

УДК 519.766.44

ЗАКОНОМЕРНОСТИ ВАРЬИРОВАНИЯ В ТЕКСТАХ РАЗЛИЧНОЙ
ПРИРОДЫ И МЕТОДИКА ИХ КОЛИЧЕСТВЕННОГО ИССЛЕДОВАНИЯ

И.В.Бахмутова, В.Д.Гусев, Т.Н.Титкова

В в е д е н и е

Многие тексты (естественноязыковые, музыкальные, генетические) образуются по следующей схеме. Имеется один или множество исходных фрагментов, повторением, варьированием и объединением которых в различном порядке формируются более сложные объекты с новыми свойствами. В свою очередь к полученным объектам вновь может быть применена указанная схема, что приводит к появлению иерархически структурированных текстов. Аппарат анализа повторов, описанный в [1], предоставляет широкие возможности для выявления структурной организации таких текстов.

В данной работе нас будет интересовать не столько выделение множества исходных элементов (фрагментов) на каждом уровне иерархии (задача сегментации текстов), сколько исследование принципов варьирования этих элементов и получение количественных характеристик варьирования. Это важно как для конструирования новых текстов (например, ДНК - молекул с заданными свойствами), так и для их идентификации, когда способ варьирования рассматривается как классификационная характеристика. Поскольку варьирование является одним из важнейших и универсальнейших компонентов творческой деятельности [2], методика количественной оценки (а в конечном счете и формализации) этого компонента выходит за рамки проблемно-ориентированного подхода и представляет интерес для широкого круга задач из области искусственного интеллекта.

На содержательном уровне под варьированием понимается повторение какой-либо ситуации в видоизмененной форме. Степень изменения исходной ситуации должна быть относительно невелика, чтобы термин "повторение" не утратил своего смысла. Хотя варьирование

неразрывно связано с понятием повтора, оно не является обязательным его атрибутом (возможен повтор "в чистом виде" - без искажений). Тем не менее количественные характеристики варьирования, полученные нами для одной из предметных областей, и многочисленные примеры - аналоги из других областей указывают на то, что при построении определенных классов текстов повторение чаще сопровождается варьированием, нежели отсутствием такового. Это дает основание утверждать, что принцип варьирования играет столь же фундаментальную роль, что и принцип повторности.

Объяснение этому факту можно искать в многоплановости той роли, которую играет варьирование. Отметим основные эффекты, достигаемые с его помощью.

1) Варьирование можно рассматривать как весьма универсальный способ "обогащения" текстов, играющий важную роль при непосредственном восприятии этих текстов (речь, музыка) человеком. В первых экспериментах по синтезу речи, к примеру, неестественность звучания ("механический голос") в значительной степени была обусловлена тем, что частота основного тона оставалась неизменной (не варьировалась). Изменение частоты (даже по случайному закону) повышало естественность звучания.

2) В подавляющем большинстве случаев, однако, варьирование осуществляется не случайным образом, а целенаправленно. В этом плане оно часто выступает как средство развития сюжетной линии (характерный пример - стихотворение "Вот дом, который построил Джек" в переводе С.Маршака).

3) Синтезируемые тексты часто должны удовлетворять не одному, а нескольким критериям одновременно. Варьирование с использованием синонимичных замен является одним из средств решения подобных задач многокритериальной оптимизации (характерный пример - синтез гена с заданными свойствами; класс синонимичных замен представлен кодонами, кодирующими одну и ту же аминокислоту).

4) Предыдущий пункт в определенных ситуациях можно трактовать как введение многозначности с помощью синонимичных замен. Посредством варьирования может быть достигнут и обратный эффект: устранение возможной неоднозначности путем использования уточняющих синонимичных повторов. Эту схему хорошо иллюстрирует фраза, принадлежащая известному физику Саймону Б.: "Сперва вы объясните,

что вы собираетесь сказать, потом говорите, потом объясняете, что вы сказали" *).

Несмотря на многообразие предметных областей, в которых встречаются текстовые данные, общезыковые закономерности просматриваются в них весьма отчетливо. В значительной степени это касается и способов варьирования, которые сходны для текстов разной природы. В разделе I перечислены основные способы варьирования и приведены примеры из различных областей. В разделе 2 введено понятие несовершенного повтора и описаны конструкции, фиксирующие определенные типы несовершенных повторов. Эти конструкции, с нашей точки зрения, наиболее адекватно отражают семантику термина "варьирование". В последующих разделах иллюстрируется возможность получения с помощью аппарата несовершенных повторов количественных характеристик варьирования. Объектом исследования выбраны музыкальные тексты (мелодии песен). Фиксируется и способ варьирования - замены. Полученные данные не имеют аналогов и могут быть использованы в программах, имитирующих деятельность композитора.

I. Способы и примеры варьирования в текстах различной природы.

Из всего многообразия приемов варьирования отметим лишь наиболее типичные: а) замену одного символа другим (иногда группой других); б) вставку или устранение символа; в) перестановку соседних символов; г) блочные перестройки (перенос фрагмента из одного места в другое; повторение элементов фрагмента в обратном порядке, перестановку фрагментов и т.д.). Часто имеет место комбинация различных приемов. Приведем примеры варьирования с использованием указанных приемов.

ПРИМЕР I. Из поэмы А.Твардовского "По праву памяти":

И даром думают, что память
Не дорожит сама собой,
Что ярской времени затянет
Любу^а был^ь
Любая боль.

Замены "любу^а - любая", "был^ь - боль" относятся к типу "а". Но если в первом случае смысл слова не меняется (замена обусловлена необходимостью согласования), то во втором случае меняется и семантика.

*) Саймон Б. Модель $R(\varphi)_2$ евклидовой квантовой теории поля. -М.: Мир, 1976.

ПРИМЕР 2. Фрагменты из вариантов мелодии песни "Вниз по ма - тушке по Волге" (№ 1 - вариант Праца (конец 18 века); №2 - вари - ант Кашина (1834 г.)):



В первом фрагменте (см. участок, выделенный надчеркиванием) высота звука изменяется скачком на 4 ступени вниз. Во втором фрагменте скачок заменяется поступенным движением к тому же звуку (так называемое "заполнение" интервала). Формально в используемой нами кодировке нотного текста (см. п.3) это выразится в замене одного элемента четырьмя с сохранением баланса по величине скачка. Подобная замена одного символа группой часто встречается, когда элементы алфавита представлены в более богатой шкале, нежели шкала наименований.

Если мы имеем дело со структурированным текстом (разбитым, например, на слова), то замену одной структурной единицы другой также можно отнести к типу "а", но уже на другом уровне иерархии.

ПРИМЕР 3 (из сказки "Сестрица Аленушка и братец Иванушка"):

"Стоит коровье копытце полно водицы"

"Стоит лошадиное копытце полно водицы"

"Стоит козье копытце полно водицы"

Повторы такого рода типичны для народных сказок. Варьирование слухит в них средством развития сюжета. Схема варьирования тем примитивнее, чем моложе слушатель, на которого она рассчитана. Вот более сложный пример (с заменами и вставками) из более "взрослой" сказки "Царевна-лягушка": "Пустил стрелу старший брат - упала она на боярский двор, ... Пустил средний брат - полетела к купцу на двор, ... Пустил младший брат - попала стрела в грязное боло - то..."

