Международный конкурс «Математика и проектирование»

Числа Каталана

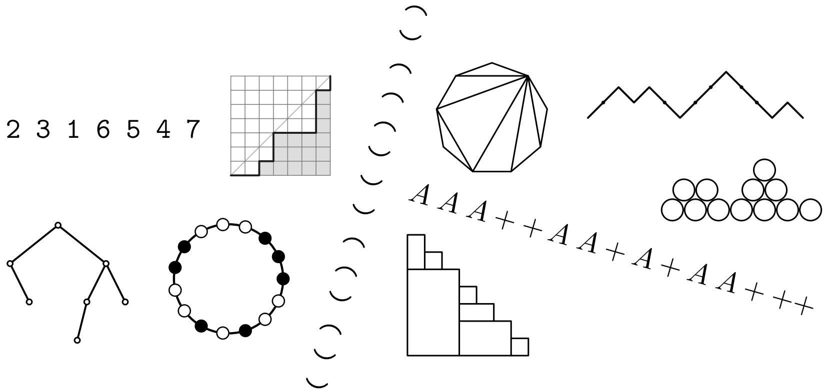
Алекс Иванов Цветанов,
Маргарита Велчева Стефанова,
Йоана Петрова Кичева

преподаватель: Ваня Данова

Софийская математическая гимназия — София, Болгария

1-е мая 2016

Что имеется общего между следующими объектами?



Этому вопросу посвящена данная работа.

Введение

В 1838 г. математик Э. Ш. Каталан исследовал вопрос о числе способов расстановки скобок в выражениях. Эти числа, для всевозможных длин n выражения, сегодня называются числами Каталана.

Впоследствии оказалось, что те же самые числа перечисляют и самые разные совокупности объектов, внешне не схожие между собой и заведомо отличающиеся от выражений.

Назовем «каталановой» любую совокупность объектов, перечисляемую числами Каталана.

Покажем некоторые из самых интересных таких совокупностей. Для каждой совокупности свойство быть «каталановой» устанавливаем указанием взаимно-однозначного соответствия (биекцию) между ней и уже известной такой совокупностью.

Обособление подвыражений

Рассмотрим задачу, схожую с рассмотренной Каталаном.

Пусть дано выражение из n операций сложения $n+1$ чисел. Сколькими способами в нем возможно обособить подвыражения?

Например, для $n=7$ и выражения

$$A+A+A+A+A+A+A+A$$

одно из возможных разбиений, указывая подвыражения подчеркиванием:

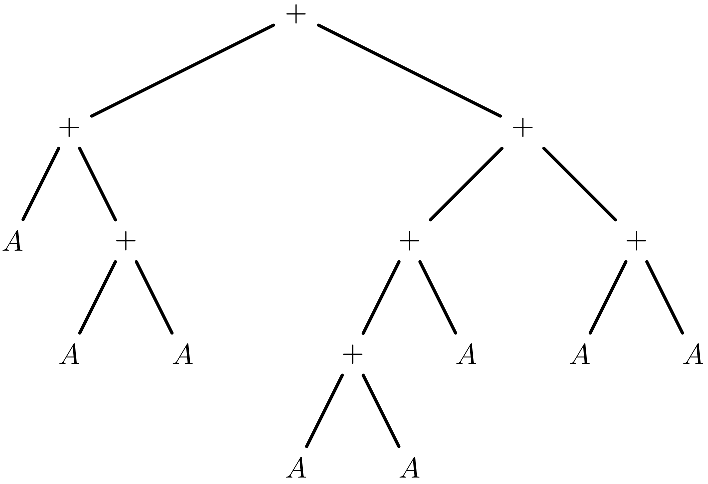
$$\underline{A+\underline{A+A}}+\underline{\underline{\underline{A+A}+A}+\underline{A+A}}$$

