

Ing. Jaroslav PLUCAR  
VÚEPE, Ostrava

## UŽIVATELSKÝ ORIENTOVANÝ SYSTÉM PRO AGREGACI INFORMACÍ Z EXISTUJÍCÍCH DATOVÝCH SOUBORŮ

Přednětem příspěvku je třetí verze programového systému UPP (Unifikované Parametrické Programy), která je implementována v jazyku PL/I na počítači IBM 370 a je využívána ve VÚEPE Ostrava pro jednoduché jednorázové agregace informací z existujících datových souborů. Třetí verze UPP vznikla modifikací verze II, která byla publikována v časopisu Mechanizace a automatizace administrativy v číslech 2 a 3 v roce 1978.

Systém je možno z hlediska vykonávané funkce považovat za běžný, problémově orientovaný "software-package". Smyslem příspěvku je ukázat na tomto konkrétním příkladu některé aspekty vytváření parametrických systémů určených pro zaškoleného uživatele - neprogramátora. Této otázce je zatím věnována jak v běžné praxi tak i v literatuře minimální pozornost. Prokázala to nakonec i diskuse na semináři "Programování 79" v minulém roce.

V první části příspěvku je uveden stručný popis systému UPP verze III, ve druhé části jsou pak rozvedeny uživatelské aspekty s konkrétními odkazy na daný popis.

### 1. Stručný popis UPP verze III, zadávání parametrů

V rámci systému pracují tři vlastní programy a jeden řízení třídící program. Pořadí, v jakém jsou programy při řešení úloh zpracovávány, odpovídá níže uvedenému popisu.

### 1.1 Program pro předzpracování parametrů

Tento program provádí formální kontrolu zadaných parametrů (Tabulky T1 - T5) a jejich konfrontaci s tabulkou popisu zpracovávaného datového souboru. Dále doplňuje načtené parametry z uvedené tabulky popisu o hodnoty, týkající se fyzického uložení údajů ve větách zpracovávaného datového souboru (pozice údaje ve větě, jeho délka v bytech, typ fyzického uložení, počet desetinných míst apod.). Z doplněných parametrů vytváří výstupní soubor parametrů, který je pak využíván výběrovým a tiskovým programem. Vytváří rovněž třídící parametry (z tabulky T3) pro firemní třídící program IERRC000.

### 1.2 Program pro výběr vět a jejich rozčlenění do skupin

Program testuje na základě tabulky T1 všechny věty vstupního souboru a ty z nich, které vyhoví uvedeným podmínkám, přebírá do dalšího zpracování. Každou vybranou větu rozšíří o 30 bytů, do kterých dosadí název skupiny, do které je věta zařazena (viz tabulka T2 - rozčlenění vět do skupin). Je-li věta zařazena do dvou nebo více skupin, pak ji program vypíše do výstupního souboru vybraných vět dva nebo vícekrát s odlišnými názvy skupin. Je tedy možno specifikovat v parametrech jak disjunktivní (oddělitelné) tak i nedisjunktivní (neoddělitelné) skupiny.

### 1.3 Firemní třídící program IERRC000

Program provede setřídění vybraných vět na základě třídících parametrů, které byly vytvořeny programem pro předzpracování parametrů z tabulky T3.

### 1.4 Program pro tisk údajů z věty a rekapitulací

Údaje i rekapitulace včetně hodnot průměrů, maximálních a minimálních hodnot u součtovaných údajů jsou tištěny na základě tabulek T3, T4 a T5 do sestavy, která má standardizovaný

tvár. Prostřednictvím parametrů, uvedených v tabulkách T3-T5 zadáváme tedy pouze obsah sestavy, nikoliv její tvár. Jinými slovy řečeno, tvár sestavy vytváří program, nikoliv uživatel. Program pracuje na principu normalizovaného programování s tím, že výroky pro obsazování a testování klíčů si generuje sám na základě informací, které získá z tabulek T3-T5 a z tabulky popisu zpracovávaného datového souboru.

