

СПИСОК ЛИТЕРАТУРЫ

1. Cooper M.D. The structure and future of the information economy // Information Proceeding and Management. — 1983. — V. 19. — P. 9-26.
2. Machlup F. Knowledge: its creation, distribution and economic significance. — Vol. 1 // Knowledge and Knowledge Production. — Princeton, 1980.
3. Takahuru H. Value and cost of information in terms of accounting // Mem. Fae. Kyus hu Univ. — 1984. — V. 44, N. 2. — P. 191-212.
4. Leeson W. Information policy: national strategies, international effects // Telemat. and Inf. — 1984. — V. 1, N. 4. — P. 395-400.
5. Turn R., Shapire N. Privacy and security in databank systems measures of effectiveness, costs and protector // Intruderactions. RAND Corporation Memo P. — 4871.
6. Hoffman L.J. Constructing security ratings for computer systems // Proceedings of the 1974 IEEE, National Telecom. Conference, San Doc Diego, CA.
7. Morehead D.R., Peitersen A.M., Rouse W.B. The value of information and computer-aided information seeking: problem formulation and application to fiction retrieval // Information Proceeding and Management. — 1984. — V. 20, N. 5-6. — P. 583-601.
8. Belis M., Guiasu S. A quantitive-qualitive measure of information in cybernetic systems // IEEE Transactions on Information Theory, IT-14. — 1968. — N. 3-4. — P. 593-594.
9. Man Mohar, Mitter J. On bounds of useful information measures // Judian Journal of Pure and Applied Mathematics. — 1979. — V. 9, N. 9 — P. 960-964.
10. Longo G. Quantitive-qualitive measure of information. — Springer-Verlag, N.Y., 1980.
11. Sharma B.D., Mitter J., Man Mohar. On generalized "useful" information // JCISS, 3, 1978. — P. 82-93.
12. Викерс П.Х. Подход стоимостного анализа к национальным информационным системам. — ЮНЕСКО, 1978.
13. Стратанович Р.А. Теория информации. — М.: Сов. Радио, 1975.
14. Шеннон К. Работы по теории информации и кибернетике. — М.: ИЛ, 1963.
15. Плэт В. Информационная работа стратегической разведки. — М.: ИЛ, 1958.

Семантическая информация: новый синтез точных и гуманитарных наук

Ю.П. Шанкин

Информационные потребности современного общества объективно стимулируют ускоренное развитие передовых информационных и телекоммуникационных технологий, что находит отражение в развертывании крупных государственных программ, таких как программа национальной информационной инфраструктуры в США (НИИ или программа "Клинтона-Гора"), президентская программа создания информационно-телекоммуникационной системы специального назначения в Российской Федерации, а также разрабатываемая федеральная программа информатизации России. В свою очередь большой и все возрастающий объем информации, предлагаемый потенциальным ее потребителям информационными системами разного рода, приводит к необходимости качественного совершенствования смысловой обработки и анализа исходной информации. Задача обработки больших массивов информации приобретает, образно говоря, экологическую окраску, так как рост объема невостребованных или избыточных данных является, по образному выражению С.Бира современным источником загрязнения окружающей среды.

Принципиально новое значение приобретает понятие безопасности применительно к информационным технологиям и собственно информации. Обеспечение достоверности и защищенности от несанкционированного доступа к информации в информационных сетях это весьма актуальная задача не только по отношению к конфиденциальной, но и к деловой информации в целом. Сегодня технологическая значимость информационного оружия (термин, интегрирующий в себе вопросы несанкционированного доступа к линиям и сетям связи, перехвата, использования, изменения информации и воздействия на систему управления сетями передачи данных, глобальными и локальными компьютерными сетями) фактически заняла "экологическую нишу" технологий эпохи программ так называемых звездных войн. Однако и здесь совмещение архитектуры взаимодействия открытых сетей с требованиями безопасности распределенных информационных систем представляет весьма сложную научно-техническую задачу, также тесно связанную с оценками объективной значимости сохраняемой информации. (Так, в рамках программы НИИ Агентство национальной безопасности США разрабатывает методику оценки степени риска доступа к информации.)

