

**ИНТЕЛЛЕКТУАЛИЗОВАННЫЙ ИНТЕРФЕЙС ПОЛЬЗОВАТЕЛЯ
ИНФОРМАЦИОННО-ПОИСКОВОЙ СИСТЕМЫ В ЗАДАЧЕ ПОИСКА
ПО КЛЮЧЕВОМУ СЛОВУ ("ОБРАЗЦУ") С УПРЕЖДАЮЩЕЙ ПОДСКАЗКОЙ**

***Анотація.** Розглядається типовий інтерфейс користувача у задачі пошуку за зразком (ключовим словом) з упереджувальною підказкою. Пропонуються математичні моделі оцінки трудомісткості варіантів інтерфейсу з покроковою та прицільною підказкою. Наводяться результати імітаційного моделювання процесів формування підказки.*

***Ключові слова:** інтерфейс користувача, пошук за ключовим словом, GOMS, упереджувальна підказка.*

***Аннотация.** Рассматривается типовой интерфейс пользователя в задаче поиска по образцу (ключевому слову) с упереджающей подсказкой. Предлагаются математические модели оценки трудоемкости вариантов интерфейса с пошаговой и прицельной подсказкой. Приводятся результаты имитационного моделирования процессов формирования подсказки.*

***Ключевые слова:** интерфейс пользователя, поиск по ключевому слову, GOMS, упереджающая подсказка.*

***Abstract.** A typical user interface for the task of sample (keyword) search with the ahead prompt is discussed. Mathematical models for estimation of interface complexity with the step-by-step and aim prompt are proposed. The results of simulations of the prompt formation are given.*

***Key words:** user interface, search keyword, GOMS, forward prompt.*

1. Введение

Типовая задача поиска по образцу, реализующая доступ к ресурсам информационно-поисковой (справочной) системы заключается в задании некоего образца (ключевого слова) и его нахождении в базовом словаре (БС) слов-эталонов. Целевые действия, выполняемые в случае нахождения (ненахождения) образца в БС, зависят от назначения системы поиска. Обычным действием общего характера является доступ к неким информационным ресурсам, связанным с соответствующим словом-эталоном.

Область приложения таких задач весьма широка – от простых систем учета товаров на складе до поисковых систем мультимедийных Web-ресурсов СППР.

В статье предлагаются подходы к оценке и снижению трудоемкости интерфейса, определяющей интеллектуальную нагрузку на пользователя (т.е. нагрузку, требующую определенных затрат умственного труда).

2. Типовой интерфейс пользователя (ИП) с упереджающей подсказкой образца

Упереджающая подсказка, являющаяся признаком "интеллектуализованности" ИП в принятом нами смысле, заключается в том, что по мере ввода начальных символов образца, вплоть до ввода "детерминанта" [1], однозначно определяющего образец, прогнозируются возможные варианты искомого эталона и предоставляются пользователю.

В основу такой подсказки положены две особенности БС:

- лексикографическая упорядоченность слов-эталонов;
- информационная избыточность.

Примем следующие обозначения:

$$A_j = (a_1 \dots a_i \dots a_n) - j - \text{е слово БС}; i = \overline{1, n}; j = \overline{1, N};$$

q – мощность множества символов (алфавита) составляющих слова – эталоны и образцы.

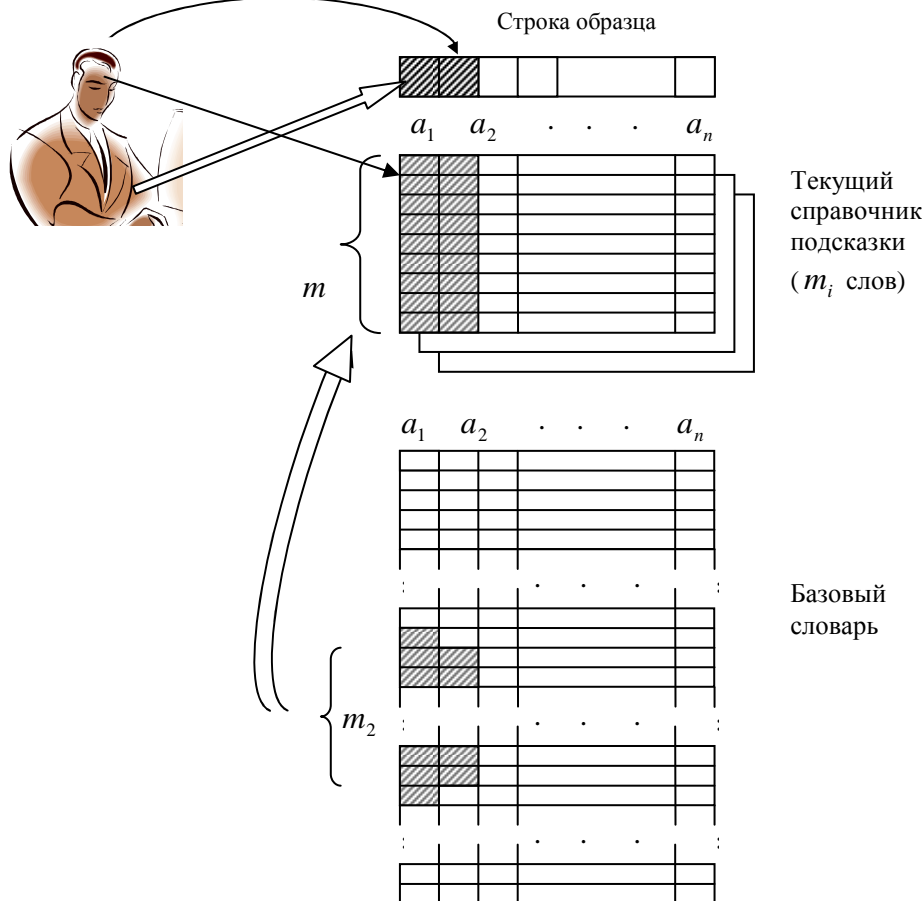
Лексикографическую упорядоченность слов БС определим следующим образом.