или, то же самое, скобками:

$$(A+(A+A))+(((A+A)+A)+(A+A)).$$

Разбиение как дерево

Каждому разбиению выражения биективно соответствует двоичное дерево, отражающее иерархию подвыражений. Для нашего примера:



Суфиксная запись и последовательности из 1 и –1

Суфиксный обход дерева (левое — правое — корень) дает суфиксную запись выражения, биективно соответствующую дереву и таким образом — данному разбиению.
Для нашего примера:

$$A\;A\;A + + A\;A + A + A\;A + + + \,.$$

Поставив 1 и –1 соответственно на местах A и $+$ в суфиксной записи, получим последовательность из $n{+}1$ единиц и n минус единиц — для нашего примера

$$1\;\;1\;\;1\;{-1}\;{-1}\;\;1\;\;1\;\;{-1}\;\;1\;{-1}\;\;1\;\;1\;\;{-1}\;{-1}\;{-1}\,,$$

у которой все частичные суммы > 0, а сумма равна 1. И наоборот, любой такой последовательности соответствует суфиксная запись некоторого разбиения исходного выражения.

Правильные последовательности из 1 и –1

Назовем «правильными порядка n» последовательности из $n+1$ единиц и n минус единиц с положительными частичными суммами.

Если найти число правильных последовательностей, найдем и число разбиений выражений с n сложениями.

Нетрудно доказать, что *любая* последовательность из $n+1$ единиц и n минус единиц можно, притом единственным образом, циклически сдвинуть так, что все частичные суммы у новой последовательности были >0 (сумма всех членов, конечно, всегда 1).

Значение n-го числа Каталана

Число всех последовательностей из $n+1$ единиц и n минус единиц — $\binom{2n+1}{n}$, поскольку для единиц можно выбрать любые n из всех $2n+1$ возможных мест. А поскольку, в силу вышесказанного, из всех последовательностей, приводимых одна в другую циклическим сдвигом, только одна правильная, для получения числа правильных последовательностей разделим $\binom{2n+1}{n}$ на число $2n+1$ возможных сдвигов.

Число разбиений выражения, или число C_n Каталана для $n>0$ тогда равно

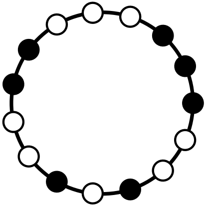
$$C_n = \frac{1}{2n+1}\binom{2n+1}{n} = \frac{(2n)!}{n!(n+1)!} = \frac{1}{n+1}\binom{2n}{n}.$$

Положим еще $C_0 = 1$.

Итак, установлено, что множество правильных строк из 1 и -1 и множество разбиений выражения — каталановы.

Из сказанного следует и что множество ожерелий из n белых и $n+1$ черных бусин тоже является каталановым (сочтем все строки вида $((n+1) \times 1, n \times -1)$ циклическими, а числа в них заменим на бусины).

Рассмотренному выше примеру соответствует ожерелье



и всего таких ожерелий $C_7 = 429$.

Пары скобок

Если из любой правильной последовательности 1 и –1 убрать начальную единицу и все оставшиеся члены 1 заменить на (, а –1 на), получится строка из n пар правильно сгрупированных (сбалансированных) скобок. Для рассмотренного выше примера

$$1\ 1\ 1\ {-1}\ {-1}\ 1\ 1\ {-1}\ 1\ {-1}\ 1\ 1\ {-1}\ {-1}\ {-1}$$

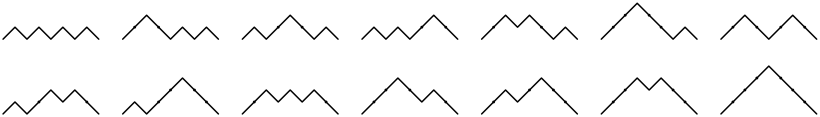
получаем

$$(\ (\ )\ )\ (\ (\ )\ (\ )\ (\ (\ )\ )\ )\ .$$

Обратным действием любая сбалансированная строка из n пар скобок превращается в правильную последовательность 1 и –1.