Семантический (смысловой, ценностный) подход к информации, есте-

ственный для любой конкретной области знаний, как правило не поддается строгой формализации и теряет содержательность при попытках распространить его на классическую теорию информации в целом (как писал ранее Шеннон: "The semantic aspects of communication are irrelevant to the engineering problems" [1] "Семантические аспекты связи не имеют отношения к инженерным проблемам.") Дихотомия понятия семантической информации связана, с одной стороны, с неизбежным личностным (т.е. в определенном смысле субъективным) восприятием и интерпретацией полученной комбинаторной (шенноновской) информации. С другой стороны, объективное содержание информации предполагает наличие общих законов ее восприятия, а следовательно и ставит содержательную научную задачу изучения этих законов. До недавнего времени столь "экзотические" вопросы не выходили за пределы гуманитарных дисциплин и рассматривались, в основном, в работах по когнитивной психологии, эпистемологии или философии. При этом кажущаяся доступность терминологии и общность суждений привлекали к этой теме многих заинтересованных наблюдателей, что в определенной степени способствовало пассивности так называемых "тяжелых ученых", т.е. специалистов в области точных естественных наук (часто предпочитающих эпигонскую строгость угрозе возможной профанации точных дисциплин).

Положение качественно изменилось с появлением новой научной парадигмы — детерминированного хаоса, позволяющего с математической строгостью говорить о существовании неслучайных систем с непредсказуемым поведением. В работах по "прикладному хаосу" описаны многочисленные примеры подобного рода систем, имеющих место в экономике, физике, социологии, медицине и других областях науки и техники (например [2,3]). Сочетание преемственности развития с непредсказуемостью действия (обусловленное одновременным существованием в системе как положительных, так и отрицательных показателей Ляпунова) открывает новую возможность компромисса точных и гуманитарных наук, позволяет структурировать процесс восприятия семантической информации, но не претендует на его точное количественное описание.

Расширение понятийного запаса теории информации для работы с семантическими структурами может происходить на базе методов анализа нелинейных самоорганизующихся диссипативных систем и развивающихся математических подходов к исследованию управляющих систем с элементами рефлексии. Семантическое восприятие информации может рассматриваться как динамический процесс самоорганизации в сложной распределенной нелинейной среде (мозге, являющемся, по-видимому, сложной аналоговой машиной, а не суперкомпьютером фон-нейманского или параллельного типа), аттракторами которого в соответствующем конфигурационном пространстве являются типичные картины-образы. По словам О.Е.Ресслера [4], "природа предпочитает гештальт", в том смысле, что хаотические аттракторы дают качество, которое наглядно и дискретно отличается от других качеств (имеются в виду фазовые портреты динамических образов и созвучность фраз

"portrait de face" и "portrait de phase"). Так, не только создание, но и восприятие литературного или музыкального произведения является творческим процессом, связанным с формированием соответствующего (фазового) образа. Вместе с тем, дискретно-формализованное представление этого произведения (текст, партитура) может рассматриваться как модифицированное возвратное отображение Пуанкаре соответствующего аттрактора, однозначно представляющее это произведение, но, одновременно, оставляющее свободу для его индивидуального творческого восприятия. Интересно также отметить, что, как показал Николис [5], даже элементарная модель дискретного хаотического процесса, порождаемого логистической зависимостью (типа $x_{n+1} = a * x_n * (1 - x_n)$), дает качественно правильное описание лингвистического закона формирования слов различной протяженности (известного как статистический закон Zipf).

Наконец, с прагматической точки зрения современных информационных технологий странный аттрактор может рассматриваться как некий хаотический процессор, обладающий большим программным (software) репертуаром, осуществляющий сжатие информации и, таким образом, экономящий объем необходимой памяти, что количественно определяется разностью между минимальной размерностью вложения и фрактальной размерностью аттрактора в битах на точку в фазовом пространстве. Практически неограниченная потенциальная информационная емкость странного аттрактора характеризуется так называемой структурной энтропией (или энтропией Колмогорова-Синяя), приводящей к линейному по времени росту объема необходимой информации о положении системы на поверхности аттрактора.

