ОПТИМИЗАЦИЯ СТРУКТУР ИЕРАРХИЧЕСКИХ РАСПОЗНАЮЩИХ АВТОМАТОВ

Г.Я. Болошин

Интерес к иерархическим (многоуровневым) распознающим автоматах в настоящее время достаточно широк. Исследователи ясно отдают себе отчет в том, что при строгом подходе иерархические автоматы нельзя оптимизировать отдельно по уровням, так как упрощение одних уровней может привести к усложнению других [1]. Поэтому необходимо рассматривать автомат в целом, пользуясь обобщенным критерием для оценки его сложности.

Критерий оптимизации

В качестве оптимизации структуры иерархического распознающего автомата наиболее целесообразно использовать его стоимость, которая определяется, в основном, следующими факторами:

- размером и конструкцией поля рецепторов,
- типом решающих правил (сложностью классификатора),
- конструкцией исполнительных органов (эффекторов),
- стоимостью ожидаемых потерь.

В результате мы можем записать выражение для вычисления общей стоимости $C$ распознающего автомата

$$C = C_{рел} + C_{кл} + C_{эфф} + C_{п}$$

где

$C_{рел}$ — стоимость поля рецепторов;
$C_{кл}$ — стоимость классификатора;
$C_{эфф}$ — стоимость эффектора;
$C_{п}$ — стоимость потерь.
Разумеется, мы предполагаем, что все компоненты в [1] измерены в одной и той же системе единиц.

При обосновании целесообразности введения промежуточных ступеней распознавания мы будем сравнивать стоимости $C'$ однородного и $C'$ иерархического (р - уровневого) распознающих автоматов, предназначенных для решения одной и той же задачи. В связи с тем, что стоимость эффектора можно пренебречь, так как она одинакова в обоих типах автоматов. Кроме того, мы будем предполагать, что $C'_{s} = C'_{s}^{r}$, поскольку снятие этого ограничения приводит к специфическим задачам, не входящим в рамки данной работы и рассмотренных отдельно [2].

В итоге нах будет интересовать величина

$$\Delta C = C' - C' = (C'_{s} - C'_{s1})(C'_{s} - C'_{s2}),$$

где $C_{s}'$ - суммарная стоимость классификаторов всех ступеней.

Введение промежуточных ступеней распознавания (иерархизация) целесообразно лишь в тех случаях, когда $\Delta C < 0$. В дальнейшем мы сосредоточим свое внимание на $\Delta C_{s}$. Что же касается $\Delta C_{s1}$, то здесь мы ограничимся лишь общими замечаниями.

Вообще говоря, в стоимость классификатора включается затраты на элементы памяти, предназначенные для хранения решающих поддири и распознаемых реализаций, а также стоимость времени $\gamma$ принятия решения. Чтобы не усложнять анализ, мы не будем обсуждать вопрос о соотношении объема памяти и времени счета, а в стоимость классификатора включим лишь затраты на элементы памяти, полагая, что $\gamma$ не лимитировано.

Однотипный распознающий автомат

Оценить стоимость классификатора однотипного (элементарного) распознающего автомата. Пусть $X_{r}$ - исходное признаковое пространство размерности $L_{r}$, $K_{r}$ - число распознаемых окончательных образов $S_{r}$. И пусть пространство $X_{r}$ разбито на кучкой из осевой на $L_{r}$ градаций. Тогда общее число ячек в $X_{r}$ будет равно

$$H_{r} = \prod_{i} L_{r}^{n_{i}}.$$

Назовем избыточность описания величиной $\theta = 1 - Q_{r}$,

где $Q_{r} = \frac{K_{r}}{H_{r}}$.

Очевидно ограничение $0 \leq Q_{r} \leq 1$.

Предположим, что относительная сложность описания всех образов из $S_{r}$ (т.е., что то же самое, сложность решаемых правил) одинакова и равна $V_{r}$. Под относительной сложностью здесь понимается количество чисел, описывающих один образ, в пересчете на одно ось пространства. Эти уровни образов (решаемые правила) хранятся в долговременной памяти, стоимость одного элемента которой обозначим через $C_{e}$. Реализация, подлежащая распознаванию, хранится в оперативной памяти, у которой стоимость одного элемента равна $C_{o}$.

