Язык запросов ИРБИС

*Поисковый язык ИРБИС* или *язык запросов ИРБИС* предназначен для формулировки запросов, с помощью которых выполняется поиск в базах данных ИРБИС.

В ИРБИС выделяются два вида поиска:

* *Прямой поиск* – быстрый поиск на основе индекса (словаря) базы данных.
* *Последовательный поиск* – медленный поиск на основе последовательного перебора (просмотра) записей базы данных с возможностью выполнения таких операций, как БОЛЬШЕ/МЕНЬШЕ, НАЛИЧИЕ/ОТСУТСТВИЕ и других.

Прямой поиск

Запрос для прямого поиска представляет собой алгебраическое (поисковое) выражение, в котором операндами являются термины словаря, а операторами – логические операторы булевой алгебры. Для изменения порядка выполнения логических операторов в поисковом выражении могут применяться скобки.

Термин словаря включает в себя собственно термин словаря и префикс, если таковой используется для данного вида терминов.

В общем виде операнд поискового выражения можно представить следующим образом:

"<префикс><термин>$"/(tag1,tag2,...tagN)

где:

* <префикс> - префикс, определяющий вид термина (вид словаря).
* <термин> - собственно термин словаря.
* $ - признак правого усечения термина. Определяет совокупность терминов, имеющих начальную последовательность символов, совпадающую с указанным термином. Может отсутствовать, в этом случае поиск идет по точному значению указанного термина.
* " – символ-ограничитель термина (двойные кавычки). Должен использоваться обязательно, если термин включает в себя символы пробел, круглые скобки, решетка (#), а также символы, совпадающие с обозначениями логических операторов (см. ниже).
* /(tag1,tag2,...tagN) – конструкция квалификации термина. Определяет метки поля, в которых должен находиться указанный термин, или точнее – вторую часть индексной ссылки. Может отсутствовать, что означает отсутствие дополнительных требований в части меток полей.

В поисковом выражении могут использоваться следующие логические операторы:

* + - оператор логического ИЛИ. Соединение двух операндов (терминов) логическим оператором ИЛИ обозначает требование поиска записей, в которых присутствует хотя бы один из терминов.
* \* - оператор логического И. Соединение двух терминов логическим оператором И обозначает требование поиска записей, в которых присутствуют оба термина.
* ^ - оператор логического НЕ. Соединение двух терминов логическим оператором НЕ обозначает требование поиска записей, в которых присутствует первый термин и отсутствует второй. Оператор НЕ не может быть одноместным (т.е. данному оператору, как и всем другим, должен ОБЯЗАТЕЛЬНО предшествовать термин).
* (G) – первый оператор контекстного И. Соединение двух терминов таким оператором контекстного И обозначает требование поиска записей, в которых оба термина присутствуют в одном и том же поле (или точнее – когда у терминов совпадают вторые части ссылок).
* (F) – второй оператор контекстного И. Соединение двух терминов таким оператором контекстного И обозначает требование поиска записей, в которых оба термина присутствуют в одном и том же повторении поля (или точнее – когда у терминов совпадают вторые и третьи части ссылок).
* . – (точка обрамленная пробелами) третий оператор контекстного И. Соединение двух терминов таким оператором контекстного И обозначает требование поиска записей, в которых оба термина присутствуют в одном и том же повторении поля друг за другом (или точнее – когда у терминов совпадают вторые и третьи части ссылок, а третьи части ссылок отличаются на единицу).

Логические операторы имеют приоритеты, которые определяют порядок их выполнения (в пределах одного уровня скобок). Ниже операторы приведены в порядке убывания приоритета:

.

(F)

(G)

\* и ^

+

Операторы одного приоритета выполняются слева направо (в пределах одного уровня скобок). Для изменения порядка выполнения логических операторов в поисковом выражении могут применяться круглые скобки. Выражения в скобках могут объединяться только операторами + \* ^.

Примеры запросов для прямого поиска:

("A=Иванов$" +"A=Петров$") \* ("V=03" + "V=05")

"K=трактор$" (F) "K=колесн$" + "K=бульдозер$" (F) "K=гусен$"

"K=очист$"/(200,922) \* "K=вод$"/(200,922)

Последовательный поиск

Запрос (или поисковое выражение) для последовательного поиска представляет собой формат на языке форматирования ИРБИС.

Собственно процесс последовательного поиска состоит в последовательном расформатировании исходных записей (т.е. тех записей, которые участвуют в поиске): если результатом расформатирования записи является строка '1' (или точнее – строка, содержащая символ '1'), то соответствующая запись удовлетворяет поисковому запросу, в противном случае – не удовлетворяет.

Таким образом, в обобщенном виде запрос для последовательного поиска можно представить как следующий формат:

if <логическое выражение> then '1' else '0' fi

где <логическое выражение> - логическое выражение, в котором можно использовать все возможности языка форматирования ИРБИС.

*Примечания:*

* В АРМ Читатель и АРМ Каталогизатор*в режиме ПОСЛЕДОВАТЕЛЬНЫЙ ПОИСК – СВОБОДНЫЙ ПОИСК в качестве запроса необходимо задавать именно <логическое выражение> (без объемлющей конструкцииif … fi).*
* *Последовательный поиск применяется только в тех случаях, когда нет возможности найти необходимые записи на основе прямого поиска (т.е. на основе словарей). Если некий последовательный поиск приходится применять регулярно, следует пересмотреть правила инверсии базы данных (<dbname>.fst) – с тем чтобы иметь возможность выполнять соответствующий поиск на основе словаря.*