Карантинный курс ОДМ.

И.В.Артамкин

1 Лекция 2.

Мы переходим к теме "булевы функции", которая прекрасно изложена все в той же в книжке Ю.И.Худак, А.И.Сирота "Основы дискретной математики". Разберите, пожалуйста, разделы 2.1 — 2.5 главы 2. (Обратите внимание на разницу в обозначениях: в книге для конъюнкции используется знак &, а здесь $\wedge$.)

Там обсуждается постановка задачи нахождения минимальной ДНФ для заданной булевой функции. (Определения всех понятий см в книге.) Ниже я попробую чуть подробнее пояснить, откуда взялась такая постановка задачи.

Одна из самых типичных задач прикладной математики выглядит так: дана функция, заданная, как правило, массивом (таблицей) ее значений, и надо найти хорошую формулу, которая эту функцию задает. Для вещественных функций эта задача, как правило, ставится и решается приближенно, а для булевых функций, которыми мы здесь занимаемся, нас интересует точное представление.

Мы с вами на последней очной лекции видели исчерпывающее решение этой задачи в случае, когда мы хотим выразить нашу функцию многочленом. Я называю это решение "исчерпывающим", поскольку получился идеально точный результат: каждая булева функция однозначно задается многочленом, в котором каждая переменная входит в степени не выше первой (это и называется "многочлен Жегалкина"), так что имеется биекция между булевыми функциями и такими многочленами. (Конечно, еще остается чисто технологический вопрос: как по данной функции найти задающий ее многочлен. Один простой способ мы уже упоминали, это метод неопределенных коэффициентов, его описание еще можно прочитать в самом начале п. 3.2, стр 92-93.)

Аналогичный вопрос возникает, если мы изменим постановку задачи: теперь мы хотим выразить булеву функцию не многочленом а формулой, использующей только операции дизъюнкции, конъюнкции (=умножения) и отрицания. Собственно, задача выражения функции многочленом тоже могла быть сформулирована ровно так же: выразить функцию при помощи операций сложения, умножения и констант (ну, констант у нас всего две: 0 и 1). Но мы сразу говорили о многочленах благодаря тому, что в школе научились преобразовывать алгебраические выражения: мы хорошо знаем, что если в любом выражении, в которое входят имена переменных, константы, знаки сложения и умножения (и скобки!), раскрыть по дистрибутивности (= "распределительный закон умножения") все скобки, а потом привести подобные члены, то останется линейная комбинация мономов (= одночленов), то есть многочлен. Поэтому мы сразу формулировали эту задачу "по-взрослому": выразить функцию многочленом.

Если действовать по аналогии с этим подходом в задаче выражения функции через булевы операции (т.е. дизъюнкции, конъюнкции (=умножения) и отрицания), то первый вопрос, который естественно задать, это вопрос о том, к какому наиболее простому и стандартному виду можно тождественными преобразованиями привести произвольную формулу, т.е. существует ли в этой теории какой-то аналог многочленов. Но перед тем, как начинать отвечать на этот вопрос, полезно будет разобраться еще в том, какой смысл мы вкладываем в слово "формула". Каждый понимает, что, скажем,

$$((a + xyz) + abc)(u + v) + a(z + u) \tag{1.1}$$

является алгебраической формулой, а, например,

$$((axy+)a(zb + +ac())u + v) + u \tag{1.2}$$

— не является. Точно так же с формулами булевой алгебры:

$$(\overline{(a \vee (x \wedge (\bar{y} \wedge z)))} \vee ((\bar{a} \wedge y) \wedge (a \vee y))) \vee (\overline{a \wedge (\overline{z \vee u}})) \tag{1.3}$$

является формулой алгебры, а, например,

$$\overline{a \vee \wedge} x(a\vee) \tag{1.4}$$

не является.

Эти утверждения можно проверить "на глаз", хотя в случае (1.3) это уже довольно кропотливая работа. Но, конечно, на самом деле нам требуется математическое определение правильной формулы, без которого, например, невозможно запрограммировать машинное определение того, является ли данный текст правильной формулой или нет. Такое определение легко дать индуктивно.

Определение 1.1. *Алгебраической формулой называется выражение, содержащее имена переменных и констант, знаки $+$ и $\cdot$ и скобки, полученные последовательным применением конечного числа следующих правил. Имена переменных и константы являются формулами. Если A и B — формулы, то $(A)\cdot(B)$ и $(A)+(B)$ также являются формулами. (При этом еще удобно договориться, что мы опускаем скобки вокруг имен переменных и констант, так что если, например, A это имя переменной x, то мы пишем $x+(B)$ вместо $(x)+(B)$ и т.д..)*

