

Modelování evolučních procesů

Radek Pelánek

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Změny v systémech

- dosavadní modely měly **statickou** strukturu (pravidla)
- komplexní systémy se **mění**: vývoj, adaptace, učení, ...
- vztah k horizontu modelu
 - vlk: reintrodukce, lov ve smečce
 - epidemie

Změny v systémech

system	rychlost zmen
nervovy system	sekundy az hodiny
imunitni system	hodiny az dny
firma	mesice az roky
zivoctisny druh	dny az století
ekosystem	roky az milenia

- evoluce – tato přednáška
- myšlení – další přednáška

- 1 pochopení přírody – jak funguje myšlení, evoluce, ...
- 2 lepší modely komplexních systémů – obohacení modelů s agenty
- 3 řešení náročných algoritmických problémů – modely jako výpočetní mechanismy

Evoluce

- **základní myšlenka** evoluce je (alespoň zdánlivě) **jednoduchá**
- **důsledky** jsou však často **komplikované a neintuitivní**
- tato přednáška:
 - výběr zajímavostí, principů
 - témata se vztahem k výpočetním modelům
- rozhodně **ne** vyčerpávající pojednání o evoluci

Evoluce: citáty

Slepice je jenom způsob, jakým vajíčko vyrábí další vajíčko.
(Samuel Butler)

Přírodní výběr je mechanismus pro generování mimořádně velkého stupně nepravděpodobnosti. (Sir Ronald Fisher)

Nothing in biology makes sense except in the light of evolution. (Theodosius Dobzhansky)

Vybrané historické poznámky

- cca 1800: Lamarck, vývoj s děděním získaných vlastností
- 1859: Darwin, Origin of species, změny + přírodní výběr
- 1865: Mendel, dědičnost
- poč. 20. stol: 'modern synthesis' předchozích dvou
- 1953: struktura DNA
- 1976: Dawkins, Selfish gene

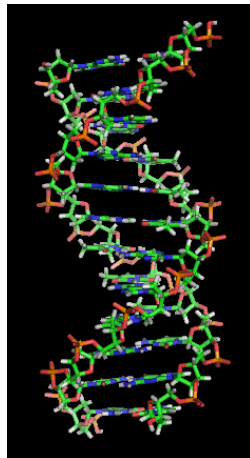
Příklad můry

- evoluce je vesměs pomalá (rychlost \sim generacím), nejde moc pozorovat
- známá výjimka: můry v okolí Manchesteru
- průběh:
 - světlá kůra stromů \rightarrow světlé můry
 - znečištění \rightarrow tmavá kůra \rightarrow tmavé můry
 - redukce znečištění \rightarrow zpět světlá kůra \rightarrow světlé můry
- NetLogo model: Biology / Evolution / Peppered Moth



Genetická informace

- genetická informace uložena v DNA
- DNA = řetězec „základních bloků“: adenine (A), cytosine (C), guanine (G), thymine (T)
- při reprodukci dochází ke kombinaci (křížení) DNA rodičů
- náhodné mutace



Základy evoluce: poznámky

- **vyvíjí se populace** jako celek (nikoliv jednotlivci)
- **diverzita** je důležitá
- schopnosti získané během života se nepřenáší (až na detaily)
srovnej: Lamarck, „kulturní evoluce“ (memy), simulovaná evoluce
- přirozený výběr \sim usměrňování vývoje

Přirozený výběr

- **přežití nejsilnějších** (survival of the fittest) nebo spíš **přežití schopných reprodukce** (survival of the reproducers)
- triviální, ale zajímavé pozorování:
Všichni naši předci žili tak dlouho, že se stihli reprodukovat.
- pozitivní zpětná vazba

Koevoluce

- organismy se adaptují na měnící se okolí
- organismy současně i ovlivňují okolí
- „Red Queen effect“ (Alenka v říši divů):
Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!

