

PROGRAMOVÍ SYSTÉM PRO VYTVÁŘENÍ A AKTUALIZACI DATOVÝCH SOUBORŮ

Ing. Šárka Jiříčková, RNDr. Čestmír Losert
Vědeckovýzkumný úhelný ústav Ostrava-Radvanice

1. Úvod

Při aplikacích v oblasti vědeckotechnických výpočtů se vyskytuje často potřeba vytvářet datové soubory, které mají danou strukturu a dané umístění položek ve větách. Může se jednat např. o soubor obsahující vstupní data pro problémově orientované programy. Častým případem je také soubor obecnějšího charakteru, který obsahuje jistý seznam objektů a jejich vlastností, tedy to, co "majitelé" těchto dat často neoprávně ale hrdě nazývají bankou dat; u takového souboru se pak většinou předpokládá zpracování několika způsoby.

Ve všech takových a podobných případech lze potřebné datové soubory vytvořit libovolným editorem, ale tento způsob není příliš pohodlný, protože je při něm nutno mít stále na paměti pořadovanou strukturu souboru. Nevýhody tohoto způsobu se však projevují zejména tehdy, když např. různé vlastnosti téhož objektu jsou dostupné v různé době a z různých zdrojů a kdy je tedy nutno již vytvořený datový soubor průběžně aktualizovat a doplňovat.

Proto byly vypracovány dva programy, které umožňují poměrně pohodlným způsobem takové soubory vytvářet a aktualizovat, zejména ve výše uvedených případech.

Ukázalo se, že funkce těchto dvou programů lze vhodně kombinovat s funkcemi jiných programů, což umožnilo přirozeným způsobem vytvořit jakýsi systém, který bude v dalším stručně popsán.

2. Program pro vytváření datových souborů

Tento program vytváří datový soubor, jehož obsahem jsou hodnoty položek zadávané z terminálu. Takový soubor může obsahovat různé typy vět a je-li to požadováno, může každá věta v prvních čtyřech pozicích obsahovat identifikaci jejího typu. Hodnoty položek jsou zapisovány daným formátem buď do nově vytvářeného nebo na konec existujícího datového souboru.

Program pracuje s tzv. souborem popisu dat, který obsahuje definice všech použitých typů vět a položek. Tento soubor může mimo jiné být vytvářen nebo doplňován přímo při zadávání hodnot položek.

Věta určitého typu je charakterizována seznamem položek v ní obsažených. Předpokládá se, že položky jsou ve větě řazeny za sebou v takovém pořadí, v jakém se jejich definice vyskytují v souboru popisu dat. K identifikaci typů vět slouží jejich jména.

Každá položka je charakterizována jménem, délkou, typem a formátem zápisu.

Jméno slouží k identifikaci položky. Délkou položky se rozumí počet pozic, které při daném formátu zápisu ve větě zaujímá.

Jsou povoleny tyto typy položek:

- T ... textová položka, která se do výstupní věty zapisuje přesně tak, jak byla zadána,
- C ... znaková položka, která se do výstupní věty zapisuje po odstranění mezer,
- I ... číselná položka typu INTEGER,
- R ... číselná položka typu REAL,
- E ... číselná položka formálně typu REAL, které se přiřadí pořadové číslo věty, nebyla-li zadána.

Pro zápis položek do výstupní věty je možno použít tyto formáty:

- I ... zápis ve tvaru INTEGER, případně po konverzi a doání požadovaného počtu nul vpravo,
- F nebo E ... zápis ve tvaru REAL s požadovaným počtem desetinných míst,
- L nebo R ... zápis po odstranění mezer a zarovnání doleva nebo doprava s ponecháním daného počtu mezer vlevo nebo vpravo.

Vlastní zadávání hodnot probíhá tak, že si program nejdříve vyžádá zadání typu věty; není-li tento typ již definován, vyžádá si jeho definici. Poté zobrazí na terminále hlavičku s označením položek, nastaví kurzor na příslušné místo a očekává zadání hodnoty položky. Při formální chybě si vyžádá znovu její zadání, formálně správnou položku zobrazí na příslušném místě terminálu tak, jak bude zapsána do výstupní věty.

