

К ПРОБЛЕМЕ ПОСТРОЕНИЯ МАШИННОГО ЯЗЫКА ДЛЯ ИНФОРМАЦИОННОЙ МАШИНЫ *)

В. А. УСПЕНСКИЙ

(МОСКВА)

§ 1

Ученый, приступающий к какому-либо исследованию, сталкивается с информацией двух типов. Информация первого типа заключена непосредственно в том фрагменте объективной действительности, который составляет предмет исследования (извлечь эту информацию—и есть задача науки). Информация второго типа—это информация, уже извлеченная из реальности предшественниками нашего ученого и заключенная (закодированная) в статьях, книгах, таблицах, карточках и т. п. Извлечь нужную информацию второго типа (только эта информация и будет рассматриваться в дальнейшем), хотя и значительно легче в подавляющем большинстве случаев (но не во всех случаях), чем нужную информацию первого типа, все же обычно довольно трудно. К тому же эта задача делается все труднее с ростом объема и сложности накопившихся знаний.

В настоящее время поиск нужной информации осуществляется «вручную», кустарными средствами. Такое положение можно было бы охарактеризовать как средневековое. Хотя со времени средних веков появилась масса указателей, каталогов и т. д., но объем сведений возрос настолько, что сейчас найти нужную справку в литературе подчас труднее, чем в средние века, когда каждый ученый мог сам хранить в голове все сведения по своей отрасли науки. Любая классификационная система принципиально ограничена и недостаточна хотя бы потому, что она рассчитана на состояние науки в момент составления системы, а пользуются системой через много лет. Все это приводит к большой затрате человеческого труда и неизбежной потере информации. Хорошо известно, что два человека, задавшись одним и тем же вопросом и обратившись с этой целью к справочникам, предметным и систематическим каталогам, библиографическим бюро и т. д., могут получить различные результаты.

Актуален поэтому вопрос об автоматизации процесса поиска нужной информации с целью быстрого и полного ее получения. Имеется в виду создание для этой цели специальной машины. Эта машина должна, таким образом, хранить в своей «памяти» значительное количество сведений и, используя их, отвечать на задаваемые ей вопросы. Предполагается, что вопросы будут такого же характера, какие ставит перед собой иссле-

*) Статья представляет собой, с незначительными изменениями, доклад, сделанный 28 мая 1957 г. на созванном Лабораторией электро моделирования АН СССР Научно-техническом совещании по комплексу вопросов, связанных с разработкой и построением информационных машин с большой долговременной памятью. [Примечание. Сейчас, через два года, некоторые положения статьи кажутся автору чересчур банальными. Автор просит читателя сделать скидку на время, отделяющее написание статьи от ее публикации.]

дователь, прежде чем обратиться к литературе, например, «какие есть книги по данной теме», или «что известно по данной теме», или «какие вещества обладают данными свойствами» и т. д. Предполагается, далее, что машина сможет совершать элементарные логические выводы, подобные тем, которые производит читатель в самом процессе чтения книги (непосредственное сопоставление фактов и т. п.). Было бы желательно, наконец, чтобы машина могла определять тождественность содержания двух различных текстов. Это немедленно нашло свое применение в патентном деле и, самое главное, это было бы чрезвычайно удобно при записи в машину поступающих сведений: машина могла бы сличать поступающие сведения с уже записанными и записывать только действительно новые факты.

Развитие современной техники позволяет сделать реальным построение подобной машины (имеются в виду, в частности, разработанные под руководством Л. И. Гутенмахера в Лаборатории электро моделирования технические средства, дающие возможность хранить большой объем информации в малом объеме пространства и извлекать эту информацию с большой скоростью). Содержание настоящей статьи составляют, однако, вопросы, связанные не с проблемой построения самой машины, а с проблемой создания специального способа записи в ней научно-технических сведений, т. е. машинного языка.

Машинный язык не следует путать с машинным кодом, который строится для данного алфавита и посредством которого любой текст, составленный из букв данного алфавита, кодируется для записи в машину в виде последовательностей нулей и единиц (если машина построена на двоичном принципе*). Речь идет о создании специального искусственного абстрактного языка для записи научных сведений с целью последующего помещения этих сведений в «памяти» машины.

Чтобы лучше изучить наш предмет, отвлечемся от всех посторонних вопросов, в том числе и от техники. Будем руководствоваться следующим принципом: «то, что может быть формализовано, может быть и автоматизировано». Итак, наша общая задача—формализовать поиск нужной информации, а на первых порах создать язык, пригодный для формализации поиска, сравнения и отождествления информации. А кто все это будет делать: машина или штат обученных (но *формально* обученных) людей—пока не важно.