Koevoluce: biologické závody ve zbrojení

- **netopýři** mají **sonar**, kterým hledají můry (sonar je sám o sobě zapeklitě komplikovaný)
- **můry** vyvinuly **měkké pokrytí** těla, které absorbuje netopýří vysílání
- **netopýři** přešli na nové **frekvence**
- **můry** přišly s novým pokrytím a s „**rušičkou**“ (vlastní signál který interferuje s netopýřím)
- **netopýři** přišli s novými leteckými **manévry** a naučili se **vypínat** sonar (čímž dělají rušení méně efektivním)

Koevoluce: biologický citát

Na počátku byla jemná křehká bylinka, kterou občas někdo sežral; na konci je trnitá a jedovatá obluda, kterou také občas někdo sežere. (J. Zrzavý, D. Storch, S. Mihulka)

Koevoluce v IT

- viry a antivirová ochrana
- spamy a antispamová ochrana

Typy evolučních modelů

- přežití nejsilnějších (bez vývoje) – „ekologické modely“ (např. Dilema vězně a třetí Axelrodův turnaj)
- přežití nejsilnějších + vývoj (křížení, mutace)
 - genetické algoritmy
 - evoluční programování
 - genetické programování
 - ...

Interpretace evolučních principů

Lze interpretovat různě:

- „nejsilnější“ mají největší šanci reprodukce
- učení se metodou pokus-omyl (a pamatování si úspěšných)
- nápodoba úspěšných

není potřeba „racionalita“ agentů, uvědomění si, co přesně dělají (srovnej dedukce, příklad piráti)

Genetické algoritmy: základní přístup

- Vyber počáteční populaci P
- Opakuj dle potřeby:
 - Vytvoř novou prázdnou populaci P'
 - Opakuj dokud P' není plná:
 - Vyber dva jedince z populace P v závislosti na kritériu **zdatnosti**
 - Volitelně: **křížení** a nahrazení potomky
 - Volitelně: **mutace**
 - Přidej do populace P'
 - $P := P'$

Reprezentace

- jedinci \sim řetězce (většinou binární)
- případně vyžadovány speciální vlastnosti řetězců (např. permutace čísel)

Výběr

- **výběr** založen na **zdatnosti** (fitness):
 - **absolutní** zdatnost (např. optimalizační problémy)
 - **relativní** zdatnost – necháme jedince spolu „soutěžit“
- **deterministický** výběr nejlepších:
 - ztráta diverzity
 - uváznutí na lokálním minimu
- používá se proto **náhodnostní** výběr s přihlédnutím k zdatnosti (ruleta)

Křížení

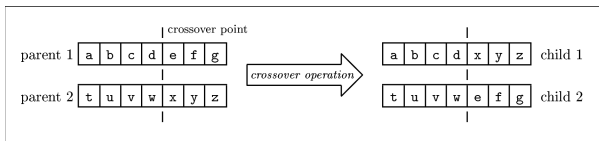


Figure 20.1 The crossover operation in action

- variace: vícebodové křížení, zachování základních bloků
- někdy speciální formy křížení, abychom zachovali požadovanou vlastnost řetězců

Mutace

- náhodná změna v řetězci, opět mohou být potřeba speciální úpravy
- většinou nevede ke zlepšení
- může pomoci překonat lokální optima
- používáno, ale s malou pravděpodobností

Praxe: volba parametrů

- základní princip genetického algoritmu jednoduchý
- avšak mnoho (skrytých) parametrů a voleb

volba reprezentace, funkce zdatnosti, přesný mechanismus křížení, mutace, výběru podle zdatnosti,

velikost populace, počet generací, pravděpodobnost mutace, ...

- není jednoduché to „rozchodit“
- existují heuristiky pro volbu parametrů, např. pro velikost populace:
 - velikost populace \times počet generací $> 100\,000$
 - velikost populace \gg počet genů

Ilustrační příklad: Hledání řetězce

Time	Average Fitness	Best Fitness	Best String
0	0.035314	0.200000	"pjrmrubynrksxiidwctxfodkodjjzfunpk "
1	0.070000	0.257143	"pjrmrubynrksxiidnybvsqc piisyexdt"
:	:	:	:
25	0.708686	0.771429	"qurmous gresn idnasvsweqt prifuseky"
26	0.724286	0.800000	"qurmous green idnasvsweqt prifuseky"
:	:	:	:
36	0.806514	0.914286	"uurious green idnas sweqt profusely"
37	0.820857	0.914286	"qurmous green ideas sweqt profusely"
:	:	:	:
41	0.895943	0.942857	"uurious green idnas sweat profusely"
42	0.908457	0.971429	"qurious green ideas sweat profusely"
:	:	:	:
45	0.927714	0.971429	"qurious green ideas sweat profusely"
46	0.936800	1.000000	"furious green ideas sweat profusely"

Table 20.2 Evolving the text string "furious green ideas sweat profusely" from an initially random pool of strings: Fitness scores shown are for the raw fitness.