ПРИМЕР 4. Фраза, иллюстрирующая способ варьирования, относящийся к типу "в": "Такие системы могут быть построены из процес -

соров любой степени сложности – от простейших однозарядных, с минимально необходимым набором логических операций, до мощных процессоров, способных выполнять операции над многозарядными числами" (Вестник АН СССР, №11, 1985 г., с.64).

Вместо подчеркнутого слова должно стоять "однозарядных". Зацепкой для редактора мог бы послужить конец фразы ("многозарядными числами"). Перестановки соседних букв, допускаемые машинисткой при печати, – пример неосознанного варьирования, обусловленного зачастую "провоцирующим" использованием слов в тексте.

Варьирование путем перестановки соседних символов часто встречается и в композиторской практике (соответствующие примеры приведены в п.6 раздела 6). Варьирование в таком случае носит осознанный характер.

ПРИМЕР 5. Эффект, достигаемый блочными перестройками, проиллюстрируем с помощью мелодий песен "Азовская партизанская" (В.Соловьев-Седой) и "На страже Родины" (С.Туликов). Эти мелодии имеют общие интонационные обороты X и Y, которые являются характерными, определяющими лицо мелодии. Используя кодировку нотного текста, описанную в п.3, их можно представить в виде семиэлементных фрагментов:

$$X = (I\text{---})(I\text{+})(I\text{---})(3\text{+})(3\text{+})(I\text{+})(I\text{+}),$$

$$Y = (2\text{+})(5\text{+})(I\text{+})(I\text{---})(I\text{+})(I\text{---})(I\text{+}).$$

В первой из мелодий фрагменты расположены в следующем порядке: ... X' ... X ... Y ... Y ... (здесь "треточие" означает вставку группы элементов, а "штрих" указывает на варьирование путем одних – ночных замен и вставок). Во второй мелодии порядок фрагментов обратный: Y Y' ... X' ... X Тиражирование фрагментов X и Y в разном количестве, варьирование их, изменение порядка следования и групповые вставки делают мелодии неузнаваемыми, хотя обе построены из довольно крупных близких блоков.

ПРИМЕР 6. На рис.1 представлены близкие в функциональном отношении фрагменты из геномов родственных бактериофагов (φX174 и φ4). Эти фрагменты характеризуют весьма сложный в функциональном отношении участок обоих геномов. Здесь заканчиваются гены E и D терминальными кодонами TGA и TAA соответственно и начинается (с иницирующего кодона ATG) ген I, которому предшествует рибосомный сайт связывания (характерная комбинация GGAG, расположенная на II позиции левее иницирующего кодона).

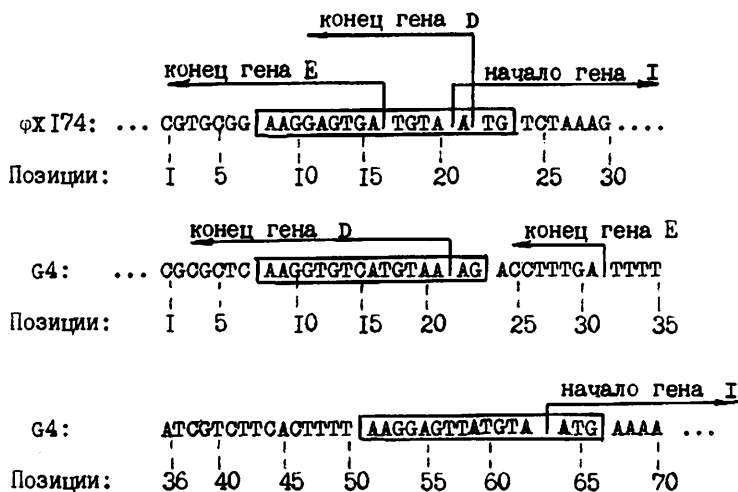


Рис. I

Представляет интерес проследить за эволюционными перестройками, которые обусловили различие фрагментов, и оценить их последствия. Отметим следующие особенности:

1) фрагмент генома G4 значительно длиннее за счет повторения (с заменами) участка из 16 нуклеотидов, заключенного в рамку, и вставки 27 нуклеотидов (позиции с 24-й по 50-ю). Указанные блочные перестройки обусловили появление некодирующего участка в G4 (позиции с 32-й по 63-ю), которого нет у ФХ I74;

2) ген E в геноме G4 удлиннен по сравнению с геном E из ФХ I74.

Причиной этого явилась одиночная замена G на C в 15-й позиции, в результате чего терминальный кодон TGA перешел в кодон TCA, кодирующий аминокислоту;

3) замена T (22-я позиция в геноме ФХ I74) на A у G4 (в той же позиции) разрушила иницирующий кодон ATG, с которого начинался ген I. Заметим, что и в рибосомном сайте связывания у G4 в 12-й позиции вместо A появилось T, что сильно ослабило функциональные возможности этого участка. Именно благодаря дупликации 16 нуклеотидов были "воссозданы" условия для инициации гена I (64-я позиция). Длина дупликации как раз соответствует размеру комплекса "рибосомный сайт связывания - иницирующий кодон".

Рассмотренный пример примечателен в следующем отношении. Участок генома $\phi X174$, заключенный в рамку, несет многофункцию — начальную нагрузку: кодирует гены E и D в разных фазах считывания, содержит два терминальных кодона (знаки окончания трансляции), один иницирующий (знак начала трансляции) и рибосомный сайт связывания. Всего лишь три замены (см. первый из выделенных участков G4) лишают его половины функций: исчезает один терминальный и иницирующий кодоны, проблематичным становится существование рибосомного сайта связывания. Дубликация со вставкой берет на себя часть функций с перегруженного в регуляторном плане участка.

Этот пример носит комплексный характер и иллюстрирует почти все виды варьирования ("а", "б", "г").

ПРИМЕР 7. Рассмотрим последние четыре строки стихотворения Антуана Арно "Листок" в разных переводах:

Стремлюсь, куда велит мне рок,
Куда на свете все стремится,
Куда и лист лавровый мчится
И легкий розовый листок. (В.Жуковский)

Несусь, куда несет суровый,
Всему неизбежный рок,
Куда летит и лист лавровый
И легкий розовый листок! (Д.Давыдов)

Я мчусь по прихоти суровой,
Куда влечет меня мой рок,
Куда несется лист лавровый
И легкий розовый листок. (С.Дуров)

Мчусь, куда уносят грезы,
В этой жизни все в свой срок:
Лепесток засохшей розы,
Как и лавровый листок. (В.Брюсов)

Приведенные переводы показывают, что одинаковые по смыслу тексты, рассматриваемые просто как последовательности элементов алфавита (нижний уровень иерархии), в принципе, могут отличаться сколь угодно сильно. Тем не менее уже на этом уровне иерархии часто обнаруживаются повторы слов (куда, рок, все (всему), суровый), словосочетаний ("и легкий розовый листок"), перестановки (лист лавровый — лавровый листок), вставки и замены. Сходство усиливает

ется при переходе к более высоким уровням иерархии, например, при замене слов или словосочетаний номерами классов эквивалентности, объединяющих синонимичные выражения (стремлюсь - несусь - мчусь; велит - влечет; лепесток - листок). Исследование закономерностей варьирования на самых высоких уровнях иерархии тесно связано с проблемой автоматического понимания текстов.

Во всех приведенных примерах у варьируемых текстов сохранялось нечто общее, что можно назвать инвариантом преобразования. Этот инвариант не всегда очевиден (о чем свидетельствует пример 7), и выявление его представляет определенные трудности. В самом общем случае таким инвариантом выступает смысл сообщения, кодируемого текстом.

