Заключительный (очный) тур

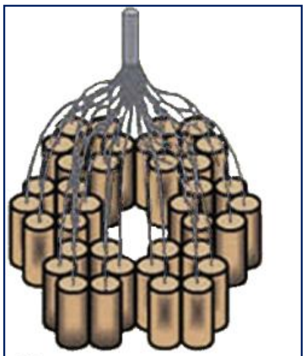
Математика

Простые задачи

1. Наноткач

(8 баллов)

Фермент целлюлозосинтаза производит одиночную целлюлозную нить, составленную из отдельных сахаридов. Ферментный комплекс (см. рисунок) путешествует по поверхности клетки, наматывая на нее производимую микрофибриллу целлюлозы как нитку на клубок. За какое время такой наноткач сделает целлюлозную оболочку массой **M** = $4.5 \cdot 10^{-5}$ нг вокруг клетки радиусом **R** = 8 мкм, если за одну секунду в микрофибриллу включается 500 сахаридов? (***2.5 балла***) Сколько оборотов вокруг клетки **k** он при этом совершит? (***5.5 балла***) Считать, что масса одного сахарида равна **m** = $2.7 \cdot 10^{-13}$ нг, длина – ***l*** = 0.5 нм, толщиной микрофибриллы пренебречь.



36 целлюлозосинтаз, объединенных в целлюлозо-синтезирующий комплекс

Ответ:

1) (<u>2.5 балла</u>) Чтобы получить оболочку массой **M**, всего необходимо синтезировать $N = \frac{M}{m} = \frac{M}{2,7 \cdot 10^{-13}}$ сахаридов.

На это понадобится $t = \frac{N}{v} = \frac{N}{500} = \frac{M}{500 \cdot 2,7 \cdot 10^{-13}} = \frac{M}{1,35 \cdot 10^{-10}}$ с.

2) (<u>5.5 балла</u>) Длина полученной микрофибриллы из 36 целлюлозных цепочек (<u>*число цепочек в микрофибрилле было дано на рисунке в условии задачи*</u>) при этом составит

$$L = \frac{N \cdot 0,5}{36} = \frac{0,5M}{36 \cdot 2,7 \cdot 10^{-13}} = \frac{M}{1,944 \cdot 10^{-11}} \text{ нм.}$$

Тогда число оборотов ферментного комплекса вокруг клетки равно

$$k = \frac{L}{2\pi R_{(\text{мкм})}} = \frac{M}{2\pi R_{(\text{нм})} \cdot 10^3 \cdot 1,944 \cdot 10^{-11}} = \frac{M}{R_{(\text{нм})} \cdot 1,2 \cdot 10^{-7}}.$$

Вариант	**R**, мкм	**M**, нг	**N**	**t**, с	**L**, нм	**k**
2	8	$4,5 \cdot 10^{-5}$	$1,67 \cdot 10^{8}$	$3,33 \cdot 10^{5}$	$2,3 \cdot 10^{6}$	46,67 (46 полных оборотов)

2. Вольвокс как кластер

(8 баллов)

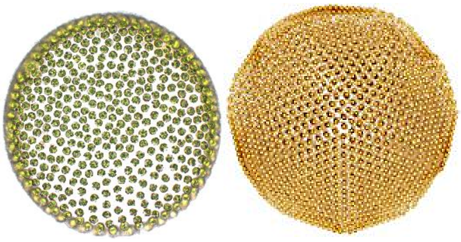
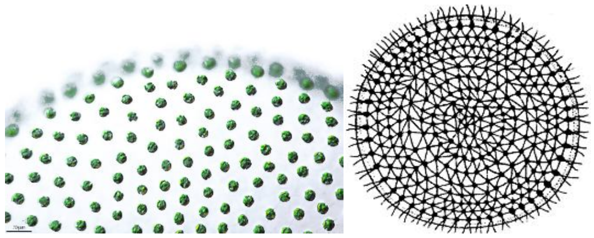
а)  б) 

Рис. а) Микрофотография вольвокса и модель поверхности нанокластера золота; б) поверхность вольвокса и ее модель.

Вольвокс представляет собой колонию одноклеточных зеленых водорослей, сгруппированных в сферу. В некотором вольвоксе каждая из клеток связана мостиком с ближайшими соседями (**x** клеток имеют 5 соседей, **y** – 6, **z** – 7). При этом формируется сеть из треугольных ячеек, покрывающая всю сферу. Найдите общее число клеток в вольвоксе **v**, а также **x** и **z**, если общее число мостиков **e** = 3660, а число клеток с шестью соседями **y** = 1050.

Ответ:

Вольвокс можно рассматривать как многогранник, имеющий $\mathbf{v} = \mathbf{x} + \mathbf{y} + \mathbf{z}$ (1) вершин, **e** ребер и **f** треугольных граней. Тогда запишем для него теорему Эйлера:

x + **y** + **z** – **e** + **f** = 2. (2)

В тоже время, исходя из свойств многогранников, мы можем записать, что:

$$e = \frac{5x}{2} + \frac{6y}{2} + \frac{7z}{2}$$ (каждое ребро принадлежит двум вершинам) (3)

и $f = \frac{5x}{3} + \frac{6y}{3} + \frac{7z}{3}$ (каждая грань имеет три вершины) (4)

Тогда, объединяя (2), (3) и (4), получаем:

$$x + y + z - \frac{5x}{2} - \frac{6y}{2} - \frac{7z}{2} + \frac{5x}{3} + \frac{6y}{3} + \frac{7z}{3} = 2. \quad (5)$$

Упрощая (5), приходим к $x = 12 + z$. (6)

Теперь подставим (6) в выражение (3):

$$e = \frac{5(12 + z)}{2} + \frac{6y}{2} + \frac{7z}{2} \quad (7)$$

$$2e = 60 + 5z + 6y + 7z \quad (8)$$

Подставляя (6) в (8), получаем $z = \frac{2e - 6y - 60}{12} = \frac{e - 3y - 30}{6}$ (9), $x = 12 + z = 7 + \frac{e - 3y}{6}$ (10) и общее число клеток в колонии $v = x + y + z = 7 + \frac{e - 3y - 30}{6} + y + \frac{e - 3y - 30}{6} = 2 + \frac{e}{3}$ (11).

