EEOB563 – Assignment #4

Dennis Psaroudakis

February 27th 2019

EEOB 563 - Assignment 4

I

1.

	Dog	Cat	Mouse	Pig	Human
Dog	0	0.1468	0.1468	0.1073	0.3295
Cat		0	0.2326	0.1073	0.3831
Mouse			0	0.1468	0.3831
Pig				0	0.3831
Human					0

2. Smallest distance:
Pig - dog

	PD	C	M	H
PD	0	0.12705	0.1468	0.3563
C		0	0.2326	0.3831
M			0	0.3831
H				0

now: PD - C

	PDC	M	H
PDC	0	0.1754	0.3652
M		0	0.3831
H			0

finally: PDC - M

	PDCM	H
PDCM	0	0.369675
H		0

Pig   Dog   Cat   Mouse   Human

0.05365   0.05365   0.063525   0.08771

0.063525
0.00987

0.02417

0.09715   0.18486
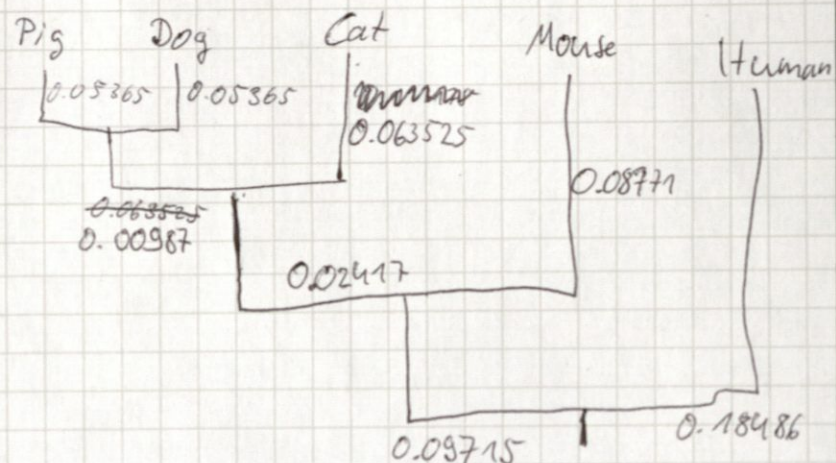
3. Neighbor-Joining

The algorithm in the notes & English Wikipedia is super complicated and confusing but in the German Wikipedia they introduced an intermediary step: average distance of each token to all others, netto divergence r_i
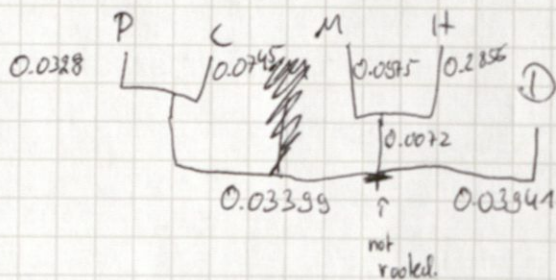
$$r_i = \frac{1}{N-2} \sum_{k=1}^{N} d_{i,k} \qquad \text{where} \quad d_{i,k} = \text{distance from } i \text{ to } k.$$

	D	C	M	P	H	r
D	0	0.1468	0.1468	0.1073	0.3295	0.2435
C		0	0.2326	0.1073	0.3831	0.2899
M			0	0.1468	0.3831	0.3031
P				0	0.3831	0.2481
H					0	0.4929

I just noticed, this $r = u$ from the lecture notes. So it's just complicated in the English WP.

Now the new matrix M:

	D	C	M	P	H
D	0	-0.3866	-0.3998	-0.3844	-0.4069
C		0	-0.3604	-0.4308	-0.3997
M			0	-0.4045	-0.4129
P				0	-0.358
H					0

smallest distance: PC

	D	PC	M	H	r
D	0	0.1468	0.1468	0.3295	0.31155
PC		0	0.2721	0.6589	0.5389
M			0	0.3831	0.401
H				0	0.68575

—) Matrix M

	D	PC	M	H
D	0	-0.704	-0.5658	-0.6678
PC		0		
M			0	
H				0

I forgot to divide b 2 »

	D	PC	M	H		r
			0.1468	0.3295		
D	0	0.0734	~~0.0734~~	~~0.1648~~		0.2749
PC		0	0.1361	0.3295		0.2695
M			0	0.3831		0.333
H				0		0.5211

↱new M

	D	PC	M	H
D	0	-0.471	-0.4611	-0.4665
PC		0	-0.4664	-0.4611
M			0	-0.471
H				0

merge M & H

	D	PC	MH
D	0	0.0734	0.04125
PC		0	0.04125
MH			0

Now D could be joined with either
PC or MH... I'll join it with PC...