Как отмечалось, семантическое восприятие информации происходит персонализированно субъектом (лицом, принимающим решение — ЛПР), поэтому, в отличие от психологической трактовки "бихейвиоризма", избегающего анализа внутренней организации субъекта, концепция "семантической информации" становится автологической, то есть теорией второго порядка, относящейся не к предмету (объекту), а к лицу-наблюдателю, его представлениям об объекте и их самооценке со стороны наблюдателя [6].

Представим ЛПР в виде "нетривиальной машины" (по терминологии Фэрстера: Non-Trivial-Machine, NTM), вход, внутреннее состояние и выход которой взаимосвязаны между собой (рисунок 1.). Предметом наблюдения Q со стороны ЛПР может быть внешний объект, процесс, формализованная запись результатов научного наблюдения или тексты на естественном языке. Восприятие (перцепция) предмета наблюдения определяется вектором ощущений ЛПР \vec{P} (ощущения могут определенное время сохраняться при исчезновении внешнего воздействия) при условии заданного на данный момент внутреннего состояния ЛПР, характеризующего его эмоциональный уровень, сосредоточенность, физическую или психологическую готовность к восприятию данного предмета и т.п.

Предание смысла ощущениям (появление интенции ЛПР, i) это весьма сложный процесс, зависящий от "внутреннего репертуара" ЛПР, т.е. его

образования, эрудиции, интеллектуальных способностей (так, например, экспериментатор способен увидеть в результатах эксперимента лишь то, возможность чего он априори допускал, хотя бы гипотетически). Предполагая выполненными необходимые метрические условия в соответствующих пространствах ощущений и представлений ЛПР, обозначим этот процесс как некоторую "гештальт- функцию" $G_g(\vec{P}|\vec{L})$, где параметр (g) играет роль параметра порядка (по терминологии Г.Хакена). В свою очередь, внутреннее состояние наблюдателя, описываемое вектором, может зависеть от результатов осмысления исходных ощущений ($\vec{L} = \vec{L}_{(i)}$), что может выражаться в обострении внимания ЛПР, появлении ассоциаций или изменении точки зрения на исходные данные.

Таким образом, возникает повторная перцепция и образуется итерационная последовательность интенций типа:

$$i_{n+1} = I(i_n) \quad (1)$$

Как показал Фейгенбаум [7] итерационные последовательности подобного рода в случае нелинейного (унимодального в локальной области) вида функции $I(i)$ приводят к трем возможным вариантам их завершения:

- к сходимости процесса (1);
- к возникновению бифуркаций и соответствующих циклов (типа $i(x^*) = x^{**}; i(x^{**}) = x^*$ или более высокой кратности);
- к возникновению хаотического режима поведения.

Количественные оценки в данном случае вряд ли возможны из-за отсутствия в явном виде используемых параметров, векторов и функций (причем в случае хаотического режима численное моделирование становится принципиально невозможным в связи с его повышенной чувствительностью к точности задания исходных данных).

В то же время структурная классификация представляется вполне уместной, при этом само отображение (1) может трактоваться как уже упоминавшееся отображение Пуанкаре непрерывного динамического процесса, происходящего в мозге наблюдателя при семантическом восприятии информации. Таким образом можно выделить три семантические области:

- Скалярная семантическая область (область I на рисунке). Характеризуется сходимостью итерационного процесса, отвечает области классических задач распознавания образов при отсутствии амбивалентных объектов.
- Векторная семантическая область (область II на рисунке). Связана с процессами, нарушающими исходную симметрию (поступок как выбор того или иного участка после бифуркации или последовательности бифуркаций). Это область принятия решений (со стороны ЛПР), приводящая к возникновению новой содержательной информации в процессе наблюдения или взаимодействия с внешним миром.
- Хаотическая семантическая область (область III на рисунке). Связана с эволюционными процессами (размышлениями). В случае топологического

соответствия странного аттрактора в конфигурационном пространстве размышлений наблюдателя и исходного аттрактора, связанного с наблюдаемым объектом или процессом (в качестве которого, в частности, может быть взято художественное или музыкальное произведение, опосредованно несущее печать подобных авторских размышлений) открывается возможность творческого восприятия объекта (процесса) со стороны ЛПР, т.е. возможность самостоятельного появления новой содержательной информации при прекращении воздействия внешнего раздражителя. Эта область является основой теоретического изучения внешнего мира и его эстетического восприятия.