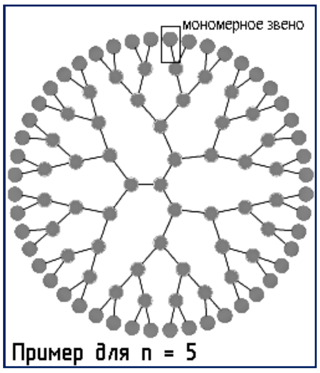
Вариант	**e**	**y**	**x**	**z**	**v**
2	3660	1050	<u>92</u>	<u>80</u>	<u>1222</u>

Примечание: приведен один из возможных вариантов решения.

Общий подход заключается в записи системы из трех уравнений ((2)-(4) или их линейных комбинаций) с тремя неизвестными и ее решении.

3. Наноконтейнер для лекарства

(8 баллов)



Молекулы дендримера – разветвленного на манер дерева полимера – имеют внутри достаточно свободного места, чтобы там, как в контейнере, поместились молекулы лекарства. Рассчитайте, во сколько раз масса переносимого лекарства может быть больше массы самого дендримера. Считать, что:

– дендример насчитывает **n** = 12 поколений ветвления, схема ветвления представлена на рисунке;

– радиус дендримера с каждым поколением увеличивается на 1 нм;

– масса одного мономерного звена $2 \cdot 10^{-22}$ г;

– плотность звеньев дендримера, молекул лекарства, а также наноконтейнера, полностью заполненного лекарством, равна 1 г/см^3.

Ответ:

1) Если лекарство занимает весь свободный объем в молекуле дендримера, значит, масса заполненного контейнера составляет

$$m_1 = V\rho = \frac{4}{3}\pi(1\cdot n)^3 \cdot 1\cdot 10^{-21} = \frac{4}{3}\pi n^3 \cdot 10^{-21} \approx 4{,}13 n^3 \cdot 10^{-21} \text{ г}$$

2) В свою очередь, масса «вещества» дендримера равна

$m_2 = b\cdot 10^{-22}\cdot S + m_3 \approx b\cdot 10^{-22}\cdot (S+1) \approx b\cdot 10^{-22}\cdot 3\cdot 2^n$ г, где b – масса одного мономерного звена x10^{22}, $S = 3+6+\ldots+3\cdot 2^{n-1} = 3\sum_{1}^{n} 2^{k-1} = 3(2^n - 1) \approx 3\cdot 2^n$ – общее число мономеров в структуре дендримера (сумма геометрической прогрессии), а m_3 – масса центральной группы дендримера (ею можно пренебречь по сравнению с массой всего дендримера).

3) Тогда на один грамм дендримера приходится

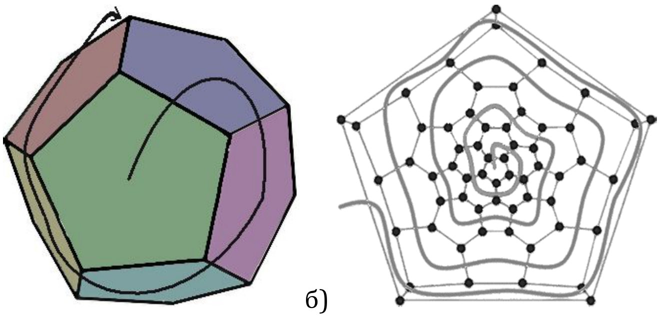
$$\frac{m_1 - m_2}{m_2} = \frac{m_1}{m_2} - 1 = \frac{4{,}13 n^3 \cdot 10^{-21}}{b\cdot 10^{-22}\cdot 3\cdot 2^n} - 1 = \frac{41{,}3 n^3}{b\cdot 3\cdot 2^n} - 1 = m \text{ г белка}$$

Вариант	**b**	**n**	**m$_1$**	**S**	**m$_2$**	**m** белка, г

2	2	12	$7{,}1\cdot10^{-18}$	12288	$2{,}5\cdot10^{-18}$	1,8 (1,9 по общей ф-ле)

4. Спирали фуллеренов

(8 баллов)



а) б)

Рис.1. Примеры спиралей на поверхности фуллеренов.

Все грани некоторого заданного фуллерена C_{32} раскрашены в разные цвета. На его поверхности можно нарисовать спираль в любом направлении из любой грани так, что она последовательно пройдет по всем его граням.

1. Какое минимальное число цветных граней в начале спирали необходимо указать, чтобы однозначно ее задать? (***2.5 балла***)

2. Найдите, сколько разных спиралей, различающихся последовательностью цветов, можно построить для этого фуллерена. (***5.5 баллов***)

Ответ:

1. (2.5 балла) Когда спираль стартует с какого-либо пятиугольника, то существует 5 (по числу соседей) вариантов выбрать второй многоугольник. На следующем шаге, от второго многоугольника, можно пойти далее либо по часовой, либо против часовой стрелки. Дальше спираль можно продолжить только одним единственным способом, т.е., спираль уже задана (*т.о., задается первыми 3-мя гранями*).

2. (5.5 баллов) Число «начал» спиралей, стартующих из любого пятиугольника, будет (см. ответ 1) $5\cdot2 = 10$. Согласно указаниям во вспомогательной информации к очному туру, фуллерен $\mathbf{C_n}$ содержит 12 пятиугольников, тогда общее число спиралей, стартующих с них, равно $10\cdot12 = 120$. Рассуждая аналогично, для шестиугольников получаем $6\cdot2\cdot\Gamma_6 = 12\cdot(n-20)/2 = 6\cdot(n-20)$. [$C_{32}$ $\Gamma_6 = 6$, C_{40} $\Gamma_6 = 10$] Тогда общее число спиралей для произвольного фуллерена $\mathbf{C_n}$ будет равно: $120 + 6\cdot(n-20) = 6n$.

Для фуллерена C_{32} получаем **192** спиралей. Для фуллерена C_{40} получаем **240** спиралей.

