Námět:

Provázanost

Motto: Velmi málo lidí by zřejmě mohlo říct, že se zajímá o počítače i genetiku. Nakonec je obojí téměř totéž až na fakt, že jedno je o strojích a druhé o živočiších, tedy i o lidech…

Příběh:
Univerzitní profesor, John Mahoney, pořádá pro veřejnost seminář o využití počítačů v genetice. Seminář však ukončí, když se dostává k závěru, že podstata genetiky se dá přirovnat k programování počítačových systémů. Nápadu se nakonec chytí skupina počítačových odborníků, kteří hodlají pomoct genetickému inženýrství. Jeden člen skupiny, James Week, začne přirovnávat tvorbu RNA k programování v jazyce C a poté tvorbu proteinů z RNA jako kompilaci spustitelných souborů z těch objektových (.o, .obj). Ovšem také zmiňuje rozdíly mezi DNA a strojovým kódem, čímž se dopracovává k názoru, že DNA má 4 pozice podle svých amino kyselin na 4 typy: Thymin (T), Adenin (A), Cytosin (C) a Guanin (G). Ovšem by se zde měla zachovat komplementarita bází, tedy spojovat jenom Cytosin s Guaninem a Thymin s Adeninem. Zachová-li se komplementarita bází, vznikají nám jen bitové položky s možnosti 0 a 1, při čemž Thymin s Adeninem může reprezentovat aktivní stav (1) a Cytosin s Guaninem naopak neaktivní stav (0). James také tvrdí další pravdivé tvrzení: „Jeden gen neobsahuje jednu látku daného typu, ale velmi mnoho takových látek. Gen obsahuje velmi mnoho bílkovin thyminu, adeninu, cytosinu i guaninu.“. Druhý člen skupiny, George Friend, uvažuje o možnosti ukládání počítačových dat do řetězce DNA. Na tento dotaz profesor Mahoney odpovídá: „Na základě délky DNA by bylo možné uložit velmi mnoho počítačových dat, okolo 3 GB, ale teoreticky by tato hodnota mohla klesnout na 750 MB.“. George přesmýšlí stále nad možnosti využití řetězce DNA pro ukládání počítačových dat. Nakonec se pozastavuje u možnosti ukládat data do struktury lidského vlasu. James přechází na Georgovo téma: „Pokud by se vytvořilo zařízení, které by umožňovalo zpracovávat řetězce DNA, to znamená, že by je to mohlo číst a zapisovat, tak by se takto mohl vytvořit ultimátní nosič počítačových dat (médium).“. Do diskuze se nyní zapojuje poslední člen skupiny, Jeofrey Walker, který zde mluví o možnosti využití nových poznatků pro pozitivní přeprogramování buněk na zvýšení jejich imunity. Všichni přítomní začnou spolupracovat na tvorbě nového média, ale postupem času zjišťují, že výroba neprobíhá přesně tak, jak předpokládali. Nakonec byl projekt pozastaven, protože George bohužel opomenul jednu substanci celého pokusu, čímž byl projekt odložen. Zbývalo jen málo do dokončení nového média a získání slávy, ale opět jen lidská nepozornost znemožnila vejít těmto vědcům v známost s jejich ultimátním médiem…

(Napsáno cca. v roce 2004)

Zamyšlení:

Počítače komunikují ve dvojkové soustavě. Jedna základní jednotka informace se nazývá bit, bit je zkratka z ang. výrazu „**bi**nary digi**t**“, tedy binární číslice. Bit může mít jen 2 stavy, nastaveno (1) či nenastaveno (0). V anglické technické literatuře se uvádí často těchto stavů jako **set / not set** . Sekvence 8 bitů nám vytváří jeden byte, což je jednotka reprezentována 2^8 hodnotami (protože máme 2 stavy na 8 bitů). Tímto získáváme maximální možnou hodnotu 256. Nicméně na základě principu fungování binární (dvojkové) soustavy, které je následující:

(A * 2^7) + (B * 2^6) + (C * 2^5) + (D * 2^4) + (E * 2^3) + (F * 2^2) + (G * 2^1) + (H * 2^0)

[proměnné A až H mohou nabývat hodnot 0 či 1], máme tedy 256 hodnot (číslic), však číslované od 0, takže rozsah je prakticky 0 až 255 znaků (interpretované číselně, při čemž tomuto číslu se říká ordinální hodnota znaku) a každý znak má svůj jistý význam – ne vše jsou však tisknutelné znaky.