Při zadávání dat je možno pomocí speciálních hodnot položek uskutečnit různé akce jako např. kopii hodnoty téže položky z předešlé věty, návrat k předešlé položce ve větě, přechod k jinému typu věty a podobně.

Před ukončením činnosti program doplní daný soubor popisem dat o všechny nově definované typy vět a položek, bylo-li to požadováno.

3. Program pro aktualizaci datových souborů

Tento program provede aktualizaci datového souboru tak, že kopíruje vstupní soubor do výstupního s respektováním nových nebo doplněných hodnot položek obsažených v souboru změn.

Program pracuje za těchto předpokladů:

- Soubor popisu dat obsahuje definice všech použitých typů vět a položek.
- Vstupní datový soubor i soubor změn může obsahovat věty konstantního nebo proměnného typu, v druhém případě musí věty obsahovat identifikaci typu.
- Je zvolena jistá položka zvaná klíčová, tuto položku musí obsahovat každá věta vstupního souboru i souboru změn.
- Vstupní datový soubor a soubor změn musí být seříděn vzestupně podle hodnoty klíčové položky.
- Pokud je - podle dále uvedeného algoritmu - věta určitého typu aktualizována hodnotami položek věty jiného typu, musí být typ "změnové" věty částí typu "aktualizované" věty ve smyslu výskytu položek v obou větách.

Vlastní aktualizace probíhá zhruba podle následujícího algoritmu:

- Věta vstupního souboru, jejíž hodnota klíčové položky se nevyskytuje v souboru změn, je přeskopírována, resp. konvertována na "vyšší" typ bez změny hodnot položek.
- U věty vstupního souboru, jejíž hodnota klíčové položky se vyskytuje v souboru změn, jsou novými hodnotami přepsány ty položky, které byly ve změnové větě zadány. Hodnoty těch položek, které jsou ve změnové větě nezadány - prázdné - nebo se tam - v důsledku "nižšího" typu - nevyskytují, se nemění.
- Z vět souboru změn, jejichž hodnota klíčové položky se nevyskytuje ve vstupním souboru, jsou vytvořeny výstupní věty tak, že z první takové věty - ve smyslu výskytu v sérii změnových vět se stejnou hodnotou klíčové položky - je vytvořena věta, která je před zápisem do výstupního souboru postupně aktualizována dalšími změnovými větami se stejnou hodnotou klíčové položky - viz předchozí pravidla.

Program lze aplikovat i v případě, kdy neexistuje vstup-

ní datový soubor. V takovém případě dojde k aktualizaci dříve se vyskytujících hodnot později se vyskytujícími hodnotami spojené s případnou konverzí na vyšší typ věty.

Během své činnosti vytváří program pro aktualizaci datového souboru tzv. soubor nezpracovaných vět, který obsahuje všechny věty vstupního datového souboru a souboru změn, které pro formální chybu nemohly být zpracovány, doplněné v prvních dvou pozicích o kód původu a důvodu nezpracování této věty.

4. Pomocné programy

- Pro snadnější práci s programy popsanými v odstavcích 2, 3 byly dále vypracovány tyto jednoduché pomocné programy:
- Program, který vytváří pracovní verzi datového souboru, která obsahuje navíc pořadová čísla vět - vzhledem k umístění v souboru. Tento soubor se používá pro třídění podle hodnot klíčové položky, aby bylo zaručeno, že při třídění nedojde ke změně pořadí u vět, které mají stejnou hodnotu klíčové položky. Téhož programu lze po setřídění použít k odstranění pořadových čísel vět.
 - Program pro výpis nebo tisk seznamu vět a položek z daného souboru popisu dat. Tento seznam obsahuje kromě vlastností položek také jejich umístění ve větě, což bývá užitečné při práci s datovými soubory.

5. Další programy systému

S datovými soubory je někdy nutno provádět různé akce nezávislé na obsahu těchto souborů. Pro tyto akce je možno použít některých programů, z nichž jsou nejdůležitější:

Změna struktury vět datového souboru. Tento program provádí vnošení položek nebo částí věty, změnu umístění položek ve větě, dodání konstantních resp. prázdných položek do věty a dodání pořadového čísla věty do věty.

