
СКОРОСТЬ ЗАПОМИНАНИЯ ПРИ РАЗНОВЕРОЯТНОМ ПОЯВЛЕНИИ СИМВОЛОВ

П. Б. Невельский, С. П. Боcharова

Харьковский госуниверситет

Как показывает анализ деятельности человека при передаче и переработке информации, роль памяти в этом случае настолько велика, что скорость передачи информации по человеческому каналу вообще определяется уровнем памяти в такой передаче. В зависимости от этого уровня можно выделить три вида каналов, передающих информацию: человеческий канал без памяти, человеческий канал с кратковременной памятью и человеческий канал с долговременной памятью (П. Б. Невельский. О скорости запоминания. «Вопросы психологии», 1966, № 1).

Канал без памяти — это канал восприятия. Информация, передаваемая по этому каналу, перерабатывается человеком последовательно, символ за символом. Пропускная способность здесь достигает десятков двоичных единиц в секунду.

О канале с памятью может идти речь, когда передача информации связана с сохранением какой-то последовательности символов, появляющихся на входе этого канала и исчезающих до начала передачи. По каналу с кратковременной памятью информация передается непосредственно с входа на выход и может быть передана полностью и без помех. Однако передача информации в данном случае ограничена объемом непосредственной памяти, который, по Миллеру (Дж. А. Миллер. Магическое число семь, плюс или минус два. Сб. «Инженерная психология», изд-во «Прогресс», 1964), составляет 9 двоичных цифр, 7 десятичных цифр, 6 букв английского алфавита или 5 односложных английских слов. Скорость передачи информации по этому каналу достигает нескольких двоичных единиц в секунду.

Если же число передаваемых символов превышает возможности непосредственной памяти, то информация либо не может быть передана полностью и без помех, либо передается с большой задержкой. Задержка обусловлена тем, что для полной и безошибочной передачи информации ее нужно воспринимать и воспроизводить неоднократно. Здесь мы видим канал с долговременной памятью: информация по нему не может быть передана непосредственно с входа на выход и нуждается в переработке.

Экспериментальное исследование пропускной способности человеческого канала с долговременной памятью в опытах по запоминанию последовательностей трехзначных чисел равной длины, но с различным содержанием информации; с равным содержанием информации на символ, но разной длины; с равным количеством передаваемой информации,

но с различным числом символов и различной информацией на символ, дало следующие результаты (П. Б. Невельский. Указ. статья).

1. При неизменном количестве предъявляемых символов общее время запоминания и скорость запоминания, выраженная числом символов в единицах информации, зависят от количества информации. При этом а) время запоминания увеличивается с увеличением информации на символ и всей информации в предъявленном ряду; б) скорость запоминания в символах уменьшается с увеличением информации на символ и всей передаваемой информации.

2. С изменением числа предъявляемых символов при равном количестве информации время и скорость запоминания показали большую зависимость от числа символов.

Скорость передачи информации по каналу с долговременной памятью не превышает десятых долей двоичных единиц в секунду, даже в том случае, когда последовательность символов не намного превосходит возможности непосредственной памяти. Эти результаты были получены в опытах по запоминанию последовательностей трехзначных чисел, отобранных из разных по длине алфавитов (2, 8, 64 или 500 трехзначных чисел) и появлявшихся с равной вероятностью.

Цель настоящего эксперимента — рассмотреть скорость долговременно запоминание в случае, когда число запоминаемых символов и длина их алфавита не изменяются, а информация на символ и вся передаваемая информация изменяются в связи с изменением вероятности появления этих символов.

Методика

Проведены три серии опытов. Запоминаемыми символами были последовательности из 15 трехзначных чисел. Во всех сериях числа отбирались из алфавита всех возможных девятисот трехзначных чисел, но с разной вероятностью. В первой серии одно из трехзначных чисел появлялось с вероятностью 0,8, а остальные 899 чисел — с вероятностями $0,2/899 = 0,00022$. Во второй серии одно из трехзначных чисел появлялось с вероятностью 0,6, а остальные 899 чисел — с вероятностью $0,4/899 = 0,00044$. В третьей серии все трехзначные числа появлялись с вероятностями $1/900 = 0,0011$.

