

І. А л е ф е л ь д І., Х е р ц б е р г е р Ю. Введение в интервальные вычисления. М., 1987. 2. В е н г е р с к и й П.С., С е н ь о П.С. Интервальный метод решения систем нелинейных уравнений, базирующийся на предельных теоремах о среднем. Львов. 1990. 23 с. Рукопись деп. в УкрНИНТИ. 3. Krawczyk R. Intervalliterationsverfahren. Ber. Math.// Statist. Sek. Forschungzent. Graz, 1982. Vol. 185-189. S. 1-49.

Стаття надійшла до редколегії 23.10.90

ДК 681.03

В.С.Костирко, О.В.Костів

ДО ПИТАННЯ

ПРО УКРАЇНІЗАЦІЮ ПЕРСОНАЛЬНИХ КОМП'ЮТЕРІВ

Широке впровадження української мови у всі сфери життя на Україні /особливо в навчальний процес, науку, діловодство/ та прискорена комп'ютеризація різноманітних галузей роблять актуальною проблему реалізації в комп'ютерах українського алфавіту /українізації комп'ютерів/.

У зв'язку з цим необхідно:

- розр'язити внутрішнє кодування українських букв у комп'ютері;
- українізувати клавіатуру пристроїв вводу і підготовки даних так, щоб при натисненні на клавіші з українськими буквами генерувалися відповідні внутрішні коди;
- українізувати пристрої виводу /дисплей, принтер/, щоб при передачі на них кодів українських букв генерувалися відповідні зображення.

Через велику різноманітність типів комп'ютерів не може бути єдиного підходу до розв'язання цих задач. Тому, з огляду на домінуюче використання в Україні персональних комп'ютерів типу IBM PC і сумісних з ними, в даному дослідженні будемо розглядати лише ці комп'ютери. Всі символи в IBM PC кодуються згідно з таблицею ASCII - числами від 0 до 255, причому латинські, службові і розділові символи кодуються від 0 до 127 /перша половина таблиці/. Друга половина ASCII-таблиці /коди від 128 до 255/ виділена для розміщення букв національних алфавітів, псевдографічних, математичних та інших спеціальних символів.

© Костирко В.С., Костів О.В., 1991

Аналогічна проблема стояла раніше при реалізації на комп'ютерах типу IBM PC російського алфавіту, тому доцільно врахувати тут досвід її розв'язання. Для розміщення в ASCII-таблиці символів російського алфавіту не знайдено єдиного загальноприйнятого рішення, тепер використовуються основний, альтернативний, болгарський та інші варіанти кодування. Однак спільним для всіх варіантів є те, що букви російського алфавіту розміщуються в другій половині таблиці в порядку А, Б, В, ...Ю, Я, а, б, в, ..., ю, я, Е, Ё. Деяку перевагу перед іншими має альтернативний варіант розміщення російських букв, оскільки в ньому не змінюється положення псевдографічних символів, які використовуються для графічного оформлення екранів, отже, він забезпечує кращу сумісність з англійським програмним забезпеченням.

Виходячи з архітектури комп'ютера IBM PC та операційної системи MS DOS, можливі наступні підходи до розв'язання цієї проблеми:

1. Замінити в ASCII-таблиці латинський алфавіт українським, розмістивши його у першій і частково - у другій половині таблиці.

2. Додати українські букви до латинського алфавіту, розмістивши їх у другій половині таблиці приблизно так, як там розміщений російський алфавіт /очевидно, відштовхуючись від альтернативного варіанту/.

3. Додати до російськомовного варіанту ASCII-таблиці * /в другій її половині/ декілька специфічних українських букв /і, ї, є, ґ/: наприклад, присвоїти букві А код 128, букві Б - код 129 і т.д.; букві п - код 175, букві р - код 224, букві с - код 225 і т.д.; букві ё - код 241, ґ - 242, ґ - 243, 6 - 244, є - 245, і - 246, і - 247, І - 248, ї - 249.

Перший варіант розміщення українських символів має той недолік, що в ньому втрачається сумісність з англійським програмним та інформаційним забезпеченням. Навіть операційна система в такому разі повинна бути українськомовною. У теперішній ситуації, коли в усьому світі напрацьовано сотні тисяч пакетів програм - в основному англійськомовних - цей варіант видається незадовільним і відкидається апіорі.