Каждому значению символа a_i поставим в соответствие число $\alpha_{k_i} = \{\alpha_0 \dots \alpha_{k_i} \dots \alpha_{(q-1)_i}\}$, такое, что $\alpha_0 = 0$; $\alpha_{q-1} = q - 1$; $\alpha_{k+1} = \alpha_k + 1$. Иными словами, символу a_i припишем значение порядкового номера этого символа в алфавите q . БС лексикографически упорядочен, если выполняется

$$\sum_{i=1}^n (\alpha_{k_i} \cdot q^{n-i}) > \sum_{j=i+1}^n (\alpha_{k_j} \cdot q^{n-i}) \quad (1)$$

Практически это означает, что слова-эталоны, интерпретируемые как числа в позиционной системе счисления с основанием q , упорядочены по убыванию значений.

Из [1] следует также, что $\frac{\alpha_{k_i}}{\alpha_{k_{i+1}}} = q$, т.е. принято, что символы $(\alpha \dots \alpha_i \dots \alpha_n)$ перенумерованы в порядке убывания старшинства.



Информационная избыточность слов БС означает, что из q^n всевозможных значений комбинаций n символов для представления реально существующих слов БС используется только N комбинаций, составляющих незначительную часть q^n ($q^n \gg N$). Это свойство позволяет идентифицировать искомое слово-эталон по части символов слова-образца (в частности, по "старшей" начальной части) и реализовать упреждающую подсказку пользователю.

Рис. 1. Общая схема интерфейса с упреждающей подсказкой

Общая схема рассматриваемых интерфейсов приведена на рис. 1.

Общий алгоритм интерфейсов по схеме рис. 1 заключается в следующем.

Пользователь последовательно вводит символы a_1, a_2, \dots, a_i образца (начиная с первого, старшего). На каждом шаге из БС в текущий справочник (возможно, виртуальный) помещаются слова с одинаковыми значениями символов $a_1, a_1 a_2, \dots, a_1 a_2 \dots a_i$.

Назовем множество слов с одинаковыми значениями символов $a_1 a_2 \dots a_i$ \hat{a}_i - множеством мощностью m_i . На схеме рис. 1 показана ситуация, когда в текущем справочнике находится \hat{a}_2 - множество. Из свойства лексикографической упорядоченности слов – эталонов ясно, что $m_1 \rangle m_2 \rangle \dots \rangle m_i$, т.е. область поиска образца сужается по мере ввода символов a_1, a_2, \dots .

В идеальном случае, при строго регулярной равномерной структуре словаря $m_i / m_{i+1} \approx q$. Практически распределение реально существующих значений слов-эталонов среди q^n всевозможных значений комбинаций символов $a_1 \dots a_n$ носит случайный характер, и, соответственно, значения m_i также носят случайный характер.

Относительно слов \hat{a}_i -множеств справедливы следующие очевидные положения.

Положение 1. \hat{a}_{i+1} -множество является подмножеством \hat{a}_i -множества.

Положение 2. Слово – эталон, принадлежащее \hat{a}_i -множеству, принадлежит и \hat{a}_{i+1} -множеству.

В рамках общей схемы рис. 1 возможны различные стратегии формирования и использования текущего справочника. Задача выбора стратегий и проектирования эффективных ИП связана с оценкой ожидаемой трудоемкости, определяющей интеллектуальную нагрузку на пользователя.

