

УДК 62.506.2.

*В. А. ЛОВИЦКИЙ*, канд. техн. наук, *Н. В. МОРОЗОВА*, *Е. И. ТИШИНА*

#### К ВОПРОСУ ОБ ИНЖЕНЕРНОЙ ПОЭТИКЕ

Введение в эксплуатацию мощных ЭВМ третьего поколения позволило подойти к созданию человеко-машинных систем, призванных быть «интеллектуальными» советчиками человека. «Советы» таких искусственных систем должны быть не только естественно-языковыми, но и, по возможности, нести как можно меньше синтаксической информации при сохранении семантики «совета». Известно [1], что рифмованные строчки несут меньше синтаксической информации, чем строчки прозы. Следовательно, для человеко-машинных систем, в которых человек вынужден воспринимать и перерабатывать большой объем информации, выдаваемых ЭВМ, проблема ее организации имеет немаловажное значение. Решением этой проблемы и занимается инженерная поэтика.

Существует несколько различных подходов к творческому процессу в искусстве. Наибольший интерес представляет так называемое пермутационное искусство [3], в котором постулируется существование атомов структуры, после чего машина порождает и исследует поле возможностей, определяемое заранее заданным алгоритмом. Под пермутацией понимается комбинирование простых, обладающих ограниченной изменчивостью элементов, открывающее колоссальное поле возможных вариаций. Таким образом, реализуется то разнообразие в единообразии, которое составляет основу свойств любого произведения искусства. При этом на первый план выступает иерархия порядков или уровней анализа. Машина, сочиняющая стихотворное произведение, может работать на уровне букв, слов, словосоче-

таний. В зависимости от этого получают пермутационные произведения различных уровней сложности.

Исследование структуры языка естественно должно привести к учету частоты повторения целых слов. Учитывая дополнительно и то, что человек сначала учится оперировать словами, определяет их смысл, и только значительно позже он учит грамматику, изучает буквы, учится писать, приходим к выводу, что лучше всего за основу работы системы взять готовые слова с учетом связи этих слов между собой.

Рассмотрим систему «Поэтесса», принцип работы которой основан на использовании структуры ближнего порядка [3]. Пример построения такой структуры, показывающий взаимосвязь слов в строке и возможность пересечения строк, приведен на рис. 1. Очевидно, что эта структура позволяет получать варианты строк, отличные от исходных.

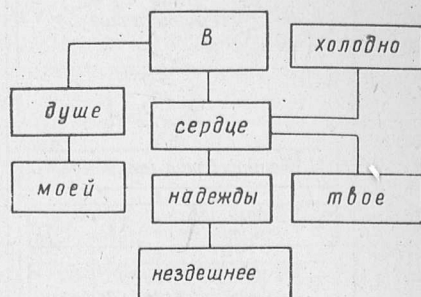


Рис. 1

Общая схема алгоритма системы «Поэтесса» дана на рис. 2. Схема разбита на блоки, которые показывают поэтапную разработку системы по мере ее усложнения.

На первом этапе работы исследовались закономерности построения новой строки. Слова из вводимой строки MSMB (блок 1 (1)) последовательно заносятся в массив-словарь MVOCAB (блок 1 (2)). В результате каждое слово занумеровано. Повторяющиеся слова заносятся в MVOCAB только один раз. Каждой вводимой строке однозначно ставится в соответствие код этой строки MSTRIN (блок 1 (3)), состоящий из порядковых номеров этих слов в словаре MVOCAB. Большую роль играет взаимосвязь слов, т. е. необходимо запомнить, какое слово за каким следует или какое слово какому предшествует, определить, какие слова могут быть начальными, конечными, и конечными и начальными в строке одновременно. Для определения начального и конечного слова в строке, в MSTRIN, вводятся специальные метки. Далее, вводятся два массива MALPH и MSTRUC. Массив MALPH содержит перечень кодов всех слов по мере формирования кодов всех вводимых строк (блок 1 (4)). Повторяющиеся коды слов заносятся только один раз. Массив MSTRUC, заполнение которого происходит параллельно с записью кодов в MALPH, сохраняет последовательность слов в строке, в нем (массиве) определяется частота встречаемости слов в числовом материале. Массивы MALPH и MSTRUC симметричны, т. е. один из параметров, определяющих их размеры, совпадает. MSTRUC состоит из двух основных частей — «что