V počítačích se pro výpis jednotlivých bytů vyjma jejich ordinární hodnoty také využívá hodnoty hexadecimální, jejíž název vzniknul z hexa – šest a decimal – deset, tedy se jedná o 6 + 10 = 16kovou soustavu. Výpočet je velmi jednoduchý, protože platí zákonitost:

(A ^ 16^1) + (B * 16^0)

Z tohoto vztahu vyplývá, že proměnné A a B mohou nabývat hodnot 0 až 15 (jelikož 15 * 16^1 + 15 * 16^0 = 240 + 15 = 255, nejvyšší ordinární hodnota znaku, a zároveň 0 * 16^1 + 0 * 16^0 = 0 + 0 = 0, tedy nejnižší ordinární hodnota znaku).

To by stačilo k teorii o počítačem využívaných numerických soustavách, nyní přejděme k teoriím ohledně genetiky. Ribonukleové kyseliny, tj. DNA (deoxyribonukleová kyselina) či RNA (ribonukleová kyselina), jsou základní stavební bloky živoucích organismů a ve své podstatě se dá říci, že jsou identické v každé buňce těla organismu (pro zjednodušení – ve skutečnosti zde dochází k mutaci buněk věkem a bylo by nutné vyextrahovat zárodečnou buňku z těla, což by bylo sice možné, ale byl by to velmi náročný úkol). Ribonukleové kyseliny jsou tvořeny řetězci tzv. dusíkatých (nukleotidových) bází, mezi které řadíme cytosin, guanin, adenin a pro případ DNA také thymin, v případě RNA je thymin nahrazen uracilem. Jednotlivé báze označujeme prvními písmeny jejich názvů, tedy C pro cytosin, G pro guanin, A pro adenin a T pro thymin (či U pro uracil - pro RNA).

Báze existují tedy 4. Pro zakódování do jednoho bytu do sekvence dusíkatých bází bychom tedy potřebovali definovat čtyřkovou číselnou soustavu. Na základě znalostí hexadecimální a binární soustavy můžeme tedy tuto soustavu definovat velmi snadno. Víme, že minimální hodnotou musí být 0 a maximální hodnotou 255. Prvně si ověříme, zda je předefinování do čtyřkové soustavy vůbec možné, tedy si odmocníme číslo 256 (počet hodnot) čtyřma. Vznikne nám tímto celé číslo 4 (4^4 = 256). Rozklad tedy možný je. Definujeme si postup převodu:

(A * 4^3) + (B * 4^2) + (C * 4^1) + (D * 4^0)

Při čemž možné hodnoty pro proměnné A, B, C a D jsou hodnoty 0 až 3 (číslujeme vždy od 0), pro ověření si můžeme udělat výpočet: 3*4^3 + 3*4^2 + 3*4^1 + 3*4^0 = 192+48+12+3=255.

Tímto však dostáváme jen zakódované řetězce znaků do čtyřkové soustavy a ne samotnou reprezentaci v dusíkatých bázích dle jejich označení. Proto si přiřadíme každé možné hodnotě nějakou bází, třeba pro 0 přiřadíme C, 1 bude G, 2 bude A a 3 bude T (nebo U pro RNA). Z jednoho bytu nám tedy vznikl řetězec (sekvence) o 4 bázích.

