

Enhanced sampling methods with PLUMED



PLUMED documentation

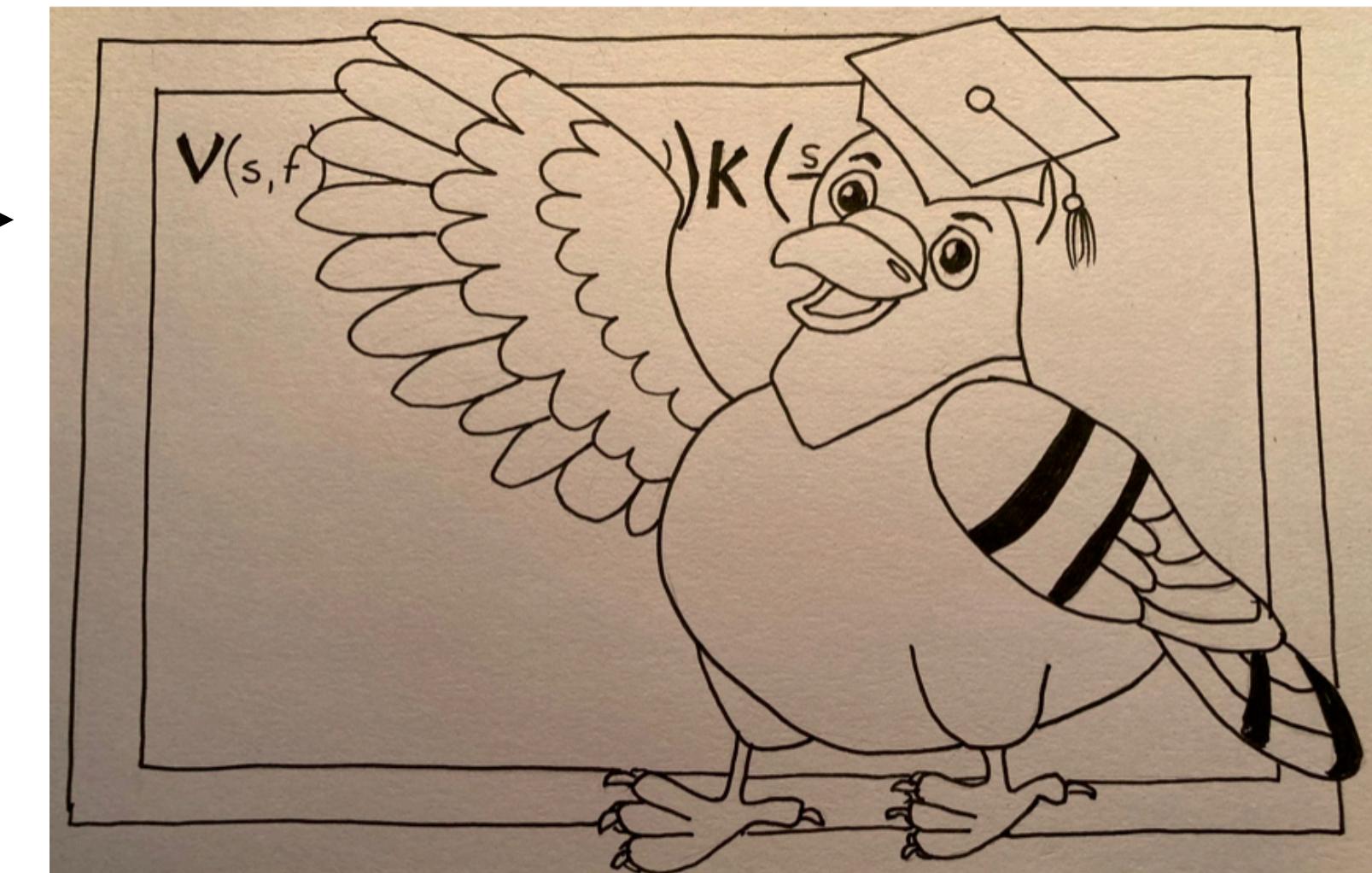
GARETH TRIBELLO

The problem with documentation

