

**ПОСТРОЕНИЕ АССОЦИАТИВНЫХ ПРАВИЛ
НА ОСНОВЕ СВЯЗИ ГЕОМЕТРИЧЕСКИХ ПАРАМЕТРОВ
ШРИФТОВ И ОБЪЕКТИВНОЙ УДОБОЧИТАЕМОСТИ**

© О. В. Токарь, к.т.н., М. А. Зильберглейт, д.х.н.,
УО «Белорусский государственный технологический
университет», Минск, Беларусь

У статті сформульовані асоціативні правила, що зв'язують геометричні параметри шрифтів, що мають високий ступінь легкості читання, визначену за методом виміру часу читання. Виявлено геометричні параметри, що впливають на хорошу читабельність шрифту.

The paper formulates the association rules, linking the geometrical parameters of fonts with a high degree of readability, a method of measuring time reading. Identified geometric parameters influencing the good readability of the font.

Постановка проблемы

Основное назначение ассоциативных правил — поиск закономерностей между связанными событиями. Одной из наиболее распространенных задач анализа данных является определение часто встречающихся сочетаний объектов в большом множестве наборов. Впервые это задача была предложена для поиска ассоциативных правил нахождения типичных шаблонов покупок, совершаемых в супермаркетах, поэтому иногда ее еще называют анализом рыночной корзины. Примером правила служит утверждение, что покупатель, приобретающий хлеб приобретет и молоко с вероятностью 72 %.

Основными характеристиками ассоциативного правила являются поддержка и достоверность правила. Поддержка является мерой статистической обоснованности результатов

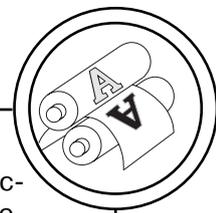
классификации. Чем выше данный показатель, тем достовернее результаты классификации [1].

Цель работы

В данной работе предпринята попытка сформировать ассоциативные правила на материале геометрических параметров шрифтов с высокой удобочитаемостью. Полученные результаты позволят сделать вывод о том, какие именно геометрические параметры влияют на хорошую удобочитаемость шрифта.

Результаты проведенных исследований

Новизна работы заключается в том, что методика построения ассоциативных правил впервые применяется для поиска зависимостей между геометрическими параметрами шрифтов с высокой степенью удобочитаемости.



Для достижения цели необходимо решить следующие задачи:

- 1) выявить шрифты с высокой удобочитаемостью;
- 2) измерить геометрические параметры этих шрифтов;
- 3) построить ассоциативные правила.

Первым этапом работы является определение удобочитаемости шрифтов по методике измерения времени чтения, которая заключается в измерении скорости (или времени) чтения испытуемого текста, оформленного данными шрифтами. Получаемые таким образом экспериментальные данные характеризуют так называемую объективную удобочитаемость исследуемой гарнитуры.

Удобочитаемость является характеристикой сравнительной, то есть ее высокая степень может быть установлена в сравнении с объектами, получившими низкие позиции по данному параметру. Поэтому изначально для оценки удобочитаемости было выбрано 35 гарнитур (с засечками и без засечек), например, Arial, BodoniCTT, BalticaCTT, BruskovayaCTT, CenturySchoolbook и др. Данные шрифты были выбраны в силу своей частоты употребления и неординарности рисунка. Пределы времени чтения для них составляют от 138 до 232 секунд. В группу с высокой удобочитаемостью входят шрифты, время чтения которых от 138 до 186 секунд: News Paper, Oliver, Free SetCTT, Calibri, NewJournal CTT, Garamond, Bruskovaya CTT, Peterburg CTT, AGPresquire, Mysl Narrow C, Benquiat GothicCTT.

