УДК 519.7

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

Н.Н. ДИДУК

Предложен способ измерения интенсивности преобразования *ограничение разнообразия*, основанный на новом понятии *ситуации заблуждения*. Рассмотрено преобразование *коллапс ситуации неопределенности*. Показано, что это преобразование может осуществляться двумя способами: *внутренним* и *внешним*. Внутренний способ коллапса ситуации связан с наступлением (в данной ситуации) случайного события или с сознательным выбором. Внешний же способ представляет собой крайний частный случай преобразования *ограничение разнообразия*. На примере коллапса ситуации показана возможность превращения внешней информации во внутреннюю. Рассмотрены меры внутренней и внешней *взаимной* информации и решен вопрос о том, с каким преобразованием ситуаций неопределенности связана последняя (внешняя) мера. Введено новое понятие *информационного канала* и показано, что информационные каналы являются неотъемлемыми участниками всех процессов, происходящих в Природе.

В первых двух частях статьи [1, 2] рассмотрен ряд преобразований информации и построены меры интенсивности некоторых из них. Но для самого важного из преобразований — *ограничения разнообразия* — меру интенсивности построить не удалось. Однако в последнем (десятом) разделе второй части введено новое понятие *ситуации заблуждения*, которое позволило сформулировать *общий подход* к построению мер интенсивности преобразований, вызванных полученной внешней информацией (предположение 2). Этот общий подход здесь (в третьей части) использован для построения меры интенсивности преобразования *ограничение разнообразия*.

Далее здесь рассмотрено преобразование *коллапс ситуации неопределенности* и показано, что его можно рассматривать либо как самостоятельное преобразование, либо как частный случай ограничения разнообразия. Эта двойственность позволила получить первое подтверждение согласованности мер внутренней и внешней информации. Две последние темы настоящей части статьи — меры внутренней и внешней *взаимной информации* и новое понятие *информационного канала*.

11. СПОСОБ ИЗМЕРЕНИЯ ИНТЕНСИВНОСТИ ПРЕОБРАЗОВАНИЯ «ОГРАНИЧЕНИЕ РАЗНООБРАЗИЯ»

Предлагаемый здесь новый подход к измерению интенсивности преобразования *ограничение разнообразия* опирается на введенное во второй части понятие *ситуации заблуждения*. Здесь мы имеем первый пример, в котором сформулированное в [2, разд. 10] предположение 2 подвергается проверке. Но рассмотрение этого примера наталкивается на следующую трудность. Предположение 2 содержит условие, что в процессе преобразования ситуации не изменяется ее *множество возможностей*. Однако преобразование *ограничение разнообразия* как раз этому условию не удовлетворяет. Действительно, его главная особенность состоит в уменьшении (урезании) множества возможностей: ситуация, описываемая пространством (X, p), преобразуется в ситуацию, описываемую пространством (A, p^A) $(A \subset X)$.

1. Количество внешней информации, необходимое для ограничения разнообразия. Покажем, что упомянутую трудность можно устранить. С этой целью воспользуемся так называемым *нулевым продолжением функций* — операцией, которая была описана в разделе 6 работы [3] (выражения (11) и (12)). В применении к ограничению разнообразия суть этой операции состоит в том, что РВ p^A (которое определено на множестве $A \subset X$) можно тривиальным образом продолжить на множество X.

Используя прием нулевого продолжения, мы можем формально построить функцию $p^A \uparrow X$, представляющую собой продолжение РВ p^A на множество X:

$$p^A \uparrow X = x \mapsto (p^A(x) : \ x \in A \,\big|\, 0) \lozenge X . \tag{1}$$

Это выражение означает, что в тех точках $x \in X$, которые принадлежат множеству A, новая функция $p^A \uparrow X$ принимает те же значения, что и функция p^A; во всех же остальных точках $x \in X$ она принимает значение 0. При этом, как нетрудно понять, новая функция $p^A \uparrow X$ тоже является распределением вероятностей, но уже на множестве X.

Итак, вместо того чтобы говорить, что после преобразования *ограничение разнообразия* заключительная ситуация описывается пространством вероятностей (A, p^A), можно сказать, что она описывается пространством вероятностей $(X,\ p^A \uparrow X)$. В самом деле, это новое пространство отличается от пространства (A, p^A) только некоторыми точками, имеющими нулевую вероятность. Меру внешней информации для преобразования *ограничение разнообразия* в [2, разд. 9, п. 4] мы обозначили $E(X, p \,|\, A)$. Следовательно, наше предположение 2 [2, разд. 10, п. 2] в применении к этому преобразованию выразится равенством

$$E(X, p \,|\, A) = E(X, p \,\|\, p^A \uparrow X) . \tag{2}$$

Теперь осталось найти выражение для величины $E(X, p \,\|\, p^A \uparrow X)$. Ввиду [2, (39)] имеем

$$E(X, p \,\|\, p^A \uparrow X) = G(X, (p^A \uparrow X) \lozenge p) - G(X, p^A \uparrow X) . \tag{3}$$

Далее, очевидно, что имеют место равенства

$$G(X, (p^A \uparrow X) \lozenge p) = \mathbf{E}\, I_p = \sum_{a \in A} p^A(a) \cdot \log \frac{1}{p(a)} , \tag{4}$$

$$G(X, p^A \uparrow X) = G(A, p^A) = \sum_{a \in A} p^A(a) \cdot I_{p^A}(a) . \tag{5}$$