Общая стоимость классификатора определяется выражением

$$C_{o} = c_{o}V_{r}K_{r} \sum_{t_{c}} \log_{2} t_{c} + c_{o} \sum_{i} \log_{2} t_{c} = (c_{o}V_{r}K_{r} + c_{o}) \log_{2} t_{c} = (c_{o}V_{r}K_{r} + c_{o}) \log_{2} \frac{H_{r}}{Q_{r}},$$

где $\lambda$ - число устойчивых состояний элемента памяти (предполагается одинаковым для оперативной и долговременной памяти).

Иерархический распознающий автомат

I. Формальная модель

Прежде чем приступить к построению формальной модели, рассмотрим один пример.

Пусть перед автоматом стоит задача распознавания напечатанных слов. Можно сконструировать автомат так, что он будет принимать решение о слове в целом по показаниям рецепторов, покрытым всеми поверхностями, на которой напечатано слово. Это будет

Пользуясь величиной $\log_{2} t_{c}$, а не ближним к ней сверху целым числом, мы получаем нижнюю оценку $C_{o}$. Верхнюю оценку нетрудно получить, подставив в (2) вместо $\log_{2} t_{c}$ величину $(\log_{2} t_{c} + 1)$. Дальнейшие выводы будем проводить лишь для нижней оценки, чтобы исследовать максимально возможный эффект от иерархизации.
одноступенный автомат. Можно пойти другим путем — разбить всю исходную поверхность на участки, в которых повсеместно по одной букве, распознавать на этих участках буквы, а затем по после —
довательности букв — слова. Это — автомат с одной промежуточной ступенью распознавания. Можно ввести ещё одну ступень, при
нимая на первом уровне решения о геометрических фигурах, по кото
рым затем распознавать буквы, а по последним — слова.

При этом исходная поверхность оказывается разбитой на ещё более мелкие участки для принятия первичных решений.

Приведенный пример наглядно иллюстрирует дальнейшие формальные построения [3]. Пусть мы имеем исходное признаковое про
странство \( X \), необходимое и достаточное для распознавания оконечных образов \( S \), с надежностью \( P \). Введем промежуточную ступень распознавания следующим образом: разобьем \( X \), на \( \mathcal{N} \), несовпадающих (в общем случае пересекающихся) подпространств и в каждом из них будем принимать решения о принадлежности реа
лизаций к образам \( S_i \) (i = 1, 2, ..., \( \mathcal{N} \)), составляющих промежуточный алфавит первой ступени. Набор решений первой ступени распознавания является признаком пространством, в котором работает вторая ступень. С этим пространством мы можем поступать так же, как и с исходным — разбить его на несовпадающие подпро
странства для принятия в них вторых промежуточных решений и т. д. Следует отметить, что для работы \( j \) —й (\( j \neq i \)) ступени распознавания (\( j \neq i \))-й ступень должна выдать не менее двух решений, ибо в противном случае \( j \)-я ступень будет прос — тым ретранслятором решений (\( j \neq i \))-й ступени и надобностью в ней, как в распознавающем устройстве, отпадает.

Последний уровень распознавания вступает в действие после того, как первый покроет все пространство \( X \).

Структурная классификация

Будем изображать распознающие автоматы в виде ориентированных конечных графов (рис. 1).

Вершина графа представляет собой элементарный распознающий автомат. Входящие дуги — параметры, по которым производятся распознавание. Выходящие дуги — решения, принятые элементарны
ми автоматами.

Множество вершин графа разбито на слои (иерархические ступени). Нулевой слой — рецепторы (исходные измерительные прибо
ры) - вершины, не имеющие входящих дуг. Последний слой - вершина, не имеющая выходящих дуг. Это - эффектор (исполнительный орган). Две вершины, непосредственно соединенные дугами, будем называть смежными.

Каждой дуге, выходящей из вершины нулевого слоя, припишем длину, равную количеству градаций данного параметра. Дугам, выходящим из вершин \( m \) -го слоя (\( m > 0 \)), припишем длины, равные количеству образцов, расположенных автоматами - вершинами, из которых эти дуги выходят. Вполне понятно, что из одной вершины могут выходить дуги только одинаковой длины.