### 1.5 Úlohy řešitelné pomocí UPP

System UPP umožňuje uživateli provádění těchto úloh:

- Výběr vět ze souboru podle toho, jak vyhoví (resp. nevyhoví) uvedeným podmínkám.
- Rozčlenění vybraných vět do skupin podle toho, jak vyhoví zadaným podmínkám. Každá skupina vět musí být uživatelem pojmenována.
- Seřídění vybraných vět podle zvolených hledisek. Třídícím hlediskem může být vždy jen některý údaj věty; údajem věty se rovněž stává jméno skupiny, které je do věty zapsáno při provádění předcházející úlohy.
- Sečtení a tisk počtu vět (se stejnou hodnotou třídícího údaje) na sumarizačním řádku výstupní sestavy.
- Sečtení a tisk součtů zvolených údajů věty na sumarizačním řádku sestavy (spolu s výše uvedeným počtem vět).
- Vypočítání a tisk průměrných hodnot součtovaných údajů věty na sumarizačním řádku sestavy (spolu s výše uvedenými součty).
- Vytisknutí jednotlivých údajů věty na běžném řádku sestavy.

### 1.6 Zadávání parametrů - příklady

Zadávání parametrů pro jednotlivé úlohy je demonstrováno na vynýšeném příkladu z oblasti jednotné evidence pracujících. Příklad z této oblasti byl zvolen proto, že význam jednotlivých údajů je obecně znám a je možno tedy příklad považovat za dostatečně ilustrativní pro vysvětlení všech funkcí

UPP. Zadávání parametrů se provádí vyplněním tabulek na uvedených formulářích. Pro vyplňování formulářů platí běžně užívané zásady: vyplňování formuláře odshora dolů bez vynechávání volných řádků, kolonky se vyplňují zleva doprava, desetinná tečka se uvádí, aj.

### Úloha "výběr vět"

Výběrem vět se rozumí zúžení množiny vět podle zadaných podmínek na podmnožinu, která uživatele zajímá. Ze základního souboru vět se vyloučí věty, jejichž testované údaje nevyhovují podmínkám uvedeným ve formuláři VÝBĚR VĚT. Hodnoty testovaných údajů je možno zadávat diskrétně nebo intervalem hodnot (od - do). Pro jeden údaj je možno uvést několik podmínek (vztah mezi podmínkami je logické "nebo"). Pro výběr věty je možno testovat několik údajů. Věta bude vybrána, pokud každý její testovaný údaj vyhoví (vztah mezi údaji je logické "a").

Výběr vět je možno dělat podle celého údaje nebo jen podle určité jeho části, kterou uživatel vyčlení takto: "od kterého znaku údaje si vezme při testování v úvahu kolik znaků".

### Příklad:

Ze souboru "pracovníci" je třeba vybrat TH pracovníky (údaj kategorie = 1) mladší padesáti let rok (narození 1930 a výše). Zápis do formuláře viz tabulka T1.

### Úloha "rozčlenění vět do skupin"

Při realizaci této úlohy počítačem dochází k označování příslušnosti jednotlivých vět k některé z vymezených skupin. Vymezení příslušnosti věty ke skupině se zajišťuje podmínkami zapsanými uživatelem do formuláře ROZČLENĚNÍ VĚT DO SKUPIN zcela obdobně jako v úloze "výběr vět". Označení věty se provede názvem skupiny, který uživatel napíše pod podmínky vymežující určitou skupinu.

## Příklad:

Soubor pracovníci je třeba rozčlenit na skupiny: muži (údaj pohlaví = 1), ženy (údaj pohlaví = 2) a mladiství (rok narození = 1959 a výše). Zapsání příkladu do formuláře viz tabulka T2.

Úlohy "seřídění vět", "tisk počtů vět", "tisk součtů údajů", "tisk průměrných hodnot součtovaných údajů", "tisk maximálních a minimálních hodnot součtovaných údajů".

Při všech těchto úlohách může být zpracováván pouze sekvenční soubor, tedy soubor, jehož věty těsně po sobě následují a jen v tomto pořadí mohou být čteny nebo tištěny. Úlohy jsou uvedeny na jednom formuláři, protože jejich zadávání spolu těsně souvisí.