Таким образом, из выражений (4) и (5) видно, что (после построения распределения вероятностей p^A) для вычисления значений $G(X,\ p^A{\uparrow}X)$ и $G(X,(p^A{\uparrow}X)\lozenge p)$ уже не играют никакой роли все точки пространства (X,p), не принадлежащие множеству A. Поэтому вместо громоздких обозначений $G(X,p^A{\uparrow}X)$, $G(X,(p^A{\uparrow}X)\lozenge p)$ и $E(X,p\|\,p^A{\uparrow}X)$ можно спокойно пользоваться обозначениями $G(A,p^A)$, $G(X,\ p^A\lozenge p)$ и $E(X,p\|\,p^A)$. Так что наше предположение 2 в применении к преобразованию *ограничение разнообразия* можно вместо (2) выразить равенством

$$E(X,p\,|\,A)=E(X,p\,\|\,p^A)=G(X,\ p^A\ \lozenge\ p)-G(A,p^A)=$$

$$=\sum_{a\in A}p^A(a)\cdot\log\frac{p^A(a)}{p(a)}. \tag{6}$$

2. Иллюстрация. С помощью графической иллюстрации покажем, что мера $E(X,p\|\,p^A)$ обладает всеми свойствами, которыми должно обладать количество внешней информации $E(X,p\,|\,A)$, вызывающей ограничение разнообразия. Такая иллюстрация представлена на рис. 2. Здесь приняты те же допущения, что и для предыдущей иллюстрации [2, разд. 9, п. 4, рис. 1]: 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* свелось к исключению из рассмотрения элемента x_1.

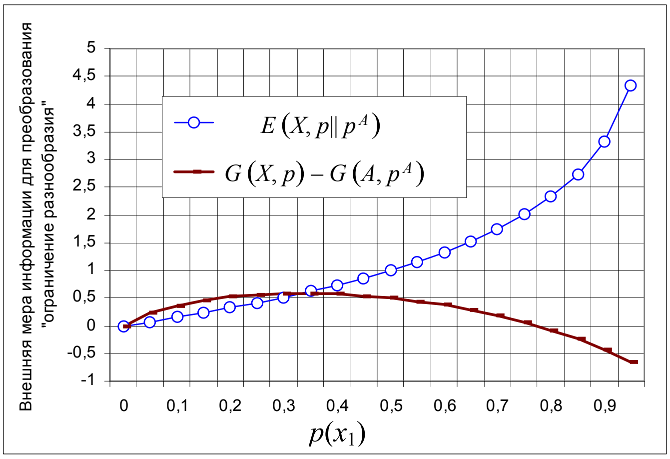

Рис. 2. Мера внешней информации $E(X,p\|\,p^A)$ в пользу гипотезы p^A против гипотезы p (информации преобразования *ограничение разнообразия*) как функция числа $p(x_1)$. Для сравнения приведена зависимость разности $G(X,p)-G(A,p^A)$ от того же числа $p(x_1)$

На рис. 2 показано, какой при этих условиях будет зависимость величины $E(X,p\|\,p^A)$ от числа $p(x_1)$ (между вероятностями $p(x_1)$, $p(x_2)$ и $p(x_3)$ имеется уже известная нам связь: $p(x_2)=p(x_3)$ и $p(x_2)+p(x_3)=1-p(x_1)$). Ради сравнения здесь же показана зависимость от числа $p(x_1)$ разности $G(X,p)-G(A,p^A)$. На диаграмме видно, что поведение величины $E(X,p\|\,p^A)$ характерно следующими особенностями:

- При условии $p(x_1)=0$ имеет место $E(X,p\|\,p^A)=0$. Это хорошо согласуется с интуитивными ожиданиями, поскольку представляется понятным, почему для исключения из рассмотрения элемента, имеющего нулевую вероятность, требуется нулевое количество информации.

- Величина $E(X, p \| p^A)$ всюду неотрицательна и возрастает с увеличением числа $p(x_1)$. Это, по-видимому, тоже хорошо согласуется с ожиданиями: чем более вероятным является исключаемый из рассмотрения элемент, тем больше для этого нужно информации.
- Наконец, имеется асимптотическое свойство, о котором мы уже говорили выше: с приближением вероятности $p(x_1)$ исключаемого из рассмотрения элемента x_1 к единице величина $E(X, p \| p^A)$ стремится к бесконечности (это можно доказать).

Таким образом, мы получили серьезное свидетельству в пользу того, что число $E(X, p \| p^A)$ действительно измеряет количество внешней информации $E(X, p \mid A)$, необходимой для преобразования ситуации неопределенности, описываемой пространством вероятностей (X, p), в ситуацию, описываемую пространством (A, p^A). Но это предположение все еще остается не доказанным. И способ, позволяющий получить его *формальное* доказательство, пока не найден. Но в следующем разделе и в четвертой части статьи будут предложены, хотя и неформальные, но фактически *окончательные* аргументы в пользу этого предположения.

12. КОЛЛАПС СИТУАЦИИ НЕОПРЕДЕЛЕННОСТИ

Это преобразование сводится к тому, что из всего множества возможностей, связанных с данной ситуацией, выделяется *одна* возможность, в результате чего сама ситуация неопределенности прекращает свое существование. Но такое выделение одной возможности может произойти двумя различными способами — *внутренним* и *внешним*. Эти способы настолько непохожи, что между ними на первый взгляд даже трудно увидеть что-либо общее.

Так, коллапс ситуации *внутренним способом* сводится к тому, что наступает событие, ранее представлявшее собой одну из возможностей данной ситуации неопределенности. Это может произойти как самопроизвольно, так и в результате чьего-то сознательного выбора. Примером самопроизвольного коллапса может служить наступление случайного события в вероятностной ситуации. Но выделение одной возможности может произойти также *внешним способом* — в виде крайнего частного случая преобразования *ограничение разнообразия*.

1. Коллапс по внутренней причине. Предположим, что по-прежнему рассматривается вероятностная ситуация, описываемая пространством вероятностей (X, p). Если в данной ситуации произошло внутреннее событие $b \in X$, это значит, что:

1) сама ситуация неопределенности, описываемая пространством вероятностей (X, p), перестала существовать (*коллапсировала*);

2) вместо ситуации неопределенности появилась *новая информация*, которая содержится в событии *b*.