3. Подход к оценке трудоемкости ИП

Одним из лучших подходов к количественному анализу интерфейса «пользователь-компьютер» считается применение семейства классических моделей GOMS [2, 3]. Применительно к рассматриваемому классу задач, связанных с клавиатурным вводом символов и анализом текстового сообщения, отображаемого пользователю на экране, в [4, 5] предлагается уточнение модели GOMS KLM (Keystroke-Level Model) [2], основанное на декомпозиции ментальных операторов, определяющих основную часть интеллектуальной нагрузки на пользователя. Подход, описанный в [4, 5], иллюстрирует рис. 2, на котором приняты следующие обозначения для микрооператоров, описывающих «атомарные» действия пользователя в рассматриваемом классе задач.

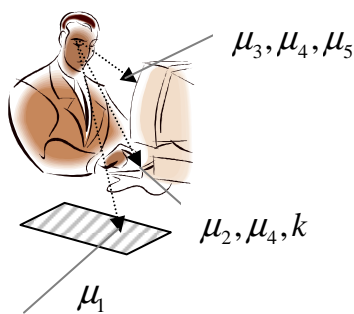


Рис. 2. Схема декомпозиции ментальных операторов

μ_1 – трудоемкость чтения текста с первичного носителя и его осмысление (запоминание) – с/симв;

μ_2 – трудоемкость поиска символов на клавиатуре и перемещение руки в позицию “над символом” – с/симв;

μ_3 – трудоемкость визуального анализа введенных символов на экране и принятие решения о дальнейших действиях (в частности, о наличии или отсутствии ошибки) – с/симв;

μ_4 – трудоемкость исправления ошибочного

символа on-line – с/симв;

μ_5 – трудоемкость визуального анализа и сравнения вводимого слова – образца с предлагаемым на экране словом – с/симв;

k – “чистая” трудоемкость нажатия клавиши рукой, расположенной над символом – с/симв.

По своей сути, операторы μ_1, μ_3, μ_5 – это чисто ментальные операторы, k – оператор движения, μ_2 и μ_4 – композиция ментальных действий и движений.

Таблица 1. Измеренные и рассчитанные значения μ, k

Язык	μ_1	μ_2	μ_3	μ_4	μ_5	k
КОД	0,47	0,38	0,214	1,06	0,0435	0,15
МНЕМОТЕКСТ	0,16	0,25	0,055	0,80	0,045	0,15

Результаты экспериментального определения значений $\mu_1 - \mu_5, k$, полученные в [5] для

неквалифицированных пользователей (не имеющих специальной подготовки в смысле машинописи) применительно к вводу и анализу цифровых кодов и мнемотекста (слов на русском языке, родном для пользователей – участников эксперимента), приведены в табл. 1.

4. Оценка трудоемкости ИП

Оценим трудоемкость N различных вариантов ИП на основе подхода и результатов, представленных в п. 2.

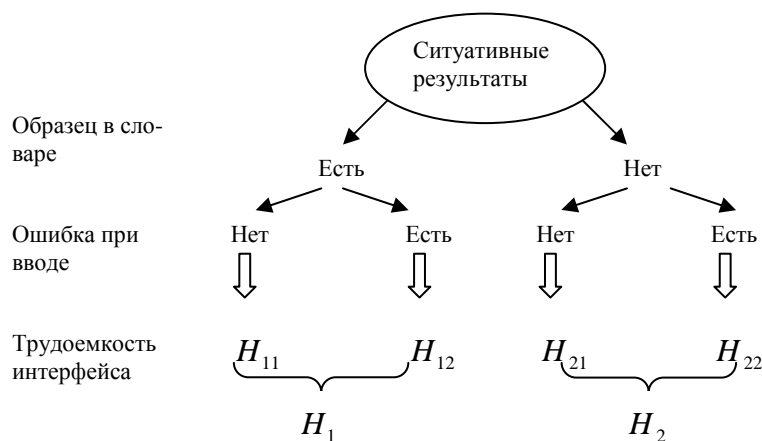


Рис. 3. Составляющие трудоемкости интерфейса искомого слова в БС).

Поскольку вероятность искажения символа при вводе данных сравнительно невелика (порядка $6-8 \cdot 10^{-3}$ [6]), ограничимся учетом относительно значимых составляющих $H_1 \approx H_{11}$ и $H_2 \approx H_{21}$, т.е. положим

$$H \approx \delta \cdot H_1 + (1 - \delta)H_2.$$

4.1. ИП с пошаговой подсказкой ("step-by-step")

В этом варианте интерфейса на каждом шаге ввода образца (т.е. ввода символов a_1, \dots, a_i) из общего текущего справочника объемом m_i слов пользователю предоставляется одна страница из m слов, выбранная по некоторому критерию. Подобный интерфейс используется, например, в поисковых системах GOOGLE, YANDEX, RAMBLER и др.