В соответствии с принятыми вероятностями ряды чисел отбирались по таблице случайных цифр следующим образом. Для первой серии каждая из восьми цифр этой таблицы (например, 0, 1, 2, 3, 4, 5, 6, 7) заменялась одним и тем же повторяющимся трехзначным числом (например, 258), каждая из остальных двух цифр (8, 9) — трехзначным числом, составленным из трех следующих за нею цифр (при условии, что первая из этих цифр не была нулем). Для второй серии каждая из шести цифр (например, 0, 1, 2, 3, 4, 5) заменялась одним и тем же повторяющимся числом, а каждая из остальных четырех цифр (6, 7, 8, 9) — трехзначным числом, составленным из трех следующих за нею цифр. Для третьей серии каждые три цифры, взятые из таблиц (если первая из них не была нулем), объединялись в трехзначное число. Отобранные последовательности символов (трехзначных чисел) для трех серий опытов выглядели примерно так:

1) 258 258 258 321 258 258 258 258 953 258 258 258 512 837 258
 2) 767 258 258 614 258 258 258 844 116 457 258 258 258 258 258
 3) 683 430 137 745 725 657 274 386 809 612 949 894 241 690 576

В первой серии на 15 символов в среднем должно было быть 12 символов повторяющихся, а во второй серии — 9. В третьей серии повторяющиеся символы — редкое явление. Конечно, в отдельных случаях в первой серии могло оказаться меньше повторяющихся символов, чем во второй, но в среднем для всех приведенных опытов и первой, и второй серий количество их приближалось к расчетному.

Для каждого опыта отбиралась новая последовательность чисел в соответствии с вероятностями их появления. Никакая замена случайно выпавших чисел не допускалось даже в том случае, если выпавший ряд оказывался более трудным или более легким для запоминания. Во всех опытах запоминаемый материал все время варьировался, но вероятности появления символов в каждой серии опытов оставались неизменными.

В разных сериях опытов изменялось количество индивидуальной информации, содержащейся в отдельном символе, которое оценивалось как отрицательный двоичный логарифм появления этого символа:

$$H_{\text{инд}} = -\log_2 p.$$

Изменялось также среднее количество информации на символ, которое оценивалось по известной формуле Шеннона

$$H_{\text{ср}} = -\sum p_i \log_2 p_i.$$

В первой серии

$$H_{\text{ср}} \equiv -\sum_{900} 0,8 \log_2 0,8 + 899 \times 0,00022 \log_2 0,00022 = 2,7 \text{ дв. ед.}$$

Во второй серии

$$H_{\text{ср}} = -\sum_{900} 0,6 \log_2 0,6 + 899 \times 0,00044 \log_2 0,00044 = 4,9 \text{ дв. ед.}$$

В третьей серии, где вероятности были равны,

$$H_{\text{ср}} = -\sum_{900} p_i \log_2 p_i = -\log_2 p = \log_2 900 = 9,8 \text{ дв. ед.}$$

Изменялось и количество всей переданной информации, содержащейся в рядах символов. Это количество оценивалось двумя способами: во-первых, как произведение средней информации, содержащейся в одном символе, $H_{\text{ср}}$, на число символов $m = 15$ (мы назвали эту оценку теоретической); во-вторых, была произведена средняя оценка индивидуальной информации, содержащейся во всех предъявленных символах по протоколам каждого из проведенных опытов каждой серии (эмпирическая оценка). То, что эти две оценки переданной информации значительно не различались, говорит о том, что было проведено достаточное число опытов (20) для того, чтобы могла проявиться вероятностная структура передаваемой информации.

Изменялась также избыточность передаваемого кода

$$R = 1 - \frac{-\sum p_i \log_2 p_i}{\log_2 n},$$

которая была тем больше, чем больше вероятности появления символов отличались от равновероятного их появления. Показатели запоминаемого материала представлены в табл. 1.