Výběr vět z datového souboru. Tento program vybírá z daného souboru ty věty, které splňují určité podmínky dané požadavkem, aby se v určitém místě věty vyskytoval alespoň jeden nebo nevyskytoval žádný z daných znakových řetězců, resp. kombinací těchto požadavků.

Rízený tisk datového souboru. Tento program provádí tisk datového souboru s volitelným výběrem položek pro tisk a s jejich volitelným umístěním v tiskové větě. Dále se připouští dodání konstantních položek a pořadových čísel vět do tiskové věty. Je-li to požadováno, interpretuje tento program danou část datové věty jako řídicí údaje pro tisk a provede pak na příslušném místě např. přechod na novou stránku, dodá daný počet prázdných vět a podobně.

Generování řídicích údajů pro tisk. Tento program vytváří modifikovanou verzi souboru vhodnou pro tisk výše uvedeným programem. Modifikace spočívá ve vkládání vygenerovaných "řídicích" vět před, za nebo místo datové věty, které splňují určité podmínky.

6. Vlastnosti systému

Přestože programy popsané v odstavci 5 byly vytvářeny v podstatě nezávisle na základních programech systému, ukazuje se, že jejich funkce vhodně doplňují funkce těchto základních programů. Vzniká tak přirozeným způsobem velmi jednoduchý, ale přitom poměrně účinný systém, který řeší řadu poměrně často se vyskytujících požadavků.

K vlastnostem tohoto systému lze uvést tyto poznámky:

Systém dovoluje poměrně pohodlným způsobem vytvářet datové soubory a danou strukturou vět a daným formátem položek.

Systém dovoluje poměrně snadno a bezpečně řešit ty případy, kdy údaje o vlastnostech jednotlivých objektů jsou získávány z různých zdrojů a v různém čase a kdy je tedy třeba příslušné datové soubory průběžně aktualizovat, doplňovat nebo slučovat.

Pevný formát položek v datovém souboru sice většinou není nutný pro aplikaci problémově orientovaných programů, má však obrovské výhody pro případ zpracování datových souborů programy popsanými v odstavci 5 nebo programy podobného charakteru. Tímto způsobem lze často některé požadavky realizovat téměř na počkání.

V důsledku silného zastoupení postupu zdola-nahoru při vytváření systému má tento kromě věcných nedostatků, které se jistě při další aplikaci projeví, také řadu formálních nedostatků. Z nich lze uvést zejména rozdílný přístup k datům u základních a doplňujících programů systému. Ten spočívá v tom, že u základních programů se pracuje s položkami pomocí jejich jmén a u doplňujících programů s číselnými věty definovanými vždy počátkem a délkou. Na druhé straně je však druhý způsob obecnější a může se někdy projevit jako výhodný nebo dokonce potřebný.

Za další formální nedostatek může být považována nejednotnost v zadávání řídicích údajů u programů z odstavce 5 - odpovědi na dotazy ve stanoveném pořadí nebo primitivní uživatelský jazyk. Tuto nejednotnost však lze do jisté míry odvodnit různým charakterem a množstvím potřebných řídicích údajů.

Popsaný systém se v současné době jeví jako relativně uzavřený. Lze však očekávat, že bude-li to třeba, nebude jeho

rozšíření o další programy obecného charakteru, resp. záměna některých programů za programy s obecnějšími funkcemi činit zásadní potíže.

7. Závěr

Popsaný systém může sloužit jako příklad toho, že i poměrně jednoduchými prostředky a dokonce s jistou mírou živelnosti lze vybudovat něco, co jednak splňuje svůj účel a jednak je schopno plnit řadu dalších funkcí, které tvoří rozsáhlé zobecnění původně zadaných požadavků.

To by na první pohled mohlo vypadat jako popření leccoho, co bylo na tomto semináři v uplynulých letech řečeno a prohlásováno. Když se však nad tím zamyslíme, zjistíme, že mimo jiné právě soustavné, byť někdy neuvědomělé dodržování některých zásad a postupů umožnilo původně samostatné objekty považovat za součásti celku s kvalitativně vyšší úrovní.

Konečně může tento příspěvek sloužit jako příklad toho, že i v oblasti vědeckotechnických výpočtů, kde to asi dosud není tak zcela samozřejmé, je užitečné se zamyslet nad otázkami, které nesouvisí přímo s věcnou stránkou řešeného problému.