Сразу же возникает вопрос о *количестве* этой новой информации. Ответ состоит в том, что оно совпадает с *количеством собственной информации* $I_p(b)$ события *b*, которое согласно [1, (2)] характеризуется выражением

$$I_p(b) = \log\frac{1}{p(b)}\,. \tag{7}$$

Теперь рассмотрим следующий вопрос: *чт***о** *такое собственная информация события b*, т.е. что именно измеряет число, названное количеством собственной информации $I_p(b)$ события *b*?

До сих пор широко распространено представление (первоначально возникшее под влиянием взглядов Норберта Винера), что информация аналогична *физической энтропии*. Как мы уже отмечали в первой и второй частях статьи, это представление нанесло значительный вред кибернетике, поскольку оно мешало изучать свойства информации. В частности, оно помешало увидеть следующее важное отличие между информацией и физической энтропией. Об информации всегда можно спросить *«о чем эта информация, к чему она относится?»*, в то время как по отношению к энтропии такой вопрос не имеет никакого смысла.

Если задать вопрос *«к чему относится собственная информация случайного события?»*, то могут быть получены даже *два* ответа (отражающие два фундаментальные свойства меры собственной информации). Вот эти ответы:

1. Собственная информация случайного события — это ***вся*** информация, которую данное событие может содержать ***о чем угодно*** (т.е. обо всех других событиях, явлениях или состояниях).

2. Собственная информация случайного события — это ***полная*** информация ***об этом событии***, т.е. это вся та информация, которую нужно было бы получить из каких угодно источников, для того чтобы убедиться, что из всех возможных событий произошло именно данное событие.

Заметим, что эти два ответа в точности соответствуют двум темам этого раздела — коллапс ситуации по внутренней и внешней причине.

Рассмотрим более подробно, чт***о*** означает первый ответ. Интуитивно кажется довольно естественным считать, что если априорная вероятность этого события была велика, то факт, что это событие произошло, несет в себе мало информации (о других событиях, явлениях или состояниях), поскольку и так было почти ясно, что оно произойдет. И обратно, если произошло очень маловероятное событие, то мы обычно склонны считать, что получили много информации (так как это может привести к существенной переоценке вероятностей каких-то других событий).

Теперь рассмотрим второй ответ. Если мы еще не знаем, что данное событие произошло, но вероятность этого велика, то для того чтобы убедиться в совершившемся факте, потребуется не очень много информации. Другое дело, если вероятность данного события очень мала. Тогда для того чтобы убедиться в том, что оно все же произошло, потребуется гораздо больше информации.

2. Коллапс по внешней причине. Превращение внешней информации во внутреннюю. В первой части статьи был сделан вывод о том, что информация, количество которой измеряется выражениями [1, (2)] и (10), является *внутренним свойством* ситуации неопределенности, описываемой пространством вероятностей (X, p). И на этом основании упомянутую информацию мы назвали ***внутренней***.

Однако особенность преобразования *коллапс ситуации неопределенности* состоит в том, что с ним можно связать также ***внешнюю*** информацию, количество которой можно затем сравнить с количеством внутренней информации (7). Действительно, выделение одной возможности может произойти также *внешним* способом — в виде крайнего частного случая преобразования *ограничение разнообразия*. Так, иногда может случиться, что информации, поступившей извне, достаточно для уменьшения всего множества возможностей до одного элемента.

Именно такую возможность подразумевал второй ответ на вопрос «*К чему относится собственная информация случайного события?*». Но в свете сказанного выше из этого ответа должно следовать, что

Количество внешней информации, необходимой для того, чтобы в пространстве вероятностей (X, p) выделить единственную точку b, должно совпасть с *количеством внутренней информации* $I_p(b)$.

Иначе говоря, в данном случае ***внешняя информация должна превратиться во внутреннюю***. Этот вывод является одновременно ожидаемым и неожиданным. И в любом случае он заслуживает проверки. Из раздела 11 нам уже известен способ вычисления количества внешней информации, связанной с преобразованием *ограничение разнообразия*. Теперь мы применим этот способ к нашему крайнему частному случаю, названному *коллапсом по внешней причине*.

3. Информационный анализ коллапса по внешней причине. Итак, предположим, что ограничение разнообразия свелось к тому, что теперь вместо множества X мы должны рассматривать его *одноэлементное* подмножество $A = b$, где $b \in X$. Для того чтобы узнать, сколько для этого требуется внешней информации, достаточно в равенство (6) вместо A подставить $\{b\}$. В результате получим

$$E(X, p \mid \{b\}) = E\left(X, p \,\|\, p^{\{b\}}\right). \tag{8}$$

Для нахождения же числа $E\left(X, p \,\|\, p^{\{b\}}\right)$ можно снова воспользоваться равенством (6), выполнив ту же подстановку и учтя два очевидных равенства $p^{\{b\}}(b) = 1$ и $G\left(\{b\}, p^{\{b\}}\right) = 0$. В результате получим

$$E(X, p \mid \{b\}) = E\left(X, p \,\|\, p^{\{b\}}\right) = G\left(X,\ p^{\{b\}} \lozenge p\right) = \log\frac{1}{p(b)} = I_p(b). \tag{9}$$

Итак, упомянутый выше ожидаемо-неожиданный вывод оказался верным: количество внешней информации $E(X, p \mid \{b\})$ оказалось равным количеству внутренней информации $I_p(b)$. Этот результат важен по следующим двум причинам. Во-первых, он дает основание утверждать, что предложенный выше (в виде предположения 2) подход к нахождению мер внешней информации, опирающийся на новое понятие *ситуация заблуждения*, выдержал очень серьезное испытание. Во-вторых, упомянутый вывод фактически представляет собой первое свидетельство полной согласованности между мерами внутренней и внешней информации.

Заметим, что коллапс ситуаций неопределенности (как внутренний, так и вызванный внешними причинами) представляет собой главный механизм ***создания новой информации***. Сколько именно информации таким образом создается, зависит от двух условий: 1) типа и характера исходной ситуации неопределенности и 2) того, какая именно из возможностей осуществилась.

