



К ВОПРОСУ О РАЗВИТИИ СЕМАНТИЧЕСКОЙ ТЕОРИИ ИНФОРМАЦИИ

Г. К. Лежава, И. Ш. Камкамидзе, З. И. Верикишвили

Институт кибернетики Грузинской академии наук, г. Тбилиси

Рассмотрен подход к решению проблемы развития семантической теории информации, основанный на применении ассоциативной системы индуктивного вывода как приемника информации и позволяющий дифференцировать два вида информации: “образующую базу знаний” информацию и “текущую актуальную” информацию. Определена количественная мера для этих видов информации.

Для развития семантической теории информации необходимо располагать возможностью описания приемника информации, преобразующего поступающие на его вход сигналы в понятия и образы. Такой подход [1] обуславливает сильную зависимость теории от “субъективного фактора”, от конкретного приемника информации и от его текущего состояния. На первый взгляд, это обстоятельство обесценивает подход, делает его малоэффективным.

Ниже будет сделана попытка показать, что данный подход продуктивен и дает возможность исследовать ряд интересных вопросов. В частности, он дает возможность четко разграничить два вида информации. Назовем их “текущей актуальной” информацией и информацией, “образующей базу знаний”.

Четкое разграничение не означает противопоставление или взаимоисключение этих двух видов информации. Один и тот же сигнал, действующий на входе приемника, может одновременно содержать как текущую актуальную, так и образующую базу знаний информацию. Поэтому в некоторых случаях целесообразно говорить не о двух видах, а о двух составляющих информации.

Подход дает возможность не только разграничить названные два вида информации, но и ввести их точную количественную меру.

Как видно из названия, текущая актуальная информация позволяет системе, содержащей приемник информации, выбрать адекватную сенсорному воздействию реакцию из своего репертуара реакций. При этом не происходит какого-либо воздействия на базу знаний, на “опыт”, накопленный приемником; не происходит и каких-либо изменений в его памяти. Эта информация никак не влияет на дальнейшие реакции, на дальнейшее поведение системы. В данном случае мы имеем в определенном смысле линейную составляющую информационного воздействия.

Что касается образующей базу знаний информации, то она вызывает изменение базовых знаний. Это приводит к изменению, корректировке характера последующих реакций системы. Поэтому в этом случае мы имеем нелинейную составляющую информационного воздействия.

Будем полагать, что простым приёмником информации, как и в работе [1] является оптоэлектронный процессор [2]. Процессор представляет собой ассоциативное устройство индуктивного выво-

да. В памяти ассоциативного устройства может быть задано множество $\{\vec{x}_1, \dots, \vec{x}_\alpha, \dots, \vec{x}_N\}$ векторов $\vec{x}_{1\alpha} = \{x_i\}$, $i = 1, \dots, d$, в виде координат точек в некотором d -мерном пространстве. При поступлении на вход устройства некоторого вектора такой же размерности $\hat{x} = \{\hat{x}_i\}$ на его выходе одновременно

возникает множество значений $\sum_{i=1}^d f(\hat{x}_i - x_i)_\alpha$, $\alpha = 1, \dots, N$. При соответствующем выборе вида функции f эти значения можно трактовать либо как множество расстояний от точки \hat{x} до всех заданных точек \vec{x}_α , либо как меру сходства некоторых описаний, заданных вектором \hat{x} и векторами \vec{x}_α .

Такой процессор может быть полезен при решении задач, где требуется быстрый индуктивный вывод, например, при реализации непараметрических алгоритмов распознавания, задач беспилотной навигации, задач управления сложными объектами и др.

Действующие на входе процессора сенсорные воздействия с разной степенью интенсивности возбуждают выходы, соответствующие различным понятиям и образам, — осуществляется преобразование пространства сенсорных воздействий в пространство понятий и образов. При этом характер преобразования полностью определяется предысторией, тем, что зафиксировано в памяти процессора.

Последовательность $\sum_{i=1}^d f(\hat{x}_i - x_i)_\alpha$ заменим по-

следовательностью условных вероятностей $P(\alpha/\hat{x})$, т. е. последовательностью условных вероятностей

того, что сенсорному воздействию \hat{x} приемник будет ставить в соответствие определенную последовательность понятий и образов с вероятностями

$P(\alpha/\hat{x}) \rightarrow \sum_{i=1}^d (\hat{x}_i - x_i)$. Естественно полагать, что ес-

ли значения $P(\alpha/\hat{x})$ близки друг к другу, воздействие \hat{x} можно считать мало информативным. И наоборот, если среди значений последовательности

$P(\alpha/\hat{x})$ имеется ярко выраженный максимум, воздействие \hat{x} можно считать высоко информативным.