Найденное взаимно-однозначное соответствие показывает, что сбалансированные строки скобок являются каталановым множеством.

Горные хребты

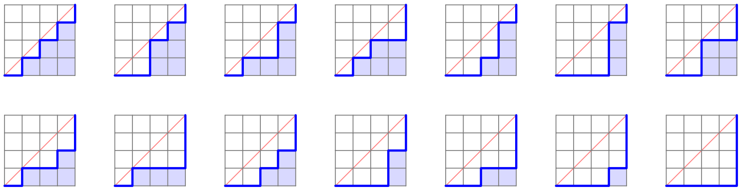


Заменив в любой сбалансированной строке (и) на наклонные линии / и \ и связав эти линии, получается фигура, известная как «горный хребет». Характеристическое свойство такой фигуры: она не пересекает горизонталь.

На рисунке показаны все горные хребты для $n = 4$.

Очевидно, что соответствие между строками скобок и хребтами взаимно-однозначно и значит множество горных хребтов — каталаново.

Поддиагональные маршруты и полиомино



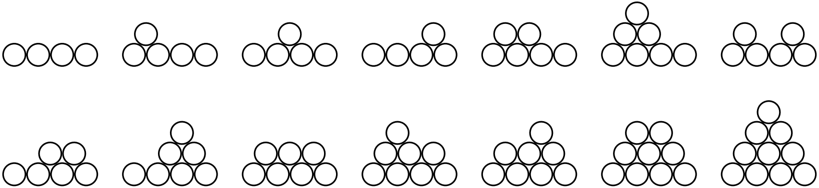
«Хребет» можно построить и на диагонали в клетчатой сетке, например снизу: вместо / и \ проводим _ и |. Получаются маршруты, непересекающие диагональ.

Каждый маршрут является и границей поддиагонального полиомино, занимающего правую нижнюю клетку или пустого.

На рисунке показаны все такие маршруты для $n = 4$.

Ясно, что совокупность поддиагональных маршрутов, также как и совокупность полиомино, являются каталановыми.

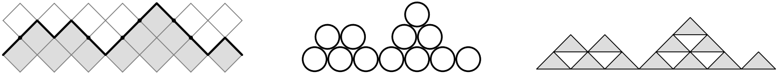
Стопки бревен (или монет)



Сколько различных стопок бревен можно возвести на основе из n плотно расположенных бревен?

Построив «горный хребет» в клетчатой сетке, как показано ниже, понятно, что «стопка бревен» — не что иное, как множество клеток под хребтом. Поэтому множество стопок — каталаново.

Клетки же, или «бревна», можно считать и треугольными.



Стековые перестановки

Перестановку некоторой последовательности $a_1, a_2, \ldots, a_n$ с участием стека производят следующим образом. Начиная с пустого стека и пустого результата, на каждом шагу делают одно из двух, пока возможно:

а. либо берут из очереди очередное a_i и кладут в стек,

б. либо берут из стека элемент и добавляют к результату.

Например, из последовательности `1 2 3 4 5 6 7` можно получить `2 3 1 6 5 4 7`, применив последовательность действий **а а б а б б а а а б б б а б**.

Заменив в последовательности действий для перестановки буквы **а** и **б** на `(` и `)`, получим сбалансированную строку из n пар скобок, а любой сбалансированной строке соответствует стековая перестановка — имеет место биекция между этими двумя множествами объектов.

Итак, стековые перестановки есть каталаново множество.

Упорядоченные леса и деревья: определения

Корневое дерево (КД) — это пара (вершина, лес). Точнее, это непустое корневое дерево. Имеется ровно одно пустое КД.

Лес — множество, возможно пустое, корневых деревьев.

Когда множество в определении леса упорядочено, лес и соответственно деревья являются упорядоченными. Именно такие деревья и леса рассматриваются дальше.

