi_{\mathcal{M}}$ *is logically equivalent with the conjunction of all its prime implicents.*

2. *Suppose* Ψ *to consist only of the variables* $p_{k_1}, \ldots, p_{k_j}$ $(1 \leqq k_1 < \ldots < k_j \leqq n)$. Ψ *holds true in* $\mathcal{M}$ *if and only if* Ψ *is the logical consequence of the conjunction of all prime implicents* $\Phi_{\mathcal{M}}$ *containing only the variables* $p_{k_1}, \ldots, p_{k_j}$.

Hence, we can limit our investigations in a given model $\mathcal{M}$ to looking for prime implicents $\Phi_{\mathcal{M}}$. The algorithm of generation and verification of all possible elementary disjunctions is described. The algorithm takes into account some elementary disjunction being a prime implicent $\Phi_{\mathcal{M}}$. A series of hypotheses which hold true in model but are not prime implicent $\Phi_{\mathcal{M}}$ is not tested. Moreover, it is possible to modify an algorithm in such a way that there will be generated and tested only "interesting" hypotheses (i.e. the hypotheses with "interesting" variables of an "interesting" form).

Finally the experience of the first practical applications of the GUHA-method is described. Some possibilities of further improvement are shown.

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