Обратите внимание, что если мы применим это определение, то окажется, что (1.1) не является формулой. Дело в том, что формулы, полученные строгим применением этого определения, очень трудно читать из-за обилия скобок, поэтому используются некоторые стандартные договоренности, позволяющие опускать некоторые пары скобок. Эти договоренности вы проходили в школе. Во-первых, если есть три формулы A, B и C, то вместо любой из двух равносильных формул $((A)+(B))+(C)$ или $(A)+((B)+(C))$ мы благодаря ассоциативности пишем просто $(A)+(B)+(C)$, и аналогично с умножением. Во-вторых, мы раз и навсегда договорились, что все умножения выполняются раньше всех сложений (если иное не следует из расстановки скобок), поэтому мы пишем $(A)\cdot(B)+(C)$ вместо $((A)\cdot(B))+(C)$, и вообще опускаем все те скобки, без которых еще есть возможность, следуя предыдущим правилам, однозначно восстановить порядок действий. (Например, в выражении $a\cdot(b+c)$ скобки опустить нельзя, потому что тогда получится совсем другое выражение $a\cdot b+c$, а в выражении $z+(x\cdot y)$ скобки не нужны, поскольку задаваемый этими скобками порядок действий восстанавливается согласно нашим договоренностям и из выражения без скобок $z+x\cdot y$.) Наконец, мы договорились опускать знак умножения — точку. Все эти договоренности мы усвоили еще в школе[1], и благодаря

[1] Мы для краткости опускаем здесь стандартные школьные сокращения для степеней $\underbrace{x\cdot x\cdot\ldots\cdot x}_{n\ \text{раз}}=x^n$, поскольку для многочленов Жегалкина они не требуются.

им алгебраические формулы относительно легко читать.

Ровно тот же подход применим и при работе с формулами булевой алгебры. Сначала дадим аналог определения 1.1.

Определение 1.2. *Формулой булевой алгебры называется выражение, содержащее имена переменных и констант (0 и 1), а также знаки дизъюнкции, конъюнкции и отрицания и скобки, полученные последовательным применением конечное число раз следующих правил. Имена переменных и константы являются формулами. Если A и B — формулы, то $(A)\vee(B)$ и $(A)\wedge(B)$ и $\overline{(A)}$ также являются формулами. (При этом еще удобно договориться, что мы опускаем скобки вокруг имен переменных и констант, так что если, например, A это имя переменной x, то мы пишем $x\vee(B)$ вместо $(x)\vee(B)$, $x\wedge(B)$ вместо $(x)\wedge(B)$ и $\bar{x}$ вместо $\overline{(x)}$ и т.д..)*

Можно проверить, что согласно этому определению (1.3) является формулой, хотя из-за обилия скобок это довольно кропотливая работа. Для сокращения этой работы можно попробовать принять договоренности, аналогичные использованным для алгебраических выражений. Во-первых, поскольку операции дизъюнкции и конъюнкции также ассоциативны, можно договориться вместо любой из двух равносильных формул $((A)\vee(B))\vee(C)$ или $(A)\vee((B)\vee(C))$ писать просто $(A)\vee(B)\vee(C)$, и аналогично для конъюнкции $\wedge$. Уже эта договоренность несколько упростит формулу (1.3):

$$\overline{(a\vee(x\wedge\bar{y}\wedge z))}\vee(\bar{a}\wedge y\wedge(a\vee y))\vee(\overline{a\wedge(\overline{z\vee u})}) \tag{1.5}$$

Далее у нас есть две возможности договориться о том, какая операция по умолчанию выполняется первой: конъюнкция или дизъюнкция. При работе с алгебраическими формулами такого вопроса не возникает, поскольку удобство общепринятого порядка действий определяется тем, что в алгебре имеется только одно правило раскрытия скобок (дистрибутивность): $x\cdot(y+z)=x\cdot y+x\cdot z$. В случае же булевой алгебры обе операции совершенно равноправны: у нас есть ДВЕ дистрибутивности $x\wedge(y\vee z)=(x\wedge y)\vee(x\wedge z)$ и $x\vee(y\wedge z)=(x\vee y)\wedge(x\vee z)$. (Пока у нас нет договоренности, какая операция выполняется первой, мы не можем в этих формулах опустить никакие скобки!)

Таким образом, оба варианта договоренности о выборе порядка действий возможны. Мы сейчас выберем один вариант, а через некоторое время разберемся, что получилось бы, если бы мы сделали противоположный выбор.

Итак, договариваемся, что, если расстановка скобок не диктует иного, то конъюнкция всегда выполняется раньше дизъюнкции. Этот выбор еще удобен и потому, что операция конъюнкции это, по существу, операция умножения, и тогда мы можем заменить знак $\wedge$ на знак умножения $\cdot$, а затем, по аналогии с алгебраическими договоренностями, вовсе его опускать. В соответствии с такими договоренностями формула (1.5): станет еще более компактной и обозримой:

$$\overline{(a \vee x\bar{y}z)} \vee \bar{a}y(a \vee y) \vee (\overline{a(\overline{z \vee u})}) \tag{1.6}$$

Далее мы все время будем пользоваться именно такой записью формул булевой алгебры. Теперь мы переходим к следующей задаче: к какому наиболее простому и стандартному виду можно равносильными преобразованиями привести произвольную формулу булевой алгебры? Напомним, что в случае алгебраических формул ответ на аналогичный вопрос нам известен со школы: в любом алгебраическом выражении мы сначала раскрываем все скобки, пользуясь дистрибутивностью $x(y+z) = xy+xz$, а затем приводим подобные члены и получаем многочлен от перемееных, входящих в нашу формулу. Нечто аналогичное (но не совсем) можно сделать и для формул булевой алгебры. Эта процедура кратко описана в начале раздела 2.4 (шаги 1-4); мы сейчас опишем ее чуть подробнее.