Схема, представленная на рис. 1, является параллельной в том смысле, что в каждом подпространстве функционирует свой распознающий автомат.

Можно построить последовательную схему, оставив на каждой иерархической ступени по одному распознающему автомату и под - ключая его поочередно ко всем подпространствам данной ступени (скалярное распознавание), возможен и смещенный (последовательно-последовательный) вариант.

Это - классификация иерархических распознающих автоматов по способу принятия решений. По типу алфавита распознаваемых об - разов иерархические автомат делются на три группы:
- с одинаковыми алфавитами в подпространствах на \( m \)-й иерархическом уровне (\( m = 1,1,3,\ldots \));
- с различными алфавитами;
- смещенный вариант.

Однако мы ограничиваемся лишь алфавитами, существенно влияющими на структуру автомата.

Очевидно, однажды сформированный иерархический распознающий автомат можно структурно упорядочить (по слоям) различными способами. Чтобы избежать этого, предлагается следующий алгоритм упорядочения.

В нулевой слой включаются все вершины, не имеющие входящих дуг.

В первый слой включаются все вершины, смежные только вершинами нулевого и первого слоев.

В \( m \)-й слой включаются все вершины, смежные только с вершинами (\( m - 1 \))-го, (\( m - 2 \))-го и т.д. слоев. Процедура продолжается до тех пор, пока не будут исчерпаны все вершины графа.

Те графы, которые допускают решение по алгоритму упорядочения, будем называть графами (автоматами) без обратных связей.

Здесь мы сталкиваемся с субъективно решаемой задачей таксономии иерархических распознающих автоматов. Как в любой задаче таксономии, здесь весьма существенную роль играют признаки, которые мы выбрали в качестве основы для классификации. В настоящей работе выбор признаков обусловлен рассмотренной выше формальной модели иерархического распознающего автомата.
Конструктивного алгоритма, пригодного для упорядочения структуры автоматов с обратными связями, к настоящему времени найти не удалось.

Из предложенного алгоритма упорядочения вытекают очевидно следствия.

СЛЕДСТВИЕ 1. В одном слое не могут находиться смежные вершины.

СЛЕДСТВИЕ 2. Граф с контурами не имеет решения по алгоритму упорядочения.

ТЕОРЕМА 1. Иерархические распознающие автоматы обладают обратными связями в том и только в том случае, если изображающие их графы имеют контуры.

Из следствия 2 известно, что при наличии контуров граф не упорядочивается. Остается доказать, что при отсутствии контуров решение по алгоритму упорядочения имеется. Поскольку граф не имеет контуров, то на любом его пути определено отношение строгого следования вершин. Если это так, то найдется хотя бы одна вершина, следующая только за вершинами нулевого слоя. По этой же причине найдется хотя бы одна вершина для второго слоя и т.д.

Так как граф конечен, то мы исчерпаем все его вершины за конечное число шагов. Твёрдом доказана.

На рис. 2а приведено изображение автомата с обратной связью. Изображение на рис. 2б не имеет обратной связи и после структурного упорядочения переходит в граф, представленный на рис. 2а.

Чтобы в дальнейшем не оговаривать длиной фразой тип рассматриваемых автоматов, мы привели ни цифры в соответствии со следующей класификационной таблицей:

<table>
<thead>
<tr>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Любой конкретный иерархический распознающий автомат можно описать структурно, взяв по одной цифре из каждого столбца.

Первый столбец: 1) Автомат с обратными связями.
2) Автомат без обратных связей.
3) Пирамидальный автомат.

Рис. 2. Типы иерархических распознающих автоматов.

-21-
Второй столбец: 1) Смещенные алфавиты в подпространствах.
2) Различные алфавиты.
3) Однаковые алфавиты.

Третий столбец: 1) Пересекающиеся подпространства любых элементарных распознающих автоматов.
2) Пересекающиеся подпространства автоматов одной и той же иерархической ступени.
3) Непересекающиеся подпространства.

Четвертый столбец: 1) Комбинированное (последовательно-палиндромное) распознавание по подпространствам.
2) Параллельное распознавание.
3) Последовательное распознавание.

