УДК 519.7

МЕРЫ ВНУТРЕННЕЙ И ВНЕШНЕЙ ИНФОРМАЦИИ (НА ПРИМЕРЕ ВЕРОЯТНОСТНЫХ СИТУАЦИЙ НЕОПРЕДЕЛЕННОСТИ). ЧАСТЬ II

Н.Н. ДИДУК

Продолжено изучение “настоящих” преобразований информации и построение соответствующих мер внешней информации, предназначенных для измерения *интенсивности преобразований*. Рассмотрены преобразования *концентрация информации*, *образование первой и второй проекций двумерной ситуации* и *ограничение разнообразия*. Последнее из них занимает особое место среди всех преобразований информации. Показано, что любое событие в Природе, вызывающее ограничение разнообразия в произвольной ситуации неопределенности, связанной с состояниями некоторой реальной системы, *содержит информацию об этой системе*. Образующиеся таким образом *информационные связи* между различными системами могут стать са*мостоятельным инструментом познания*, альтернативным к традиционному инструменту познания, каковым являются причинные связи. Введено понятие *ситуации заблуждения*, которое используется для вычисления количества внешней информации, требуемого для разнообразных преобразований ситуаций неопределенности.

В первой части статьи [1] показано, что положенные в основу современной кибернетики представления об информации, о способах ее изучения и об ее преобразованиях, *неверны*. Рассмотрен ряд общих вопросов, связанных с возможными способами изучения информации и ее преобразований. Впервые обнаружена непосредственная связь информации с *ситуациями неопределенности* и получены два важных вывода. Во-первых, ***нельзя изучать информацию, не рассматривая какую-либо ситуацию неопределенности*** (там, где нет неопределенности, не может быть и информации). Во-вторых, ***преобразования информации — это преобразования ситуаций неопределенности*** (а вовсе не преобразования текстов). Эти выводы опровергают основную идею, содержащуюся в концепции “переработки информации в компьютерах”, — идею, согласно которой преобразования текстов — это и есть преобразования информации.

Однако полученное опровержение порождает новую проблему: теперь необходимо начать систематическое изучение “настоящих” преобразований информации. В [1] рассмотрен первый пример “настоящего” преобразования — *квантование*. Это преобразование связано с *потерей информации*. А здесь — во второй части статьи — рассмотрены еще два преобразования этого типа — *концентрация информации* и *образование первой и второй проекций двумерной ситуации неопределенности*.

Кроме того, здесь начато изучение преобразований, в которых информация не теряется, а *приобретается* (и, следовательно, она должна откуда-то поступить). Одним из таких преобразований является *ограничение разнообразия* (в данной ситуации неопределенности). Это преобразование занимает особое место среди всех преобразований информации, так как оно позволило прояснить смысл самого понятия *информация* и обнаружить неизбежность возникновения *информационных связей* в Природе. Существование же информационных связей означает ни много, ни мало, а необходимость в радикальном изменении наших представлений о возможных способах изучения Природы. Действительно, очевидно, что информационные связи могут стать самостоятельным инструментом познания, отличным от единственного известного до настоящего времени инструмента познания *причинно-следственных связей*.

7. КОНЦЕНТРАЦИЯ ИНФОРМАЦИИ

После рассмотрения преобразования *квантование* [1, разд. 6] интересно было бы ответить на следующий вопрос: всегда ли потеря информации, связанная с квантованием, является нежелательной? Для ответа необходимо принять во внимание, с какой целью (или по какой причине) проводилось квантование. Действительно, обычно квантование рассматривают как вынужденную потерю разрешающей способности ради упрощения каких-либо процедур. Однако возможны также другие мотивы проведения квантования. Например, цель этого преобразования информации может заключаться в том, чтобы потерять *лишнюю* (ненужную) информацию и получить в результате *концентрацию* нужной информации.

1. Описание преобразования «концентрация информации». Пусть на множестве X определена некоторая функция f, область значений которой обозначим Y. Как известно, знания элемента x множества X *формально* достаточно для того чтобы узнать значение $f(x)$ этой функции. Но если элемент множества X неизвестен (т.е. на множестве X имеется ситуация неопределенности), а нас интересуют именно *значения* функции f, то может оказаться выгоднее перейти от ситуации неопределенности, заданной на множестве X, к ситуации, порождаемой функцией f на области ее значений Y. Такой переход мы будем называть **концентрацией информации** (относительно заданной функции f).

Рассмотрим преобразование *концентрация информации* (относительно функции f) для вероятностной ситуации, описываемой *пространством вероятностей* (X, p) [1, разд. 5] (где p — распределение вероятностей (РВ), *действующее* на множестве X [1, разд. 3]). Построение описания такого преобразования начинается с квантования пространства (X, p) относительно разбиения $\mathscr{X}_f$ множества X, порождаемого заданной функцией f. Известно, что всякая функция f, определенная на множестве X, порождает на этом множестве *каноническое разбиение* $\mathscr{X}_f$ следующего вида:

$$\mathscr{X}_f = \{f^{-1}(y) : y \in Y\}, \tag{1}$$

где Y — область значений функции f, а $f^{-1}(y)$ — множество всех таких $x \in X$, что $f(x) = y$.

На множестве $\mathscr{X}_f$ можно задать распределение вероятностей $\boldsymbol{p}_f$, характеризуемое условием: для каждого класса $A \in \mathscr{X}_f$ должно выполняться равенство

$$\boldsymbol{p}_f(A) = \sum_{a \in A} p(a), \tag{2}$$

аналогичное равенству [1, (6)].

Мы получили, таким образом, факторпространство $(\mathscr{X}_f, \boldsymbol{p}_f)$ пространства (X, p), описывающего исходную ситуацию неопределенности. Но этого недостаточно, так как мы хотели получить описание той ситуации неопределенности, которую функция f порождает на области своих значений Y. До получения искомого описания остался один шаг. А само это описание имеет вид пространства вероятностей (Y, q), *изоморфного* факторпространству $(\mathscr{X}_f, \boldsymbol{p}_f)$. Искомое распределение вероятностей q на множестве Y характеризуется выражением

$$q = y \mapsto \boldsymbol{p}_f(f^{-1}(y)) \Diamond Y \tag{3}$$

(РВ q представляет собой такую функцию, определенную на множестве Y, которая каждому элементу $y \in Y$ ставит в соответствие вероятность $\boldsymbol{p}_f(f^{-1}(y))$).