Если иметь дело со сведениями, записанными на каком-либо реальном (например, русском) языке, то будет трудно выполнить чисто формально (т. е. без обращения к содержанию) поиск, сравнение и т. п. нужных сведений, ибо реальный язык обладает двумя существенными недостатками:

1. Он неоднозначен, субъективен; он имеет, как говорят, «прагматическую» функцию (т. е. предусматривает индивидуальную реакцию того или иного организма).

2. Он не формализован, не имеет ясной структуры, богат исключениями и потому неудобен для машины**).

В какой мере эти недостатки являются препятствием к использованию реального (скажем, русского) языка в качестве машинного? Пер-

*) С этой точки зрения телеграмма, записанная азбукой Морзе,—это все еще текст на русском, английском или другом реальном языке, только иначе закодированный. Говоря приблизительно, мы понимаем под языком некоторую систему записи фактов объективной действительности, а под кодированием—какое-либо обратимое преобразование уже имеющегося языка, причем язык рассматриваем «с точностью до кодирования» (т. е. два языка, один из которых получается из другого кодированием, мы считаем одним и тем же языком).

**) Реальный язык можно рассматривать как некоторую идеальную структуру плюс накладываемые на нее «шумы» (т. е. неправильности); при восприятии человеком языкового сообщения эти «шумы», по-видимому, устраняются.

вое препятствие носит принципиальный характер, хотя его можно было бы устранить путем выделения из русского языка «непрагматической части», т. е. допущения лишь некоторых стандартных терминов и конструкций (нечто в роде Basic English) с точно установленными правилами их семантического истолкования. Второе препятствие носит более практический характер: принципиально возможно заложить в машину всю грамматику русского языка со всеми исключениями, и при создании автоматического перевода так и приходится поступать (если желать переводить любые тексты), но для задачи создания информационной машины такой способ окажется практически бесполезным (можно будет производить лишь простейший поиск, например, поиск всех фраз, в которые входит заданное слово; более сложные задачи будут решаться медленно и потребуют громоздких программ*).

Выход подсказывается отчетливым пониманием того, что наша цель — поместить в машину не тексты какого-либо языка сами по себе, а информацию, которая записана в этих текстах. Коль скоро реальные языки обладают с точки зрения наших целей серьезными недостатками, не следует использовать ни один из них в качестве машинного языка. Нужно создать искусственный абстрактный язык, наилучшим образом пригодный для записи естественнонаучных сведений, подобно тому как шахматная нотация удобнее, чем реальные языки, для записи шахматных партий. К этому абстрактному языку предъявляются, таким образом, следующие требования:

1) он должен быть недвусмысленным; каждая запись должна допускать однозначное истолкование;

2) он должен быть удобным для формализации следующих функций:

а) выбор нужной информации;

б) отождествление различным образом записанных фактов;

в) дедукция, т. е. производство логического вывода.

Эти требования отчасти противоречат друг другу и делают задачу создания оптимального (с точки зрения этих требований) языка задачей экстремального характера.

В качестве первой задачи надо создать частные абстрактные языки, обслуживающие отдельные отрасли науки (такие, как математика, химия) или даже ветви этих отраслей (геометрия, органическая химия). В какой-то степени эти языки уже созданы (ведь говорят же о «языке формул» математики или химии). Создавать частные абстрактные языки должны представители соответствующей конкретной науки совместно с логиками.

На абстрактном языке будет храниться информация в машине (мы опять-таки отвлекаемся здесь от того, что в машине все тексты будут закодированы в виде последовательностей нулей и единиц) и на первых порах будет происходить «разговор» человека с машиной (человеку ведь легче выучиться машинному языку, чем машине — человеческому). Как правило, человек будет задавать вопросы, а машина — отвечать, но иногда и машина может задать человеку вопрос, например, попросить его уточнить свой вопрос.

В дальнейшем необходимо разработать формальные правила перевода с абстрактного языка на реальные и обратно.

Автоматический перевод с реальных языков на абстрактный даст возможность автоматически записывать в машину факты, выраженные на тех или иных реальных языках (быть может, на первых порах тексты долж-

*) Эти недостатки реальных языков, заключающиеся в неадекватности их структуры структуре человеческих знаний, совершенно естественны. Ведь языки довольно поздно стали использоваться для записи сложных научных фактов (уже словарный запас языка не приспособлен для обозначения научных понятий).

