

Ц 840Г  
Г-577

СООБЩЕНИЯ  
ОБЪЕДИНЕННОГО  
ИНСТИТУТА  
ЯДЕРНЫХ  
ИССЛЕДОВАНИЙ

ДУБНА



6/II-78

711/2-78

10 - 11052

Н.Н.Говорун, Ю.П.Залаторюс, В.И.Никитина,  
В.М.Сумароков, Г.Н.Тентюкова

ЯЗЫК ВХОДНЫХ СООБЩЕНИЙ И ЯЗЫК ЗАПРОСА  
ИНФОРМАЦИОННОЙ СИСТЕМЫ "КАДРЫ"

**1977**

10 - 11052

Н.Н.Говорун, Ю.П.Залаторюс, В.И.Никитина,  
В.М.Сумароков, Г.Н.Тентюкова

ЯЗЫК ВХОДНЫХ СООБЩЕНИЙ И ЯЗЫК ЗАПРОСА  
ИНФОРМАЦИОННОЙ СИСТЕМЫ "КАДРЫ"

Институт  
ядерной физики  
БИБЛИОТЕКА

Говоруя Н.Н. и др.

10 - 11052

Язык входных сообщений и язык запроса информационной системы "КАДРЫ"

Дано описание языка входных сообщений и языка запросов фактографической информационной системы для кадрового учета. Первый предназначен для ввода и редактирования информации, образующей базу данных. Вторым используется для задания предписаний на поиск и выдачу информации пользователем-непрограммистом.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1977

Govorun N.N. et al.

10 - 11052

Input Communication Language and Request Language for Information System "Personnel"

Input communication language and that for requests of factographic information system for personnel registration are described. The first language is intended for the input and editing the data which forms the data basis, the other is used for instructing a user-non-programmer to search and output the data.

The investigation has been performed at the Laboratory of Automation, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1977

Входной язык (язык входных сообщений) и язык запроса предназначены для удобного взаимодействия пользователя с системой.

Входной язык используется для ввода и редактирования информации, образующей базу данных.

Язык запроса служит для формулирования пользователем задания на поиск информации и выдачу справок.

Оба вида языка требуют от абонента системы знания структуры данных /1,2/, т.е. представления пользователя о данных.

В некоторых системах оба языка используют одни и те же синтаксические конструкции и различаются только разделом процедур (действий), поскольку входной язык указывает процедуру формирования или обновления информации, а язык запроса требует поиска и выдачи информации.

В первом варианте системы "КАДРЫ" входной язык и язык запроса различаются по синтаксису. Это объясняется, главным образом, тем, что на ранних этапах работы системы, особенно при формировании начального экземпляра базы данных, пользователь, который вводит данные, оперирует с гораздо большим объемом информации и должен быть лучше знаком со структурой хранения данных, чем пользователь, запрашивающий справку из системы.

В работе описаны только форматы языков. Процессоры запроса и ввода данных, а также программы формирования массивов и принципы работы поисковой системы здесь не обсуждаются.

## ЯЗЫК ВХОДНЫХ СООБЩЕНИЙ

Исходным документом, источником сведений об объекте является анкета.

Количество признаков, характеризующих объект системы и значения которых заносятся в анкеты, может быть значительным.

Хотя набор признаков и их последовательность во входном документе фиксированы, из-за их большого количества нецелесообразно применять язык разделителей /1/. По этой же причине неудобен и повествовательный язык. Поэтому в качестве языка входных сообщений (описания документов) применены языки двух типов: язык фиксированных позиций и ключевых слов. Выбор языка в каждом конкретном случае определяется видом задачи (ввод или редактирование данных) и режимом работы (оперативный или пакетный).

Язык фиксированных позиций характеризуется жестким форматом ввода. В этом случае каждый признак анкеты занимает фиксированную позицию входного носителя (т.е. колонку перфокарты или позицию строки экрана дисплея). На самом деле фиксированные позиции занимают только простые признаки /2/ и подпризнаки внутри каждого сложного. Порядок задания самих сложных признаков не фиксирован. Формат ввода задается в системных таблицах.