Опыты проводились индивидуально. Испытуемыми были 20 студентов университета. С каждым испытуемым ставилось по одному опыту

каждой серии. Все три серии проводились в один день. Порядок следования серии определялся по таблице случайных цифр. Экспериментатор с одинаковой скоростью и интонацией читал ряд из 15 чисел, после чего испытуемый повторял то, что запомнил. Сразу же после первого повторения экспериментатор читал ряд еще раз и испытуемый воспроизводил последовательность вторично. Так продолжалось до первого правильного воспроизведения всего ряда. Воспроизведение считалось правильным только в том случае, если были воспроизведены не только все числа, но и их порядок.

Таблица 1

Показатели передаваемой информации

Показатели	Серии опытов		
	1	2	3
1. Число переданных символов m	15	15	15
2. Длина алфавита символов n	900	900	900
3. Номинальная информация на символ в двоичных единицах $H_{\text{ном}} = \log_2 n$	9,8	9,8	9,8
4. Среднее число кусков информации	6	10	15
5. Вероятность часто появляющегося символа $p_{\text{ч}}$	0,8	0,6	0,0011
6. Вероятность редко появляющегося символа $p_{\text{р}}$	0,00022	0,00044	0,0011
7. Индивидуальная информация часто появляющегося символа в двоичных единицах $H_{\text{инд}} = -\log_2 p_{\text{р}}$	0,32	0,74	9,8
8. Индивидуальная информация редко появляющегося символа в двоичных единицах $H_{\text{инд}} = -\log_2 p_{\text{ч}}$	12,1	11,1	9,8
9. Средняя информация на символ в двоичных единицах $H_{\text{ср}} = -\sum_{i=1}^n p_i \log_2 p_i$	2,7	4,9	9,8
10. Переданная информация в двоичных единицах (теоретическая оценка) $H = -m \sum p_i \log_2 p_i$	40	74	147
11. Переданная информация в двоичных единицах (эмпирическая оценка) $H = \sum^m H_{\text{инд}}$	38	72	147
12. Избыточность в процентах $R = 100(1 - H_{\text{ср}})H_{\text{ном}}$	73	50	0

Результаты опытов

В результате проведенных опытов получены оценки времени запоминания и скорости запоминания.

Измерялось время запоминания в момент начала предъявления информации, момент начала восприятия и момент завершения первого правильного воспроизведения всей предъявленной информации. Ско-

рость предъявления информации составляла около 0,6 символа в секунду с некоторыми отклонениями в сторону увеличения при чтении символов, содержащих меньше информации, и в сторону уменьшения для символов с большим количеством информации.

Для оценки скорости передачи информации принимались две разные меры:

1) скорость в символах в секунду как отношение 15 символов ко времени запоминания в секундах:

$$C_{\text{симв/сек}} = m/t \text{ сек.}$$

2) скорость в двоичных единицах в секунду, которая оценивалась как отношение суммы 15 индивидуальных значений информации для каждого переданного символа ко времени запоминания в секундах:

$$C_{\text{дв. ед/сек}} = \frac{\sum_{\text{инд}}^m H_{\text{инд}}}{t_{\text{сек}}}$$

Опыты показали, что с увеличением средней информации на символ и уменьшением избыточности символов увеличивается время запоминания и уменьшается скорость запоминания, выраженная числом символов в единицу времени (*симв./сек*) и числом единиц информации в единицу времени (*дв. ед/сек*).