ны будут предварительно препарироваться человеком). Это позволит записать в машину имеющиеся уже сейчас в литературе сведения (без создания автоматического перевода является, конечно, практически невозможным ввести в машину уже имеющуюся литературу). Автоматический перевод с абстрактного языка на реальный позволит воспроизводить сведения на желаемом языке. Таким образом, при разговоре человека с машиной человек получит право не знать машинного языка. Наконец, абстрактный язык можно будет рассматривать в качестве языка-посредника, и тем самым будет решена задача автоматического перевода научно-технических текстов: чтобы перевести текст с языка *A* на язык *B* надо перевести текст с языка *A* на машинный язык и потом с машинного языка на язык *B*.

§ 2

Создание частного абстрактного языка для данной теории должно начаться с создания соответствующей метатеории.

Предметом теории является определенный фрагмент реального мира, предметом метатеории—сама система положений данной теории. Отношение теории к метатеории аналогично отношению искусства к искусствоведению. Художник отображает объективную действительность в художественных образах, затем приходит искусствовед и изучает саму систему этих образов и отношение этой системы к действительности, но не отображаемую действительность. Точно так же ученый отображает действительность в научных образах, а метатеоретик изучает систему этих образов и ее связь с отображаемой действительностью.

Создать метатеорию легче всего для таких «теорий», предметом которых являются не природные процессы, а процессы, происходящие по законам, установленным людьми. Таковы, например, спортивные игры и уличное движение.

Предмет теории уличного движения—наилучшая организация уличного движения. Предмет метатеории уличного движения—уже выработанная в теории система правил уличного движения. Рассмотрим для примера несколько подробнее, в чем должна была бы состоять эта метатеория (если бы она существовала). Прежде всего исследуется система понятий, фигурирующих в правилах. Эти понятия классифицируются в зависимости от их отношения друг к другу. Таким образом, выделяются, например, такие группы понятий: то, что движется (транспорт, пешеходы); то, по чему движутся (улицы, площади и т. д.); то, что регулирует (регулирующие, светофоры, дорожные знаки и т. д.). Выделяются скрытые, не явно содержащиеся в правилах понятия («здание», «прямое направление» и т. д.). Выделяются понятия неопределяемые («светофор») и определяемые («перекресток»). Затем происходит исследование предложений теории. Они подразделяются на дефинитивные, в которых определяется какое-либо понятие, и позитивные, в которых формулируется какой-либо факт. Вся система предложений исследуется затем с точки зрения ее непротиворечивости (не противоречат ли какие-либо правила уличного движения друг другу), полноты (все ли случаи предусмотрены правилами уличного движения) и однозначности (не допускает ли какое-либо правило разных по смыслу истолкований). Заметим, что значительную часть этих исследований можно провести, вовсе не зная, что такое уличное движение, или зная о нем очень мало. (Можно, например, считать, что «светофор»—это живое существо, а «регулирующий»—механизм.)

Вопросы о связи понятий и предложений теории между собой без обращения к их смыслу составляют синтаксическую часть метатеории. (Таковы, например, вопросы непротиворечивости.) Вопросы о смысловом

истолковании понятий и предложений теории составляют семантическую часть метатеории.

Метатеория, таким образом, четко очерчивает круг понятий и предложений данной теории. Она классифицирует понятия по их характеру, по степени общности, по семантическим связям. Она выясняет общелогические понятия, требующиеся для данной теории (для теории уличного движения, например, не нужно общее понятие натурального числа, а для химии оно нужно). Метатеория должна фиксировать и устранять омонимию, т. е. употребление одного и того же термина в разных смыслах. (Например: «водород горюч» и «водород воспламенился»; в первой из этих фраз термин «водород» означает водород как вещество, а во второй— данный конкретный объем водорода.) Метатеория должна изучать способы, какими вводятся в данной науке новые понятия. В математике, например, они вводятся, по крайней мере, двумя способами—либо явным определением, либо как неопределяемое понятие (но тогда перечисляются его свойства). Определения в нематематических науках имеют свою специфику, поскольку используют непосредственный опыт (определение таких, например, понятий, как энергия, представляет значительные трудности). Изучая систему предложений теории, метатеория должна выделить среди них аксиомы, а также указать правила (правила вывода), по которым разрешается переходить от одних предложений к другим.

Метатеория призвана выработать точные нормы «непрагматического языка», на котором должны писаться сочинения по данной теории. В дальнейшем, быть может, следовало бы при научно-технических издательствах учредить органы, подобные отделам технического контроля (ОТК) на заводах; эти органы должны были бы контролировать печатную продукцию с точки зрения установленных синтаксических и семантических норм. При этом, конечно, вовсе не должно проверяться, верна или не верна данная статья (так же, как не дело ОТК проверять, нужно или не нужно данное изделие); должны контролироваться лишь такие формальные вещи, как недвусмысленность, непротиворечивость и т. д.

