УДК 519.7

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

Н.Н. ДИДУК

Показано, что внешняя информация, вызывающая преобразования ситуаций неопределенности, может поступать по информационным каналам. Это происходит в два этапа. Сначала само существование канала создает внешнюю информацию, которая приводит к *индукции* ситуации неопределенности на выходе канала и к *образованию* информационной связи между двумя системами. После этого становится возможным *функционирование* образовавшейся информационной связи. Материальный аспект функционирования сводится к порождению сигнала на входе канала и последующей его передаче по каналу. Информационный же аспект состоит в том, что каждое событие появления сигнала на входе или на выходе канала создает внешнюю информацию, которая ведет к *вынужденному* изменению ситуации неопределенности на противоположном конце канала. Подвергнуты изучению три преобразования информации, связанные с информационными каналами: *индукция* ситуации неопределенности на выходе канала, *образование информационной связи* между двумя системами и *вынуждение* ситуации неопределенности на выходе канала. Получено свидетельство того, что внутренняя информация может превращаться во внешнюю. Сформулированы выводы из статьи.

В первых трех частях статьи [1–3] рассмотрен ряд преобразований информации и построены соответствующие меры интенсивности этих преобразований. Эти преобразования были связаны как с потерей информации, так и с ее приобретением. Однако вопрос о том, откуда берется приобретаемая (внешняя) информация, до сих пор оставался открытым. В третьей части статьи введено новое понятие *информационного канала*. А в настоящей четвертой части показано, что внешняя информация может быть либо *передана* по информационному каналу, либо *индуцирована* самим существованием канала.

Итак, в этой (заключительной) части статьи изучаются некоторые (но не все) преобразования информации, связанные с информационными каналами. Рассматриваются две системы A и B, для которых определено понятие *состояния*, и предполагается, что конкретные состояния, в которых находятся обе системы, нам неизвестны. Это значит, что для нас существуют только *ситуации неопределенности*, касающиеся состояний обеих этих систем, и до тех пор, пока между системами A и B не образуется информационная связь, эти ситуации будут *независимыми*. А что нужно, чтобы информационная связь образовалась? Для этого достаточно, например, чтобы эти системы были соединены *информационным каналом*. Поскольку мы уже знаем, что канал может действовать только в каком-нибудь одном направлении, для определенности предположим, что речь идет о канале, который будет действовать в направлении от системы A к системе B.

Сразу подчеркнем, что следует отличать ***образование*** информационной связи от ***функционирования*** этой связи. Это разные преобразования, которые происходят под воздействием разных порций внешней информации: сначала информационная связь должна образоваться, а затем она может начать функционировать. *Образование* информационной связи не предполагает, что стало известным состояние какой-либо из систем A и B (т.е. еще не предполагает появления на входе или на выходе канала какого-либо сигнала). Вопросы же, касающиеся изменений в системах A и B после того как стало известным состояние одной из них, относятся уже к *функционированию* информационной связи. Это изменение (которое зависит от ситуации неопределенности в системе A и от соединяющего обе системы канала) здесь названо ***индукцией*** ситуации неопределенности на выходе канала (вызванной ситуацией на его входе). *Внутренний* же (наиболее важный) аспект образования информационной связи состоит в том, что системы A и B перестают быть информационно независимыми.

Начнем с рассмотрения индукции.

16. ИНДУКЦИЯ СИТУАЦИИ НЕОПРЕДЕЛЕННОСТИ НА ВЫХОДЕ КАНАЛА, ВЫЗВАННАЯ СИТУАЦИЕЙ НА ЕГО ВХОДЕ

Для того чтобы канал мог действовать в направлении от системы A к системе B, он должен быть “подключен” к системе A своим входом, а к системе B — выходом. Таким образом, множество возможных состояний системы A должно совпасть с входным алфавитом канала, а множество возможных состояний системы B — с его выходным алфавитом. Тогда (по терминологии теории информации) система A станет источником информации, а система B — ее приемником. Точнее говоря, источником и приемником информации станут *ситуации неопределенности* в системах соответственно A и B.

1. Простейшие каналы, описываемые на вероятностном языке. Прежде чем непосредственно переходить к рассмотрению индукции, необходимо сделать несколько замечаний. Мы и дальше в качестве ведущего примера используем *вероятностный* тип неопределенности. Однако теперь для этого недостаточно ограничиться рассмотрением таких ситуаций неопределенности, которые можно описать пространствами вероятностей. Теперь необходимо аналогичное ограничение наложить также на рассматриваемые информационные каналы. А это значит, что мы ограничимся рассмотрением таких каналов, которые можно описать с помощью *переходных распределений вероятностей*.

Как это ни смешно, но эта примитивная модель (которая, конечно, ни на что не годна в практическом отношении) дает возможность рассмотреть основные проблемы, связанные с преобразованиями информации в каналах. Заметим, что в работе [4] предложен способ формального описания очень широкого спектра информационных каналов.

Напомним также о выводе, к которому мы пришли в первой части статьи (разд. 2), — о том, что в теории ситуаций неопределенности недопустимо отождествлять описания ситуаций неопределенности с самими этими ситуациями (как это делается в теории вероятностей). А поскольку каждый информационный канал может быть представлен как семейство ситуаций неопределенности, этот вывод должен естественным образом распространяться и на каналы.

В связи с недопустимостью смешивания понятий *канал* и *описание канала* может возникать (и действительно возникает) проблема некоторой тяжеловесности языка. Для того чтобы эту проблему несколько смягчить, описания каналов иногда будем называть **формальными каналами**.

Итак, пусть X и Y — два *дискретных* множества и пусть s есть *переходное распределение* (ПР) с множества X на Y. Известно, что любое такое ПР можно представить как *семейство* вида

$$s = \left(s(\bullet \mid x) \mid x \in X\right) \tag{1}$$

(с множеством индексов X). Для каждого $x \in X$ элемент $s(\bullet \mid x)$ семейства s представляет собой *распределение вероятностей* на множестве Y (а для каждого $y \in Y$ число $s(y \mid x)$ обычно называют *условной вероятностью* элемента y относительно элемента x). Сказанное выше означает, что $s(\bullet \mid x)$ представляет собой *функцию*, определенную на Y. Подчеркнем также, что x есть *не аргумент* этой функции $s(\bullet \mid x)$, а *параметр*. Иначе говоря, если $x, z \in X$ и $x \neq z$, то $s(\bullet \mid x)$ и $s(\bullet \mid z)$ — *разные* (вообще говоря) функции.