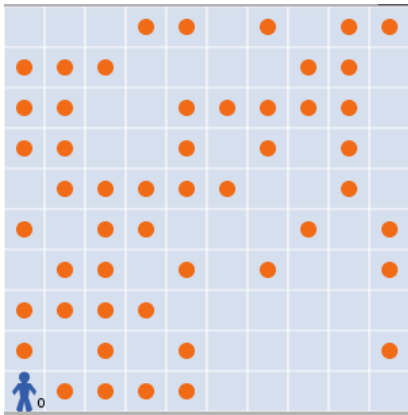
Demo příklady

- NetLogo:
 - Computer Science / Simple Genetic Algorithm
 - Biology / Evolution / Sunflower biomorphs
- mnoho demo příkladů (Java aplety) na Internetu — stačí hledat „Genetic algorithm“
- typická ilustrace: problém obchodního cestujícího (TSP, traveling salesman problem)

Robot sbírá jídlo

- NetLogo: Computer Science / Robby the Robot
- mřížka, zdi, jídlo
- vstup: robot vidí bezprostřední okolí
- výstup: přesun na vedlejší pole
- úkol: posbírat co nejvíce jídla
- genetický algoritmus zvládne vymyslet neintuitivní „finty“

Robby the Robot



Robby the Robot

```

generation 54:
  best fitness = 29.5
  best strategy: [+ ● + † - x ● + † x x † - † - x ● + - † x + † ● + † - † -
generation 55:
  best fitness = 31
  best strategy: [+ ● + † - x ● - † x x † - † - x ● + - † x + † ● + - † -
generation 56:
  best fitness = 30.5
  best strategy: [+ ● + † - x ● + † x x † - † - x ● + - † x + † ● + † - † -
generation 57:
  best fitness = 32
  best strategy: [+ ● + † + x ● + † x x † ● † - † - x ● + - † † + † ● + † - † -
generation 58:
  best fitness = 30.5
  best strategy: [+ ● + † - x ● + † x ● † - † - † - x ● ● - † x + † ● + - † -
generation 59:
  best fitness = 33
  best strategy: [+ ● + † - x ● + † x ● † - † - † - x ● ● - † x + † ● - † - † -
generation 60:
  best fitness = 34.5
  best strategy: [+ ● + † + x ● + † x x † ● † - † - x ● + - † x + † ● + † x † -
generation 61:
  best fitness = 35.25
  best strategy: [+ ● + † - x ● + † x ● † - † - † - x ● ● - † x + † ● + - † -

```

Optimalizační problém

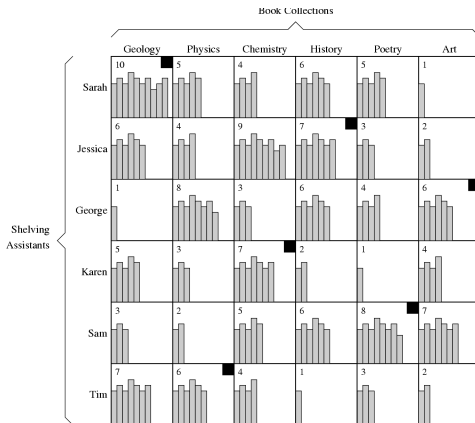


Figure 18.6 The task assignment problem: Black squares in the entries denotes the optimal assignment with a total shelving rate of 44.