Из приведенных рассуждений следует, что для оценки количества текущей актуальной информации, получаемой приемником при сенсорном воздействии \hat{x} , может быть применено выражение, аналогичное формуле Шеннона [3]:

$$H = - \sum_{\alpha=1}^N P(\alpha/\hat{x}) \log P(\alpha/\hat{x}). \quad (1)$$

Такое определение количества текущей актуальной информации имеет существенное преимущество. Правда, в классической теории количества информации определено с помощью аналогичной формулы, однако там фигурирует понятие “ансамбля”; необходимо иметь конечную полную таблицу вероятностей. Это не создает никаких трудностей при решении задач связи и подобных задач, т. е. задач, для которых и создавалась шенноновская теория. Телеграфный аппарат не сопоставляет поступающим сигналам какие-либо понятия и образы. Он просто печатает на бумаге символы, которые взаимно однозначно соответствуют поступающим сигналам. Поэтому классическая теория способна оценить лишь величину поступающего на вход приемника разнообразия, а содержание, закодированное в этом разнообразии, остается вне этой теории. Не менее важно и то обстоятельство, что составление конечных полных таблиц вероятностей, легко осуществляемое для классических задач связи, для подавляющего большинства процессов, связанных с обменом информацией, не только затруднительно, но и лишено какого-либо смысла [4]. В системе излагаемых представлений всегда существует конечная полная таблица вероятностей, формула (1) имеет совершенно иной смысл и может быть применена для описания значительно более широкого класса информационных процессов.

Теперь введем количественную меру информации, образующей базу знаний. Система индуктивного вывода формирует понятия и образы путем наложения отображений описаний отдельных реализаций этих понятий и образов в процессе накопления опыта или обучения.

В памяти оптоэлектронного процессора [2] запись такого описания \hat{x}_j , состоящего из d параметров, занимает отдельную строку страницы данных, содержащей d столбцов. Значения отдельных компонентов x_i записываются в виде интерполирующих функций $f(x - x_i)_j$, $i = 1, \dots, d$.

Когда $j = 1$, т. е. когда в памяти записано описание только одной реализации некоторого понятия или образа, характер оценки меры близости (или сходства) \hat{x}_j и действующего на сенсоры воздействия \hat{x} определяется видом функции f .

При дальнейшем обучении, когда в память будут введены описания K реализаций, оценка меры близости осуществляется с помощью функции

$$\frac{1}{K} \sum_{j=1}^K f(x - x_i)_j, \quad i = 1, \dots, d, \quad (2)$$

и вид этой функции определяет характер распределения интенсивностей сигналов на N выходах приемного устройства. Таким образом, после ввода описания каждой очередной реализации производится корректировка, уточнение выражения (2) и соответствующим образом корректируются значения $P(\alpha/\hat{x})$.

Процедура такой “корректировки знаний” в значительной степени схожа с процедурой построения оценки плотности распределения случайной величины методом обобщенных гистограмм (методом окон Парзена) [5, 6]. Отличается она тем, что дисперсия интерполирующих функций остается постоянной, в то время как в случае окон Парзена при $j \rightarrow \infty$ она медленно стремится к нулю. Можно показать, что в случае постоянной дисперсии процедура ведет не к построению функции плотности распределения, а к свертке функции плотности распределения и интерполирующей функции.

Так как на практике почти всегда мы имеем дело с малыми значениями K , результаты, получаемые с помощью процедур “корректировки знаний” и окон Парзена, получаются идентичными. Отметим, что процедура “корректировки знаний” в отличие от процедуры Парзена, по своей сути не асимптотическая. Поэтому она требует своей нестатистической интерпретации.

Таким образом при вводе в память каждого последующего описания происходит “корректировка знаний”. Коррекцию δ' для каждой i -й компоненты можно выразить следующим образом:

$$\delta' = \left| \frac{1}{K} \sum_{j=1}^K f(x - x_i)_j - \frac{1}{K+1} \sum_{j=1}^{K+1} f(x - x_i)_j \right|,$$

а полную коррекцию —

$$\delta = \sum_{i=1}^d \left| \frac{1}{K} \sum_{j=1}^K f(x - x_i)_j - \frac{1}{K+1} \sum_{j=1}^{K+1} f(x - x_i)_j \right|.$$

Правомерно считать, что чем больше δ , т. е. чем больше влияние на базу знаний оказывает используемое при обучении описание \hat{x}_j , тем больше информации несет оно для данного приемника. Следовательно, величина δ может служить мерой информации, образующей базу данных.

ЛИТЕРАТУРА

1. Lezhava G., Tevzadze R. For the Subject of Estimation of Semantic Information // 7-th international multiconference “System engineering, cybernetics and computer science” 2003, Orlando, Florida, USA.
2. Lezhava G. Optical Computations and Artificial Intellect // Ibid.
3. Shannon C. E. A Mathematical Theory of Communication // Bell System Techn. J., 1948. — Vol. 27, № 3—4.
4. Marian Masur. The Qualitative theory of the information. — Warsaw, 1970 (on Polish).
5. Parzen E. On estimation of a probability density function and mode // Ann. Math. stat. 1962. — Vol. 33, № 3.
6. Duda R. O., Hart P. E. Pattern Classification and Scene Analysis. — New York, 1973.

E-mail: glezhava@hotmail.com

□