5. Переходник графен-трубка. Часть 2: индексы хиральности трубки

(8 баллов)

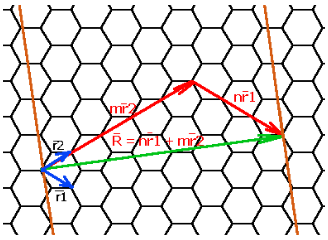


Рис. 1. Для получения нанотрубки (**n**, **m**), графитовую плоскость надо разрезать по пунктирным линиям и свернуть вдоль направления вектора $\vec{R} = n\vec{r}_1 + m\vec{r}_2$. В этом примере **n** = 3 **m** = 5.

Различают следующие типы нанотрубок:

- «зубчатые», **n** = **m**

- зигзагообразные, **m** = 0 или **n** = 0

- спиральные или хиральные нанотрубки (все остальные значения **n** и **m)**

На заочном туре многие из вас познакомились с нанопереходником графен-трубка, который получается, если в разрезы на листе графена добавить 6 графеновых треугольников (см. рисунок 2 а,б).

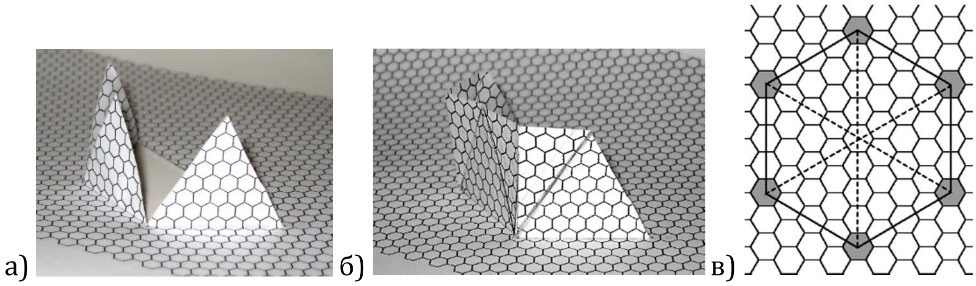


Рис.2. а,б) Промежуточные этапы построение нанопереходника графен-трубка (пример); в) расположение разрезов и семиугольников на листе графена для некоторого переходника.

Для нанотрубки, получающейся из продолжения переходника, схема которого представлена на рис. 2в, определите:

1) ее тип;
2) индексы хиральности (**n**, **m**);
3) диаметр (считать, что трубка имеет форму цилиндра, длина С–С связи равна 0.14 нм).

Ответ:

1) Сворачивание нанотрубки происходит вдоль вектора **R** (рис. 1), поэтому периметр очерченного на рис. 2в шестиугольника равен длине вектора **R**. Сравнивая рисунки, очевидно, что периметр этого шестиугольника проходит параллельно единичному вектору (т.е. один из индексов хиральности трубки равен 0), трубка – **зигзагообразная (n,0)**.

2) Теперь необходимо найти, сколько единичных векторов содержит вектор R. Расстояние между центрами шестиугольников равно длине единичных векторов. Значит, необходимо сосчитать по рисунку 2в число расстояний между центрами шестиугольников, укладывающихся в периметр обведенного шестиугольника. Между двумя соседними вершинами обведенного шестиугольника их укладывается (в вариантах 2 и 4, соответственно) 4 и 6, т.е. это нанотрубки (24,0) и (36,0).

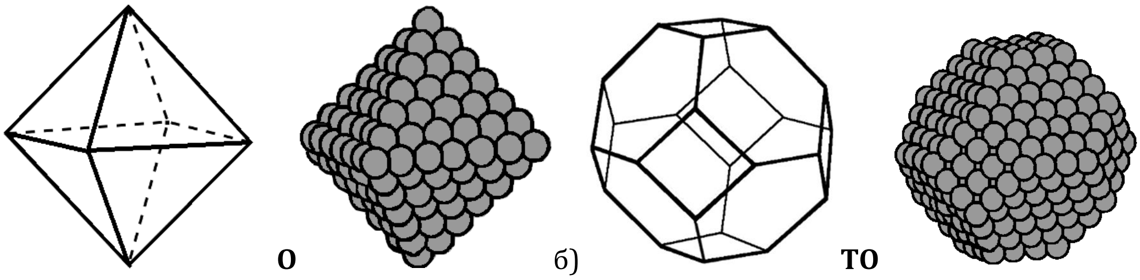
3) Обозначив длину связи С-С как **a**, находим длину единичного радиус вектора – $\mathbf{a}\sqrt{3}$ (как длину малой диагонали углеродного шестиугольника). Зная периметр цилиндрического

сечения углеродной трубки несложно найти ее диаметр: $D = \frac{a\sqrt{3}}{\pi} n = 0{,}077n$. Для вариантов 2 и 4 находим D = 1,9 и 2,8 нм.

Более сложные задачи

6. Золотые кластеры

(20 баллов)



а) **O** б) **TO**

Атомы золота могут образовывать кластеры в форме а) октаэдра **O** с ребром **n** атомов и общим числом атомов **O**(**n**); б) правильного усеченного октаэдра **TO** с ребром **m** атомов и общим числом атомов **TO**(**m**).

На рисунке приведены примеры для **n** = 7 и **m** = 4.

1. Слои в форме какого многоугольника при сложении стопкой формируют кластер **O**? (***0.5 балла***) Сколько и каких фигур мы должны «отсечь» от исходного октаэдра, чтобы получить **TO**? Выразите через **m** число атомов на их ребрах. (***2.5 балла***)

2. Выведите формулы **O**(**n**) (***2 балла***) и **TO**(**m**) (***4 балла***).

3. Оцените размеры кластера **TO** с ребром **m** = 5 и кластера **O**, усечением которого он получен, как радиусы описанных вокруг них сфер. (***5 баллов***). Радиус атома золота считать равным 0.14 нм.