Optimalizační problém

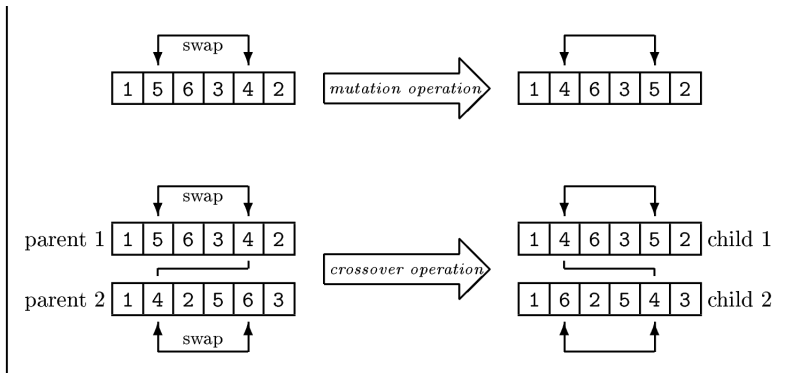


Figure 20.3 The crossover and mutation operations applied to candidate solutions of a combinatorial optimization problem

Evoluční dilema vězně: Axelrodova studie

- otázka: jak moc byly výsledky turnajů ovlivněny tím, že lidé očekávali určité složení odeslaných strategií?
- strategie uvažující poslední 3 tahy, začíná z náhodných
- strategie, které se vyvinou připomínají charakteristiky TFT – tj. dominance principů, na kterých je TFT založena není způsobena lidskými očekáváním, kulturními hodnotami, ...

Dilema vězně

- strategie zohledňující posledních k kol – snadná reprezentace řetězcem
- např. pro $k = 1$ řetězec 5 znaků:
 - 1 tah v prvním kole
 - 2 co dělat když minule: oba spolupracovali
 - 3 co dělat když minule: já spolupracoval on zradil
 - 4 co dělat když minule: já zradil on spolupracoval
 - 5 co dělat když minule: oba zradili
- např. TFT je “SSZSZ”
- zdatnost = průměrný bodový zisk

Dilema vězně

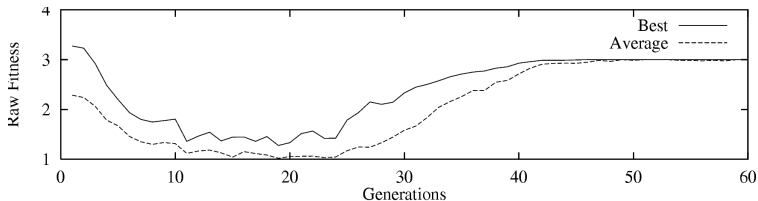


Figure 20.4 Average and best raw fitness scores for the IPD-playing GA

G. Flake, *The Computational Beauty of Nature*

Další aplikace genetických algoritmů

- optimalizační problémy: rozvrhy, protein folding
- učení a plánování: robotika, hry
- návrh: hardware, materiály

Evoluce a míra zdatnosti

- „přežití nejsilnějších“ – ale co to znamená „nejsilnější“?
- existuje absolutní míra zdatnosti? **ne**
- zdatnost organismu závisí na prostředí, ve kterém se nachází a které sám ovlivňuje (viz koevoluce)
- **evolučně stabilní strategie**

Evolučně stabilní strategie

- pokud by do populace tvořené čistě touto strategií přišla cizí strategie, tak by se nerozšířila
- zjemnění Nashova ekvilibria (teorie her)

Jestřáb a holubice

- agenti bojují o zdroje, úspěšní se množí, neúspěšní vymírají
- každý dodržuje jednu ze dvou strategií:

jestřáb vždy útočí, ze souboje utíká jen při velkých poraněních

holubice nikdy neútočí (vyčkává dokud to druhý nevzdá), ze souboje utíká

Jestřáb a holubice: tabulka

Model teorie her, **průměrný** zisk ze vzájemného setkání:

	jestřáb	holubice
jestřáb	-2; -2	10; 0
holubice	0; 10	3; 3

- konkrétní hodnoty nejsou příliš důležité, hlavní je uspořádání hodnot
- jde v podstatě o hru „Kuře“ (viz slidy Spolupráce)

Kdo je nejsilnější?

- Kdo je „silnější“? Co je výhodnější strategie: jestřáb nebo holubice?

Kdo je nejsilnější?