Особенностью методики ассоциативных правил в классическом варианте является то, что для их построения необходимо располагать очень большим множеством наборов (например, базой покупок в супермаркете за весьма значительное количество времени). Однако применение ее для данного объекта исследования позволяет обойтись значительно меньшим количеством множеств (в их роли выступают шрифты с набором параметров), поскольку диапазон значений геометрических параметров текстовых шрифтов значительно уже ассортимента покупательской корзины в супермаркете.

Второй этап работы связан непосредственно с измерением геометрических параметров. На рис. 1 указаны элементы буквы, позволяющие рассчитать геометрические параметры шрифта.

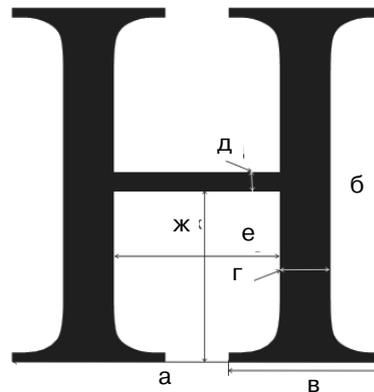
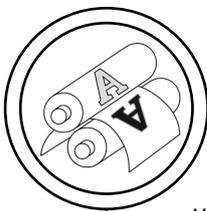


Рис. 1. Элементы буквы «н» шрифта Times New Roman

Для каждого из шрифтов для букв «н» и «а» были измерены следующие параметры [2]:

1. Пропорциональность — отношение высоты б к ширине а.



2. Контрастность — отношение ширины основного штриха $г$ к ширине дополнительного штриха $д$.

3. Отношение максимальной толщины штриха к минимальной толщине (целесообразно измерять для буквы «а»).

4. Отношение величины кегля (24,7) к высоте $б$.

5. Отношение основного штриха $г$ к внутрибуквенному просвету. Внутрибуквенный просвет строчной буквы «н» — расстояние между правым и левым вертикальными элементами $е$, измеренное на уровне середины расстояния между горизонтальным элементом и базовой линией $ж$. Внутрибуквенный просвет для строчной буквы «а» измеряется как расстояние между основным штрихом и наплывом, для чего ширина буквы, а затем и расстояние между дугами (изгибами) делятся пополам.

6. Размер засечки — отношение ширины засечки $в$ к ширине основного штриха $г$.

7. Отношение площади к периметру. Данный параметр был

определен с помощью свободно распространяемого в интернете макроса CurveInfo для программы CorealDRAW. На рис. 2 отражены результаты для буквы «а» в одном из шрифтовых оформлений.

Полученные геометрические параметры для изучаемых шрифтов были разбиты на две или три группы (в зависимости от разброса значений) и переведены в словесную форму, например, «контрастность большая», «контрастность средняя» и «контрастность маленькая» или «пропорциональность большая», «пропорциональность маленькая». Аналогично были переработаны и остальные геометрические параметры для удобства построения ассоциативных правил в программе Deductor Studio Academic 5.2. На рис. 3 представлен пример построения ассоциативных правил в данной программе.

Кроме расшифровки в ассоциативном правиле самой полученной записи геометрических параметров требуется дешиф-

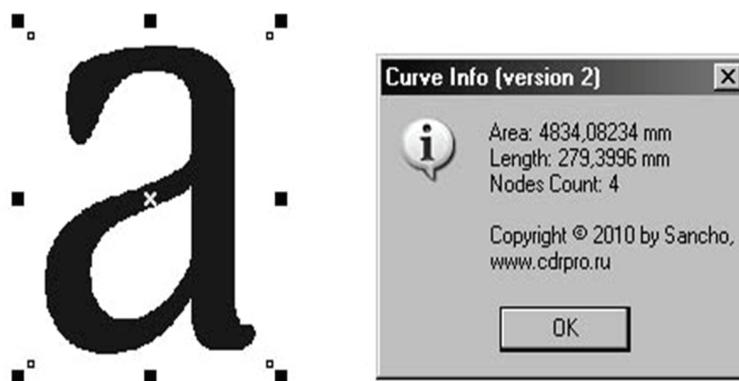
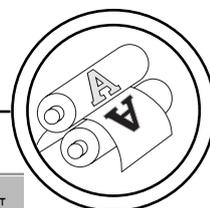