Сначала надо расправиться со всеми вхождениями знака отрицания: несколько раз пользуясь формулами де Моргана $\overline{xy} = \bar{x} \vee \bar{y}$ и $\overline{x \vee y} = \bar{x}\bar{y}$ мы можем добиться того, что знаки отрицания будут стоять только над именами переменных. Вот как это выглядело бы в случае (1.5): на первом шаге получаем

$$\bar{a}(\overline{x\bar{y}z}) \vee \bar{a}y(a \vee y) \vee (\bar{a} \vee (z \vee u)) \tag{1.7}$$

(тут мы еще воспользовались формулой двойного отрицания $\bar{\bar{x}} = x$), замечаем еще, что теперь еще две пары скобок оказываются лишними, и еще раз применяя формулу де Моргана, получаем

$$\bar{a}(\bar{x} \vee y \vee \bar{z}) \vee \bar{a}y(a \vee y) \vee \bar{a} \vee z \vee u \tag{1.8}$$

Следующий шаг ровно такой, как и в случае алгебраических формул: последовательно раскрываем все скобки, пользуясь ОДНОЙ дистрибутивностью, а именно $x(y \vee z) = xy \vee xz$. В результате получим выражение совсем без скобок, состоящее из нескольких мономов, соединенных знаками дизъюнкции, где под мономами мы понимаем произведение некоторого количества имен переменных и имен переменных с отрицаниями. В случае нашего примера получится так:

$$\bar{a}\bar{x} \vee \bar{a}y \vee \bar{a}\bar{z} \vee \bar{a}ya \vee \bar{a}yy \vee \bar{a} \vee z \vee u \tag{1.9}$$

Логика алгебраического случая подсказывает нам, что дальше надо привести подобные члены, но перед этим стоит разобраться, какие мономы у нас могут быть. Во-первых, понятно, что если какая-то переменная входит в произведение два раза, то выражение можно упростить по формуле $xx = x$. Далее, если в какой-то моном входит и переменная и ее отрицание, то, поскольку $x\bar{x} = 0$, этот моном является на самом деле нулевым. Поскольку $x \vee 0 = x$, это означает, что данный моном можно просто отбросить. Тем самым получается, что в моном каждая переменная может входить не более одного раза — либо с отрицанием, либо без. Следовательно, каждый моном является *элементарной конъюнкцией* в смысле определения 60 главы 2. В нашем примере в (1.9) имеется один моном $\bar{a}ya$, в который входят и переменная a и ее отрицание, так что его можно опустить, и один моном $\bar{a}yy$, который упрощается до $\bar{a}y$, так что получается

$$\bar{a}\bar{x} \vee \bar{a}y \vee \bar{a}\bar{z} \vee \bar{a}y \vee \bar{a} \vee z \vee u \tag{1.10}$$

Остается лишь приведение подобных членов, которое в нашем случае сводится к применению формулы $a \vee a = a$ (в нашем случае (1.10) дублируются два монома), и полученное выражение будет как раз дизъюнктивной нормальной формой (ДНФ) в смысле определения 61. В нашем случае получилась такая ДНФ:

$$\bar{a}\bar{x} \vee \bar{a}y \vee \bar{a}\bar{z} \vee \bar{a} \vee z \vee u \tag{1.11}$$

Таким образом, ДНФ для формул булевой алгебры является до некоторой степени аналогом многочленов для формул обычной алгебры: это стандартный вид, к которому можно привести любую формулу. Но есть очень большая разница: как мы видели, каждую булеву функцию можно однозначно задать простейшим многочленом (многочленом Жегалкина), так что имеется биекция между булевыми функциями и задающими их многочленами Жегалкина. А с ДНФ ситуация совсем другая: лемма 6 на стр 58 показывает, что различных ДНФ на много порядков больше, чем булевых функций (от того же числа переменных), поэтому ни о какой однозначности представления речь не идет — надо искать наиболее удобные (экономные) ДНФ для данной функции. Об этом — следующая лекция.

На первый взгляд, правда, может показаться, что проблема однозначности представления легко решается, если мы решим ограничиться совершенными ДНФ (п. 2.3). Действительно, их ровно столько, сколько булевых функций, и представление в виде СДНФ однозначно — чего же еще хотеть? Но все портит одна неприятность: СДНФ это, в некотором смысле, самая громоздкая из всех ДНФ, задающих данную функцию, так что если мы озабочены оптимизацией вычислений, то она нам точно не годится.