4. При равном объеме более предпочтительной является форма кластера, имеющая меньшую площадь поверхности. Для октаэдра и правильного усеченного октаэдра равных объемов найдите соотношение площадей поверхностей и сделайте вывод, какая форма кластера будет более предпочтительной. (***6 баллов***)

Ответ:

1. *Слои в форме какого многоугольника при сложении стопкой формируют кластер **O**? (**0.5 балла**) Сколько и каких фигур мы должны «отсечь» от исходного октаэдра, чтобы получить **TO**? Выразите через **m** число атомов на их ребрах. (**2.5 балла**)*

2. *Выведите формулы **O(n)** (**2 балла**) и **TO(m)** (**4 балла**).*

1) Октаэдрический кластер представляет собой стопку <u>квадратов</u> из атомов золота, ребро которых сначала растет от 1 до **n**, а затем снова уменьшается до 1.

$$O(n) = \sum_{1}^{n} k^2 + \sum_{1}^{n-1} k^2 = 2\sum_{1}^{n} k^2 - n^2 = 2\frac{2n^3 + 3n^2 + n}{6} - n^2 = \frac{2n^3 + n}{3}$$

2) На каждой из <u>шести</u> квадратных граней **ТО** можно «достроить» <u>правильные квадратные пирамиды</u> из атомов золота, на ребро которых приходится **<u>m – 1</u>** атом (на один меньше, чем ребро **ТО**). При этом **ТО** будет дополнен до «исходного» неусеченного октаэдра, на ребро которого придется **n** = **m** + 2(**m** – 1) = 3**m** – 2 атома (ребро **ТО** плюс удвоенное ребро «отсеченного» кластера в виде четырехугольной пирамиды). Таким образом, общее количество атомов в **ТО** будет равно разности числа атомов в «исходном» октаэдре и суммарного числа атомов золота в шести «отсеченных» пирамидах:

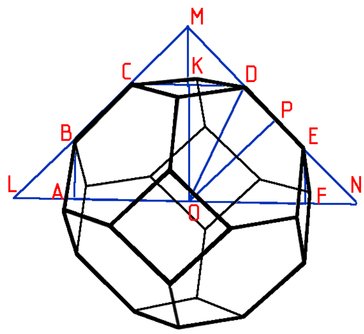
$$TO(m) = O(3m-2) - 6P(m-1)$$

где $P(m) = \sum_{1}^{m} k^2 = \frac{2m^3 + 3m^2 + m}{6}$ - число атомов в одной квадратной пирамиде.

$$TO(m) = \frac{2(3m-2)^3 - (3m-2)}{3} - 6\frac{2(m-1)^3 + 3(m-1)^2 + (m-1)}{6}$$

Упрощая до многочлена, получаем: $TO(m) = 16m^3 - 33m^2 + 24m - 6$

3. *Оцените размеры кластера **ТО** с ребром **m** = 5 и кластера **О**, усечением которого он получен, как радиусы описанных вокруг них сфер. (**5 баллов**). Радиус атома золота считать равным 0.14 нм.*



1) Рассмотрим фигуру **ABCDEF** (рис.). Здесь **BC** = **DE** = **a** – ребро правильного усеченного октаэдра, **AB** = **CK** = **KD** = **EF** = $\frac{\sqrt{2}}{2}a$ – половина диагонали квадратной грани правильного усеченного октаэдра.

2) Чтобы рассчитать радиус сферы, описанной около усеченного октаэдра, **OD**, нам необходимо найти длину отрезка **KO**, являющегося высотой рассматриваемой фигуры **ABCDEF**.

3) Треугольники **CMD** и **LMN** подобны по первому признаку подобия ($\angle LMN = \angle CMD$, $\angle MLN = \angle MCD$ как соответственные при параллельных прямых **CD** и **LN**) . При этом **MD** = **DE** = **EN** = ***a*** = 1/3 **MN** = 1/3***A***. Следовательно, **MK** = 1/3 **MO**, и **KO** = 2/3 **MO**.

4) **MO** = $\frac{\sqrt{2}}{2}$ **MN** = $\frac{\sqrt{2}}{2}$ ***A*** (как катет равнобедренного прямоугольного треугольника **MON**, отвечает радиусу сферы, описанной около октаэдра).

Тогда **KO** = 2/3 **MO** = $\frac{\sqrt{2}}{2} \cdot \frac{2}{3}$ **MN** = $\frac{\sqrt{2}}{3} \cdot 3a = a\sqrt{2}$.

5) $OD = \sqrt{KD^2 + KO^2} = \sqrt{\left(\frac{\sqrt{2}}{2}a\right)^2 + \left(a\sqrt{2}\right)^2} = a\sqrt{\frac{2}{4} + 2} = a\sqrt{\frac{5}{2}}$.

6) Для кластера **ТО**: $a = 2r(m-1)$ и $R_{TO} = 2r(m-1)\sqrt{\frac{5}{2}} + r$ (рассматриваем радиус сферы, проходящей через центры атомов, лежащих в вершинах кластера, и прибавляем поправку на радиус самого атома).

m = 5 и **r** = 0,14 нм: $R_{TO} = 2 \cdot 0{,}14(5-1)\sqrt{\frac{5}{2}} + 0{,}14 = 1{,}9$ нм.

7) Для кластера **О**: $A = 2r(n-1)$ и $R_O = \sqrt{2}r(n-1) + r$ (рассматриваем радиус сферы, проходящей через центры атомов, лежащих в вершинах кластера, и прибавляем поправку на радиус самого атома).

$n = 3m - 2 = 15 - 2 = 13$ и **r** = 0,14 нм: $R_O = \sqrt{2} \cdot 0{,}14(13-1) + 0{,}14 = 2{,}5$ нм

Примечание: Второй вариант расчета радиуса описанной сферы для усеченного октаэдра:

$$OD = \sqrt{OP^2 + PD^2} = \sqrt{\left(\frac{A}{2}\right)^2 + \left(\frac{a}{2}\right)^2} = \frac{1}{2}\sqrt{(3a)^2 + a^2} = \frac{a}{2}\sqrt{10}$$

$$R_{TO} = r(m-1)\sqrt{10} + r = 0{,}14(5-1)\sqrt{10} + 0{,}14 = 1{,}9 \text{ нм}$$

или $OD = \frac{1}{2}\sqrt{(2r(n-1))^2 + (2r(m-1))^2} = r\sqrt{(n-1)^2 + (m-1)^2}$ и $R_{TO} = 0{,}14\left(\sqrt{144+16} + 1\right) \approx 1{,}9$ нм

4. *При равном объеме более предпочтительной является форма кластера, имеющая меньшую площадь поверхности. Для октаэдра и правильного усеченного октаэдра равных объемов найдите соотношение площадей поверхностей и сделайте вывод, какая форма кластера будет более предпочтительной. (**6 баллов**)*

1) Объем кластера **О** равен $V_O = 2 \cdot \frac{1}{3} \cdot A^2 \cdot A\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{3}A^3$

2) Объем кластера **ТО** равен $V_{TO} = \frac{\sqrt{2}}{3}(3a)^3 - 6 \cdot \frac{\sqrt{2}}{6}a^3 = 8\sqrt{2}a^3$

3) При равенстве объемов ($\frac{\sqrt{2}}{3}A^3 = 8\sqrt{2}a^3$) ребра кластеров соотносятся как $A = 2\sqrt[3]{3}a$.

