



# *Matematika v biologiji: iskanje popolnih filogenetskih dreves*

Martin Milanič

Fakulteta za matematiko, naravoslovje in informacijske tehnologije  
Univerza na Primorskem

MARS 2009, Koper  
19. avgust 2009

# Razvoj moderne biologije

- ◆ Watson in Crick (1953):
  - **DNK** ima obliko dvojne vijačnice
  - vsaka od obeh strani enolično določa drugo.

No. 4356 April 25, 1953

NATURE

737

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

<sup>1</sup> Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1925).

<sup>2</sup> Lougnot-Higgins, M. S., *Mon. Not. Roy. Astron. Soc., Geophys. Supp.*, **5**, 285 (1949).

<sup>3</sup> Van der Waals, J. D., *Woods Hole Papers in Phys. Oceanogr. Meteor.*, **11** (5) (1950).

<sup>4</sup> Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

## MOLECULAR STRUCTURE OF NUCLEIC ACIDS

### A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey<sup>1</sup>. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining  $\beta$ -D-deoxyribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's<sup>2</sup> model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis.

is a residue on each chain every 3.4 Å. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 Å. The distance of a phosphorus atom from the fibre axis is 10 Å. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally<sup>3,4</sup> that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data<sup>5,6</sup> on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

# *Razvoj moderne biologije*

- ◆ **Razvoj biologije** vse bolj temelji na uporabi matematičnih modelov in dobrih računalniških orodij.
- ◆ **Bioinformatika:** razvijajoča se interdisciplinarna znanost, ki proučuje razvoj in uporabo matematičnih modelov in učinkovitih algoritmov za reševanje bioloških problemov.

# *Evolucija*

- ◆ Vse vrste organizmov na Zemlji se (zelo počasi) razvijajo skozi čas. Temu procesu pravimo **evolucija**.
- ◆ Eden osrednjih bioloških problemov:  
*Za dano množico (obstoječih) živalskih ali rastlinskih vrst ugotovi, kako so se razvile iz skupnega prednika.*

# *Filogenetska drevesa*

- ◆ Tak razvoj po navadi ponazorimo s t.i. *razvojnimi ali filogenetskimi drevesi*.
- ◆ Kadar so drevesa preveč omejujoča, uporabimo namesto njih t.i. **filogenetska omrežja**.

## ***Problem rekonstrukcije filogenetskih dreves:***

*Za dano množico obstoječih živalskih  
ali rastlinskih vrst konstruiraj  
filogenetsko drevo, ki ponazarja razvoj  
danih vrst iz skupnega prednika.*



# *Prve evolucijske študije*

- ◆ Od Darwina pa vse do 60-tih let prejšnjega stoletja so znanstveniki ugotavljali razvojne relacije med vrstami **na osnovi njihovih anatomskih in morfoloških lastnosti**
- ◆ Take študije so temeljile predvsem na subjektivnih ocenah.  
*Posledično:*
  - *nepopolni zaključki*
  - *nepravilni zaključki*



## *Modernejši pristopi*

- ◆ Pred približno 40 leti sta Emile Zuckerkandl in Linus Pauling predlagala, da bi morala razvojna drevesa temeljiti na podobnostih in razlikah v **zaporedjih nukleotidov v DNK**
- ◆ Sledile so vroče razprave, ali je kaj takega smiselno.
- ◆ Danes pa je to osrednji pristop za študij evolucije.

# *Modernejši pristopi*

- ◆ Rekonstrukcija filogenetskih dreves danes temelji predvsem na:
  - ◆ **zaporedjih nukleotidov v DNK**
  - ◆ **zaporedjih genov.**
- ◆ Zaporedja so seveda ***matematični objekti.***
- ◆ Pri obeh metodah so **matematični modeli** in **računalniška orodja** bistvenega pomena.
- ◆ Za ogrevanje si pogledjmo dva primera!

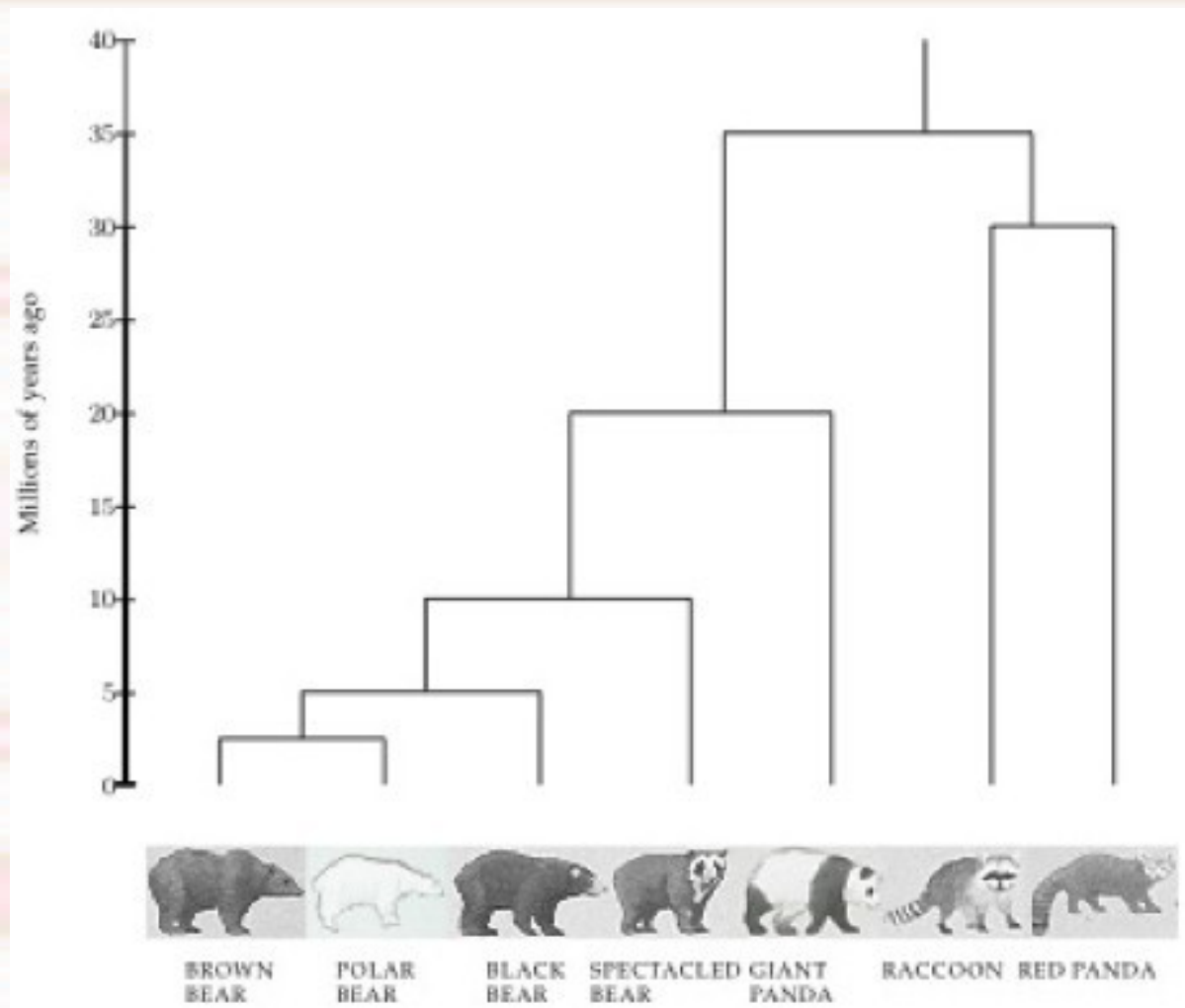
# *Panda: medved ali rakun?*

- ◆ Skoraj sto let so bili znanstveniki v dvomih, kateri družini pripada **veliki panda**.
- ◆ Veliki pande so
  - na videz *podobni medvedom*,
  - imajo *lastnosti, značilne za rakune* (npr.: pande ne poznajo zimskega spanja).

## *Panda: medved ali rakun?*

- ◆ Medtem ko je veliki panda veselo prebival v gozdovih Sečuana, je bil dolga desetletja predmet vročih razprav med biologi...
- ◆ ... vse dokler ni leta 1985 skupina znanstevnikov (Steven O'Brien in njegovi kolegi) rešila uganko - s pomočjo **analize DNK**.