2. Представление текстов в терминах "повторов".

В основе нашего подхода к анализу текстов лежит понятие "повтора". Для классификации повторов используем два независимых признака: а) степень близости фрагментов, образующих повтор; б) степень их "разнесенности" по тексту или группе текстов. В соответствии с первым признаком будем различать повторы совершенные и несовершенные, а со вторым - повторы первого и второго рода. Заметим, что каждый из признаков допускает большее (чем два) число градаций и соответственно возможна более детальная классификация, однако степень детализации, принятая в данной работе, достаточна для иллюстрации подхода.

Совершенным повтором назовем два идентичных фрагмента текста, расположенных относительно друг друга произвольным образом. Если число совпадающих фрагментов F больше двух, повтор называем совершенным кратности F . Совершенный повтор кратности F можно трактовать как совокупность из $F(F-1)/2$ совершенных повторов.

Несовершенным повтором назовем пару фрагментов текста, близких в определенном смысле (например, в смысле хэминггова расстояния, редакционного и т.д.). Если отношение близости характеризуется свойством ассоциативности, можно ввести понятие несовершенного повтора кратности $F (F > 2)$. К примеру, если близкими считать все фрагменты длины 1, отличающиеся друг от друга не более чем по k заранее фиксированным позициям ($1 \leq i_1 < i_2 < \dots < i_k \leq 1$), понятие несовершенного повтора кратности F имеет следующий содержательный смысл: оно характеризует ситуации, когда в варьируемых фрагментах существуют "неустойчивые" позиции, наиболее часто за -

трагиваемые варьированием, и относительно "стабильные", несущие основную функциональную нагрузку.

Если фрагменты принадлежат одному тексту, будем говорить о повторе первого рода, в противном случае (фрагменты в разных текстах) - о повторе второго рода.

В соответствии с этим делением будем различать четыре типа повторов: совершенные первого рода, совершенные второго рода, несовершенные первого рода и несовершенные второго рода.

Для исследования количественных характеристик варьирования нам понадобятся лишь конструкции, основанные на понятии несовершенного повтора. Более того, в данной работе мы фиксируем свое внимание только на классе повторов, где в качестве меры близости используется хэммингово расстояние. Это соответствует способу варьирования, допускающему лишь замены символов в произвольных позициях фрагмента.

Основанием для нашего выбора послужили следующие соображения:

а) по частоте использования (встречаемости) замены стоят на первом месте во многих приложениях (мутационные перестройки в ДНК-молекулах, орфографические ошибки в текстах на естественном языке и т.д.);

б) некоторые другие типы варьирования, например, перестановки соседних символов, формально сводятся к заменам рядом расположенных элементов;

в) алгоритмические трудности, связанные с организацией направленного перебора при получении количественных оценок, преодолеваются для данного типа варьирования достаточно просто.

2.1. Назовем $(1, k)$ -повтором текста T (или $(1, k)$ -повтором первого рода) пару произвольных фрагментов длины l из T , отличающихся друг от друга по k позициям ($k < l$). Совокупность всех возможных $(1, k)$ -повторов, содержащихся в тексте T , назовем $(1, k)$ -характеристикой текста и обозначим $\Phi_{1, k}(T)$.

ПРИМЕР 8. Пусть $T = abcabda$. Фрагмент текста T , расположенный в позициях с i -й по j -ю включительно, будем обозначать $T[i:j]$. Тогда

$$\Phi_{3, 1}(T) = \left\{ \left(T[1:3] \right); \left(T[2:4] \right) \right\} = \left\{ \left(\begin{array}{ccc} a & b & c \\ \parallel & \parallel & \\ a & b & d \end{array} \right); \left(\begin{array}{ccc} b & c & a \\ \parallel & \parallel & \\ b & d & a \end{array} \right) \right\};$$

$$\Phi_{4, 1}(T) = \left\{ \left(T[1:4] \right) \right\} = \left\{ \left(\begin{array}{cccc} a & b & c & a \\ \parallel & \parallel & & \parallel \\ a & b & d & a \end{array} \right) \right\}.$$

Фрагменты, составляющие $(1, k)$ -повтор, расположены по вертикали, а совпадающие элементы помечены знаком "||".

Параметр k характеризует расстояние между фрагментами, образующими повтор. Его величина должна быть мала по сравнению с l , иначе речь будет идти не о варьировании уже имеющегося фрагмента, а о создании нового. Если параметры l и k подобраны правильно, количество обнаруживаемых в тексте $(1, k)$ -повторов невелико и они, как правило, являются функционально значимыми. В худшем случае число $(1, k)$ -повторов в тексте может иметь порядок N^2 , где N - длина текста.

Обнаруженный $(1, k)$ -повтор часто допускает расширение: его длину можно увеличить до значения $l' > l$ при сохранении числа несовпадений k . Процедура расширения является аналогом перехода от отдельной $(1, k)$ -характеристики к спектру таких характеристик и обеспечивает компактность представления результатов.

2.2. Переход от несовершенных повторов первого рода к несовершенным повторам второго рода делается по следующей схеме. Назовем $(1, k)$ -повтором группы текстов $\bar{T} = (T_1, T_2, \dots, T_m)$ (или $(1, k)$ -повтором второго рода) пару фрагментов длины l , разнесенных по разным текстам группы и отличающихся друг от друга по k позициям. Совокупность всевозможных $(1, k)$ -повторов второго рода, содержащихся в группе текстов \bar{T} , назовем $(1, k)$ -характеристикой группы текстов и обозначим $\Phi_{k,1}(\bar{T})$.

ПРИМЕР 8. Пусть $\bar{T} = (T_1, T_2, T_3)$, где $T_1 = \text{abcabda}$, $T_2 = \text{dbcscb}$, $T_3 = \text{bdcscb}$. Тогда

$$\Phi_{k,1}(\bar{T}) = \left\{ \begin{pmatrix} T_1[2:5] \\ T_2[2:5] \end{pmatrix}, \begin{pmatrix} T_2[2:5] \\ T_3[1:4] \end{pmatrix} \right\} = \left\{ \begin{pmatrix} b & c & a & b \\ || & || & || & || \\ b & c & c & b \end{pmatrix}, \begin{pmatrix} b & c & c & b \\ || & || & || & || \\ b & d & c & b \end{pmatrix} \right\}.$$

Заметим, что пара фрагментов $T_1[1:4]$ и $T_1[4:7]$, также образующая $(4, 1)$ -повтор, не входит в $\Phi_{k,1}(\bar{T})$, поскольку оба фрагмента принадлежат одному тексту. Аналогично предыдущей конструкции осуществляется расширение фрагментов в ситуациях, когда оно не влечет за собой увеличения числа несовпадений. Каждый повтор сопровождается информацией о номерах текстов, содержащих фрагменты, и местах вхождения фрагментов в текст.

2.3. Анализ используемых в реальных текстах схем варьирования требует обработки достаточно представительных текстовых массивов. Проблема организации направленного перебора выходит здесь на первый план.

Прямой алгоритм вычисления $\Phi_{1,k}(T)$ путем попарного сопоставления всех фрагментов длины 1 из T имеет трудоемкость $O(N^2 \cdot l)$, где N - длина текста T . Рекуррентная схема вычисления со "скачком" [3] позволяет избежать сопоставления части фрагментов, которые заведомо не могут образовать $(1,k)$ -повтор. Это снижает трудоемкость до $O(N^2/c(1,k))$, где величина делителя всегда больше единицы, а с уменьшением отношения k/l может стать много больше единицы.