4) Площадь поверхности кластера **О** равна $S_O = 8 \cdot \frac{\sqrt{3}}{4}A^2 = 2\sqrt{3}A^2$

5) Площадь поверхности кластера **ТО** равна $S_{TO} = 6a^2 + 8 \cdot 6 \cdot \frac{\sqrt{3}}{4}a^2 = 6a^2\left(1 + 2\sqrt{3}\right)$

6) Тогда $\frac{S_{TO}}{S_O} = \frac{6a^2\left(1+2\sqrt{3}\right)}{2\sqrt{3}A^2} = \frac{3a^2\left(1+2\sqrt{3}\right)}{\sqrt{3}\left(2\sqrt[3]{3}a\right)^2} = \frac{3\left(1+2\sqrt{3}\right)}{4\sqrt{3}\left(\sqrt[3]{3}\right)^2} = 0{,}93$ (0,99 если брать округленные значения корней)

То есть, более предпочтительной является форма правильного усеченного октаэдра.

7. Чехарда и фуллерены

(20 баллов)

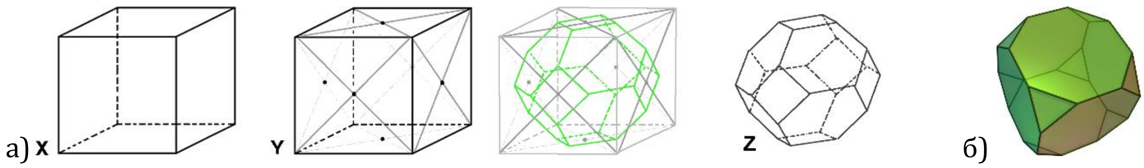


Рис.1 а) Пример применения операции чехарды к кубу; б) усеченный куб

Простая операция чехарды позволяет «перепрыгнуть» от многогранника **X** к многограннику **Z** (рис. 1а). Для этого в центры граней **X** необходимо поставить точки и соединить их с близлежащими вершинами: получится промежуточный многогранник **Y**, соединив центры граней которого получим **Z**.

1. Из какого многогранника **X** в результате «чехарды» получается усеченный куб (рис. 1б)? (***2.5 балла***)

Операция чехарды часто используется при моделировании фуллеренов, поскольку для любого исходного фуллерена $\mathbf{C_x}$ она позволяет построить новый фуллерен – $\mathbf{C_z}$.

2. Опишите (проиллюстрировав рисунками), какие структурные элементы $\mathbf{C_z}$ образуются внутри пяти- и шестиугольников исходного фуллерена $\mathbf{C_x}$, а также на его связях и вокруг его вершин. (***3 балла***)

3. Выразите число вершин нового фуллерена **z** через число вершин исходного **x**. (***4 балла***)

4. Чему равно минимально возможное значение **z** при чехарде фуллеренов? (***2.5 балла***) Может ли при этом исходный фуллерен $\mathbf{C_x}$ быть получен в результате чехарды из некоторого другого многогранника? Поясните. (***3 балла***)

5. Какая главная структурная особенность расположения пятиугольников существует для всех полученных чехардой фуллеренов? (***2 балла***) На рисунке 2 найдите структурный изомер C_{72}, который получается операцией чехарды. Ответ обоснуйте. (***3 балла***)

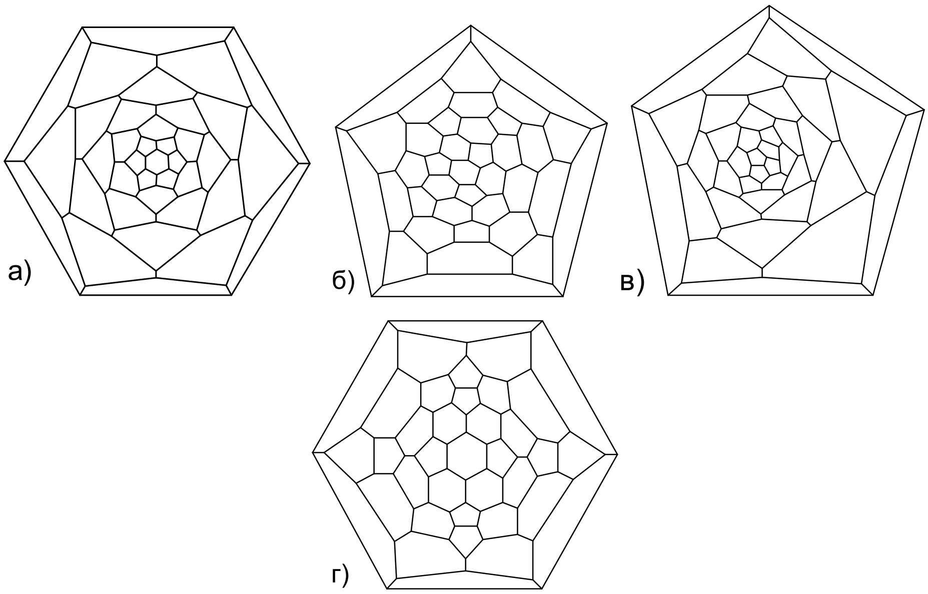
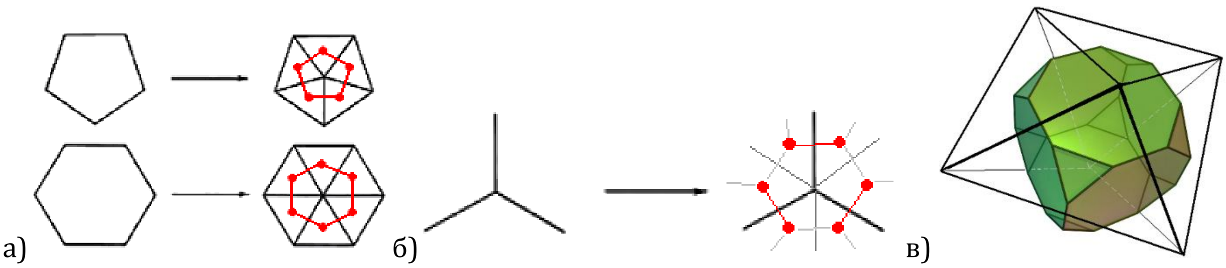


Рис. 2 Диаграммы Шлегеля для разных структурных изомеров фуллерена C_{72}.

