

**Меняев М.Ф., Овчинникова Н.Е.**

### **Определение авторского знака с помощью интеллектуальной системы**

**Описание программы определения авторского знака библиографической записи – подсистемы библиотечной интеллектуальной системы БАРС, разработанной в Московском государственном техническом университете им. Н.Э. Баумана.**

Несмотря на широкое использование компьютерной технологии в библиотеках, обработка новых поступлений остается трудоемкой. Поэтому совершенствование автоматизированных систем, берущих на себя некоторые функции библиотекаря при обработке новых поступлений, становится важнейшей задачей организации библиотечных процессов на современном этапе.

Для учета и распределения новых поступлений используется множество сведений об издании. К ним относятся заглавие, автор, год издания, авторский знак и др. Набор таких сведений составляет основу различных форматов описания библиографических данных, составляющих семейство MARC-форматов.

Остановимся на авторском знаке издания, который используется при различных системах расстановки библиотечного фонда и расстановке карточек в ряде каталогов.

Существенное облегчение и ускорение процесса определения авторского знака издания достигается применением таблиц авторских знаков, составленных Л.Б. Хавкиной. Авторские знаки, из которых состоит таблица, позволяют расставить книги в точном алфавитном порядке, почти не задумываясь над тем, какая из них должна быть по алфавиту раньше, а какая позже, руководствуясь только буквенно-числовой последовательностью. Это обусловлено тем, что авторский знак входит в состав шифра хранения докумен-

та. Напомним, что авторский знак — это условное обозначение первого слова, с которого начинается первый элемент библиографического описания. Он состоит из начальной буквы этого слова и числа, соответствующего первым, а иногда и последующим буквосочетаниям этого же слова. В некоторых случаях область действия авторского знака распространяется не на одно слово, а на два и более.

При определении авторского знака нередко возникают ошибки, так как этот процесс достаточно трудоемкий и утомительный, а его успех зависит от квалификации и опыта библиотекаря. Неверное определение авторского знака издания значительно усложняет работу библиотеки.

Применение компьютерной технологии решило много задач по автоматизации обработки информации, но оставило без внимания проблему определения авторского знака. Настало время еще раз обратиться к ней.

Разработана программа, использующая новые подходы к построению автоматизированных библиотечных систем. Она позволяет определять авторский знак издания по ходу ввода информации об издании без использования брошюр с таблицами. Такая программа самостоятельно определяет авторский знак для каждого введенного слова и предоставляет библиотекарю возможность согласиться с выбранным решением или указать другой, верный с его точки зрения авторский знак. Программа, осуществляющая процесс определения авторского знака, является системой искусственного интеллекта.

Применение элементов искусственного интеллекта дает ряд важных преимуществ. Одно из его положительных качеств — постоянство. Человеческая компетентность со временем ослабевает. Человек должен постоянно практиковаться, чтобы сохранить свой профессиональный уровень. Любой значительный перерыв в деятельности человека-эксперта может серьезно отразиться на его профессиональных качествах. Компетентность искусственного интеллекта, однажды созданная, сохраняется навсегда (без учета аварий аппаратуры). Другое преимущество такой компетентности — легкость, с которой ее можно передавать или воспроизводить. Передача знаний от одного человека к другому — сложный, дорогой и трудоемкий процесс. У искусственного интеллекта более устойчивые и воспроизводимые результаты. Эксперт-человек может принимать различные решения в тождественных ситуациях, что никогда не случится с искусственным интеллектом.

Программа определения авторского знака представляет собой экспертную систему с соответствующей базой знаний. Экспертная система — это некоторый набор программ, решающий задачи в интересующей нас предметной области. Этот набор программ называется системой, потому что содержит несколько компонентов, в том числе базу знаний и базу данных. Дадим некоторые пояснения.

Знания — информация, необходимая для достижения определенного результата. Эта информация имеет форму правил. Выделенные знания о предметной области называются базой знаний. База знаний экспертной системы содержит данные и правила, использующие эти данные как основу для принятия решений. База знаний накапливается в процессе построения экспертной системы. Знания в системе выражены в явном виде и организованы так, чтобы упростить принятия решений.

Рассмотрим базу правил, используемую экспертной системой для определения авторского знака. Она разбита на две группы: первая — главная таблица базы правил; вторая — таблицы правил для конкретных случаев.

Главная таблица базы правил представляет собой файл, имя которого BASE RUL, а в каждой его строке определяется правило выбора файла из второй группы.

Суть каждого правила в следующем: первым символом в каждой строке является ключ — первая буква в фамилии автора книги. Вторым символом должен стоять символ ":" — как разделитель. Далее следует имя файла базы правил второй группы.

Каждый файл базы правил второй группы имеет схожую структуру: первая строка файла используется для проверки соответствия значения ключа данному файлу правил.

Начиная со второй строки в базе записываются правила для нахождения номеров записей в файлах базы данных.