Алгоритм вычисления $\Phi_{1,k}(T)$ является модификацией алгоритма из [3]. Сохраняется и приведенная выше оценка трудоемкости с той лишь разницей, что под N следует понимать сумму длин текстов T_1, T_2, \dots, T_m . Программы вычисления $\Phi_{1,k}(T)$ и $\Phi_{1,k}(T)$ реализованы в рамках модифицированной версии пакета прикладных программ СИМВОЛ [4].

2.4. Под количественными характеристиками рассматриваемой нами схемы варьирования будем понимать классификацию типов замен, их относительные частоты, распределение замен по длине фрагментов, образующих несовершенный повтор, специфичность (способность характеризовать тот или иной класс текстов) и ряд других. Эти данные необходимы, к примеру, для оценки функциональной значимости тех или иных участков текста (консервативные участки более значимы), прогнозирования развития эволюционного процесса, моделирования новых текстов с помощью ЭВМ. Методику получения количественных характеристик варьирования и характер выявляемых закономерностей проиллюстрируем на материале песенных мелодий.

3. Кодирование нотного текста. Пусть $T = z_1 z_2 \dots z_N$ - музыкальный текст, состоящий из последовательности N нот. Следуя Р.Х.Зарипову [5], для представления мелодий будем использовать интервально-метрическую структуру. Определим ее как последовательность пар $I_k S_k$, $k = 1, 2, \dots, (N-1)$, где I_k - количество ступеней между высотами k -го и $(k+1)$ -го звуков мелодии, а S_k - метрическая характеристика, дающая представление об относительной силе этих звуков. Формально $I_k = W_{k+1} \ominus W_k$, где W_k - высота k -го звука, а " \ominus " - символ специальной операции вычитания;

$$S_k = \begin{cases} + & \text{при } p_{k+1} > p_k, \\ - & \text{при } p_{k+1} \leq p_k, \end{cases}$$

где p_k - относительная сила доли такта, соответствующей k -му зву-

ку. Подчеркнем, что значение стопы $S_k = (-)$ характеризует три ситуации:

- а) переход от слабой доли к сильной при $P_{k+1} < P_k$,
 - б) переход от сильной доли к сильной
 - в) переход от слабой доли к слабой
- } при $P_{k+1} = P_k$.

Приписывание знака "-" стопе S_k в случаях "б" и "в" можно мотивировать тем, что каждый звук имеет определенную продолжительность, фиксируемую нотным знаком, и сила его звучания постепенно ослабевает, что напоминает ситуацию, с которой мы имеем дело в случае "а".

При перекодировке мелодии из нотной записи в (IS)-представление принимаем следующее соглашение: сначала указываем значение интервала (т.е. величину $|I|$), затем его знак ("+" соответствует движению звуковысотной линии вверх, "-" - вниз), потом S_k - знак стопы ("+" соответствует переходу от сильной доли к слабой, "-" - наоборот). Если $|I| = 0$, знак интервала формально не определен, но для единообразия записи ставится "+". Код (2+-), например, трактуется как скачок на 2 ступени вверх с одновременным увеличением силы звука, а код (0++) - как сохранение высоты звука при уменьшении его силы. Рис.2 иллюстрирует способ получения (IS)-представления из нотной записи на примере отрывка из мелодии "Комсомольской песни" (музыка В.Соловьева-Седого).

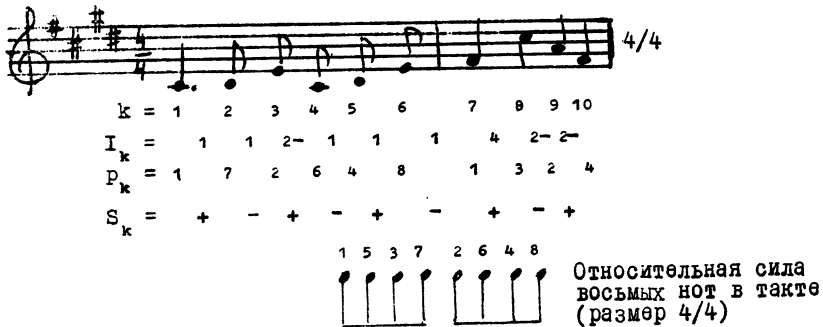


Рис. 2

(IS)-характеристика приведенного музыкального фрагмента имеет вид (I++I+-2-+I+-I++I+-4++2-2-+).

Используемое описание обеспечивает, на наш взгляд, желательный компромисс между двумя противоречивыми требованиями. С одной стороны, оно достаточно полно, так что мелодия не утрачивает своей индивидуальности, остается узнаваемой. С другой стороны, оно не слишком детально за счет устранения малосущественных (для многих приложений) факторов, маскирующих наличие элемента повторности в мелодии. В частности, такой вид музыкального варьирования, как секвентное повторение (повтор мелодического оборота на другой высоте с сохранением ритма и внутренних интервальных соотношений), полностью нивелируется в данной системе описания.

4. Классификация замен в музыкальных текстах. Если элементы трехзначного кода $\{|I|, \text{знак } I, S\}$ обозначить соответственно цифрами 1, 2, 3, то можно выделить 8 типов замен, нумерация которых указывает, какие элементы кода варьируют в данном типе. В заменах типа 1 изменяется лишь величина интервала - первый элемент кода. Поскольку при $|I| = 0$ знак интервала формально не определен, целесообразно из типа 1 выделить подтип 0, включив в него все те случаи, когда ненулевое значение величины интервала (с любым знаком!) заменяется нулевым (или наоборот). К типу 1 будем тогда относить лишь замены с $|I| \neq 0$ (ненулевое значение интервала заменяется ненулевым).

Кроме рассмотренных двух типов, выделим: тип 2 - изменяется лишь второй элемент кода (знак I); тип 3 - изменяется лишь третий элемент кода (знак стопы S); тип 1, 2 - изменяются первый и второй элементы кода одновременно, но $|I| \neq 0$ (соответствующие замены отнесены к типу 0); тип 1, 3 - изменяются первый и третий элементы кода при сохранении знака интервала (сюда же относим случаи, когда знак одного из интервалов не определен, например, (3+) заменяется на (0+)); тип 2, 3 - изменяются второй и третий элементы кода; тип 1, 2, 3 - изменяются все три элемента кода.

5. Описание эксперимента. Экспериментальный материал включал в себя около 600 мелодий (песни советских композиторов [7], русские народные песни [8], песни Шуберта [9]), из которых формировались однородные в определенном смысле подборки по разным принципам: по композитору (Шуберт, Дунаевский, Блантер и т.д.), по времени создания (песни советских композиторов 50-х годов, русские народные песни), по тематике (песни Победы [10]). Объем закодированного и обработанного материала составил примерно $5 \cdot 10^4$ нот.