Действительно, нетрудно показать (опять же с помощью теории структур Бурбаки), что пространства вероятностей (Y, q) и $(\mathscr{X}_f, \boldsymbol{p}_f)$ *изоморфны*, а функция

$$g = y \mapsto f^{-1}(y) \Diamond Y \tag{4}$$

является изоморфизмом пространства (Y, q) на факторпространство $(\mathscr{X}_f, \boldsymbol{p}_f)$. Причем, для каждого $y \in Y$ имеет место равенство

$$q(y) = \boldsymbol{p}_f(f^{-1}(y)). \tag{5}$$

Таким образом, мы совершили преобразование *концентрация информации*, состоящее в переходе от ситуации неопределенности, заданной на множестве X (описываемой пространством вероятностей (X, p)) к ситуации, порождаемой функцией f на области ее значений Y. Эта новая ситуация описывается пространством вероятностей (Y, q). Переход состоял из двух шагов: 1) *квантования* исходной ситуации (описание этого шага состояло в построении разбиения $\mathscr{X}_f$ множества X, порожденного функцией f, а затем — в построении факторпространства $(\mathscr{X}_f, \boldsymbol{p}_f)$ исходного пространства (X, p)); 2) перехода к ситуации неопределенности на области значений Y функции f (эта ситуация описывается пространством вероятностей (Y, q), *изоморфным* факторпространству $(\mathscr{X}_f, \boldsymbol{p}_f)$.

2. Мера концентрации информации. Ввиду *изоморфии* пространств (Y,q) и $(\mathcal{X}_f, \boldsymbol{p}_f)$ внешнюю **меру концентрации информации** $E(X,p\,|\,f)$ (относительно функции f) можно ввести следующим образом:

$$E(X,p\,|\,f) = E(X,p\,|\,\mathcal{X}_f)\,. \tag{6}$$

Очевидно, что введенная мера удовлетворяет неравенству

$$E(X,p\,|\,f) \geq 0\,, \tag{7}$$

которое следует из неравенства [1, (13)] и изоморфии пространств (Y,q) и $(\mathcal{X}_f, \boldsymbol{p}_f)$.

8. ОБРАЗОВАНИЕ ПЕРВОЙ И ВТОРОЙ ПРОЕКЦИЙ ДВУМЕРНОЙ СИТУАЦИИ НЕОПРЕДЕЛЕННОСТИ

Эти преобразования являются частными случаями квантования. Они хорошо иллюстрируют мысль, что потеря информации при квантовании может быть полезной (и может иногда рассматриваться не как потеря, а как *концентрация* информации).

1. Описание преобразований проектирования. Пусть задана некоторая *двумерная* ситуация неопределенности (т.е. ситуация с двумерным множеством возможностей). Известно, что если задано двумерное множество, то всегда можно построить две его проекции. Так, например, первой и второй проекцией произведения $X \times Y$ являются соответственно множества X и Y (более общий случай мы здесь рассматривать не будем).

Итак, пусть имеется некоторая двумерная ситуация неопределенности, которая описывается (тоже двумерным) пространством вероятностей $(X \times Y, \pi)$ (где π есть РВ, действующее на произведении $X \times Y$ дискретных множеств X и Y). И предположим, что для любых $x \in X$ и $y, y' \in Y$ мы не хотим различать состояния природы (x, y) и (x, y'). Фактически это значит, что теперь мы хотим рассматривать в качестве нового множества возможностей множество X (вместо множества $X \times Y$). Очевидно, что в таком случае мы *автоматически* оказались в новой (одномерной) ситуации неопределенности. Это произошло независимо от того, понимаем ли мы смысл происшедшего и умеем ли мы новую ситуацию описать.

Элементарная теория вероятностей предлагает способ, позволяющий найти описание новой ситуации (однако *не предлагает* ни подходящий для этих целей язык, ни систему обозначений). Эта ситуация может быть описана (одномерным) пространством вероятностей $(X, \mathbf{pr_1}\pi)$, где $\mathbf{pr_1}\,\pi$ — РВ, действующее на множестве X , которое характеризуется следующим условием: для каждого $x \in X$ имеет место равенство

$$\mathbf{pr_1}\,\pi(x) = \sum_{y \in Y} \pi(x, y)\,. \tag{8}$$

Аналогичный *автоматический* переход к новой ситуации неопределенности нас ждет и в том случае, если в качестве нового множества возможностей мы выберем множество Y . И снова способ, позволяющий найти описание новой ситуации, можно извлечь из элементарной теории вероятностей: эта ситуация может быть описана (одномерным) пространством вероятностей $(Y, \mathbf{pr_2}\,\pi)$, где $\mathbf{pr_2}\,\pi$ — РВ, действующее на множестве Y, которое характеризуется условием: для каждого $y \in Y$ имеет место равенство

$$\mathbf{pr_2}\,\pi(y) = \sum_{x \in X} \pi(x, y)\,. \tag{9}$$

Распределения $\mathbf{pr_1}\,\pi$ и $\mathbf{pr_2}\,\pi$ мы будем называть соответственно **первой и второй проекциями** распределения π. А построенные выше (одномерные) пространства вероятностей $(X, \mathbf{pr_1}\pi)$ и $(Y, \mathbf{pr_2}\pi)$ мы назовем соответственно **первой и второй проекциями** (двумерного) пространства вероятностей $(X \times Y, \pi)$.

Итак, отталкиваясь от двумерной ситуации неопределенности, описываемой пространством вероятностей $(X \times Y, \pi)$, мы пришли к описаниям $(X, \mathbf{pr_1}\pi)$ и $(Y, \mathbf{pr_2}\pi)$ двух новых ситуаций неопределенности (которые тоже будем называть **первой и второй проекциями** исходной ситуации). Обе эти ситуации тоже оказались вероятностными (поскольку их описаниями являются пространства вероятностей).

Теперь еще раз подчеркнем специально, что нужно раз навсегда оставить распространенную моду *отождествлять* преобразование распределения π в распределение $\mathbf{pr_1}\,\pi$ (или в распределение $\mathbf{pr_2}\,\pi$) с преобразованием *реальной* двумерной ситуации в *реальную* одномерную. Действительно, выражения (8) и (9) *не создают* преобразования ситуаций неопределенности, а только являются их *описаниями*. Реальные же преобразования ситуаций (как уже отмечалось выше) происходят независимо от того, умеем ли мы их описать, т.е. известны ли соотношения (8) и (9) (и даже независимо от того, известна ли вообще теория вероятностей). Причинами же, вызывающими эти преобразования ситуаций, являются преобразования множества возможностей.