# *Razvojno drevo medvedov in rakunov*



<http://bix.ucsd.edu/bioalgorithms/downloads/bookfigs/>

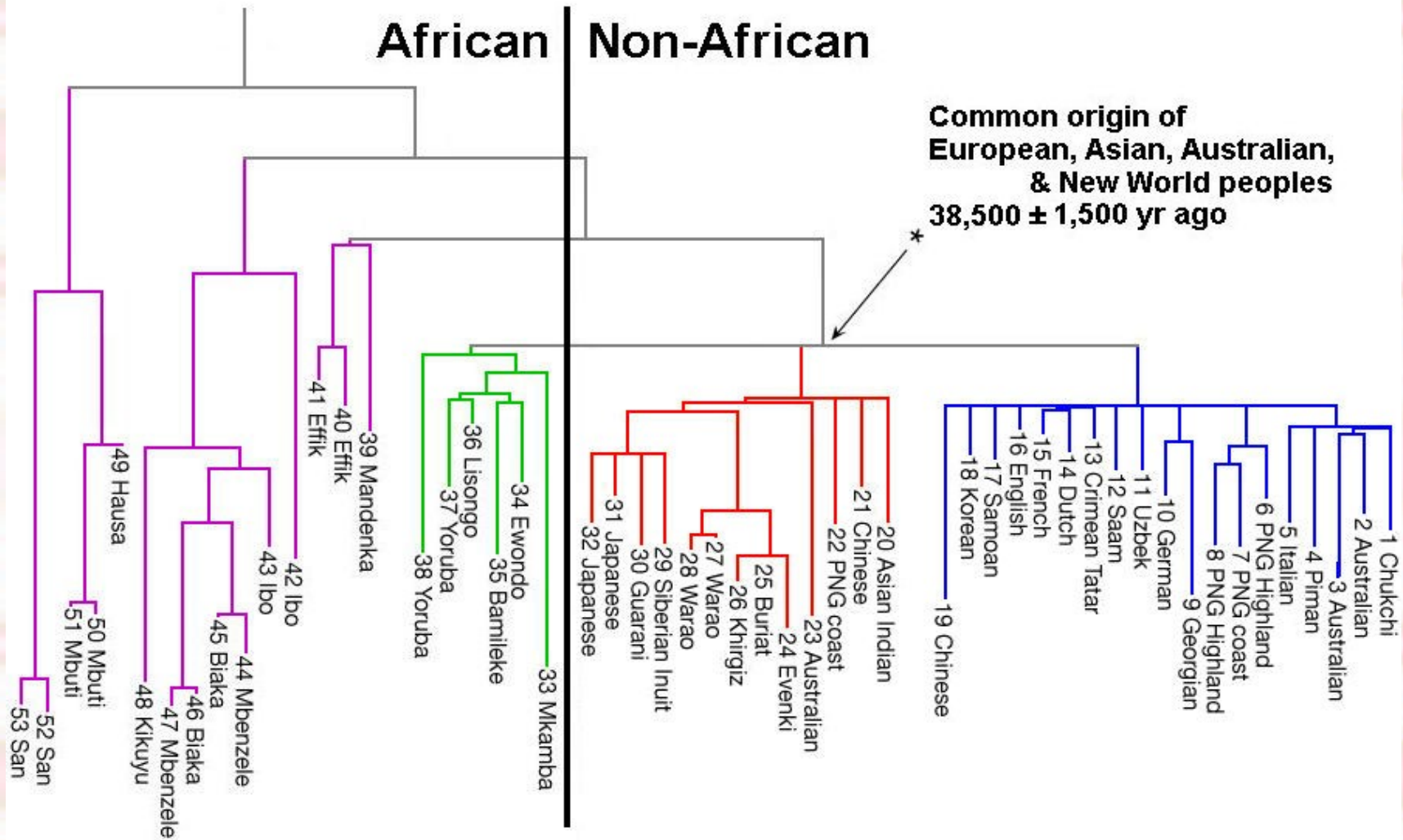
## *Od kod izvira skupni prednik vseh človeških ras?*

- ◆ Približno v času, ko so znanstveniki rešili uganko velikega pande, pa je prišla na dan nova uganka.
- ◆ Rekonstrukcija razvojnega drevesa človeških ras s pomočjo **analize DNK** je postavila naslednjo domnevo:  
*Skupni prednik vseh človeških ras je živel v Afriki pred približno 200.000 leti.*

## *Od kod izvira skupni prednik vseh človeških ras?*

- ◆ Ta domneva je v nasprotju s prej veljavno domnevo, ki pravi, da se je človek razvil pred 2 milijonoma let istočasno v različnih delih sveta.
- ◆ Znanstveniki so sedaj zedinjeni, da je domneva o afriškem izvoru najverjetneje pravilna.

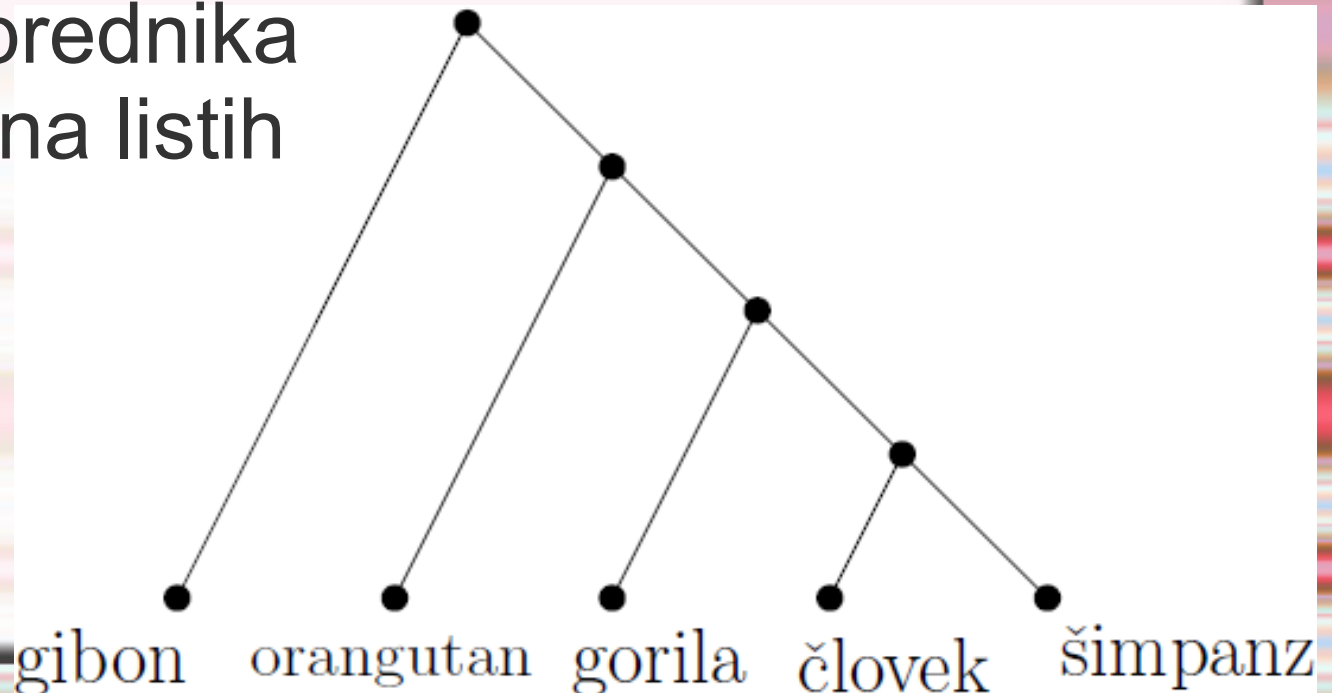
# Razvojno drevo človeških ras





# *Kako iz zaporedij DNK gradimo filogenetska drevesa*

- ◆ listi predstavljajo obstoječe organizme
- ◆ notranja vozlišča predstavljajo njihove prednike
- ◆ koren predstavlja skupnega razvojnega prednika organizmov na listih



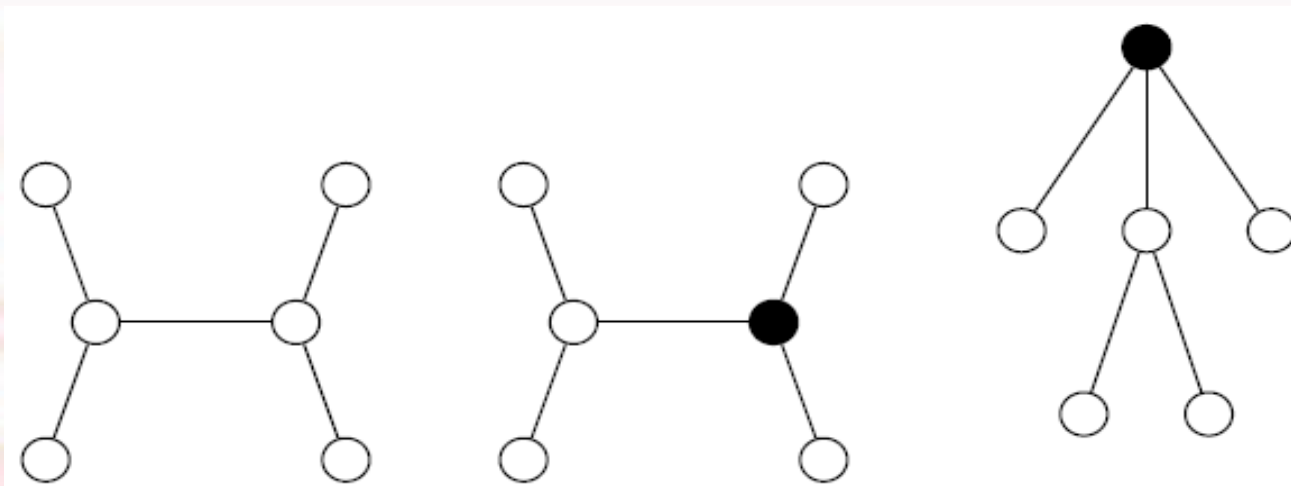
# *Filogenetska drevesa: matematična definicija*

*(za tiste, ki ste se že srečali z grafi)*

- ◆ Kot matematični objekt so filogenetska drevesa *dvojiška drevesa*
- ◆ **drevo** je povezan graf brez ciklov
- ◆ drevo s korenom je **dvojiško**, če ima vsako vozlišče, ki ni list, natanko dva sinova
- ◆ drevo brez korena je **dvojiško**, če je vsako vozlišče stopnje 1 ali 3