При анализе несовершенных повторов первого рода использовались (7,1)- и (9,2)-характеристики каждой мелодии; при анализе повторов второго рода - (9,1)- и (11,2)-характеристики групп мелодий (с расширениями в обоих случаях). Выбор указанных значений параметров l и k в большинстве случаев обеспечивает нам "неслучайность" обнаруживаемых закономерностей (повторов). В свою очередь это означает, что фигурирующие в них замены можно рассматривать как результат осознанного варьирования (в случае повторов второго рода речь идет о варьировании типовых (бытующих) интонаций, либо оригинальных (заимствованных) [6]). Анализ сводился к подсчету числа замен каждого из 8 типов в каждой из групп мелодий. Если какая-либо конкретная замена встречалась внутри мелодии многократно, она учитывалась всего один раз. Аналогично поступали и в отношении повторов второго рода. Замены, приходящиеся на края сравниваемых фрагментов, не принимались во внимание, т.е. учитывались только "внутренние" искажения. Наряду с подсчетом числа и разновидностей замен, анализировались места их расположения внутри сравниваемых фрагментов. Из краткой характеристики проведенных экспериментов (табл.1) видно, что суммарное число замен, зафиксированное в каждом эксперименте, весьма велико (даже без учета кратности использования замен в мелодиях) и коррелирует с объемом текста. Одна и та же замена может встречаться в разных мелодиях. В случае повторов первого рода отношение суммарного числа замен к числу разновидностей характеризует среднюю (абсолютную) частоту использования замен в анализируемом массиве мелодий. В русских народных песнях [8], к примеру, каждая разновидность, в среднем, встречается в двух мелодиях, в песнях советских композиторов [7] - в трех, у Шуберта [9] - в четырех.

Фактические частоты использования той или иной замены могут сильно отличаться от средних. В табл.2 приведены наиболее характерные (часто используемые) разновидности замен по каждому из 8 типов. Конкретная замена представляется в виде пары трехзначных кодовых комбинаций, расположенных одна под другой и заключенных в круглые скобки. Такая запись интерпретируется следующим образом: в рассматриваемой группе музыкальных текстов существует мелодия, содержащая (1, k)-повтор с указанной парой кодов в одной из тех позиций, где зафиксировано несовпадение (один из кодов заменяется другим). Число над скобками обозначает, сколько раз встретилась такая замена.

Сводная таблица по описанию экспериментов (ш - число мелодий, N - объем текста)

№ экс- перим.	Классификация экспери- мента	Суммар- ное чи- сло за- мен	Код-во разно- видно- стей	Распределение замен по типам (%)									
				0	1	2	3	1,2	1,3	2,3	1,2,3		
1.	(7,1) - и (9,2) - повторы 1-го рода в русских народных пес- нях; ш = 172; N = 7402.	126	62	21,4	29,4	19,8	4	19	1,6	2,4	2,4	2,4	
2.	(9,1) - и (11,2) - повторы 2-го рода в русских народных пес- нях; ш = 172; N = 7402.	97	51	33,1	19,6	12,3	6,2	14,4	11,3	3,1	0	0	
3.	(7,1) - и (9,2) - повторы 1-го рода в песнях советских ком- позиторов; ш = 216; N = 16905.	488	145	21,3	36,1	15,4	5,7	16,2	3,9	1	0,4	0,4	
4.	(9,1) - и (11,2) - повторы 2-го рода в песнях советских ком- позиторов; ш = 216; N = 16905.	306	117	23,8	29,7	15,7	3,8	16,6	8,8	0,3	1,3	1,3	
5.	(7,1) - и (9,2) - повторы 1-го рода в песнях Шуберта (т. I, 2, 6); ш = 122; N = 24084.	732	173	31,5	27,6	9,3	8,6	11,2	8,7	1,1	2	2	
6.	(10,1) - (11,2) - и (15,3) - повторы 2-го рода в песнях Шуберта; ш = 122; N = 24084.	331	90	43,2	15,4	7	5,8	11,2	12,3	3,6	1,5	1,5	

Т а б л и ц а 2

Фрагмент сводной (для трех массивов) таблицы замен,
упорядоченных по убыванию частоты их встречаемости
(эксперименты I,3,5)

	29	24	22	20	19	17	13	
Тип С	Песни Шуберта	$\begin{smallmatrix} (1+-) \\ (0+-) \end{smallmatrix}$	$\begin{smallmatrix} (0++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} (0++) \\ (1++) \end{smallmatrix}$	$\begin{smallmatrix} (1--) \\ (0+-) \end{smallmatrix}$	$\begin{smallmatrix} (0+-) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} (0++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} (0+-) \\ (2--) \end{smallmatrix}$, ...
	Песни сов. композит.	$\begin{smallmatrix} 16 \\ (0++) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 13 \\ (0++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 13 \\ (0++) \\ (1--) \end{smallmatrix}$	$\begin{smallmatrix} 12 \\ (0++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 9 \\ (0+-) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (0+-) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 7 \\ (0++) \\ (1++) \end{smallmatrix}$, ...
	Русск. нар. песни	$\begin{smallmatrix} 4 \\ (0++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (0++) \\ (1++) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2++) \\ (0++) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2+-) \\ (0+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (0++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (0+-) \\ (1--) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (4++) \\ (0++) \end{smallmatrix}$, ...
Тип I	Песни Шуберта	$\begin{smallmatrix} 22 \\ (1+-) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 19 \\ (2+-) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 14 \\ (1+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 14 \\ (2+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 10 \\ (1--) \\ (2--) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (2+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (1++) \\ (2++) \end{smallmatrix}$, ...
	Песни сов. композит.	$\begin{smallmatrix} 12 \\ (1+-) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 10 \\ (1+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 10 \\ (3+-) \\ (5+-) \end{smallmatrix}$	$\begin{smallmatrix} 9 \\ (1+-) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (1+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (1--) \\ (3--) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (2+-) \\ (4+-) \end{smallmatrix}$, ...
	Русск. нар. песни	$\begin{smallmatrix} 6 \\ (1--) \\ (3--) \end{smallmatrix}$	$\begin{smallmatrix} 6 \\ (1+-) \\ (3+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (3+-) \\ (5+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (2+-) \\ (4+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (1+-) \\ (4+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (2+-) \\ (5+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (1++) \\ (2++) \end{smallmatrix}$, ...
Тип 2	Песни Шуберта	$\begin{smallmatrix} 26 \\ (1+-) \\ (1--) \end{smallmatrix}$	$\begin{smallmatrix} 26 \\ (1++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (2++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2+-) \\ (2--) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (3--) \\ (3+-) \end{smallmatrix}$, ...		
	Песни сов. композит.	$\begin{smallmatrix} 35 \\ (1--) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 29 \\ (1++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 6 \\ (2++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (3--) \\ (3--) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (2--) \\ (2+-) \end{smallmatrix}$ - других нет		
	Русск. нар. песни	$\begin{smallmatrix} 13 \\ (1--) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 8 \\ (1++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2++) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 1 \\ (3+-) \\ (3--) \end{smallmatrix}$ - других нет			
Тип 3	Песни Шуберта	$\begin{smallmatrix} 19 \\ (0++) \\ (0+-) \end{smallmatrix}$	$\begin{smallmatrix} 17 \\ (1+-) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 10 \\ (1++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 5 \\ (2--) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 4 \\ (2++) \\ (2+-) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (7++) \\ (7++) \end{smallmatrix}$, ...	
	Песни сов. композит.	$\begin{smallmatrix} 6 \\ (0++) \\ (0+-) \end{smallmatrix}$	$\begin{smallmatrix} 6 \\ (1--) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (1++) \\ (1+-) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2+-) \\ (2--) \end{smallmatrix}$	$\begin{smallmatrix} 3 \\ (2+-) \\ (2++) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (4+-) \\ (4--) \end{smallmatrix}$, ...	
	Русск. нар. песни	$\begin{smallmatrix} 3 \\ (1+-) \\ (1--) \end{smallmatrix}$	$\begin{smallmatrix} 2 \\ (0++) \\ (0+-) \end{smallmatrix}$ - других нет					