2. Образование проекций двумерных ситуаций как частный случай концентрации информации. Теперь осталось показать связь между образованием проекций и преобразованием *концентрация информации*. Для этого достаточно вспомнить, что было сказано о последнем преобразовании в разделе 7. Пусть на множестве X имеет место некоторая ситуация неопределенности. И пусть нас интересуют не элементы множества X, а значения некоторой функции f, которая определена на этом множестве. Но мы не можем узнать конкретное значение этой функции, поскольку неизвестно значение ее аргумента (неизвестно состояние природы $x \in X$). Иначе говоря, как на области определения X функции f, так и на области ее значений, имеют место ситуации неопределенности. В этом случае лучшее, что мы можем сделать, сводится к тому, чтобы перейти от первой ситуации неопределенности ко второй, т.е. от исходной ситуации на множестве X к ситуации на множестве *значений* функции f. Такой переход и был назван ***концентрацией информации***.

Покажем, что образование проекций двумерных ситуаций — это частный случай концентрации информации. Пусть исходная ситуация неопределенности двумерна и описывается пространством вероятностей $(X \times Y, \pi)$. На множестве $X \times Y$ можно определить две стандартные функции, отображающие произведение $X \times Y$ на каждый из сомножителей. Эти функции обычно обозначаются pr_1 и pr_2 и называются **функциями проектирования** произведения $X \times Y$ на множества соответственно X и Y. Они имеют вид:

$$\mathrm{pr}_1 = (x, y) \mapsto x \Diamond X \times Y, \tag{10}$$

$$\mathrm{pr}_2 = (x, y) \mapsto y \Diamond X \times Y. \tag{11}$$

Эти выражения означают, что каждому элементу (x, y) произведения $X \times Y$ функция pr_1 ставит в соответствие элемент x, а функция pr_2 — элемент y. Иначе говоря, для каждой пары $(x, y) \in X \times Y$ выполняются соотношения

$$\mathrm{pr}_1(x, y) = x, \tag{12}$$

$$\mathrm{pr}_2(x, y) = y. \tag{13}$$

Заметим, что необходимо отличать функции проектирования $\mathbf{pr_1}$ и $\mathbf{pr_2}$, заданные с помощью выражений (8) и (9), от функций проектирования pr_1 и pr_2 (выражения (10) и (11)), поскольку это *разные* функции. Действительно, функции pr_1 и pr_2 определены на множестве $X \times Y$, а функции $\mathbf{pr_1}$ и $\mathbf{pr_2}$ — на множестве всех *распределений вероятностей* на множестве $X \times Y$.

Как отмечалось в разделе 7 (п. 1), преобразование *концентрация информации* (относительно заданной функции) состоит из двух шагов: 1) *квантования* исходной ситуации (описание этого шага состояло в построении факторпространства исходного пространства относительно канонического разбиения области определения заданной функции); и 2) перехода к ситуации неопределенности на области значений заданной функции.

Покажем, что две определенные выше функции проектирования pr_1 и pr_2 как раз и являются теми функциями, относительно которых концентрация информации обеспечивается преобразованиями проектирования. Действительно, по образцу выражения (1) можно построить каноническое разбиение $\mathcal{X}_{\mathrm{pr}_1}$ множества $X \times Y$ (на классы неразличимости), порождаемое функцией pr_1. Поскольку функция pr_1 отображает множество $X \times Y$ *на* множество X, можно написать

$$\mathcal{X}_{\mathrm{pr}_1} = \{\mathrm{pr}_1^{-1}(x) : x \in X\}. \tag{14}$$

Легко понять, что для каждого $x \in X$ класс неразличимости $\mathrm{pr}_1^{-1}(x)$ представляет собой множество всех пар вида (x, y), где $y \in Y$. Иначе говоря, множество $\mathrm{pr}_1^{-1}(x)$ имеет вид

$$\mathrm{pr}_1^{-1}(x) = \{(x, y) : y \in Y\} = \{x\} \times Y. \tag{15}$$

Поэтому вместо (14) можем написать

$$\mathcal{X}_{\mathrm{pr}_1} = \{\{x\} \times Y : x \in X\}. \tag{16}$$

Имея разбиение $\mathcal{X}_{\mathrm{pr}_1}$ множества $X \times Y$, нетрудно построить факторпространство $(\mathcal{X}_{\mathrm{pr}_1}, \boldsymbol{p}_{\mathrm{pr}_1})$ по образцу факторпространства $(\mathcal{X}_f, \boldsymbol{p}_f)$ (разд. 7). Покажем теперь, что полученное таким образом факторпространство $(\mathcal{X}_{\mathrm{pr}_1}, \boldsymbol{p}_{\mathrm{pr}_1})$ *изоморфно* пространству вероятностей $(X, \mathbf{pr_1}\pi)$. Это значит, что множества X и $\mathcal{X}_{\mathrm{pr}_1}$ равномощны и существует такая *биекция* (взаимно однозначное отображение) f_1 множества X на $\mathcal{X}_{\mathrm{pr}_1}$, что для каждого $x \in X$ имеет место равенство

$$\mathbf{pr_1}\,\pi(x) = \boldsymbol{p}_{\mathrm{pr}_1}(f_1(x)). \tag{17}$$

Легко сообразить, что упомянутая биекция f_1 должна иметь вид

$$f_1 = x \mapsto \{x\} \times Y \lozenge X. \tag{18}$$

А это значит, что для каждого $x \in X$ выполняется

$$f_1(x) = \mathrm{pr}_1^{-1}(x) = \{x\} \times Y. \tag{19}$$

Теперь, имея образец построения факторпространства $(\mathcal{X}_{\mathrm{pr}_1}, \boldsymbol{p}_{\mathrm{pr}_1})$, изоморфного первой проекции $(X, \mathbf{pr_1}\pi)$, нетрудно построить и факторпространство $(\mathcal{Y}_{\mathrm{pr}_2}, \boldsymbol{p}_{\mathrm{pr}_2})$, изоморфное второй проекции $(Y, \mathbf{pr_2}\pi)$ (мы этого делать здесь не будем). А изоморфия пространств $(X, \mathbf{pr_1}\pi)$ и $(Y, \mathbf{pr_2}\pi)$ соответственно пространствам $(\mathcal{X}_{\mathrm{pr}_1}, \boldsymbol{p}_{\mathrm{pr}_1})$ и $(\mathcal{Y}_{\mathrm{pr}_2}, \boldsymbol{p}_{\mathrm{pr}_2})$ означает, что образование первой и второй проекций двумерной ситуации неопределенности является примерами концентрации информации.