Наконец, создание информационной машины требует от метатеории установления характера вопросов, который ставит перед собой исследователь, обращаясь к литературе, и которые теперь будут задаваться машине.

Из сказанного видно, что создание метатеории—задача серьезная, лишь немного уступающая по трудности созданию самой теории. Думается, однако, что необходимость создания метатеории диктуется не только созданием информационной машины, но и всем развитием науки. Ведь создание метатеории имеет огромное значение для самой науки, поскольку метатеория контролирует чистоту понятий и помогает лучше разобраться в системе наших представлений о мире. Метатеории значительно помогут и в решении таких повседневных практических задач, как составление рефератов, указателей, энциклопедий и т. д. Некоторые метатеории уже построены или строятся. Так созданы основы метаматематики, видны контуры метакимии, положено начало метабиологии. В дальнейшем можно было бы приступить к созданию метамедицины.

Самое сложное при построении частного абстрактного языка для данной теории—это создание метатеории. В метатеории каждое предложение теории приводится к стандартному виду: «такие-то объекты находятся в таких-то отношениях», причем и возможные объекты и возможные отношения перечислены заранее. Построение языка состоит, далее, в том, что происходит символическая запись этих приведенных предложений.

Методы символической записи заимствуются из математической логики, в которой далеко разработаны средства выражения синтаксических отношений.

В математической логике имеются специальные знаки, выражающие общелогические связи и операции: «и», «или», «существует» и т. д. К этим знакам теперь надлежит прибавить знаки для объектов рассматриваемой теории (в том числе для понятий), т. е. термины (постоянные и переменные), знаки для отношений (предикатные знаки), знаки для функций (функторы). Например, если знак G означает отношение «лежит на» (в геометрии), то предложение «точка a лежит на прямой p » запишется в виде $G(a, p)$.

Если условиться точки обозначать латинскими буквами, а прямые— греческими, то предложение «для любых двух точек найдется проходящая через них прямая» надлежит привести к виду «для всякого объекта a и всякого объекта b существует такой объект x , что a лежит на x и b лежит на x »; символическая запись этого предложения такова:

$$(a)(b)(\exists x)[G(a, x) \& G(b, x)].$$

Можно не вводить различий между записями объектов различных типов, но тогда придется ввести предикаты «быть объектом такого-то типа». Пусть, например, T означает «быть точкой» и P — «быть прямой». Тогда предложение «для любых двух точек найдется проходящая через них прямая» приведет к виду «для всякого объекта a и всякого объекта b , если a есть точка и b есть точка, то существует такой объект c , что c есть прямая, a лежит на c и b лежит на c »; запись такова:

$$(a)(b)\{T(a) \& T(b) \rightarrow (\exists c) \rightarrow [P(c) \& G(a, c) \& G(b, c)]\}.$$

Еще один пример. Пусть t означает температуру, Tv — предикат «быть в твердом состоянии» и \approx — воду. Предложение «точка замерзания воды есть 0° » можно привести к виду «утверждение, что температура воды меньше 0° , равносильно утверждению, что вода находится в твердом состоянии». Запись такова:

$$(t(\approx)) < 0^\circ \sim Tv(\approx).$$

Создание символических способов записи возможно (и даже довольно несложно) при условии, что уже разработана метатеория (здесь видно значение выявления в метатеории скрытых понятий; достаточно попробовать записать символически возможность движения транспорта на зеленый свет).

При построении абстрактного языка, помимо математической логики, целесообразно использовать формальную семантику. Следует различать семантическую философию, семантику в языкознании, иначе называемую семасиологией, как учение о значении лингвистических элементов (слов, выражений и т. п.) и семантику в логике как учение о значении логических элементов (понятий, суждений и т. п.). Здесь нас интересует именно семантика в логике, в частности, формальная семантика, разработанная Карнапом и другими учеными. Необходимо различать философские взгляды этих ученых и их конкретные достижения в построении на базе математической логики формальных символических систем для записи естественнонаучных фактов. Эти формальные системы при всем их несовершенстве и неокончателности могут с успехом быть использованы для построения машинного языка. Существенны также разработанные представителями этого направления в логике методы формального (без обращения к содержанию!) анализа символических выражений; необходимо учесть, что машине доступен лишь такой, формальный и по необходимости «бессодержательный» язык. Приходится с сожалением признать, что у нас до последнего времени

имели место недооценка и даже неправильная оценка семантики, которую к тому же часто смешивали с семантической философией (недавнее положение тут можно сравнить с недавним же положением кибернетики).

