



## Математика для школьников 7 – 11 класса (заочный тур) Решение задачи 10. Триpletный Код Полуэкта

1. Число вариантов триплетов составляет  $4^3 = 64$ , этого хватит, чтобы закодировать необходимые Полуэктору  $26 + 26 + 1 + 7 = 60$  символов. Даже останется 4 неиспользованных триплета.

При этом возможно всего

$$A_{64}^{60} = \frac{64!}{4!} \approx 5,29 \cdot 10^{87}$$

вариантов триpletных кодировок.

2. Для решения необходимо будет или менять в тексте все угаданные символы аминокислот на разгаданные символы (либо в электронном виде в редакторе, либо в распечатке с помощью корректора), или подписывать разгаданные символы на распечатке между строк карандашом.

Однако, поскольку число аминокислот (20) меньше числа букв в английском алфавите, при расшифровке будут неизбежны коллизии (одной аминокислоте могут соответствовать несколько букв), остается только надеяться на то, что Полуэкт подбирал кодировку так, чтобы коллизии не сильно затрудняли расшифровку. Также важно не забывать, что по условию каждому символу текста соответствует аминокислота и каждая аминокислота соответствует символу текста.

Для начала в закодированном тексте необходимо разгадать разделители слов. Самый простой способ: догадаться, что последний символ “Y” – скорее всего, точка или другой знак препинания, поскольку он стоит в самом конце. В тексте мы видим еще 4 таких символа, и после каждого из них стоит символ “M” – это, вероятно, пробел.

После подстановки получаем следующий текст:

CNCWFVSANCD SNG RVRDSVP SPVNHDWC DP WCV WV HAV FWP H LPWFDPDCA  
RDPV SHDWCP WV HAV CNCWHVSACGWADVP. CW RWCRVP HANH HAV PVSVCH  
CWKVG L PDSV RNP NRNPRVR VWP RVRVGWLFVCH WV HAV CNCWFNSADCVP  
SPVNHVR KT H AV KWHHWF IQL NLLPWNSA. DH RDGG NGGWR SPVNHDCA  
CNCWRVRDPVP RDHA SW CHPWGGNKG V FWRV FVCHP NH HAV CNCWPSNGV.  
RADSA SWQGR GVNR HW VQHQP V PVRWGQHDWC DC CVR FNHVDPNGP  
SPVNHDWC NCR RDPVNPV HPVNH FVCH. NCR FNT GVNR HW HAV AWGT  
APNDG WV CNCWHVSACGWADVP I CNCWPWKWHP.

Бросается в глаза обилие коротких слов HAV – это, вероятно, определенный артикль the. Слово “nano” примечательно тем, что первая и третья буквы у него одинаковы. Заметим, что первое слово начинается с “CNCW”, так же начинаются еще 7 других слов. Делаем замену H => t A => h V => e, C => n, N => a, W => o:

nanoFeShanDSaG ReRDSeP SPeatDon DP one oe the FoPt LPoFDPDnh  
RDPeStDonP oe the nanoteShnoGohDeP. no RonReP that the PeSent  
noKeG LPDSe RaP aRaPReR eoP ReReGoLFent oe the nanoFaShDneP  
SPeateR KT the KottoFIQL aLLPoaSh. Dt RDGG aGGoR SPeatDnh  
nanoReRDPeP RDth SontPoGGaKGe FoReFentP at the nanoPSaGe.  
RhDSh SoQGR Gear to eQtQPe PeRoGQtDon Dn neR FatePDaGP  
SPeatDon anR RDPeaPe tPeatFent. anR FaT Gear to the hoGT  
hPaDG oe nanoteShnoGohDeP I nanoPoKotP.

Текст начинает угадываться, но также становятся видны последствия «слияния» букв, (например: “one oe the” вместо “one of the” – при этом очевидно, слились две *последовательные* буквы алфавита «e» и «f»).

Расшифровываем дальше. Слово nanoteShnoGohDeP похоже на «nanotechnolohies» = “nanotechnologies” (опять слились *последовательные* буквы «g» и «h»), подставляя S => c, G => l, D => i, P => s, получаем:

nanoFechanical ReRices cseation is one oe the Fost LsoFisinh  
Risectons oe the nanotechnolohies. no RonRes that the secent  
noKel Lsice Ras aRasReR eos ReReloLFent oe the nanoFachines  
cseateR KT the KottoFIQL aLLsoach. it Rill alloR cseatinh  
nanoReRises Rith contsollaKle FoReFents at the nanoscale.  
Rhich coQlR leaR to eQtQse seRoLQtion in neR Fatesials  
cseation anR Risease tseatFent. anR FaT leaR to the holt  
hsail oe nanotechnolohies I nanosoKots.

nanoFechanical = nanomechanical (F => m)