- Kdo je „silnější“? Co je výhodnější strategie: jestřáb nebo holubice?
- nelze jednoduše odpovědět – záleží na složení populace
- převládají jestřábi \Rightarrow je lepší být holubice
- převládají holubice \Rightarrow je lepší být jestřáb

Analýza hry

- podíl jestřábů v populaci je p (a podíl holubic tedy $1 - p$)
- průměrný zisk jestřába je:
$$z_j = -2p + 10(1 - p) = 10 - 12p$$
- průměrný zisk holubice je:
$$z_h = 0p + 3(1 - p) = 3 - 3p$$
- převažují jestřábi (p je blízko jedné) \Rightarrow je výhodnější být holubicí ($z_h > z_j$)
- převažují holubice (p je blízko nuly) \Rightarrow je výhodnější být jestřábem ($z_h < z_j$)

Evolučně stabilní strategie

- jestřáb ani holubice nejsou evolučně stabilní
- jaká strategie je evolučně stabilní?

Evolučně stabilní strategie

- jestřáb ani holubice nejsou evolučně stabilní
- jaká strategie je evolučně stabilní?
 - mixovaná strategie: chovej jako jestřáb s pravděpodobností $7/9$, chovej se jako holubice s pravděpodobností $2/9$
 - odvetník: neútočí, ale útoky oplácí

Analýza hry

- lze realizovat systémovou dynamikou (Stella) i pomocí agentů (NetLogo)
- analýza pevných bodů, dynamiky
- jeden z námětů na projekty

Altruismus a evoluce

altruismus – silný, slabý; komplikovanější pojem než se zdá
altruistické chování:

- zisk pro ostatní
- ztráta pro mě (nebo neutrální stav)

Může být altruismus (evolučně) výhodný? Jak se mohlo vyvinout a udržet altruistické chování při „výběru nejsilnějších“?

Altruismus a evoluce

roli hraje více jevů, zde na modelech ilustrujeme dva:

- **vliv podmínek prostředí**
- **populační viskozita** (omezené šíření populace v prostoru)

Altruismus – buněčný automat

- NetLogo: Social Science / Altruism
- 2D model, pravděpodobnostní buněčný automat
- každá buňka používá jednu ze dvou strategií: **altruistická**, **sobecká**
- výběr strategie na další kolo: losováním s přihlédnutím k úspěšnosti **agentů v okolí** („ruleta“)
- podmínky prostředí: **nemoci**, **využitelnost**

Altruismus: ohodnocení

- c „cena altruismu“, jak moc mě altruistické chování stojí
- b „zisk z altruismu souseda“, jak moc mi pomůže altruistické chování mého souseda
- N_A počet altruistických buněk v okolí, včetně sebe samé

zisk:

- pro sobeckou buňku: $1 + b \cdot N_A/5$,
- pro altruistickou buňku: $1 - c + b \cdot N_A/5$.

Altruismus: analýza modelu

- za **dobrých** podmínek prostředí dominuje **sobecká** strategie
- za **zhoršených** podmínek prostředí dominuje **altruistická** strategie

Krávy: popis modelu

- NetLogo: Social Science / Cooperation
- model s agenty: krávy žerou travu
- **různá rychlost dorůstání trávy** dle délky
- **sobecké** krávy: sežerou všechnu travu, co je k dispozici
- **spolupracující** krávy: vždycky trochu trávy nechají (aby rostla rychleji)

Chování

- dle podmínek prostředí (rychlost růstu trávy, energie trávy, rychlost přesunu krav) vítězí buď sobecké či spolupracující
- větší „populační viskozita“ \Rightarrow spolupráce

Otevřená evoluce

- evoluce nesměřuje k předem danému „vrcholu“
- **koevoluce** – každý se snaží „vylézt“ co nejvýš v aktuální krajině (která se však mění)
- modely evoluce (např. genetické algoritmy) – většinou dodáváme **externí** míru zdatnosti
- jak modelovat opravdu otevřenou evoluci?