Очевидно, в автомат IIII входят как частный случай все остальные 80 типов автоматов.

Вполне понятно, что тип 3333 - самый простой, а потому он является целью, к которой мы должны стремиться при конструировании иерархических распознающих автоматов. В настоящей работе мы ограничимся детальным рассмотрением только этой "цели" и ближайшей к ней "окрестности", а именно:

3 3 3 2 3

3. Избыточность описания в иерархических распознающих автоматах

Вершину графа, соответствующую эфектору, будем называть наибольшим элементом (см. [4]).

Обозначим через $b_i$ количество дуг, выходящих из $i$-й вершины. Произведение $B = \prod b_i$, где $M$ - количество вершин графа без наибольшего элемента, будем называть общим коэффициентом пересечения подпространства иерархического распознающего автомата.

Справедлива следующая

ТЕОРЕМА П. Если одноступенный и иерархический распознающий автомат решает одну и ту же задачу, то коэффициент избыточности описания в одноступенчатом автомате равен общему коэффициенту пересечения подпространств, умноженному на произведение коэффициентов избыточности описания во всех элементарных автоматах, входящих в иерархический распознающий автомат, то есть

$$Q_o = B \prod Q_P,$$

где $K_o$ - суммарное количество вершин нулевого слоя (эффектора и рецепторов).

ДОКАЗАТЕЛЬСТВО.

$$K_o = Q_r \prod C_{j_i}$$

где $j_o$ - последняя степень распознавания (вершину, предшествующую наибольшему элементу), $n_r$ - количество входящих дуг, $x_i$ - длина $j_i$-й входящей дуги.

Но поскольку любая входящая дуга выходит из другой вершины, то

$$C_{j_i} = Q_i \prod C_{j_{i-1}} \prod C_{j_{i-1}}$$

где $x_i$ - длина $j_i$-й входящей дуги, $j_{i-1}$ - предпоследняя степень распознавания и т.д.

Выполнив все подстановки, мы, очевидно, получим

$$K_o = \prod Q_r \prod C_{j_i} = H \prod Q_r$$

Но для одноступенного автомата $K_o = Q_r H_o$, поэтому

$$Q_o = B \prod Q_P,$$

что и требовалось доказать.
4. Автомат типа 3332

Поскольку в рассматриваемом автомате распознавание параллельное, то
Собщ = Срасс и Δ С = Срасс - Ссообщ.
На м-й ступени в j-м подпространстве стоимость Ckм, mj определяется (по аналогии с выражением (2))

Ckм, mj = Σ Ck mj + Ck mj + Ck mj,

Здесь, как и ранее, принято естественное предположение об одновременности ячеек памяти (одинаковые Л, C3 и C0, для всех m и j).

При параллельном распознавании для р ступеней
\[ C_k = \sum_{m=1}^{r} C_{km} = \sum_{m=1}^{r} (C_{km} V_{km} + C_{km} V_{km} + C_{km} V_{km}) \log \frac{k_{m}}{Q_{m}^{1}} \]

где нижний индекс r обозначает последнюю ступень распознавания; \( \beta_{m} \) - количество подпространств на m-й ступени.

Очевидно, что \( r = r \). Выделение р -то слагаемого в (4) потребовалось потому, что в него входит \( C_{km} \) (число окончных образов) - величина, которая задана условиями задачи и которую мы не можем изменить.

При одинаковых алфавитах в подпространствах:
\[ C_{km} = C_{km}, Q_{km} = Q_{km}, V_{km} = V_{km} (j = 1, 2, ..., \beta_{m}) \]

\[ C_{km} = \sum_{m=1}^{r} (C_{km} V_{km} + C_{km} V_{km} + C_{km} V_{km}) \log \frac{k_{m}}{Q_{m}^{1}} \] (3)

Введем обозначение \( Q_m = \frac{\beta_{m}}{Q_{m}^{1}} \), где \( Q_m \) - коэффициент общей избыточности описания на m-й ступени распознавания.

Для пираимдального автомата \( Q_m = Q_m \beta_{m} \) и \( B = 1 \). По-этому (в соответствии с теоремой II) \( Q_m = Q_m \beta_{m} \).