Если информационный канал описывается переходным распределением s с множества X на множество Y, то эти множества будем называть соответственно **входным** и **выходным алфавитами** как упомянутого информационного канала, так и *формального канала* s.

2. Описание индукции. Выше было сказано, что преобразование *индукция* не связано с появлением или передачей каких-либо сигналов, а предполагает только само существование канала, *соединяющего* две системы. Для описания индукции достаточно иметь: 1) описания двух ситуаций неопределенности в системах A и B до образования информационной связи между ними и 2) описание соединяющего их канала. Оставаясь в рамках вероятностного типа неопределенности, мы предположим, что исходные ситуации в системах A и B описываются пространствами вероятностей соответственно (X, p) и (Y, q), а канал — переходным распределением s.

Заметим, что после образования информационной связи между системами A и B *совокупная* ситуация в этих системах будет полностью характеризоваться парой распределений (p, s) (т.е. она не будет зависеть от распределения q). Пары вида (p, s) являются частным случаем так называемых *сигнальных пар* (в работе [5] показано, что это новое понятие является очень емким; оно позволяет значительно усовершенствовать язык, используемый при обсуждении работы каналов). Поэтому и пару (p, s) здесь тоже будем называть (вероятностной) **сигнальной парой** (с входным и выходным алфавитами соответственно X и Y). РВ p и ПР s будем называть (формальным) **источником** и (формальным) **каналом** сигнальной пары (p, s).

Что же произойдет с системами A и B при образовании информационной связи? Ситуация в системе A не изменится (как и до образования связи она будет описываться пространством вероятностей (X, p)). А вот в системе B возникнет новая ситуация, которая не будет иметь никакого отношения к той, которая существовала до возникновения связи. Это преобразование и названо здесь **индукцией**.

Новую — ***индуцированную*** — ситуацию в системе B уже нужно будет описывать пространством вероятностей $(Y, p*s)$, где $p*s$ — распределение вероятностей на множестве Y следующего вида:

$$p*s = y \mapsto \sum_{x \in X} p(x) \cdot s(y \mid x) \lozenge Y . \tag{2}$$

Распределение РВ $p*s$ назовем **сверткой** сигнальной пары (p, s).

Подчеркнем, что распределение $p*s$ начинает *действовать* на множестве Y с момента образования информационной связи между системами.

3. Внутренние информационные меры индуцированной ситуации. Найдем выражения для количества собственной информации точек пространства $(Y, p*s)$ и его меры неопределенности. Аналогично выражению [1, (2)] количество собственной информации $I_{p*s}(y)$ каждой точки $y \in Y$ пространства вероятностей $(Y, p*s)$ имеет вид

$$I_{p*s}(y) = \log \frac{1}{p*s(y)} . \tag{3}$$

Мера неопределенности ситуации, описываемой пространством $(Y, p*s)$, характеризуется выражением

$$G(Y, p*s) = \sum_{y \in Y} p*s(y) \cdot \log \frac{1}{p*s(y)} , \tag{4}$$

аналогичным выражению [1, (5)].

17. ОБРАЗОВАНИЕ ИНФОРМАЦИОННОЙ ЗАВИСИМОСТИ МЕЖДУ СИСТЕМАМИ НА ВХОДЕ И НА ВЫХОДЕ КАНАЛА

Выше мы уже подчеркнули, что необходимо отличать ***образование*** информационной связи между системами A и B от ***функционирования*** этой связи. Вопросы *функционирования* информационной связи (возникшей в результате появления канала, соединяющего системы) будем рассматривать в двух следующих разделах. А здесь ограничимся только вопросами, связанными с *образованием* информационной связи.

1. Описание преобразования. В разделе 16 было рассмотрено преобразование *индукции*, которое исходную ситуацию неопределенности в системе B переводит в некоторую новую ситуацию (эту ситуацию канал s *индуцирует* в систему B из системы A). Как было уже отмечено выше, индукция представляет собой только внешний (наблюдаемый) эффект образования информационной связи. Здесь же мы сосредоточимся на рассмотрении внутреннего (наиболее важного) эффекта этого преобразования. Внутренний эффект имеет *информационный смысл*.

Как этот информационный смысл можно выразить формально? Для этого сначала нужно перейти к рассмотрению *совокупных* (для обеих систем A и B) ситуаций неопределенности и их формальных описаний. Под *исходной* совокупной ситуацией в системах A и B можно понимать *двумерную* ситуацию, первая и вторая проекции которой информационно независимы и описываются пространствами вероятностей соответственно (X, p)

и (Y, q). А под *результирующей* совокупной ситуацией в этих двух системах можно понимать (тоже *двумерную*) ситуацию, первая и вторая проекции которой описываются пространствами вероятностей соответственно (X, p) и $(Y, p*s)$.

В третьей части статьи показано, как построить произведение $(X\times Y, p\times q)$ пространств (X, p) и (Y, q) [3, разд. 14, п. 1]. Это построение опирается на соотношение [3, (18)]:

$$p\times q\,(x, y) = p(x)\cdot q(y), \tag{5}$$

справедливое для любых $x\in X$ и $y\in Y$. Легко понять, что произведение $(X\times Y, p\times q)$ и является описанием исходной совокупной ситуации в системах A и B, поскольку имеют место соотношения [3, (18) и (19)]. Покажем, как описать заключительную совокупную ситуацию.

2. Переход к совместному рассмотрению двух информационно связанных ситуаций. Итак, мы хотим найти способ описания *двумерной* ситуации неопределенности, первой и второй проекциями которой были бы соответственно ситуация в системе A (описываемая пространством вероятностей (X, p)) и индуцированная ситуация в системе B (описываемая пространством вероятностей $(Y, p*s)$). Для описания такой ситуации потребуется двумерное пространство вероятностей вида $G(X\times Y, p\square s)$, где РВ $p\square s$ определяется так:

$$p\square s = (x, y)\mapsto p(x)\cdot s(y\,|\,x)\,\Diamond\, X\times Y. \tag{6}$$

Это значит, что вероятность $p\square s(x, y)$ любой пары $(x, y)\in X\times Y$ должна удовлетворять равенству

$$p\square s(x, y) = p(x)\cdot s(y\,|\,x). \tag{7}$$

РВ $p\square s$ (на произведении двух множеств $X\times Y$), мы назовем **композицией** сигнальной пары (p, s).