Úlohou "seřídění vět" rozumíme uspořádání vět podle hodnot znaků zvolených údajů ve vzestupném nebo sestupném pořadí. Třídění je třeba zadávat s ohledem na požadovaný tisk počtů vět: je-li požadován tisk počtu vět se stejnými hodnotami určitého údaje, pak soubor vět musí být podle tohoto údaje seříděn. Po vytištění počtu vět je možno zadat odkok na novou stránku. Této možnosti se využívá pro zvýšení přehlednosti výstupní sestavy nebo je-li potřeba výstupní sestavu rozdělit na části a rozeslat na jednotlivé útvary podniku.

"Tisk součtů údajů" lze zadávat pro všechny údaje, jejichž hodnota je v datovém souboru zapsána numerickými znaky. Pokud v některých údajích se vyskytnou znaky nenumernícké (např. místo nul mezery), budou tyto nenumernícké znaky automaticky nahrazeny nulou a sečítání se provede s takto modifikovanými údaji. Na výstupní sestavě bude na tuto modifikaci upozorněno. U sčítaných údajů je rovněž možno zadávat tisk jeho průměrné hodnoty, která vznikne podělením součtu hodnot údaje počtem vět. Rovněž je možno zadat tisk maximálních a minimálních hodnot součtovaného údaje.

### T1 - VÝBĚR VĚT

Pořad. číslo	Údaj dle kterého se provádí výběr				Hodnota testovaného údaje	
	Název údaje	Od znaku	Počet znaků	R E L A C I E	Dolní mez intervalu hodnot údaje nebo diskrétní hodnota	Horní mez intervalu hodnot
V 0 1	KATEGORIE			= 1		
V 0 2	BATMAN	5	2	= 25		99
V 0 3						
V 0 4						

### T2 - ROZČLENĚNÍ VĚT DO SKUPIN

Pořad. číslo	Údaj dle kterého se provádí výběr				Hodnota testovaného údaje	
	Název údaje	Od znaku	Počet znaků	R E L A C I E	Dolní mez intervalu hodnot nebo diskrétní hodnota údaje	Horní mez intervalu hodnot údaje
R 0 1	BATMAN	5	2	= 59		99
R 0 2	SKUPINA				MLADISTVI	
R 0 3	POHLAVI			= 1		
R 0 4	SKUPINA				MUZI	
R 0 5	POHLAVI			= 2		
R 0 6	SKUPINA				ZENY	

### T5 - TISK JEDNOTLIVÝCH ÚDAJŮ VĚTY

Název sestavy	
T 0 0	SPECIÁLNÍ PRŮZKUM ZAMĚSTNANCŮ

Poř. číslo	Tisk údajů na běž. řádku sestavy		
	Název údaje	Od znaku	Počet znaků
T 0 1	EVIZCIS		
T 0 2	JMENN		
T 0 3	BATMAN		
T 0 4	UZBA		

### T3 - SETŘÍDĚNÍ VĚT - TISK POČTŮ VĚT

TŘÍDĚNÍ VĚT						POČET VĚT														
Stupeň třídění	Třídící údaj				Druh tříd.	Tisk	Odskok													
	Od znaku	Počet znaků	VZS = vzestupně VZS = sest.	ANO = tisknout ANO=netisk.	ANO = pro- vést odskok ANO=neprov.															
S 0 1	S	K	U	P	I	N	A													
S 0 2	S	T	R	E	D															
S 0 3	S	Z	A	N	E	S	T													
S 0 4	S	J	H	E	N	E														

KDY TISKNOUT SOUČET

1 = tisknout

2 = netiskn.

### T4 - TISK SOUČTŮ ÚDAJŮ

Pořadové číslo	Sečítaný údaj				Tisk prům.	Tisk max.	Tisk min.													
	Název údaje	Od znaku	Počet znaků	ANO = tisknout ANO=netiskn.	ANO = tisknout ANO=netiskn.	ANO = tisknout ANO=netiskn.	ANO = tisknout ANO=netiskn.													
T 0 1	N	Z	B	A																
T 0 2	N	Z	B	Z	A	R														
T 0 3																				
T 0 4																				
T 0 5																				
T 0 6																				
T 0 7																				
T 0 8																				

## Úloha "tisk jednotlivých údajů věty"

Při tisku jednotlivých údajů věty dochází k prostému výpisu hodnot zvolených údajů na běžném řádku sestavy. Tisk se provádí pouze v jednom řádku a pokud požadovaný počet tištěných znaků přesáhne rozměr stránky, přebytečné znaky se nevytisknou. Názvy údajů uvedené ve formuláři se tisknou jako záhlaví výstupní sestavy. Jednotlivé údaje jsou automaticky odělovány dvěma mezerami.