Согласно определениям КД и леса, каждому лесу с $n \geq 0$ вершинами однозначно соответствует КД с $n+1$ вершинами, и наоборот — каждому КД с $n \geq 1$ вершинами однозначно соответствует лес с $n-1$ вершинами. Таким образом, для каждого $n \geq 0$ имеется биективное соответствие между множествами $n+1$-вершинных корневых деревьев и n-вершинных лесов.

Леса и строки скобок

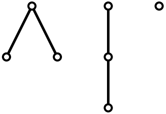
Сбалансированной строке из n пар скобок поставим в соответствие лес по следующему правилу:

> *соседним парам скобок соответствуют соседние (под)деревья, а вложенным скобкам — деревья, наследующие одно другое.*

Например строке

$$(\; (\;) \; (\;) \;) \; (\; (\; (\;) \;) \;) \; (\;)$$

соответствует лес



.

Соответствие биективно: любому лесу из n вершин сопоставляется единственная строка из n сбалансированных пар скобок.

Поэтому лес является каталановым множеством.

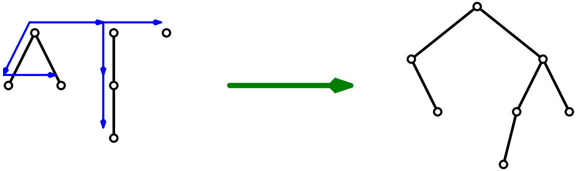
Леса и двоичные деревья

Двоичное дерево (ДД) — это дерево, которое либо пусто, либо состоит из корня и двух наследников, отличаемых как левый и правый, которые сами есть ДД.

Любому лесу сопоставим двоичное дерево согласно правилу:

- *корням деревьев леса соответствуют вершины, справа наследующие одна другую по порядку следования; также и с наследниками любой вершины;*
- *первому из наследников любой вершины соответствует левый наследник данной вершины.*

Пример получения ДД из леса:

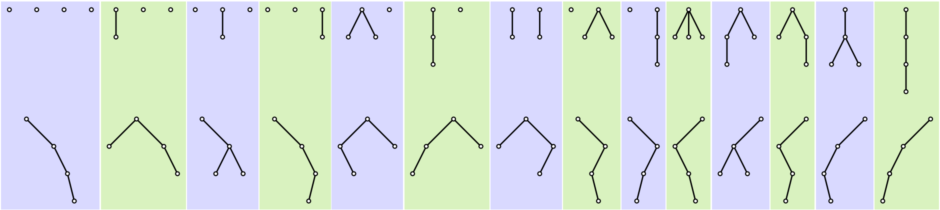


Леса и двоичные деревья — продолжение

Соответствие имеет силу и в обратном направлении: любому ДД единственным образом сопоставляется некоторый лес.

Этим установлено, что и множество двоичных деревьев является каталановым.

Сверху на рисунке показаны все леса порядка 4, а внизу — соответствующие им двоичные деревья с 4 вершинами.



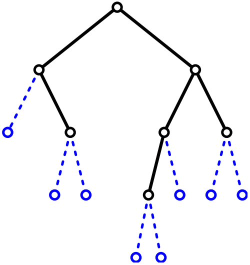
Строго двоичные деревья

Строго двоичное дерево (СДД) — двоичное дерево, каждая вершина которого имеет либо 0, либо 2 наследника.

Нетрудно сообразить, что у каждого СДД имеется листьев на один больше, чем внутренних вершин.

Деревья арифметических выражений с двуместными операциями, разбитых на подвыражения, которые рассматривали вначале — именно СДД.

Из непустого двоичного дерева можно получить СДД, добавив 1 или 2 наследника каждой вершине, у которой они 1 или 0. Обратное преобразование — из СДД удалить все листья и соответствующие им ребра. Этими двумя преобразованиями устанавливается биекция между двоичными деревьями с $n \geq 1$ вершинами и СДД с $n+1$ листьями (и n внутренними вершинами).