Рис. 2. Измерение площади (Area) и периметра (Length) в программе CorealDRAW



№	Номер правила	Условие	Следствие	Поддержка		Достоверность	Лифт
				Кол-во	%		
1	1	к/в а маленькое	контр а маленькое	9	81,82	90,00	1,100
2	2	к/в а маленькое	макс/мин а маленькое	9	81,82	90,00	1,100
3	3	к/в а маленькое	ш/п а маленькое	9	81,82	90,00	0,990
4	4	ш/п а маленькое	к/в а маленькое	9	81,82	90,00	0,990
5	5	ш/п а маленькое	ш/п н маленькое	9	81,82	90,00	1,100

Рис. 3. Пример записи ассоциативных правил

ровка понятий «маленькое», перевод их из словесной в «большое» и «среднее», т. е. цифровую форму (таблица).

Ассоциативные правила для группы шрифтов с высокой удобочитаемостью

Номер правила	Условие	Следствие	Поддержка, %	Достоверность, %
Наиболее достоверные правила				
1	Кегль на высоту «а» маленькое (от 1,74 до 2,08)	Контрастность «а» маленькое (от 0,84 до 2,56)	81,82	90
2	Кегль на высоту «а» маленькое (от 1,74 до 2,08)	Макс. на мин. «а» маленькое (от 1,1 до 3,73)	81,82	90
3	Кегль на высоту «а» маленькое (от 1,74 до 2,08)	Основной штрих на внутрибуквенный просвет «а» маленькое (от 0,19 до 0,64)	81,82	90
4	Основной штрих на внутрибуквенный просвет «а» маленькое (от 0,19 до 0,64)	Кегль на высоту «а» маленькое (от 1,74 до 2,08)	81,82	90
5	Основной штрих на внутрибуквенный просвет «а» маленькое (от 0,19 до 0,64)	Основной штрих на внутри-буквенный просвет «н» маленькое (от 0,19 до 0,64)	81,82	90
Правила со средней достоверностью				
6	Контрастность «н» маленькая (от 0,79 до 2,33)	Пропорциональность «а» большая (от 1,18 до 1,42)	63,64	87,5
7	Пропорциональность «н» большая (от 0,15 до 1,5)	Засечек нет	45,45	83,33
8	Пропорциональность «н» большая (от 0,15 до 1,5)	Площадь на периметр «н» маленькое (от 31,09 до 36,54)	45,45	83,33
9	Пропорциональность «н» маленькая (от 24,66 до 32,47)	Контрастность «н» средняя (от 2,32 до 3,86)	27,27	75
Правила с низкой достоверностью				
10	Основной штрих на внутрибуквенный просвет «н» большое (от 0,36 до 0,60)	Засечки «н» большие (от 0,57 до 0,88)	9,09	50
11	Контрастность «а» среднее (от 2,56 до 4,28)	Кегль на высоту «а» большое (от 2,08 до 2,42)	9,09	50
12	Макс. на мин. «а» среднее (от 3,73 до 6,37)	Кегль на высоту «а» большое (от 2,08 до 2,42)	9,09	50
13	Основной штрих на внутрибуквенный просвет «н» большое (от 0,57 до 0,88)	Площадь на периметр «н» среднее (от 36,53 до 41,98)	9,09	50
14	Основной штрих на внутрибуквенный просвет «н» большое (от 0,57 до 0,88)	Основной штрих на внутрибуквенный просвет «а» большое (от 0,64 до 1,09)	9,09	50

Примечание. Наиболее важные правила выделены в таблице полужирным начертанием.