Язык ключевых слов имеет форму последовательности пар "признак-значение". Признак задается его ключом, роль которого выполняет специальный номер, а значение - строкой символов. Пары отделяются друг от друга разделителем, и весь фрагмент ввода (предложение) заканчивается признаком конца. Порядок задания пар и их положение на входном носителе внутри записи произвольны. Используя метаязык Бэкуса, можно дать следующее описание фрагмента ввода исходной информации, который включает данные для ввода или редактирования одного объекта системы:

< фрагмент ввода > ::= < режим работы > < имя объекта > < последовательность пар > < КЗ >

< режим работы > ::= редактирование | изменение имени объекта | аннулирование объекта | < пусто >

< пусто > ::=

Если режим работы не задан (< пусто >), то производится первоначальный ввод анкеты.

< имя объекта > ::= фамилия имя отчество имя - подразделения | < пусто >

< последовательность пар > ::= < пара > | < последовательность пар > , < пара >

< пара > ::= < ключ > = < значение >

< ключ > ::= < номер > | < номер > . < номер >

< номер > ::= < цифра > | < номер > < цифра >

< цифра > ::= 0|1|2|3|4|5|6|7|8|9

< значение > ::= < последовательность любых символов, которые воспринимаются программой ввода системы "Дубна" /3/, кроме символов , : >

< КЗ > ::= жконец

Кроме описанных входных языков в оперативном режиме для редактирования данных может быть использован язык мультитайпного редактирования /4/.

## ЯЗЫК ЗАПРОСА

Для того чтобы пользователь мог указать системе "КАДРЫ", какую информацию и в какой форме он желает получить, разработан специальный язык запроса. Он имеет повествовательную форму, но с жесткими синтаксическими ограничениями. Язык запроса рассчитан на пользователя - непрограммиста, от которого требуется знание только структуры данных и словаря, используемого системой.

Прежде чем перейти к описанию языка, введем некоторую терминологию.

На языке запроса пользователь задает поисковое предписание и предписание на форму и содержание выдачи. Поисковое предписание - это выраженное в терминах информационнопоискового языка смысловое содержание информационного запроса.

Словарный состав языка определяется двумя наборами слов или словосочетаний.

Первый набор - это имена значений признаков. Они же являются ключами, которые функционально определяют разбиение файла анкет на подфайлы, образуя списковую структуру /2/. Ключи входят в Справочник ключей, который имеет в системе имя VNRES, а имена значений признаков (ключи) обозначаются МКβ (мнемокод β - части тезауруса).

В Справочник ключей входят не все значения признаков, а только те, которые выбраны для выполнения роли ключей при создании списковой структуры. Включение значения признака в Справочник

ключей зависит от частоты его участия в запросах. Количество ключей определяется наличием свободной памяти на диске для хранения "линеек" /2/. В процессе эксплуатации системы Справочник ключей подвергается модификации, для чего созданы специальные программы /5/.

Второй набор слов словарного состава языка - это имена признаков. Все признаки (и подпризнаки), характеризующие объект системы, имеют имена, которые заносятся в таблицу ATNES. Имена признаков (подпризнаков) обозначаются МК  $\alpha$  (мнемокод  $\alpha$ -части тезауруса).

Кроме двух описанных наборов слов в словарь системы включено слово ФИО. Оно является идентификатором имени объекта системы, т.е. фамилии, имени, отчества сотрудника.

При формулировании запроса пользователь в качестве допустимых слов употребляет или слово из первой части словаря (МК  $\beta$ ), или слово из второй части, т.е. имя признака (МК  $\alpha$ ), и тогда он указывает его значение, или ФИО и его значение.