следует» и «что предшествует». В MSTRUC в столбце, соответствующем коду записываемого в MALPH слова, в части «что следует» записывается номер кода следующего за ним слова. В части «что предшествует» записывается номер кода пред-

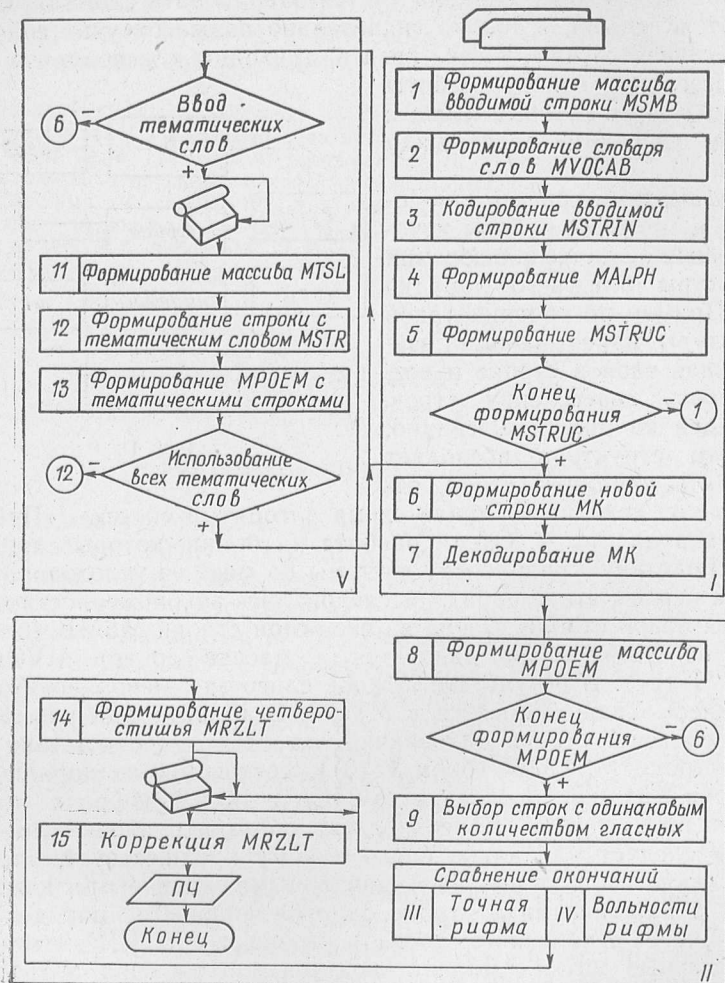


Рис. 2

шествующего слова. В процессе поступления числового материала массив MSTRUC заполняется полностью (блок 1 (5)). Построение новой строки осуществляется с помощью MSTRUC в режиме генератора и начинается с наиболее часто встречающегося слова в MALPH. Если код этого слова несет одновременно и признак начала строки, то он записывается в МК. Выбор вто-

рого слова осуществляется с помощью MSTRUC. В столбце с номером, соответствующем первому слову, в части «что следует» по частоте встречаемости выбирается слово, следующее за первым. Процесс построения строки продолжается до тех пор, пока вновь выбранное слово не будет нести признака конца строки. Вновь созданная строка МК (блок 1 (6)) раскодируется (блок 1 (7)) и каждому коду в МК однозначно ставятся в соответствие слова из словаря MVOCAB. Формирование последующих строк производится аналогично. Как пример работы этого алгоритма можно привести следующие строки:

Глубокий снег в избушке отдохнуть  
Повисла на жатве ночи  
Сверкают в хрустальное ведро.