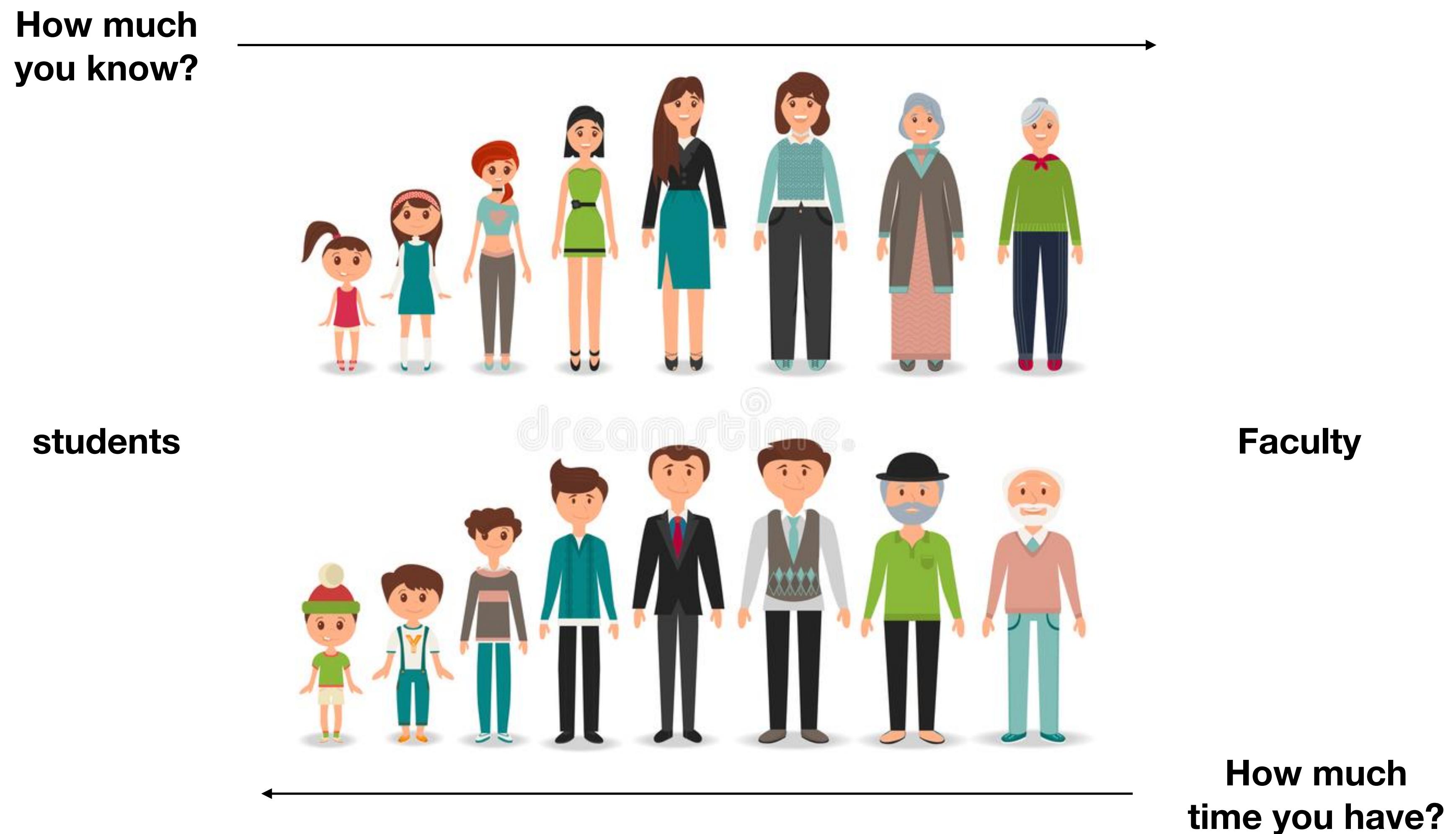


I am trying to do X with PLUMED but it is not working/is slow/
PLUMED doesn't read my input file

