

УДК 531.625

## СТРУКТУРНЫЕ СВОЙСТВА ИНФОРМАЦИИ

Б. Д. Лазебник, Л. А. Шелепин

*Рассматривается взаимосвязь информации с немарковскими процессами и на этой основе анализируются структурные свойства информации. Обсуждается проблема информационного обеспечения управления.*

Современная теория информации получила широкое применение в оптимизации технических средств связи, автоматическом регулировании, системах управления, социально-экономической сфере. В физике информация была введена в анализ процессов через негэнтропийный принцип [1], а в философском плане трактовалась как некоторая третья сущность наряду с материей и энергией.

Однако в последнее время значительное внимание стало уделяться трудностям, которые, как оказалось, лежат в самых основах теории информации. Подробный анализ представлен в книге Корогодина [2], где, в частности, рассматриваются противоречия негэнтропийного принципа, отождествляющего физическую энтропию и информацию, дана критика положения о всеобщности информации и указывается на ее связь с феноменом жизни, обсуждаются понятия ценности, смысла, эффективности информации, которые пока являются чем-то внешним для теории.

Первоначально теория информации была ориентирована на анализ отдельных сообщений в каналах связи. Количество содержащейся в них информации  $I_i$  определялось через вероятность  $P_i$  появления того или иного символа  $i$  в сообщении ( $K$  – коэффициент пропорциональности)

$$I_i = -K \log P_i. \quad (1)$$

В более общем плане информация  $I$ , содержащаяся в сообщении, записывалась в виде

$$I = K \log(P/P_0), \quad (2)$$

где  $P$  – вероятность события после приема сообщения,  $P_0$  – вероятность до приема. При анализе отдельных сообщений количество информации играет определяющую роль и

можно абстрагироваться от ее структуры. Но сообщения в реальных системах – это отдельные элементы, вычлененные из крупных информационных массивов, обладающих сложной структурой. Поэтому последовательный информационный подход должен с самого начала учитывать, наряду с количественными, структурные свойства. Именно бесструктурность существовавшего подхода и обуславливала основные трудности теории.

Ключевой момент здесь заключается в том, что информация возникает как неотъемлемое свойство немарковских процессов, изучение которых находится еще в начальной стадии. В настоящее же время в основе современной парадигмы и прежде всего классической и квантовой физики лежит анализ марковских процессов, или процессов без последствия. Последние характеризуются тем, что, зная состояние системы в какой-либо момент времени  $t_0$ , можно, в принципе, определить вероятностную картину поведения системы в будущем, причем эта картина не изменяется от дополнительных сведений о событиях при  $t < t_0$ .

Для немарковских процессов характерна зависимость от предыстории, от событий при  $t < t_0$ . Знание о прошлом меняет вероятности в настоящем. Информация коррелирует процессы и направляет будущее в определенное русло. Сигналы из прошлого могут идти разными путями. Получая и преобразуя их, можно изменять спонтанный ход событий. Символически это отражается оператором [2]

$$(R) \xrightarrow{Q}_{p,P} (Z), \quad (3)$$

где  $R$  – исходная ситуация,  $Z$  – одно из событий с вероятностью осуществления  $p$ ,  $Q$  – ”механизм”, или оператор, приводящий к возрастанию вероятности достижения  $Z$  до значения  $P$ . Т.е. можно говорить об определенном действии, отличающемся от спонтанного наличием оператора  $Q$ . Совокупность приемов, правил, сведений, необходимых для построения оператора  $Q$ , названа в [2] информацией. Эта совокупность должна обладать определенной структурой. Ее сложность возрастает с усложнением оператора (информационной системы, организма).

Немарковские процессы включают структурные изменения, причем, в отличие от синергетики, не отдельные переходы между конкретными структурами, а преобразования совокупности структур, т.е. цепочки взаимосвязанных структурных переходов в единой системе. Если марковские процессы определяются начальными условиями и дифференциальными уравнениями, то немарковские – интегродифференциальными уравнениями для функции распределения  $f(x, t)$  по пространству и времени (оператор  $\hat{L}$  и функция

$g$  определяются конкретным процессом)

$$\frac{\partial f(x, t)}{\partial t} = \int \hat{L}[f(x, t - s)]g(s)ds. \quad (4)$$

Характерная особенность этих уравнений – нелокальность во времени, приводящая в конечном счете к процессам, включающим операции с информацией.

Для немарковских процессов можно выделить три типа равновесных распределений, имеющих, по существу, эмпирический характер. Распределения первого типа связаны с негэнтропией  $S$ , являющейся мерой сложности и упорядоченности структур. По своей форме они аналогичны больцмановскому распределению по энергии для марковских процессов

$$W = W_0 \exp(-S/\theta). \quad (5)$$

Здесь  $W$  – конкретный параметр, например, биомасса конкретных видов в биоценозе [3], величина  $\theta$  – аналог температуры.

Второй тип равновесных распределений связан с известным золотым сечением и числами Фибоначчи. Как показано в [4], этим распределениям соответствует определенное значение  $\theta$  в (5).

Третий тип распределений – это распределения Ципфа – Парето, изучение которых имеет почти вековую историю [5]. Получен целый массив самых разнообразных эмпирических зависимостей: числа людей от получаемых ими доходов, числа видов от числа особей, числа биологических родов от количества видов в них, числа хозяев от количества паразитов в них, числа ученых от количества написанных ими статей и т.п. Оказалось, что все эти зависимости имеют гиперболический характер

$$P(x) = A/x^\alpha. \quad (6)$$

Обычно  $1 < \alpha \lesssim 2$ , хотя для целого ряда видов человеческой деятельности  $\alpha$  может лежать вне этих пределов [5]. Распределение (6) преобразованием к новой переменной  $S = \ln x$  приводится к зависимости (5). Величина  $x$  играет роль вероятностного параметра, логарифм которого соответствует негэнтропии.