Tierra

- <http://life.ou.edu/tierra/>
- model digitálních organismů
- prostředí = virtuální počítač
- organismy = jednoduché programy
- soutěží o dostupné zdroje:
 - procesorový čas \sim energie
 - procesorová paměť \sim materiál

Programy

- sekvence jednoduchých instrukcí: 32 různých 5-bitových instrukcí v **zjednodušeném assembleru**
- **prapředeck** – jednoduchý program, který **kopíruje sám sebe** (ručně vytvořen)
- dál klasický genetický algoritmus (křížení, mutace)
- kopírování a míra **zdatnosti** zadány **implicitně** – zdatný je ten, kdo se zvládne množit

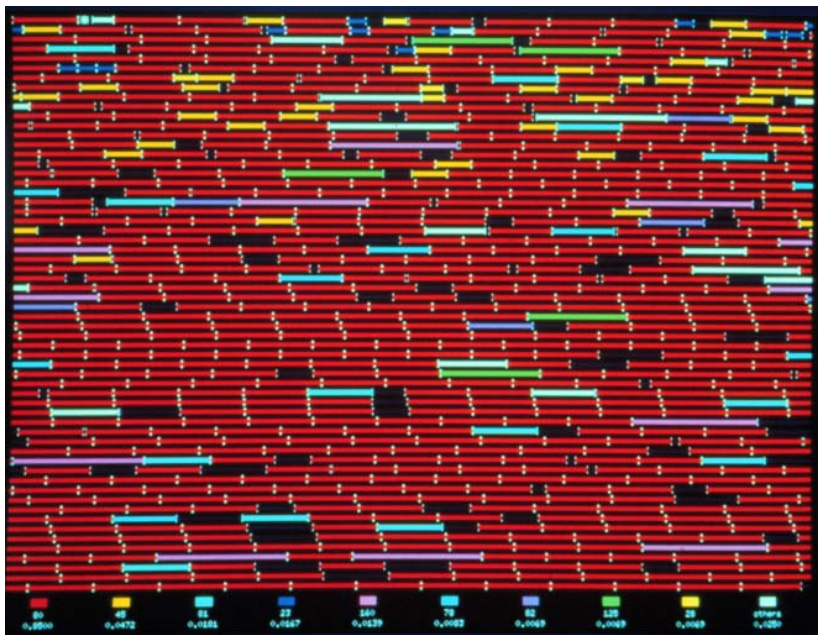
Dynamika

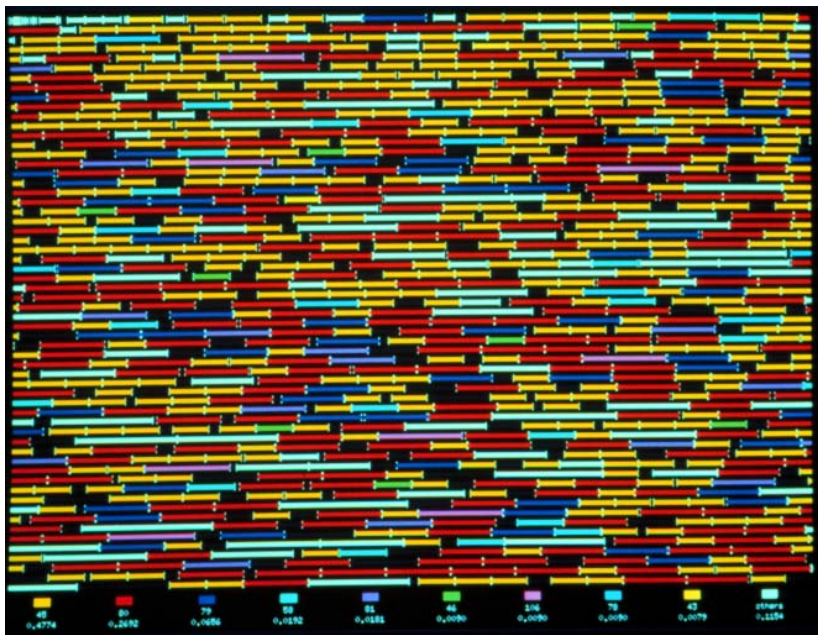
- paraziti
 - využívají pro svoje kopírování zdrojový kód jiných programů
 - jejich vlastní kód je kratší a tím jsou úspěšnější
- hyperparaziti
 - zneužívají parazitů
 - využijí toho, že parazit zavolal jejich kód, a pak už mu volání nevrátí zpět