13. МЕРЫ ВНУТРЕННЕЙ И ВНЕШНЕЙ ВЗАИМНОЙ ИНФОРМАЦИИ

Сейчас мы рассмотрим две меры внутренней и внешней информации, которые играют очень важную роль в теории информации. Вместе с тем, может показаться, что наши представления о двух видах отношения информации к ситуациям неопределенности, выражаемых словами *«внутренняя и внешняя информация»*, в данном случае оказываются несостоятельными. В самом деле, в теории информации эти две меры вводятся таким образом, что ту из них, которую по нашей терминологии следует именовать *мерой внешней взаимной информации*, не связывают ни с каким преобразованием ситуаций неопределенности.

Нарушая принятый в этой статье порядок изложения, мы здесь поступим точно так же, т.е. сначала построим эти две меры чисто формально. И только после этого рассмотрим вопрос о том, с каким преобразованием действительно связана мера внешней взаимной информации.

Пусть задана двумерная ситуация неопределенности (как в [2, разд. 8]), которая описывается (тоже двумерным) пространством вероятностей $(X \times Y, \pi)$ (где π есть некоторое РВ, *действующее* на произведении дискретных множеств $X \times Y$).

1. Мера внутренней взаимной информации. Для всяких $x \in X$ и $y \in Y$ число

$$I_{\pi}(x \leftrightarrow y) = \log \frac{\pi(x,y)}{\mathbf{pr}_1\pi(x) \cdot \mathbf{pr}_2\,\pi(y)} \tag{10}$$

будем называть **количеством внутренней взаимной информации** между элементами x и y в пространстве $(X \times Y, \pi)$.

Заметим, что x и y являются не элементами пространства $(X \times Y, \pi)$, а элементами соответственно первой и второй его *проекций* $(X, \mathbf{pr_1}\,\pi)$ и $(Y, \mathbf{pr_2}\,\pi)$ [2, разд. 8]. Вообще говоря, проекции (двумерного) пространства вероятностей $(X \times Y, \pi)$ *не являются независимыми*, а мера $I_{\pi}(x \leftrightarrow y)$ измеряет степень зависимости между элементами x и y этих проекций.

Число $I_{\pi}(x \leftrightarrow y)$ измеряет количество: 1) той информации, которую элемент x содержит об элементе y; 2) той информации, которую элемент y содержит об элементе x. Можно показать, что для всех $x \in X$ и $y \in Y$ имеют место соотношения

$$I_{\pi}(x \leftrightarrow y) = I_{\mathbf{pr_1}\pi}(x) + I_{\mathbf{pr_2}\,\pi}(y) - I_{\pi}(x,y), \tag{11}$$

$$I_{\mathbf{pr_1}\pi}(x) \ge I_{\pi}(x \leftrightarrow y) \le I_{\mathbf{pr_2}\,\pi}(y). \tag{12}$$

А из неравенств (12) следует, что число $I_{\pi}(x \leftrightarrow y)$ не может превысить ни количество $I_{\mathbf{pr_1}\,\pi}(x)$ собственной информации элемента x, ни количество

$I_{\mathbf{pr}_2\,\pi}(y)$ собственной информации элемента y. Это хорошо согласуется с теми разъяснениями по поводу свойств количества собственной информации, которые были даны выше (разд. 12, п. 1).

2. Мера внешней взаимной информации. Рассмотрим *информационную функцию*

$$I_{\pi}(\bullet \leftrightarrow \bullet) = (x, y) \mapsto I_{\pi}(x \leftrightarrow y) \lozenge X \times Y\,. \tag{13}$$

Очевидно, что функция $I_{\pi}(\bullet \leftrightarrow \bullet)$ является случайной величиной относительно действующего на множестве $X \times Y$ распределения π. Математическое ожидание этой случайной величины мы здесь обозначим $E(X \leftrightarrow Y \mid \pi)$ и назовем **количеством внешней взаимной информации** между проекциями (двумерной) ситуации неопределенности, описываемой пространством вероятностей $(X \times Y, \pi)$:

$$E(X \leftrightarrow Y \mid \pi) = \mathbf{E}\, I_{\pi}(\bullet \leftrightarrow \bullet) = \sum_{(x,y)\in X \times Y} \pi(x,y) \cdot \log \frac{\pi(x,y)}{\mathbf{pr}_1\pi(x) \cdot \mathbf{pr}_2\,\pi(y)}\,. \tag{14}$$

В теории информации эта мера получила несколько неудачных названий и еще более неудачных обозначений. Так, в книгах [4, 5] она была названа *средней взаимной информацией* ([4, разд. 2.9], [5, разд. 2.2]), а в книгах [6, 7] — даже взаимной информацией (или информацией связи) *двух случайных величин* ([6, § 1.1], [7, § 6.2]). Последнее название является вообще неадекватным, поскольку эта мера относится не к величинам, а к *событиям*. А все традиционные обозначения этой меры были совершенно неудовлетворительными по двум разным причинам. Во-первых, отсутствовал подходящий для создания обозначений язык. Во-вторых, не было возможности учесть отличие между *внутренними* и *внешними* мерами ввиду отсутствия этих понятий.

Величина $E(X \leftrightarrow Y \mid \pi)$ играет очень важную роль в теории информации. Можно показать, что имеет место равенство, аналогичное (11):

$$E(X \leftrightarrow Y \mid \pi) = G(X, \mathbf{pr}_1\pi) + G(Y, \mathbf{pr}_2\pi) - G(X \times Y, \pi)\,. \tag{15}$$

Кроме того, очевидно, справедливы также неравенства

$$G(X, \mathbf{pr}_1\pi) \ge E(X \leftrightarrow Y \mid \pi) \le G(Y, \mathbf{pr}_2\pi)\,, \tag{16}$$

аналогичные неравенствам (12).

Однако необходимо помнить, что, несмотря на вышесказанное, величины $I_{\pi}(x \leftrightarrow y)$ и $E(X \leftrightarrow Y \mid \pi)$ отнюдь не аналогичны, так как относятся к разным *категориям*. И в то время как количество *внешней* взаимной информации всегда неотрицательно:

$$E(X \leftrightarrow Y \mid \pi) \ge 0\,, \tag{17}$$

количество *внутренней* взаимной информации $I_{\pi}(x \leftrightarrow y)$ может быть как положительным, так и отрицательным.