3. Информационный аспект образования проекций. Теперь наша задача состоит в сравнении исходного двумерного пространства вероятностей $(X \times Y, \pi)$ с его проекциями $(X, \mathbf{pr_1}\pi)$ и $(Y, \mathbf{pr_2}\pi)$. Количество собственной информации произвольного элемента $(x, y) \in X \times Y$ пространства $(X \times Y, \pi)$ характеризуется выражением

$$I_\pi(x, y) = \log \frac{1}{\pi(x, y)}. \tag{20}$$

А выражения

$$I_{\mathbf{pr_1}\pi}(x) = \log \frac{1}{\mathbf{pr_1}\pi(x)} \quad \text{и} \quad I_{\mathbf{pr_2}\pi}(y) = \log \frac{1}{\mathbf{pr_2}\pi(y)} \tag{21}$$

характеризуют количества собственной информации соответственно элемента $x \in X$ пространства $(X, \mathbf{pr_1}\pi)$ и элемента $y \in Y$ пространства $(Y, \mathbf{pr_2}\pi)$.

Нетрудно показать, что для любых $x \in X$ и $y \in Y$ меры собственной информации удовлетворяют неравенствам

$$I_{\mathbf{pr_1}\pi}(x) \le I_\pi(x, y) \ge I_{\mathbf{pr_2}\pi}(y), \tag{22}$$

$$I_\pi(x, y) \le I_{\mathbf{pr_1}\pi}(x) + I_{\mathbf{pr_2}\pi}(y). \tag{23}$$

Аналогичным неравенствам удовлетворяют и меры неопределенности пространств вероятностей $(X \times Y, \pi)$, $(X, \mathbf{pr_1}\pi)$ и $(Y, \mathbf{pr_2}\pi)$:

$$G(X, \mathbf{pr_1}\pi) \le G(X \times Y, \pi) \ge G(Y, \mathbf{pr_2}\pi), \tag{24}$$

$$G(X \times Y, \pi) \le G(X, \mathbf{pr_1}\pi) + G(Y, \mathbf{pr_2}\pi). \tag{25}$$

4. Две меры концентрации информации. По аналогии с мерой концентрации информации $E(X, p\,|\,f)$, введенной в разделе 7 (п. 2) можно ввести две внешние **меры концентрации информации** при преобразовании двумерной ситуации неопределенности в две проекции этой ситуации:

$$E(X \times Y,\ \pi \mid \mathrm{pr}_1) = G(X \times Y, \pi) - G(X,\ \mathbf{pr_1}\pi), \tag{26}$$

$$E(X \times Y,\ \pi \mid \mathrm{pr}_2) = G(X \times Y, \pi) - G(Y,\ \mathbf{pr_2}\pi), \tag{27}$$

а из неравенств (24) следуют неравенства

$$E(X \times Y,\ \pi \mid \mathrm{pr}_1) \geq 0 \quad \text{и} \quad E(X \times Y,\ \pi \mid \mathrm{pr}_2) \geq 0, \tag{28}$$

аналогичные неравенству (7).

9. ПРЕОБРАЗОВАНИЕ «ОГРАНИЧЕНИЕ РАЗНООБРАЗИЯ»

Теперь мы переходим к рассмотрению примеров другого рода преобразований ситуаций неопределенности — преобразований, при которых информация не теряется, а, наоборот, *приобретается* (а это возможно только в том случае, если информация откуда-то *поступает*). Первым примером такого рода преобразования является *ограничение разнообразия*. Как известно, это понятие ввел У. Росс Эшби.

1. Достаточное условие наличия информационной связи. Согласно У. Эшби понятие *информация* связано не только с понятием *разнообразие*, но еще более тесно — с понятием *ограничение разнообразия*, которое Эшби рассматривает во второй части книги «Введение в кибернетику». Он написал:

«Тот факт, что нечто “предсказуемо”, подразумевает наличие ограничения разнообразия» [2, с. 187].

Иначе говоря, наличие ограничения разнообразия в системе означает ее “предсказуемость”. Но чт***о*** здесь следует понимать под “предсказуемостью”? Сейчас мы покажем, что под “предсказуемостью” системы можно понимать наличие полученной откуда-то *информации* об этой системе. Связав все это вместе, получим следующий вывод.

Если в системе наблюдается ограничение разнообразия (т.е. сокращение множества) возможных ее состояний, то это является достаточным свидетельством того, что о системе получена информация от внешнего источника.

Это последнее утверждение можно следующим образом уточнить.

Достаточное условие. Рассмотрим систему, для которой определено понятие *состояния* и известно множество возможных состояний. И предположим, что где-то произошло событие, или наблюдается явление, которое *несовместимо* (согласно существующим представлениям) с некоторыми из возможных для данной системы состояний. Иначе говоря, *логическим следствием* утверждения об этом событии или явлении (в сочетании с существующей системой знаний) является ***ограничение разнообразия*** возможных состояний системы. Тогда есть основание утверждать, что данное событие или явление содержит (внешнюю) ***информацию*** о состояниях системы, или что в направлении от данного события или явления к состояниям системы существует ***информационная связь***. ■

Это *достаточное условие* существования внешней информации не является также и необходимым условием, поскольку внешняя информация не обязательно ведет к уменьшению разнообразия в данной ситуации неопределенности, а может вызывать другие преобразования этой ситуации. Но зато наше *достаточное условие* может выступать в совершенно необычной роли — в роли своеобразного доказательства того, что возникновение информационных связей в Природе представляет собой массовое явление, а не что-то уникальное, поскольку случаи, когда имеют место ограничения разнообразия, наблюдаются буквально на каждом шагу (множество примеров рассмотрел Эшби, и здесь нет нужды все это повторять).