Нетрудно убедиться в том, что пространство $G(X\times Y, p\square s)$ действительно описывает искомую совокупную ситуацию неопределенности в системах A и B. Для этого достаточно проверить выполнение двух равенств

$$\mathbf{pr_1}\, p\square s = p, \tag{8}$$

$$\mathbf{pr_2}\, p\square s = p*s. \tag{9}$$

3. Мера внутренней взаимной информации между входом и выходом канала. Пусть задана точка (x, y) пространства $G(X\times Y, p\square s)$. Из теории информации хорошо известна (с точностью до обозначений) мера

$$I_{p,s}\,(x\leftrightarrow y) = \log\frac{s(y\,|\,x)}{p*s(y)}, \tag{10}$$

которую здесь мы назовем **количеством внутренней взаимной информации** между состоянием $x\in X$ на входе формального канала s и состоянием $y\in Y$ на его выходе (при ситуации неопределенности на входе канала, описываемой пространством вероятностей (X, p)).

Заметим, что мера $I_{p,s}(x \leftrightarrow y)$ является частным случаем меры $I_{\pi}(x \leftrightarrow y)$ [3, (10)]. Действительно, для получения выражения (10) достаточно в [3, (10)] вместо π подставить $p \square s$ и использовать равенство (7):

$$I_{p \square s}(x \leftrightarrow y) = \log \frac{p \square s(x,y)}{p(x) \cdot p * s(y)} = I_{p,s}(x \leftrightarrow y). \qquad (11)$$

4. Количество внешней информации, необходимое для образования информационной связи. Итак, после данных выше разъяснений образование информационной связи представляется как преобразование (двумерной) ситуации неопределенности, описываемой пространством вероятностей $(X \times Y, p \times q)$ (с независимыми проекциями), в (двумерную) ситуацию, описываемую пространством $G(X \times Y, p \square s)$. Поэтому естественным образом возникает вопрос, сколько для этого преобразования необходимо внешней информации. Ответ можно получить, обращаясь к определению 3 и предположению 2 [2, разд. 10, п. 2]: искомое количество должно совпадать с количеством внешней информации в пользу гипотезы $p \square s$ против гипотезы $p \times q$. В соответствии с определением 3 эта величина должна быть обозначена $E(X \times Y, p \times q \,\|\, p \square s)$. А согласно [2, (40)] она должна удовлетворять соотношению

$$E(X \times Y, p \times q \,\|\, p \square s) = \sum_{(x,y) \in X \square Y} p \square s(x,y) \cdot \log \frac{p \square s(x,y)}{p \times q(x,y)}. \qquad (12)$$

Если воспользоваться соотношениями (5) и (7), то после простых преобразований получим выражение

$$E(X \times Y, p \times q \,\|\, p \square s) = \sum_{x \in X} p(x) \cdot \sum_{y \in Y} s(y \mid x) \cdot \log \frac{s(y \mid x)}{q(y)}. \qquad (13)$$

5. Связь с количеством внешней взаимной информации между входом и выходом канала. Можно показать, что каким бы ни было распределение q, будет выполняться неравенство

$$E(X \times Y, p \times q \,\|\, p \square s) \ge E(X \times Y, p \times (p * s) \,\|\, p \square s) =$$
$$= \sum_{x \in X} p(x) \cdot \sum_{y \in Y} s(y \mid x) \cdot \log \frac{s(y \mid x)}{p * s(y)}. \qquad (14)$$

Заметим, что правая часть выражения (14) представляет собой частный случай меры внешней взаимной информации $E(X \leftrightarrow Y \mid \pi)$ [3, (14)]. Действительно, подставив в [3, (14)] распределение $p \square s$ вместо π и использовав соотношения (8) и (9), получим

$$E(X \leftrightarrow Y \mid p \square s) = \sum_{(x,y) \in X \square Y} p \square s(x,y) \cdot \log \frac{p \square s(x,y)}{p(x) \cdot p * s(y)}. \qquad (15)$$

Сравнение выражений (14) и (15) дает равенство

$$E(X \leftrightarrow Y \mid p \square s) = E(X \times Y, p \times (p * s) \,\|\, p \square s) =$$
$$= \sum_{x \in X} p(x) \cdot \sum_{y \in Y} s(y \mid x) \cdot \log \frac{s(y \mid x)}{p * s(y)}. \qquad (16)$$

Итак, мы пришли к следующему выводу.

Количество внешней информации, необходимое для образования информационной связи, не может оказаться меньшим, чем количество внешней ***взаимной*** *информации между входом и выходом канала (при заданной ситуации неопределенности на входе канала).*

18. ВЫНУЖДЕНИЕ СИТУАЦИИ НЕОПРЕДЕЛЕННОСТИ НА ВЫХОДЕ КАНАЛА ИЗВЕСТНЫМ СОСТОЯНИЕМ ЕГО ВХОДА

В предыдущем разделе показано, чт***о*** происходит при *образовании* информационной связи между системами A и B благодаря наличию связывающего их канала. Теперь мы рассмотрим, как эта связь может *работать*. Заметим, что выделение этих двух вопросов для отдельного рассмотрения не является искусственным (как могло бы показаться), поскольку между ними имеется принципиальное отличие. Действительно, как видно из раздела 17, образование информационной связи между системами A и B сводится к тому, что ситуации неопределенности в обеих системах перестают быть независимыми. С другой стороны, *функционирование* информационной связи состоит в том, что становится известным состояние системы A или системы B. При этом, любое состояние x системы A может рассматриваться как ***сигнал***, который появился на входе канала, а затем будет передан по каналу. Аналогично, любое состояние y системы B может рассматриваться как сигнал, который появился на выходе канала и представляет собой результат *доставки* некоторого сигнала x к системе B.

Заметим, что степень отличия принимаемых сигналов от переданных может колебаться в широких пределах. К тому же, не существует простой связи между этими отличиями, с одной стороны, и степенью *потери информации* в канале, с другой. Так, канал может, например, менять способ *представления* информации практически без ее потери. С другой стороны, возможны значительные потери информации без изменения способа ее представления. Далее, время доставки сигнала от системы A к системе B может колебаться от очень маленького (как в обычных системах связи) до очень большого (вплоть до миллиардов лет — как при наблюдении далеких космических объектов). Смысл введения понятия *информационный канал* как раз и состоит в том, чтобы выделить существенные характеристики каналов, которые должны отражаться в их описаниях.