* Б р я б р и н В.М. Програмное обеспечение ЭВМ. М., 1989

Другий варіант кодування українських букв забезпечує хорошу сумісність з англomовним програмним забезпеченням, але не забезпечує сумісності з російськомовним. У теперішній конкретній ситуації цей варіант поки що виглядає неприйнятним, оскільки більшість програмного забезпечення, яке офіційно поширюється і продається в Україні, з російськомовним. Однак з часом, коли сформується ринок україномовних програмних засобів, цей варіант, можливо, стане найбільш привабливим.

Отже, на даний час найбільш реальним видається третій варіант розміщення букв українського алфавіту. Проте цей варіант не забезпечує такої принципової особливості кодування, як лексикографічне впорядкування слів, тому його варто розглядати як проміжний етап на шляху до другого варіанта. Найпростіший спосіб переходу від третього до другого варіанта кодування полягає в обробці початкових текстів /програм, даних, документів і т.п./ спеціальними перекодувальниками.

Відзначимо, що жоден зі згаданих варіантів розміщення букв російського алфавіту не забезпечує правильного лексикографічного впорядкування слів з двох причин: по-перше, через розміщення букв Ё, ё після всіх інших букв, - тому слово "ёж" буде розміщуватися в кінці словника, після слів "яблоко" та "ягода"; по-друге, через те, що всі великі букви /окрім Ё/ передують всім маленьким буквам, - тому всі власні імена /Полтава, США, Мерседес/ будуть знаходитися на початку словника, перед словами "американець" чи "абажур".

З цієї причини при другому варіанті кодування українських букв видається доцільним розмістити їх в ASCII-таблиці в строго алфавітному порядку: А, а, Б, б, В, в, ..., Я, я, Ь, ь. З огляду на перспективу відновлення в українській мові букви ґ у цій таблиці також доцільно передбачити місце для символів ґ, ґ, розмістивши їх між буквами г і д. В результаті одержимо наступний варіант кодування букв українського алфавіту: А - 128, а - 129, Б - 130, б - 131, ..., у - 174, у - 175, Ф - 224, ф - 225, Х - 226, х - 227, ..., Ь - 240, ь - 241.

Для українізації пристроїв вводу /виводу інформації залежно від можливостей пристроїв застосовуються програмні або апаратні засоби /драйвери і "перешивання" мікросхем/. З міркувань можливого переходу в майбутньому на інший варіант кодування українських

букв, а також з міркувань забезпечення сумісності з різноманітними пакетами програм, при наявності відповідних технічних можливостей, бажано зберегти гнучкість у використанні зовнішніх пристроїв машини, віддаючи перевагу програмним засобам. При наявності технічних можливостей також слід надавати перевагу нерезидентним програмним засобам, які меншою мірою змінюють операційне середовище і тому забезпечують більший ступінь сумісності з іншими програмними забезпеченнями.

Найбільш популярним підходом до українізації клавіатури є резидентні драйвери. Проте, оскільки клавіатура належить до числа дешевих пристроїв, в майбутньому більш доцільним видається виробництво спеціалізованих українських клавіатур. Незалежно від способу реалізації і способу кодування букв, виходячи з інтересів кінцевого користувача, розміщення українських букв на клавіатурі повинно відповідати стандарту української друкарської машинки.

Перемикання клавіатури на український шрифт повинно здійснюватися натисненням на таку клавішу або комбінацію клавіш, яка б дозволила мінімізувати випадкове небажане перемикання. Оскільки момент перемикання шрифту дуже важливий для користувача, то перемикання варто супроводжувати короткою, характерною для кожного шрифту, мелодією. Для того щоб користувач не був змушений пам'ятати, в якому мовному середовищі він знаходиться, перемикання на латинський шрифт слід здійснювати іншою клавішею /або комбінацією клавіш/. З метою полегшення ідентифікації користувачем активного шрифту бажано постійно так чи інакше відображати його стан на екрані /наприклад, відображенням в лівому верхньому кутку символ-ідентифікатора шрифту, певним забарвленням рамки екрана/. Бажано, щоб драйвер мав засоби власної дезактивації /більшість поширених драйверів клавіатури можна дезактивувати лише перезавантаженням системи/.

Дисплей IBM PC може працювати у двох основних режимах: символному та графічному, причому основним є символний режим. Українізація дисплея програмними засобами для роботи в символному режимі можлива тільки при наявності у складі адаптера дисплея достатнього об'єму власної оперативної пам'яті /типу EGA, VGA і аналогічних/. При цьому драйвер може бути нерезидентним. Для полегшення ідентифікації букв у задачах, пов'язаних з сортуванням слів, бажано, щоб всі українські та російські букви відрізнялися за написанням від латинських букв /наприклад, і від i, р від p, Н від n/.