Тип 1,2	Песни Шуберта	$\overset{10}{(I\rightarrow)}$, $\overset{10}{(I\leftarrow)}$, $\overset{7}{(I\rightarrow)}$, $\overset{6}{(I\leftarrow)}$, $\overset{5}{(I\rightarrow)}$, $\overset{4}{(I\leftarrow)}$, ...
	Песни сов. композит.	$\overset{8}{(I\rightarrow)}$, $\overset{7}{(I\rightarrow)}$, $\overset{7}{(I\leftarrow)}$, $\overset{7}{(I\leftarrow)}$, $\overset{6}{(I\rightarrow)}$, $\overset{6}{(I\leftarrow)}$, ...
	Русск. нар. песни	$\overset{4}{(I\leftarrow)}$, $\overset{3}{(I\rightarrow)}$, $\overset{3}{(I\rightarrow)}$, $\overset{3}{(I\leftarrow)}$, $\overset{2}{(I\rightarrow)}$, $\overset{2}{(I\rightarrow)}$, ...
Тип 1,3	Песни Шуберта	$\overset{8}{(0\rightarrow)}$, $\overset{6}{(0\rightarrow)}$, $\overset{5}{(2\rightarrow)}$, $\overset{5}{(0\rightarrow)}$, $\overset{4}{(2\rightarrow)}$, ...
	Песни сов. композит.	18 замен (все по разу)
	Русск. нар. песни	$\overset{1}{(I\rightarrow)}$, $\overset{1}{(0\rightarrow)}$ - других нет
Тип 2,3	Песни Шуберта	$\overset{6}{(I\rightarrow)}$, $\overset{2}{(I\rightarrow)}$ - других нет
	Песни сов. композит.	$\overset{5}{(I\rightarrow)}$ - единственный вариант
	Русск. нар. песни	$\overset{2}{(I\rightarrow)}$, $\overset{1}{(I\rightarrow)}$ - других нет
Тип 1,2,3	Песни Шуберта	$\overset{4}{(2\rightarrow)}$, $\overset{3}{(I\rightarrow)}$, $\overset{2}{(2\rightarrow)}$, $\overset{1}{(I\rightarrow)}$, ...
	Песни сов. композит.	$\overset{1}{(2\rightarrow)}$, $\overset{1}{(I\rightarrow)}$ - других нет
	Русск. нар. песни	$\overset{1}{(I\rightarrow)}$, $\overset{1}{(I\rightarrow)}$, $\overset{1}{(I\rightarrow)}$ - других нет

6. Обсуждение результатов анализа музыкальных текстов. По итогам проведенных экспериментов можно сделать следующие выводы.

6.1. Повтор – важнейший структурный элемент песенных мелодий. Из 122 анализировавшихся мелодий Шуберта лишь 13 не имели повторов с указанными значениями l и k (условно их можно назвать "бесповторными"). Из 216 мелодий советских композиторов в класс бесповторных вошли 37. Соответствующий показатель для подборки русских народных песен (172 мелодии) равен 67. Условность термина "бесповторный" означает, что при уменьшении значения l или увеличении k неслучайные повторы могут обнаружиться и в этих мелодиях.

Из двух типов повторности (первого и второго рода) превалирующей является повторность первого рода. Косвенное подтверждение этому дает сопоставление числа замен, фиксируемых в повторах первого и второго рода на одном и том же исходном материале (см. табл. I).

6.2. Варьируемый повтор количественно преобладает над неварьируемым (совершенным). Число мелодий, в которых встречаются только неварьируемые повторы ($k = 0$) длины 7 и выше, невелико: 7 у Шуберта [9], 19 – у советских композиторов [7] и лишь подборка русских народных песен [8] представляет относительное исключение – 48 (заметим, что русские народные песни в среднем значительно короче мелодий из двух первых подборок). Ситуации, когда имеет место варьирование, наблюдаются гораздо чаще (Шуберт – 102 мелодии, сборник [7] – 160 мелодий, подборка русских народных песен – 57 мелодий). Следует учесть, что речь идет только об одном способе варьирования. Добавление других может лишь расширить класс мелодий с варьируемыми повторами. Это и обуславливает интерес к изучению закономерностей варьирования.

6.3. Состав и количественные характеристики замен в повторах первого и второго рода в целом довольно близки между собой. Отмечается некоторое преобладание замен типа 0 в повторах второго рода над аналогичными заменами в повторах I рода; в отношении замен типа I картина обратная.

6.4. Из трех элементов кодовой комбинации изменению чаще всего (примерно в 50–60% случаев) подвергается лишь один первый элемент – величина интервала (замены типа 0 и I). Примерно в 15% случаев изменению подвергается лишь второй элемент кодовой комбинации (знак I), а в (15–20)% случаев – первый и второй элементы одновременно (замены типа 1,2). Замены типа 3, 1,3, 2,3 и

1,2,3, связанные с изменением метрических характеристик, довольно редки (в сумме 10-15%). Наиболее сильные отклонения от приведенных средних значений наблюдаются в мелодиях Шуберта (замены типа 2 по повторам второго рода составляют всего лишь 7%, зато замены, меняющие метрические характеристики, - 23%). Для русских народных песен доля замен типа 1,3 в повторах второго рода в 7 раз выше, чем в повторах первого рода.

6.5. Наиболее часто используемые замены (см. табл.2) связаны с изменением направления движения мелодической линии, т.е. принадлежат типу 2, который в целом не является самым распространенным. Эти же замены наиболее часты и в повторах второго рода (сборники [7,8]). Исключение вновь составляют мелодии песен Шуберта [9], для которых наиболее распространенными являются замены типа 0, и самая частая замена ((1+)/(0+-)) принадлежит данному типу. Это объясняется тем, что интервал "прима" (I = 0) - самый частый в мелодиях Шуберта, который активно использовал рецитатив, близкий к традиционному оперному, особенно в своих ранних балладах. Кроме того, лирической мелодии Шуберта свойственна большая напевность и непрерывность течения. Новая фраза часто вступает на том самом звуке, который завершал предыдущую (в нашей кодировке такому переходу (как и рецитативу) соответствует значение I = 0). Зрелый Шуберт все более отходит от применения рецитатива традиционного типа, разнообразит его, что и приводит к появлению большого числа замен типа 0. Этот же тип замен Шуберт использует для сглаживания скачков в мелодической линии (когда ненулевое значение интервала заменяется нулевым).

Самыми разнообразными по составу являются замены типа I, которые составляют 30-40% от общего многообразия замен. На втором месте идут замены типа 0 (15-25%). Наименьшее число разновидностей - в заменах типа 2 и 3. В целом эти цифры коррелируют со значениями мощностей каждого класса замен, получаемыми из чисто комбинаторных соображений.

Что касается конкретного вида замен, то анализ табл.2 показывает, что замены в мелодиях песен сборников [7-9] весьма похожи: соответствующие ранжирования имеют высокую степень корреляции.

6.6. Если рассматриваются (1,k) - повторы со значением $k > 1$, то наблюдается кластеризация замен внутри фрагментов, образующих повтор. Так, если $k = 2$, то примерно в половине случаев несоответствия расположены рядом друг с другом, образуя своего рода "ган -

дем". Это говорит о неслучайности распределения замен по длине повтора. Данный вывод относится как к повторам первого, так и второго рода.