Крайне важно понимать, что понятие *логического следствия* (которое фигурирует в достаточном условии) не совпадает с понятием *причинно-следственной связи*. Логическое следствие — чисто лингвистическое понятие, которое нельзя применять непосредственно к реальным событиям, а можно применять только к *утверждениям* об этих событиях. С другой стороны, причинно-следственные связи могут существовать только между реальными событиями или явлениями.

Это различие, которое может показаться вполне безобидным, ведет к следующим двум ***принципиальным*** отличиям информационных связей от более привычных причинно-следственных связей. Причинно-следственные связи *материальны* и всегда действуют только *в одну сторону* — в направлении от причины к следствию. Информационные же связи ***нематериальны*** и всегда действуют ***в обе стороны***. Непонимание этих отличий ведет к огромному количеству недоразумений и разочарований, когда информационные явления пытаются анализировать — *безуспешно* — путем поиска причин и следствий. Такой подход является в корне ошибочным, так как… (сейчас будет раскрыта страшная тайна, которую скрывают от трудящихся) ***информационные явления не подчиняются принципу причинности***.

Наиболее остро эта проблема стоит в биологии. В качестве примера приведем короткую цитату из статьи «Причина и следствие в биологии» профессора Эрнста Майра:

«…мне хотелось бы сосредоточить внимание на особых трудностях, связанных с классической концепцией причинности в биологии. Эти трудности начинают терзать исследователей при первых же попытках создать единое понятие причины» [3, с. 47].

Объясняются подобные трудности тем, что ввиду двустороннего характера информационных связей, их направление часто оказывается *противоположным* направлению *причина-следствие*. Например, если состояние системы A содержит информацию о состоянии системы B, то и обратно, состояние системы B содержит информацию о состоянии системы A. Причем, если речь идет о вероятностных ситуациях, то, как известно из теории информации, количество той и другой информации одинаково. При этом может случиться даже так, что, хотя между состояниями систем A и B имеются информационные связи, между ними нет (или даже вообще не может быть) причинно-следственных связей. Такие информационные связи будем называть ***косвенными***.

Образование косвенных информационных связей возможно тогда, например, когда речь идет о состояниях систем A и B *в один и тот же момент времени*, а сами эти системы находятся на некотором расстоянии друг от друга.

Как известно, в этом случае причинно-следственная связь между состояниями систем исключена. Существование же информационных связей (двусторонних), тем не менее, возможно при условии, что состояния обеих систем A и B являются следствием *общей причины*, находящейся в некоторой третьей системе.

Существует один важный частный случай (двусторонних) информационных связей, о котором следует упомянуть специально. Любой материальный объект во Вселенной содержит огромное количество разнообразной информации. В том числе он содержит информацию как о своем будущем, так и о своем прошлом. Ее можно использовать для получения вполне обоснованных выводов как в одну, так и в другую сторону. Однако современная наука, которая ***целиком опирается на принцип причинности***, признает законными только выводы о будущем (кроме тривиальных случаев при рассмотрении обратимых явлений). Причинный характер объяснений в настоящее время рассматривается как признак их научности. Вот что пишет по этому поводу С. Крымский в книге «Запити філософських смислів»:

«Важливим критерієм науковості є введення причинної матриці пояснення явищ… Отже, пошук причинного пояснення є атрибутивним для визначення наукового підходу» [4, с. 154, 155].

Таким образом, как это ни смешно, но современная наука не может обосновать какие бы то ни было выводы о прошлом (поскольку настоящее не может быть причиной прошлого).

Тем не менее существует огромная (и очень пестрая) армия ученых, занимающихся исключительно добыванием выводов о прошлом. Кроме *историков* в узком понимании этого слова, к этой армии принадлежат астрофизики; специалисты по космологии; геологи; археологи; палеонтологи; палеоботаники; исследователи, пытающиеся узнать, как возникла Солнечная система, Земля, другие планеты, их спутники, астероиды и кометы, или выяснить, как возникла жизнь на Земле. И все эти ученые вынуждены работать в научном вакууме, поскольку никакие "научные" методы получения выводов о прошлом ***неизвестны***.

Как до этого могло дойти? Ведь понятие информации вошло в науку уже более чем полвека назад. Предполагается также, что информация все это время интенсивно изучалась. Однако в разделе 1 (первая часть статьи) мы дали краткую справку о том, ***как она изучалась***. В кибернетике была также провозглашена программа изучения *законов преобразования информации*. Но предложенный *способ* осуществления этой программы был совершенно негодным — он сводился к изучению преобразований *носителей информации* — текстов! Но изучая преобразования текстов, ничего нельзя узнать о преобразованиях информации. Ясно также, что при упомянутых выше "способах изучения" информации и ее преобразований невозможно прийти к выводу о существовании информационных связей в Природе и невозможно увидеть те перспективы кибернетики (и вообще естествознания), которые открываются вследствие существования таких связей.

Повышению общей информационной культуры не могли способствовать и широко распространенные представления о том, что информация якобы аналогична *физической энтропии*. Ложная аналогия — не помощник, а серьезная помеха. В самом деле, опираясь на эту "аналогию", тоже невозможно прийти к выводу о существовании информационных связей в Природе, поскольку физическая энтропия никакие подобные связи образовывать не может.

2. Описание преобразования «ограничение разнообразия». Рассмотрим снова вероятностную ситуацию неопределенности, описываемую пространством вероятностей (X, p). Предположим, что где-то произошло событие, которое *несовместимо* (согласно существующей системе знаний) с некоторыми состояниями природы из множества X. Иначе говоря, в результате полученной внешней информации выяснилось, что некоторые элементы множества X осуществиться не могут. Так что теперь вместо множества X мы должны рассматривать его *собственное* подмножество, которое обозначим A. Согласно предложенному выше достаточному условию наличия внешней информации сказанное можно рассматривать как пример *действия информационной связи* в направлении от данного события к рассматриваемой ситуации неопределенности.

Попытаемся понять, чт***о*** в этом случае должно произойти с исходной ситуацией неопределенности на множестве X? Как только определилось подмножество A, до которого оказалось ограниченным исходное множество возможностей X, мы *автоматически* оказались в новой ситуации неопределенности, независимо от того, понимаем мы это или нет. Единственное, чт***о*** нам теперь осталось сделать, — это правильно *описать* эту новую ситуацию.