Примем следующие обозначения:

$\bar{v}_1^{(1)}, \bar{v}_2^{(1)}$ – среднее количество символов образца, вводимых до завершения процесса поиска в случае положительного и отрицательного результатов поиска соответственно;

$\bar{m}_1^{(1)}, \bar{m}_2^{(1)}$ – среднее количество слов, участвующих в процессе визуального анализа предъявляемых страниц текущего справочника в случае положительного и отрицательного результатов поиска.

Определим ориентировочные значения $H_1^{(1)}$ и $H_2^{(1)}$ на основе значений μ_2, μ_3, μ_5, k , приведенных в табл. 1.

$$H_1^{(1)} \approx (\mu_2 + k + \mu_3) \bar{v}_1^{(1)} + m_1^{(1)} \cdot n \cdot \mu_5 + (P + BB), \quad (2)$$

$$H_2^{(1)} \approx (\mu_2 + k + \mu_3) \bar{v}_2^{(1)} + m_2^{(1)} \cdot n \cdot \mu_5 + (\mu_2 + k). \quad (3)$$

Первое слагаемое в (2) и (3) определяет затраты времени на ввод и визуальную верификацию символов образца, второе – на визуальный анализ предъявляемых страниц текущего справочника, третье – на подтверждение того или иного результата поиска (P, BB – стандартные операторы GOMS подвода курсора и клика мышью).

Определим условную вероятность $P_{i_i}^{(1)}$ того, что на шаге i в предъявляемой случайной странице текущего справочника после ввода символа a_i будет найден образец, имеющийся в БС, при условии, что он не был найден на предыдущих шагах.

Предполагая случайным распределение N реально существующих слов БС среди q^n всевозможных значений, рассмотрим следующую модель.

В некую урну "с узким отверстием" вбрасывается M шаров. Вероятность того, что очередной шар попадает в отверстие урны, равна $r = \frac{N}{q^n}$. Поскольку эта вероятность не зависит от результатов предыдущих бросаний, общие вероятностные результаты вбрасываний могут быть описаны схемой и соотношениями процесса независимых испытаний Бернулли, в соответствии с которыми вероятность $P\left[\left(\overline{o, x}\right), r, M\right]$ получить в точности x или менее удачных исходов равна

$$P\left[\left(\overline{o, x}\right), r, M\right] = \sum_{g=0}^x C_M^g \cdot r^g \cdot (1-r)^{M-g}, \quad (4)$$

а среднее количество удачных исходов \bar{N}_M равно

$$\bar{N}_M = \sum_{g=0}^M g \cdot C_M^g \cdot r^g \cdot (1-r)^{M-g} = rM. \quad (5)$$

В соответствии с такой интерпретацией для $M = q^{n-i}$ примем

$$P_{i_i}^{(1)} \approx \frac{m}{r \cdot q^{n-i}}.$$

Безусловная вероятность $P_1^{(1)}(i)$ того, что образец будет найден именно на шаге i , равна

$$P_1^{(1)}(i) = P_{i_i} \prod_{s=1}^{i-1} (1 - P_{1_s}^{(1)}).$$

Если образец не был найден за $(I_m - 1)$ шагов, то на шаге I_m , когда область поиска сузится до значений $m_{1m} \leq m$, образец определенно будет обнаружен, т.е. мы можем положить $P_{1m}^{(1)} = 1$. Значение I_m определяется из условия $\frac{m}{r \cdot q^{n-1m}} \geq 1$.

Таким образом,

$$\bar{v}_1^{(1)} = \sum_{i=1}^{I_m-1} i \cdot P_1^{(1)}(i) + I_m \cdot 1 \cdot \prod_{s=1}^{I_m-1} (1 - P_{1_s}^{(1)}), \quad (6)$$

$$\bar{m}_1^{(1)} = \sum_{i=1}^{I_m-1} \left[(i-1) \cdot m + \frac{m+1}{2} \right] \cdot P_1^{(1)}(i) + \left[(I_m-1) \cdot m + \frac{m+3}{4} \right] \cdot 1 \cdot \prod_{s=1}^{I_m-1} (1 - P_{1s}^{(1)}), \quad (7)$$

где I_m – ближайшее целое, большее или равное $\log_q \left(\frac{q^n \cdot r}{m} \right) = \log_q \frac{N}{m}$.