Так как \( k_{m} = Q_{km} \), то \( Q_{km}^{1} = \frac{k_{m}}{Q_{m}^{1}} \).

\[ L_m \log \frac{k_{m}}{Q_{m}^{1}} = \log \frac{k_{m}}{Q_{m}^{1}} \] (4)

Легко видеть, что при увеличении \( m \) от 1 до р выражение (4) не возрастает и остается в пределах от \( \log \frac{k_{m}}{Q_{m}^{1}} \) до \( \log \frac{k_{m}}{Q_{m}^{1}} \). Верхний предел мы изменить не можем, поэтому что в него входят только заданные величины. Нижний же предел можно уменьшать увеличением \( r \), а остальное \( r \) можно брать таким способом, чтобы (4) как можно быстрее достигло своего минимального значения. Таким образом, в иерархических распознавающих автоматах целесообразно максимум избыточности сосредоточить на первой ступени распознавания. Из (2), (3) и (4) получаем

\[ \Delta C = C_{km}^{1} - C_{km}^{2} = \log \frac{C_{km}^{1} V_{km} + C_{km}^{2} V_{km}}{C_{km}^{1} V_{km} + C_{km}^{2} V_{km}} \frac{1}{Q_{km}^{1}} \frac{1}{Q_{km}^{2}} \] (5)

Из (5) видно, что нам необходимо минимизировать \( r \). Указано, что при этом увеличивается количество элементарных распознавающих автоматов на первой ступени и

\[ \frac{1}{Q_{km}^{1}} \frac{1}{Q_{km}^{2}} \]

Легко видеть, что при \( r \geq 2, \beta_{m} \geq 1 \), причем \( \beta_{m} = 1 \) при \( r = 2 \). Итак, оптимальный вариант иерархического распознавающего автомата типа 3332 должен иметь две ступени распознавания, два элемента алфавита в подпространствах первой ступени и одну избыточность на второй ступени распознавания. В целом, эти требования являются теоретически пределом, причем часто принципиально недостижимыми, например, при \( k_{m} = 2 \) (при любом целом \( \beta > 1 \)).

Обратимся к идеального случаю: \( r = 2, Q = 1, k_{m} = 2 \).

\[ \Delta C = \left[ C_{km}^{1} \left( \frac{V_{km}^{1} - V_{km}^{2}}{V_{km}^{1} + V_{km}^{2}} \right) \right] C_{km}^{1} \log \frac{k_{m}}{Q_{m}^{1}} \] (6)

Выражение (6) определяет максимальный выигрыш в стоимости автомата при иерархизации процедур распознавания по типу 3332.

Для большей наглядности приведем несколько частных случаев (см. табл. 1, \( \delta_{1} \)).

1) При \( V_{km}^{1} = V_{km}^{2} = V \) и при двух распознаваемых обратах \( k_{m} = 2 \), иерархизация нецелесообразна (\( \Delta C > 0 \)). Чем большие \( k_{m} \), тем эффективнее иерархизация.

2) При \( V_{km}^{1} > V_{km}^{2} \) и при отсутствии избыточности исходного описания (\( Q = 1 \)), иерархизация нецелесообразна (\( \Delta C > 0 \)). Чем больше избыточность, тем эффективнее иерархизация.

Приведем численный пример, иллюстрирующий возможный выигрыш от иерархизации (\( V_{km}^{1} = V_{km}^{2} = V \), \( k_{m} = 2 \)). Чем большие группы из стандартных отсчетов значений энергий в 12 частотных поло-
5. Автомат типа 3333

В автомате типа 3333 эффект от иерархизации может быть боль-ше, чем в 3332, поскольку в 3333 распознавание скользящее и входная реализация может последовательно (по частям) подаваться на вход и то же поле рецепторов, размер которого равен \( \frac{n}{\tau} \), где \( n \) - число рецепторов в автомате типа 3332, а \( \tau \) - количество предшествующих на первом уровне распознавания.

Кроме того, уменьшаются затраты и на классификатор.