14. С КАКИМ ПРЕОБРАЗОВАНИЕМ СИТУАЦИЙ НЕОПРЕДЕЛЕННОСТИ СВЯЗАНА МЕРА ВНЕШНЕЙ ВЗАИМНОЙ ИНФОРМАЦИИ?

Для ответа на этот вопрос необходимо сначала рассмотреть одно очень простое преобразование ситуаций неопределенности.

1. Переход к совместному рассмотрению двух информационно независимых ситуаций неопределенности. Рассмотрим две ситуации неопределенности, (дискретные) множества возможностей которых обозначим X и Y. И предположим, что вместо отдельного рассмотрения событий $x \in X$, относящихся к первой ситуации, и событий $y \in Y$, относящихся ко второй, мы хотим рассматривать *комплексные события*, которые могут изображаться парами вида (x, y). Переход к рассмотрению пар (x, y) в качестве новых комплексных событий означает, что теперь вместо двух ситуаций неопределенности мы фактически рассматриваем одну ***двумерную*** ситуацию, множеством возможностей которой является произведение $X \times Y$ множеств X и Y.

Теперь наша задача состоит в том, чтобы суметь эту новую ситуацию описать. Когда обе исходные ситуации неопределенности являются вероятностными, способы построения описания новой ситуации известны (но эти способы различны в зависимости от того, имеется ли информационная связь между этими ситуациями, или они независимы). Пока мы займемся описанием новой (двумерной) ситуации при условии независимости исходных ситуаций.

Пусть на множествах X и Y *действуют* распределения вероятностей (РВ) p и q соответственно. Тогда обе рассматриваемые ситуации могут быть исчерпывающим образом описаны двумя пространствами вероятностей (X, p) и (Y, q). Предположим, что нет никаких сведений о связи между распределениями p и q, так что эти ситуации можно считать независимыми. Теория вероятностей предлагает способ, позволяющий найти описание новой (двумерной) ситуации (однако *не предлагает* ни подходящий для этих целей язык, ни систему обозначений). Если исходные ситуации между собой *не связаны* (информационно независимы), то на множестве $X \times Y$ должно действовать РВ $p \times q$, которое характеризуется следующим условием: для любых $x \in X$ и $y \in Y$ событие (x, y) может произойти с вероятностью $p \times q\,(x, y)$, определяемой равенством

$$p \times q(x, y) = p(x) \cdot q(y). \tag{18}$$

Таким образом, мы имеем полное описание новой ситуации неопределенности (которая тоже является вероятностной). Она характеризуется распределением вероятностей $p \times q$, действующим на множестве возможностей $X \times Y$. Формальным описанием такой ситуации является (двумерное) пространство вероятностей $(X \times Y, p \times q)$. Распределение $p \times q$ будем называть **произведением** пары распределений (p, q), а пространство вероятностей $(X \times Y, p \times q)$ естественно назвать **произведением пространств** (X, p) и (Y, q) (взятых в указанном порядке).

Легко убедиться в том, что произведение $p \times q$ пары распределений (p, q) обладает следующими свойствами:

$$\mathbf{pr_1}\, p \times q = p, \quad \mathbf{pr_2}\, p \times q = q, \tag{19}$$

где $\mathbf{pr_1}$ и $\mathbf{pr_2}$ — операции проектирования, описанные во второй части статьи [2, разд. 8, п. 1].

2. Две интерпретации меры внешней взаимной информации. Пусть имеется некоторая двумерная ситуация неопределенности, которая описывается пространством вероятностей $(X \times Y, \pi)$ (где π — есть РВ, действующее на произведении дискретных множеств $X \times Y$). Рассмотрим две проекции этой ситуации, описания которых имеют вид $(X, \mathbf{pr_1}\, \pi)$ и $(Y, \mathbf{pr_2}\, \pi)$ [2, разд. 8, п. 1]. И рассмотрим новую (двумерную) ситуацию неопределенности, которая может быть описана пространством вероятностей $(X \times Y,\ \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi)$. Ввиду сказанного выше, проекции этой новой ситуации *информационно независимы*. Следовательно, условием информационной независимости проекций исходной ситуации, описываемой пространством $(X \times Y, \pi)$, является равенство

$$\pi = \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi . \tag{20}$$

Теперь можно предложить две интерпретации меры внешней взаимной информации $E(X \leftrightarrow Y \mid \pi)$ (14). Применив соглашение (18) к выражению (14), получим следующее:

$$E(X \leftrightarrow Y \mid \pi) = \sum_{(x,y) \in X \times Y} \pi(x,y) \cdot \log \frac{\pi(x,y)}{\mathbf{pr}_1 \pi \times \mathbf{pr}_2\, \pi(x,y)} . \tag{21}$$

Иначе говоря, если вспомнить выражение [2, (37)], то можно написать

$$E(X \leftrightarrow Y \mid \pi) = G(X \times Y, \pi \Diamond \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi) - G(X \times Y, \pi) , \tag{22}$$

где

$$G(X \times Y, \pi \Diamond \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi) = \sum_{(x,y) \in X \times Y} \pi(x,y) \cdot \log \frac{1}{\mathbf{pr}_1 \pi \times \mathbf{pr}_2\, \pi(x,y)} \tag{23}$$

— мера неопределенности *ситуации заблуждения*, описываемой пространством двойных вероятностей $(X \times Y,\ \pi \Diamond \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi)$ [2, разд. 10, п. 1]. Эта ситуация такова: в действительности на множестве $X \times Y$ действует РВ π, но тот, кто принимает некое решение, думает, что действующим является распределение $\mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi$. Это значит, что имеет место равенство

$$E(X \leftrightarrow Y \mid \pi) = E(X \times Y,\ \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi \,\|\, \pi) . \tag{24}$$

Из этого равенства и вытекают упомянутые две интерпретации меры $E(X \leftrightarrow Y \mid \pi)$. Эти интерпретации настолько важны, что заслуживают отдельного неформального разъяснения.