Приходится снова и снова повторять одно и то же — что преобразования ситуаций неопределенности (а следовательно — и преобразования информации) происходят не на бумаге и не в каких-то специальных *“преобразователях”*, а *реально* — в той *действительности*, в которой мы живем, принимаем решения и действуем.

Итак, смысл рассматриваемого преобразования сводится к тому, что теперь вместо множества X мы должны рассматривать его *собственное* (т.е. не совпадающее с самим множеством X) подмножество A. Ясно, что это подмножество A не может оказаться *пустым*, так как в этом случае мы бы пришли к *противоречию*, состоящему в том, что теперь *не осталось ни одного состояния природы, которое могло бы осуществиться*.

Но это еще не все — аналогичное противоречие получилось бы и в том случае, если бы *суммарная вероятность*

$$\overline{p}(A) = \sum_{x \in A} p(x) \tag{29}$$

всех элементов множества A равнялась нулю. Действительно, это означало бы, что все оставшиеся якобы возможные состояния природы в действительности имеют нулевые вероятности и, следовательно, фактически тоже осуществиться не могут.

Есть все основания предполагать, что никакие события во Вселенной не могут приводить к подобным противоречиям. Поэтому, если противоречие все же возникает (т.е. оказывается, что $\overline{p}(A) = 0$), остается возможным только одно объяснение: противоречивой является та *система знаний*, на основании которой был получен вывод о составе множества A.

Итак, мы должны предположить, что $\overline{p}(A) > 0$. Преобразование *ограничение разнообразия* сводится к тому, что при упомянутом условии $\overline{p}(A) > 0$ мы *автоматически* оказываемся в новой ситуации неопределенности (с множеством возможностей A). Подчеркнем специально, что сам факт осуществления этого преобразования совершенно не зависит от того, известна ли уже теория вероятностей. Действительно, от знания теории вероятностей зависит только то, сумеем ли мы правильно *описать* это преобразование.

Итак, если исходная ситуация неопределенности описывалась пространством вероятностей (X, p), то, как следует из теории вероятностей, ситуация, образовавшаяся в результате преобразования *ограничение разнообразия*, тоже должна описываться некоторым пространством вероятностей, которое мы обозначим (A, p^A) (можно показать, что (A, p^A) есть *подпространство* пространства (X, p)). А для того чтобы найти описание (A, p^A), достаточно построить распределение вероятностей p^A (как функцию, определенную на множестве A). Согласно известным соотношениям теории вероятностей эта функция должна иметь вид

$$p^A = x \mapsto \frac{p(x)}{\overline{p}(A)} \lozenge A . \tag{30}$$

Это значит, что для всякого элемента $a \in A$ должно выполняться равенство

$$p^A(a) = \frac{p(a)}{\overline{p}(A)} . \tag{31}$$

(заметим, что в выражениях (30) и (31) не возникает деления на ноль, поскольку, как мы показали выше, имеет место $\overline{p}(A) > 0$).

3. Информационный аспект преобразования. Теперь займемся вопросом о преобразовании, которому подверглась мера собственной информации и мера неопределенности всей ситуации. Найдем выражение для *количества собственной информации* $I_{p^A}(a)$ элемента $a \in A$ новой ситуации неопределенности, описываемой подпространством (A, p^A). Опираясь на выражения [1, (2)] и (31), получим

$$I_{p^A}(a) = \log \frac{1}{p^A(a)} = \log \frac{\overline{p}(A)}{p(a)} . \tag{32}$$

Как видно из выражения (32), для каждого $a \in A$ должно выполняться неравенство $I_p(a) \ge I_{p^A}(a)$ (равенство будет иметь место только в случае $\overline{p}(A) = 1$). При этом разность $I_p(a) - I_{p^A}(a)$ для всех $a \in A$ одна и та же:

$$I_p(a) - I_{p^A}(a) = \log \frac{1}{\overline{p}(A)} \ge 0 . \tag{33}$$

Таким образом, в результате получения информации от постороннего события количество собственной информации для каждого элемента множества A *уменьшилось*. Теперь запишем выражение для меры неопределенности новой ситуации:

$$G(A, p^A) = \mathbf{E} I_{p^A} = \sum_{a \in A} p^A(a) \cdot I_{p^A}(a) = \sum_{a \in A} p^A(a) \cdot \log \frac{1}{p^A(a)} . \tag{34}$$

4. Опровержение давно укоренившихся представлений. Теперь мы попытаемся найти способ измерения количества внешней информации, необходимой для преобразования ситуации неопределенности, описываемой пространством вероятностей (X, p), в ситуацию, получающуюся в результате *ограничения* множества X до его собственного подмножества A. Обозначим это количество $E(X, p \mid A)$. Может показаться, что по аналогии с преобразованием *квантование* число $E(X, p \mid A)$ должно равняться разности

$$\delta G = G(X, p) - G(A, p^A) . \tag{35}$$

В самом деле, кажется почти очевидным, что степень неопределенности $G(A, p^A)$ новой ситуации здесь также не должна превышать неопределенность $G(X, p)$ исходной ситуации, т.е. должно выполняться неравенство $G(X, p) \geq G(A, p^A)$. Для проверки этого предположения можно рассмотреть два частных случая: 1) когда суммарная вероятность $\overline{p}(A)$ множества A равна единице; и 2) когда мощность $|A|$ множества A равна единице (т.е. множество A сводится к одному элементу).

Легко показать, что в первом случае (когда $\overline{p}(A) = 1$) мы получим равенство $G(X, p) = G(A, p^A)$ (из равенства $\overline{p}(A) = 1$ следует, что вероятность $p(x)$ каждой точки $x \in X \setminus A$ пространства вероятностей (X, p), не принадлежащей множеству A, равна нулю). Во втором же случае (когда множество A сводится к одному элементу) получим равенство $G(A, p^A) = 0$. После такого ограничения разнообразия ситуация неопределенности фактически прекращает свое существование, т.е. возникает так называемы *коллапс ситуации неопределенности*. Этому преобразованию посвящен раздел 12 (третья часть статьи).

Оба эти частные случаи не противоречат предположению $G(X, p) \geq$ $\geq G(A, p^A)$. Однако сейчас мы покажем, что в общем случае это предположение является ***ошибочным***. Рассмотрим такой пример. Пусть множество X состоит из трех элементов: $X = \{x_1, x_2, x_3\}$, а множество A — из двух: $A = \{x_2, x_3\}$. Иначе говоря, преобразование *ограничение разнообразия* в данном случае свелось к исключению из рассмотрения элемента x_1.

