РОССИЙСКАЯ АКАДЕМИЯ НАУК
ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
ИМЕНИ М.В. КЕЛДЫША

С.В. Попов

ЛОГИЧЕСКАЯ ЭНТРОПИЯ

Москва, 2005г.

УДК УДК 004.62

Попов С.В. Логическая энтропия. Препринт Института прикладной математики им. М.В. Келдыша РАН. Москва, 2005г.

Вводится понятие логической энтропии как меры неопределенности и сложности информационных моделей, описываемых булевскими функциями. Приводятся содержательные интерпретации логической энтропии. Демонстрируется отличие логической энтропии от энтропии Теории информации. Выделяется класс физических систем, энтропия которых совпадает с энтропией их информационных моделей.

Popov S.V. Logic entropy. Preprint of the Keldysh Institute of Applied Mathematics of RAS. Moscow, 2005.

It is entered notion of logical entropy as measures to uncertainties and difficulties of the information models, described logical functions. Profound interpretation to logical entropy is given. It is demonstrated difference between the logical entropy and entropy of Theory of information. Stands out the class of the physical systems, entropy which comply with entropy their information models.

De non apparetibus et non existentibus eadem est ratio.
О том, чего не видно и о том,
чего не существует, судят одинаково. (лат.)

Введение. Содержательное обоснование логической энтропии, как меры неопределенности информационных моделей дано в [1], где физические системы представляются логическими функциями, аргументы которых используются для кодирования объектов. В результате объекты в информационных моделях кодируются единичными означиваниями логических функций. И если различные означивания приводят к логически эквивалентным функциям, то они определяют одну характеристическую функцию некоторой совокупности объектов.

Более формально это выглядит так.

Пусть булевская функция $f(x_1, x_2, \ldots, x_n)$ служит информационной моделью физической системы. Ее объекты кодируются наборами вида $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ признаков, $\sigma_i \in \{0, 1\}$, $i = 1, 2, \ldots, n$, при котором функция f истинна. Тем самым, число различных объектов не превосходит мощности множества

$$\{(\sigma_1, \sigma_2, \ldots, \sigma_n) | \, f(\sigma_1, \sigma_2, \ldots, \sigma_n) = 1\}$$

единичных означиваний функции f.

В [1] введено понятие неопределенности физической системы $\varphi(x_1, x_2, \ldots, x_n)$, определяемой переменными $x_1, x_2, \ldots, x_i$, $i \le n$. Аналогично определяется неопределенность ее информационной модели $f(x_1, x_2, \ldots, x_n)$, определяемой переменными $x_1, x_2, \ldots, x_i$, $i \le n$. Как будет показано неопределенность функции $f(x_1, x_2, \ldots, x_n)$, определяемая переменными $x_1, x_2, \ldots, x_i$, обладает следующими свойствами.

1. Ее значение зависит от числа k подфункций, которые получаются в результате разложения функции по переменным $x_1, x_2, \ldots, x_i$. При увеличении k (при прочих равных условиях) неопределенность возрастает.

2. Неопределенность определяется относительными долями мощностей соответствующих множеств $\sum_1$, $\sum_2$, .., $\sum_k$ означиваний переменных x_1, x_2, …, x_i, приводящих к эквивалентным функциям. Если все множества $\sum_1$, $\sum_2$, .., $\sum_k$ одинаковы и доля каждого равна $1/k$, то неопределенность монотонно возрастает с ростом k.

3. При переходе от означивания переменных x_1, x_2, …, x_i, к означиванию переменных x_1, x_2, …, x_i, x_{i+1} их неопределенность зависит от неопределенности, определяемой переменными x_1, x_2, …, x_i.

Последнее свойство легко объяснимо с содержательной точки зрения, если заметить, что между аргументами функции f могут существовать зависимости, когда значения одних переменных в той или иной степени определяют значения других. Например, в случае функциональной зависимости одному значению некоторых переменных соответствует в точности одно значение других. И тогда порядок означивания существенно определяет последовательность разбиений объектов на классы. В результате, свойство аддитивности неопределенности нарушается.

1. Неопределенность булевских функций. Определим понятие энтропии логической функции и установим ее связь со структурой функций. Формальное определение логической энтропии мы введем, основываясь на следующем.

Пусть $f(\boldsymbol{Y}, \boldsymbol{w})$ есть булевская функция, множество $\boldsymbol{Y}$ ее аргументов назовем входными переменными и $\boldsymbol{w}$ – выходными. При определенных значениях $\boldsymbol{\sigma_Y}$ ее входных переменных и $\boldsymbol{\sigma_w}$ - выходных $f(\boldsymbol{\sigma_Y}, \boldsymbol{\sigma_w}) = 1$. Тем самым, функция f определяет отображение, ставящее в соответствие входным значениям - выходные. В общем случае, один входной кортеж определяет несколько выходных.

Напомним, что логическая функция $f(\boldsymbol{Y}, \boldsymbol{w})$ *представляет* вычислимую двоичную функцию $\varphi(\boldsymbol{Y})$ если: $f(\boldsymbol{Y}, \boldsymbol{w}) = 1 \Leftrightarrow \varphi(\boldsymbol{Y}) = \boldsymbol{w}$.

Назовем кортеж $\boldsymbol{\sigma}_{\boldsymbol{y}}$ означиваний (не обязательно всех) входных переменных $\boldsymbol{y}$ *$\boldsymbol{y}$-набором*. Тем самым, для логической функции $f(\boldsymbol{Y}, \boldsymbol{w})$, представляющей вычислимую функцию, верно, что всякий ее $\boldsymbol{y}$-набор определяет некоторое множество $\boldsymbol{w}$-наборов.

Нам потребуется следующее определение.

Бинарной программой π_f для логической функции $f(x_1, x_2, \ldots, x_n)$ называется ор-дерево с единственным корнем (*истоком*), которому приписана функция $f(x_1, x_2, \ldots, x_n)$, и двумя висячими узлами, одному из которых приписано значение 1 (1-*сток*), другому 0 (0-*сток*), а дуги помечены литерами x_i или $\bar{x}_i$, $i = 1, 2, \ldots, n$. Если два узла сети соединяются путем, дуги которого не содержат ортогональных меток, то назовем его *проводящим*. Каждый узел дерева достижим из истока. Проводящий путь, дуги которого помечены метками, образующими множество $\{x_{j_1}^{\sigma_1}, x_{j_2}^{\sigma_2}, \ldots, x_{j_m}^{\sigma_m}\}$ литер, где $\sigma_1, \sigma_2, \ldots, \sigma_m \in \{0, 1\}$, назовем *определяющим* это множество.

Внутренние узлы и дуги программы помечены следующим образом.

1. Каждой дуге приписана в точности одна литера x_i или $\bar{x}_i$, $i = 1,2, \ldots, n$.

2. Из каждого внутреннего узла ведут в точности две дуги, которым приписаны литеры x_i и $\bar{x}_i$, $i = 1,2, \ldots, n$.

3. Если в узел N ведет проводящий путь, дуги которого помечены метками $x_{j_1}^{\sigma_1}, x_{j_2}^{\sigma_2}, \ldots, x_{j_m}^{\sigma_m}$, то этому узлу приписана функция, которая получается из исходной в результате присваивания переменным $x_{j_1}, x_{j_2}, \ldots, x_{j_m}$ значений соответственно $\sigma_1, \sigma_2, \ldots, \sigma_m$. Узлы, соответствующие эквивалентным функциям, склеиваются. При этом узлу приписан в точности один представитель класса эквивалентности.

4. Для каждого из 2^n двоичных наборов $\boldsymbol{\sigma_x}$ переменных имеется путь проводящий из истока либо в 1-сток либо в 0-сток, множество меток дуг которого покрывается компонентами $\boldsymbol{\sigma_x}$.

Так как логическая функция однозначно характеризуется перечислением своих 1-проводящих путей, то будем мыслить бинарные программы, состоящими лишь из 1-проводящих путей. Тогда бинарная программа представляет собой двухполюсник с истоком, которому приписана логическая функция *f*, и стоком, который помечен 1.