Таким образом, три рассмотренных типа распределений для немарковских систем – это разные проявления единого равновесного распределения, отображающего иерархию структур, распределение их набора по степени сложности. Важным его свойством является устойчивость, т.е. композиция двух таких распределений приводит к распределениям того же типа. Имеется три устойчивых распределения, выражающихся через

элементарные функции: гауссово и Коши – симметричные, двухсторонние, используются в физике; рассматриваемое гиперболическое распределение – асимметричное, одностороннее и описывает биологические, экономические и социальные процессы, а также практически любые процессы и явления, связанные с человеческим поведением.

Информационные структуры также характеризуются негэнтропией  $S$ . Распределение количества информации задается соотношением

$$I = \exp(-S/\rho). \quad (7)$$

Величина  $\rho$ , подобно  $\theta$  в уравнении (5), служит индикативным показателем состояния информационной системы. Распределение (7) может быть использовано при составлении набора проектной документации, в информационном обеспечении управления. Как показал анализ информационных потоков в системе управления [6], в среднем для подразделений разного уровня 35% всего объема информации являются лишними и не используется. Более всего лишних показателей получают подразделения высшего уровня – 76% всего объема информации. Для эффективного информационного обеспечения управления необходимо соответствие объема информации распределению (7).

Распределения (5) – (7) характеризуются резкой асимметрией. Например, в распределении (6) для ученых по числу написанных ими статей основная часть – малопродуктивна. Из общего числа авторов 10% высокопродуктивных пишут половину статей. Это свойство концентрации позволяет проводить эффективное управление экономикой на основе прямого регулирования предприятий и объединений с высоким значением  $S$ , т.е. крупным и сложным производством. Для основной массы предприятий достаточны относительно мягкие методы регулирования. В соответствии с соотношением (6) и наблюдающимся возрастанием сложности производства, характеризуемой  $S$ , общая тенденция современного развития экономики – постепенное уменьшение числа производственных объединений и фирм, выпускающих основную массу внутреннего валового продукта.

Эффект концентрации может быть использован и в информационном воздействии. Существует иерархическое распределение понятий, и среди них ведущие, обладающие высоким значением  $S$ . Этим ведущим немного, но вокруг них и в зависимости от них концентрируются другие понятия. Поэтому успех воздействия средств массовой информации может быть обусловлен концентрацией именно на этих понятиях.

Заметим, что использование (7) позволяет снять отмечавшееся выше противоречие негэнтропийного принципа Бриллюэна [1], связывающего физическую негэнтропию с

информацией, т.е. объекты разной природы, структурные и количественные характеристики.

Особенность немарковских процессов – возникновение фактора воздействия, обусловленного прошлым и меняющего взаимодействия в системе. В результате развивается процесс, обусловленный появлением нелинейных вероятностей типа [7]

$$P[(x + 1) \rightarrow x, \Delta t] \sim \lambda x \Delta t \quad (8)$$

(где  $\lambda$  – константа,  $x$  – характерный параметр) и приводящий к возникновению положительной обратной связи. Динамика процессов установления равновесных распределений в информационной системе связана с понятием поризма. Поризм в [2] определяется как утверждение, оформившееся в ходе решения какой-либо задачи, которое по своему содержанию охватывает более широкий круг явлений, чем относившийся к первоначальной задаче. По существу, это переход на более высокий уровень негэнтропии. Такой переход становится возможным в результате перекомбинирования информации в системе. Здесь имеется аналогия с установлением равновесного бальцмановского распределения по энергии при ее накачке на нижние уровни атомов и молекул. Возникновение новых идей, творчество являются, по существу, одним из проявлений динамики установления равновесия, когда при избытке информации с малым значением  $S$  в результате ее преобразования возникает распределение типа (7), содержащее более высокие значения  $S$ .

При анализе структурных свойств информации выше была использована обобщенная характеристика – негэнтропия  $S$ , но для многих практических приложений существенное значение имеют конкретные структуры информационных систем, имеющих свой адекватный язык. В принципе, для любой системы можно выделить три типа информационных структур: отражающую внешний мир, внутрисистемную, прогностическую. Каждая из них строится адекватно своему назначению. В экономике им соответствуют научно-техническая, технико-экономическая, планово-экономическая информации. В работе [8] было показано, что управляющая информация должна включать все три указанные типа структур. В настоящее же время информационное обеспечение управления строится на основе структур научно-технической информации. Последовательное использование немарковского подхода и структурных свойств информации открывают новые возможности в информатике, где мы имеем дело с крупными информационными массивами.

Авторы выражают благодарность РФФИ (грант N 97-06-80045) за поддержку.

ЛИТЕРАТУРА

- [1] Бриллюэн Л. Наука и теория информации. М., Физматгиз, 1960.
- [2] Корогодин В. И. Информация и феномен жизни. Пущино, 1991.
- [3] Быстрова Т. В., Шелепин Л. А. Труды ФИАН, **218**, 60 (1994).
- [4] Харитонов А. С., Шелепин Л. А. Краткие сообщения по физике ФИАН, N 7-8, 79 (1996).
- [5] Петров В. М., Яблонский А. И. Математика и социальные процессы. М., Знание, 1980.
- [6] Мамиконов А. Г. Управление и информация. М., Наука, 1975.
- [7] Феллер В. Введение в теорию вероятностей и ее приложения. М., Мир, 1967.
- [8] Лазебник Б. Д. Научно-техническая информация, N 6, 1 (1993).

Поступила в редакцию 22 мая 1997 г.