Итог «каталановости» деревьев и лесов

В силу сказанного, для каждого $n \geq 0$, число

- n-вершинных лесов,
- $n+1$-вершинных корневых деревьев,
- n-вершинных двоичных деревьев,
- строго двоичных деревьев с $n+1$ листьями (или с n внутренними вершинами)

одно и то же и является числом Каталана C_n.

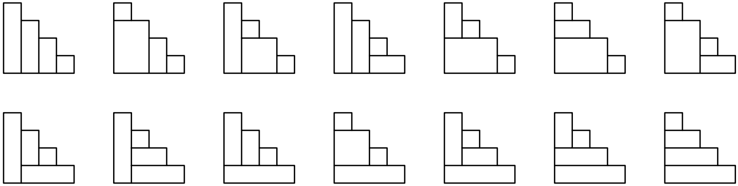
Разрезание лестницы на прямоугольники

Если из фигуры типа «лестница» вырезать прямоугольник с одним углом — угол лестницы, а другой — одна из ступеней, останется одна или две лестницы. Если их разрезать дальше, всего получится n прямоугольников, по числу ступеней.

Каждому такому разрезанию соответствует двоичное дерево с корнем данный прямоугольник и наследниками — лестницы-соседи с обеих сторон прямоугольника. Также и каждому двоичному дереву соответствует разрезание.

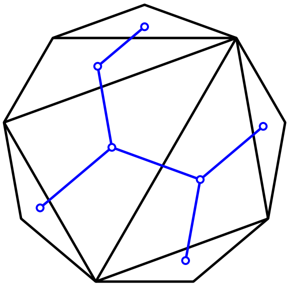
Из этого следует, что разрезания есть каталаново множество.

На рисунке — все разрезания лестницы с 4 ступенями.



Триангуляция выпуклого многоугольника

Для любой триангуляции выпуклого n-угольника построим двойственный граф — с вершинами для треугольников и ребрами для соседства треугольников. У этого графа $n-2$ вершин и $n-3$ ребер и если его «подвесить» за любую вершину, получается двоичное дерево. (Можно наперед пометить какую-нибудь сторону многоугольника, чтобы выбрать примыкающий к ней треугольник, какой бы он не оказался для конкретной триангуляции.)

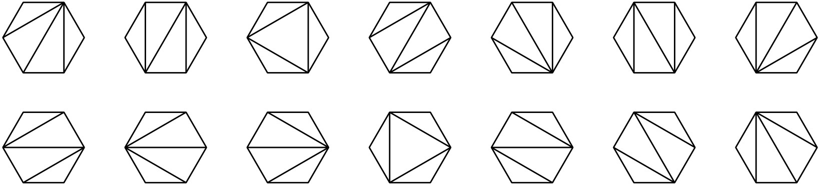


Вместе с тем, любому $n-2$-вершинному двоичному дереву соответствует ровно одна такая триангуляция.

В силу этого соответствия, множество триангуляций является каталановым.

Триангуляции шестиугольника

На рисунке — все триангуляции выпуклого шестиугольника.
(Их число — $C_{6-2} = C_4 = 14$.)



Разрезание окружности хордами

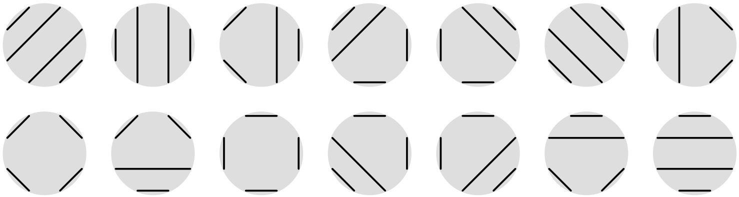
$2n$ точек на окружности соединены парами без пересечений.

Пометим одну из точек. Хорда, на которой точка лежит, объявим корнем двоичного дерева. Поддеревьями будут выпуклые фигуры по обеим сторонам от хорды — пустые либо разрезанные хордами. При наличии хорд одна из них выбирается вершиной поддерева и идет дальнейшее ветвление. В конце процесса получим n-вершинное двоичное дерево.