Пусть функция $f(\boldsymbol{Y}, \boldsymbol{w})$ представляет вычислимую двоичную функцию φ. В бинарной программе π_f каждое вычисление $\varphi(\boldsymbol{\sigma_Y}) = \boldsymbol{\sigma_w}$ представлено единственным путем из истока в сток. Ограничимся рассмотрением программ, для которых все пути из истока на начальных отрезках содержат дуги, помеченные литерами, образованными только из входных переменных $\boldsymbol{y}$, и после этих отрезков дуг с такими метками не встречается. Такие бинарные программы назовем *однородными*.

Введем еще одно определение.

***y**-сечением* однородной бинарной программы назовем множество всех ее узлов, которыми завершаются все пути, начинающиеся в истоке и дуги которых помечены литерами, образованными лишь из переменных некоторого подмножества $\boldsymbol{y}$ входных переменных.

Применительно к вычислениям, реализуемым бинарными программами, можно говорить о неопределенности выходных значений от входных. Действительно, если задан $\boldsymbol{y}$-набор $\boldsymbol{\sigma_y}$, то он задает единственный путь программы π_f из истока в некоторый узел *N*. Нам удобно мыслить узел *N* как некоторый *промежуточный вычислитель*. Он может соединяться со стоком несколькими путями, каждый из которых характеризуется собственными выходными значениями. И только по входным значениям нельзя определить, какой именно выходной вектор появится на выходе такого вычислителя. В связи с этим можно говорить лишь о различении

множеств выходных векторов, которые определяются различными узлами ***y***-сечения.

Мы рассмотрим несколько различных подходов к вычислению неопределенности и дадим их содержательные интерпретации. Это прояснит связь настоящего подхода с традиционным в Теории информации.

В последующем, главным образом мы будем рассматривать логические функции, представляющие всюду определенные вычислимые функции. Во избежание путаницы, мы всегда будем указывать, какие логические функции рассматриваем, если это не вытекает из контекста. Очевидно, что *если логическая функция* $f(\boldsymbol{Y}, \boldsymbol{w})$ *представляет всюду определенную вычислимую функцию* $\varphi(\boldsymbol{Y})$, *то при любом* ***y****-наборе* $\boldsymbol{\sigma}_y$ *функция* $f(\boldsymbol{\sigma}_y, \boldsymbol{w})$ *не является тождественным нулем*. В этом случае $\varphi(\boldsymbol{\sigma}_y) = \{\boldsymbol{\sigma}_w: f(\boldsymbol{\sigma}_y, \boldsymbol{\sigma}_w) = 1$ при любых подстановках на места не означенных переменныъ$\}$.

Если полагать, что входные переменные независимы и их значения 0 и 1 одинаково возможны, то можно говорить о неопределенности зависимости выходных значений от входных, как о характеристике ***y***-сечения. Действительно, чем меньше ***y***-сечение, тем больше определенность, какой промежуточный вычислитель вычислеят выходные значения. В предельном случае, когда сечение содержит в точности один узел, при любом входном векторе все вычисления осуществляются одним вычислителем. Если сечение имеет $2^{|y|}$ узлов, то неопределенность максимальна, так как имеется возможность выбора из наибольшего числа доступных вычислителей.

Выберем в качестве показателя *неопределенности* выходных значений в зависимости от входных ***y***, двоичный логарифм от числа узлов ***y***-сечения программы π_f. Полагаем, что все входные переменные ***y*** независимы, ***y***-сечение состоит из k узлов, каждый его узел соединен с истоком одним и тем же числом $\bar{t}$ путей и число всех путей программы из истока в ***y***-сечение равно t.

Справедливо равенство $\log k = \log (t / \bar{t})$. Доля путей, ведущих в один узел ***y***-сечения среди всех путей программы π_f из истока в ***y***-сечение равна $p_y = \bar{t}/t$. Тогда $\log k = - \log p_y$. Если предположить, что доли $p_y(i)$ путей, ведущих в i-ый узел ***y***-сечения программы π_f различны, наложив требование, чтобы выполнялось равенство $\sum_{i=1,k} p_y(i) = 1$, то получим более общую формулу $- \sum_{i=1,k} p_y(i) \log p_y(i)$, которая выражает неопределенность означенных переменных. Еще раз подчернем, что рассматриваемое сечение определяется в результате означивания подмножества входных переменных ***y***. Из этого следует, что

$$\log k \geq - \sum_{i=1,k} p_y(i) \log p_y(i).$$

Доля $p_y(i)$ в общем случае зависит от способа построения бинарной программы выше ***y***-сечения. Чтобы в этом убедиться, рассмотрим два примера.

***Пример* 1.** Пусть функция $f(x_1, x_2, y_1, y_2, \boldsymbol{z}) =$

$$(x_1\ \bar{x}_2 \vee \bar{x}_1\ \bar{x}_2)(y_1y_2 \vee \bar{y}_1\ \bar{y}_2)f_1(\boldsymbol{z}) \vee (\bar{x}_1x_2)(y_1y_2 \vee \bar{y}_1\ \bar{y}_2)f_1(\boldsymbol{z}) \vee x_1x_2(\bar{y}_1y_2 \vee y_1\ \bar{y}_2)f_2(\boldsymbol{z}).$$

Здесь $\boldsymbol{z}$ – это не пустой набор переменных отличных от x_1, x_2, y_1, y_2. Начальный фрагмент бинарной программы для этой функции приведен на рис.1. Здесь для каждого узла вначале указан его номер, затем (в скобках) - доли путей, ведущих из истока в этот узел, от числа всех приведенных путей. Так в узел **1** ведут два пути, которые получаются в результате означиваний переменных $\{x_1 = 1, x_2 = 0\}$ и $\{x_1 = x_2 = 0\}$. На рис. 1 этим путям соответствует одна дуга, помеченная $\{x_1\ \bar{x}_2\} \cup \{\bar{x}_1\ \bar{x}_2\}$. Соответствие между меткой дуги и соответствующими означиваниями очевидно. Всего же путей из истока, которые получаются в результате означиваний переменных x_1 и x_2, четыре. Следовательно, доля всех путей из истока в узел **1** среди всех путей равна ½. Аналогично вычисляются доли путей, ведущие в остальные узлы.

Построим для этой функции бинарную программу с другим порядком означиваний переменных, как на рис.2. Видно, что доли путей ведущих в узлы, которым приписаны одинаковые функции для первой программы и для второй (узлы **4** и **5** первой программы соответствуют узлам **3** и **4** второй) отличаются.