В подборке русских народных песен среди повторов I рода встретилось 30 тандемных замен, в песнях советских композиторов - 107, песнях Шуберта - 189. Для повторов 2-го рода эти параметры равны 29, 48 и 72 соответственно. Массовость тандемных замен (особенно в повторах первого рода) позволяет говорить о варьировании тандемного типа, как о приеме, часто имеющем место в композиторской практике.

Анализ тандемных замен (а более широко - преобразований, которые затрагивают сразу группу следующих друг за другом элементов) позволяет выделить в качестве типовых следующие преобразования.

I. Обострение или сглаживание пика в мелодической линии как в восходящем, так и нисходящем движении при сохранении (за редким исключением) метрических акцентов. Формально, если замена имеет вид:

$$\left(\begin{array}{cc} I_1 & S_1 \\ I_2 & S_2 \end{array} \right), \text{ то } S_1 = S_1', S_2 = S_2', \text{ а } I_1 + I_2 = I_1' + I_2'.$$

Разновидностями такого преобразования являются:

а) замена повтора на одной ноте скачком (вверх или вниз) от этой ноты на какой-либо интервал и возврат к ней:

$$\left(\begin{array}{c} 0+ \\ 5- \end{array} \right) \begin{array}{c} 0++ \\ 5++ \end{array}, \left(\begin{array}{c} 0+ \\ I+- \end{array} \right) \begin{array}{c} 0++ \\ I+- \end{array};$$

б) замена одного скачка другим, усиливающая или уменьшающая динамику мелодии:

$$\left(\begin{array}{c} 3++ \\ 5++ \end{array} \right) \begin{array}{c} 3- \\ 5-- \end{array}, \left(\begin{array}{c} 2+ \\ I++ \end{array} \right) \begin{array}{c} 2+- \\ I-- \end{array};$$

в) движение к одному и тому же звуку за счет опевания его сверху или снизу:

$$\left(\begin{array}{c} 5- \\ 2+- \end{array} \right) \begin{array}{c} 4++ \\ 3-+ \end{array}, \left(\begin{array}{c} 2+ \\ I+- \end{array} \right) \begin{array}{c} 4-+ \\ 3-+ \end{array}, \left(\begin{array}{c} I- \\ I++ \end{array} \right) \begin{array}{c} I-- \\ 3-- \end{array};$$

г) заполнение интервала, когда скачок на какой-либо интервал заменяется двумя последовательными интервалами, сумма величин которых равна величине скачка:

$$\left(\begin{array}{c} 0+ \\ 2+- \end{array} \right) \begin{array}{c} 7++ \\ 5++ \end{array}, \left(\begin{array}{c} 2+ \\ 4-+ \end{array} \right) \begin{array}{c} 2-+ \\ 0++ \end{array}, \left(\begin{array}{c} 0+ \\ I++ \end{array} \right) \begin{array}{c} 2++ \\ I++ \end{array}, \left(\begin{array}{c} 6++ \\ 2+- \end{array} \right) \begin{array}{c} I++ \\ 5+- \end{array}.$$

2. Перестановка соседних элементов, иногда сопровождаемая незначительным их варьированием. Формально, если замена имеет вид:

$$\begin{pmatrix} I_1 & S_1 & I_2 & S_2 \\ I'_1 & S'_1 & I'_2 & S'_2 \end{pmatrix}, \text{ то } I_1 S_1 = I'_2 S'_2 \quad \text{и} \quad I_2 S_2 = I'_1 S'_1.$$

Встречается и вариант перестановки со вставкой:

$$\begin{pmatrix} I_1 & S_1 & [I_2 & S_2] & I_3 & S_3 \\ I'_1 & S'_1 & [I'_2 & S'_2] & I'_3 & S'_3 \end{pmatrix}, \text{ где } I_1 S_1 = I'_3 S'_3, I_3 S_3 = I'_1 S'_1, \text{ а } I_2 S_2 = I'_2 S'_2.$$

Примеры перестановок:

$$\begin{pmatrix} 5+-0++ \\ 0++5+- \end{pmatrix}; \begin{pmatrix} 4+-I-- \\ I--3+- \end{pmatrix}; \begin{pmatrix} I+- [I++ 7-- \\ 7-- [I++ I+- \end{pmatrix}; \begin{pmatrix} I+- [I-- 0++ \\ 2++ [I-- I+- \end{pmatrix}.$$

точная перестановка

варьируемая перестановка

перестановка со вставкой

варьируемая перестановка со вставкой

3. Изменение направления движения мелодической линии. Формально, если замена имеет вид $\begin{pmatrix} I_1 & S_1 & I_2 & S_2 \\ I'_1 & S'_1 & I'_2 & S'_2 \end{pmatrix}$, то знак $(I_1 + I_2)$ противоположен знаку $(I'_1 + I'_2)$, а значения стоп чаще всего совпадают. Среди замен такого типа самыми распространенными, хотя и немногочисленными, являются:

$$\begin{pmatrix} I \rightarrow I \rightarrow \\ I \rightarrow I \rightarrow \end{pmatrix}, \begin{pmatrix} I \rightarrow I \rightarrow \\ I \rightarrow I \rightarrow \end{pmatrix}, \text{ в которых } |I_1 + I_2| = |I'_1 + I'_2|.$$

Остальное многообразие замен такого типа представимо в вышеприведенной общей форме:

$$\begin{pmatrix} 2++3+- \\ I \rightarrow I \rightarrow \end{pmatrix}; \begin{pmatrix} 2+-0++ \\ I \rightarrow I \rightarrow \end{pmatrix}; \begin{pmatrix} I \rightarrow I \rightarrow \\ 0++0+- \end{pmatrix}; \begin{pmatrix} 2--2++ \\ 3--4+- \end{pmatrix} \quad \text{и т.д.}$$

Естественно, что отмеченные выше три типа преобразований не исчерпывают всего их разнообразия, но отчетливо проявляют тенденции варьирования.

Предположение, что варьирование часто затрагивает сразу группу следующих друг за другом элементов, подтверждается анализом несовершенных повторов с числом различий $k > 2$. Например, из 18 обнару-

женных в песнях Шуберта (15,3)–повторов в 7 случаях все три замены располагались подряд, в 5 случаях отмечались тандемы, в 6 случаях – тандемы со вставкой. Следует заметить, что тандемы, образуемые тремя последовательными заменами, часто сводимы к вышеперечисленным типам преобразований. Например, замена $\begin{pmatrix} I \rightarrow I \rightarrow 5 \rightarrow \\ 2 \rightarrow I \rightarrow 0 \rightarrow \end{pmatrix}$ сводится к преобразованию I "в", а замена $\begin{pmatrix} I \rightarrow I \rightarrow 0 \rightarrow \\ 0 \rightarrow 0 \rightarrow I \rightarrow \end{pmatrix}$ является перестановкой с варьированной вставкой.

6.7 Разбиение множества замен на 8 указанных выше типов основывается на специфике используемой кодировки и в первом приближении выглядит несколько формальным. Возникает вопрос, насколько согласуется данное разбиение с тем, которому осознанно (а может быть, и подсознательно) следуют сами композиторы?