**Выводы**

Итак, для шрифтов с высокой удобочитаемостью можно сделать следующие выводы:

1. В основном наиболее достоверные правила сформировали параметры, относящиеся к букве «а» за исключением одного параметра. Стало быть, на удобочитаемость, прежде всего, влияет буква «а» (группа округлых букв), а «н» — во втором случае, хотя она и считается базовой для прорисовки прямоугольных букв и символов в шрифте. Это не удивительно, поскольку буква «а» имеет более индивидуальный шрифтовой рисунок, а «н» более стандартна.

2. Сильная зависимость наблюдается в области маленьких значений геометрических параметров (в измеренном диапазоне), а средняя и низкая зависимость — в области средних и больших значений.

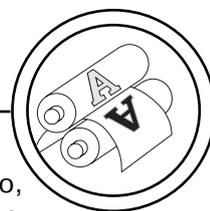
3. Наибольшую достоверность получили правила с параметрами «Кегль на высоту», «Основной штрих на внутрибуквенный просвет», «Контрастность», «Максимальная ширина штриха на минимальный» (последний параметр можно не учитывать, так как он близок к контрастности). Рассмотрим сами правила, и расшифруем входящие в них параметры:

— Правило «Если отношение кегельной площадки на высоту буквы «а» имеет небольшое значение (от 1,74 до 2,08), то контрастность «а» тоже имеет небольшое значение (от 0,84 до 2,56)». Условие подразумевает, что разница между площадкой, на которой располагается буква, и высотой буквы небольшая,

т. е. буква достаточно большая, а белого пространства вокруг нее мало. Следствие означает, что разница между основным штрихом и дополнительным тоже относительно небольшая, хотя и существует, и даже может быть обратной (дополнительный штрих шире основного). Правило подсказывает, что удобочитаемы шрифты с символами, занимающими значительную часть кегельной площадки, и при этом малоконтрастные с соотношением штрихов до 2,56.

— Правило «Если отношение кегельной площадки на высоту буквы «а» имеет небольшое значение (от 1,74 до 2,08), то основной штрих на внутрибуквенный просвет «а» маленькое (от 0,19 до 0,64)». Условие аналогично, а следствие подразумевает, что разница между основным штрихом и внутрибуквенным просветом стандартна (просвет всегда значительно больше) для почти всех шрифтов, участвовавших в эксперименте. Правило подсказывает, что удобочитаемы шрифты с символами, занимающими значительную часть кегельной площадки, и при этом имеющими значительное внутрибуквенное белое пространство.

— Правило «Если основной штрих на внутрибуквенный просвет «а» маленькое (от 0,19 до 0,64), то основной штрих на внутрибуквенный просвет «н» тоже маленькое (от 0,19 до 0,64). Правило показывает, что для удобочитаемости шрифта внутрибуквенный просвет в «а» и «н» должен на определенную величину превышать основной штрих (от 0,19 до 0,64).



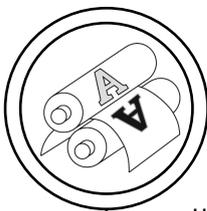
4. Правила со средней достоверностью (зависимость есть, но не так ярко выражена) в основном сформировали параметры с буквой «н», т. е. они не так важны для удобного считывания буквы «н», хотя тоже имеют значение. Это параметры «Контрастность», «Пропорциональность», «Засечки» и «Площадь на периметр». Рассмотрим сами правила, и расшифруем входящие в них параметры:

— Правило «Если контрастность «н» маленькая (от 0,79 до 2,33), то пропорциональность «а» большая (от 1,18 до 1,42)». Поскольку выше установлено, что соотношение между основным и дополнительным штрихом для «а» должно быть небольшим, то логично предположить (так как шрифт — это единое целое), что и для «н» это соотношение тоже должно быть небольшим в указанных пределах (это и является условием). Следствие выражено большой пропорциональностью, под чем подразумевается, что высота и ширина буквы должны различаться с указанного соотношения, не быть одинаковыми или близкими, ширина не должна быть больше высоты. Правило показывает, что для удобочитаемости шрифта в принципе значимо, что разница между шириной основного и дополнительного штриха для «н» небольшая, а высота символа должна быть больше его ширины.