Абстрактный язык, или формальная семантическая система, состоит из списка элементарных символов (знаков), правил образования (устанавливающих, какие комбинации знаков допускаются), правил преобразования (устанавливающих, какие допускаются преобразования выражений с целью получения логического вывода) и правил интерпретации (устанавливающих, какой смысл надлежит приписывать выражениям, составленным по правилам образования).

В дальнейшем создание частного абстрактного языка и метатеории должно идти параллельно (с опережением языка метатеорией).

§ 3

Специальные вопросы возникают в связи с записью информации в машину.

Прежде всего вопрос о том, что записывать. Дело в том, что можно записать числовую таблицу (которая займет значительный объем памяти), а можно записать формулу, из которой таблица может быть получена простой подстановкой. Формула, конечно, требует меньшего объема памяти для своей записи, но тогда надо предусмотреть выполнение машиной таких операций, как подстановка чисел в формулу. Можно пойти еще дальше и записывать в памяти машины не все факты данной теории, а только некоторые базисные факты, из которых остальные факты могут быть получены посредством логического вывода (при условии, что машина сможет сама осуществлять этот логический вывод). Возможно, далее, что окажется целесообразным избрать некоторую специальную структуру фраз машинного языка, более приспособленную для поиска информации. Пусть, например, имеется фраза, содержащая термины a_1, a_2, \dots, a_s ; обозначим ее $\mathfrak{A}(a_1, a_2, \dots, a_s)$. Если ведется поиск по признакам a_1, a_2, \dots, a_s (т. е. нас интересуют фразы, содержащие термины a_1, a_2, \dots, a_s), то придется при поиске проникать внутрь фразы. Этого можно было бы избежать и, следовательно, упростить поиск, если указанную фразу переписать в другом виде: вынести все термины вперед, а внутри фразы заменить эти термины их номерами: a_1, a_2, \dots, a_s ; $\mathfrak{A}(1, 2, \dots, s)$. (Например, «стул стоит на полу» переходит в «стул, пол; 1 стоит на 2».) Тогда поиск можно будет проводить лишь по приставкам, стоящим впереди фразы. Точно так же можно выносить в приставку не только предметы, но и отношения: «стул стоит на полу» переходит в «стул, пол, стоять на; 3(1, 2)». Такая обработка фразы удобна при осуществлении поиска фраз, содержащих данное отношение. Можно, конечно, сделать и так, чтобы машина сама при вводе производила необходимую перестройку фразы.

Несколько замечаний о размещении сведений в памяти машины. Память машины можно считать организованной по иерархическому принципу: «дом», в «доме» — «этажи», на «этажах» — «коридоры», в «коридорах» — «комнаты», в «комнатах» — «шкафы», в «шкафах» — «полки», на «полках» — «папки». Каждое хранимое в памяти сведение имеет свой адрес, который состоит из номера соответствующего «дома», «этажа» и т. д. Примерный состав хранимого в памяти таков:

1. Предложения науки на абстрактном языке: аксиомы, теоремы, определения.

2. Правила логического вывода — правила преобразования (т. е., по существу, часть программы машины).

3. Сведения, записанные на обыкновенном языке — цитаты, библиографические ссылки, не формализуемые определения и т. д.

Рациональное размещение всего этого обширного и разнородного материала представляет собой сложную задачу. Критерием рациональности размещения должны служить занимаемый в памяти объем и затрачиваемое на поиск время. При этом следует учесть, что по причинам, о которых говорилось вначале, мы лишь в ограниченной степени можем использовать классификационный принцип.

Размещение информации в памяти должно предусмотреть возможность динамического поиска, т. е. поиска по определенной программе, с использованием промежуточных результатов. Например, нам нужно найти вещества, вредно действующие на бациллу туберкулеза (пример заимствован у В. В. Серпинского и Г. М. Влэдуда). Для этого можно запрограммировать такой поиск: сначала найти вещества, необходимые для жизни бациллы, а затем—вещества, вступающие с первыми веществами в устойчивые соединения. Промежуточные результаты поиска могут без надобности и не выдаваться наружу. При организации памяти машины могут быть учтены, в частности, следующие положения:

1°. Вопрос организации размещения сведений встает лишь в связи с необходимостью экономить место и время. Принципиально возможно разместить сведения как попало, так как машина обладает возможностью сплошного просмотра всей памяти. Программа поиска в этом случае будет такова: «просматривай последовательно все имеющиеся сведения, отбирая то, что нужно».

2°. Можно использовать классификацию по неизменным признакам и различать, например, аксиомы, теоремы, определения, цитаты и т. д., помещая каждый из этих видов сведений в своем участке памяти.