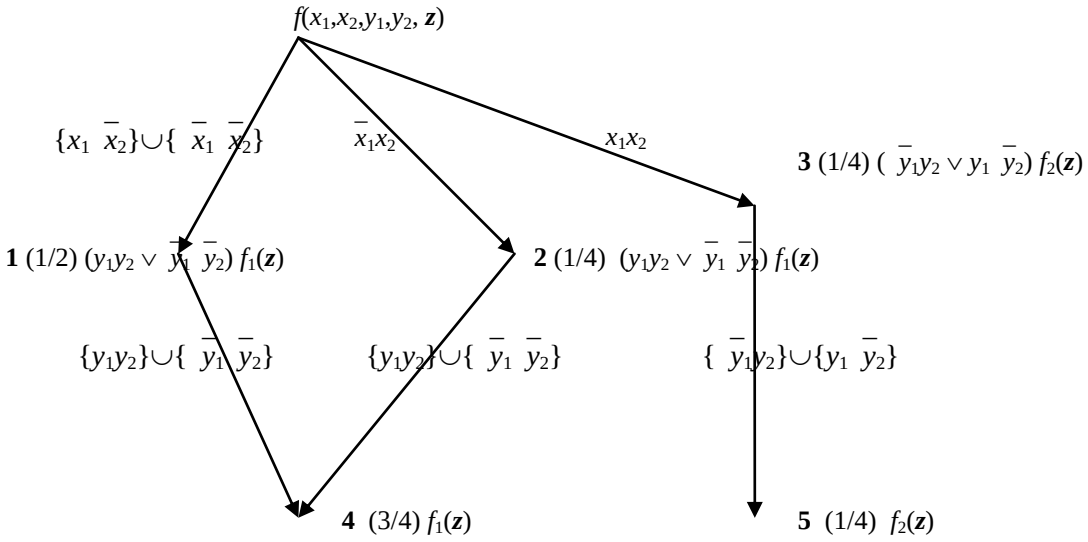

Рис. 1

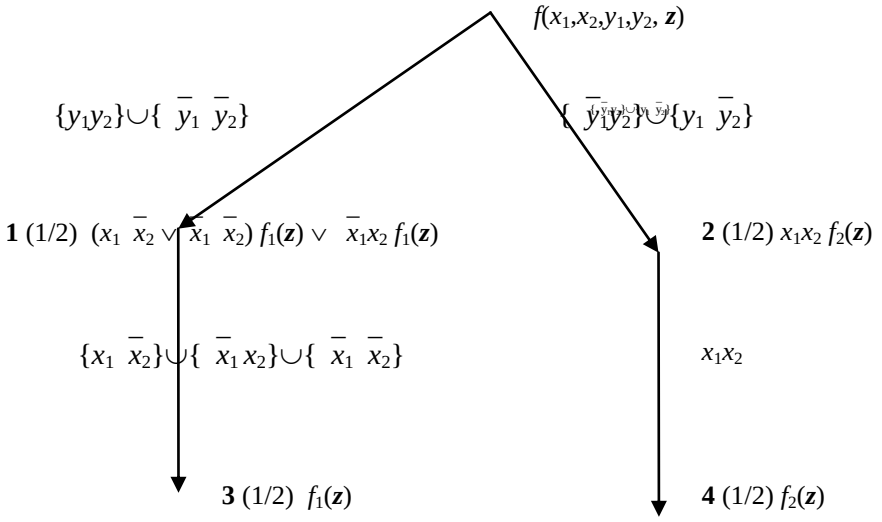

Рис. 2

***Пример* 2.** Пусть логическая функция $f(x_1, x_2, y_1, y_2, \mathbf{z}) =$

$(x_1\ \bar{x}_2\vee\ \bar{x}_1\ \bar{x}_2)(y_1y_2\vee\ \bar{y}_1y_2\vee\ \bar{y}_1\ \bar{y}_2)f_1(\mathbf{z})\vee(\ \bar{x}_1x_2)(y_1y_2\vee y_1\ \bar{y}_2)f_1(\mathbf{z})\vee x_1x_2\ \bar{y}_1\ \bar{y}_2f_2(\mathbf{z})$.

Здесь $\boldsymbol{z}$ – это не пустой набор переменных отличных от x_1, x_2, y_1, y_2. Начальный фрагмент бинарной программы для этой функции приведен на рис.3.

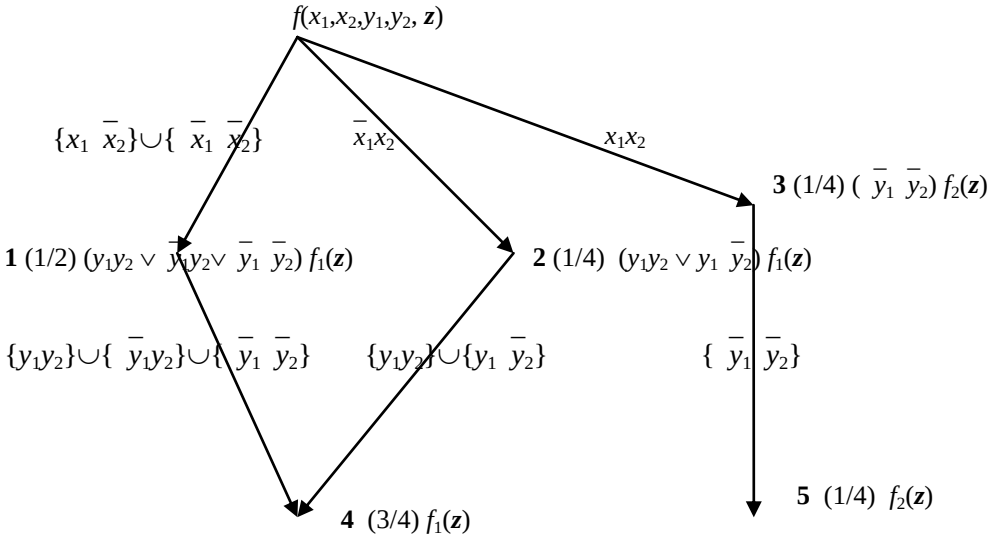

Рис. 3

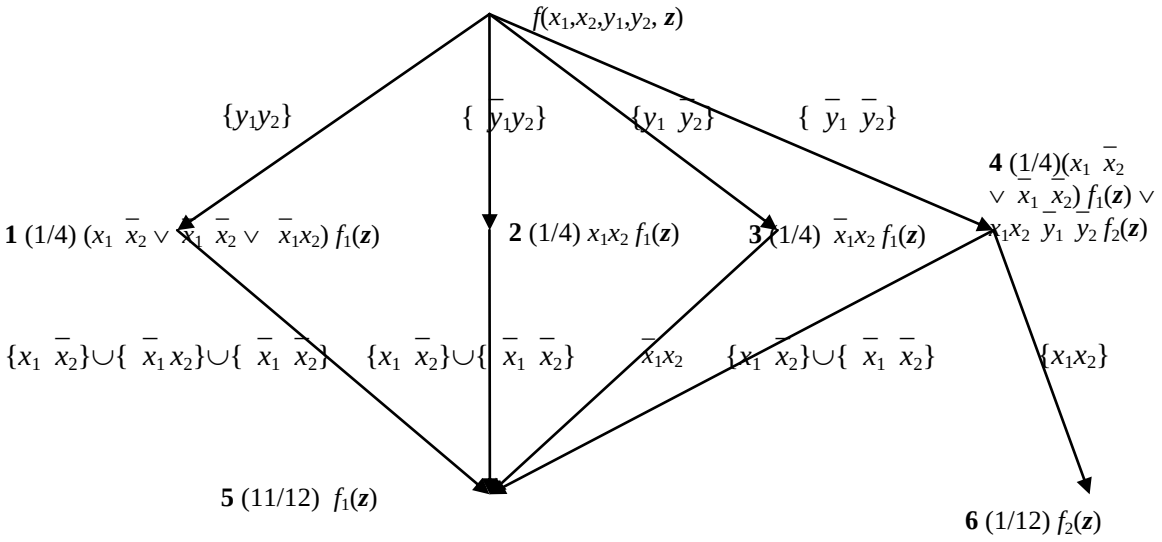

Рис. 4

Таким образом, если мы хотим ввести меру неопределенности как функцию бинарной программы, расположенной выше $\boldsymbol{y}$-сечения, то необходимо учитывать последовательности означиваний переменных в бинарной программе.

Введем ряд определений.

Пусть i-ый узел $\boldsymbol{y}$-сечения соединен $N_{\boldsymbol{yz}|\boldsymbol{y}}(j_1|i)$, $N_{\boldsymbol{yz}|\boldsymbol{y}}(j_2|i)$, …, $N_{\boldsymbol{yz}|\boldsymbol{y}}(j_k|i)$ путями в точности с узлами соответственно j_1, j_2, …, j_k $\boldsymbol{yz}$-сечения и $N_{\boldsymbol{yz}} = N_{\boldsymbol{yz}|\boldsymbol{y}}(j_1|i) + N_{\boldsymbol{yz}|\boldsymbol{y}}(j_2|i) + \ldots + N_{\boldsymbol{yz}|\boldsymbol{y}}(j_k|i)$ – это число всех путей, ведущих из i-го узла в $\boldsymbol{yz}$-сечение. Тогда величина $p_{\boldsymbol{yz}|\boldsymbol{y}}(j_s|\, i) = N_{\boldsymbol{yz}|\boldsymbol{y}}\,(j_s|\, i)/N_{\boldsymbol{yz}}$ – есть доля путей, ведущих из i-го узла $\boldsymbol{y}$-сечения в j_s-ый узел $\boldsymbol{yz}$-сечения, $s = 1, 2, \ldots, k$ среди всех путей из i-го узла $\boldsymbol{y}$-сечения в $\boldsymbol{yz}$-сечение. Понятно, что $\sum_{s=1,k} p_{\boldsymbol{yz}|\boldsymbol{y}}(j_s|\, i) = 1$, так как $j_1, j_2, \ldots, j_k$ – это все узлы $\boldsymbol{yz}$-сечения, в которые имеются пути из i-го узла $\boldsymbol{y}$-сечения.