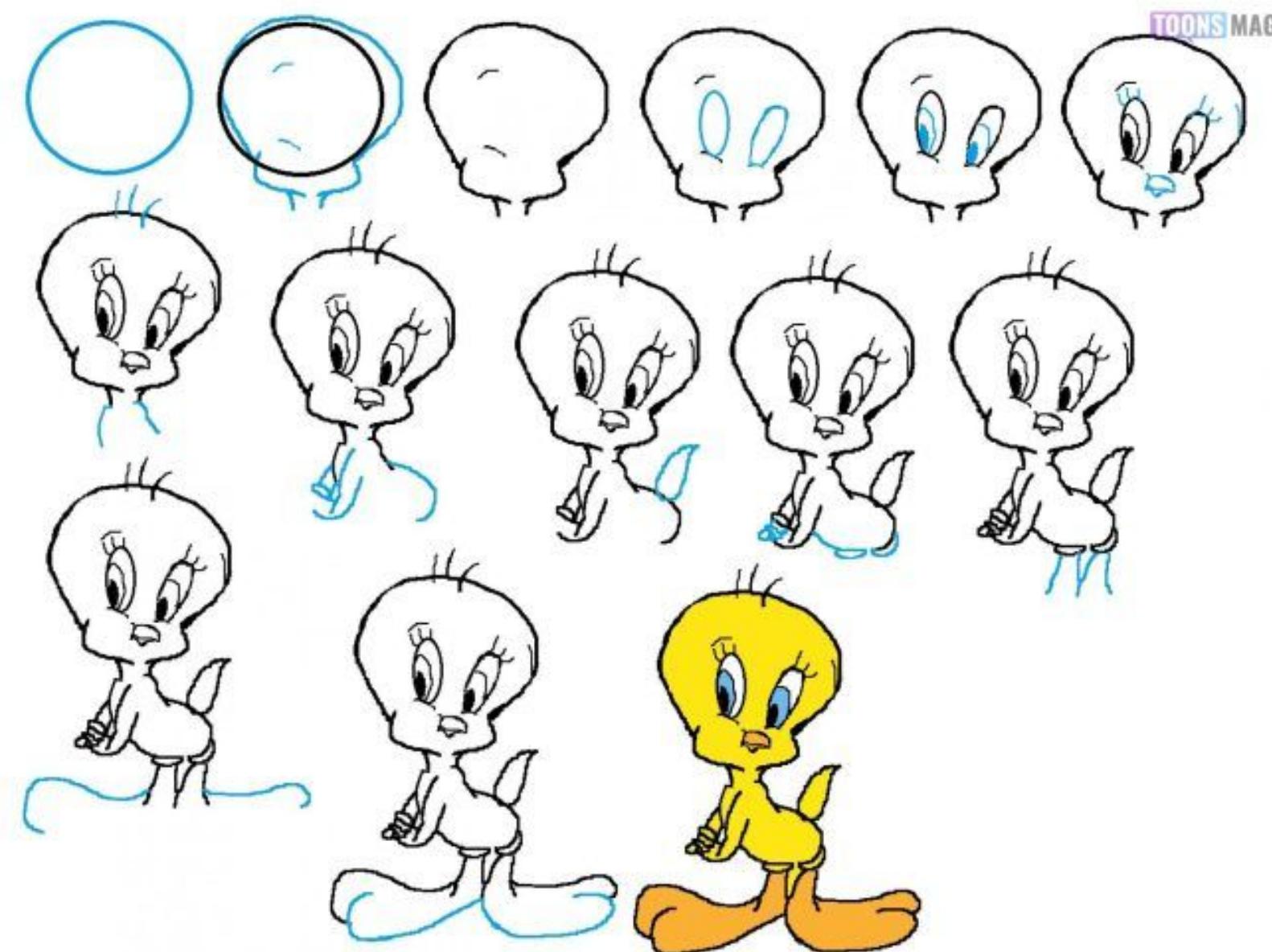
Read the manual.



Reproducibility crisis



Documentation design principles



Easy to follow, step by step instructions on completing specific, well-defined tasks that connect to the background theory.