Jakákoli počítačová informace (vzhledem ke své reprezentaci pomocí sekvence bytů) může být tedy zakódována do sekvence dusíkatých bází. Nicméně nám nastává nyní jiná otázka? Je možné tyto báze kombinovat v libovolném pořadí? Uvažme situaci, kdy se v našem těle vytváří (či lze uměle syntetizovat) aminokyseliny. Z čeho jsou tvořeny aminokyseliny? Právě z dusíkatých bází a každá aminokyselina (dále jen AMK) je tvořena tripletem bázi tvořící kodón, triplet bází, jak již název napovídá je sekvence 3 bází a tedy dostáváme teoretické číslo 64 (4^3). Nicméně žádná odmocnina čísla 256 nám nedá číslo 64, proto je nutné využít sekvence více bytů pro vytvoření více kodónů, čímž se postupně dopravujeme ke vztahu 64^4 = 256^3 (16777216, dle jednotek bychom toto v bytech popsali jako hodnotu 16 MB). Tedy uvažme počítačovou aplikaci, která čte soubor dat po 3 bytech a každé 3 byty implementuje jako 4 kodóny (implementované jako čísla se spodní hranicí 0 a horní hranicí 63). To uděláme následovně: uvážíme ordinární hodnoty všech 3 přečtených znaků k získání pomocné součtové hodnoty, např. tedy pro ordinární hodnoty 125, 132 a 165 získáme číslo 125 * 65536 + 132 * 256 + 165 = 8225957.

[obecně tedy A * 256^2 + B * 256^1 + C * 256^0]

Nyní je nezbytné provést zpětný rozklad, však do naší kodónové soustavy dle algoritmu:

A * 64^3 + B * 64^2 + C * 64^1 + D * 64^0

vznikne nám:
A = 31, B = 24, C = 18, D = 37

Tedy rozklad 3 bytů do sekvence 4 kodónů již máme hotov. Celou proceduru budeme opakovat pro všechna data, která chceme zakódovat.

Dostáváme se k závěru, který je bohužel zatím (doufám, že toto slovo zde zmizí někdy) nedořešený. Bylo by totiž k němu nutné sestavit zařízení, které na základě vstupních hodnot (které jsou zde pojmenované jako proměnné A,B,C a D) dokáže sestavit nějaký organismus. Toto by bylo velmi složité a prakticky by zde mohlo pomoci velmi genetické inženýrský, nicméně mluvíme o vytváření umělých bytostí, které by mohly být na základě svých stavebních prvků živé. Toto zařízení by muselo zvládat i instrukce opačné – tedy takovýto organismus „přečíst“ a rozložit na jednotlivé báze, které by přes definované rozhraní opět přeneslo do připojeného počítače a aplikace na tomto PC by dešifrovala data do potřebné binární podoby.

Však kdyby takové zařízení existovalo, bylo by teoreticky možné provádět na jeho základě teleportaci. Člověk by byl rozložen do sekvence kodónů a AMK a přenesen z bodu A např. formou posílání dat přes internet do bodu B, kde by stejné zařízení člověka složilo do jeho původní podoby. Problémem je, že originál by musel být zničen, jinak by došlo ke klonování jedince.

Samotná teleportace však nese i jiná úskalí, např. co si lze představit pod pojmem duše a jak by se ta přenesla? Jak by se přenesl lidský engram a vzpomínky? Toť otázka, aneb: 2B || !2B = ? (☺)