3°. Можно было бы помешать при каждом «доме», «этаже» и т. д. некоторое заглавие (нечто вроде списка жильцов). Это позволило бы просматривать сначала лишь заглавия, а в дальнейшем анализировать только участки памяти с подходящими заглавиями.

4°. Целесообразна, по-видимому, известная автономия отдельных участков памяти с тем, чтобы поиск мог вестись параллельно на нескольких участках с ограниченным обменом информации между ними.

5°. Необходимо широко использовать систему перекрестных ссылок. Например, при каждом термине может стоять следующий адрес этого же термина или адреса связанных с ним (семантически) терминов.

6°. Из предыдущего пункта видно, что при организации памяти важно изучение семантических связей. Машина сама может выявлять устойчивые семантические связи (обнаруживающиеся, например, при динамическом поиске) и запоминать их для последовательного использования.

7°. При организации памяти должны использоваться вероятностные соображения (должно затрачиваться в среднем небольшое время на поиск). Тут могут пригодиться и методы исследования операций и даже методы теории игр.

8°. В дальнейшем надо предусмотреть возможность автоматической реорганизации памяти машины в зависимости от накопленного ею «опыта» и мыслить, таким образом, информационную машину как обучающуюся.

§ 4

Построение информационной машины и создание информационного языка ставит специфические проблемы перед математикой и прежде всего перед математической логикой.

При создании машинного языка приходится решать три основные задачи: 1. Символическая запись понятий. 2. Символическая запись суждений. 3. Символическая запись контекста.

В математической логике разработаны методы решения второй задачи: записать символически логические отношения между понятиями, составляющими суждения (об этом уже говорилось ранее).

Задача о записи контекста и не вставала в математической логике; ведь, формально рассуждая, любой текст может быть записан одним предложением, одним высказыванием. Но при этом мы, конечно, что-то теряем. Известно, например, что в языке значение тех или иных выражений зависит от контекста (при этом получается экономия в языковых средствах, так как разные смыслы записываются одним и тем же выражением). Было бы целесообразно формализовать эти особенности контекста. Это облегчило бы и перевод с реальных языков на абстрактный, поскольку абстрактный язык при формализации в нем контекста был бы ближе к реальному.

Очень важно решение первой задачи, т. е. выбор рациональной терминологии машинного языка. Было бы желательно, чтобы связи между понятиями выражались самой терминологией и, следовательно, не нужно было бы специальных предложений, фиксирующих эту связь. Если использовать обычную, скажем, русскую терминологию и задать машине какой-нибудь вопрос об антибиотиках, то чтобы машина включила в круг своих рассуждений также и пенициллин, необходимо иметь в памяти предложение, гласящее, что пенициллин является антибиотиком. Такая фраза не будет нужна, если, скажем, обозначением для антибиотиков будет служить Q , для пенициллина Qa , для синтомицина— Qb и т. д. Здесь может помочь космоглотика (наука об искусственных языках); ведь еще в XVII в. создавались так называемые философские языки, терминология которых строилась по описанному только что иерархическому принципу. Отчасти поможет здесь математической логике и химия с ее разработанной системой названий органических веществ (замечу, что проблема тождества для названий химических веществ представляет несомненный математический интерес, особенно в свете неразрешимости общей проблемы тождества для ассоциативных исчислений). В математической логике еще не созданы, но должны быть созданы способы символической записи понятий, отражающие построение одних понятий из других, подобные уже созданным способам символической записи высказываний, отражающим построение одних высказываний из других.

Среди отдельных проблем, встающих перед математической логикой, можно упомянуть еще следующие:

1°. Абстрактный язык для данной теории можно считать построенным, когда будут заданы правила, позволяющие по любому факту данной теории находить его языковые выражения. Речь идет, следовательно, об алгоритмизации соответствия «смысл»—«формальное выражение». Встает необходимость точного определения того, что значит в этом случае алгоритмизация, для чего в свою очередь требуется формализация понятия смысла.

2°. Математическая логика обслуживает сейчас главным образом математику. Необходимо создать «прикладные» разделы математической логики, обслуживающие конкретные науки (такие, как химия).

3°. В естественнонаучных построениях фигурируют такие не свойственные математике и потому не формализованные в классической математической логике категории, как возможность, правдоподобность и т. д. Необходимо поэтому формализовать эти понятия и развивать, стало быть, многозначную логику, модальную логику и другие направления интенциональной логики, или логики содержания (наряду с экстенциональной логикой, или логикой объема).

4°. Наконец, надо заняться выделением в множестве доказуемых формул базисных сетей, т. е. множеств формул, из которых все другие доказуемые формулы выводимы в данное число шагов (эта идея принадлежит Р. Л. Добрушину).