Если образец в БС отсутствует, то на шаге i это становится известным, если при вбрасывании q^{n-i} шаров в урну попадает 0, 1, ..., m шаров. Следовательно,

$$P_{2i}^{(1)} = P\left[\overline{[0, m]}, r, q^{n-i}\right],$$

$$P_2^{(1)}(i) = P_{2i}^{(1)} \cdot \prod_{s=1}^{i-1} (1 - P_{2s}^{(1)})$$

и
$$\bar{v}_2^{(1)} = \sum_{i=1}^n i \cdot P_2^{(1)}(i). \quad (8)$$

Отсюда

$$\bar{m}_2^{(1)} = \sum_{i=1}^n \left[(i-1) \cdot m + \frac{m}{2} \right] \cdot P_2^{(1)}(i). \quad (9)$$

В выражениях (7), (9) принята равновероятность различных вариантов заполненности последней проверяемой страницы и результатов перебора представленных слов при поиске образца.

В табл. 2, 3 (МНЕМОТЕКСТ; $m = 10, q = 32, n = 8$) и 4 (КОД; $m = 10, q = 10, n = 12$) в качестве примера приведены результаты расчетов значений $P_{li}^{(1)}, P_1^{(1)}(i)$ и $H^{(1)}, \bar{v}_1^{(1)}, \bar{m}_1^{(1)}, \bar{v}_2^{(1)}, \bar{m}_2^{(1)}$ для выбранных типовых значений параметров q и N , соответствующих словам русского языка и цифровым кодам. В частности, для русского языка принято: средняя длина слов БС $n = 8$ символов, количество слов в типовом словаре $N = 10^5$ [7]. Отсюда для $q = 32, n = 8, N = 1,1 \cdot 10^5$ величина $r \approx 10^{-7}$. Для иллюстрации трендов взяты два значения $N = 1,1 \cdot 10^6$ ($r = 10^{-6}$) и $N = 1,1 \cdot 10^4$ ($r = 10^{-8}$). Аналогичные значения r выбраны и для гипотетических цифровых кодов.

Таблица 2. Значения $p_{li}^{(1)}, p_1^{(1)}(i)$: МНЕМОТЕКСТ

r	N	$P^{(1)}$	i			
			1	2	3	4
10^{-6}	$1,1 \cdot 10^6$	$P_{li}^{(1)}$	0,0003	0,0093	0,2894	1
		$P_1^{(1)}(i)$	0,0003	0,0093	0,2866	0,7038
10^{-8}	$1,1 \cdot 10^4$	$P_{li}^{(1)}$	0,0290	0,8519	1	-
		$P_1^{(1)}(i)$	0,0290	0,8272	0,1433	-

Таблица 3. Значения $H^{(1)}, \bar{v}_1^{(1)}, \bar{m}_1^{(1)}, \bar{v}_2^{(1)}, \bar{m}_2^{(1)}$: МНЕМОТЕКСТ

r	N	δ		$\bar{v}_1^{(1)}$	$\bar{m}_1^{(1)}$	$\bar{v}_2^{(1)}$	$\bar{m}_2^{(1)}$
		1,0	0,75				
10^{-6}	$1,1 \cdot 10^6$	14,08	14,31	3,69	30,85	4,00	35,50
10^{-8}	$1,1 \cdot 10^4$	8,13	8,34	2,11	16,31	2,51	20,58

Таблица 4. Значения $H^{(1)}, \bar{v}_1^{(1)}, \bar{m}_1^{(1)}, \bar{v}_2^{(1)}, m_2^{(1)}$: КОД

r	N	δ		$\bar{v}_1^{(1)}$	$\bar{m}_1^{(1)}$	$\bar{v}_2^{(1)}$	$\bar{m}_2^{(1)}$
		1,0	0,75				
10^{-6}	$1 \cdot 10^6$	28,31	28,85	4,96	44,67	5,42	49,67
10^{-8}	$1 \cdot 10^4$	16,50	17,02	2,96	24,90	3,42	29,67

Как видно из данных табл. 3, 4, трудоемкость рассматриваемого интерфейса непропорционально велика. Например, в поисковых системах GOOGLE, YANDEX, RAMBLER и др. поиск и предоставление пользователю информационных ресурсов, связанных с образцом, требует менее секунды, а ввод и идентификация образца – до десяти и более секунд. Правда, на одном из шагов ввода до идентификации образца пользователь может обнаружить в словах-эталонах предъявляемой страницы что-то подходящее по смыслу и, таким образом, ввести в итоге меньше символов. Однако это требует дополнительных размышлений, более длительных, чем просто сравнение эталона с задуманным образцом ($n \cdot \mu_5$). Кроме того, подобные размышления практически неприемлемы при вводе кодов.

Поэтому в целом трудоемкость ИП в рассматриваемом варианте представляется уместным оценить как относительно высокую. Тенденции, определяющие зависимости $\bar{v}_1^{(1)}, H_1$ от значений r, N, δ и алфавита образца, ясны из данных табл. 2–4.