Обратив процесс, из каждого n-вершинного ДД получаем разрезание окружности n хордами без пересечений.

В силу биекции между разрезаниями и двоичными деревьями, первое множество является каталановым.

На рисунке — все разрезания с 4 хордами.



Треугольник Каталана

Последовательностей бесконечно много и каждая из них бесконечна. Каждая последовательность начинается вторым членом предыдущей, а каждое число является суммой верхнего и левого от него.

1	1	1	1	1	1	1	1	...
	1	2	3	4	5	6	7	...
		2	5	9	14	20	27	...
			5	14	28	48	75	...
				14	42	90	165	...
					42	132	297	...
						132	429	...
							429	...
								...

Начальные элементы строк треугольника есть не что иное, как C_n для $n = 0, 1, 2, \ldots$. Вообще, j-й элемент в строке i для $i, j = 0, 1, 2, \ldots$ равен числу последовательностей из i членов -1 и $i+j$ членов 1 с неотрицательными частичными суммами.

Программное определение треугольника и чисел C_n

Простота образования треугольника Каталана дает удобный способ запрограммировать порождение его содержания.

В первом из следующих определений на языке Haskell порождается последовательность последовательностей, т. е. сам треугольник. Второе — последовательность чисел Каталана.

```
catrows = iterate (scanl1 (+) . tail) [1,1..]
cats = map head catrows
```

При помощи данных определений можно проводить например такие вычисления:

```
cats !! i                -- i-е число Каталана
take n cats              -- n первых чисел Каталана
catrows !! i !! j        -- j-й член строки i
take n (catrows !! i) -- n первых чисел строки i
```

Задачи

Доказать следующие рекуррентные равенства для $n \geq 0$:

- $C_{n+1} = \frac{4n+2}{n+2}\, C_n$;
- $C_{n+1} = C_0 C_n + C_1 C_{n-1} + \cdots + C_n C_0$.

Доказать «каталановость» следующих множеств:

- последовательности натуральных чисел $1 \leq a_1 \leq a_2 \leq \ldots \leq a_n$, для которых $a_i \leq i$;
- последовательности целых чисел $a_1, \ldots, a_n$, такие что a_i равно количеству j, для которых $j < i$ и $a_j \leq a_i$;
- перестановки чисел $1, 2, \ldots, n$, у которых длины всех убывающих подпоследовательностей не больше 2.

Доказать указанное выше свойство чисел треугольника Каталана, вкл. то, что среди них — числа Каталана.

Программное порождение каталановых объектов

В целях изучения каталановых множеств нами построено несколько компьютерных программ.

Одна из них порождает все правильные скобочные строки (СС), притом с минимальной разницей: каждая строка отличается от предыдущей обменом скобок только в двух местах. Другая программа переводит скобочную строку в соответствующее двоичное дерево (ДД).

Дальше, имеются программы для переводов:

— из СС в пермутацию со стеком;
— из СС в «стопку бревен»;
— из ДД в разрезание лестницы;
— из ДД в триангуляцию многоугольника;
— и др.

Отдельный набор программ обеспечивает графическое представление деревьев, маршрутов, ожерелий, триангуляций и пр. объектов.

Заключительные замечания

Числа Каталана являются предметом неослабевающего интереса в математике, в том числе в школе. Об этом свидетельствует обилие литературы по теме — только в последние годы вышло несколько монографий — а также то, что статья о числах Каталана считается длиннейшей в энциклопедии OEIS (`oeis.org`).

Исключительно разнообразие объектов, перечисление которых приводит к числам Каталана.

Во многих случаях изучение чисел Каталана и связанных с ними объектов не требует специфических математических знаний, а поэтому оно может проводиться и в школе.