Not a lot of work to produce



Registering keywords

```
void Coordination::registerKeywords( Keywords& keys ) {  
    CoordinationBase::registerKeywords(keys);  
    keys.add("compulsory","NN","6","The n parameter of the switching function ");  
    keys.add("compulsory","MM","0","The m parameter of the switching function; 0 implies 2*NN");  
    keys.add("compulsory","D_0","0.0","The d_0 parameter of the switching function");  
    keys.add("compulsory","R_0","The r_0 parameter of the switching function");  
    keys.add("optional","SWITCH","This keyword is used if you want to employ an alternative to the continuous switching function defined above. "  
            "The following provides information on the \ref switchingfunction that are available. "  
            "When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.");  
}
```

Glossary of keywords and components

The atoms involved can be specified using

GROUPA

First list of atoms. For more information on how to specify lists of atoms see [Groups and Virtual Atoms](#)

GROUPB

Second list of atoms (if empty, N*(N-1)/2 pairs in GROUPA are counted). For more information on how to specify lists of atoms see [Groups and Virtual Atoms](#)

Compulsory keywords

NN

(default=6) The n parameter of the switching function

MM

(default=0) The m parameter of the switching function; 0 implies 2*NN

D_0

(default=0.0) The d_0 parameter of the switching function

R_0

The r_0 parameter of the switching function

Options

NUMERICAL_DERIVATIVES

(default=off) calculate the derivatives for these quantities numerically

NOPBC

(default=off) ignore the periodic boundary conditions when calculating distances

SERIAL

(default=off) Perform the calculation in serial - for debug purpose

PAIR

(default=off) Pair only 1st element of the 1st group with 1st element in the second, etc

NLIST

(default=off) Use a neighbor list to speed up the calculation

NL_CUTOFF

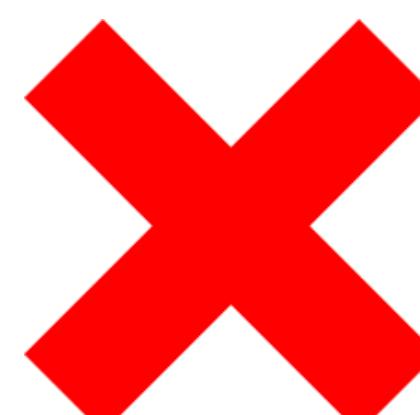
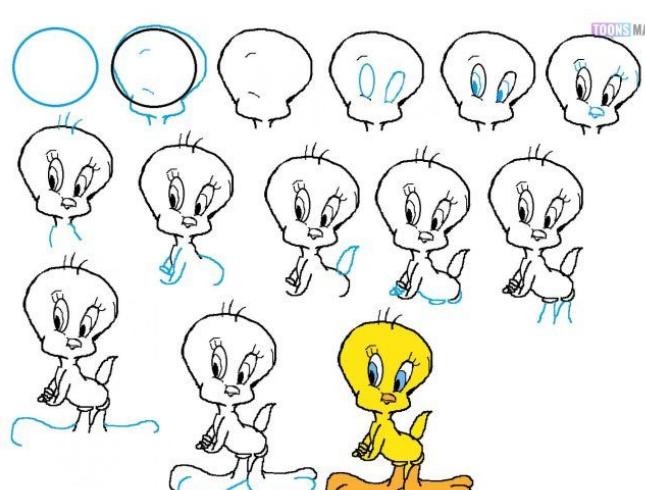
The cutoff for the neighbor list

NL_STRIDE

The frequency with which we are updating the atoms in the neighbor list

SWITCH

This keyword is used if you want to employ an alternative to the continuous switching function defined above. The following provides information on the [switchingfunction](#) that are available. When this keyword is present you no longer need the NN, MM, D_0 and R_0 keywords.

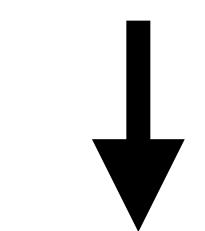


Annotated inputs

Regularly testing to ensure input still works

Link to manual page

Click on the labels of the actions for more information on what each action computes



v2.8 passing

```
d1: DISTANCE ATO ..... the pair of atoms that we are calculating the distance between  
d2: DISTANCE ATOMS=2,4  
d2c: DISTANCE ATOMS=2,4 COMPONENTS  
PRINT ARG=d1,d2,d2c.x
```