Пусть в бинарной программе i-ый узел $\boldsymbol{y}$-сечения соединен путями в точности с узлами $j_1, j_2, \ldots, j_k$ $\boldsymbol{yz}$-сечения; $h_1, h_2, \ldots, h_m$ – суть все узлы $\boldsymbol{yzu}$-сечения, в которые мы попадаем из узлов $j_1, j_2, \ldots, j_k$ $\boldsymbol{yz}$-сечения (см.рис.5). Тем самым из i-го узла $\boldsymbol{y}$-сечения в узлы $h_1, h_2, \ldots, h_m$ $\boldsymbol{yzu}$-сечения мы можем попасть только через какой-либо узел из $j_1, j_2, \ldots, j_k$ $\boldsymbol{yz}$-сечения.

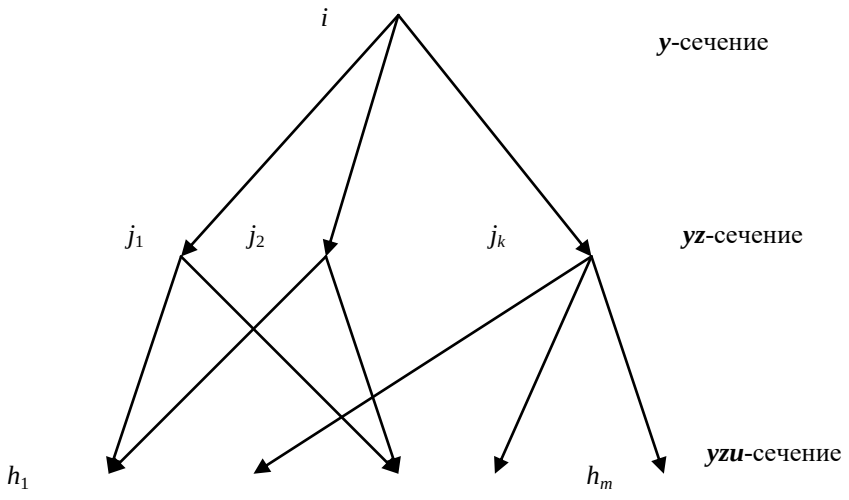

Рис.5

Из определения следует равенство:

$$\sum_{t=1,k} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \,|\, j_t)\; p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t \,|\, i) = p_{\boldsymbol{yzu}|\boldsymbol{y}}(h_s \,|\, i).$$

Действительно, i-ый узел $\boldsymbol{y}$-сечения соединен путями с узлами j_1, j_2, …, j_k $\boldsymbol{yz}$-сечения и доля таких путей, ведущих в j_t–ый узел равна $p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t \,|\, i)$. В свою очередь j_t–ый узел соединен с некоторыми узлами из совокупности

h_1, h_2, …, h_m ***yzu***-сечения, причем доля таких путей, ведущих из j_t–го узла в h_s-ый узел равна $p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \mid j_t)$. Но тогда доля путей из i-го узла ***y***-сечения в h_s-ый узел ***yzu***-сечения равна указанной сумме.

Справедливы следующие равенства:

$\sum_{s=1,m} p_{\boldsymbol{yzu}|\boldsymbol{y}}(h_s \mid i) = 1$; $\sum_{t=1,k} p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t \mid i) = 1$; $\sum_{s=1,m} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \mid j_t) = 1$, $t = 1, 2, \ldots, k$.

Последнее равенство выполняется для тех пар узлов, которые соединены путями.

Теорема* 1.** *Для всяких, не пересекающихся множеств* ***y, ***z***, ***u*** *аргументов имеет место равенство* $\sum_{t=1,k,\ s=1,m} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s|j_t)\, p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t|i) = 1$.

Доказательство. По определению $\sum_{t=1,k} p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t \mid i) = 1$.

Для каждого из узлов j_1, j_2, …, j_k ***yz***-сечения имеем равенство $\sum_{s=1,m} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \mid j_t) = 1$, $t = 1, 2, \ldots, k$.

Умножая каждый член первой суммы на $\sum_{s=1,m} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \mid j_t)$, получаем равенство $\sum_{s=1,m} \sum_{t=1,k} p_{\boldsymbol{yzu}|\boldsymbol{yz}}(h_s \mid j_t)\, p_{\boldsymbol{yz}|\boldsymbol{y}}(j_t \mid i) = 1$.

Теорема доказана.

Определим рекурсивно величину $p_{\boldsymbol{z}}(i)$ как долю путей, проходящих из истока в i-ый узел ***z***-сечения бинарной программы.

Пусть $\boldsymbol{z} = z_1 z_2 \ldots z_m$ – последовательность аргументов и однородная бинарная программа имеет z_1-сечение, z_1z_2-сечение, …, $z_1z_2\ldots z_m$-сечение, причем множества путей заданы следующими значениями: $N_{z_1z_2\ldots z_m|\, z_1z_2\ldots z_{m-1}}(i|\, j_{m-1})$, $N_{z_1z_2\ldots z_{m-1}|\, z_1z_2\ldots z_{m-2}}(j_{m-1}|\, j_{m-2})$, …, $N_{z_1z_2|\, z_1}(j_2|\, j_1)$, $N_{z_1}(j_1)$ при соответствующем именовании узлов сечений. Нетрудно увидеть, что число $N_{z_1z_2\ldots z_m}(i)$ всех путей из истока программы в i-ый узел ***z***-сечения равно сумме

$$\sum_{j_1,j_2,\ldots,j_{m-1}} N_{z_1z_2\ldots z_m|z_1z_2\ldots z_{m-1}}(i|j_{m-1})\, N_{z_1z_2\ldots z_{m-1}|z_1z_2\ldots z_{m-2}}(j_{m-1}|j_{m-2})\ldots$$

$$\ldots\, N_{z_1z_2|\, z_1}(j_2|\, j_1)\, N_{z_1}(j_1),$$

где суммирование ведется по всем узлам указанных сечений. Используя частичные суммы, получим, что эта сумма равна

$$\sum_{j_{m-1}} N_{z_1 z_2 \ldots z_m | z_1 z_2 \ldots z_{m-1}}(i | j_{m-1})\, N_{z_1 z_2 \ldots z_{m-1}}(j_{m-1}).$$

Здесь $N_{z_1 z_2 \ldots z_{m-1}}(j_{m-1})$ – это число путей, ведущих из истока в j_{m-1}-ый узел $z_1 z_2 \ldots z_{m-1}$-сечения.

Обозначим $N_{\boldsymbol{z}}$ общее число различных путей из истока в $\boldsymbol{z}$-сечение и разделим каждый член последней суммы на величину $N_{\boldsymbol{z}}$. Тогда отношение

$$N_{z_1 z_2 \ldots z_m | z_1 z_2 \ldots z_{m-1}}(i | j_{m-1})\, N_{z_1 z_2 \ldots z_{m-1}}(j_{m-1}) / N_{\boldsymbol{z}}$$

представляет собой долю всех путей из истока в i-ый узел $\boldsymbol{z}$-сечения, проходящих через j_{m-1}-ый узел $z_1 z_2 \ldots z_{m-1}$-сечения среди общего множества путей из истока в $\boldsymbol{z}$-сечение.

Доля $p_{z_1 z_2 \ldots z_{m-1}}(j_{m-1})$ путей из истока в $z_1 z_2 \ldots z_m$-сечение, проходящих через узел j_{m-1} $z_1 z_2 \ldots z_{m-1}$-сечения также определяется общим числом таких путей среди остальных, ведущих в $z_1 z_2 \ldots z_m$-сечение. Причем только часть $p_{z_1 z_2 \ldots z_m | z_1 z_2 \ldots z_{m-1}}(i | j_{m-1})$ из них ведет далее в i-ый узел $z_1 z_2 \ldots z_m$-сечение. Эти рассуждения позволяют ввести такое индуктивное определение.

Базис конструкции. Если множество аргументов пусто, то $p_{\lambda}(1) = 1$. В этом случае узел λ-сечения - это исток бинарной программы.

Индуктивное построение. Пусть для некоторой последовательности $\boldsymbol{z}$ аргументов $p_{\boldsymbol{z}}(j)$ есть доля путей из истока в j-ый узел $\boldsymbol{z}$-сечения, y – это новая переменная и $p_{\boldsymbol{z}y|\boldsymbol{z}}(i \mid j)$ – доля путей, ведущих из в j-го узла $\boldsymbol{z}$-сечения в i-ый узел $\boldsymbol{z}y$-сечения. Тогда $p_{\boldsymbol{z}y}(i) = \sum_j p_{\boldsymbol{z}y|\boldsymbol{z}}(i \mid j)\, p_{\boldsymbol{z}}(j)$.