Удачно «каталановость» множеств устанавливать путем нахождения взаимно-однозначного соответствия с уже знакомыми множествами, как мы и делали. Это более полезно по сравнению со счетными методами ввиду того, что позволяет не только найти размер множества, но и получить представление о его строении. Метод приведения также развивает способность к абстрактному мышлению: для изучения свойств одних объектов их заменяют другими.

Заключительные замечания (продолжение)

Приведение одного объекта к другому часто имеет алгоритмический характер, с условиями и повторениями, что позволяет применять компьютерные программы для порождения и приведения самого разного вида объектов. Если, например, составлен алгоритм для последовательного порождения объектов данного множества, он может применяться и для всех других множеств, к которым данное можно привести.

В большинстве случаев множествами, к которым удобно привести данный вид объектов, являются выражения со скобками и деревья, т. е. общие представления иерархии. Тогда уместно ориентироваться на разработку алгоритмов порождения именно таких множеств.

Eugène Charles Catalan

Э. Каталан родился в г. Брюж (Бельгия), получил образование в Парижской политехнической школе и работал в области геометрии, теории чисел, цепных дробей и комбинаторики.

Рассмотренная им задача, в связи с которой появилось название «числа Каталана», была такой: сколькими способами можно расставить n пар скобок в цепочке из $n \geq 2$ букв так, чтобы внутри каждой пары было ровно два терма (буквы или выражения в скобках).

В такой постановке нам сразу заметно прямое соответствие с двоичными деревьями, но в первой половине 19 в. математическое понятие дерева было незнакомо.

До Каталана Л. Эйлер, решая задачу о триангуляции выпуклого многоугольника, фактически нашел последовательность чисел, названную именем Каталана.

Литература

- Conway, J. H., Guy R. K. *The book of numbers*. Copernicus, 1996, pp. 96-106.
- Grimaldi R. P. *Fibonacci and Catalan numbers: an introduction*, John Wiley & Sons, 2012.
- Koshy Th., *Catalan numbers with applications*, Oxford University Press, 2009.
- Roman S., *An introduction to Catalan numbers*, Birkhäuser, 2015.
- Stanley R.P., *Catalan numbers*, Cambridge University Press, 2015.
- *Catalan number*. http://en.wikipedia.org/wiki/Catalan_number.
- *Catalan's triangle*. http://en.wikipedia.org/wiki/Catalan%27s_triangle.
- Stanley, R. and Weisstein, E .W. *Catalan Number. From MathWorld – A Wolfram Web Resource*. http://mathworld.wolfram.com/CatalanNumber.html.
- Weisstein, E .W. Catalan's Triangle. *From MathWorld – A Wolfram Web Resource*. http://mathworld.wolfram.com/CatalansTriangle.html.
- Sloane N. J. A. *Числа Каталана (последовательность A000108 в OEIS)*. http://oeis.org/A000108.
- Sloane N. J. A. *Треугольник Каталана (последовательность A009766 в OEIS)*. http://oeis.org/A009766.
- *Числа Каталана на The Math Forum*. http://mathforum.org/advanced/robertd/catalan.html.
- *Числовые треугольники, связанные с разбиениями*. http://en.wikiversity.org/wiki/Partition_related_number_triangles.

Программные технологии

В работе использованы:

- язык и система LaTeX[1] с пакетом для изготовления презентаций BEAMER[2]
- программа для создания чертежей `sp`[3]
- языки программирования и описания рисунков JavaScript[4], SVG[5] и Haskell[6]

[1] http://ctan.org/pkg/latex

[2] http://ctan.org/pkg/beamer

[3] http://www.math.bas.bg/bantchev/sp

[4] http://ecma-international.org/publications/standards/Ecma-262.htm

[5] http://w3c.org/Graphics/SVG

[6] http://haskell.org

Благодарности

Выражаем благодарность

Ване Дановой,
Бойко Банчеву и
Василу Тинчеву

за оказанную помощь в работе по теме и изготовлении презентации.

Спасибо за внимание!