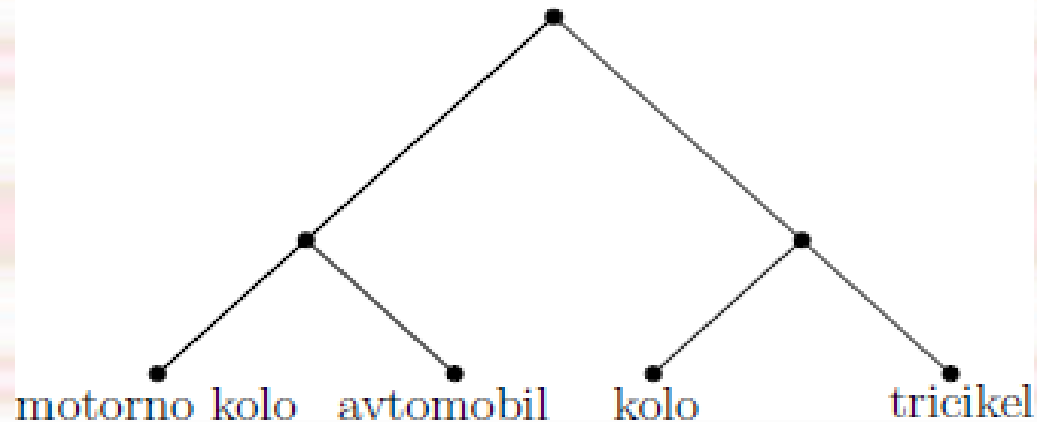
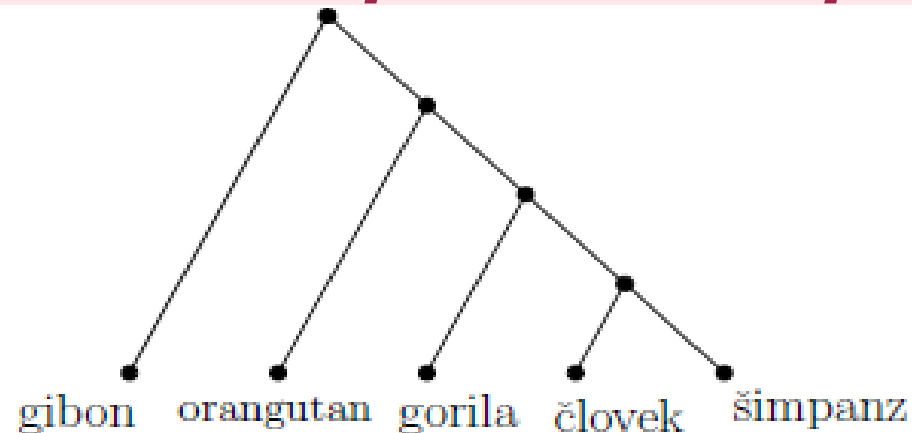
# *Filogenetska drevesa*

- ◆ Obstaja več vrst filogenetskih dreves:
  - s korenom ali brez njega
  - dolžine povezav so pomembne ali pa ne



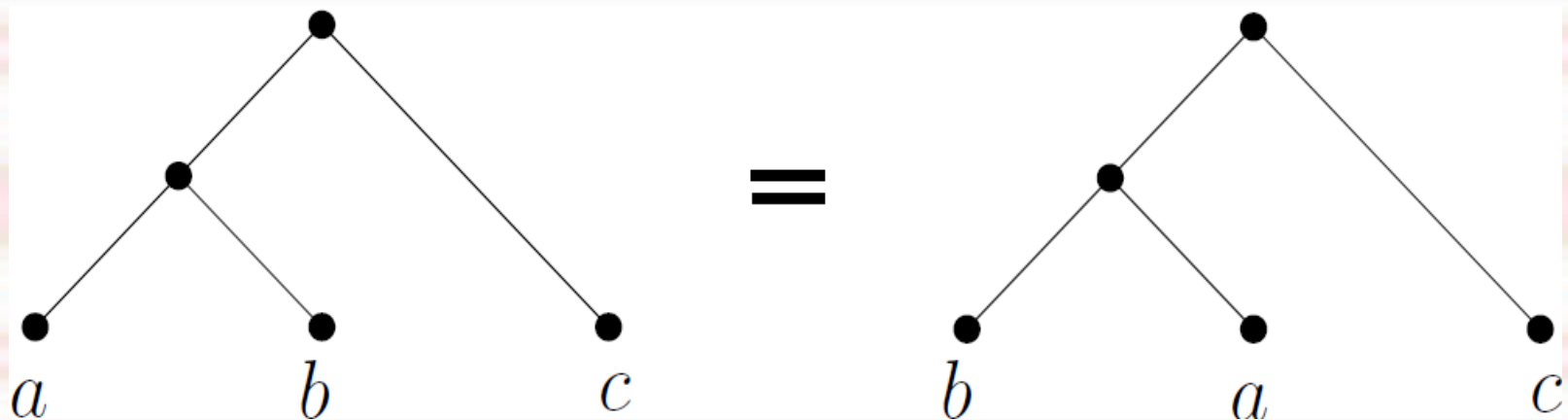
drevo brez korena in drevo s korenom  
(narisano na dva načina)

# *Dve zelo enostavni filogenetski drevesi (s korenom, dolžina povezav ni pomembna)*



# Število filogenetskih dreves

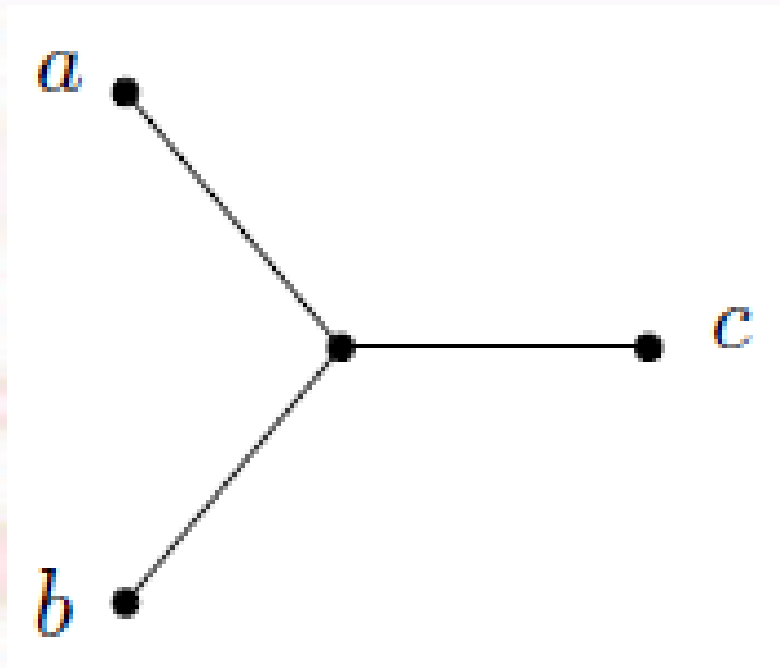
- ◆ Preštejmo, koliko je filogenetskih dreves (brez korena) z  $n$  listi.
- ◆ Pozor! Dreves, ki se razlikujejo le v vrstnem redu listov pod istim notranjim vozliščem, ne razlikujemo:



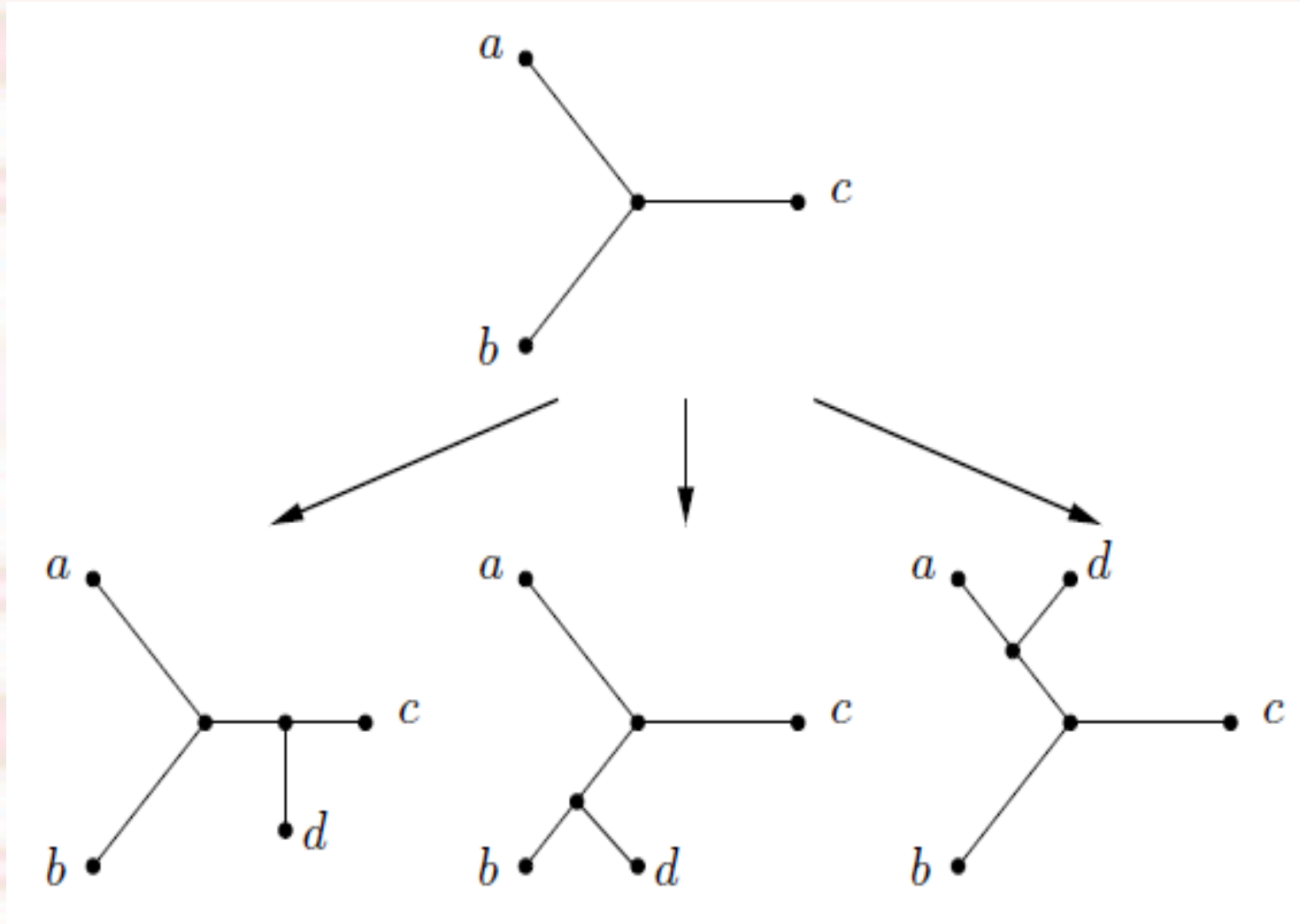
podobno kot pri visečih igračah



- ◆ Uganimo število in ga dokažimo z indukcijo na  $n$ .
- ◆  $n = 3$ : obstaja eno samo drevo (brez korena) s **tremi listi**:



- ◆ Iz tega drevesa lahko tvorimo drevo s 4 listi na 3 načine:





# Število filogenetskih dreves

- ◆ Drevo z  $n$  listi lahko tvorimo iz drevesa z  $n-1$  listi tako, da mu dodamo en nov list na eno izmed vej.
- ◆ Z indukcijo lahko pokažemo, da ima dvojiško drevo z  $n$  listi **natanko  $2n-3$  vej** (povezav).
- ◆ Torej imamo za zgornji korak  
$$2(n-1)-3 = 2n-5$$
 možnosti.

# Število filogenetskih dreves

- ◆ Ugibamo:

Za vsak  $n \geq 3$  je število dvojiških dreves brez korena z  $n$  listi podano z izrazom

$$(2n-5) \times (2n-7) \times \dots \times 5 \times 3.$$

- ◆ Dokaz: za vajo (indukcija).

- ◆ Primer:

Za  $n = 7$ :  $11 \times 9 \times 7 \times 5 \times 3 = 10395$

Za  $n = 20$ : več kot  $2 \times 10^{20}$  dreves

# *Število filogenetskih dreves*

- ◆ Filogenetskih dreves brez korena je torej **zelo veliko!**
- ◆ Dreves s korenom je pa še več kot dreves brez korena, saj imamo na voljo še poljubno izbiro korenskega vozlišča.
- ◆ K sreči nam (vsaj v nekaterih primerih) k rekonstrukciji filogenetskih dreves pomagajo matematični razmisleki.

# *Metode rekonstrukcij filogenetskih dreves*

- ◆ Filogenetska drevesa lahko rekonstruiramo na več načinov:
  - na osnovi poznanih oz. izračunanih **razvojnih razdalj med danimi organizmi** (manjša kot je razdalja med dvema organizmoma, bolj sta si organizma sorodna);
  - na osnovi **lastnosti danih organizmov**.
- ◆ Poglejmo si primer rekonstrukcije na osnovi lastnosti danih organizmov.

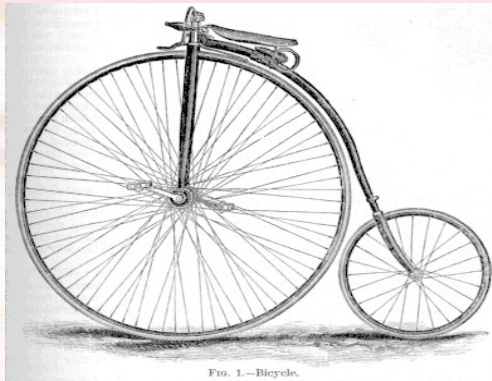
## ***Filogenetska drevesa, ki temeljijo na lastnostih organizmov***

- ◆ Danih je  $n$  organizmov in  $m$  numeričnih lastnosti kot npr.:
  - število nog,
  - prisotnost živčnega sistema,
  - prisotnost določenega podzaporedja nukleotidov v DNK,
  - prisotnost določenega gena.
- ◆ Lastnosti organizmov ponazorimo z matriko.

## *Primer: „evolucija“ avtomobila*

	<b>prisotnost motorja</b>	<b>število koles</b>
<b>kolo</b>	<b>0</b>	<b>2</b>
<b>tricikel</b>	<b>0</b>	<b>3</b>
<b>motorno kolo</b>	<b>1</b>	<b>2</b>
<b>avtomobil</b>	<b>1</b>	<b>4</b>