Это определение распространяется на произвольные множества аргументов $\boldsymbol{y}$ и $\boldsymbol{z}$ следующим образом.

$$p_{\boldsymbol{yz}}(i) = \sum_j p_{\boldsymbol{yz}|\boldsymbol{y}}(i \mid j)\, p_{\boldsymbol{y}}(j).$$

Пусть $\boldsymbol{z} = \boldsymbol{z}'\, u$ и мы уже получили, что $p_{\boldsymbol{yz}'}(h) = \sum_j p_{\boldsymbol{yz}'|\boldsymbol{y}}(h \mid j)\, p_{\boldsymbol{y}}(j)$. По определению

$$p_{\boldsymbol{yz}'u}(i) = \sum_{h=1,q} p_{\boldsymbol{yz}'u|\boldsymbol{yz}}(i \mid h)\, p_{\boldsymbol{yz}}(h) = \sum_{h=1,q} p_{\boldsymbol{yz}'u|\boldsymbol{yz}}(i \mid h) \sum_{j=1,k} p_{\boldsymbol{yz}'|\boldsymbol{y}}(h \mid j)\, p_{\boldsymbol{y}}(j) =$$

$$\sum_{h=1,q,j=1,k} p_{\boldsymbol{yz}'u\,|\boldsymbol{yz}}(i \mid h)\; p_{\boldsymbol{yz}'\,|\boldsymbol{y}}(h \mid j)\; p_{\boldsymbol{y}}(j).$$

Известно, что $\sum_{h=1,q} p_{\boldsymbol{yz}'u\,|\boldsymbol{yz}}(i \mid h)\; p_{\boldsymbol{yz}'\,|\boldsymbol{y}}(h \mid j) = p_{\boldsymbol{yz}'u\,|\boldsymbol{y}}(i \mid j)$. Следовательно, $p_{\boldsymbol{yz}'u}(i) = \sum_{j=1,k} p_{\boldsymbol{yz}'u\,|\boldsymbol{y}}(i \mid j)\; p_{\boldsymbol{y}}(j)$.

Справедливо утверждение.

Теорема* 2.** *Справедливо равенство* $\sum_{i=1,h} p_{\boldsymbol{z}}(i) = 1$, где суммирование ведется по всем узлам ***z-сечения.

Доказательство. Пусть $p_{\boldsymbol{u}}(j)$ есть известная доля путей из истока в узел j ***u***-сечения, y – это переменная, $\boldsymbol{z} = \boldsymbol{u}y$ и $\sum_{j=1,k} p_{\boldsymbol{u}}(j) = 1$. По определению, $p_{\boldsymbol{u}y}(i) = \sum_{j=1,k} p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j)\; p_{\boldsymbol{u}}(j)$, где $p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j)$ – относительная доля путей, ведущих из j-го узла ***u***-сечения в i-ый узел ***u****y*-сечения. Следовательно, произведение $p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j)\; p_{\boldsymbol{u}}(j)$ есть доля путей из истока в в i-ый узел ***u****y*-сечения, проходящих через j-ый узел ***u***-сечения.

$\sum_{i=1,h} p_{\boldsymbol{u}y}(i) = \sum_{i=1,h} \sum_{j=1,k} p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j)\; p_{\boldsymbol{u}}(j) = \sum_{j=1,k}(\sum_{i=1,h} p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j))\; p_{\boldsymbol{u}}(j) = \sum_{j=1,k} p_{\boldsymbol{u}}(j) = 1$.

При доказательстве мы использовали равенства $\sum_{i=1,h} p_{\boldsymbol{u}y|\boldsymbol{u}}(i \mid j) = 1$ и $\sum_{j=1,k} p_{\boldsymbol{u}}(j) = 1$.

Теорема доказана.

Назовем *логической энтропией* функции f, определяемой переменными ***y***, величину

$$H^f_{\boldsymbol{y}} = -\sum_{i=1,k} p_{\boldsymbol{y}}(i) \log p_{\boldsymbol{y}}(i),$$

где суммирование ведется по всем узлам ***y***-сечения.

Из определения логической энтропии следует, что в общем случае ее значение зависит от порядка означивания переменных ***y*** в бинарной программе. Это поясняют примеры 1 и 2. Опишем класс логических функций, для которых логическая энтропия $H^f_{\boldsymbol{y}}$ определяется лишь видом множества ***y*** переменных и не зависит от порядка их означивания.

Справедливо следующее утверждение.

***Лемма* 3**. *Пусть* $f(\boldsymbol{x}, \boldsymbol{w})$ *представляет* всюду вычислимую функцию $\boldsymbol{w} = \varphi(\boldsymbol{x})$ и $\boldsymbol{y} \subseteq \boldsymbol{x}$. *Тогда разложение этой функции по переменным* $\boldsymbol{y}$ *приводит к одному выражению с точностью до перестановки конъюнктивных и дизъюнктивных членов.*

Доказательство следует из того, что при любом означивающем $\boldsymbol{y}$-наборе $\boldsymbol{\sigma}_y$ в разложении присутствует конъюнкция, первым членом которой является конъюнкт $\boldsymbol{y}^{\boldsymbol{\sigma y}}$ и вторым функция, которая получается из $f(\boldsymbol{x}, \boldsymbol{w})$ в результате этого означивания.

***Теорема* 4**. *Значение* H^f_y *не зависит от порядка означиваний переменных* $\boldsymbol{y}$, *а определяется только множеством* $\boldsymbol{y}$.

Из этого следует, что логическая энтропия H^f_y функции $f(\boldsymbol{Y}, \boldsymbol{w})$, представляющей физическую систему $\boldsymbol{w} = \varphi(\boldsymbol{Y})$, где φ - всюду определенная функция совпадает с ее энтропией, определяемой переменными $\boldsymbol{y}$, как она введена в [1].

Если же функция $f(\boldsymbol{Y}, \boldsymbol{w})$ представляет не всюду определенную вычислимую функцию, то логическая энтропия зависит от порядка означивания переменных.

Если ясно, о какой функции идет речь, то в обозначении H^f_y будем опускать верхний индекс.

Понятие логической энтропии введено, исходя из представления бинарной программы как некоторого вычислителя. Тем самым она выступает мерой неопределенности процесса вычисления. Ту же формулу можно получить в результате несколько иных рассуждений. Как следствие получим, что логическая энтропия служит и мерой сложности функции.

Определим по функции $f(\boldsymbol{y}, \boldsymbol{z})$ отношение $\boldsymbol{y}$–эквивалентности $\boldsymbol{y}$-наборов:

$$\boldsymbol{\sigma}^1_y \approx \boldsymbol{\sigma}^2_y \Leftrightarrow f(\boldsymbol{\sigma}^1_y, \boldsymbol{z}) = f(\boldsymbol{\sigma}^2_y, \boldsymbol{z}).$$

При этом говорим, что функция $f(\boldsymbol{\sigma}^1_y)$ *определяется* тем классом эквивалентности, которому принадлежит $\boldsymbol{y}$-набор $\boldsymbol{\sigma}^1_y$.

Пусть число всех классов ***y***-эквивалентности равно k: $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^k$; $\boldsymbol{\sigma}^1{}_1 \in \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2{}_1 \in \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^k{}_1 \in \boldsymbol{\sigma}^k$ суть представители этих классов и Σ_i есть множество всех единичных означиваний функции $f(\boldsymbol{\sigma}^i_y, \boldsymbol{z})$, $i = 1, 2, \ldots, k$. Назовем множество $\boldsymbol{\sigma}^i \times \Sigma_i$ означиваний - *порождаемым* i-ым классом ***y***-эквивалентности. Пусть Σ есть множество всех единичных означиваний функции $f(\boldsymbol{y}, \boldsymbol{z})$. Тогда ***y***–эквивалентность определяет разбиение множества Σ на k не пересекающихся подмножеств: $\Sigma = \cup_{i=1,k} \boldsymbol{\sigma}^i \times \Sigma_i$, $i = 1, 2, \ldots, k$ – по числу классов эквивалентности.