Подчеркнем, что каждый раз, когда в системе A или в системе B появляется новый сигнал, и, следовательно, эта система переходит в новое состояние, ситуация неопределенности в данной системе прекращает свое существование (*коллапсирует*). Вторая же ситуация при этом изменяется в зависимости от того, какой сигнал был порожден в системе A или принят в системе B.

Классическая теория информации была ориентирована на многократное и бесперебойное создание сигналов источником информации и их передачу каналом связи. Такая модель системы связи требовалась для доказательства асимптотических теорем. Здесь же у нас совершенно другие цели, и модель *многократного* порождения и передачи сигналов здесь пока не нужна. Поэтому мы ограничиваемся рассмотрением ***одного шага*** функционирования информационной связи.

В данной статье мы рассматриваем только первый этап функционирования информационной связи — когда становится известным состояние системы A (т.е. сигнал, порожденный на входе канала). Это событие приводит к возникновению на выходе канала (т.е. в системе B) новой ситуации неопределенности.

1. Описание вынуждения. Итак, тема этого раздела — последствия возникновения сигнала на *входе* информационного канала. Пусть стало известно, что система A оказалась в состоянии x. Тогда для наблюдателя, находящегося в этой же системе A (и знающего, что она оказалась в состоянии x), ситуация неопределенности в системе B должна измениться: должна возникнуть ситуация, которую необходимо описывать пространством вероятностей $(Y, s(\bullet \mid x))$, где РВ $s(\bullet \mid x)$ (действующее на множестве Y) имеет вид

$$s(\bullet \mid x) = y \mapsto s(y \mid x) \lozenge Y . \tag{17}$$

Будем говорить, что эта ситуация в системе B является **вынужденной** (т.е. представляет собой результат *вынуждения* состоянием x системы A).

Таким образом, преобразование *вынуждения* приводит к тому, что ситуация неопределенности, образовавшаяся на выходе канала в результате индукции и описываемая пространством вероятностей $(Y, p * s)$, превращается в вынужденную ситуацию, описываемую пространством $(Y, s(\bullet \mid x))$.

2. Внутренние информационные меры вынужденной ситуации. Количество собственной информации $I_{s(\bullet \mid x)}(y)$ каждой точки $y \in Y$ пространства вероятностей $(Y, s(\bullet \mid x))$ имеет вид

$$I_{s(\bullet \mid x)}(y) = \log \frac{1}{s(y \mid x)} , \tag{18}$$

а мера неопределенности порожденной ситуации, описываемой пространством $(Y, s(\bullet \mid x))$, характеризуется выражением

$$G(Y, s(\bullet \mid x)) = \sum_{y \in Y} s(y \mid x) \cdot \log \frac{1}{s(y \mid x)} . \tag{19}$$

3. Условная неопределенность. Напомним, что для описания рассмотренного в разделе 16 преобразования *индукции* ситуации в системе B использовалась *свертка* $p * s$ сигнальной пары (p, s), которая, как показывает выражение (2), представляла собой результат усреднения распределений $s(\bullet \mid x)$ по всем $x \in X$. Однако усреднение распределений $s(\bullet \mid x)$ — не единственный возможный вид усреднения. Можно, например, усреднить (тоже по всем $x \in X$) меры неопределенности $G(Y, s(\bullet \mid x))$. В результате получится новая *информационная мера*

$$G(Y, s \parallel X, p) = \sum_{x \in X} p(x) \cdot G(Y, s(\bullet \mid x)), \tag{20}$$

которую мы здесь назовем **условной неопределенностью** состояния системы B относительно состояния системы A.

Если в выражение (20) вместо $G(Y, s(\bullet \mid x))$ подставить правую часть выражения (19), то получится следующее:

$$G(Y,s \parallel X,p) = \sum_{x \in X} p(x) \cdot \sum_{y \in Y} s(y \mid x) \cdot \log \frac{1}{s(y \mid x)}. \qquad (21)$$

Легко показать, что какой бы ни была сигнальная пара (p,s), будет выполняться неравенство

$$G(Y, p*s) \geq G(Y,s \parallel X,p), \qquad (22)$$

где $G(Y, p*s)$ определено в (4).

Неравенство (22) явно имеет какой-то серьезный смысл. И этот смысл состоит в том, что имеет место равенство

$$E(X \leftrightarrow Y \mid p\square s) = G(Y, p*s) - G(Y,s \parallel X,p) \qquad (23)$$

(которое нетрудно проверить). Это равенство добавляет еще одну грань к тому, чт***о*** нам уже известно о мере внешней взаимной информации $E(X \leftrightarrow Y \mid p\square s)$.

19. ИНФОРМАЦИОННЫЙ АНАЛИЗ ВЫНУЖДЕНИЯ СИТУАЦИЙ НЕОПРЕДЕЛЕННОСТИ

В третьей части статьи было показано, что внешняя информация может превращаться во внутреннюю [3, разд. 12, пп. 2 и 3]. Теперь мы покажем, что и обратно, внутренняя информация может превращаться во внешнюю.

1. Откуда берется внешняя информация преобразования? Как следует из раздела 16, до тех пор, пока состояние системы A остается неизвестным, ситуация неопределенности в системе B представляет собой результат *индукции* из системы A и должна, следовательно, описываться пространством вероятностей $(Y, p*s)$ (где РВ $p*s$ определяется выражением (2)). Но после того, как стало известно, что система A находится в некотором состоянии $x \in X$, ситуация неопределенности в системе B должна измениться. Вместо исходной (индуцированной) ситуации возникнет новая ситуация, *вынужденная* состоянием x системы A. Эта новая ситуация уже должна описываться пространством вероятностей $(Y, s(\bullet \mid x))$.

Таким образом, мы можем теперь указать возможное *логическое основание* для изменения ситуации неопределенности в системе B. Этим основанием может быть *сообщение* о таком событии: «Система A оказалась в состоянии $x \in X$ ». Более кратко это событие будем обозначать $x \in X$ или даже просто x.

Сообщение об этом событии может, в частности, оказаться несовместимым (при существующей системе знаний) с некоторыми возможными состояниями системы B. И тогда преобразование исходной ситуации неопределенности в этой системе в новую ситуацию будет содержать в себе также черты преобразования *ограничение разнообразия*, которое рассматривалось в [2, разд. 9].