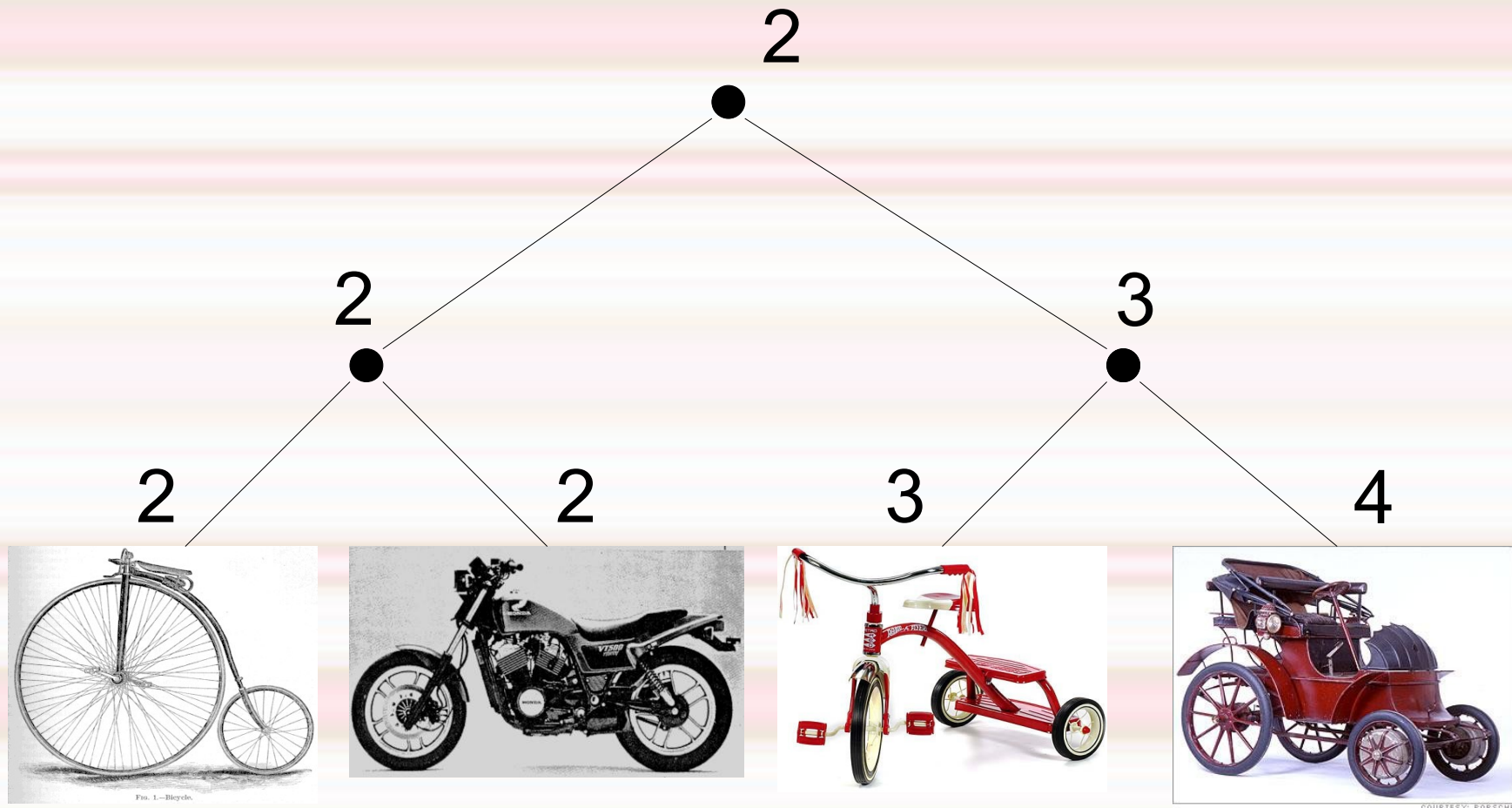
# Primer: „evolucija“ avtomobila







# Možna „evolucija“ avtomobila glede na število koles

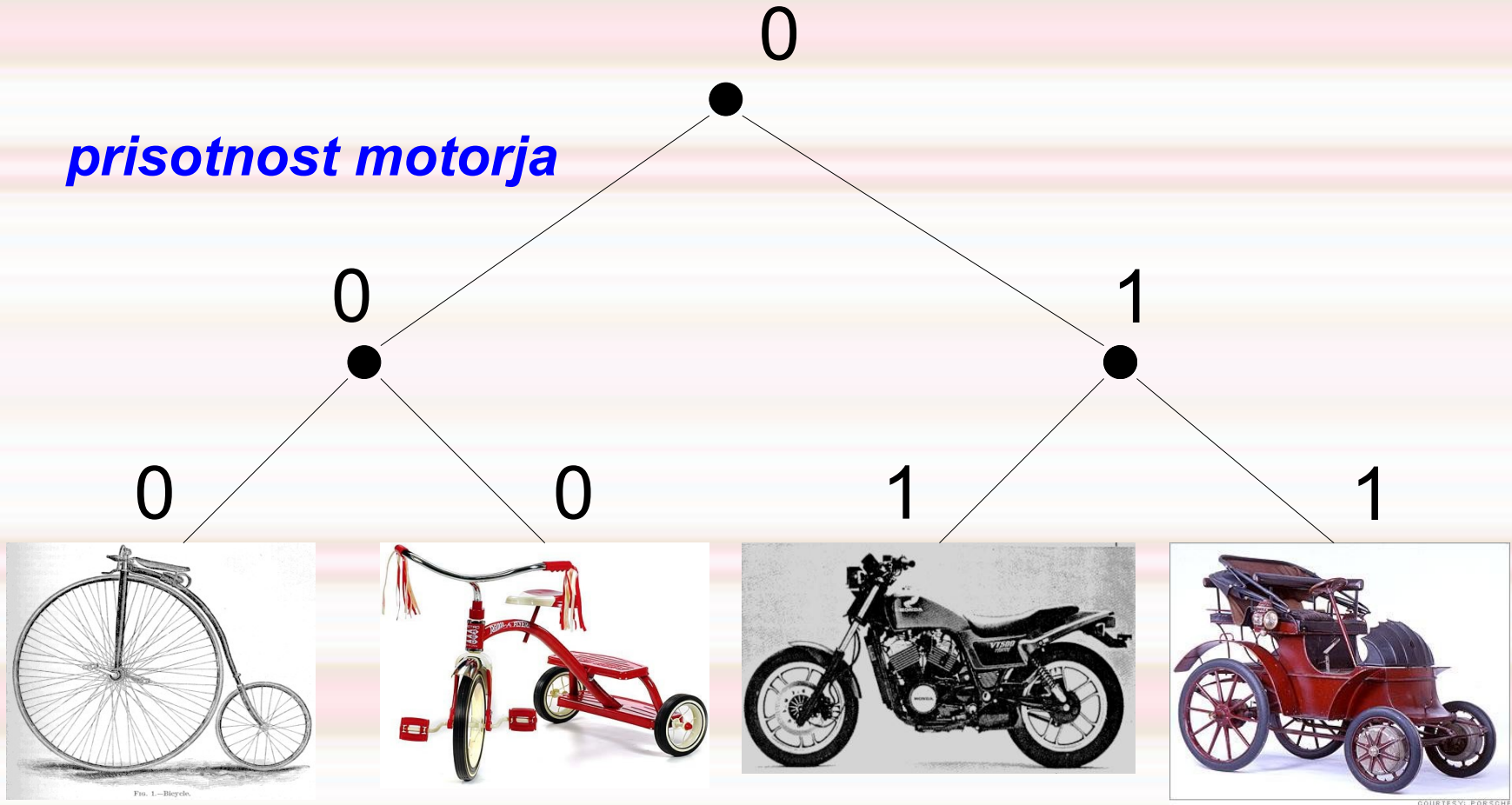


# *Popolna filogenetska drevesa*

- ◆ **Popolno filogenetsko drevo** je filogenetsko drevo, v katerem *za vsako od  $m$  lastnosti velja, da vozlišča, ki imajo isto vrednost lastnosti, tvorijo poddrevo (povezan podgraf)*
- ◆ Popolna filogenetska drevesa prikazujejo razvoj organizmov na posebno enostaven način.

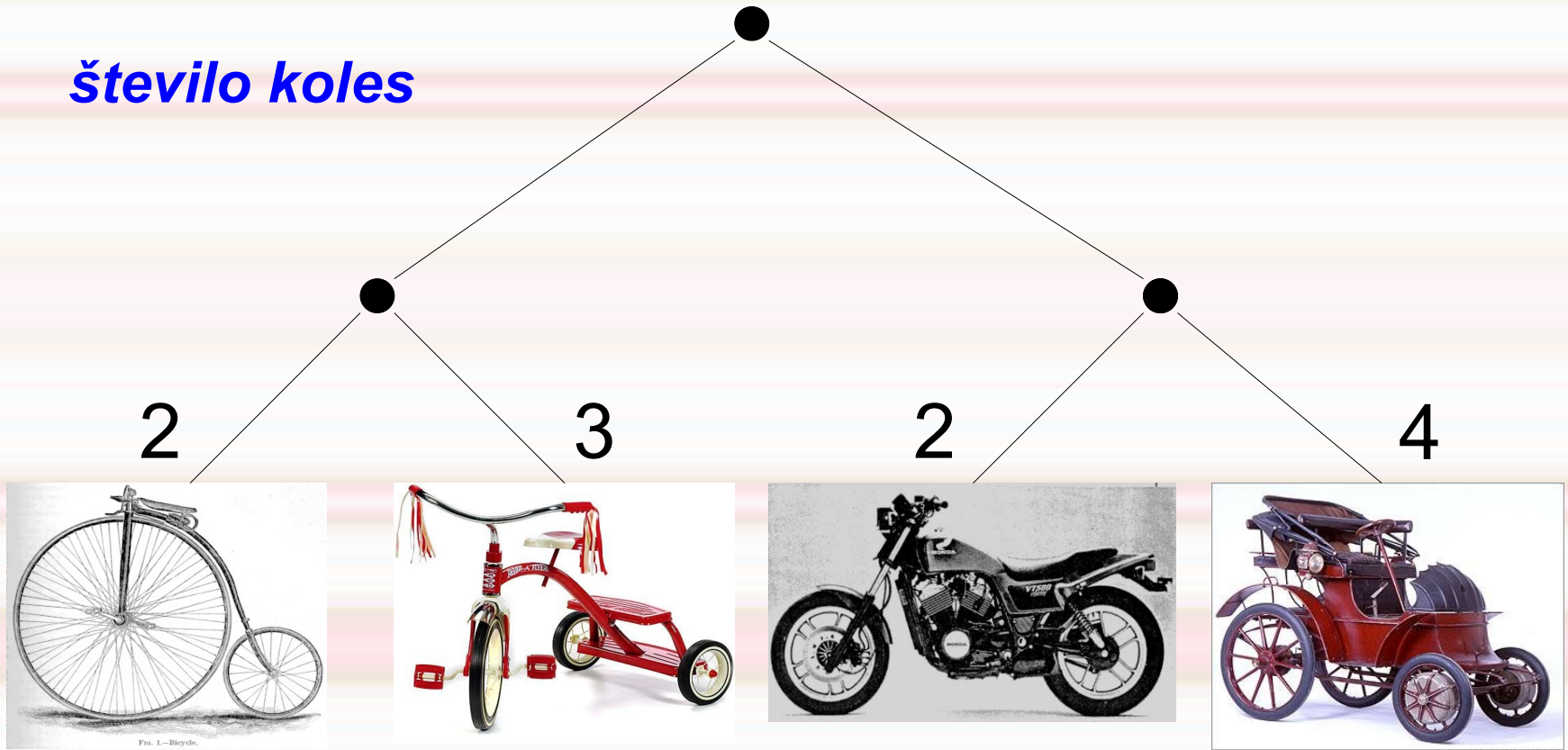
- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

*prisotnost motorja*



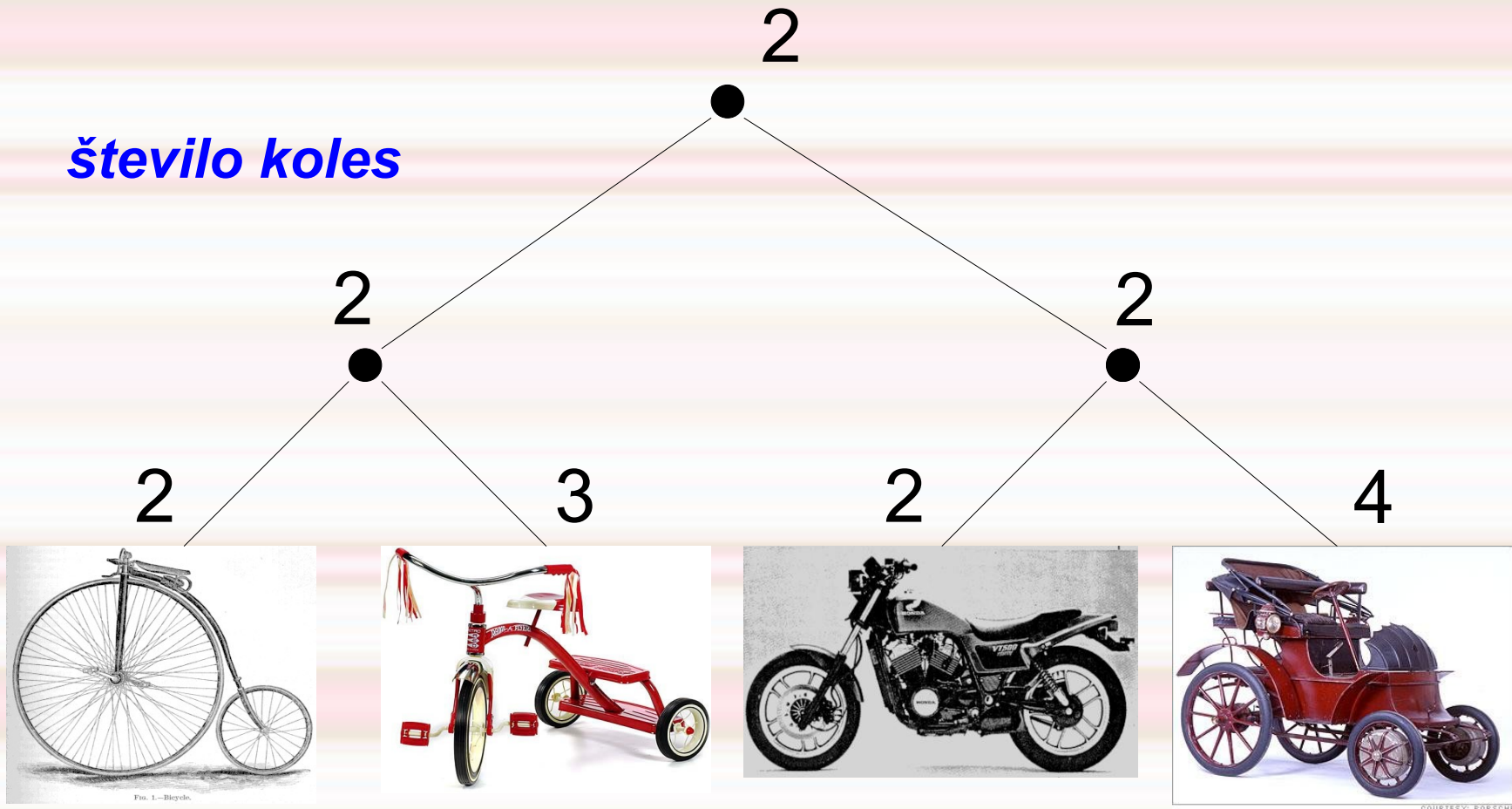
- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