Click here

The DISTANCE action with label **d2c** calculates the following quantities:

v2.8 passing

Quantity Description

- d2c.x the x-component of the vector connecting the two atoms
- d2c.y the y-component of the vector connecting the two atoms
- d2c.z the z-component of the vector connecting the two atoms

```
d1: DISTANCE ATOMS=3,5  
d2: DISTANCE ATOMS=2,4  
d2c: DISTANCE ATOMS=2,4 COMPONENTS  
PRINT ARG=d1,d2,d2c.x
```

The future of annotated inputs

```
c: COORDINATIONNUMBER SPECIES=1-64 SWITCH={RATIONAL D_0=3.0 R_0=1.5} MEAN  
PRINT ARG=c.* FILE=colv2
```

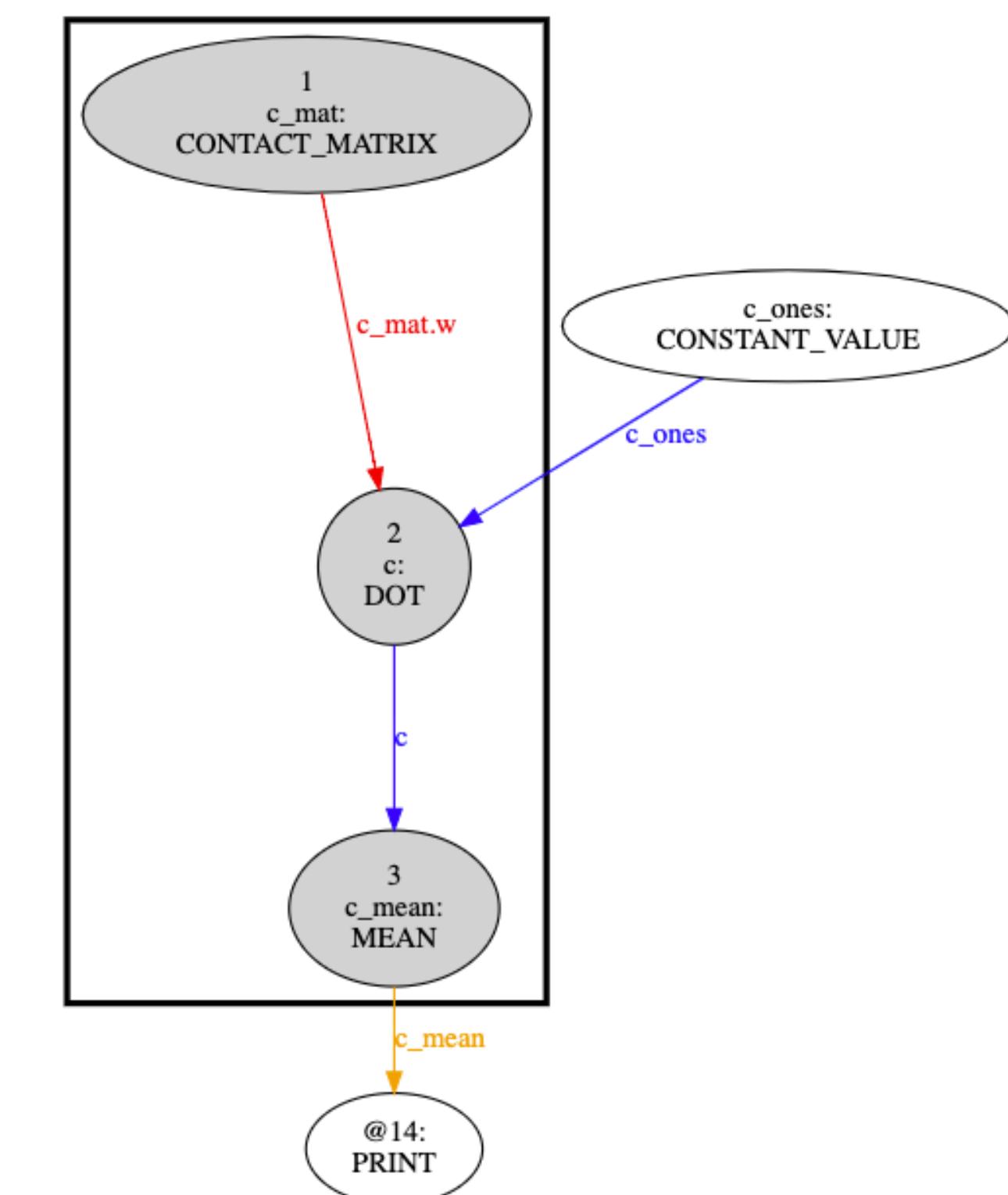
```
c_mat: CONTACT_MATRIX GROUP=1-64 SWITCH={RATIONAL D_0=3.0 R_0=1.5}
```

```
c_ones: ONES SIZE=64
```

```
c: MATRIX_VECTOR_PRODUCT ARG=c_mat.w,c_ones
```

```
c_mean: MEAN ARG=c PERIODIC=NO
```