Решение:



Внимательно рассматривая рисунки, можно заметить закономерности чехарды:

– внутри **n**-угольной грани исходного многогранника **X** получается **n**-угольник конечного многогранника **Z**, повернутый относительно исходного (рис. а);

– при этом в **Z** появляется еще один тип многоугольников. Если в исходном многограннике **X** в какой-то вершине сходилось **m** ребер, то вокруг этой вершины **X** образуется 2**m**-угольник многогранника **Z** (рис. б);

– вместо ребер старого многогранника **X** возникают «повернутые» на 90º ребра нового многогранника **Z** (рис. б).

1. *Из какого многогранника **X** в результате «чехарды» получается усеченный куб (рис. 1б)?* (***2.5 балла***)

Треугольные грани **Z** на рис. 1б условия могли появиться только из треугольников исходного **X** (поскольку 3 – нечетное число), значит, исходный многогранник имеет 8 треугольных граней. Оставшиеся 6-ть 8-миугольных граней усеченного куба – это упомянутый выше «появившийся» вокруг 6-ти вершин исходного **X** новый тип граней. Т.е. в каждой из 6-ти вершин **X** сходилось по 8/2 = 4 ребра (4 грани). Многогранник **X**, в котором 6 вершин, в каждой из которых сходится по 4 треугольных грани – это четырехугольная бипирамида, состоящая из равносторонних треугольников – **октаэдр** (рис. в).

2. *Опишите (проиллюстрировав рисунками), какие структурные элементы* $\boldsymbol{C_z}$ *образуются внутри пяти- и шестиугольников исходного фуллерена* $\boldsymbol{C_x}$*, а также на его связях и вокруг его вершин. (**3 балла**)*

Внутри пятиугольников возникают пятиугольники, внутри шестиугольников – шестиугольники; на месте <u>связей</u> возникают «повернутые» связи; вокруг <u>вершин</u> старого фуллерена возникают шестиугольники нового фуллерена (рис. а,б).

3. *Выразите число вершин нового фуллерена **z** через число вершин исходного **x**.* (***4 балла***)

Поскольку внутри пятиугольников и шестиугольников старого фуллерена $\mathbf{C_x}$ образуются соответственно пятиугольники и шестиугольники нового фуллерена (рис.а), значит, формула нового фуллерена будет:

$\mathbf{z} = 5\Gamma_5 + 6\Gamma_6 = 5\cdot 12 + 6\cdot(\mathbf{x}-20)/2 = 60 + 3\mathbf{x} - 60 = \mathbf{3x}$

где Γ_5 и Γ_6 – количество 5-ти и 6-ти угольников старого фуллерена C_x, соответственно; согласно указаниям в конце задач можно было использовать без вывода $\Gamma_5 = 12$ и $\Gamma_6 = (\mathbf{x}-20)/2$

4. *Чему равно минимально возможное значение **z** при чехарде фуллеренов?* (***2.5 балла***)
Очевидно, полученный в результате чехарды фуллерен с минимальным числом вершин может быть получен из фуллерена $\mathbf{C_x}$ с минимальным числом вершин. Поскольку для любого фуллерена $\Gamma_5 = 12$ и $\Gamma_6 = (\mathbf{x}-20)/2$, то это имеющий форму додекаэдра фуллерен C_{20} (содержит 0 шестиугольников, и 12 пятиугольников). Из него чехардой получится фуллерен, в котором в 3 раза больше вершин, т.е. $\mathbf{C_{60}}$ (это именно бакибол с разделенными пятиугольниками, а не какой-то его другой его изомер).

Может ли при этом исходный фуллерен $\boldsymbol{C_x}$ *быть получен в результате чехарды из некоторого другого многогранника? Поясните.* (***3 балла***)

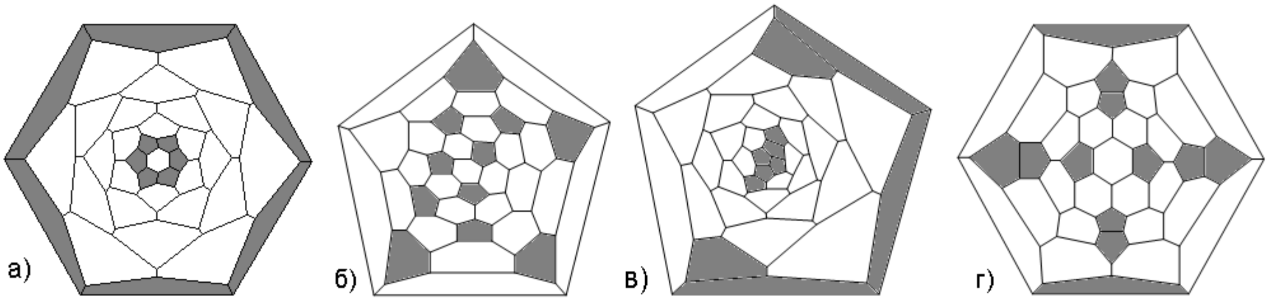
При чехарде каждая вершина исходного многогранника **X** приводит к граням **Z** с <u>четным</u> числом вершин (рис. б). Поскольку в додекаэдре нет ни одной такой грани, то он заведомо не может быть получен в результате чехарды (несмотря на то, что существует соответствующий додекаэдру промежуточный двойственный многогранник **Y** – икосаэдр).