4.

The generated matrices and trees can be found here. I chose Platypus (20) as the outgroup, this is the generated tree:

```
1   (((((Elephant,Sloth,(Flying_fox,(Whale,Pig),(Human,Cat),Armadillo,
2   Hedgehog,(Rabbit,Aardvark),Dog,Tenrec,Tree_shrew),Mouse),(Opossum,
3   Bandicoot,Wallaby),Echidna),Platypus);))))
```
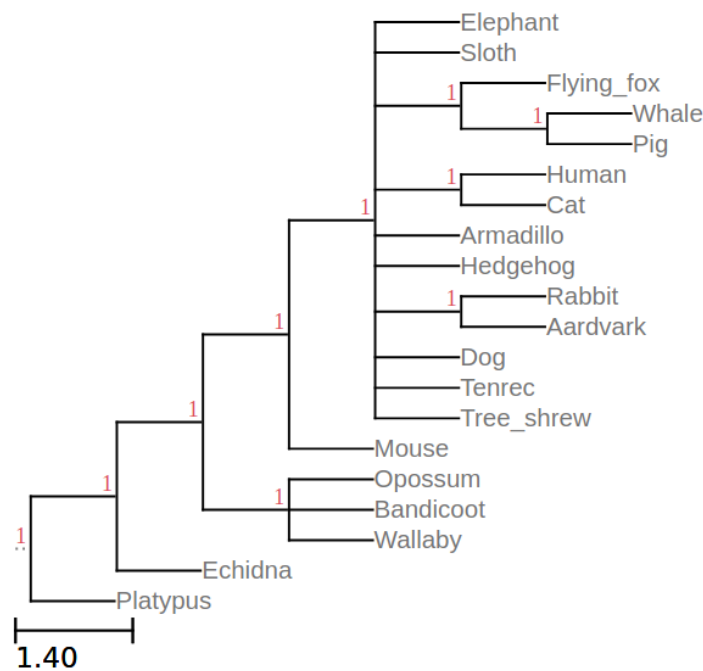

Figure 1: Strict consensus tree

5.

I used the Jukes-Cantor distance model, just because I used it for Part I so I knew most about it. This is the generated consensus tree from all the bootstrapped alignments:

```
1   ((((((((((Cat:1.00,Rabbit:1.00):1.00,Mouse:1.00):1.00,Dog:1.00):1.00,((Elephant:1.00,
2   Sloth:1.00):1.00,Human:1.00):1.00):1.00,(((Armadillo:1.00,Whale:1.00):1.00,Pig:1.00)
    :1.00,
```

```
3   (Flying_fox:1.00,Tree_shrew:1.00):1.00):1.00):1.00,Aardvark:1.00):1.00,(Hedgehog
        :1.00,
4   Tenrec:1.00):1.00):1.00,(Opossum:1.00,(Wallaby:1.00,Bandicoot:1.00):1.00):1.00):1.00,
5   Platypus:1.00):1.00,Echidna:1.00);
```
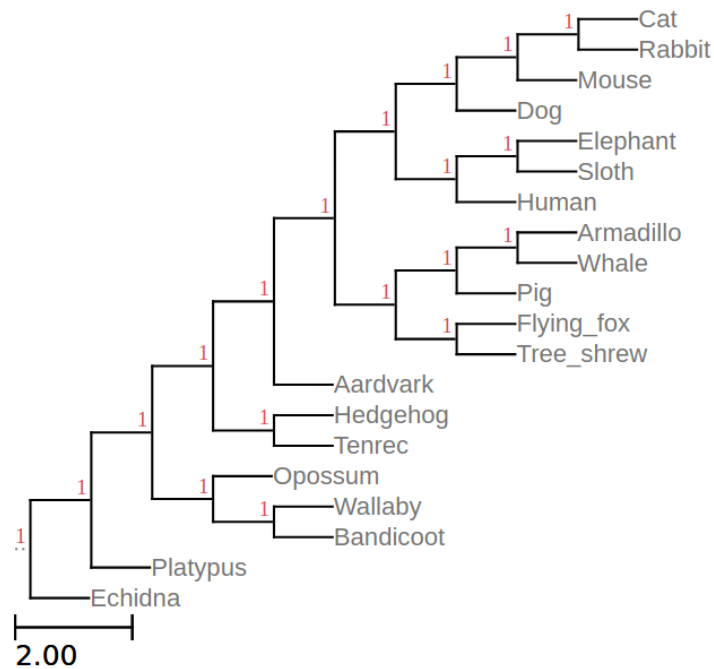

Figure 2: Majority consensus tree

8.