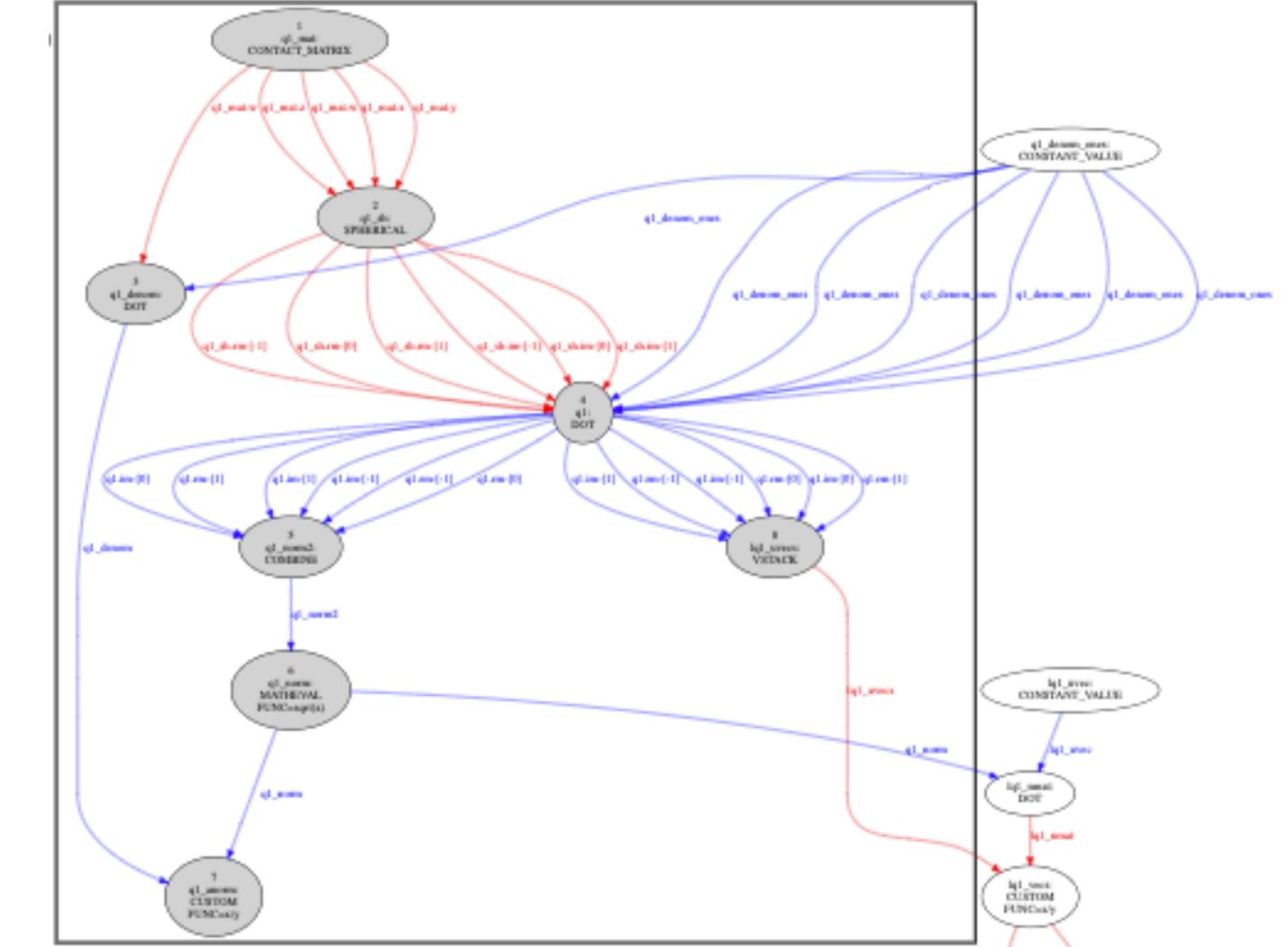
```
PRINT ARG=c.* FILE=colv2
```



q1: Q1 SPECIES=1-64 SWITCH={RATIONAL D_0=3.0 R_0=1.5}

lq1: LOCAL_Q1 SPECIES=q1 SWITCH={RATIONAL D_0=3.0 R_0=1.5} MEAN

PRINT ARG=lq1.* FILE=colv2



q1_mat: CONTACT_MATRIX GROUP=1-64 COMPONENTS SWITCH={RATIONAL D_0=3.0 R_0=1.5}

q1_denom_ones: CONSTANT_VALUE VALUES=1,1,1,1,1...

q1_denom: MATRIX_VECTOR_PRODUCT ARG=q1_mat.w,q1_denom_ones ← Just coordination number again