Использование в поисковом предписании МК  $\beta$  (вместо альтернативного МК  $\alpha$  = <значение>) повышает эффективность поиска, поскольку МК  $\beta$  являются ключами и для них созданы "линейки" /2/.

Некоторое усложнение в язык, а также в разработку Процессора запросов и процедур формирования выдачи вносит наличие сложных списковых признаков /2/. Сложный списковый признак задается набором подпризнаков, значения которых образуют строку, причем каждый такой признак может иметь несколько строк (т.е. список строк) в соответствии с несколькими наборами значений подпризнаков. Если в поисковое предписание входят разные подпризнаки одного сложного спискового признака и их значения указаны для одной и той же строки признака, то такая сложная конструкция имеет в языке специальную структуру (СФСЗ - скобочная форма списковой записи).

Поисковое предписание задается в виде булевой функции, представленной в дизъюнктивной нормальной форме (ДНФ):

$$ПП = \bigvee_{i=1}^L T_i$$

где

$$T_i = \bigwedge_{j=1}^{M_i} t_{ij} \quad - \text{терм,}$$

- L - количество термов поискового предписания,
- $t_{ij}$  - структурная единица поискового предписания,
- $M_i$  - количество структурных единиц в конъюнктивной

группе.

Далее дается описание синтаксиса языка.

< ЗАПРОС > ::= НАЙТИ : < ПП > : ВЫДАТЬ : < ПВ > : |  
 НАЙТИ : < ПП > : ВЫДАТЬ АЦПУ : < ПВ > : |  
 ВЫДАТЬ : < ПВ > : | ВЫДАТЬ АЦПУ : < ПВ > :

< ПП > ::= поисковое предписание.

< ПП > ::= < ДНФ > :

< ДНФ > ::= < терм > | < ДНФ > + < терм >

< терм > ::= < структурная единица ПП > | < терм > , < структурная единица ПП >

+ - знак логического сложения,

, - знак логического умножения.

< структурная единица ПП > ::= < Э > | N < Э > | < Ф > | < СФСЗ > | N < СФСЗ >

N - знак логической инверсии.

< Э > ::= < МК  $\beta$  > | < МК  $\alpha$  > = < набор значений > | < сл.сл. > \* < МК  $\alpha$  > = < набор значений >

< МК  $\beta$  > ::= мнемокод ВТНЕС (имя значения признака)

< МК  $\alpha$  > ::= мнемокод АТНЕС (имя признака)

< сл.сл. > ::= ГОД | ВОЗРАСТ | СТАЖ

< набор значений > ::= < интервал > | < набор значений > S < интервал >

< интервал > ::= < число > | < интервал > + < число >

< число > ::= < дата > | < целое >

< целое > ::= < цифра > | < целое > < цифра >

< дата > ::= < две цифры > . < две цифры > . < две цифры >

< две цифры > ::= < цифра > < цифра >

< цифра > ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

< Ф > ::= ФИО = < текст ФИО >

< текст ФИО > ::= < Ф > < И > < О >

< Ф > ::= < фамилия > | < начальный фрагмент фамилии >

< И > ::= < имя > | < начальный фрагмент имени > | < пусто >

< О > ::= < отчество > | < начальный фрагмент отчества > | < пусто >

Фамилия, имя, отчество и их фрагменты задаются строками русских символов.

<Э> и <Ф> являются элементарными единицами структуры ПП.  
 <СФСЗ> ::= (<последовательность элементов>)  
 <последовательность элементов> ::= <Э> | N <Э> | <последова-  
 тельность элементов> , <Э> | <последовательность элементов> ,  
 N <Э>

Все <Э> , входящие в <СФСЗ> , должны относиться к подпри-  
 знакам одного и того же признака.

<ПВ> ::= предписание для выдачи.