Итак, исходная ситуация неопределенности в системе B, которая может быть описана пространством вероятностей $(Y, p*s)$, преобразуется в новую ситуацию, описываемую пространством вероятностей $(Y, s(\bullet \mid x))$. Так как это преобразование происходит вследствие сообщения о событии $x \in X$, мы приходим к выводу, что информация, необходимая для данного преобразования, создается этим событием.

2. Количество информации, необходимое для преобразования «вынуждение». По аналогии с тем, что было сказано в [2, разд. 10, п. 2], можем узнать, какое количество внешней информации нужно было получить для того, чтобы *исходная ситуация*, для описания которой использовалось пространство вероятностей $(Y, p*s)$, превратилась в *заключительную ситуацию* (которая описывается пространством $(Y, s(\bullet\,|\,x))$. Ответ на этот вопрос дает предположение 2, согласно которому требуемое количество внешней информации должно совпадать с количеством внешней информации $E(Y, p*s \,\|\, s(\bullet\,|\,x))$ в пользу гипотезы $s(\bullet\,|\,x)$ против гипотезы $p*s$ [2, разд. 10, п. 2, определение 3]. Последнее же ввиду [2, (39)] должно характеризоваться выражением

$$E(Y, p*s \,\|\, s(\bullet\,|\,x)) = G(Y, s(\bullet\,|\,x) \lozenge p*s) - G(Y, s(\bullet\,|\,x)), \qquad (24)$$

где $G(Y, s(\bullet\,|\,x) \lozenge p*s)$ — мера неопределенности *ситуации заблуждения*, описываемой пространством двойных вероятностей $(Y, s(\bullet\,|\,x) \lozenge p*s)$.

Мы уже знаем, что мера неопределенности такой ситуации заблуждения должна измеряться следующим выражением, аналогичным выражению [2, (37)]:

$$G(Y, s(\bullet\,|\,x) \lozenge p*s) = \mathbf{E}\, I_{p*s} = \sum_{y\in Y} s(y\,|\,x) \cdot I_{p*s}(y) =$$

$$= \sum_{y\in Y} s(y\,|\,x) \cdot \log \frac{1}{p*s(y)}. \qquad (25)$$

Поэтому ввиду (19) окончательно получим

$$E(Y, p*s \,\|\, s(\bullet\,|\,x)) = \sum_{y\in Y} s(y\,|\,x) \cdot \log \frac{s(y\,|\,x)}{p*s(y)}. \qquad (26)$$

3. Количественное сопоставление двух ролей информации. Теперь мы можем еще раз сопоставить две роли информации: *внутреннюю информацию*, связанную с наступлением события $x \in X$ в системе A (а именно — собственную информацию этого события), и *внешнюю информацию* преобразования ситуации неопределенности в системе B.

Согласно выражению [1, (2)] количество

$$I_p(x) = \log \frac{1}{p(x)} \qquad (27)$$

собственной информации события x зависит от его вероятности $p(x)$. Можно предположить, что от этой же вероятности $p(x)$ должно зависеть и количество $E(Y, p*s \,\|\, s(\bullet\,|\,x))$ внешней информации преобразования ситуации в системе B, вызванного событием x.

Сравнение выражений (26) и (27) показывает, что прямого подтверждения нашего предположения мы не имеем. Действительно, в выражении (26) не видно даже намека на зависимость величины $E(Y, p*s \,\|\, s(\bullet\,|\,x))$ от вероятности $p(x)$. Поэтому мы проиллюстрируем на простом примере параллельное поведение двух упомянутых величин — количества $I_p(x)$ собственной информации события x и числа $E(Y, p*s \,\|\, s(\bullet\,|\,x))$ — в зависимости от вероятности $p(x)$. Пусть каждое из множеств X и Y имеет по два элемента: $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$. Тогда свойства формального источника p полностью описываются одним числом $p(x_1)$, а свойства формального канала s — двумя числами: $s(y_1 \mid x_2)$ и $s(y_2 \mid x_1)$. Кроме того, примем еще допущение, что канал s является симметричным, т.е. выполняется условие $s(y_1 \mid x_2) = s(y_2 \mid x_1)$.

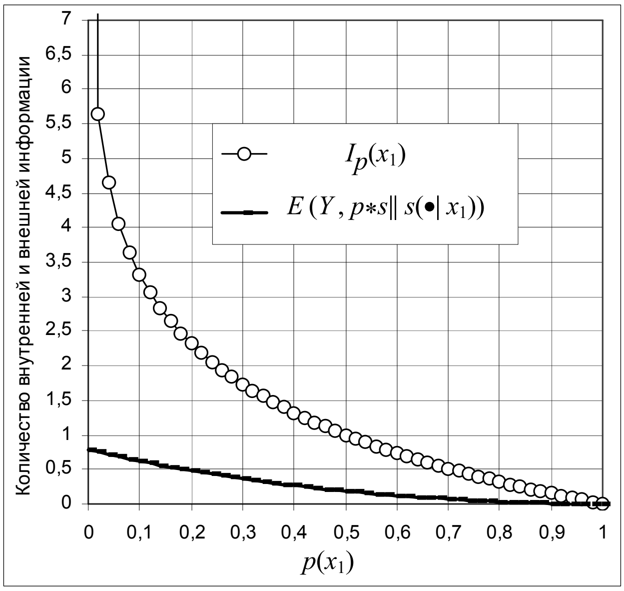

Рис. 5. Зависимость от числа $p(x_1)$ количества собственной информации $I_p(x_1)$ точки x_1 пространства (X, p) и количества внешней информации $E(Y, p{*}s \| s(\bullet \mid x_1))$ при условии, что s есть двоичный симметричный канал, характеризуемый равенствами $s(y_2 \mid x_1) = s(y_1 \mid x_2) = 0{,}25$.