1. Согласно равенству (24) *количество внешней взаимной информации* $E(X \leftrightarrow Y \mid \pi)$ совпадает с количеством $E(X \times Y,\ \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi \,\|\, \pi)$ *внешней информации* в пользу гипотезы о том, что верным описанием рассматриваемой ситуации является пространство вероятностей $(X \times Y, \pi)$, против гипотезы о том, что ее нужно описывать пространством вероятностей $(X \times Y,\ \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi)$ (проекции которого ***независимы***).

2. Согласно равенству (24) и предположению 2 [2, разд. 10, п. 2] число $E(X \leftrightarrow Y \mid \pi)$ измеряет также количество внешней информации, которое необходимо для преобразования ситуации неопределенности, описываемой пространством $(X \times Y,\ \mathbf{pr_1}\, \pi \times \mathbf{pr_2}\, \pi)$ (с независимыми проекциями) в ситуацию, описываемую пространством $(X \times Y, \pi)$.

15. ПОНЯТИЕ ИНФОРМАЦИОННОГО КАНАЛА

Обычно, когда говорят о передаче информации (или о передаче сигналов), имеют в виду передачу по специально созданному для этой цели *каналу связи*. Такое понимание объясняется известными постановками классической теории информации. Однако с момента создания теории информации прошло более чем полстолетия. И сейчас уже можно с уверенностью утверждать, что сигналы, несущие информацию, передаются не только по искусственно созданным каналам связи, и не только по каналам, созданным биологической эволюцией, но что подобные процессы передачи происходят повсеместно в нашей Вселенной.

Иначе говоря, информация во Вселенной не пассивна. Она ничем не напоминает тот мрачный образ, который лелеял Норберт Винер и который существенным образом повлиял на его понимание созданной им новой науки — кибернетики. Под информацией, существующей во Вселенной, Винер понимал некий реликт прошлого, неспособный обновляться или увеличивать свое количество, а способный только постепенно деградировать. Вот цитата из его книги «Кибернетика и общество»:

«Тот факт, что информация может быть потеряна, а не приобретена, является, как мы видели, кибернетической формой второго закона термодинамики… Подобно тому, как в замкнутой системе энтропия [термодинамическая] стихийно стремится к увеличению, точно так же информация стремится к уменьшению» [8, с. 87, 123].

Сейчас уже ясно, что представления Винера об информации, как следует из приведенного отрывка, фактически противоречили идеям и результатам Шеннона. В самом деле, очевидно, что введенные Шенноном понятия *источника информации* и *канала связи* имеют аналоги в Природе. Так, источником информации может служить любое событие во Вселенной, любое изменение состояния какого-либо объекта, любое столкновение или трансмутация элементарных частиц (в частности, распад элементарной частицы или, наоборот, ее рождение). Всё это случаи, когда *создается новая информация*. А поскольку такого (или какого-нибудь другого) рода события происходят все время и повсеместно, это значит, что и новая информация создается постоянно и бесперебойно. Из сформулированного же во второй части настоящей статьи *достаточного условия наличия информационной связи* [2, разд. 9, п. 1] следует, что эта новая информация может передаваться в другие системы.

Поэтому появилась необходимость в выяснении механизма (или способа) *возникновения* информационных связей, а также механизма *функционирования* этих связей (заметим, что упомянутое *достаточное условие* никаких сведений об этих двух механизмах не содержит).

1. **Информационный канал и информационная связь.** Итак, каким же образом устанавливается информационная связь между двумя системами? Для этого достаточно, чтобы существовал *информационный канал* (в дальнейшем иногда будем говорить просто «канал»), по которому могли бы передаваться *сигналы* (содержащие *информацию*) от одной из этих систем к другой. Уже из одной только этой фразы следует, что информационный канал и информационная связь — не одно и то же. Имеются два принципиальных отличия:

1) канал материален, а информационная связь — нет;

2) канал действует только в одну сторону (в направлении от причины к следствию), а информационная связь — в обе стороны.

Известно, что *передача сигналов* (*а, следовательно, и информации*) — *это материальный процесс*. С другой стороны, до сих пор никто и никогда не обратил внимания на то, что верно и обращение этой фразы:

Любой материальный процесс ***всегда*** *связан с передачей некоторых сигналов и, следовательно, с передачей информации. Это значит, что* ***любой материальный процесс является информационным каналом***.

Это совершенно новый поворот. И он требует осмысления. Рассмотрим примеры.

Любая частица, летящая в космосе, несет в себе информацию о разнообразных вещах. Тип частицы, ее импульс, энергия, заряд — это далеко не полный список того, что можно о ней узнать, располагая соответствующей аппаратурой. А некоторая часть содержащейся в данной частице информации может быть (при определенных условиях) использована для того, чтобы кое-что узнать и о тех процессах, которые породили эту частицу (или отправили ее в путь). Далее, пронизывающие всю Вселенную электромагнитные волны — свет, радиоволны, рентгеновское и гамма-излучение — все это примеры материальных процессов, непосредственно связанных с передачей информации из одной части Вселенной в другую (специалисты активно используют эти процессы для изучения Вселенной).

Одним из наиболее впечатляющих достижений XX века явилось приобретение знаний о том, чт***о*** происходило с нашей Вселенной буквально с самого момента ее рождения (по современным представлениям это случилось около 14-и миллиардов лет назад). Пристальное изучение происходивших тогда процессов и привело к современным представлениям об основных этапах развития Вселенной. Этой теме посвящена обширная литература, в том числе и популярная. Здесь можно сослаться на следующие книги: [9–13] (особенно впечатляют названия книг С. Вайнберга «Первые три минуты» [9] и Дж. Нарликара «Неистовая Вселенная» [10]).