<ПВ> ::= <оператор выдачи> :

<оператор выдачи> ::= ФИО | ФИН | S { ВСЕ | <перечень призна-  
 ков> | ФИО, <перечень признаков> | ФИН, <перечень признаков> /  
 ФИОАН, <перечень признаков>

<перечень признаков> ::= <МК Δ> | <перечень признаков> , <МК Δ>

Предписание на выдачу задает форму и содержание выдачи информации  
 для документов, отобранных согласно заданию поискового предписа-  
 ния. Выдача производится на экран дисплея или АЦПУ ЭВМ в зависи-  
 мости от задания

ВЫДАТЬ: или ВЫДАТЬ АЦПУ:

Допустимые типы выдачи.

- 1) S . Выдается S=n , где n - количество отобранных докумен-  
 тов.
- 2) ФИО или ФИН. Выдается S=n и затем в алфавитном порядке спи-  
 сок сотрудников. Каждая строка содержит фамилию, имя и отчест-  
 во при задании ФИО или фамилию и инициалы при задании ФИН.
- 3) ФИО, <перечень признаков>  
 ФИН, <перечень признаков>  
 ФИОАН, <перечень признаков>  
 Выдается упорядоченная по ключу ФИО или ФИН таблица, каждая стро-  
 ка которой содержит в первом столбце фамилию, имя, отчество или  
 фамилию и инициалы сотрудника (а для ФИОАН еще и номер анкеты),  
 а в следующих столбцах - значения признаков, заданных в перечне.  
 Таблица снабжается заголовком.
- 4) ВСЕ. Выдается вся информация, известная системе относительно  
 найденных объектов.
- 5) <перечень признаков> . Выдаются фамилия, имя, отчество и  
 значения признаков для отобранных объектов. Этот тип выдачи  
 отличается от типа ФИО, <перечень признаков> только формой,  
 но не содержанием выдачи.

### Примечание

Пробелы игнорируются всюду в поле "запроса", кроме < текст ФИО>,  
 поскольку здесь пробел является разделителем.

Примеры запросов:

- 1) НАЙТИ: ФИО = ИВАНОВ НИКОЛ. СЕМ.:  
 ВЫДАТЬ АЦПУ: ВСЕ:
- 2) НАЙТИ: ГОД\*Д-ОКОНЧ-УЧ-ЗАВ = 68,ЛВТА,ИНЖЕНЕР:  
 ВЫДАТЬ: ФИН, Д-РОЖДЕНИЯ, ПОЛ:
- 3) НАЙТИ: (АНГЛ-ЯЗ, СВ-ВЛАД) + (ИСП-ЯЗ, СО-СЛОВАРЕМ):  
 ВЫДАТЬ: ФИО, ПОДРАЗДЕЛЕНИЕ:
- 4) НАЙТИ: ВОЗРАСТ\*Д-РОЖДЕНИЯ = 16#18S55#60:  
 ВЫДАТЬ: ФИО, ТИП-УЧ-ЗАВ:
- 5) НАЙТИ: ЛВТА, ВЕСТЬ-НАГРАДЫ:  
 ВЫДАТЬ: ФИН, ПОДРАЗДЕЛЕНИЕ, ДОЛЖНОСТЬ-ПРОФ, ТИП-НАГРАДЫ,  
 НАЗВ-НАГРАДЫ:

### Литература

1. Информационные системы общего назначения (аналитический обзор  
 систем управления базами данных). М., "Статистика", 1975.
2. Говорун Н.Н. и др. Организация информационных массивов системы  
 "КАДРЫ". ОИЯИ, IO-II05I, 1977.
3. Мазный Г.Л. Мониторная система "Дубна". ОИЯИ, II-5974, 1971.
4. Веретенев Ю.В. и др. Мультидоступная система "МУЛЬТИТАИП" на  
 БЭСМ-6. ИАЭ-2409, М., 1974.
5. Говорун Н.Н. и др. Программное обеспечение системы "КАДРЫ",  
 ОИЯИ, IO - IO95I , 1977.

Рукопись поступила в издательский отдел  
 I ноября 1977 года.