5. *Какая главная структурная особенность расположения пятиугольников существует для всех полученных чехардой фуллеренов?* (***2 балла***)

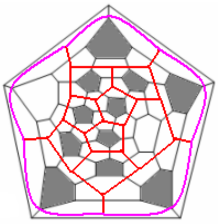
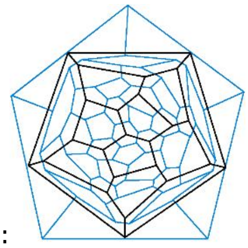
Поскольку пятиугольники нового фуллерена $\mathbf{C_z}$ оказываются внутри пятиугольников исходного, и вокруг каждой вершины пятиугольников исходного фуллерена после чехарды образуются шестиугольники (см. рис. 1 а,б) то <u>все</u> пятиугольники нового фуллерена $\mathbf{C_z}$ оказываются **разделены** шестиугольниками.

На рисунке 2 найдите структурный изомер C_{72}*, который получается операцией чехарды. Ответ обоснуйте.* (***3 балла***)

Значит, фуллерены, которые не содержат разделенных пятиугольников, *заведомо* не могут быть получены в результате чехарды. Для ответа на вопрос необходимо найти пятиугольники на приведенных проекциях Шлегеля (например, закрашивая их). Все изомеры кроме (б) не содержат разделенных пятиугольников, значит, только он мог быть получен в результате операции чехарды из фуллерена C_{24}.



Так же этот пункт можно было решать графически: попробовать построить вокруг каждого пятиугольника на данных в условии проекциях Шлегеля изомеров C_{72} пятиугольники исходного фуллерена C_x. Несложно убедится, что построенные таким образом пятиугольники смогут «сложится» в фуллерен только для (б).


или, более наглядно:


8. Биоинформатика, часть 2: Архиватор для хомячка

(20 баллов)

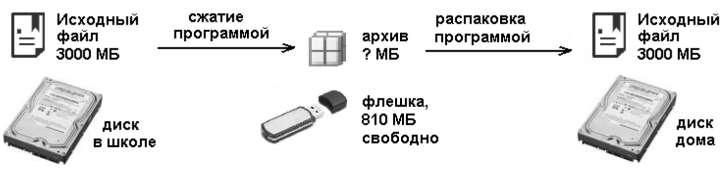
Информация на компьютере хранится в ячейках памяти, каждая из которых может принимать значение 0 или 1 и несет один бит информации. Для записи (кодирования) любого символа в текстовом файле используются 8 таких ячеек (8 бит).

У юного нанотехнолога Полуэкта есть текстовый файл размером ровно 3 000 мегабайт, в котором только четырьмя символами, отвечающими нуклеотидам (**A**, **C**, **G**, **T**), записан фрагмент генома сирийского хомячка (*Mesocricetus auratus*).

ACTCATTGACACTATAGTAGACACATAGCTCGTATGTCTACAGACACACTCCGTCAGAGTGAGTACAGACATCTG
AAGGGCGGGGACAGCTGCAGCTCGTGCAGACGAGGATTACCTTCGTGAGTATATCATTGTACTAGCTCTCCGTGA

1. Рассчитайте общее число символов нуклеотидов в файле. (***2 балла***)

Для домашней работы Полуэкт решил скопировать этот файл к себе на флешку. Однако на ней оставалось свободно только 810 мегабайт, а имеющийся у него популярный архиватор смог заархивировать (сжать)[1] файл лишь до ~843.2 мегабайт. Сначала Полуэкт расстроился, что ему придется удалить часть информации с флешки, затем он быстро смекнул, как написать простейший архиватор, который сможет сжать файл, а затем распаковать его обратно дома.



2. Какое минимальное количество информации требуется, чтобы закодировать произвольный нуклеотид в последовательности? Найдите размер архива, получившегося у Полуэкта, и вкратце опишите алгоритм работы архиватора. (***4 балла***)

Дома Полуэкт случайно открыл архив в текстовом редакторе и, листая, обнаружил слово «нано»[2].

3. Какова длина нуклеотидной последовательности, закодированной в этом слове? Рассчитайте, сколько раз слово «нано» будет встречаться в архиве, полученном сжатием по такому же алгоритму случайной нуклеотидной последовательности размером 1 терабайт. (***5 баллов***)

4. Сколько разных нуклеотидных последовательностей может кодировать это слово? (***3 балла***)

Открыв архив в шестнадцатеричном редакторе (позволяет увидеть в шестнадцатеричном виде находящуюся в ячейках памяти компьютера информацию), Полуэкт обнаружил, что буквам *н*, *а*, *о* соответствуют числа ED_{16}, $E0_{16}$, EE_{16}.

```
FF CD B4 FA 5E 55 75 42 27 ED E0 ED EE 2E B9 2B  яНґъ^UuB'нано.№+
CB C5 2E 64 35 F0 7C CB 5A CF 2C 87 7A 4A A6 CC  ЛЕ.d5p|ЛZП,‡zJ¦М
```

5. Расшифруйте соответствующую этому слову нуклеотидную последовательность, если новые символы нуклеотидов в ней появляются в обратном алфавитному порядке[3]. (***6 баллов***)

[1] Сжатие (упаковка) данных – алгоритмическое преобразование данных, производимое с целью уменьшения занимаемого ими объёма. Программа, осуществляющая процедуры сжатия и распаковки, называется архиватором.

[2] Написанное строчными русскими буквами. Коды больших и маленьких, а также русских и английских букв различаются.

[3] Примером последовательности нуклеотидов с таким же «порядком появления» символов является **T**T**G**TG**C**CTCG**A**A

При решении не забывайте, что $0_{16} = 0_{10} \ldots F_{16} = 15_{10}$

<u>Вспомогательная информация:</u>

Фуллерен – каркасная молекула углерода. Может быть описан как выпуклый многогранник, составленный из правильных пяти- и шестиугольников, в каждой вершине которого сходятся по три ребра. Любой фуллерен может быть представлен в виде диаграммы Шлегеля – плоской проекции на одну из граней.

Число пятиугольников в произвольном фуллерене C_n составляет $Г_5 = 12$, шестиугольников – $Г_6 = 0.5(n – 20)$.

$\pi = 3.1$; $\sqrt{2} \approx 1.4$, $\sqrt{3} \approx 1.7$, $\sqrt{5} \approx 2.2$, $\sqrt{7} \approx 2.6$, $\sqrt[3]{2} \approx 1.3$, $\sqrt[3]{3} \approx 1.4$, $\sqrt[3]{5} \approx 1.7$.