4.2. ИП с прицельной подсказкой (“taking aim”)

Возможность снижения трудоемкости ИП с пошаговой подсказкой основана на следующих предпосылках.

Рассмотрим дискретную функцию $P_i(i)$, представленную в табл. 2.

Из данных табл. 2 видны следующие явные свойства рассмотренного процесса пошаговой подсказки

1. Значения $P_1^{(1)}(1)$, а для $r = 10^{-6}$ $P_1^{(1)}(1)$ и $P_1^{(2)}(1)$ малы, так что ввод начальных символов образца делает относительно малый вклад в общую вероятность успешного завершения процесса (т.е. обнаружения и идентификации образца).

2. Вероятность успешного завершения процесса максимальна в районе значений i , близких к $\bar{v}_1^{(1)}$.

Отсюда следует простой, чисто пользовательский, прием ускорения – ввод начальных символов образца “вслепую” и задержка начала просмотра страниц подсказки (из Положений 1,2 ясно, что образец “никуда не денется”).

Эффективной программной реализацией этого приема (обозначим трудоемкость соответствующего ИП через $H^{(2)}$) должно быть предоставление пользователю для просмотра лишь последней страницы, объем которой $\leq m$, или, по крайней мере, информирование пользователя о появлении этой страницы (например, звуковым сигналом).

Еще более эффективным представляется решение, заключающееся в автоматическом сужении области поиска до одного искомого слова (обозначим трудоемкость ИП через $H^{(3)}$).

Оценим значения $H^{(2)}$ и $H^{(3)}$ для отмеченных вариантов ИП на основе базового выражения (1).

4.2.1. Оценка $H^{(2)}$

Для оценки значений $H^{(2)}$ справедливы выражения (2), (3) при условии подстановки соответствующих значений $\bar{v}_1^{(2)}, \bar{v}_2^{(2)}, \bar{m}_1^{(2)}, \bar{m}_2^{(2)}$. Определим эти значения. Для этого вернемся к модели вбрасывания шаров в урну.

Поскольку пользователь обращается к экрану лишь при выполнении условия $m_i \leq m$, условная вероятность $P_i^{(2)}$ завершения процесса на шаге i определяется через вероятность того, что на шаге i при вбрасывании q^{n-i} шаров в урне окажется не более m шаров. При определении значений $P_{1i}^{(2)}$ и $P_{2i}^{(2)}$ учтем следующее.

В случае потенциально положительного результата поиска в БС есть искомый образец, и среди m_i слов одно имеет не случайное, а детерминированное происхождение. Это означает, что один шар кладется в урну заранее, а $0, 1, \dots, m-1$ шаров попадают случайным образом, определяя удачные исходы испытаний Бернулли (т.е. в выражении (4) $x := m-1$).

В случае отрицательного результата поиска детерминированный шар в урне отсутствует, и следует $x := m$.

Таким образом,

$$P_{1i}^{(2)} = P\left[\left(\overline{0, m-1}\right), r, q^{n-i}\right],$$

$$P_{2i}^{(2)} = P\left[\left(\overline{0, m}\right), r, q^{n-i}\right],$$

$$P_{1,2}^{(2)}(i) = P_{1,2,i}^{(2)} \prod_{s=1}^{i-1} (1 - P_{1,2,s}^{(2)}).$$

Полагая $\prod_{s=1}^n (1 - P_{1,2,s}^{(2)}) \approx 0$ (т.е. процесс определенно заканчивается при $i \leq n$) и принимая прежнее упрощающее допущение о равновероятности вариантов заполненности проверяемой страницы и результатов перебора представленных слов, запишем следующие выражения для искомых значений:

$$\bar{v}_{1,2}^{(2)} = \sum_{i=1}^n i \cdot P_{1,2}^{(2)}(i); \quad \bar{m}_1^{(2)} \approx \frac{m+3}{4}; \quad \bar{m}_2^{(2)} \approx \frac{m}{2}.$$

В табл. 5 и 6 приведены результаты расчетов для образцов вида МНЕМОТЕКСТ и КОД и прежних наборов значений параметров m, q, n .