Анализ характера замен в повторах первого рода показывает, что довольно часто встречаются мелодии, в которых превалирует какой-то один тип варьирования, иногда довольно узкий. Среди мелодий Шуберта, к примеру, встречается немало таких, где преобладают замены типа 0. Дунаевский в песне "Наша молодость идет" пользуется лишь заменами типа 1,3, а в мелодии "Если Волга разольется" [7] только заменами типа 3. Заметим, что это весьма узкие (специфичные) классы замен, которые составляют примерно 4–5% от общего числа замен. В песне А.Новикова "Смуглянка" [7] используются лишь замены типа I с $|\Delta I| = 1$, где под ΔI понимается разность величин интервалов у кодов, описывающих замену (например, (2++) заменяется на (3++) или (3—) на (4—) и т.п.). В песнях Б.Соловьева-Седого "Услышь меня, хорошая" и В.Захарова "И кто его знает" [7] встречаются только замены типа 2. Число подобных примеров нетрудно увеличить.

Нередки случаи, когда преобладает и какой-то один способ размещения замен (разнесенные или тандемные). Например, в песне Г.Новосова "Далеко-далеко" 7 тандемных замен и лишь одна одиночная, а в песне Н.Богословского "Лизавета" [7] все замены расположены тандемами.

Рассмотренные примеры, с нашей точки зрения, подтверждают, что, по крайней мере, такие типы замен, как 0, 1, 2, 3, 1,2, 1,3, существуют объективно, а не только обусловлены используемой кодировкой. Каждый из типов служит достижению определенной выразительной цели. Эффект, получаемый за счет использования замен того или иного типа, по-видимому, осознается отдельными ком-

позиторами, о чем свидетельствует нарочитое, подчеркнутое предпочтение, отдаваемое определенным типам замен в отдельных мелодиях.

Что же касается замен типа 2,3 и 1,2,3, то они столь малочисленны, что делать какие-либо выводы в их отношении пока преждевременно. Заметим, только, что они довольно часто фигурируют не в одиночном варианте, а в тандемном. Это наводит на мысль, что замены такого типа носят в основном искусственный характер и возникают в результате применения "групповых" преобразований (например, перестановки, заполнения интервалов и т.д.).

6.8. Наблюдаемая предрасположенность отдельных композиторов к тому или иному типу замен (например, Шуберта к заменам типа 0) уже в определенной степени характеризует стиль композитора. Еще более дифференцированным является использование тех или иных вариантов замен внутри каждого типа. Отдельные замены носят иногда характер "визитной карточки" композитора. К примеру, замена (2→) на (2—) весьма редкого типа 3 встретилась всего лишь в трех из 216 мелодий сборника [7], все три принадлежат В.Соловьеву-Седому. Замена (0++) на (5++) в том же массиве песен встретилась дважды, причем оба раза у И.Дунаевского ("Физкультурный марш" и "Спортивный марш"). Контексты, в которых используется эта замена в обоих мелодиях, также весьма похожи:

$$\begin{array}{l}
 \text{"Физкультурный марш"} \\
 \left(\begin{array}{c} \text{I} \rightarrow \text{I} \rightarrow 5 \rightarrow \text{I} \rightarrow \text{I} \rightarrow \text{I} \rightarrow 0 \rightarrow \text{I} \\ \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \\ \text{I} \rightarrow \text{I} \rightarrow 0 \rightarrow \text{I} \rightarrow \text{I} \rightarrow \text{I} \rightarrow 0 \rightarrow \text{I} \end{array} \right);
 \end{array}
 \quad
 \begin{array}{l}
 \text{"Спортивный марш"} \\
 \left(\begin{array}{c} 2 \rightarrow 3 \rightarrow 5 \rightarrow 0 \rightarrow \text{I} \rightarrow \text{I} \rightarrow \text{I} \rightarrow \\ \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \quad \parallel \\ 2 \rightarrow \text{I} \rightarrow 0 \rightarrow 0 \rightarrow \text{I} \rightarrow \text{I} \rightarrow \text{I} \rightarrow \end{array} \right).
 \end{array}$$

Эти и подобные им примеры показывают, что исследование интегральных и локальных характеристик замен открывает интересные возможности в области идентификации мелодий и количественной стилистики.

З а к л ю ч е н и е

Повторы являются важными структурными элементами текстов различной природы: литературных и музыкальных произведений, первичных структур ДНК - молекул и белков, программ, написанных на алгоритмических языках. Повторение отдельных фрагментов часто сопровождается их варьированием. Роль варьирования в формировании указанных текстов не менее значительна, чем роль повторности. Использование ЭВМ и разработанного авторами аппарата анализа повторов открывает широкие возможности для количественного исследования закономерностей варьирования.

В первой части работы сформулированы основные цели, достигаемые с помощью варьирования, и рассмотрены наиболее типичные приемы варьирования, которые оказываются общими для текстов различной природы. Введено понятие несовершенного повтора, который выступает в качестве индикатора каждого элементарного акта варьирования. Кратко описаны конструкции, позволяющие фиксировать все элементарные акты варьирования определенного вида. Алгоритмы направленного перебора, реализующие эти конструкции, составляют основу методики количественного исследования закономерностей варьирования.

Возможности методики проиллюстрированы во второй части работы на примере анализа песенных мелодий. В эксперименте было использовано свыше 600 мелодий, разбитых на три однородные в определенном смысле подборки. Наибольший интерес представляют количественные оценки повторности и варьированной повторности, выявление стилевых особенностей на уровне варьирования, а также тот факт, что варьирование часто затрагивает сразу группу следующих друг за другом элементов.

Возможности дальнейшего продвижения в этой области авторы видят в рассмотрении других способов варьирования (особенно на семантическом уровне), разработке новых эффективных алгоритмов перебора, без которых невозможно получение количественных оценок, и в использовании качественных и количественных закономерностей варьирования в программах синтеза текстов.

Л и т е р а т у р а

1. ГУСЕВ В.Д. Механизмы обнаружения структурных закономерностей в символьных последовательностях // Проблемы обработки информации. - Новосибирск, 1983. - Вып. 100: Вычислительные системы. - С. 47-66.
2. ЗАРИПОВ Р.Х. Машинный поиск вариантов при моделировании творческого процесса. - М.: Наука, 1983. - 232 с.
3. ГУСЕВ В.Д., КУЛИЧКОВ В.А., НИКУЛИН А.Е. Алгоритмы поиска несовершенных повторов в генетических текстах // Анализ символьных последовательностей. - Новосибирск, 1985. - Вып. 113: Вычислительные системы. - С. 107-122.
4. ГУСЕВ В.Д., КОСАРЕВ Ю.Г., ТИМОФЕЕВА М.К. и др. Пакет прикладных программ для анализа произвольных символьных последовательностей значительной длины (СИМВОЛ) // Структурный анализ символьных последовательностей. - Новосибирск, 1984. - Вып. 101: - Вычислительные системы. - С. 3-21.
5. ЗАРИПОВ Р.Х. Построение частотных словарей музыкальных интонаций для анализа и моделирования мелодий // Проблемы кибернетики. 1984. Вып. 41. - С. 207-252.

6. БАХМУТОВА И.В., ГУСЕВ В.Д., ЗАРИПОВ Р.Х., ТИТКОВА Т.Н. Выявление и анализ сходных фрагментов в музыкальных произведениях // Анализ символьных последовательностей. - Новосибирск, 1985. - Вып. IIЗ: Вычислительные системы. - С. 3-45.
7. Русские песни. Выпуск 3. Песни советских композиторов. - Л.: Музыка, 1965. - 622 с.
8. Русские народные песни. - М.: Музыка, 1985. - 125 с.
9. Шуберт. Избранные песни, т. I, 2, 6. - Л.: Музыка, 1965.
10. День победы. Песни великого подвига. - М.: Худ.литература-Музыка, 1985. - 175 с.

Поступила в ред.-изд.отд.
15 июня 1987 года