Виведення текстової інформації на принтер теж може здійснюватися у символному та графічному режимах. Оскільки виведення текстів у символному режимі реалізуються в декілька разів швидше, ніж у графічному, то будемо обговорювати тут саме цей варіант роботи. Українізація принтера нерезидентними програмними засобами можлива тільки при наявності у нього достатньої кількості завантажуваних символів, наприклад, принтери сімейства Star (Star ND-15, Star NX-10 і т.п.) дозволяють завантажити у власну оперативну пам'ять зображення мало не для всіх символів таблиці ASCII, а принтери сімейства Amstrad - зображення не більше 32 символів, чого недостатньо для розв'язання цієї задачі.

На жаль, більшість пакетів програм спеціальним чином інтерпретують деякі символи другої половини таблиці ASCII, що накладає певні обмеження на використання української мови при роботі з цими пакетами. Наприклад, пакети сімейств dBase та FoxBase розглядають коди I38 і I4I як керуючі, пакет Norton Commander - код 224, пакет PC Write - коди 246, 249, 250, пакет Data Easy та ряд інших пакетів - коди більші від 240. Зображення символів з кодами 244 та 245 не можуть бути змінені завантаженням у принтери сімейства Star. Для спілкування українською мовою з такими пакетами програм та принтерами можна піти двома шляхами:

- 1/ змінити кодування букв українського алфавіту так, щоб задовольнити вимоги пакетів програм чи принтерів;
- 2/ адаптувати програмне забезпечення так, щоб задовольнити вимоги кодування.

Оскільки кодування символів є інтерфейсом між багатьма користувачами, ключем для розуміння текстів та тісно пов'язане з пакетами програм, воно повинно бути продуманим і рідко змінюватися. Як і операційна система, це один з найбільш консервативних елементів програмного та інформаційного забезпечення комп'ютерів. Тому перший шлях видається неприйнятним та й, можливо, нездійсненним з огляду на велику кількість масових пакетів програм і суперечливість їх вимог до кодування.

Більш перспективним може стати другий шлях. Але для його реалізації з самого початку доцільно стати на шлях співробітництва з фірмами, які розробляють і поширюють відповідне програмне забезпечення, - в рамках сумісних підприємств контрактів, бартерних угод і т.п. Історія невдалого копіювання популярної серії

IBM 360/370 у вітчизняній серії машин ЕС ще раз підтверджує перспективність саме такого шляху українізації комп'ютерів.

Стаття надійшла до редколегії 16.07.90

УДК 519.21

І.Д.Квіт
ПОРІВНЯННЯ ДВОХ ВИБІРОК

Нехай

$$x_{(1)} < x_{(2)} < \dots < x_{(l)} < \dots < x_{(K)} \quad /1/$$

та

$$y_{(1)} < y_{(2)} < \dots < y_{(j)} < \dots < y_{(l)} \quad /2/$$

варіаційні ряди як зазгодом зрізаних вибірок обсягів відповідно m і n незалежних спостережень над абсолютно неперервними випадковими змінними ξ та η , де $\bar{l}, \bar{j}$ - середні ранги. Кожен з рядів /1/ та /2/ може бути побудований на основі, наприклад, повної вибірки, підмножини повної вибірки, справа зрізаної вибірки, з двох боків зрізаної вибірки і т.п. Слід перевірити гіпотезу про те, що вибірки однорідні, тобто, що ξ і η стохастично еквівалентні та, отже, мають однакову функцію розподілу ймовірностей

$$P\{\xi \leq x\} \equiv P\{\eta \leq x\}. \quad /3/$$

Якщо гіпотеза /3/ істинна, то перетворення рядів /1/ і /2/ за допомогою відповідно підбраного того самого розподілу ймовірностей $F(x)$ утворює два однорідні ряди з інтервалу $[0,1]$. На основі методики, описаної в праці [1], кожному з рядів $F(x_{(i)})$, $(i = 1, \dots, K)$ та $F(y_{(j)})$, $(j = 1, \dots, l)$ ставиться у відповідність варіаційний ряд повної вибірки

$$r_{x(1)}^* < r_{x(2)}^* < \dots < r_{x(i)}^* < \dots < r_{x(K)}^* \quad /4/$$

та

$$r_{y(1)}^* < r_{y(2)}^* < \dots < r_{y(j)}^* < \dots < r_{y(l)}^* \quad /5/$$

обсягів відповідно K та l .