В разработанной экспертной системе по определению авторского знака в качестве базы данных используются таблицы Л.Б. Хавкиной, представленные в электронной форме. Применяются две базы данных: для русского и английского языков. При вводе фамилии автора система автоматически определяет используемый язык и открывает соответствующую базу данных. В случае, если таковая не найдена, система выдаст предупреждение об ошибке; и до тех пор, пока не будет установлена необходимая база данных, дальнейшая работа по определению авторского знака невозможна.

Формат базы данных следующий: определенным буквосочетаниям различной длины соответствует некоторый набор цифр. Пер-

вая буква и набор цифр образуют авторский знак. База данных создана в .dbf – формате.

Фрагмент базы данных для английского языка:

G	10	H
Ga	11	Ha
Gad	12	Haс
Gae	13	Haе
Gai	14	Haі
Gal	15	Haг
Gam	16	Haһ
Gan	17	Hał
Gap	18	Hall

Для удобства работы в одном столбце двум буквосочетаниям соответствует один и тот же авторский знак.

Определение авторского знака происходит при заполнении соответствующего поля записи "Фамилия" или, если первым полем является заглавие, при его заполнении. При вводе фамилии автора или его псевдонима автоматически заполняется и поле "Авторский знак". В это поле помещается определенный системой знак. Например, при вводе букв "Вено" в поле "Авторский знак" будет выведена запись "В – 294", что соответствует авторскому знаку заданной фамилии. В программе предусмотрена защита поля "Авторский знак"; он не доступен библиотекарю, т. е. нельзя изменить его содержание при помощи клавиатуры или "мыши".

Программа позволяет библиотекарю работать непосредственно с таблицами. Можно просмотреть не только значения кода, полученного системой, но и соседние с ним по таблице знаки. С помощью клавиш или "мыши" можно перемещать по экрану курсор (синюю полосу в таблице), автоматически устанавливая на код, выбранный библиотекарем.

При установке авторского знака на экране показаны не менее трех записей из таблицы: авторский знак введенной фамилии, знаки, расположенные до и после определенного. Например, при введенной фамилии "Агрос" таблица имеет вид:

Агрон	263
Агрос	264
Агрот	265

При желании можно увеличить размер таблицы, используя соответствующую клавишу управления системой. При этом на экран будут выведены еще две записи, расположенные в таблице до и после уже выделенных. Тогда приведенная выше таблица приобретает вид:

Агрономо	262
Агроп	263
Агрос	264
Агрот	265
Агроч	266

Таким образом можно расширить таблицу до любого желаемого размера. Дополнительные записи на экране позволяют не только просматривать близлежащие знаки, но и корректировать установленный системой авторский знак. Это весьма удобно в том случае, когда библиотекарь не согласен с решением системы. С помощью курсора или "мыши" библиотекарь выбирает верный с его точки зрения авторский знак и устанавливает его в поле записи. После этого поле "Авторский знак" будет содержать код, выбранный пользователем.

Для корректного определения авторского знака необходимо соблюдение нескольких условий: вводимая фамилия автора должна начинаться с русской или латинской заглавной буквы, первым символом должна быть буква. Если эти правила нарушены, то система выдаст предупреждение об ошибке, авторский знак определен не будет, соответствующее поле и фрагмент таблицы на экране останутся пустыми. В этих случаях авторский знак необходимо установить вручную.

Процесс установки авторского знака происходит при побуквенном вводе информации, т. е. сначала авторский знак определяется для первой буквы, затем для первых двух и т. д. Например, при вводе слова "код" на экране можно увидеть следующую последовательность знаков:

"К"	К-11
"Ко"	К-55
"Код"	К-575

Не всегда авторский знак можно определить однозначно. Это связано с тем, что число всевозможных буквосочетаний очень велико и авторские таблицы охватывают лишь их малую часть. По-

этому может возникнуть конфликтная ситуация. Рассмотрим одну из них. Ниже приведен фрагмент таблиц Л.Б. Хавкиной для русского языка:

Вен	29
Вене	291
Вени	292
Венк	293
Вено	294
Венс	295

В том случае, если вводимая фамилия начинается с "Вена" или "Венр" авторский знак нельзя определить однозначно. Тогда система определит ближайшее к заданному буквосочетанию и выдаст соответствующий знак. Например:

Вена – В – 291  
Венр – В – 295

Определение авторского знака по заглавию происходит только в том случае, когда фамилия автора отсутствует. Каждый раз при вводе информации в поле "Заглавие" происходит проверка: была ли введена ранее в соответствующее поле фамилия автора? Если фамилия отсутствует, происходит определение авторского знака для заглавия издания. При этом выполняются все те же действия, что и при определении авторского знака для фамилии.

Программа определения авторского знака фамилии автора или заглавия издания является подсистемой библиотечной интеллектуальной системы БАРС, разработанной в МГТУ им. Н.Э. Баумана.