Однако снова всплывает уже поднимавшийся ранее вопрос: как вообще возможно изучение (да еще *пристальное*!) тех процессов, которые происходили так давно? Ведь, как мы уже отмечали в [2, разд. 9, п. 1], с точки зрения современной философии науки и теории познания ***научное изучение прошлого невозможно***. Конечно, *буквально так никто не говорит*! Однако до сих пор считается неоспоримой истиной, что научные знания можно получать, только опираясь на *принцип причинности* (опора на этот принцип считается признаком научности). С другой стороны, какие-либо выводы о прошлом таким способом получить нельзя, поскольку *настоящее не может быть причиной прошлого*. Отсюда непосредственно следует выделенное выше утверждение.

Тем не менее изучение прошлого нашей Вселенной дало поразительные результаты. Сейчас уже известно, каким образом и почему возникло вещество (сначала наиболее элементарная его разновидность — водород); как возникли звезды первого поколения, которые со временем превратились в автоматические “заводы” по переработке водорода в другие химические элементы; как появлялись звезды дальнейших поколений, в области притяжения которых уже могли образовываться планеты; как возникали галактики и их скопления; и, наконец, с чего все это началось — с Большого Взрыва. Почему оказалось возможным все это узнать? Потому, что все упомянутые процессы и порождаемые ими сигналы, которые (после их доставки) были доступны непосредственному наблюдению, несли в себе *информацию* (так что их изучение позволяло часть этой информации превращать в *знания*).

И все же, приведенные примеры (несмотря на их вселенский масштаб) могут показаться недостаточно убедительной иллюстрацией к утверждению, что *любой* материальный процесс связан с передачей информации, поскольку эти примеры ни в малейшей степени не отражают представления о "любом материальном процессе". Поэтому теперь рассмотрим несколько примеров из повседневной жизни.

Никого не удивишь заявлением, что в разнообразных приборах (таких как микроскоп, телескоп, барометр, весы, термометр) происходит передача информации.

Несомненно, найдутся желающие внести в это заявление поправку, утверждая, что приборы не передают, а "перерабатывают" информацию. В такой поправке есть доля истины. Но в этой статье нет возможности разобраться в том, какова эта доля.

Однако лишь немногие были бы готовы признать, что информация передается от прошлых состояний к будущим даже в таких машинах, как мясорубка, соковыжималка, стиральная машина, мельница. Предположим, например, что в некой лаборатории проводится химический эксперимент. Можно ли сказать, что результат только что прошедшей химической реакции содержит в себе информацию об исходном состоянии системы, которое существовало до реакции? Разумеется! Ведь зная результат реакции, можно довольно много (но, обычно, не все) узнать об исходном состоянии системы. А это было бы совершенно невозможно, если бы в упомянутом результате реакции *объективно* не содержалась информация об исходном состоянии!

Еще пример: процесс приготовления пищи. Легко понять, что этот процесс тоже связан с передачей информации. Действительно, конечный результат (т.е. готовая пища) несет в себе информацию об исходных продуктах. А сами исходные продукты несут информацию о тех условиях, в которых они выросли (лес, поле, огород), или о тех процессах, с помощью которых они были произведены.

Из всего этого следует, что сама Вселенная является гигантским клубком переплетенных между собой информационных каналов. А примеры, подтверждающие наш тезис, можно было бы приводить без конца. Однако в этом нет необходимости, поскольку очевидно, что любая материальная система содержит огромное количество информации, в том числе — информацию о своем прошлом и о своем будущем. Но для того чтобы система содержала информацию о своем прошлом, необходимо, чтобы информация непрерывно передавалась из прошлого в будущее. И это действительно происходит, независимо от того, какие материальные процессы идут внутри системы. Фактически информация передается именно через эти материальные процессы (какими бы они ни были).

В первой части статьи (разд. 1) отмечено, что широко распространенные представления об информации сводятся к тому, что информацию отождествляют либо со *сведениями* («Словарь по кибернетике» [14], «Энциклопедия кибернетики» [15]), либо с *текстами* (концепция так называемой “переработки информации” в компьютерах). Для тех, кто придерживается таких представлений, все вышесказанное может звучать дико. Действительно, им должно быть совершенно непонятно, какие такие *сведения* или, тем более, *тексты* могут передаваться из прошлого в будущее, скажем, в процессе приготовления пищи. Но рассмотренные выше примеры как раз и демонстрируют еще раз всю нелепость упомянутых распространенных представлений об информации.

2. Две роли материальных процессов. Итак, из сказанного следует, что любой материальный процесс исполняет одновременно как бы две роли. Первая из этих ролей широко известна, так как ею уже давно заинтересовалась наука (физика, химия, биология). Эта роль и подразумевается всегда, когда говорят о материальном процессе. В самом общем понимании эта роль состоит в следующем:

1. Любой материальный процесс переводит некоторую материальную систему из одного состояния в другое.

Но нас здесь будет интересовать как раз *вторая роль* материальных процессов. В наиболее общей формулировке эта роль состоит в следующем:

2. Новое состояние, в которое перешла материальная система в результате некоторого материального процесса, может рассматриваться как ***сигнал****, несущий как информацию об этом новом состоянии системы, так и информацию (обычно неполную) о причинах перехода системы в новое состояние (в частности, об исходном ее состоянии).*

Необходимо заметить, что в связи с этой второй ролью материальных процессов мы оказались в совершенно незнакомой ситуации, не имевшей прецедентов в истории науки. Действительно, первое, что сразу становится очевидным, — это то, что о способах изучения материальных процессов в этой второй роли неизвестно буквально ничего. Начинать придется с нуля, а точнее — с выработки хотя бы некоторых базовых понятий.

Заметим, что здесь намечается целый пласт совершенно новых сложнейших теоретических и экспериментальных проблем, решением которых (наряду с большим количеством других чисто *кибернетических* проблем) уже давно должна была заняться кибернетика, вместо того чтобы заниматься не своим делом — разработкой компьютеров.

Основным из искомых базовых понятий является понятие *информационного канала*.