*število koles*



- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

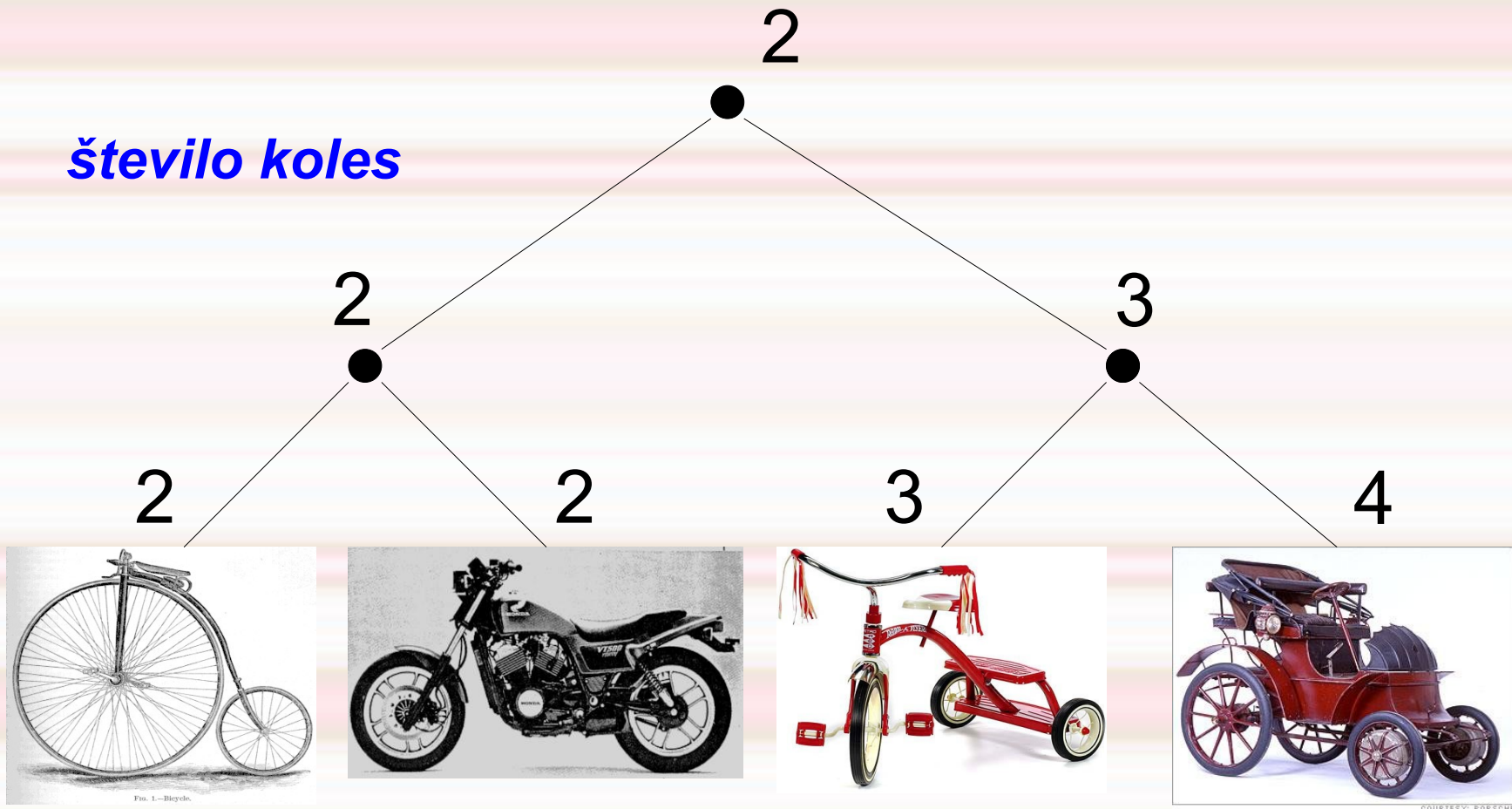
število koles





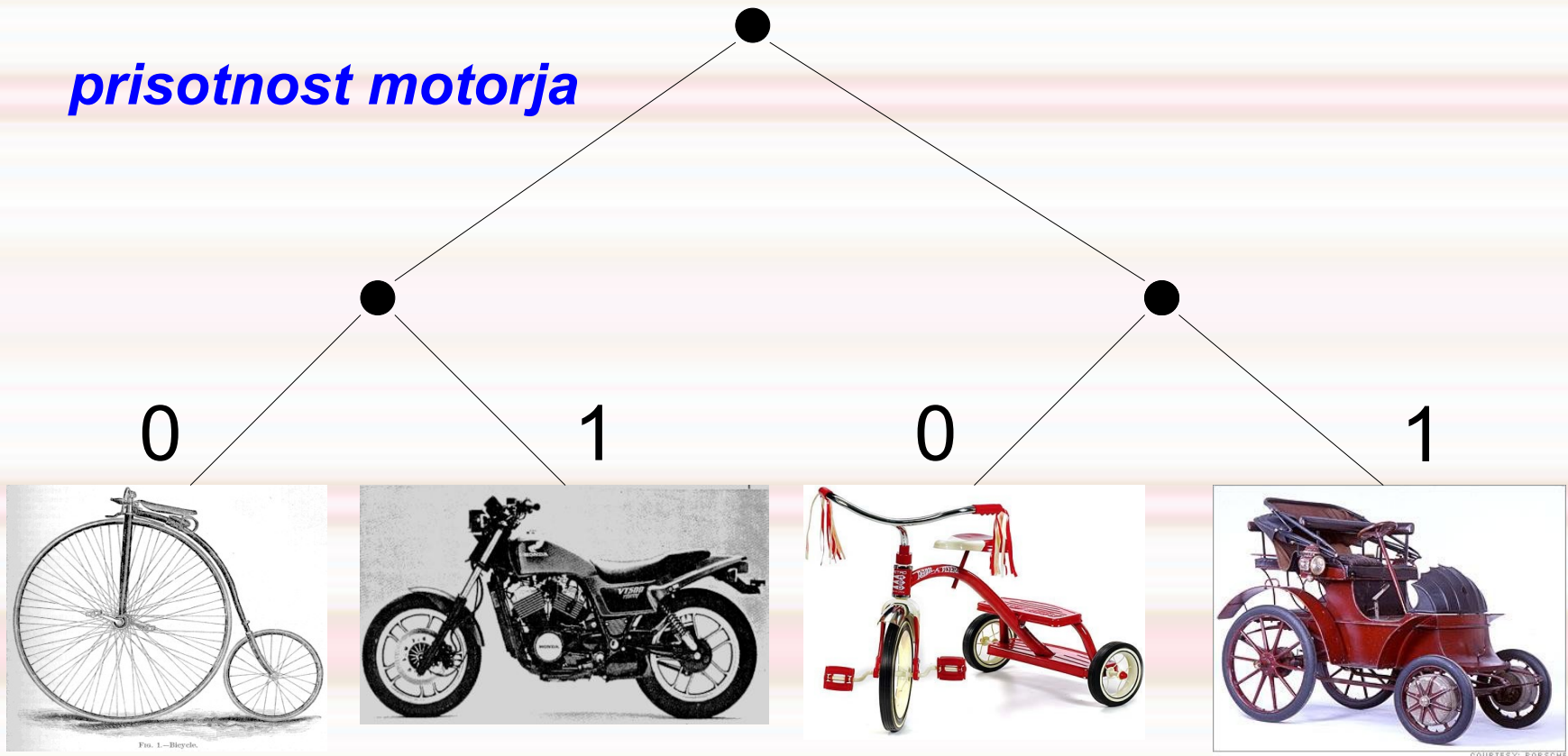
- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