Мы не будем повторять всех рассуждений предыдущего раздела, а приведем лишь конечный результат:

\[
\Delta C = C_{ea} - C_{eo} = \left( C_{e} + C_{o} \right) \log_2 \frac{K_0}{\eta_0} \left( \frac{V}{\eta} - \frac{V}{\eta_0} \right) \log_2 \frac{Q_0}{Q_0}
\]

Некоторые частные случаи приведены в таблице I (\( \delta_1 \)). В отличие от 3332, в 3333 при \( V_1 = V_2 = K_0 = V \) и при двух распознаваемых образах иерархизация целесообразна, если

\[
Q_0 < 2^{\frac{2V}{\eta_0 - \eta}}
\]

Обсуждение результатов

Полученные выше данные показывают, что в тех случаях, когда велики избыточность исходного описания и количество образов, введение промежуточных ступеней распознавания приводит к весьма существенному сокращению стоимости классификатора распознавающего автомата. Причем, чем выше \( K_0 \) и \( \delta_1 \), тем эффективнее иерархизация.

Теоретическим оптимумом является двухступенчатый пирамидальный автомат со скользящим распознаванием двух образов в каждом из непересекающихся подмножеств и с отсутствием избыточности на второй ступени распознавания. Оптимум не всегда достижим. В частности, \( K_0 \) может быть таким, что невозможно обес-печение

\[\Lambda = 2^\Lambda = \Lambda = 2^\Lambda = \Lambda\]

Используя упомянутые в (\( \delta_2 \)) в (\( \delta_2 \)) и (\( \delta_3 \)) формулы для значений максимальной иерархизации автоматов типа 3332, в таблице представлены следующие значения:

<table>
<thead>
<tr>
<th>( \delta_1 )</th>
<th>( \delta_2 )</th>
<th>( \delta_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
</tr>
<tr>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
<td>( \frac{2V}{\eta_0 - \eta} )</td>
</tr>
</tbody>
</table>

Примечание: \( \delta = \Delta C \)

Вместе с тем необходимо учесть, что потребуется определенные затраты на построение коммутирующих устройств.
чить \( q_i \in \mathbb{Q} \) и целесообразно ввести дополнительные промежуточные ступени.

Кроме того, могут быть наложены ограничения на размерность подпространств, в которых принимаются промежуточные решения. И при этом может оказаться, что целесообразно введение более двух ступеней распознавания. Например, если необходимо распознавать 16 образов в 32-мерном пространстве с 32 градациями по каждой из осей, причем на первом уровне решения могут приниматься не более чем 4-мерных подпространств, то оптимальной оказывается трехступенчатая структура. Она в 5 раз дешевле одноступенчатой и на 20% двухступенчатой:

\[ a = V_x V_y, \quad V_x = \frac{v_x - 1}{2} = \frac{1}{2}, \quad V_y = \frac{v_y - 1}{2} = \frac{1}{2}, \quad C = C_0, = C. \]

До сих пор мы считали, что исходная размерность столь велика, что не накладывает дополнительных ограничений на структуру распознавающего автомата. В ряде случаев это оказывается не так.

Легко видеть, что для автоматов типа 333х выполняется не равенство

\[ \frac{32}{2^5} \leq \frac{32}{2^5} \leq \frac{32}{2^5} \leq \frac{32}{2^5} \leq \ldots \leq 1 \]

и если \( n \) достаточно мало, то с ним приходится считаться.

Это приводит к необходимости либо использовать пересекающиеся подпространства, либо распознавать в подпространствах более чем по два образа.

Итак, ясно, что для каждой конкретной задачи распознавания имеется своя оптимальная структура распознавающего автомата. В настоящей же работе изучен лишь предел оптимизации, к которому нужно стремиться при разработке распознавающих автоматов и алгоритмов.

**Литература**

2. Г.Я. ВОЛОШИН. О решающих правилах в иерархических распознавающих автоматах. - Данный сборник, стр. 3.
3. Г.Я. ВОЛОШИН. Модель иерархического распознающего автомата и ее свойства. - Конференция по теории автоматов и искусственному мышлению. Аннотации к докладам, Ташкент, 1968.
4. К. БЕРЛ. Теория графов и её применения. ИД, М., 1962.

*Поступила в редакцию 8.1.1968 г.*