Строго говоря, имеет место еще одна область — область семантической индифферентности, когда итерационный процесс (1) не возникает (отсутствуют действительные корни уравнения $i = I(i)$). Данная область обусловлена неготовностью наблюдателя к осмыслению воспринимаемой им (шенноновской) информации. Однако положение может измениться при модификации "гештальт-функции" наблюдателя, что является целью профессиональных воспитательных и образовательных общественных институтов.

Автологическое моделирование процесса восприятия семантической информации касается не только представлений наблюдателя (ЛПР) об объекте наблюдения, но и его представлений о себе самом, а также оценки (самооценки) этих представлений. В случае социологического или психологического эксперимента, когда объект наблюдения, также обладает рефлексией, модель ситуации может включать в себя представления ЛПР о представлениях объекта и т.д. То есть структура модели становится многоуровневой и, хотя результаты попыток формализованного описания подобных ситуаций напоминают "улыбку чеширского кота", в отдельных случаях (например, как показал В.Лефевр при выборе предиката, связывающего различные уровни рефлексивной модели, в виде импликации или антиимпликации) возможно получение нетривиальных содержательных результатов моделирования [8,9]. Причем для анализа рефлексивных систем и описания представлений ЛПР более адекватным является аппарат нечеткой математики, первоначально предложенный Л. Заде для работы именно с "лингвистическими переменными". Подобный подход, учитывающий возможность зависимости, возникающую в мозге наблюдателя, образа объекта от мнения окружающих или этических установок общества в целом, в большей степени значим для хаотической области, наиболее чувствительной к возмущениям внутреннего мира наблюдателя. Вместе с тем известны психологические эксперименты с "навязыванием" субъекту (ЛПР) мнения окружающих в достаточно тривиальных случаях распознавания простых объектов (т.е. в скалярной семантической области).

Функциональная зависимость типа (1) может быть многомерной, т.е. может связывать различные компоненты вектора интенции (размерность которого определяется конкретными биофизическими процессами, протекающими в мозге наблюдателя. Кроме того, как отмечалось выше, отображение Пуанкаре понижает размерность реального динамического непрерывного про-

цесса, переводя его в режим так называемого дискретного времени. Особенностью многомерных систем является множественность аттракторов (в том числе — странных аттракторов, отвечающих различным режимам детерминированного хаотического поведения), а конкретное положение системы в фазовом пространстве зависит от начальных условий, определяемых в нашем случае исходными ощущениями и начальным "состоянием" наблюдателя. Таким образом, в многомерном случае может быть реализована селективная память, соответствующая различным семантическим областям восприятия информации.

В качестве примера на рисунке 2 приведены возможные зоны "притяжения" (т.е. аттракторы) в фазовой плоскости двумерного отображения Хеннона:

$$\begin{cases} x_{n+1} = A - Bx_n^2 + y_n \\ y_n = Cx_{n+1} \end{cases}$$

при $A = 2,0$, $B = 0,5$ и $C = 0,5$, отвечающие различным областям начальных условий:

$$x_0 = 0,5; y_0 = 0,5 \text{ (рисунок 2a)}$$

$$x_0 = 0,4; y_0 = 1,0 \text{ (рисунок 2б)}$$

Динамика систем, находящихся в состоянии детерминированного хаотического поведения, в частотной области подобна шумовому процессу, т.е. характеризуется непрерывным спектром. Вместе с тем, существуют численные алгоритмы (например, предложенные П.Грассбергером и И.Прокачиа [10]), позволяющие по анализу конечной выборки выявлять функциональную природу наблюдаемого хаотического процесса путем определения фрактальной размерности соответствующего странного аттрактора динамической системы или минимальной размерности его вложения. При этом анализу подвергаются результаты конечного числа измерений непрерывного процесса.

Применительно к произвольным конечным символьным цепочкам (например, в виде текста), подобные алгоритмы определяют размерность (величину) максимальной повторяющейся последовательности символов в тексте. В этом случае последовательности большей размерности будут однозначно определять свое место в заданном тексте, т.е. имеет место своего рода функциональная зависимость всех последующих символов текста от таким образом найденного числа заранее заданных символов. Данный результат, однако, не имеет отношения к хаотическим системам или семантике, а лишь определяет синтаксические свойства заданного текста.