število koles



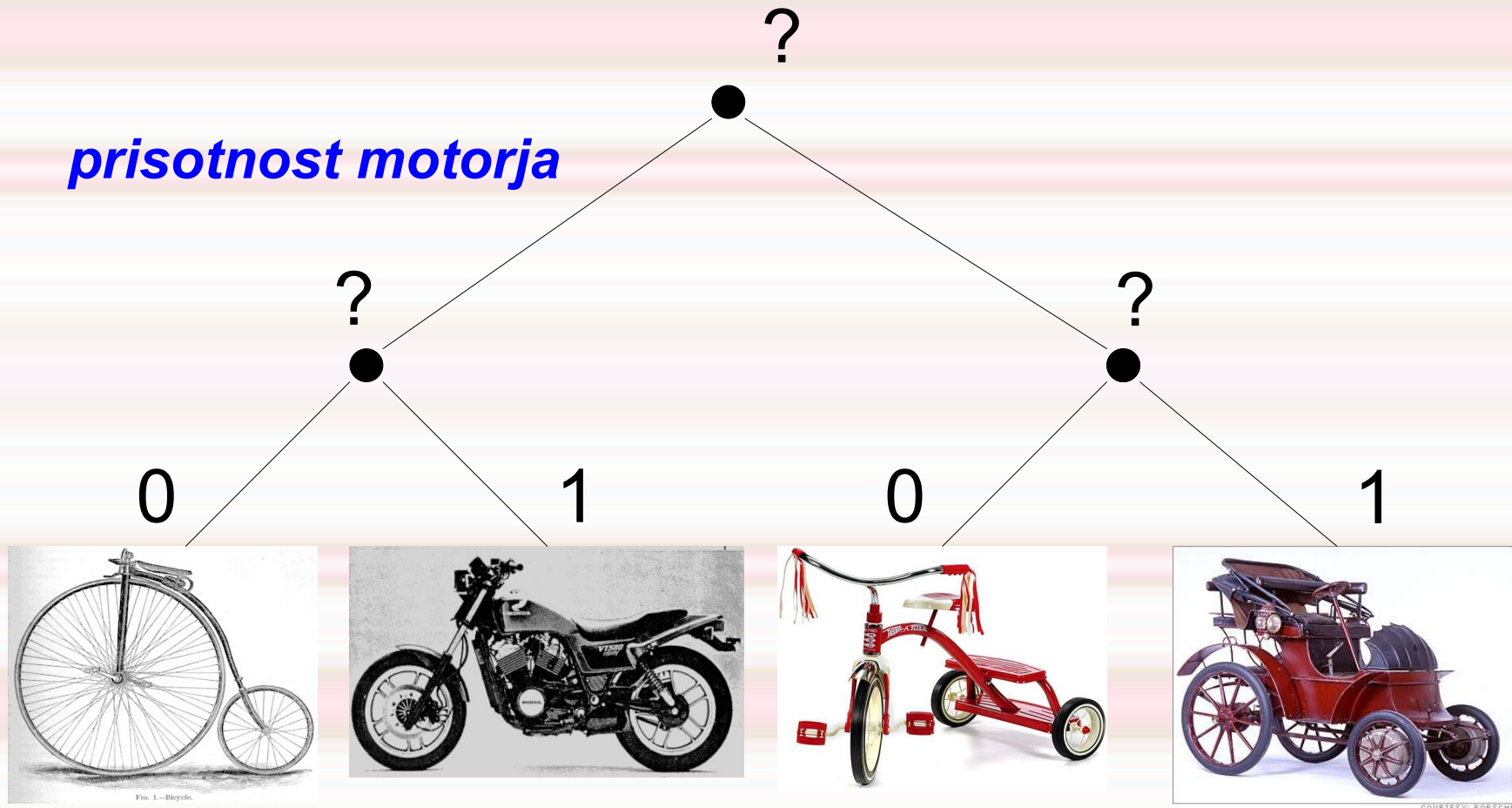
- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

*prisotnost motorja*



- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

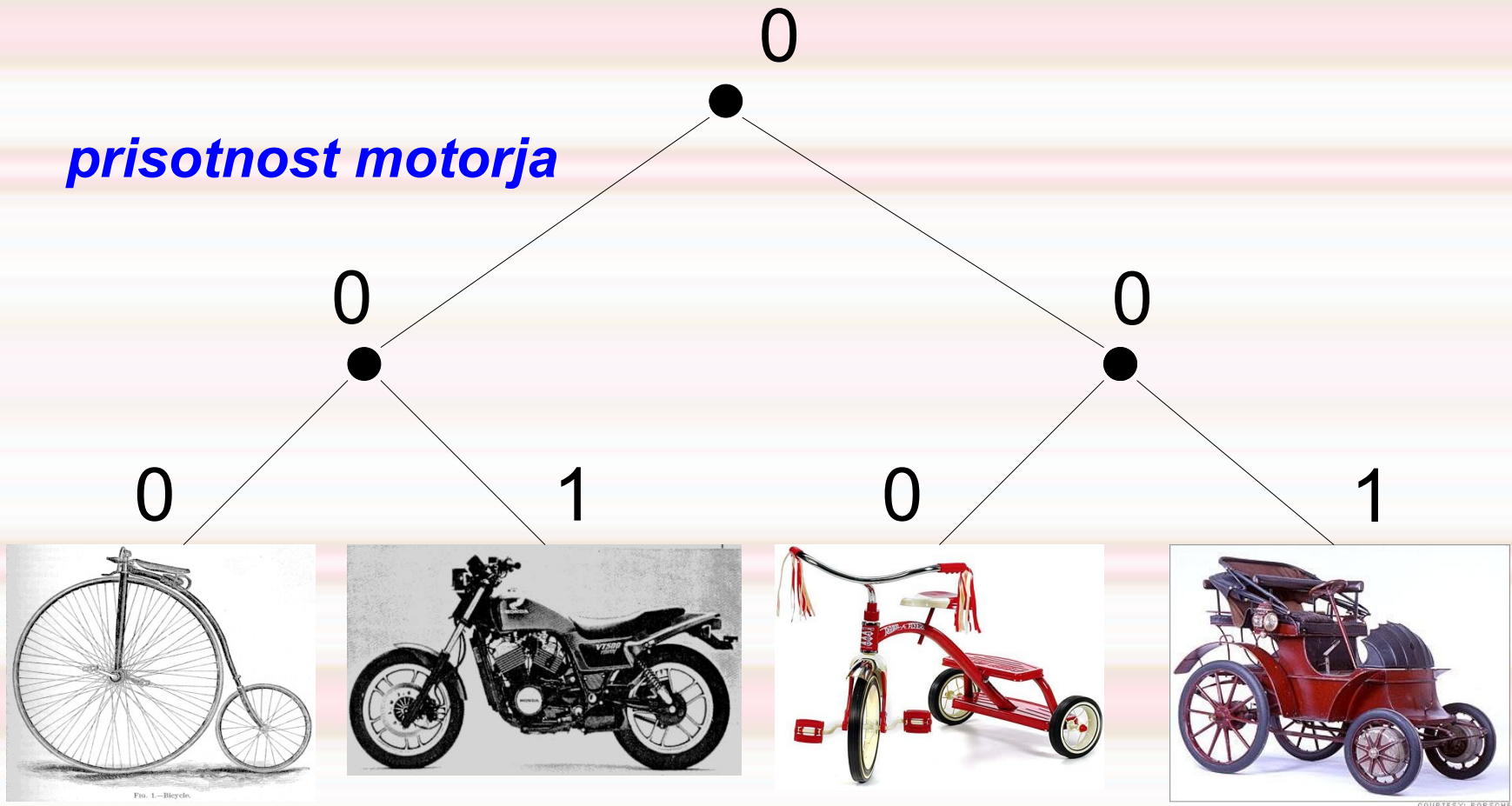
*prisotnost motorja*





- ◆ Od prikazanih možnih evolucij avtomobila je le prva popolno filogenetsko drevo!

*prisotnost motorja*



# *Problem rekonstrukcije popolnih filogenetskih dreves*

- ◆ Dana je  $m \times n$  matrika.  
vrstice: organizmi  
stolpci: (celoštevilske) lastnosti
- ◆ Element v  $i$ -ti vrstici in  $j$ -tem stolpcu prikazuje vrednost  $j$ -te lastnosti, ki jo ima  $i$ -ti organizem.
- ◆ ***Problem: določi, ali obstaja popolno filogenetsko drevo, skladno z dano matriko!***

	<b>prisotnost motorja</b>	<b>število koles</b>
<b>kolo</b>	<b>0</b>	<b>2</b>
<b>tricikel</b>	<b>0</b>	<b>3</b>
<b>motorno kolo</b>	<b>1</b>	<b>2</b>
<b>avtomobil</b>	<b>1</b>	<b>4</b>

- ◆ Za to matriko popolno filogenetsko drevo obstaja.

# *Problem rekonstrukcije popolnih filogenetskih dreves*

- ◆ V splošnem je problem **težak**: ne poznamo rešitev, ki bi bile bistveno boljše od preizkušanja vseh možnih dreves.
- ◆ Problem pa je enostavno rešljiv, če ima vsaka lastnost samo **dve možni vrednosti**.
  - ◆ prisotnost motorja: 0/1 (0 = NE, 1 = DA)
  - ◆ prisotnost določenega gena
  - ◆ prisotnost določenega podzaporedja v DNK

## *Dvojiške lastnosti*

- ◆ Predpostavimo, da skupni prednik nima nobene od danih lastnosti.
- ◆ Recimo, da imamo dano naslednjo matriko:

	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>	<b>L5</b>	<b>L6</b>
<b>O1</b>	0	0	0	0	1	0
<b>O2</b>	0	1	0	0	1	0
<b>O3</b>	0	0	1	1	0	0
<b>O4</b>	1	0	1	1	0	0
<b>O5</b>	0	0	1	0	0	1

	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>	<b>L5</b>	<b>L6</b>
<b>O1</b>	0	0	0	0	1	0
<b>O2</b>	0	1	0	0	1	0
<b>O3</b>	0	0	1	1	0	0
<b>O4</b>	1	0	1	1	0	0
<b>O5</b>	0	0	1	0	0	1

- ◆ Označimo z  $S(i)$  množico vseh organizmov, ki imajo  $i$ -to lastnost.
- ◆  $S(1) = \{4\}$ ,  $S(2) = \{2\}$ ,  $S(3) = \{3,4,5\}$  itd.

	L1	L2	L3	L4	L5	L6
O1	0	0	0	0	1	0
O2	0	1	0	0	1	0
O3	0	0	1	1	0	0
O4	1	0	1	1	0	0
O5	0	0	1	0	0	1

{4} {2} {3,4,5} {3,4} {1,2} {5}

### Izrek (Gusfield, 1991):

*Za dano 0/1 matriko obstaja popolno filogenetsko drevo natanko tedaj, ko za vsak par  $i, j$  velja ena od naslednjih trditev:*

- ◆  $S(i)$  in  $S(j)$  nimata skupnih elementov;
- ◆  $S(i)$  vsebuje  $S(j)$ ;
- ◆  $S(j)$  vsebuje  $S(i)$ .