— Правило «Если пропорциональность «н» большая (от 1,17 до 1,5), то шрифт без засечек». Условие дает понять, что высота и ширина «н» должны различаться с указанного соотношения,

при этом высота, конечно, должна быть больше ширины. Следствие означает, что эта зависимость важна для шрифта без засечек. Правило показывает, что для удобочитаемости шрифта без засечек значимо, чтобы именно высота символа «н» была больше ширины на указанное соотношение.

— Правило «Если пропорциональность «н» большая (от 1,15 до 1,5), то площадь на периметр «н» маленькое» (от 31,09 до 36,54). Условие аналогично предыдущему, а следствие говорит о том, что разница между площадью и периметром наименьшая из возможных (в рамках рассмотренных шрифтов). Вспомним, что площадь — часть поверхности, ограниченная замкнутым контуром, то есть размер самого символа, а периметр — это полная длина замкнутой кривой линии, описывающей внешний контур символа. Можно предположить, что чем эти величины больше, тем более сложен рисунок шрифта, хотя на размеры может влиять и высота с шириной символа. Наименьшая разница между площадью и периметром получилось в данном случае при не очень большой площади и наименьшей величине периметра. Такой параметр в основном соответствует шрифтам с простым рисунком, без засечек и декоративностей. Правило показывает, что для удобочитаемости шрифта значимо, чтобы высота символа «н» была больше ширины на указанное соотношение, если шрифт с простым рисунком, вероятнее всего без засечек.



— Правило «Если пропорциональность «н» маленькая (от 0,81 до 1,16), то контрастность «н» средняя» (от 2,34 до 3,86) тоже весьма интересно для анализа. Условие говорит, что в удобочитаемом шрифте может быть маленькая пропорциональность, т. е. высота и ширина могут быть и одинаковые и даже ширина больше высоты. Следствие добавляет, что при этом у шрифта ширина основного штриха может быть намного больше дополнительного (в измеренных пределах). Правило показывает, что шрифт может получить высокую оценку по удобочитаемости, если высота

и ширина символов почти одинаковые и шрифт достаточно контрастен.

5. Правила с низкой достоверностью уже менее значимы, стоит лишь отметить, что их составляют в разных вариациях уже упоминавшиеся параметры «Основной штрих на внутрибуквенный просвет», «Контрастность», «Площадь на периметр», «Кегль на высоту», «Засечки» как для «а», так и для «н».

Полученные результаты подтверждают, что высокой удобочитаемости шрифта действительно способствуют определенные геометрические параметры.

1. Зильберглейт М. А. Методика и техника подготовки курсовых и дипломных работ / М. А. Зильберглейт, Л. И. Петрова. — Минск : Бел. наука, 2003. 2. Токарь О. В. Удобочитаемость современных текстовых шрифтов : научное издание / О. В. Токарь. — Минск : ООО «Современная школа», 2007. 3. Райхман Э. П. Экспертные методы в оценке качества товаров / Э. П. Райхман, Г. Г. Азгальдов. — М., 1974.

1. Zil'berglejt M. A. Metodika i tehnika podgotovki kursovyh i diplomnyh rabot / M. A. Zil'berglejt, L. I. Petrova. — Minsk : Bel. navuka, 2003. 2. Tokar' O. V. Udobochitaemost' sovremennyh tekstovyh shriftov : nauchnoe izdanie / O. V. Tokar'. — Minsk : ООО «Sovremennaja shkola», 2007. 3. Rajhman Je. P. Jekspertnye metody v ocenke kachestva tovarov / Je. P. Rajhman, G. G. Azgal'dov. — M., 1974.

Рецензент — О. В. Зоренко, к.т.н.,
доцент, НТУУ «КПІ»

Надійшла до редакції 26.03.13