### Příklad:

Soubor pracovníků, který byl v předcházejícím příkladu rozčleněn na skupiny MUŽI, ŽENY, MLADISTVÍ, seřídíme podle skupin, uvnitř skupiny podle středisek, uvnitř středisek podle profesí (profese = zaměstnání vykonávané v organizaci) a v rámci jednotlivých profesí, střediskem a skupinou. Za skupinou provedeme odskok na novou stránku. Vytiskneme rovněž součty údajů: mzdová sazba - základní plat (zkrácený název MZDA), mzdová třída (mzdová třída je součástí ukazatele mzdové sazby, zkrácený název MZDZAR) za každou profesi. U jednotlivých profesí spočteme průměrnou mzdovou třídu. Jmenovitě vypíšeme všechny pracovníky s udáním těchto údajů na běžném řádku sestavy: evidenční číslo, jméno, datum narození, mzdová sazba - základní plat. Zápis do formulářů viz tabulka T5.

## 2. Uživatelská orientace parametrického systému

V této kapitole jsou uvedeny základní požadavky kladené na systém určený pro zaškoleného uživatele-neprogramátora, tj. pro parametrického uživatele. Rovněž jsou zde uvedeny některé doporučené postupy, jak tyto požadavky realizovat. Níže uvedená doporučení si rozhodně nečiní nárok na úplnost a všeplatnost. Sama problematika implementace uživatelsky orientovaných parametrických systémů je v současné době teprve v počátcích. Jde spíše o úvod do problematiky, o nastínění a prvotní klasifikaci problémů s naznačením cest, jak tyto problémy řešit.

## 2.1 Forma a rozsah zadávaných parametrů

Při volbě formy, jakou budou parametry do systému zadávány, existují dvě základní možnosti: verbální forma, jakou jsou zapisovány např. instrukce programu ve vyšších programovacích jazycích, nebo tabulková forma, jaká byla použita u popsaného systému UPP.

V literatuře se často objevuje názor o nutnosti přiblížení jazyka počítače mluvenému lidskému jazyku. Většina lidí si pak automaticky odvozuje že to musí být jazyk hovorový. Pro takový jazyk by pak byla verbální forma ideální formou. Vývoj takových jazyků, blízkých hovorovému, je však teprve v počátcích a dosavadní poznatky naznačují, že jejich syntax i sémantika budou zřejmě dosti složité.

Pro jednoduché komunikační jazyky, známé v současné době (jsou založeny na práci se jménem údaje, hodnotou údaje a vazbami mezi údaji), plně postačí tabulková forma zadávání parametrů. Tabulková forma zobrazování informací je pro svou náročnost, přehlednost a ilustrativnost známa a používána už po staletí. Uživatelé jsou s ní z vlastní praxe důvěrně obeznáni, při saškolení je tedy nemusíme učit prakticky ničemu novému. Tabulkové formy využila mj. také jedna z nejúspěšnějších programovacích firem Informatics Inc. v systému pro řízení práce s datovými soubory MARK IV.

Jedním z nejdůležitějších momentů při návrhu parametrického systému je volba okruhu zpracovávaných úloh a s této volby vyplývající rozsah zadávaných parametrů. Při návrhu musíme najít a navrhnout správné přeporce mezi univerzálností navrhovaného systému a mezi jednoduchostí jeho obsluhy a efektivností jeho práce. Rozhodně není správné, aby do okruhu úloh zpracovávaných systémem byly zahrnuty všechny úlohy, které se vyskytují v rámci zpracovávané problematiky. Do okruhu úloh by měly být zahrnuty jen "nejtypičtější, nejčastěji se vyskytující úlohy" v uživatelských požadavcích.