На этом промежуточном этапе результат представляет собой строку с определенным набором слов. Как видно из приведенного примера, совокупность этих и подобных им строк ничем не напоминает поэтическое произведение. Это объясняется тем, что поэтическая речь отличается от обыкновенной определенными приемами ритмизации речи и для получения поэтического текста необходимо учитывать внутреннюю организацию стиха. Школьная метрика различает три системы стихосложения — метрическую, основанную на длительности; тоническую, построенную на ударении; силлабическую, исходящую из счета слогов [4]. Остановимся немного подробнее на силлабическом способе стихосложения, который задает первичную ритмическую структуру при построении текста в системе «Поэтесса». Размер стиха здесь определяется только количеством слогов, причем все слоги признаются равными между собой. Различий между слогами не делается никаких — ни по долготе и краткости, ни по ударности и неударности [4]. Только в школе учат, что такое ударение, а в быту ударение как бы не существует. Поскольку ударные и неударные слоги не различаются, остается только один признак — самое количество слогов. Но стих держится не только на размере, но и на выделении фразовых единиц. И в силлабическом стихосложении учитывается фразовое ударение, которое стоит в конце полустушия. Это единственное ударение, влияющее на силлабическое стихосложение.

Дополнительным приемом ритмизации речи служит расчленение произведения на строфы: если основным признаком стиха является последовательность равновеликих отрезков, создающих первичную ритмическую волну, то последовательность равновеликих групп таких отрезков, да еще упорядоченно размещенных в группе, углубляет речевую размерность [5]. Учитывая указанное выше, систему строили генерацией текстов в виде строф, каждая из которых содержит четыре стихотворных строки. Различные варианты строк, полученные на первом этапе обучения системы, содержатся в массиве МРОЕМ (блок II (8)).

Система подсчитывает и запоминает суммарное количество гласных букв в каждой строке, которые однозначно определяют число слогов в строке. Из массива МРОЕМ необходимо построить максимально возможное число четверостиший. Порядок строк с одинаковым количеством гласных (блок II (9)) выбирается по одному из трех возможных способов перестановки строк в данной строфе:  $(aa, bb)$ ,  $(ab, ab)$ ,  $(ab, ba)$ , где  $a$  и  $b$  обозначают строки с различным числом гласных. Порядок строк задается заранее и при необходимости может быть изменен. Построение четверостишья (блок II (14)) по описанному алгоритму осуществляется в блоке II. Результатом работы алгоритма являются следующие два примера.

**Пример 1.** Блестит на охоту  
Мороз и иней  
Живее сладкого  
Плоды дороги.

**Пример 2.** Ночью мороз не сдули  
Который мне дороги  
Мертвой дороги  
Безлюдный простор.

Порядок строк в первом примере —  $(ab, ab)$ , во втором —  $(aa, bb)$ . Вероятно, какое-либо из двух четверостиший понравится читателю. Значит, и при данной, столь несовершенной, работе системы по моделированию сочинения стихотворных произведений может быть получен результат, в какой-то мере удовлетворяющий художественную потребность читателя [6].

Дальнейшее совершенствование системы связано с таким понятием, как рифма. Рифма — не обязательное условие русского стиха. Рифма — созвучие двух слов, стоящих в определенном месте ритмического построения стихотворения. В русском стихе рифма должна находиться на конце стиха. Именно концевые созвучия, дающие связь между двумя стихами, именуются рифмой. Следовательно, у рифмы есть два качества: ритмическая организация, потому что она (рифма) отличает концы стихов, и созвучие. Рифма — явление звуковое, а не графическое, поэтому не важно, как пишутся два разных слова, а важно, чтобы они звучали одинаково [7]. В русском языке различают несколько типов рифм, один из них так называемая точная рифма — совпадение всех звуков в окончании [7]. В системе «Поэтесса» при выделении окончаний считается достаточным взять для сравнения четыре последних символа в последнем слове каждой строки. Сравнение окончаний производится в строках, содержащих одинаковое количество слогов. Так как рифма — это всякий звуковой повтор, несущий организующую функцию в метрической композиции стихотворения, то различают три типа рифм, которые встречаются в строфе из четырех стихотворных строк [7]:

смежные (АА, ВВ), перекрестные (АВ, АВ), опоясывающие (АВ, ВА), где А и В строки с различными окончаниями последних слов.