Таблица 2

Результаты опытов

Испытуемые	Время запоминания в секунду			Скорость запоминания в символах за секунду			Скорость запоминания в двоичных единицах за секунду		
	Серии опытов*			Серии опытов			Серии опытов		
	1	2	3	1	2	3	1	2	3
1	360	516	847	0,042	0,0291	0,0177	0,144	0,142	0,174
2	185	401	1800	0,081	0,0374	0,0083	0,153	0,131	0,082
3	155	387	1054	0,097	0,0388	0,0142	0,258	0,189	0,139
4	371	510	2150	0,040	0,0294	0,0070	0,108	0,144	0,068
5	70	485	1085	0,214	0,0309	0,0138	0,405	0,129	0,135
6	57	360	907	0,263	0,0417	0,0166	0,498	0,232	0,162
7	120	463	960	0,125	0,0324	0,0157	0,237	0,158	0,158
8	187	240	1255	0,080	0,0625	0,0120	0,278	0,262	0,117
9	132	605	1115	0,114	0,0248	0,0134	0,303	0,172	0,132
10	34	203	717	0,442	0,0738	0,0209	0,834	0,361	0,205
11	30	505	1800	0,500	0,0297	0,0083	0,160	0,125	0,082
12	45	194	580	0,334	0,0773	0,0259	0,631	0,378	0,254
13	135	520	1440	0,111	0,0289	0,0104	0,210	0,181	0,102
14	92	360	600	0,163	0,0417	0,0250	0,373	0,218	0,245
15	103	255	1408	0,146	0,0589	0,0106	0,503	0,287	0,104
16	158	473	1455	0,095	0,0317	0,0103	0,253	0,155	0,101
17	118	517	819	0,127	0,0290	0,0183	0,241	0,122	0,179
18	571	540	1089	0,026	0,0278	0,0139	0,132	0,136	0,136
19	108	167	914	0,139	0,0954	0,0161	0,427	0,433	0,161
20	370	140	950	0,041	0,1070	0,0158	0,172	0,227	0,155
Среднее значение	170	392	1147	0,159	0,046	0,015	0,316	0,209	0,145

* Примечание. Различия в сериях 1—2, 1—3 и 2—3 по всем показателям значимы. Применялся критерий Вилкоксона для сопряженных пар, $p < 0,01$.

Увеличение средней информации на символ в 3,6 раза и уменьшение избыточности от 73% до 0 привело к увеличению времени запоминания в девять раз и к уменьшению скорости запоминания в символах в секунду в 10 раз, а в двоичных единицах в секунду — в два раза.

Анализ результатов

Опыты подтвердили следующие полученные ранее результаты для случая, когда число символов неизменно: 1) время запоминания увеличивается с увеличением информации на символ и всей передаваемой информации; 2) скорость запоминания символов уменьшается с увеличением информации на символ и всей передаваемой информации.

Самым важным из показателей, полученных в настоящих опытах, является скорость запоминания в единицах информации, переданной за единицу времени (*дв. ед/сек*), так как эта мера скорости характеризует продуктивность процесса запоминания самого разнообразного по своему содержанию материала. Как уже было указано, наибольшее влияние на скорость в двоичных единицах за секунду, а следовательно и на продуктивность запоминания, оказывает число запоминаемых символов, с увеличением которого скорость эта значительно падает и продуктивность запоминания уменьшается.

Для случая, когда число передаваемых символов не изменяется, в предшествующих опытах были получены следующие результаты. С увеличением средней информации с 3 до 9 дв. ед. на символ скорость запоминания информации увеличивается, а с увеличением средней информации с 1 до 3 дв. ед. на символ, по крайней мере для 8 символов, наоборот, уменьшается. Последнее может объясняться либо тем, что 8 символов (по 1 дв. ед. на символ) — количество, не превосходящее объем непосредственной памяти, либо тем, что с уменьшением информации увеличивается скорость сенсорных реакций.

Если исключить редко встречающиеся в человеческой практике равновероятные последовательности бинарных символов (информация на символ равна 1 дв. ед.), то можно считать, что с увеличением информации на символ при равновероятном безызбыточном появлении символов из разных алфавитов скорость запоминания информации увеличивается.

Настоящие опыты дали другие результаты. Оказалось, что при равновероятном появлении символов из одного алфавита с увеличением информации на символ и уменьшением избыточности скорость запоминания информации уменьшалась. Таким образом, при равновероятном появлении символов одного алфавита скорость запоминания в большей мере зависит от количества информации, чем от числа символов. При этом количество информации не непосредственно определяет скорость запоминания, а направляет деятельность субъекта с определенной структурой запоминаемого материала. Этот материал в данных опытах не различался числом запоминаемых символов, но среднее число кусков запоминаемого материала или кусков информации в разных сериях было различным. Кусок информации — это объединение нескольких символов в один более крупный символ, т. е. символ образованный в результате группирования других символов. Мы полагаем, что два разных трехзначных числа не образуют куска и что способ запоминания шестизначного числа не имеет никаких преимуществ перед способом двух разных трехзначных чисел. Вместе с тем, любое число повторяющихся символов можно рассматривать как один кусок запоминаемого материала. Так,