На рис. 1 показано, какой в этом случае будет зависимость величин $G(X, p)$ и $G(A, p^A)$ от числа $p(x_1)$ при условии, что между вероятностями $p(x_1)$, $p(x_2)$ и $p(x_3)$ имеется следующая связь: $p(x_2) = p(x_3)$ (и, конечно, $p(x_2) + p(x_3) = 1 - p(x_1)$). На графике хорошо видно, что неравенство $G(X, p) \geq G(A, p^A)$ справедливо не для всех значений $p(x_1)$ (при значениях $p(x_1) >\approx 0.8$ оно не выполняется).

Рассмотренный пример позволяет сделать следующие выводы.

1. Для исключения из рассмотрения возможного состояния природы x_1 требуется поступление информации извне (достаточное условие наличия внешней информации; п. 1).

2. При вполне определенных условиях (когда $p(x_1) >\approx 0.8$) эта полученная информация ведет не к уменьшению, а к увеличению степени неопределенности рассматриваемой ситуации.

Таким образом, этот пример фактически опровергает целую систему представлений, возникших после появления теории информации — представлений, согласно которым *информация — это то, что устраняет* (*или хотя бы уменьшает*) *неопределенность*. В чем же причина того, что упомянутые представления считались истинными более полувека? Причина — в двух специфических особенностях классической теории информации.

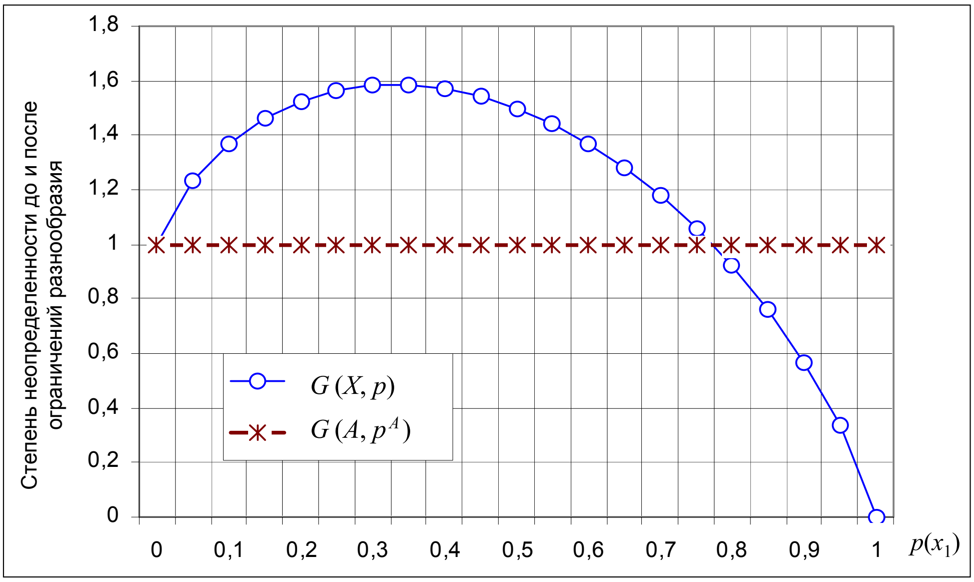

Рис. 1. Зависимость величин $G(X, p)$ и $G(A, p^A)$ от числа $p(x_1)$ при условиях: 1) $X = (x_1, x_2, x_3)$; 2) $A = (x_2, x_3)$; 3) $p(x_2) = p(x_3)$; 4) $p(x_2) + p(x_3) = 1 - p(x_1)$

Во-первых, эта теория не содержит признаков того, что было замечено отличие между двумя *ролями* информации, о которых здесь мы говорим как о *внутренней* и *внешней* информации. Но без учета этого отличия вообще невозможно выработать непротиворечивые представления об информации, поскольку информация в процессе исполнения этих двух ролей демонстрирует настолько непохожие свойства, что вообще отсутствует возможность их рассмотрения с единой точки зрения.

Во-вторых, главные результаты теории информации были *асимптотическими* и относились к случаю *безостановочной передачи сообщений* от источника к приемнику. В таких условиях наиболее важную роль играют не одиночные события, а их *массовый эффект*. А массовый эффект, связанный с передачей информации от источника к приемнику, состоит в том, что *в среднем* переданная информация действительно уменьшает степень неопределенности состояния приемника. Таким образом, ошибочные представления держались за счет того, что под результатом действия информации на ситуацию неопределенности всегда понимали только *средний результат*.

10. СИТУАЦИИ ЗАБЛУЖДЕНИЯ

Итак, разность δG (35), по-видимому, не может рассматриваться как количество $E(X, p \mid A)$ внешней информации, необходимой для *ограничения разнообразия*, т.е. той информации, которая требуется для преобразования ситуации, описываемой пространством (X,p), в ситуацию, описываемую пространством (A,p^A). И дело здесь не только в том, что эта разность в некоторых случаях становится отрицательной. Есть еще два важных свойства, которыми должна обладать мера $E(X,p\,|\,A)$, но которыми *не обладает* разность δG.

Одно из этих обязательных свойств является совершенно очевидным. Мы знаем, что преобразование *ограничение разнообразия* для ситуации неопределенности, описываемой пространством вероятностей (X,p), представляет собой исключение из рассмотрения некоторых элементов множества X. Представляется естественным, что количество информации, необходимое для исключения этих элементов, должно зависеть от их вероятностей (относительно исходного распределения p). Причем, чем б***о***льшими вероятностями обладают эти элементы, тем больше необходимо информации для того, чтобы их вообще исключить из рассмотрения.

Второе из упомянутых обязательных свойств является асимптотическим. Если суммарная вероятность исключаемых элементов стремится к единице, то количество информации, требующейся для исключения их из рассмотрения, должно, по-видимому, стремиться к бесконечности.

1. Ситуации заблуждения и способ их описания. Для построения меры, обладающей перечисленными свойствами, мы рассмотрим другой подход к информационному анализу преобразования *ограничение разнообразия*. С этой целью предположим, что смысл преобразования *ограничение разнообразия*, т.е. перехода от ситуации неопределенности, описываемой пространством вероятностей (X,p), к ситуации, описываемой пространством (A,p^A), может быть выражен с помощью следующего утверждения:

Первоначальное предположение о том, что пространство (X,p) представляет собой верное описание имеющей место ситуации неопределенности, не подтвердилось. Оказалось, что эту ситуацию нужно описывать пространством (A,p^A).