Systém by tedy neměl být koncipován tak, aby splnil 100 %

uživatelových požadavků. Správně navržený systém by měl splnit cca 90 - 95 % požadavků a tím, že ostatní budou realizovány jinými způsoby.

Zkušenosti totiž ukazují, že s nárůstem univerzality každého systému nad určitou rozumnou hranicí neúměrně nárůstají jak náklady a práce, spojené s implementací systému, tak i náročnost na zaškolení uživatele. Toto se zvláště nepříznivě projevuje u uživatele-neprogramátora.

Problém je možno demonstrovat na systému UPP. Při koncipování tabulky T1 nebyla záměrně uplatněna celá řada možností logických operací, jak je známe z vyšších programovacích jazyků. Může např. uplatnit logické "nebo" mezi dvěma různými údaji, toto zúžení možností je však vyváženo jednoduchostí zadávaných parametrů. S pomocí pouhých dvou hodnot dvou parametrů (rovnost-nerovnost, diskrétní hodnota-interval) jsou pokryty, jak ukázaly praktické zkušenosti, téměř všechny v praxi se vyskytující požadavky.

Klademe-li takový důraz na jednoduchost a co největší omezení rozsahu zadávaných parametrů, je třeba při implementaci takového systému klást velký důraz na jeho flexibilitu, tzn. předem vytýpat okruh možných budoucích změn. Je to důležité jak pro prvotní implementaci, tak i pro pozdější zásahy do programů.

Flexibilitu systému (a s ní spojené zásahy do programů) je třeba zajišťovat průběžně, nejlépe v osobním kontaktu s autory parametrického systému. Praktické zkušenosti ukazují, že bez osobního kontaktu dochází k nejasnostem a s nich vyplývajícím chybám i při sebelépe vedené dokumentaci takového systému.

## 2.2 Charakter údajů v parametrech

V uživatelsky orientovaných parametrech bychom měli v co největší míře omezit (pokud je nemožno úplně vyloučit) údaje, které nemají "uživatelský" charakter. Jde o údaje, které

souvisejí s programovým zpracováním úlohy, jako je pozice údaje ve větě, délka údaje v bytech apod. S touto problematikou se UPP vyrovnává pomocí tabulky popisu zpracovávaného datového souboru. Uživatel uvede do parametrů pouze symbolické jméno údaje a program pro předzpracování parametrů si už zjistí charakteristiky tohoto údaje (pozice ve větě, délka v bytech atd.) z uvedené tabulky popisu.

"Neuživatelské" údaje v parametrech jsou podstatně závažnějším problémem, než by se na první pohled zdálo. Zařazením takových údajů do parametrů nutíme totiž uživatele-programátora, aby zvládal problematiku, která je mu cizí a kterou pro vlastní vyřešení úlohy nepotřebuje a potřebovat nebude. Tato defakto sbytečná práce navíc vede někdy až k negativní motivaci uživatele na využívání systému.

### 3. Závěr

Problematika uživatelsky orientovaných parametrických systémů se oproti parametrickým systémům orientovaným programátorsky liší podstatněji pouze ve formě a obsahu zadávaných parametrů. Po stránce implementační a funkční zde velké rozdíly nejsou.

Základní problém spočívá v tom, abychom dokázali přiblížit formu zadávání parametrů myšlení uživatele. K demonstraci odchylek od běžných programátorských postupů postačí např. srovnání parametrů pro výběr vět v tabulce T1 s "výběrovými výroky" ve vyšších programovacích jazycích.

Systém UPP představuje pouze jakési první přiblížení k uživatelské formě parametrického systému. Zkušenosti s ním však naznačují, že implementace a využívání obdobných systémů je v praxi možné a že tyto systémy mohou být značným přínosem jak při vlastním zpracování úloh pomocí počítače (zpracování úloh, které by jinak zpracovány nebyly v důsledku omezené programátorské kapacity), tak i při "osvětové činnosti". Je totiž podstatný rozdíl v přístupu k využívání výpočetní techniky u pracovníků, kteří jen pasivně přijímají sestavy s počítače, a u pracovníků, kteří si možnosti počítače sami aktivně ověřili, byť i jen formou jednoduchých parametrů.