

УДК 004.5

О ВЫБОРЕ ОБЪЕМА ПОРЦИИ ПОДСКАЗКИ В ЗАДАЧЕ ПОИСКА ПО КЛЮЧЕВОМУ СЛОВУ

В.А. Литвинов, С.Я. Майстренко, И.Н. Оксанич

Институт проблем математических машин и систем НАН Украины

e-mail: io2004@bigmir.net

Один из наиболее распространенных быстрых способов доступа пользователя к информационным ресурсам заключается в задании ключевого слова («образца»), с которым через некий базовый словарь (БС) связаны ключевые слова, индексы и ресурсы (в простейшем случае ресурсом может быть сам словарь, дополненный атрибутами, характеризующими каждое слово).

Типовым механизмом повышения эффективности (usability) интерфейса пользователя (ИП) является подсказка возможных вариантов ключевого слова по мере ввода начальных символов, имеющих совпадающее начало с введенными символами. В [1] рассмотрены модели «пошаговой» и «прицельной» подсказки, приводятся сравнительные оценки трудоемкости (производительности) соответствующих вариантов ИП на основе общей технологии GOMS-KLM [2] и уточняющих результатов [3, 4].

Представляет интерес вопрос о выборе объема «порции» (страницы) подсказки m . Частичный ответ на этот вопрос (в контексте оценок трудоемкости ИП) дают модели [1] и их обобщение на случай известного (заданного) распределения вероятностей обращений к словам БС.

Данные таблицы иллюстрируют зависимости от m существенных параметров ИП для неопытного пользователя. В таблице приняты следующие обозначения:

q - алфавит представления ключевых слов (КС);

N - объем базового словаря КС;

n - количество символов КС;

$\bar{H}$ - средняя трудоемкость ввода КС;

$\bar{v}_1$ - среднее количество введенных символов КС в процессе подсказки;

$\bar{m}_1$ - среднее количество слов, просмотренных пользователем;

λ^0 - параметр экспоненциальной функции, аппроксимирующей плотность распределения вероятностей обращения к КС. Для равномерного распределения принято $\lambda^0=0$; значение $\lambda^0=10$ примерно соответствует закону Парето 20/80.

Как следует из данных таблицы (и аналогичных данных для других значений N), изменение объема страницы подсказки вызывает двоякий эффект.

Так, уменьшение m , с одной стороны, влечет за собой увеличение среднего количества введенных символов, а, с другой, – относительно более значительное уменьшение суммарного количества просматриваемых слов; в результате значение $\bar{H}$ уменьшается. В частности, для $m=1$ значение $\bar{H}$ приближается к соответствующим минимальным оценкам прицельной подсказки с точностью до искомого слова [1]. Иными словами, чем меньше m , тем меньше трудоемкость ввода ключевого слова.

Для полного ответа на поставленный вопрос необходимо учесть трудно формализуемые ментальные аспекты визуального анализа порции подсказки. Эти аспекты заключаются в возможности обозрения пользователем лексикографически и ментально «близких» вариантов задаваемого ключевого слова и выбора более подходящего (при недостаточно четком представлении о том, что же именно пользователю нужно). Эта возможность минимально при $m=1$ для мнемотекстов [3, 4] и не имеет практического значения для кодов.

Значения H , $\bar{v}_1$, $\bar{m}_1$ в зависимости от m , r , λ^0

Образец	Подсказка	λ^0	m	$\bar{V}_1$	$\bar{m}_1$	H
ТЕКСТ $q = 32$ $N = 1,1 \cdot 10^4$ $n = 8$	Пошаговая	0	1	3,14	2,93	3,78
			5	2,55	9,87	6,01
			10	2,11	15,69	7,91
			15	1,96	18,94	9,01
		10	1	2,39	1,83	3,04
			5	1,88	6,31	4,42
			10	1,75	11,9	6,38
			15	1,65	14,8	7,38
	Прицельная с точностью до m	-	1	3,29	1	3,16
			5	2,98	2	3,38
			10	2,63	3,25	3,67
			15	2,13	4,5	3,89
	Прицельная с точностью до 1	-	1	3,29	1	3,16
			5	3,29	1	2,26
			10	3,29	1	2,26
			15	3,29	1	2,26
КОД $q = 10$ $N = 10^4$ $n = 12$	Пошаговая	0	1	4,38	4,35	6,83
			5	3,46	14,72	11,55
			10	2,96	24,85	16,48
			15	2,82	32,40	20,31
		10	1	3,25	2,7	5,13
			5	2,54	10,06	8,44
			10	2,24	17,29	12,00
			15	2,06	22,61	14,63
	Прицельная с точностью до m	-	1	4,69	1	5,31
			5	3,97	2	5,3
			10	3,54	3,25	5,63
			15	3,08	4,5	5,94
	Прицельная с точностью до 1	-	1	4,69	1	5,31
			5	4,69	1	4,54
			10	4,69	1	4,54
			15	4,69	1	4,54

Литература

1. Интеллектуализованный интерфейс пользователя информационно-поисковой системы в задаче поиска по ключевому слову («образцу») с упреждающей подсказкой / Г.Е. Кузьменко, В.А. Литвинов, С.Я. Майстренко, И.Н. Оксанич // Математичні машини і системи. - 2011. - №1. - С. 61-71.
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/kieras/GOMS/KLM.pdf>.
3. 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 // Девятая междунар. науч. конф. им. Т.А. Таран «Интеллектуальный анализ информации ИАИ-2009»: сб. труд. - Киев, 2009. - С. 212-218.
4. Оксанич И.Н. Модель декомпозиции ментальных операторов в проблемно-ориентированном интерфейсе пользователя и ее экспериментальное исследование / И.Н. Оксанич // Математичні машини і системи. - 2010. - №1. - С. 105-112.