Решение перечисленных задач требует развития теоретических исследований в различных областях математической логики, в частности, по теории исчислений, теории моделей и т. д. Одновременно необходимо приступить к созданию специальной машинной логики, изучающей в общем виде строение современных машин и взаимодействие их элементов.

Для записи в машину выражения абстрактного и реального языков должны кодироваться в виде последовательностей двоичных знаков. При этом кодировании возникает ряд проблем, характерных для теории информации. Одной из таких проблем является создание оптимального по длине кода. Другая проблема состоит в создании помехоустойчивого кода (надо предусмотреть возможность сбоев в машине); здесь предполагается использование теории так называемых «кодов с обнаружением ошибки».

Удовлетворение требованиям теории информации возможно лишь на базе известного компромисса с теорией алгоритмов, ибо есть основания полагать, что наиболее рациональный с теоретико-информационной точки зрения код будет нерационален с точки зрения теории алгоритмов. Здесь имеются в виду алгоритмы поиска, алгоритмы дедукции и алгоритмы перевода. Изучение этих алгоритмов, по существу, смыкает теорию алгоритмов с работами по теоретическому программированию, ведущимися коллективом под руководством А. А. Ляпунова. Требования теории алгоритмов окажут, по-видимому, значительное влияние на окончательную структуру машинного языка.

Важные задачи встают и перед лингвистикой. С сожалением надо признать, что у нас недооценивается прикладное значение лингвистики (так же как и логики). Лингвистика рассматривается главным образом (помимо своего чисто теоретического аспекта) как основа для составления школьных грамматик, для обучения иностранному языку и т. п. А ведь лингвистика имеет важные практические применения при разработке способов рациональной записи фактов. Работы в этом направлении у нас насчитываются единицами (здесь имеются в виду прежде всего работы Н. Ф. Яковлева и Е. Д. Поливанова по созданию алфавитов для бесписьменных языков). Характерно, что для развития прикладного аспекта лингвистики необходимо развитие самых теоретических ее отраслей и прежде всего ее связей с семиотикой.

§ 5

Построение информационной машины и создание машинного языка есть, по существу, кибернетическая задача. Кибернетика, как известно, изучает живые и искусственные организмы и коллективы таких организмов с точки зрения их способности воспринимать, хранить, перерабатывать и передавать информацию. Поиск информации в литературе—кибернетический процесс, в котором участвуют и обмениваются информацией такие организмы, как человек, книги, библиотеки, библиографические бюро и т. п. Поиск информации при помощи машины—также кибернетический процесс, моделирующий предыдущий.

Современная вычислительная машина моделирует мышление и в меньшей степени память; по своей структуре она моделирует скорее машинно-счетную станцию, чем мозг. Информационная машина наряду с мышлением моделирует также и память; по своей структуре она моделирует библиотеку с достаточным штатом библиографов и метатеоретиков (плюс машинно-счетную станцию для обработки библиографических карточек).

Однако и для вычислительной и для информационной машины мозг является идеалом. Поэтому чрезвычайно важно подчеркнуть связь всей затронутой проблематики с психоневрологией. Изучение поиска информации в мозге может дать многое для построения машины. Достаточно указать на необычайно быстрое протекание процесса вспоминания у человека, на роль в этом процессе ассоциаций и опыта. Очень желательно смоделировать в машине процессы вспоминания по ассоциации, а также накопление опыта (обучающаяся машина, о чем уже говорилось выше). С этой целью необходимо усилить изучение процессов, связанных с хранением и переработкой информации в мозгу, а также параллели между этими процессами и процессами, происходящими в машинах (заметим, кстати, что часто проводимую аналогию между нейтронами и электронными лампами в качестве хранителей информации нельзя считать достаточно обоснованной).

В дальнейшем, с передачей информационной машине все большего количества «человеческих функций» будет, по-видимому, происходить постепенная «прагматизация» абстрактного языка; в частности, машина сможет воспринимать контекст.

Значительный интерес с точки зрения создания информационной машины представляет такая отрасль кибернетики, как общая теория знаковых систем (семиотика). Как правило, знаковые системы, которыми пользуется человек, являются избыточными. Например, знаки c c o o воспринимаются человеком как буква c ; можно думать, что во всех этих знаках содержится некий общий инвариант (аналогично, смысл является инвариантом при выражении одного и того же суждения на разных языках; язык-посредник и есть формализация этого инварианта). Подобная избыточность проявляется, конечно, не только в восприятии отдельных букв: известно, что примерно половину букв текста можно зачеркнуть, не нарушая возможности однозначного опознавания содержания. Эту избыточность можно сократить, уменьшив, таким образом, объем памяти.