Again, I used the Jukes-Cantor matrix. In my case, the two trees basically were the same, only the distances
were slightly different:

Without NNI/SPR:

```
1   (((((Cat:0.114165981250,Human:0.162999018750):0.010559694336,((((Hedgehog
        :0.178496817708,Tenrec:0.157865182292):0.012008114955,(Mouse:0.121806644097,(((
        Bandicoot:0.115205470588,Opossum:0.129977529412):0.001976093750,Wallaby
        :0.119019406250):0.017233485577,(Echidna:0.102001583333,Platypus:0.112883416667)
        :0.066429764423):0.010216355903):0.004562978795):0.005054285156,(Armadillo
        :0.119819296875,(Sloth:0.152378928571,Elephant:0.163353071429):0.012417703125)
```

```
        :0.006888199219):0.002261112559,(Rabbit:0.131889476562,Aardvark:0.120376523438)
        :0.005459932129):0.002273915039):0.003076952637,(Dog:0.127646692383,Tree_shrew
        :0.148635307617):0.001928684082):0.011074565918,Flying_fox:0.117479167614)
        :0.013291332386,Pig:0.093116691667,Whale:0.119435308333);
```

With NNI/SPR:

```
1   (((((Cat:0.112028724609,Human:0.165136275391):0.009859859863,((((Hedgehog
        :0.177491945312,Tenrec:0.158870054687):0.009443020020,(Mouse:0.123529160645,(((
        Bandicoot:0.113613665527,Opossum:0.131569334473):0.004081156738,Wallaby
        :0.116914343262):0.018083876465,(Echidna:0.105270403320,Platypus:0.109614596680)
        :0.065579373535):0.008493839355):0.007128073730):0.005542280762,(Armadillo
        :0.119178175781,(Sloth:0.143488730469,Elephant:0.172243269531):0.013058824219)
        :0.006400203613):0.002117241699,(Rabbit:0.130965949219,Aardvark:0.121300050781)
        :0.005603812988):0.002973749512):0.003076952637,(Dog:0.127646692383,Tree_shrew
        :0.148635307617):0.001928684082):0.011074565918,Flying_fox:0.117013964355)
        :0.013756535645,Pig:0.087929280762,Whale:0.124622719238);
```

Left: Without NNI/SPR, right: with NNI/SPR
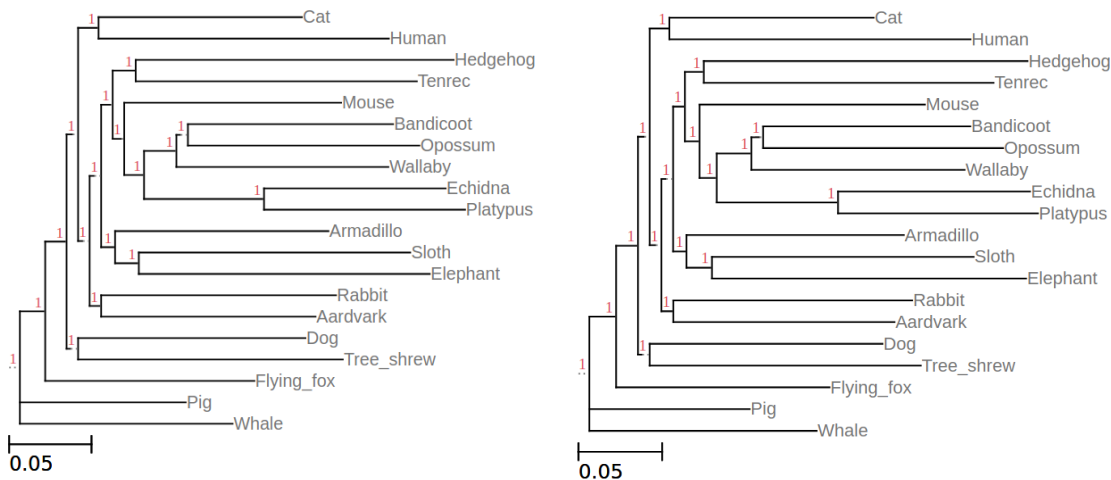
I may have done something wrong, you can check my parameters and input here.