Очевидно, что организация внутренней структуры строфы и типы рифмы в строфе совпадают. Определенное количество слов в строке, величина и внутренняя структура строки, количество строк, рифмующиеся окончания придают стихотворной системе ритмичность. Результаты, полученные при работе системы с учетом точной рифмы (блок III), имеют уже большую близость с моделируемым объектом, чем те результаты, которые были приведены ранее:

Вдалеке я мечтала  
Сухими жесткими словами  
Осень солнцу отдала  
Легла пушистыми коврами.

Но в русском языке существуют еще так называемые вольности рифмы, основанные на несоответствии правописания и произношения [7]. Подобными по произношению считаются согласные Д — Т, Б — П, Ж — Ш, Г — Х. Буквы А — Я, У — Ю, О — Ё обозначают одинаковый гласный звук. Исходя из этого, сравнение окончаний при формировании четверостишья осуществляется способом, несколько отличающимся от описанного ранее (блок IV). Выбор способа подбора окончаний выполняется в режиме диалога. В зависимости от полученного ответа на свой вопрос, система «Поэтесса» выбирает тот или иной способ сравнения окончаний. Следуя по второму пути выбора рифмующихся окончаний, получаем:

Летучей лаской снегопад  
Сердце занято мечтами  
Пусть снежинки свой заводят  
Поздний вечер над прудами.

Академик А. Н. Колмогоров проанализировал соотношение запаса слов с рифмой. Оказалось, что запас в 100 слов дает тройную рифму, в 200 — четырехкратную, а в 500 слов — в изобилии дает поэтам десятикратные рифмы. Эти данные позволили определить минимальный словарный запас системы «Поэтесса».

Следующий этап работы по усовершенствованию системы — формирование человеку классифицировать литературные произведения по темам? По нашему мнению, это возможно благодаря определенному набору слов, наиболее характерных для той или иной темы. Система «Поэтесса» предусматривает генерацию четверостишья на определенную тему, которая задается набором тематических слов (блок V (11)). Ввод этих слов осуществляется в режиме диалога. Если в процессе работы какое-либо из тематических слов не было найдено в словарном запасе

МВОСАВ системы, то в режиме диалога возможна замена этого слова другим, относящимся к данной теме словом (блок V). После ввода тематических слов свою работу система осуществляет следующим образом: генерируется новая строка, содержащая одно из тематических слов (блок V (12)). Эта строка заносится в массив МРОЕМ (блок V (13)). Процесс продолжается до тех пор, пока не будут использованы все тематические слова. Дальнейшая работа системы аналогична описанной. Как пример приведем четверостишие, составленное системой «Поэтесса» на тему «Зима»:

Мороз и иней.  
Снег в любви моей  
Живее сугробов таял город  
Снежинки свой заводят хоровод.

По мере формирования строфы просто в режиме генератора либо в тематическом режиме осуществляется последовательный вывод их на печать. Далее, в режиме диалога возможна корректировка полученного текста при необходимости изменения какой-либо строки. В системе осуществляется подбор нескольких новых вариантов строк, согласующихся с изменяемой по числу гласных и окончанию (блок II (15)). После выбора одного из наиболее подходящих вариантов скорректированное четверостишие выдается на печать.