На рис. 5 показано поведение величин $I_p(x_1)$ и $E(Y, p * s \| s(\bullet \mid x_1))$ в зависимости от числа $p(x_1)$ при условии, что канал s характеризуется соотношением $s(y_1 \mid x_2) = = s(y_2 \mid x_1) = 0{,}25$. Из рисунка видно, что поведение обеих изучаемых величин в каком-то смысле является согласованным: обе они убывают с увеличением числа $p(x_1)$ и обе достигают значения 0 при условии $p(x_1) = 1$. И при этом все время остается верным неравенство

$$I_p(x_1) \ge E(Y, p * s \| s(\bullet \mid x_1)). \tag{28}$$

4. Зависимость от параметров канала. Похоже, что на основании рис. 5 ничего больше утверждать нельзя. Но зато можно спросить: *что означает* и *чем объясняется* неравенство (28)? Не может ли быть так, что между двумя видами информации, фигурирующей в левой и правой части неравенства, существует тесная связь? Или даже так, что слева и справа в соотношении (28) речь идет об ***одной и той же информации*** — той, которая создается в результате коллапса ситуации неопределенности в системе A, а затем с помощью канала передается в систему B? Но тогда возникает вопрос, почему вместо неравенства (28) мы не получили равенства? Это могло бы объясняться тем, что не вся информация (в количестве $I_p(x_1)$), создаваемая на входе сигнальной пары (p, s), доходит до ее выхода. И происходит это из-за того, что в информационном канале имеется ***шум*** (в примере, рассмотренном на рис. 5, этот шум довольно значителен, поскольку вероятность ошибки при передаче принята равной 0,25).

Если это предположение является верным, то интересно посмотреть, как ведут себя величины $I_p(x_1)$ и $E(Y, p * s \| s(\bullet \mid x_1))$ в зависимости не от числа $p(x_1)$, характеризующего источник p сигнальной пары (p, s), а от числа $s(y_1 \mid x_2) = s(y_2 \mid x_1)$, характеризующего количество шума в (симметричном) формальном канале s. Такая иллюстрация представлена на рис. 6, где показаны упомянутые зависимости для трех вариантов значений $p(x_1)$: $p(x_1) = 0{,}3$; 0,5 и 0,8. (Поскольку величина $I_p(x_1)$ не зависит от числа $s(y_1 \mid x_2) = s(y_2 \mid x_1)$, ее поведение изображается горизонтальными прямыми.)

И здесь мы получили полное подтверждение выдвинутого предположения. Оказывается, что в тех случаях, когда в информационном канале полностью исчезает шум (а это происходит, когда вероятность ошибки ($s(y_1 \mid x_2) = s(y_2 \mid x_1)$ равняется нулю или единице), имеет место равенство

$$I_p(x_1) = E(Y, p * s \,\|\, s(\bullet \mid x_1)). \qquad (29)$$

(Необходимо заметить, что равенство (29) является *асимптотическим*, так как непосредственно вычислить значения величины $E(Y, p * s \,\|\, s(\bullet \mid x_1))$ при *строгом* отсутствии шума в канале оказывается невозможным ввиду необходимости деления на нуль.)

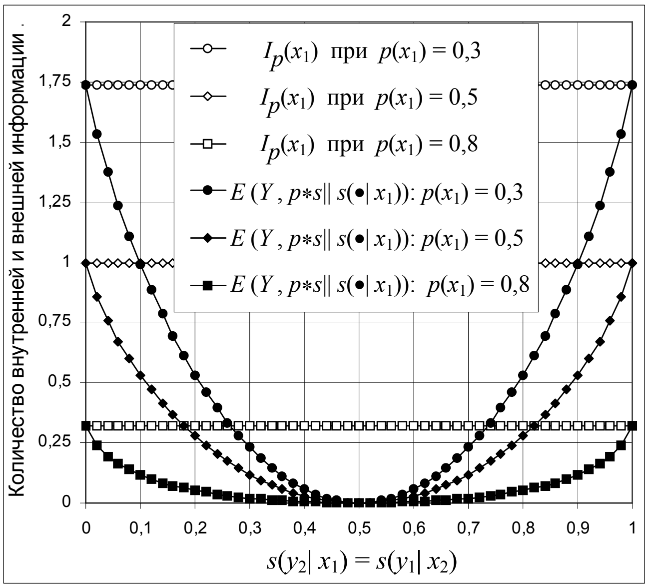

Рис. 6. Зависимость количества внешней информации $E(Y, p{*}s \| s(\bullet| x_1))$ от параметра $s(y_2| x_1) = s(y_1| x_2)$ симметричного формального канала s для сигнальной пары (p, s) при условиях $p(x_1) = 0{,}3$; 0,5 и 0,8. Для сравнения приведены значения количества $I_p(x_1)$ собственной информации элемента x_1 при тех же условиях.

В связи с таким поворотом представляется интересным опять посмотреть на зависимости величин $I_p(x_1)$ и $E(Y, p * s \,\|\, s(\bullet \mid x_1))$ от числа $p(x_1)$, но уже при различных значениях вероятности ошибки в формальном канале s. Такая иллюстрация представлена на рис. 7.

Здесь тоже рассматривается симметричный канал. Поэтому вероятность ошибки, которая здесь обозначена a, имеет вид $a = s(y_1 \mid x_2) = \; = s(y_2 \mid x_1)$. При вероятности ошибки $a = 0{,}5$ количество внешней информации $E(Y, p * s \,\|\, s(\bullet \mid x_1))$ равно нулю для всех значений $p(x_1)$. Это объясняется тем, что при $a = 0{,}5$ информационный канал вообще неспособен передавать информацию. Видно также, что с уменьшением вероятности ошибки кривая, изображающая поведение количества внешней информации $E(Y, p * s \,\|\, s(\bullet \mid x_1))$, все больше приближается к кривой для количества собственной информации $I_p(x_1)$. В случае же, если вероятность ошибки близка к нулю, обе кривые практически совпадают. Так, из таблицы значений, по которой построен график рис. 7, видно, что отличие между значениями $E(Y, p*s \,\|\, s(\bullet \,|\, x_1))$ и $I_p(x_1)$ при $p(x_1) = 10^{-10}$ улавливается не раньше, чем в восьмом знаке после запятой (напомним, что идеально точное совпадение невозможно, так как для $a = 0$ величину $E(Y, p*s \,\|\, s(\bullet \,|\, x_1))$ вычислить нельзя). Заметим также, что, как следует из рис. 6, аналогичная картина наблюдалась бы и при *увеличении* вероятности ошибки a от значения 0,5 до единицы.

5. Две гипотезы. Несмотря на то, что три последние иллюстрации касаются простейшего варианта сигнальной пары (p, s) (когда каждое из множеств X и Y имеет по два элемента), можно предположить, что обнаруженная на них связь между количеством собственной информации $I_p(x)$ элемента x и количеством внешней информации $E(Y, p*s \,\|\, s(\bullet \,|\, x))$ имеет общий характер. Поэтому, вероятно, можно сформулировать и доказать две теоремы о существовании такой связи для произвольных вероятностных сигнальных пар. Но в данной статье нет места для доказательств подобных теорем. Поэтому мы здесь ограничимся только формулировками, а сами теоремы будем пока называть *гипотезами*.