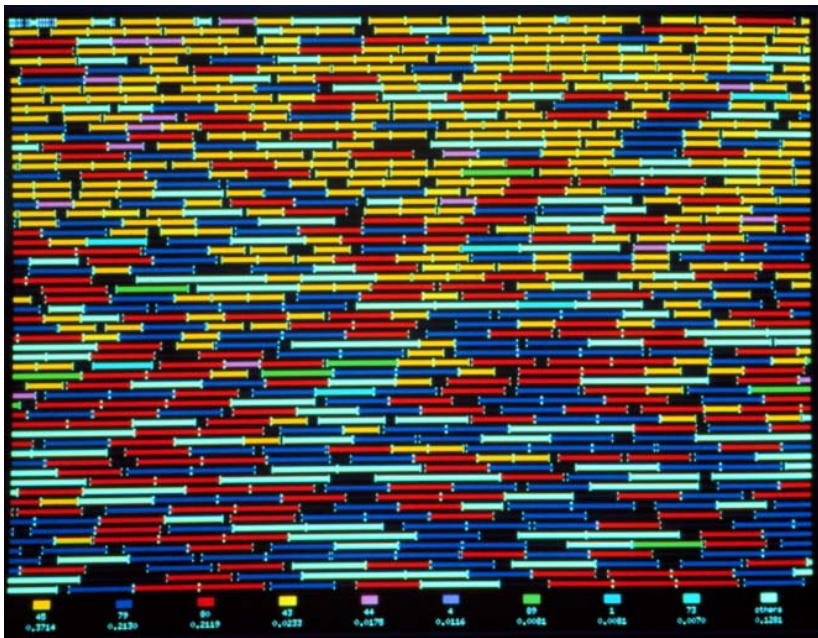
Ilustrace

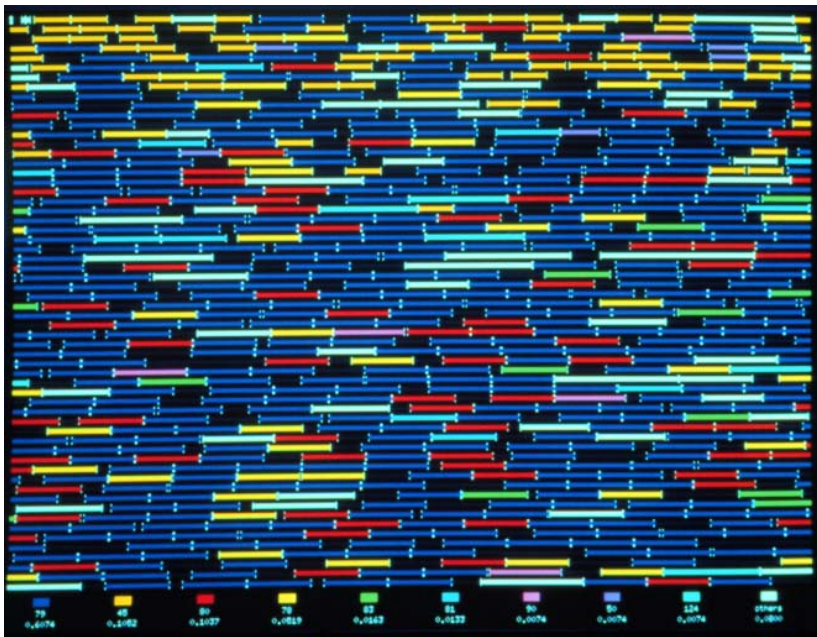
Následující ilustrace:

- červená: základní organismus
- žlutá: parazit
- modrá: imunní organismus









Tierra: shrnutí

- bohatá a zajímavá dynamika
- existují rozšíření
- ale i tak se vývoj vždy zastaví
- nedochází tedy k opravdu otevřené evoluci

Shrnutí

- pojmy, principy:
 - koevoluce
 - evolučně stabilní strategie
 - evoluce a altruismus
 - otevřená evoluce
- modelovací přístupy:
 - genetické algoritmy
 - agenti
 - teorie her
 - matematické analýzy, systémová dynamika