Вместе с тем, многомерный процесс типа (1) применительно к дискретным (булевым) переменным может породить режим "поведения" (т.е. вычислительный режим как функцию времени), обладающий свойствами, не содержащимися в исходно запрограммированных элементах. Как отмечается в [11], информационно-осмысленные символические структуры могут генерироваться соответствующим нелинейным механизмом, находящимся в удаленной от равновесия хаотической области.

Рис. 1. ЛПР-NTM

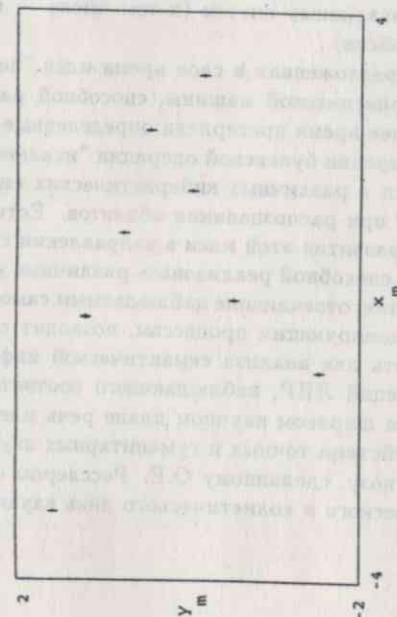
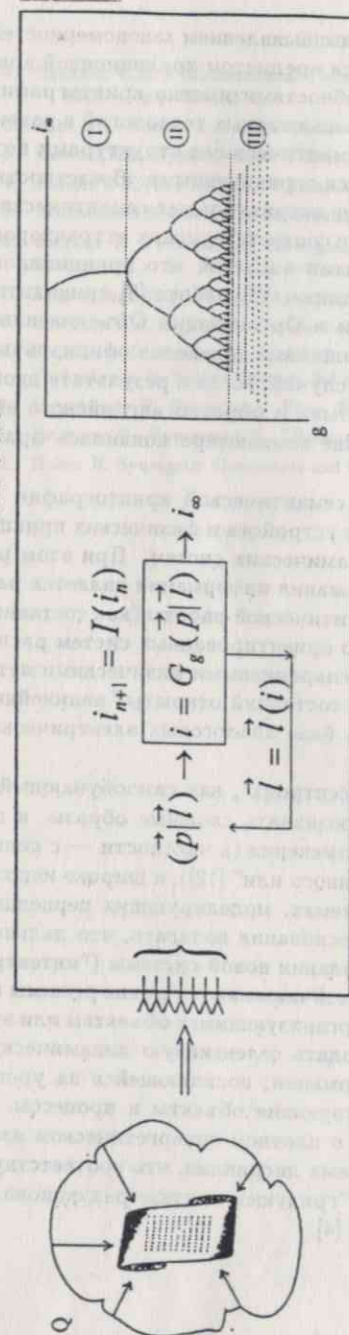


Рис. 2а.

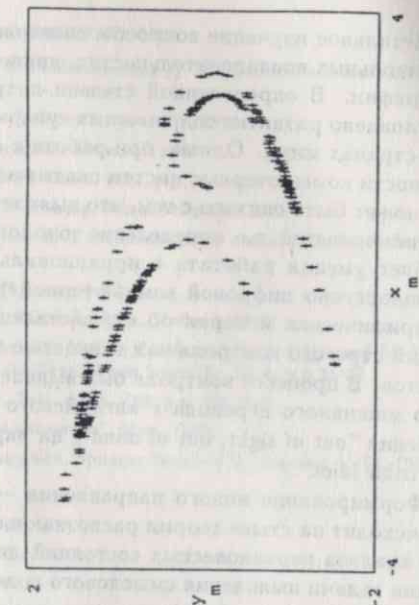
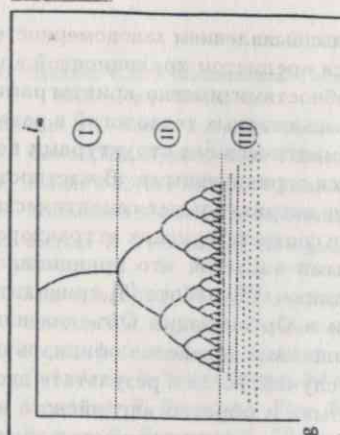


Рис. 2б.