ReRices = devices (R => слившиеся d и v, видим, что также могут сливаться и *непоследовательные* буквы, причина этого станет ясна позже)

nanomechanical dedices cseation is one oe the most Lsomisinh  
disectons oe the nanotechnolohies. no dondes that the secent  
noKel Lsice das adasded eos dedeloLment oe the nanomachines  
cseated KT the KottomIQL aLLsoach. it dill allod cseatinh  
nanodedises dith contsollaKle modements at the nanoscale.  
dhich coQld lead to eQtQse sedolQtion in ned matesials  
cseation and disease tseatment. and maT lead to the holt  
hsail oe nanotechnolohies I nanosoKots.

coQld = could (Q => u)

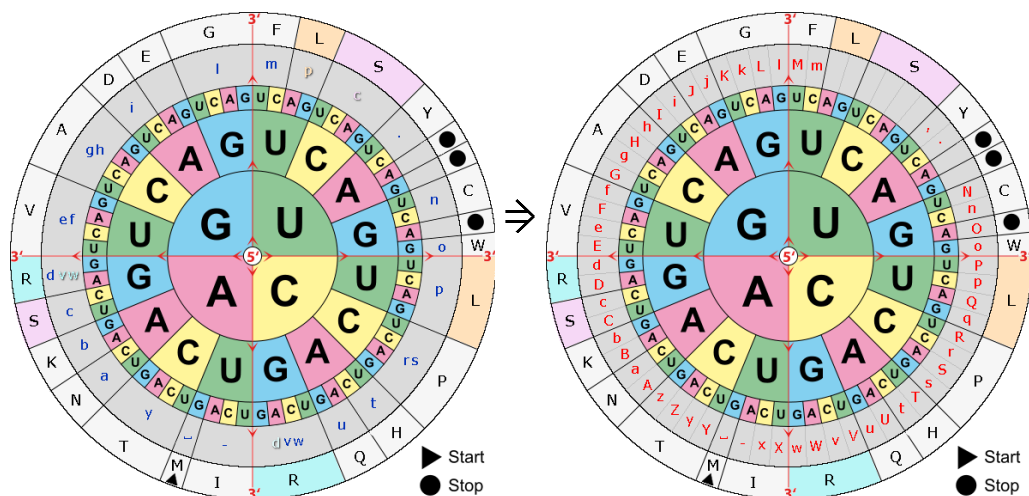
KT the Kottom = by the bottom (K=b, T=y)

aLLsoach, Lsomisinh (L = p, s,r <= P)

nanomechanical dedices cseation is one oe the most psomisinh  
disectons oe the nanotechnolohies. no dondes that the secent  
nobel psice das adasded eos dedelopment oe the nanomachines  
cseated by the bottom-up appsoach. it dill allod cseatinh  
nanodedises dith contsollable modements at the nanoscale.  
dhich could lead to eutuse sedolution in ned matesials  
cseation and disease tseatment. and may lead to the holy  
hsail oe nanotechnolohies - nanosobots.

das adaded = was awarded (d, w <= R)

Удобно отмечать разгаданные буквы (включая сливающиеся) на карте:



Становится видна система в расположении букв: они почти всегда располагаются по алфавиту по часовой стрелке. Коллизии происходят, когда одну аминокислоту кодируют более 2-х триплетов. Считая, что алфавитный порядок букв по часовой стрелке не нарушается, мы сможем сопоставить триплеты буквам, при этом становятся понятны коллизии не идущих по алфавиту букв (относятся к 3 аминокислотам, S, R, L, которые встречаются более чем в одном непрерывном секторе).

Теперь, зная, какие буквы сливались друг с другом, мы можем (перебирая ограниченное число комбинаций, а также исправив пару опечаток) восстановить полный текст:

Nanomechanical devices creation is one of the most promising directions of the nanotechnologies. No wonder that the recent Nobel Prize was awarded for development of the nanomachines created by the bottom-up approach. It will allow creating nanodevices with controllable movements at the nanoscale, which could lead to future revolution in new materials creation and disease treatment, and may lead to the Holy Grail of nanotechnologies - nanorobots.

Похожие методы расшифровки описаны в популярной художественной литературе, см. Эдгар Аллан По «Золотой жук» и Артур Конан Дойл «Пляшущие человечки».

3. Если считать, что заглавные и строчные буквы расположены на схеме единообразно (т.е. что первым триплетом по часовой стрелке всегда кодируется либо заглавная, либо строчная буква) – то можно. На помощь нам приходит буква “o”. Поскольку в тексте она везде встречается в середине слов, можно утверждать, что это именно строчная буква. Тогда для соответствующей ей заглавной буквы “O” остается только единственное место – под идущим ранее стоп кодоном (=> первыми по часовой стрелке кодируются заглавные буквы).

4. Нуклеотиды РНК:

AUU GAU AUG GAU UGC CAC UAC AUG UGA GGG ACC UUC CUC UAC AUG UGG UGC AUG  
UGU AAC UGC UGG CAC GUC AGC GCG UGC UGG GGG UGG GCC GAC GUC CCG

5. Полипептид:

IDMDCHYM (поскольку букве "O" соответствует стоп-кодон UGA, то, прочитав его согласно указанному в условии коду, рибосома остановит сборку).