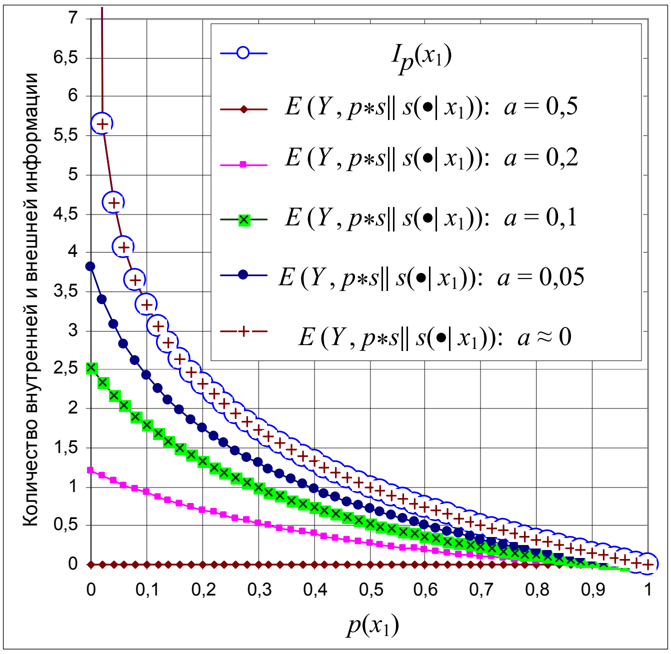

Рис. 7. Зависимость количества внешней информации $E(Y, p*s\| s(\bullet| x_1))$ от вероятности $p(x_1)$ для сигнальной пары (p, s) с симметричным формальным каналом s при условиях $s(y_2| x_1) = s(y_1| x_2) = a$ и $a = 0,5$; 0,2; 0,1; 0,05 и $a \approx 0$. Для сравнения приведена зависимость количества $I_p(x_1)$ собственной информации элемента x_1 от того же числа $p(x_1)$.

Гипотеза 1. Для всякой вероятностной сигнальной пары (p, s) с входным алфавитом X и для любого $x \in X$ имеет место неравенство

$$I_p(x) \geq E(Y, p*s \,\|\, s(\bullet \,|\, x)) \,. \blacksquare \tag{30}$$

Вторая гипотеза является асимптотической, и для ее формулировки необходимо сначала дать определение понятия *информационного канала без шума*.

Под *шумом* здесь понимается одна из разновидностей помех, такая, которая снижает способность канала передавать информацию. Шум противопоставляется некоторым другим видам помех — например, разного рода искажениям, — которые, хотя тоже мешают передаче информации и создают различные проблемы, но не уменьшают *количества* передаваемой информации.

Определение 4. Будем говорить, что в информационном канале **отсутствует шум**, если любое состояние на выходе канала позволяет однозначно восстановить то состояние на его входе, которое было причиной данного состояния на выходе. В этом случае будем говорить также, что шум отсутствует и в *формальном канале*, с помощью которого описывается данный информационный канал. ■

Нумерация определений в статье сквозная. Первое определение появилось в первой части статьи, а второе и третье — во второй части.

Простейший пример канала без шума можно построить при условии, когда входной и выходной алфавиты канала равномощны. Поскольку между такими алфавитами всегда можно установить взаимно однозначное соответствие, эти алфавиты можно просто отождествить. Поэтому предположим, что множество X является как входным, так и выходным алфавитом рассматриваемого канала. И рассмотрим такое переходное распределение $i = \left(i(\bullet \mid x) \mid x \in X\right)$ с множества X на себя, что для каждого $x \in X$ распределение вероятностей $i(\bullet \mid x)$ на множестве X *тривиально* и характеризуется условием $i(x \mid x) = 1.$ Из этого условия следует, что для любых двух различных элементов $x, y \in X$ должно выполняться равенство $i(y \mid x) = i(x \mid y) = 0.$ Вот такое переходное распределение и является описанием простейшего канала без шума.

Но понятие (вероятностного) канала без шума ни коим образом не ограничивается тем случаем, когда его входной и выходной алфавиты равномощны. Опираясь на определение 4, можно выяснить особенности конструкции любого канала без шума.

Здесь пока имеются в виду исключительно каналы вероятностного типа. Более общий подход к понятию *информационного канала без шума* предложен в статье [4].

Пусть вероятностный информационный канал с входным и выходным алфавитами соответственно X и Y описывается переходным распределением s с множества X на Y. Для каждого $x \in X$ пусть $s[x]$ есть множество всех таких $y \in Y$, что $s(y \mid x) \neq 0$. Формально это выглядит так:

$$s[x] = \left\{y \in Y : s(y \mid x) \neq 0\right\}. \tag{31}$$

И пусть переходное распределение s удовлетворяет следующему условию: для любых двух *различных* элементов $x, x' \in X$ множества $s[x]$ и $s[x']$ не пересекаются.

Можно показать, что это условие является необходимым и достаточным для того, чтобы в канале *отсутствовал шум*. Одно из следствий этого условия состоит в том, что у канала, в котором отсутствует шум, мощность выходного алфавита не может быть меньше мощности входного (в противном случае потери информации в канале будут неизбежными).

Будем считать, что у читателя уже сложилось неформальное представление о том, чт***о*** такое (вероятностный) канал без шума. Так что мы уже можем перейти к формулировке нашей гипотезы.