Детальное изучение вопросов, связанных с выявлением закономерностей в символьных последовательностях, является предметом традиционной криптографии. В определенной степени потребностями именно криптографии обусловлено развитие современных суперкомпьютерных технологий в развитых странах мира. Однако при работе с семантическими структурами возможности компьютерных систем оказываются ограниченными. В частности, это может быть связано с тем, что выявление нетривиальных семантических закономерностей, т.е. определение топологии соответствующих аттракторов, требует умения работать с иррациональными числами, что принципиально недоступно цифровой компьютерной технике. (В работе [3] приводится апокрифическая история об осуществлении в Организации Объединенных Наций строгого контроля над точностью машинных переводов официальных текстов. В процессе контроля был выявлен случай, когда в результате двойного машинного перевода с английского языка и обратно английского выражения "out of sight, out of mind", на экране компьютера появилась фраза "Invisible idiot".)

Формирование нового направления — семантической криптографии — происходит на стыке теории распознающих устройств и физических принципов анализа неравновесных состояний динамических систем. При этом решение задачи выявления смыслового содержания информации является развитием традиционной информационно-аналитической работы (как составной части системной криптографии) и машинно ориентированных систем распознавания образов в направлении синтеза с современными физическими методами изучения неравновесных устойчивых состояний открытых нелинейных диссипативных систем (в том числе — на базе аналоговых электрических устройств).

Предложенная в свое время идея "персептрона", как самообучающейся кибернетической машины, способной распознавать сложные образы, в последнее время претерпела определенные изменения (в частности — с целью реализации булевой операции "исключенного или" [12]), и широко используется в различных кибернетических системах, моделирующих перцепцию ЛПР при распознавании объектов. Есть основания полагать, что дальнейшее развитие этой идеи в направлении создания новой системы ("интенетрона"), способной реализовать различные устойчивые хаотические режимы поведения, отвечающие наблюдаемым самоорганизующимся объектам или эволюционирующим процессам, позволит создать селективную динамическую память для анализа семантической информации, появляющейся на уровне интенций ЛПР, наблюдающего соответствующие объекты и процессы. В более широком научном плане речь идет о плотном синергетическом взаимодействии точных и гуманитарных научных дисциплин, что соответствует прогнозу, сделанному О.Е. Ресслером о "грядущем синтезе редукционалистического и холистического лика науки" [4].

СПИСОК ЛИТЕРАТУРЫ

1. Shannon C.E. A Mathematical Theory of Communication // The Bell System Journal. XXVI. — 1948. — p.p. 370-423.
2. Kim J.H., Stinger J. Applied Chaos. — J. Wiley & Sons, NY, 1992.
3. Cambel A.B. Applied Chaos Theory, Academic Press, NY, 1993.
4. Rossler O.E. The Future of Chaos, Applied Chaos, 1992, p.p. 457-465.
5. J.S.Nicolis J.S. Chaos and Information Processing, World Scientific, 1991.
6. Foerster H. Principles of self-organization in a sociomanagerial context, Springer Series in Synergetics, v. 26, 1984, p.p. 2-25.
7. Feigenbaum M. // Statist Phys, 19 (1978), p.p. 25-52.
8. Lefebvre V.A. Algebra of Conscience, D.Reidel Publ.Co., Holland, 1982.
9. Shankin Y.P. Autological Model of a Decision Making Problem / Fuzzy Logic and Intelligent Technologies in Nuclear Science, Ed.by Da Ruan, World Scientific, 1994, p.p.34-37.
10. Grassberger P., Procaccia I., Phys. Rev. Lett., 1983, V50, p.p. 346-349.
11. Николис Г., Пригожин И., "Познание сложного", Мир, 1990.
12. Haken H. Synergetic Computers and Cognition, Springer Series in Synergetics, V. 50, 1991.