Таблица 5. Значения $H^{(2)}, \bar{v}_1^{(2)}, \bar{m}_1^{(2)}, \bar{v}_2^{(2)}, \bar{m}_2^{(2)}$: МНЕМОТЕКСТ

r	N	δ		$\bar{v}_1^{(2)}$	$\bar{m}_1^{(2)}$	$\bar{v}_2^{(2)}$	$\bar{m}_2^{(2)}$
		1,0	0,75				
10^{-6}	$1.1 \cdot 10^6$	4,29	4,27	4,00	3,25	3,99	5,00
10^{-8}	$1.1 \cdot 10^4$	3,67	3,63	2,63	3,25	2,51	5,00

Таблица 6. Значения $H^{(2)}, \bar{v}_1^{(2)}, \bar{m}_1^{(2)}, \bar{v}_2^{(2)}, \bar{m}_2^{(2)}$: КОД

r	N	δ		$\bar{v}_1^{(2)}$	$\bar{m}_1^{(2)}$	$\bar{v}_2^{(2)}$	$\bar{m}_2^{(2)}$
		1,0	0,75				
10^{-6}	$1 \cdot 10^6$	7,12	7,20	5,54	3,25	5,42	5,00
10^{-8}	$1 \cdot 10^4$	5,63	5,71	3,54	3,25	3,42	5,00

4.2.2. Оценка $H^{(3)}$

Рассматриваемый вариант интерфейса является частным случаем предыдущего. Здесь $x = 0$ как для положительного, так и отрицательного результата поиска. При попадании в пресловутую урну 0 случайных шаров в случае положительного результата в урне оказывается только один детерминированный шар, а в случае отрицательного – 0 шаров.

Таким образом,

$$P_{1,2i}^{(3)} = P\left[\left(\overline{0,0}\right), r, q^{n-i}\right],$$

$$P_{1,2}^{(3)}(i) = P_{1,2i}^{(3)} \prod_{s=1}^{i-1} (1 - P_{1,2,s}),$$

$$\bar{v}_{1,2}^{(3)} = \sum_{i=1}^n i \cdot P_{1,2}^{(3)}(i),$$

$$\bar{m}_1^{(3)} = 1, \bar{m}_2^{(3)} \approx 0.$$

В табл. 7 и 8 приведены результаты расчетов для прежних видов образцов и наборов значений параметров

Таблица 7. Значения $H^{(3)}, \bar{v}_1^{(3)}, \bar{m}_1^{(3)}, \bar{v}_2^{(3)}, \bar{m}_2^{(3)}$: МНЕМОТЕКСТ

r	N	δ		$\bar{v}_1^{(3)}$	$\bar{m}_1^{(3)}$	$\bar{v}_2^{(3)}$	$\bar{m}_2^{(3)}$
		1,0	0,75				
10^{-6}	$1,1 \cdot 10^6$	2,89	2,80	4,67	1	4,67	~ 0
10^{-8}	$1,1 \cdot 10^4$	2,26	2,17	3,29	1	3,29	~ 0

Таблица 8. Значения $H^{(3)}, \bar{v}_1^{(3)}, \bar{m}_1^{(3)}, \bar{v}_2^{(3)}, \bar{m}_2^{(3)}$: КОД

r	N	δ		$\bar{v}_1^{(3)}$	$\bar{m}_1^{(3)}$	$\bar{v}_2^{(3)}$	$\bar{m}_2^{(3)}$
		1,0	0,75				
10^{-6}	$1 \cdot 10^6$	6,03	5,9	6,69	1	6,69	~ 0
10^{-8}	$1 \cdot 10^4$	4,54	4,41	4,69	1	4,69	~ 0

5. Заключение

Исходя из полученных результатов, мы можем сделать следующие выводы.

1) Рассмотренные модели позволяют получить ориентировочные значения существенных параметров $P(i), \bar{v}(i), \bar{m}$ и общие значения трудоемкости H для различных видов ИП с упреждающей подсказкой и различных значений параметров БС.

Как видно из данных табл. 3–8, на фоне общей полезности упреждающей подсказки как способа снижения трудоемкость и повышение usability ИП в задачах поиска по образцу, прицельная подсказка обеспечивают существенно меньшую трудоемкость, чем пошаговая (в частности, с произвольным выбором предъявляемой пользователю страницы текущего справочника). Этот вывод иллюстрирует табл. 9, обобщающая результаты расчетов. В табл. 9 представлены относительные значения коэффициентов $h_{21} = H^{(2)} / H^{(1)}$ и

$h_{31} = H^{(3)} / H^{(1)}$, мало зависящие от квалификации пользователя.

Таблица 9. Обобщенные результаты расчетов

h	r	МНЕМОТЕКСТ			КОД		
		N	δ		N	δ	
			1,0	0,75		1,0	0,75
h_{21}	10^{-6}	$1,1 \cdot 10^6$	0,30	0,30	$1 \cdot 10^6$	0,25	0,25
	10^{-8}	$1,1 \cdot 10^4$	0,45	0,43	$1 \cdot 10^4$	0,34	0,34
h_{31}	10^{-6}	$1,1 \cdot 10^6$	0,20	0,20	$1 \cdot 10^6$	0,21	0,20
	10^{-8}	$1,1 \cdot 10^4$	0,28	0,26	$1 \cdot 10^4$	0,27	0,26

Данные таблиц 3-8 также определяют общие тенденции зависимостей H от r, N, q .