Пусть задано пространство вероятностей (X, p). И пусть задана некоторая (бесконечная) последовательность

$$\sigma = (s_1, s_2, s_3, ...) \tag{32}$$

переходных распределений с множества X на множество Y (т.е. последовательность вероятностных формальных каналов с входным и выходным алфавитами X и Y). Тогда для каждого $x \in X$ получим следующую последовательность мер внешней информации:

$$\Theta_x = \big(E(Y, p * s_i \,\|\, s_i(\bullet \mid x)) : i = 1, 2, 3, \ldots\big) . \qquad (33)$$

Гипотеза 2 формулируется так:

Гипотеза 2. Пусть заданы: пространство вероятностей (X, p) и последовательность σ (32) (вероятностных) *формальных* каналов s_i. Тогда, если последовательность σ сходится к некоторому формальному каналу без шума, то для каждого $x \in X$ последовательность Θ_x внешних мер информации $E(Y, p * s_i \,\|\, s_i(\bullet \mid x))$ сходится к внутренней мере $I_p(x)$ — количеству собственной информации элемента x. ■

Эта гипотеза имеет интересное следствие. Сравнение выражений (16) и (26) позволяет сделать вывод, что меры внешней информации $E(Y, p * * s \,\|\, s(\bullet \mid x))$ для всех $x \in X$ тесно связаны с еще одной внешней мерой — количеством внешней взаимной информации $E(X \leftrightarrow Y \mid p, s)$ между входом и выходом сигнальной пары (p, s). Эта связь выражается соотношением

$$E(X \leftrightarrow Y \mid p, s) = \sum_{x \in X} p(x) \cdot E(Y, p * s \,\|\, s(\bullet \mid x)) . \qquad (34)$$

Для того чтобы понять, чт***о*** это значит, рассмотрим информационную функцию

$$E(Y, p * s \,\|\, s(\bullet \mid \bullet)) = x \mapsto E(Y, p * s \,\|\, s(\bullet \mid x)) \ \Diamond\ X . \qquad (35)$$

Очевидно, что функция $E(Y, p * s \,\|\, s(\bullet \mid \bullet))$ является случайной величиной (относительно действующего на X распределения p). А ввиду (34) количество внешней взаимной информации $E(X \leftrightarrow Y \mid p, s)$ представляет собой математическое ожидание этой случайной величины, т.е.

$$E(X \leftrightarrow Y \mid p, s) = \mathbf{E}\, E(Y, p * s \,\|\, s(\bullet \mid \bullet)) . \qquad (36)$$

С другой стороны, как видно из [1, (5)], математическим ожиданием информационной функции I_p является мера неопределенности $G(X, p)$ пространства вероятностей (X, p). Таким образом, если для каждого $x \in X$ последовательность Θ_x мер внешней информации $E(Y, p * s_i \,\|\, s_i(\bullet \mid x))$ сходится к количеству $I_p(x)$ собственной информации, то при тех же условиях последовательность

$$\Omega = \big(E(X \leftrightarrow Y \mid p, s_i) \mid i = 1, 2, 3, \ldots\big) \qquad (37)$$

должна сходиться к мере неопределенности $G(X, p)$. Вот в этом и состоит следствие гипотезы 2:

Следствие. Пусть заданы: пространство вероятностей (X, p) и последовательность σ (вероятностных) формальных каналов s_i. Тогда, если последовательность σ сходится к некоторому каналу без шума, то последовательность Ω мер внешней взаимной информации $E(X \leftrightarrow Y \mid p, s_i)$ ($i = 1, 2, 3, \ldots$) сходится к мере неопределенности $G(X, p)$. ■

* * * * *

Итак, мы обнаружили связи между всеми мерами информации: между *внутренними* и *внешними* мерами; между мерами, родословная которых восходит к работам Клода Шеннона, и теми, которые происходят от идей Соломона Кульбака. Обнаружение таких связей является крайне важным особенно в свете того, что нам предстоит искать *расширения* всех этих мер на другие типы неопределенности, отличные от вероятностного типа, и даже такие, которые фактически не имеют с вероятностным типом ничего общего. Проблема поиска таких расширений является очень непростой, и знание упомянутых связей может очень помочь при попытках ее решения.

20. ВЫВОДЫ ИЗ СТАТЬИ

- Распространенные представления об *информации* нуждаются в радикальном пересмотре. Неверными являются как представления о том, чт***о*** такое информация, так и о том, какова роль информации в Природе.
- Информация не является внутренним свойством текстов (Эшби), но она является внутренним свойством ***ситуаций неопределенности*** (это касается ситуаций ***всех типов неопределенности***, а не одного только вероятностного типа). Отсюда следует, что нельзя изучать информацию, не рассматривая какую-либо ситуацию неопределенности.
- Ситуации неопределенности — часть нашей реальности. Поэтому проблема их изучения, так же как и проблема изучения их внутреннего свойства — информации — является не математической, а ***естественнонаучной проблемой***.
- Распространенные представления о *преобразованиях информации* тоже нуждаются в радикальном пересмотре. Преобразования информации — это ***преобразования ситуаций неопределенности***. К преобразованиям текстов они не имеют никакого отношения.
- Введение двух разновидностей информации — ***внутренней*** и ***внешней*** (по отношению к ситуациям неопределенности) позволило поставить вопрос о разработке мер внешней информации для измерения ***интенсивности преобразований*** ситуаций неопределенности. Возможность построения таких мер продемонстрирована на ряде примеров.
- Внешняя информация может превращаться во внутреннюю, а внутренняя — во внешнюю. Эти превращения происходят с сохранением *количества информации*. Это доказывает полную ***согласованность*** предложенной здесь ***системы мер внутренней и внешней информации***. Но это также позволяет сделать еще более важный вывод, что ***существует одно понятие информации*** (а не два, как могло бы показаться).
- Преобразование ***ограничение разнообразия*** занимает особое место среди остальных преобразований. Показано, что любое событие в Природе, сообщение о котором влечет ограничение разнообразия в ситуации неопределенности, связанной с состояниями некоторой реальной системы, *содержит информацию об этой системе*. Это значит, что существует ***информационная связь*** между такого рода событиями и состояниями системы.

• Информационные связи могут стать самостоятельным инструментом познания. Они принципиально отличаются от единственного известного сейчас инструмента познания — причинно-следственных связей — двумя особенностями: они *нематериальны* и *двусторонни*. Эти отличия ведут к тому, что информационные явления ***не могут подчиняться принципу причинности***.

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. — № 4 — С. 94–110.
3. *Дидук Н.Н.* Меры внутренней и внешней информации (на примере вероятностных ситуаций неопределенности). Часть III // Системні дослідження та інформаційні технології. — 2013. — № 2 — С. 127–142.
4. *Дидук Н.Н.* Информационные каналы как развитие представлений о каналах связи // Системні дослідження та інформаційні технології. — 2007. — № 1. — С. 129–141.
5. *Дидук Н.Н.* Сигнальные пары и их применение // Системні дослідження та інформаційні технології. — 2008. — № 2 — С. 128–143.

Поступила 12.06.2012

Статья напечатана под редакцией автора