Функции $f(\boldsymbol{\sigma}^i_y, \boldsymbol{z})$, $i = 1, 2, \ldots, k$, связаны с исходной функцией $f(\boldsymbol{y}, \boldsymbol{z})$ очевидным образом:

$$f(\boldsymbol{y}, \boldsymbol{z}) = \vee_{i=1,k} (\boldsymbol{y}^{\boldsymbol{\sigma}1^i} \vee \boldsymbol{y}^{\boldsymbol{\sigma}2^i} \vee \ldots \vee \boldsymbol{y}^{\boldsymbol{\sigma}m^i}) f(\boldsymbol{\sigma}^i{}_1, \boldsymbol{z}).$$

Здесь $\boldsymbol{\sigma}_1{}^i, \boldsymbol{\sigma}_2{}^i, \ldots, \boldsymbol{\sigma}_m{}^i$ – означивания переменных ***y***, составляющие i-ый класс эквивалентности, $i = 1, 2, \ldots, k$, они определяют одну функцию $f(\boldsymbol{\sigma}^i{}_1, \boldsymbol{z})$, (обозначим её f_i).

Представим функцию $f(\boldsymbol{y}, \boldsymbol{z})$ следующим образом:

$$f(\boldsymbol{y}, \boldsymbol{z}) = \vee_{i=1,k} [\boldsymbol{y}^{\boldsymbol{\sigma}^i}] f_i,$$

где $[\boldsymbol{y}^{\boldsymbol{\sigma}^i}]$ обозначает дизъюнкцию конъюнктов, определяемых всеми ***y***–наборами одного класса эквивалентности $\boldsymbol{\sigma}^i = \{\boldsymbol{\sigma}_1{}^i, \boldsymbol{\sigma}_2{}^i, \ldots, \boldsymbol{\sigma}_m{}^i\}$, $i = 1, 2, \ldots, k$.

Разлагая каждую функцию $f(\boldsymbol{\sigma}^i{}_1, \boldsymbol{z})$ по переменным ***z***, получим равенство

$$f(\boldsymbol{y}, \boldsymbol{z}) = \vee_{i=1,k\ j=1,h} [\boldsymbol{y}^{\boldsymbol{\sigma}^i}] [\boldsymbol{z}^{\boldsymbol{\lambda}^j}] f(\boldsymbol{\sigma}_1{}^i, \boldsymbol{\lambda}_1{}^j) = \vee_{i=1,k\ j=1,h} [\boldsymbol{y}^{\boldsymbol{\sigma}^i}] [\boldsymbol{z}^{\boldsymbol{\lambda}^j}] f_{ij}.$$

Сокращенно это разложение обозначим так: $f(\boldsymbol{y}, \boldsymbol{z}) = \vee_{i=1,k\ j=1,h} [\boldsymbol{y}^{\boldsymbol{\sigma}^i} \boldsymbol{z}^{\boldsymbol{\lambda}^j}] f_{ij}$. Множества $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^k$ попарно не пересекаются, но множества $\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \ldots, \boldsymbol{\lambda}^h$ могут пересекаться.

Мы рассматриваем логические функции, представляющие лишь всюду определенные вычислимые функции. Поэтому каждая функция f_{ij} не

равна тождественному нулю, $i = 1, 2, \ldots, k$, $j = 1,2, \ldots, h$. Несколько различных пар (i, j) индексов могут определять эквивалентные логические функции f_{ij}. Перечислим все классы ***yz***-эквивалентности. Пусть t-ый класс ***yz***-эквивалентности характеризуется множеством $\{(i, j_1), (i, j_2), \ldots, (i, j_q)\}$ пар индексов в разложении функции $f(\boldsymbol{y}, \boldsymbol{z})$. Тогда мы говорим, что t-ый класс ***yz***-эквивалентности *порожден* i-ым классом ***y***-эквивалентности.

Разложение функции $f(\boldsymbol{y}, \boldsymbol{z})$ по переменным ***y*** и ***z***, можно представить ее в виде матрицы M_f, как на рис.5.

z / ***y***	λ^1	λ^2	λ^3	…
σ^1	f_{11}	f_{12}	f_{13}	…
σ^2	f_{21}	f_{22}	f_{23}	…
σ^3	f_{31}	f_{32}	f_{33}	…
…	…	…	…	…

Рис. 5

Здесь f_{ij} обозначает функцию $f(\sigma_1^{\,i}, \lambda_1^{\,j})$, i-ая строка соответствует i-му классу ***y***-эквивалентности, i=1, 2, …, k, j-ый столбец - j-му классу ***z***-эквивалентности, j=1, 2, …, h. В общем случае нескольким парам (i, j), где i – номер столбца и j – номер строки, может соответствовать один класс эквивалентности ***yz***-наборов. Но число классов ***yz***-эквивалентности не больше числа таких различных пар. Все порожденные одним i–м ***y***-классом ***yz***-классы характеризуются эквивалентными функциями из одной i–ой строки.

Введем теперь показатель, аналогичный тому, который использовался при исследовании бинарных программ.

Определим $p_{yz|y}(j|i)$ как условную долю j-го ***yz***-класса, порожденного i-м ***y***-классом среди всех ***yz***-классов, порожденных i-м ***y***-классом. Из разло-

жения функции f по переменным и определения бинарной программы понятно, что все введенные ранее равенства относительно условных и абсолютных долей сохраняются при новой интерпретации. В частности, определение абсолютной доли $p_{\boldsymbol{y}}(i)$ i-го $\boldsymbol{y}$-класса среди остальных $\boldsymbol{y}$-классов выглядит так.

Базис конструкции. При пустом множестве аргументов имеется единственный класс эквивалентности, который определяется самой функцией f и поэтому доля его $p_{\boldsymbol{\lambda}}(1) = 1$.

Индуктивное построение. Пусть для некоторой последовательности $\boldsymbol{z}$ аргументов $p_{\boldsymbol{z}}(j)$ есть доля j-го класса $\boldsymbol{z}$-эквивалентности среди остальных $\boldsymbol{z}$-классов, y – это новая переменная и $p_{\boldsymbol{z}y|\boldsymbol{z}}(i \mid j)$ – доля i-го $\boldsymbol{z}y$-класса среди всех $\boldsymbol{z}y$-классов, порожденных j-м $\boldsymbol{z}$-классом. Тогда $p_{\boldsymbol{z}y}(i) = \sum_j p_{\boldsymbol{z}y|\boldsymbol{z}}(i \mid j)\, p_{\boldsymbol{z}}(j)$.

Рассмотрим однородную бинарную программу π_f для функции $f(\boldsymbol{y})$. Пусть $N_1, N_2, \ldots, N_k$ – суть все узлы $\boldsymbol{y}$-сечения и $T_1, T_2, \ldots, T_k$ – совокупности путей, проходящих из истока программы в узлы соответственно $N_1, N_2, \ldots, N_k$, причем мощности множеств $T_1, T_2, \ldots, T_k$ равны соответственно $t_1, t_2, \ldots, t_k$ и $t = \sum_{i=1,k} t_i$. Каждый узел N_i соответствует, с одной стороны, в точности одному классу $\boldsymbol{y}$-эквивалентности, а с другой, - единственной логической функции, определяемой всяким $\boldsymbol{y}$-набором из этого класса $\boldsymbol{y}$-эквивалентности. По определению, доля $p_{\boldsymbol{y}}(i)$ всех наборов, определяемых одним классом $\boldsymbol{y}$-эквивалентности, соответствующего узлу N_i, совпадает с долей $p_{\boldsymbol{y}}(i)$ путей из истока в i-ый узел $\boldsymbol{y}$-сечения. $i = 1, 2, \ldots, k$.

При таком определении абсолютной доли, выполняется теорема, аналогичная Теореме 1.

***Теорема* 5.** *Имеет место равенство* $\sum_i p_{\boldsymbol{z}}(i) = 1$.

Если мы рассматриваем логические функции, представляющие всюду определенные вычислимые функции, то верно следующее утверждение.