Uvážíme-li reálné pokusy ohledně kvantové teleportace, musíme se zabývat i dalšími věcmi jakým jsou např. superpozice stavů a další věci z hlediska kvantové fyziky. Fyzika je velmi spojená s biologií i vzhledem k tomu, že se jedná o vědu podobnou matematice, je alfou i omegou všeho vědění, stejně jako matematika. Sama o sobě kvantová fyzika má navíc i jiné stavy než jsou popsané ve výše uvedených stránkách, např. má tzv. qubit (kvantový bit, při čemž bit je též jednotkou informace, jak je již popsáno na straně 2 tohoto dokumentu). Qubit je analogií k výše uvedenému bitu a je popsán stavovým vektorem dvouvrstvého kvantově mechanického systému (což je prakticky jen systém o 2 stavech, jak již je výše uvedeno – stavy označujeme jako 0 či 1, popř. nenastaveno a nastaveno, resp. ang. Set a not set). Nicméně samotný qubit zahrnuje i superpozice těchto stavů, čímž rozšiřuje jejich možný rozsah. Kvantové mechaniky a qubitů lze využít pro tzv. kvantovou kryptografii (kryptografie je způsob ochrany dat před zneužitím, především pomocí šifrování a podobných algoritmů). Vzhledem k tomu, že sama o sobě kvantová kryptografie využívá jevu známého jako kolaps vlnové funkce při svém měření, byla by kvantové zašifrovaná informace bezpečná i za předpokladu, že někdo odchytí naši komunikaci, jelikož by nemohl zákonitě získat kompletní validní informaci. Pravdou je, že bychom tuto informaci na základě kolapsu vlnové funkce nezískali ani my, nicméně bychom jako oprávnění příjemci o informaci mohli zažádat znovu. Jelikož by data nedošla v pořádku, měli bychom dobrou metodu na zjištění, zda naše komunikace byla odposlouchávána či ne. Při důkazu odposlouchávání bychom využili jiné metody přenosu dat a vyžádali si informaci znovu (v naději, že tato druhá cesta nebude kompromitována jako cesta první). Útočník sám o sobě by nezískal tedy také kompletní informaci a vzhledem ke kolapsu vlnové funkce by ani nemohl vědět, zda jím získaná informace byla získána v pořádku či nikoli, tedy ani kompletní odposlech komunikace by mu nemusel být nic platný, jelikož citlivé údaje, např. hesla pro přístup k bankovním účtům by byla přenesena v nečitelné podobě. Metoda kvantové kryptografie zpravidla využívá tzv. propletenosti fotonů (angl. photon entanglement), což je metoda spočívající v propletenosti dvojice částic, kdy nemluvíme o stavu jednotlivých částic jako společně o stavu částic. Měření na jedné částici z dvojice tímto způsobí kolaps vlnové funkce, kterou je popsána daná dvojice a to změní stav druhé částice, tedy provede změnu qubitu pro danou dvojici částic. I když si mnozí čtenáři tohoto dokumentu mohou myslet, že je o kompletní science-fiction, pravda je odlišná. Od října roku 2008 existuje projekt SECOQC (Secure Communication based on Quantum Cryptography, tj. Bezpečná komunikace založená na Kvantové Kryptografii, stránky projektu jsou k nalezení na http://www.secoqc.net), do které je zapojena i Česká republika (ač jen Palackého univerzita v Olomouci). Samotný projekt pojednává především o využití QKD (Quantum Key Distribution, kvantová distribuce klíče), jelikož objem dat přenášených běžnými uživateli může být enormní, proč tedy nevyužít jen náhodně vygenerovaného klíče pro každý přenos? Prakticky to jde přirovnat k SSL (Secure-Socket Layer) vrstvě v informačních technologiích. Jak tato komunikace probíhá? Je položen dotaz na stránku, která je zabezpečená SSL vrstvou. Stránka pošle do klienta (prohlížeče) veřejnou část klíče, který je použit na zašifrování dat pomocí asymetrické kryptografie a dešifrován na straně serveru pomocí privátního klíče. Samotná komunikace již poté probíhá na šifrovaná tímto způsobem. Proč tedy nenahradit SSL certifikáty kvantově přenášenými klíči za využití QKD metody? Zřejmě někoho z čtenářů napadne, zda to vážně je potřebné, pokud již existuje právě zmiňovaná SSL vrstva. Odpověď zní, že stále je, protože SSL vrstva je prolomitelná a má-li potenciální útočník dost potřebných vzorků dat, může sestavit komunikaci a rozšifrovat i data přenesená přes tuto vrstvu – práce to není jednoduchá ani v žádném případě rychlá, nicméně zde riziko existuje. V případě využití kvantové distribuce klíče zde toto riziko opadá kvůli náhodnosti generování klíče, který bude bezpečně přenesen přes kvantovou datovou síť a jehož odchycení bude znamenat kolaps vlnové funkce, na základě čehož poté ani jedna strana nebude mít klíč k dešifrování. Oprávněný uživatel si může opětovně poté zažádat o klíč (nejlépe nějakou jinou metodou, která není kompromitována), nicméně útočník v tomto

kontextu nemůže nic. Ani nemá možnost odposlechem zjistit, jaký je správný klíč např. pro prolomení bezpečnosti nějakého bankovního účtu, což z pochopitelných důvodů zvyšuje bezpečnost informací. Samotná stránka SECOQC obsahuje velmi zajímavé informace o tomto tématu, což je věc, kterou jedině doporučuji přečíst, pokud někoho tato problematika zajímá. Stránka je v anglickém jazyce, nicméně většina dostupných informací o těchto technologiích také, ale to snad žádného potenciálního vědce neodradí :)