	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>	<b>L5</b>	<b>L6</b>
<b>O1</b>	0	0	0	0	1	0
<b>O2</b>	0	1	0	0	1	0
<b>O3</b>	0	0	1	1	0	0
<b>O4</b>	1	0	1	1	0	0
<b>O5</b>	0	0	1	0	0	1

$\{4\}$   $\{2\}$   $\{3,4,5\}$   $\{3,4\}$   $\{1,2\}$   $\{5\}$

Z uporabo izreka se zlahka prepričamo, da za to matriko popolno filogenetsko drevo obstaja!



## *Ta rešitev je zelo učinkovita:*

- ◆ Recimo, da bi radi preverili, ali obstaja filogenetsko drevo za nabor 20 organizmov, glede na 10 različnih lastnosti.
- ◆ Število vseh možnih filogenetskih dreves z 20 listi je ogromno (**več kot  $2 \times 10^{20}$  dreves**)
- ◆ Da preverimo pogoj v izreku, pa moramo med seboj primerjati le **10 podmnožic množice števil  $\{1, 2, \dots, 20\}$ .**

## *Drevo lahko tudi rekonstruiramo*

- ◆ Z malce več truda lahko popolno filogenetsko drevo, če obstaja, tudi **učinkovito rekonstruiramo**.
- ◆ Najprej uredimo stolpce glede na število enic.

	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>	<b>L5</b>	<b>L6</b>
<b>O1</b>	0	0	0	0	1	0
<b>O2</b>	0	1	0	0	1	0
<b>O3</b>	0	0	1	1	0	0
<b>O4</b>	1	0	1	1	0	0
<b>O5</b>	0	0	1	0	0	1

postane

	<b>L3</b>	<b>L4</b>	<b>L5</b>	<b>L1</b>	<b>L2</b>	<b>L6</b>
<b>O1</b>	0	0	1	0	0	0
<b>O2</b>	0	0	1	0	1	0
<b>O3</b>	1	1	0	0	0	0
<b>O4</b>	1	1	0	1	0	0
<b>O5</b>	1	0	0	0	0	1

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

{1,2,3,4,5}

{3,4,5} {3,4} {1,2} {4} {2} {5}

Začnemo z drevesom, ki vsebuje le koren. Ta predstavlja množico vseh organizmov {1,2,3,4,5}.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

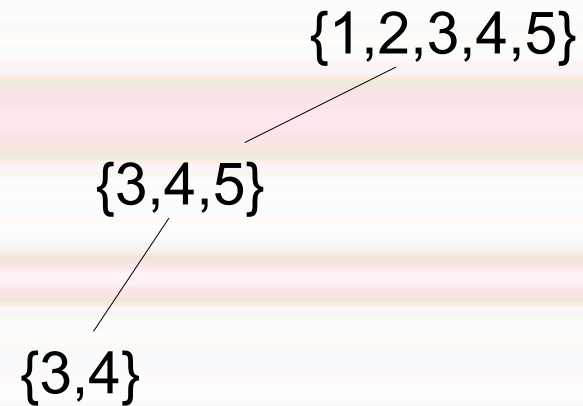
**{3,4,5} {3,4} {1,2} {4} {2} {5}**

**{1,2,3,4,5}**  
**{3,4,5}**

Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

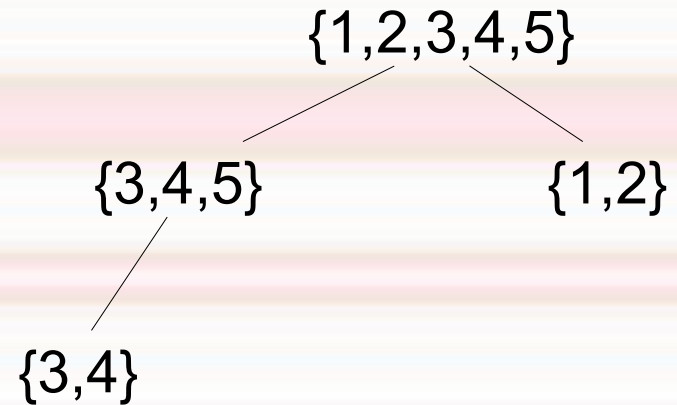
**{3,4,5} {3,4} {1,2} {4} {2} {5}**



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

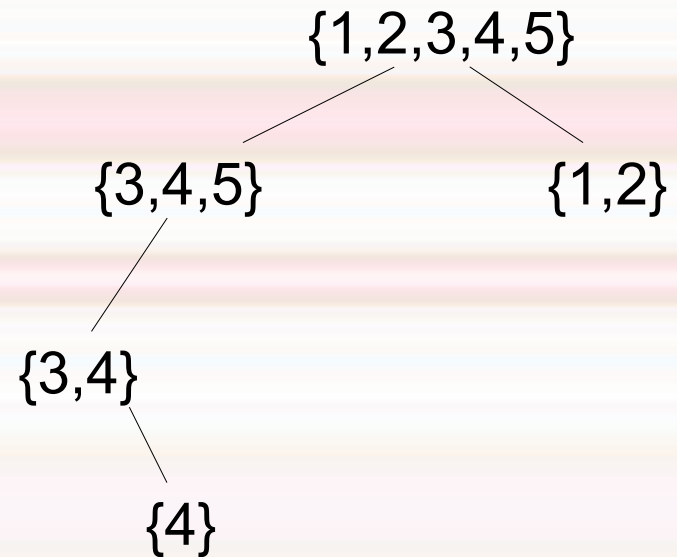
{3,4,5} {3,4} {1,2} {4} {2} {5}



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

{3,4,5} {3,4} {1,2} {4} {2} {5}



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.



	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

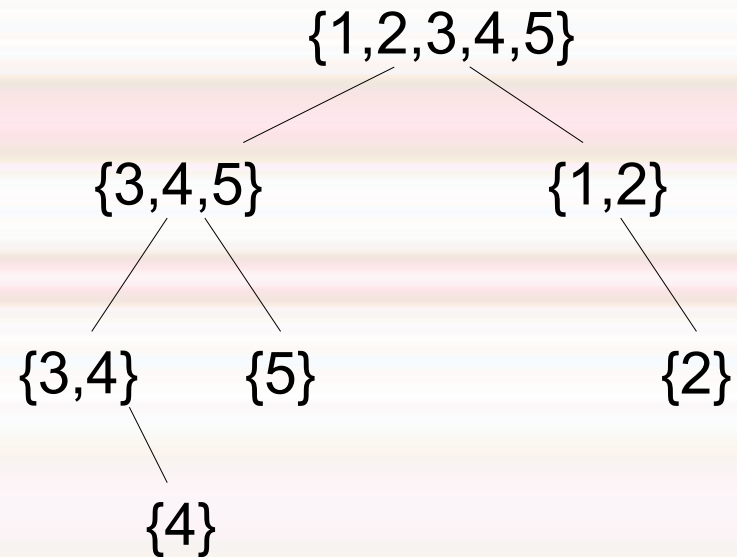
{3,4,5} {3,4} {1,2} {4} {2} {5}



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

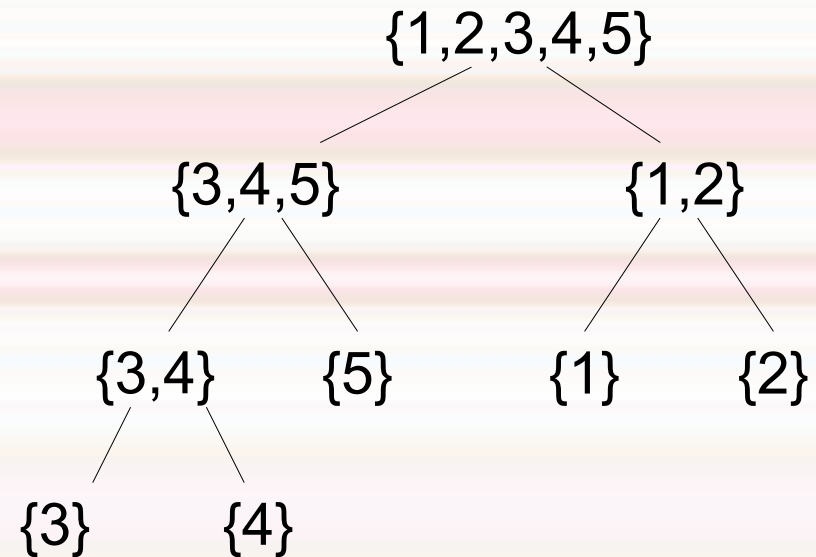
{3,4,5} {3,4} {1,2} {4} {2} {5}



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

	L3	L4	L5	L1	L2	L6
O1	0	0	1	0	0	0
O2	0	0	1	0	1	0
O3	1	1	0	0	0	0
O4	1	1	0	1	0	0
O5	1	0	0	0	0	1

{3,4,5} {3,4} {1,2} {4} {2} {5}



Urejene stolpce dodajamo  
v drevo po vrsti, drugega  
za drugim.

Nazadnje dodamo še  
manjkajoče liste.

# Sklep

- ◆ rekonstrukcija filogenetskih dreves je eno osrednjih bioloških vprašanj
- ◆ obstajajo številni pristopi za rešitev tega problema, ki temeljijo na matematičnih metodah
- ◆ prešteli smo filogenetska drevesa (dvojiška drevesa brez korena) z danim številom listov
- ◆ rekonstrukcija popolnih filogenetskih dreves je v splošnem težek problem
- ◆ enostavna rešitev za 0/1 lastnosti

***Vprašanja?***