***Лемма* 6.** *Пусть* $p_{\boldsymbol{y}}(i)$ *есть доля* i*-го класса* $\boldsymbol{y}$*-эквивалентности,* $q_{\boldsymbol{z}}(j)$ *– доля* j*-го класса* $\boldsymbol{z}$*-эквивалентности. Тогда доля* $\boldsymbol{yz}$*-наборов, которые ха-*

рактеризуются принадлежностью ***y****-наборов* *i-му классу* ***y****-эквивалентности и* ***z****-наборов* *j-му классу* ***z****-эквивалентности среди всех* ***yz****-наборов равна произведению* $p_y(i)q_z(j)$, $i = 1, 2,\ldots, k, j = 1, 2, \ldots, h$.

Доказательство. Все логические функции f_{ij}, $i=1, 2, \ldots, k, j=1, 2, \ldots. h$ не являются тождественно нулевыми. Следовательно, для каждого ***y***-набора $\boldsymbol{\sigma}_r^i \in \boldsymbol{\sigma}^i$ и для каждого ***z***-набора $\boldsymbol{\lambda}_q^j \in \boldsymbol{\lambda}^j$ все конъюнкты вида $\boldsymbol{y}^{\sigma r^i}\boldsymbol{z}^{\lambda q^j}$ входят в разложение исходной функции.

Лемма доказана.

С другой стороны, если рассматривать логические функции, представляющие не всюду определенные функции, то утверждение места не имеет. Действительно, в этом случае некоторые функции матрицы M_f могут быть тождественно нулевыми и, следовательно, для ***y***-набора $\boldsymbol{\sigma}_r^i \in \boldsymbol{\sigma}^i$ для некоторых ***z***-наборов $\boldsymbol{\lambda}_q^j \in \boldsymbol{\lambda}^j$ конъюнкты вида $\boldsymbol{y}^{\sigma r^i}\boldsymbol{z}^{\lambda q^j}$ могут не входить в разложение исходной функции. Следовательно, не все ***yz***-наборы, у которых ***y***-наборы из класса $\boldsymbol{\sigma}^i$, характеризуются принадлежностью ***z***-наборов какому-либо классу ***z***-эквивалентности.

Сложность зависимости функции $f(\boldsymbol{y})$ от переменных ***y*** характеризуется фактор-множеством ***y***-эквивалентности. Действительно, чем больше фактор-множество, тем больше различных подформул встречается в разложении логической функции по этим переменным. В терминах бинарной программы мощности фактор-множества ***y***–эквивалентности соответствует число узлов ***y***–сечения.

Если $p_y(i)$ есть доля путей, проходящих из истока в i-ый узел y-сечения, то обратная величина $t_y(i) = 1/ p_y(i)$ пропорциональна числу путей, ведущих из истока в этот узел, $i = 1,2, \ldots, k$. $\sum_{i=1,k} p_y(i) \log t_y(i)$ есть математическое ожидание случайной величины $\log t_y(i)$, $i = 1,2, \ldots, k$, с распределением $p_y(i)$.

Введем следующее определение.

Эффективной пропускной способностью одного узла ***y***–сечения бинарной программы назовем величину τ, такую, что $\log \tau = \sum_{i=1,k} p_y(i) \log t_y(i)$. Таким образом, эффективную пропускную способность можно представлять как математическое ожидание величины, пропорциональной числу путей, проходящих через один узел ***y***-сечения. Поэтому эффективная пропускная способность отличается от «реальной» некоторой аддитивной константой. Если доли $p_y(i)$ одинаковы, то $\log \tau = \log \bar{t}$. В этом случае эффективная пропускная способность совпадает с числом путей из истока, которые проходят через один узел ***y***-сечения.

Введем величину $t = \sum_{i=1,k} t_y(i)$ которая пропорциональна общему числу путей, ведущих из истока в ***y***-сечение. Величину $H_y = \log (t/\tau)$ назовем *эффективным* ***y****–сечением.* Тем самым эффективное ***y***–сечение есть мера сложности бинарной программы, расположенной выше ***y***–сечения: при большем эффективном сечении H_y и при постоянном числе всех путей из истока в ***y***–сечение, среднее число путей, проходящих через один узел сечения, мало, а число узлов в сечении - велико. Но чем шире ***y***–сечение, тем сложнее зависимость логической функции от переменных ***y***, так как число различных подформул, участвующих в разложении функции по переменным ***y*** больше. Поэтому зависимости исходной функции от переменных ***y*** выражается сложнее.

$$H_y = \log (t/\tau) = \log t - \log \tau = \log t - \sum_{i=1,k} p_y(i) \log t_y(i)$$
$$= -\sum_{i=1,k} (t_y(i)/t) \log (t_y(i)/t) = -\sum_{i=1,k} p_y(i) \log p_y(i).$$

Как видно, величина эффективного ***y***-сечения и энтропии логической функции совпадают. Заметим, что энтропия получена из определения неопределенности выходных значений от входных, а эффективное сечение получено из анализа сложности зависимости функции от ее аргументов.

2. Некоторые интерпретации логической энтропии. Поясним с содержательной точки зрения понятие логической энтропии.

Пусть $f(\boldsymbol{y}, \boldsymbol{z})$ есть логическая функция и $\xi(\boldsymbol{y})$ есть функция, значение которой от аргументов $\boldsymbol{\sigma}_y$ равно номеру того класса $\boldsymbol{y}$-эквивалентности, которому принадлежит кортеж $\boldsymbol{\sigma}_y$. Обозначим долю таких $\boldsymbol{y}$-наборов, при которых значение функции $\xi(\boldsymbol{y})$ равно i, $i = 1, 2, \ldots, k$, среди всех наборов через p_i. Будем говорить, что переменные $\boldsymbol{y}$ имеют распределение $p_y(1)$, $p_y(2)$, …, $p_y(k)$. Понятно, что $\Sigma_{i=1,k}\, p_y(i) = 1$. Назовем функцию, которая получается из $f(\boldsymbol{y}, \boldsymbol{z})$ подстановкой $\boldsymbol{y}$–наборов из i-го класса эквивалентности - *соответствующей* значению p_i. Следовательно, показатель неожиданности $-\log_2 p_i$ функции $\xi(\boldsymbol{y})$ зависит от доли наборов, определяемых i-м классом $\boldsymbol{y}$-эквивалентности. Чем больше наборов в i-м классе, тем меньше неожиданность того, что конкретное означивание ему принадлежит. В терминах бинарной программы это выглядит так: чем больше путей определяется одним классом $\boldsymbol{y}$–эквивалентности, тем большая доля вычислений осуществляется промежуточным вычислителем, который соответствует данному классу и ассоциируется с соответствующим узлом $\boldsymbol{y}$-сечения программы. Логическая энтропия, определяемая переменными $\boldsymbol{y}$, есть усредненная неожиданность того, что конкретное вычисление будет реализовываться промежуточным вычислителем, соответствующим конкретному классу $\boldsymbol{y}$-эквивалентности (в терминах бинарной программы – некоторому узлу $\boldsymbol{y}$-сечения).

В теории информации энтропия служит мерой *свободы* системы: чем больше у системы степеней свободы, т.е. чем меньше на нее наложено ограничений, тем больше ее энтропия. Поэтому энтропия максимальна при одинаковой доле наблюдаемых событий, а всякое отклонение от него приводит к ее уменьшению. В пределе, когда доля одного события равна 1, энтропия равна нулю. Применительно к логической энтропии, под степенями свободы понимается число классов $\boldsymbol{y}$-эквивалентности. Чем их меньше, тем больше определенность, какой промежуточный вычислитель используется для продолжения вычисление. И наоборот, чем больше классов эквивалент-

ности и чем однороднее доли p_1, p_2, …, p_k, тем больше возможностей для выбора того или иного промежуточного вычислителя и, следовательно, больше неопределенность.

В терминах сложности логической функции эти рассуждения выглядят так: если переменные $\boldsymbol{y}$ определяют небольшое фактор-множество, то они описывают лишь небольшое число различных свойств логической функции. Очевидно, что для описания небольшого числа свойств требуется меньше параметров (в нашем случае - аргументов) и наоборот, чем разнообразнее проявления системы, тем больше параметров необходимо для ее описания. При прочих равных условиях, более сложная система наблюдателю кажется менее определенной и наоборот, чем больше определенности у наблюдателя, тем проще зависимость демонстрируемого поведения системы от входных параметров. В данном контексте под сложностью зависимости системы от аргументов $\boldsymbol{y}$ понимается число классов $\boldsymbol{y}$-эквивалентности. Это согласуется с определением неопределенности функции от переменных $\boldsymbol{y}$, как сложности зависимости от этих аргументов. Чем меньше классов $\boldsymbol{y}$-эквивалентности, тем, с одной стороны, меньше средний показатель неожиданности, а с другой, - тем проще выражается зависимость функции от этих аргументов. Если же классов эквивалентности много, то больше средний показатель неожиданности и функция сложнее зависит от аргументов $\boldsymbol{y}$. Тем самым, чем больше энтропия H_y, тем сложнее выражается зависимость функции от аргументов $\boldsymbol{y}$ и наоборот.