Наконец, абстрактный машинный язык сможет в будущем выполнять важную кибернетическую информацию в качестве средства непосредственного обмена информацией между машинами.

§ 6

Построение машины и создание языка должны идти одновременно, оказывая влияние друг на друга. Вероятно, первый, ограниченный этап деятельности машины будет характеризоваться следующим:

1°. Машинный язык еще не будет окончательно создан. Сведения будут храниться на русском языке. Кроме того, будет разработан особый «телеграфный» стиль рефератов и значительная часть материалов будет храниться на «телеграфном» языке.

2°. Поиск будет происходить по отдельным признакам, без проникновения во внутреннее содержание предложений и, быть может, на первых порах без проникновения даже во внутреннее содержание статьи. Такой поиск можно представлять себе как поиск по заглавию (естественному или искусственному; пример искусственного заглавия — шифр по десятичной классификации, который ставит на книге или статье библиограф).

Построение информационной машины и машинного языка требует углубленных теоретических исследований, и прежде всего исследований, нужных для решения основной задачи — создания метатеорий. Думается, что необходимо как можно скорее покончить с пренебрежением к формализации теории. Хочу высказаться также в защиту таких страшных слов, как «формализм» и «упрощенчество». Иногда бывает необходимо

прибегать и к формализму, и к упрощенчеству хотя бы потому, что только формальное и упрощенное мы можем передать машине. Чтобы создать машину, необходимо стать на ее точку зрения.

Для развития метатеорий, для построения машинного языка и связанных с ним алгоритмов перевода необходимо не только идейное содружество наук, но и организационное содружество научных учреждений. Необходимо также организовать соответствующую подготовку кадров.

Решение поставленных задач, научных и организационных, требует приложения значительных сил и средств. Затрата этих сил и средств оправдывается общенаучной значимостью указанных задач. Ведь их решение приведет не только к возможности быстрого и полного получения нужных сведений и тем самым к повышению производительности умственного труда, но и к тому, что мы станем лучше разбираться в самой системе человеческих знаний.

ЛИТЕРАТУРА *)

к § 1

- [1] Perry J. W., Kent A., Berry M., *Machine literature searching*. New York—London, 1956.
 [2] Гутенмахер Л. И., Статистические и информационные машины нового типа, *Вестник АН СССР*, № 10 (1956), 12—21.

к § 2

- [3] Полтев К. М., *Пособие по правилам движения автотранспорта*. М., 1957.
 [4] Клини С. К., *Введение в метаматематику*. М., ИЛ, 1957.
 [5] Woodger J., *Biology and language*. Cambridge, Great Britain, 1952.
 [6] Carnap R., *Introduction to semantics*. Cambridge, USA, 1946.

к § 4

- [7] Гильберт Д., Аккерман В., *Основы теоретической логики*. М., ИЛ, 1947.
 [8] Дрезен Э., *В поисках всеобщего языка*. М.—Л., 1925.
 [9] *Автоматы (сборник статей)*. М., ИЛ (1957).
 [10] Shannon C. E., Weaver W., *The mathematical theory of communication*, Urbana, 1949.
 [11] *Коды с обнаружением и исправлением ошибок (сборник статей)*. М., ИЛ, 1956.
 [12] Яковлев Н. Ф., *Математическая формула построения алфавита (опыт практического приложения лингвистической теории)*, *Культура и письменность Востока*, кн. 1. М., 1928, 41—64.
 [13] Иванов В. В., *Лингвистические взгляды Е. Д. Поливанова*, *Вопросы языкознания*, № 3, 1957, 55—76. (См. в списке научных работ Е. Д. Поливанова, приведенном в указанной статье В. В. Иванова, работы на № 26, 32, 35, 83.)

к § 5

- [14] Wiener N., *Cybernetics*, Paris, 1948. [Русский перевод: Винер Н., *Кибернетика*, Изд. «Советское радио», М., 1958.]

Поступило в редакцию 15 X 1957

*) Обсуждаемую в настоящей статье проблематику затрагивает ряд докладов на упомянутом в подстрочном примечании на стр. 39 Совещании (перечень докладов на этом Совещании приведен в 4-м выпуске «Проблем кибернетики» на стр. 268—268). Изложения двух из них были опубликованы за время печатания данной статьи:

Гутенмахер Л. И., *Электрическое моделирование некоторых процессов умственного труда*, *Вестник АН СССР*, № 10 (1957), стр. 88—95.

Иванов В. В., *Лингвистические вопросы создания машинного языка для иформационной машины*, *Материалы по машинному переводу*, сб. 1, ЛГУ, 1958, стр. 10—33.