Для того чтобы лучше понять, как это утверждение может повлиять на решение нашей задачи, полезно сначала рассмотреть более общую задачу. Предположим, что мы столкнулись с ситуацией, которая имеет так называемую “вероятностную природу”, но, тем не менее, не относится к тем ситуациям, к которым применима теория вероятностей. На (дискретном) множестве X действует РВ q, но тот, кто хочет принять некое решение, не знает верного описания этого распределения. Он думает, что действующим на X распределением является РВ p (которое может как совпадать с распределением q, так и отличаться от него). Такого рода ситуации демонстрируют особую, специфическую разновидность неопределенности, которую можно квалифицировать как ***заблуждение***. Поэтому подобные ситуации мы и будем называть **ситуациями заблуждения**.

Ясно, что ситуация заблуждения не может быть формально описана пространством вероятностей (X,q). Но она, конечно, не может быть описана и пространством вероятностей (X,p). Для ее описания необходимо каким-то образом учесть оба распределения вероятностей: p и q. Мы условимся описывать эту ситуацию символом $(X, q \Diamond p)$, *формальный* смысл которого сейчас поясним.

Оба распределения p и q, входящие в конструкцию $q \Diamond p$, являются математическими структурами, заданными на множестве X. А Н. Бурбаки показал, что произвольное *конечное* число математических структур, (заданных на одном и том же множестве) может быть представлено в виде *одной* структуры. Вот эту результирующую структуру, порожденную структурами p и q, мы и обозначили символом $q \Diamond p$. Так что формально конструкция $(X, q \Diamond p)$ представляет собой еще один пример математического пространства.

Определение 2. Математические пространства вида $(X, q \Diamond p)$ будем называть **пространствами двойных вероятностей**. ■

2. Количество внешней информации в пользу одной гипотезы против другой. Теперь возникает такой вопрос: можно ли для ситуации заблуждения, описываемой пространством двойных вероятностей $(X, q \Diamond p)$, предложить некую *меру неопределенности*? Для решения этого вопроса заметим, что с точки зрения принимающего решения количество собственной информации, содержащейся в некоторой точке $x \in X$ этого пространства, должно измеряться выражением [1, (2)]. Иначе говоря, информационная функция принимающего решения должна иметь тот же вид

$$I_p = x \mapsto I_p(x) \Diamond X, \tag{36}$$

что и в выражении [1, (3)].

Поэтому, если под мерой неопределенности ситуации по-прежнему понимать математическое ожидание (относительно действующего распределения) информационной функции (как в [1, (5)]), то ***мерой неопределенности ситуации заблуждения***, описываемой пространством двойных вероятностей $(X, q \Diamond p)$, следует считать число

$$G(X, q \Diamond p) = \mathbf{E}\, I_p = \sum_{x \in X} q(x) \cdot I_p(x) = \sum_{x \in X} q(x) \cdot \log \frac{1}{p(x)}, \tag{37}$$

которое не совпадает с математическим ожиданием $\mathbf{E}\, I_p$ в [1, (5)] ввиду того, что здесь у нас действующим распределением является не p, а q.

Свойства правой части выражения (37) хорошо изучены в теории информации. Известно, что какими бы ни были распределения p и q, выполняется неравенство

$$G(X, q \Diamond p) \geq G(X, q), \tag{38}$$

причем это неравенство превращается в равенство тогда и только тогда, когда $p = q$. Более того, как видно из выражения (37), если для некоторого $x \in X$ число $p(x)$ устремить к нулю (при условии $q(x) \neq 0$), то степень неопределенности $G(X, q \Diamond p)$ устремится к бесконечности.

На основании неравенства (38) можно ввести следующую ***меру внешней информации***:

$$E(X, p \| q) = G(X, q \lozenge p) - G(X, q) . \tag{39}$$

Раскрыв разность в правой части (39), получим

$$E(X, p \| q) = \sum_{x \in X} q(x) \cdot \log \frac{q(x)}{p(x)} . \tag{40}$$

А глядя на выражение (40), приходим к выводу, что внешняя мера $E(X, p \| q)$ представляет собой не что иное, как дискретный вариант знаменитой меры С. Кульбака, которую он положил в основу статистических решений в книге «Теория информации и статистика» [5, с. 13–16].

Однако здесь мы получили эту меру естественным путем, исходя из других соображений (как ее обозначение и наименование у С. Кульбака, так и способ ее обоснования и формального описания совершенно не соответствуют смыслу решаемой здесь задачи). И здесь эту меру мы обозначили иначе. Мы также дадим ей другое имя.

Определение 3. Меру $E(X, p \| q)$ назовем **количеством внешней информации** в пользу гипотезы q против гипотезы p . ■

На нескольких примерах, рассматриваемых в этой статье дальше, мы покажем, что число $E(X, p \| q)$ имеет следующий смысл.

Предположение 2. Пусть *в результате полученной внешней информации* некоторая вероятностная ситуация неопределенности, описываемая пространством вероятностей (X, p), преобразовалась в вероятностную ситуацию (с тем же множеством возможностей), описываемую пространством вероятностей (X, q). Тогда величина $E(X, p \| q)$ (40) измеряет количество внешней информации, которое необходимо для этого преобразования. ■

*****

В следующей части статьи будет, в частности, показано, как понятие ситуации заблуждения позволяет найти способ измерения интенсивности преобразования *ограничение разнообразия*.

ЛИТЕРАТУРА

1. *Дидук Н.Н.* Меры внутренней и внешней информации (на примере вероятностных ситуаций неопределенности). Часть I // Системні дослідження та інформаційні технології. — 2012. — № 3 — С. 107–124.
2. *Эшби У.Р.* Введение в кибернетику. — М.: Изд. ин. лит., 1959. — 432 с.
3. *На пути* к теоретической биологии. I. Пролегомены. — М.: Мир, 1970. — 182 с.
4. *Кримський С.* Запити філософських смислів. — Київ: Парапан, 2003. — 240 с.
5. *Кульбак С.* Теория информации и статистика. — М.: Наука, 1967. — 408 с.

Поступила 01.06.2009

Статтю надруковано в редакції автора