Приведем теперь интерпретацию логической энтропии с позиций статистической модели равновероятных последовательностей. Для этого представим логическую функцию $f(\boldsymbol{y}, \boldsymbol{z})$ таблицей истинности, в которой приведены лишь ее единичные означивания. Разложение по $\boldsymbol{y}$ выглядит следующим образом:

$$f(\boldsymbol{y}, \boldsymbol{z}) = \vee_{i=1,k} [\boldsymbol{y}^{\boldsymbol{\sigma}^i}] \, f(\boldsymbol{\sigma}^i_1, \boldsymbol{z}).$$

Полагаем, что все переменные $\boldsymbol{y}$ случайны и независимы, и $p_y(i)$, $i = 1, 2, \ldots, k$, есть доля наборов, включающих $\boldsymbol{y}$-поднаборы из i-го класса $\boldsymbol{y}$-эквивалентности в достаточно большом случайно порожденном множестве единичных означиваний, мощность которого равна N. Тогда число наборов, обладающих $\boldsymbol{y}$–поднаборами из i-го класса $\boldsymbol{y}$-эквивалентности, равно $N_i = N\, p_y(i)$.

По закону больших чисел во всяком достаточно большом подмножестве единичных означиваний доля означиваний, порожденных i-м классом $\boldsymbol{y}$-эквивалентности совпадает с $p_y(i)$, $i = 1, 2, \ldots, k$ и доли наборов, порожденных каждым из k классов $\boldsymbol{y}$-эквивалентности, не зависят от вида множества единичных означиваний. Поэтому из N наборов $N_i = N\, p_y(i)$ порождены i-м классом и вероятность выбора такого множества единичных означиваний, которая характеризуется распределением $p_y(i)$, $i = 1, 2, \ldots, k$, по классам $\boldsymbol{y}$-эквивалентности, равна $q = p_y(1)^{N_1} p_y(2)^{N_2} \ldots p_y(k)^{N_k} = (p_y(1)^{p_y(1)} p_y(2)^{p_y(2)} \ldots p_y(k)^{p_y(k)})^N$. Неопределенность H^* всего множества единичных означиваний равна $-\log q$. Поэтому $H^* = -N \sum_{i=1,k} p_y(i) \log p_y(i)$. Но тогда $-\sum_{i=1,k} p_y(i) \log p_y(i) = H^*/N$ представляет собой среднюю неопределенность, которая приходится на единственное единичное означивание.

Тем самым, получили еще одну интерпретацию логической энтропии: формула $-\sum_{i=1,k} p_y(i) \log p_y(i)$ определяет среднюю неопределенность, которая приходится на одно единичное означивание функции $f(\boldsymbol{y}, \boldsymbol{z})$, если их классификация осуществляется с помощью $\boldsymbol{y}$-эквивалентности при достаточно большом числе независимо порожденных единичных означиваний.

Наконец, еще одна трактовка логической энтропии основывается на следующих рассуждениях.

Пусть $\boldsymbol{y}$-эквивалентность порождает k классов. Присвоим каждому из N единичных означиваний функции $f(\boldsymbol{y}, \boldsymbol{z})$ номер того класса из {1, 2, …,

$k\}$, порождением которого он является. Число различных перестановок N единичных означиваний из которых $N_i = N\,p_i$ обладают номером i, равно

$$M = \frac{N!}{\prod_{i=1,k} N_i!}.$$

Тогда энтропия этого случайного множества

$$H^* = \log M = \log N! - \sum_{i=1,k} \log N_i!.$$

По формуле Стирлинга

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Полагая, что N достаточно велико, имеем

$$H^* = N \log N - \sum_{i=1,k} N_i \log N_i = \sum_{i=1,k} N_i \log N - \sum_{i=1,k} N_i \log N_i = - N \sum_{i=1,k} p_y(i) \log p_y(i).$$

В этом случае энтропия

$$- \sum_{i=1,k} p_y(i) \log p_y(i) = (1/N) \log M$$

есть средняя неопределенность, которая приходится на одно единичное означивание при их случайном порождении.

Последние интерпретации логической энтропии проясняют ее поведенческую природу. Под N понимается число проводящих путей бинарной программы из истока в сток. Тогда N_i есть число таких путей, проходящих через i-ый узел $\boldsymbol{y}$-сечения, p_i – доля этих путей. M есть число способов распределения вычислений по k промежуточным вычислителям, соответствующих узлам $\boldsymbol{y}$-сечения, при условии, что на i-ый вычислитель приходится p_i-ая доля всех вычислений. Но тогда $\log M / N$ есть средняя неопределенность того, на каком вычислителе будет обрабатываться отдельная последовательность аргументов.

Если $k = 1$, то неопределенность равна 0. Если k возрастает, то возрастает и величина $\max(-\sum_{i=1,k} p_y(i) \log p_y(i))$. То есть неопределенность отнесения того или иного единичного означивания к соответствующему классу растет.

Из последнего определения следует, что логическая энтропия представляет собой информацию, которую мы получаем об одном единичном означивании, при известном разложении функции по переменным ***y***. Единичные означивания, порожденные одним классом ***y***-эквивалентности на промежуточном этапе разложения по ***y*** не различимы с точностью до ***y***-эквивалентности. Они характеризуются одним номером класса. Следовательно, информация, которая приходится на одно означивание при известном разложении по переменным ***y*** касается именно принадлежности к классу эквивалентности.

Достаточно очевидна аналогия такого представления с представлением энтропии в Теории информации как меры информации, которая передается одним символом сообщения,. В нашем случае каждое единичное означивание является носителем сигнала, который указывает на принадлежность соответствующего ***y***-набора конкретному классу эквивалентности.

Пример* 3**. Опишем два класса логических функций. Первый называется локальным и характеризуется энтропией, не зависящей от переменных ***y, определяющих сечение бинарной программы. Второй обладает энтропией, линейно зависящей от мощности множества ***y*** переменных не зависимо от порядка их означивания.

Первый класс функций описан в [2]. Для них доказана ограниченность сечения бинарных программ при некотором порядке означивания переменных. Содержательно это значит, что между переменными функций из этого класса имеется большое число зависимостей или, что эквивалентно, такие функции обладают небольшим числом подфункций.

С другой стороны в [2] описан класс не локальных функций, для бинарных программ которых любое ***y***-сечение (когда число переменных ***y*** не превосходит половины от числа всех аргументов) содержит не менее $2^{c|y|}$ узлов, где c – положительная константа. При этом доли путей, ведущих в

разные узлы $\boldsymbol{y}$-сечения, совпадают. Отсюда следует, что логическая энтропия, определяемая такими множествами аргументов, пропорциональна $|\boldsymbol{y}|$. Содержательно это обозначает, что для таких функций, их аргументы по большей части попарно независимы или, что эквивалентно, они обладают большим числом подфункций.

Если говорить о том, какова логическая энтропия большинства логических функций, то отметим, что для почти всех логических функций сложность реализации контактными схемами ограничена снизу экспонентой от числа переменных. Следовательно, почти все логические функции при любом построении бинарных программ обладают максимальной энтропией, сравнимой с общим числом их переменных.

Литература

1. Брошкова Н.Л., Попов С.В., О проектировании информационных систем. Препринт ИПМ РАН им. М.В.Келдыша, 2005.

2. Брошкова Н.Л., Попов С.В., О локальности информационных систем. Препринт ИПМ РАН им. М.В. Келдыша, 2005.

3. Шеннон К. Работы по теории информации и кибернетике, М.: ИЛ. 1963, - 830 с.

4. Колмогоров А.Н. Теория информации и теория алгоритмов, М.: Наука, 1987, - 303 с.

5. Файнстейн А. Основы теории информации. М.: ИЛ, 1960. – 140 с.