2) Выражения (6) ÷ (9) дают заведомо приближенные значения вычисляемых переменных, в частности, для ИП с пошаговой подсказкой, поскольку при определении $P_{i_i}^{(1)}$ мы оперируем не вероятностями случайных значений количества удачных исходов, т.е. попавших в урну шаров, а их средними значениями (расчеты по более точным формулам неприемлемо сложны из-за высокой размерности). Результаты имитационного моделирования ИП с пошаговой и целевой подсказками, проведенные для оценки погрешностей, связанных с отмеченными выше и другими принятыми допущениями, показали относительно небольшое расхождение между расчетными и экспериментальными данными. В табл. 10 для $\delta = 1,0$ сведены значения \bar{v}_1 , принятые в качестве основного критерия сравнительной оценки. Как видно, расхождение не превышает 3 ÷ 4%. Примерно в таких же пределах находится и расхождение значений \bar{m}_1 .

Таблица 10. Результаты имитационного моделирования интерфейсов

Образец	r	N	Расчетные значения			Экспериментальные значения		
			$\bar{v}_1^{(1)}$	$\bar{v}_1^{(2)}$	$\bar{v}_1^{(3)}$	$\bar{v}_1^{(1)}$	$\bar{v}_1^{(2)}$	$\bar{v}_1^{(3)}$
МНЕМО-ТЕКСТ	10^{-6}	$1.1 \cdot 10^6$	3,69	4,00	4,67	3,70	4,02	4,74
	10^{-8}	$1.1 \cdot 10^4$	2,11	2,63	3,29	2,17	2,65	3,30
КОД	10^{-6}	$1 \cdot 10^6$	4,96	5,54	6,69	5,02	5,55	6,76
	10^{-8}	$1 \cdot 10^4$	2,96	3,54	4,69	3,03	3,55	4,75

3) Трудоемкость ИП с пошаговой подсказкой может быть снижена благодаря учету вероятностей обращений к словам БС и выбору на каждом шаге страницы с максимально вероятными искомыми словами. Степень снижения зависит от гипотетической функции распределения вероятностей обращений. Для оценки ожидаемых характеристик в этом случае требуется соответствующее усложнение рассматриваемой выше модели. Соответственно усложняется и интерфейс.

4) В ИП с прицельной подсказкой пользователь ограничивается в возможностях обозрения лексикографически и ментально «близких» вариантов задаваемого образца и выбора более подходящего (при недостаточно четком представлении о том, что же именно пользователю нужно). В связи с этим возможен гипотетический компромиссный вариант ИП – представление прицельной страницы, дополненной наиболее вероятными, как объект поиска, словами, близкими к запрошенному образцу.

СПИСОК ЛИТЕРАТУРЫ

1. Белоус Л.В. Модель упреждающей подсказки в интерфейсе пользователя / Л.В. Белоус, В.А. Литвинов, С.Я. Майстренко // Математичні машини і системи. – 2004. – № 3. – С. 156 – 163.
2. Kieras D. Using the Keystroke-Level Model to Estimate Execution Times [Электронный ресурс] / D. Kieras. – University of Michigan. – Режим доступа: <ftp://www.eecs.umich.edu/people/rchong/kieras/GOMS/KLM.pdf>.
3. A Guide to GOMS Model Usability Evaluation using GOMSL and GLEAN4 Revision [Электронный ресурс]. – 2006. – March 31. – Режим доступа: ftp://www.eecs.umich.edu/people/kieras/GOMS/GOMSL_Guide.pdf
4. Кузьменко Г.Е. Декомпозиция ментальных операторов в моделях GOMS-KLM применительно к интерфейсу пользователя в задачах ввода и контроля данных / Г.Е. Кузьменко, В.А. Литвинов, И.Н. Оксанич // Интеллектуальный анализ информации. IX междунар. конф. им. Т.А. Таран. ИАИ-2009, (Киев, 19–22 мая 2009 г.). – К., 2009. – С. 212 – 218.
5. Оксанич И.Н. Модель декомпозиции ментальных операторов в проблемно-ориентированном интерфейсе пользователя и ее экспериментальное исследование / И.Н. Оксанич // Математичні маши-

ни і системи. – 2010. – № 1. – С. 105 – 112.

6. Литвинов В.А. Контроль достоверности и восстановления информации в человеко-машинных системах / В.А. Литвинов, В.В. Крамаренко. – Киев: Техника, 1986. – 200с.

7. Широков В.А. Інформаційна теорія лексикографічних систем / Широков В.А. – Київ: Довіра, 1990. – 331 с.

Стаття надійшла до редакції 11.10.2010