Полученные результаты нельзя считать окончательными. Дальнейшее совершенствование системы будет идти по пути еще большего усложнения внутренней структуры строфы, а именно: учет ударения, расстояния между ударными гласными, т. е. моделирование тонического стихосложения [7]. Это усложнение внутренней структуры позволит получать поэтические тексты любого размера — ямб, хорей, дактиль, амфибрахий, анапест. Огромные возможности улучшения работы системы «Поэтесса» открывает режим диалога. Он позволяет пополнить словарный запас системы, задать способ и размер стихосложения, контролировать систему непосредственно в процессе ее работы и при необходимости направить ее по пути, представляющему для нас наибольший интерес.

На оценку машинных результатов влияют многие факторы и среди них один из наиболее значительных — психологическая установка. Человек не может «объективно» оценивать стихи, написанные искусственной системой и сравнивать их с сочинениями поэтов, что вполне объяснимо психологически. Наибольшее различие возникает при оценке машинного творчества у представителей гуманитарного и инженерно-математического типов мышления.

Мы провели небольшой психологический эксперимент с участием учащихся студенческой группы специальности «Прикладная математика». Им было предложено два примера.

**Пример 1.** Я кровь ковал  
Я Вас любил.  
Любовь еще быть может...  
Не Вас, не к вам.

**Пример 2.** Под бурями судьбы жестокой  
Остались мне одни страдания  
Плоды сердечной пустоты  
Я разлюбил свои мечты.

Требовалось определить, какое из этих четверостиший написано человеком, а какое системой «Поэтесса».

В результате эксперимента 90% опрошенных отдали предпочтение второму примеру (система «Поэтесса»), а не первому (В. Сосюра [8]). Причем критерий оценки был один — «что хуже» с их точки зрения, то и написано ЭВМ. Нельзя, конечно, сказать, что это четверостишие лучше любого другого четверостишия, написанного человеком. Но ясно, что оно уже выиграло по сравнению с четверостишием среднего качества. Часто можно слышать, что машинные сочинения лишены смысла. Но что понимается под смыслом художественного произведения? Дать определение смысла, его психологическую оценку трудно. Анализируя художественное произведение, человек сопоставляет свой жизненный опыт, свои взгляды с описываемыми событиями. И если описываемым событиям человек не может найти объяснения, то он говорит, что это произведение бессмысленно. И наоборот, если предложить человеку какой-либо набор слов, то он подсознательно создает некоторый образ или ситуацию, подобную его жизненным идеалам, попытается найти какой-то смысл даже в бессмысленных на первый взгляд сочетаниях слов. Примером этого могут служить приведенные в [8] четверостишия.

При построении программного обеспечения системы «Поэтесса» использовался модульный принцип. Все модули написаны на языке Фортран-4 и Ассемблере и реализованы на ЭВМ ЕС-1050.

Программное обеспечение вместе с массивами занимает 240 К оперативной памяти ЭВМ.

**Список литературы:** 1. Кондратов А. М. Теория информации и поэтика (Энтропия ритма русской речи).— Проблемы кибернетики, 1963, вып. 9, с. 279—286. 2. Грекова И. Полемика и ее издержки (По поводу спора «Машина и творчество»).— Новый мир, 1973, № 7, с. 225—235. 3. Моль А., Фукс В., Касслер М. Искусство и ЭВМ.— М.: Мир, 1975.— 556 с. 4. Томашевский Б. В. Стилистика и стихосложение.— М.: Учпедгиз, Ленингр. отд-ние, 1959.— 535 с. 5. Шенгели Г. А. Техника стиха.— М.: Госполитиздат, 1960.— 312 с. 6. Лобанцев Ю. Л., Соколов В. В. О влиянии художественной установки на восприятие поэзии.— В кн.: Эстетику в жизнь.— Свердловск, 1973, № 4, с. 45—50. 7. Жирмунский В. М. Теория стиха.— Л.: Сов. писатель, 1975.— 664 с. 8. Соложенкина С. У забавы кофта — ситчик.— Правда, 1979, 3 февраля, с. 3.

Поступила 22 февраля 1979 г.