Формула суммы квадратов последовательности натуральных чисел 1, 2, …, n: $\sum_{m=1}^{m=n} m^2 = \frac{n(n+1)(2n+1)}{6}$.

Решение:

1. *Рассчитайте общее число символов нуклеотидов в файле.* (***2 балла***)

1 килобайт = 1024 (2^{10}) байт, 1 мегабайт = 1024 килобайт, 1 символ = 8 бит = 1 байт. Значит, файл содержит столько же символов, сколько и байт. Т.е. в файле 3 000 мегабайт записана информация о $3000 \cdot 1024 \cdot 1024 = 3000 \cdot (1000+24)^2 = 3000 \cdot (1000^2+2 \cdot 24000+24^2) = 3000 \cdot 1048576 = 3\,145\,728\,000$ нуклеотидах.

2. *Какое минимальное количество информации требуется, чтобы закодировать произвольный нуклеотид в последовательности? Найдите размер архива, получившегося у Полуэкта, и вкратце опишите алгоритм работы архиватора.* (***4 балла***)

Поскольку символов нуклеотидов всего 4, то для их кодирования можно использовать меньшее число бит, чем для кодирования *произвольного* текстового символа. 4-м вариантам нуклеотида можно сопоставить 4 значения двух ячеек памяти (два бита, $\log_2 4 = 2$): 00, 01, 10, 11. Значит, число кодирующих символ бит можно уменьшить в 8/2 = 4 раза, размер файла при этом составит 3000/4 = **750** Мб. Архиватор Полуэкта читает символ из файла и в соответствии с выбранной Полуэктом таблицей соответствия символов пишет в итоговый файл соответствующие символу значения 2-х бит. При распаковке файла, наоборот: читает по 2 бита, и в соответствии с той же таблицей пишет в распакованный файл текстовый символ нуклеотида.

3. *Какова длина нуклеотидной последовательности, закодированной в этом слове? Рассчитайте, сколько раз слово «нано» будет встречаться в архиве, полученном сжатием по такому же алгоритму случайной нуклеотидной последовательности размером 1 терабайт.* (***5 баллов***)

Поскольку каждой букве текстового файла соответствуют 8 бит, а заархивированному нуклеотиду 2, то в коде одной буквы архива содержится информация о 4-х нуклеотидах. Тогда слову «нано» в архиве соответствует последовательность из $4 \cdot 4$ = **16**-ти нуклеотидов.

Нуклеотидная последовательность, занимающая 1 **тера**байт, содержит 2^{40} нуклеотидов (килобайт - 2^{10}, мегабайт – 2^{20}, гигабайт – 2^{30}, терабайт - 2^{40}). Поскольку в каждом положении последовательности может быть один из 4-х нуклеотидов, то вероятность того, что произвольная последовательность 16-ти нуклеотидов случайно совпадет с предварительно заданной последовательностью, составит $(1/4)^{16} = 2^{-32}$.

В случайной последовательности из 2^{40} нуклеотидов содержится $2^{40}-16 \approx 2^{40}$ последовательностей из 16-ти нуклеотидов, каждая из которых имеет вероятность 2^{-32} совпасть с заданной (приводящей к слову «нано»). Следовательно, после описанного архивирования 1 терабайта случайных нуклеотидов в среднем получается $2^{40} \cdot 2^{-32} = 2^8$ = **256** слов «нано».

4. *Сколько разных нуклеотидных последовательностей может кодировать это слово?* (***3 балла***)

Количество последовательностей равно числу возможных вариантов таблиц соответствия 4-м буквам нуклеотидов (A, C, G, T) 4-х значений 2-х бит (00, 01, 10, 11) – т.е. числу размещений из 4 по 4 (числу перестановок порядка 4): $P_4 = 4! = 4 \cdot 3 \cdot 2 \cdot 1 =$ **24**

5. *Расшифруйте соответствующую этому слову нуклеотидную последовательность, если новые символы нуклеотидов в ней появляются в обратном алфавитному порядке.* (***6 баллов***)

Найдем, какие последовательности нулей и единиц сопоставлены каждой русской букве. Для этого переведем шестнадцатеричные коды в двоичные (через десятичную систему)

$\mathbf{ED}_{16} = 14 \cdot 16^1 + 13 \cdot 16^0 = 237_{10}$

237/2=118,5 (1); 118/2=59 (0); 59/2=29,5 (1); 29/2 = 14,5 (1); 14/2 = 7 (0); 7/2 = 3,5 (1); 3/2=1,5 (1); 1/2=0,5 (1)

$\mathbf{ED}_{16} = 237_{10} = 11101101_2$

$\mathbf{E0}_{16} = 14 \cdot 16^1 + 0 \cdot 16^0 = 224_{10}$

224/2=112 (0); 112/2=56 (0); 56/2=28 (0); 28/2 = 14 (0); 14/2 = 7 (0); 7/2 = 3,5 (1); 3/2=1,5 (1); 1/2=0,5 (1)

$\mathbf{E0}_{16} = 224_{10} = 11100000_2$

$\mathbf{EE}_{16} = 14 \cdot 16^1 + 14 \cdot 16^0 = 238_{10}$

238/2=119 (0); 119/2=59,5 (1); 59/2=29,5 (1); 29/2 = 14,5 (1); 14/2 = 7 (0); 7/2 = 3,5 (1); 3/2=1,5 (1); 1/2=0,5 (1)

$\mathbf{EE}_{16} = 238_{10} = 11101110_2$

Получается, что слову «нано» соответствует следующая последовательность бит:

```
   н        а        н        о
/----------\ /----------\ /----------\ /----------\
11101101|11100000|11101101|11101110
```

поскольку для кодирования каждого нуклеотида требуется 2 бита, а новые нуклеотиды в последовательности, согласно условию, появляются в обратном алфавитному порядке, то можем сопоставить:

```
11 10 11 01    11 10 00 00    11 10 11 01    11 10 11 10
 T  G  T  C     T  G  A  A     T  G  T  C     T  G  T  G
```

Отметим, что Полуэкт присвоил алфавитному порядку символов нуклеотидов двоичные коды в порядке возрастания: A ↔ 00, C ↔ 01, G ↔ 10, T ↔ 11