Информационный канал — это произвольный *материальный процесс* (*материальная система* или *среда* — естественная или искусственная), который рассматривается исключительно с точки зрения возможности *передачи сигналов*, а также возможности изучения *информационных свойств* процесса передачи (т.е. характера *помех*, влияющих на качество передачи).

Эта формулировка, конечно, не претендует на то, что она способна полностью раскрыть смысл нового понятия *информационный канал*. Следовательно, ее нельзя рассматривать и как определение этого понятия. И вообще, ни о каких определениях пока речь идти не может, поскольку сейчас с понятием *информационный канал* еще связано слишком много неясностей. Эти неясности могут быть устранены только в процессе дальнейшей работы — работы, объем которой даже трудно себе представить.

Что это за работа? Необходимо уже сейчас наметить хотя бы некоторые задачи. Пока имеет смысл выделить всего три *главные задачи*.

1. Научиться *описывать* разнообразные информационные каналы, а также выяснить, возможна ли единая схема такого описания.

2. Разработать систему *информационных мер*, отражающих свойства каналов.

3. Дать примеры описания некоторых конкретных каналов и применения информационных мер.

Заметим, что первые две из трех выделенных задач фактически представляют собой сложнейшие проблемы. С другой стороны, нетрудно сообразить, что в классической теории информации был достаточно подробно рассмотрен лишь один (очень узкий) частный случай обеих проблем. Действительно, изучавшиеся в теории информации *каналы связи* являются специфической разновидностью информационных каналов. Поэтому можно было бы подумать, что для решения упомянутых проблем удастся использовать опыт теории информации.

Однако на пути использования такого опыта имеется одно препятствие, которое кажется непреодолимым. Главная трудность состоит в следующем. Наивно было бы рассчитывать на то, что информационные свойства упомянутых материальных систем и процессов удастся описать на том языке — *вероятностном*, — который использовался в теории информации для описания каналов связи. Действительно, хорошо известно, что в Природе вероятностные ситуации (в точном значении этих слов) встречаются крайне редко — чаще всего возникают ситуации неопределенности неизвестных и неизученных типов. Это значит, что возможность *корректного* применения теории вероятностей для описания разнообразных ситуаций неопределенности может возникать только в исключительных случаях. С другой стороны, мы уже показали, что нельзя изучать информацию, не рассматривая какие-либо ситуации неопределенности. Сказанное в полной мере относится и к информационным свойствам каналов. Но в таком случае мы приходим к выводу, что проблема изучения информационных каналов не может быть решена без развития соответствующих разделов *теории ситуаций неопределенности*.

К счастью выяснилось, что проблема нахождения подхода к формальному описанию информационных каналов для своего решения не требует каких-то новых радикальных идей. Действительно, как показано в работе [16], всякий информационный канал можно рассматривать как некий *пучок ситуаций неопределенности* (как бы растущий из множества возможных состояний на входе этого канала). Поэтому для получения языка, пригодного для формального описания информационных каналов, достаточно суметь формально описать два понятия: *пучок* и *ситуация неопределенности*. В работе [16] было показано, что для формального описания первого из них вполне подходит математическое понятие *семейства*. А предложенному автором *универсальному* способу описания разнообразных ситуаций неопределенности посвящена большая серия статей. Этот способ сводится к понятию *пространство неопределенности* [17, 18].

*****

В следующей (заключительной) части статьи будут рассмотрены два преобразования информации, связанные с информационными каналами, — *индукция* и *вынуждение* ситуаций неопределенности на выходе канала.

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. *Дидук Н.Н.* Меры внутренней и внешней информации (на примере вероятностных ситуаций неопределенности). Часть II // Системні дослідження та інформаційні технології. — 2012. — № — С. 94–110.
3. *Дидук Н.Н.* Прообразы пространств неопределенности. Простые подпространства // Системні дослідження та інформаційні технології. — 2005. — № 1 — С. 127–142.
4. *Фано Р.* Передача информации. Статистическая теория связи. — М.: Мир, 1965. — 440 с.
5. *Галлагер Р.* Теория информации и надежная связь. — М.: Советское радио, 1974. — 720 с.
6. *Чисар И., Кёрнер Я.* Теория информации. — М.: Мир, 1985. — 400 с.
7. *Стратонович Р.Л.* Теория информации. — М.: Советское радио, 1975. — 424 с.
8. *Винер Н.* Кибернетика и общество (The Human use of Human Beings). — М.: ИЛ, 1958. — 200 с.
9. *Вайнберг С.* Первые три минуты. Современный взгляд на происхождение Вселенной. — М.: Энергоиздат, 1981. — 208 с.
10. *Нарликар Дж.* Неистовая Вселенная. — М.: Мир, 1985. — 256 с.
11. *Новиков И.Д.* Как взорвалась Вселенная. — М.: Наука. Гл. ред. физ.-мат. лит., 1988. — 176 с.
12. *Хокинг С.* Краткая история времени. От Большого Взрыва до черных дыр. — Санкт-Петербург: Амфора, 2003. — 270 с.
13. *Хокинг С.*, *Млодинов Л.* Кратчайшая история времени. — Санкт-Петербург: Амфора, 2007. — 180 с.
14. *Словарь* по кибернетике. — Киев: Гл. ред. УСЭ, 1979. — 624 с.
15. *Энциклопедия* кибернетики. — Киев: Гл. ред. УСЭ, 1974. Том 1. — 608 с. Том 2. — 624 с.
16. *Дидук Н.Н.* Информационные каналы как развитие представлений о каналах связи // Системні дослідження та інформаційні технології. — 2007. — № 1. — С. 129–141.
17. *Дидук Н.Н.* Пространства неопределенности и изоморфизм // Системні дослідження та інформаційні технології. — 2002. — № 4. — С. 128–143.
18. *Дидук Н.Н.* Пространства неопределенности. Энтропия и теорема кодирования // Кибернетика. — 1984. — № 2. — С. 69–73.

Поступила 12.06.2010

Статтю надруковано під редакцією автора