q1_sh: SPHERICAL_HARMONIC ARG1=q1_mat.x ARG2=q1_mat.y ARG3=q1_mat.z ARG4=q1_mat.w L=1

q1: MATRIX_MATRIX_PRODUCT ...

ARG1=q1_sh.rm-[-1],q1_sh.im-[-1],q1_sh.rm-[0],q1_sh.im-[0],q1_sh.rm-[1],q1_sh.im-[1]

ARG2=q1_denom_ones,q1_denom_ones,q1_denom_ones,q1_denom_ones,q1_denom_ones,q1_denom_ones

...

q1_norm2: COMBINE ...

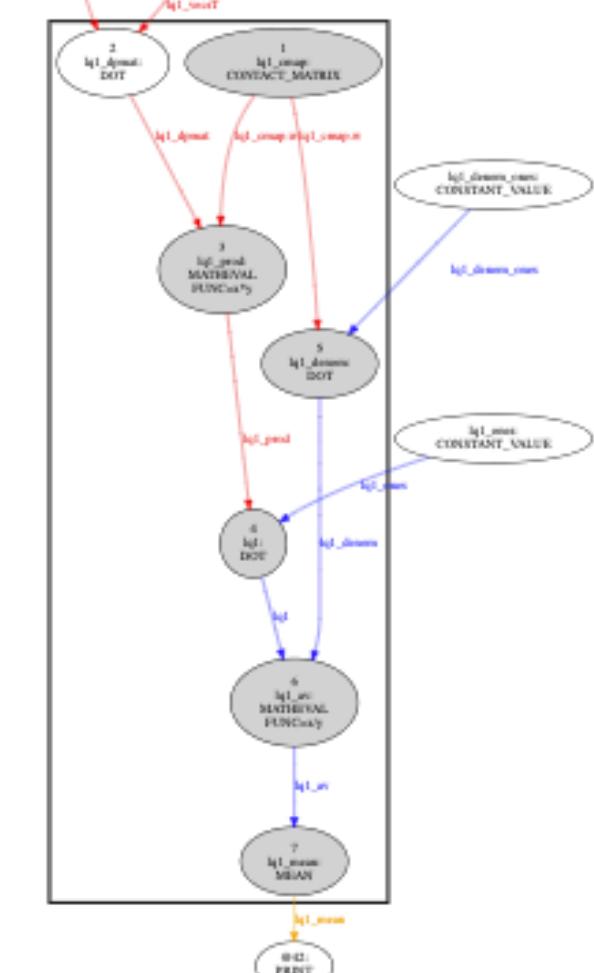
ARG=q1.rm-[-1],q1.im-[-1],q1.rm-[0],q1.im-[0],q1.rm-[1],q1.im-[1]

POWERS=2,2,2,2,2,2

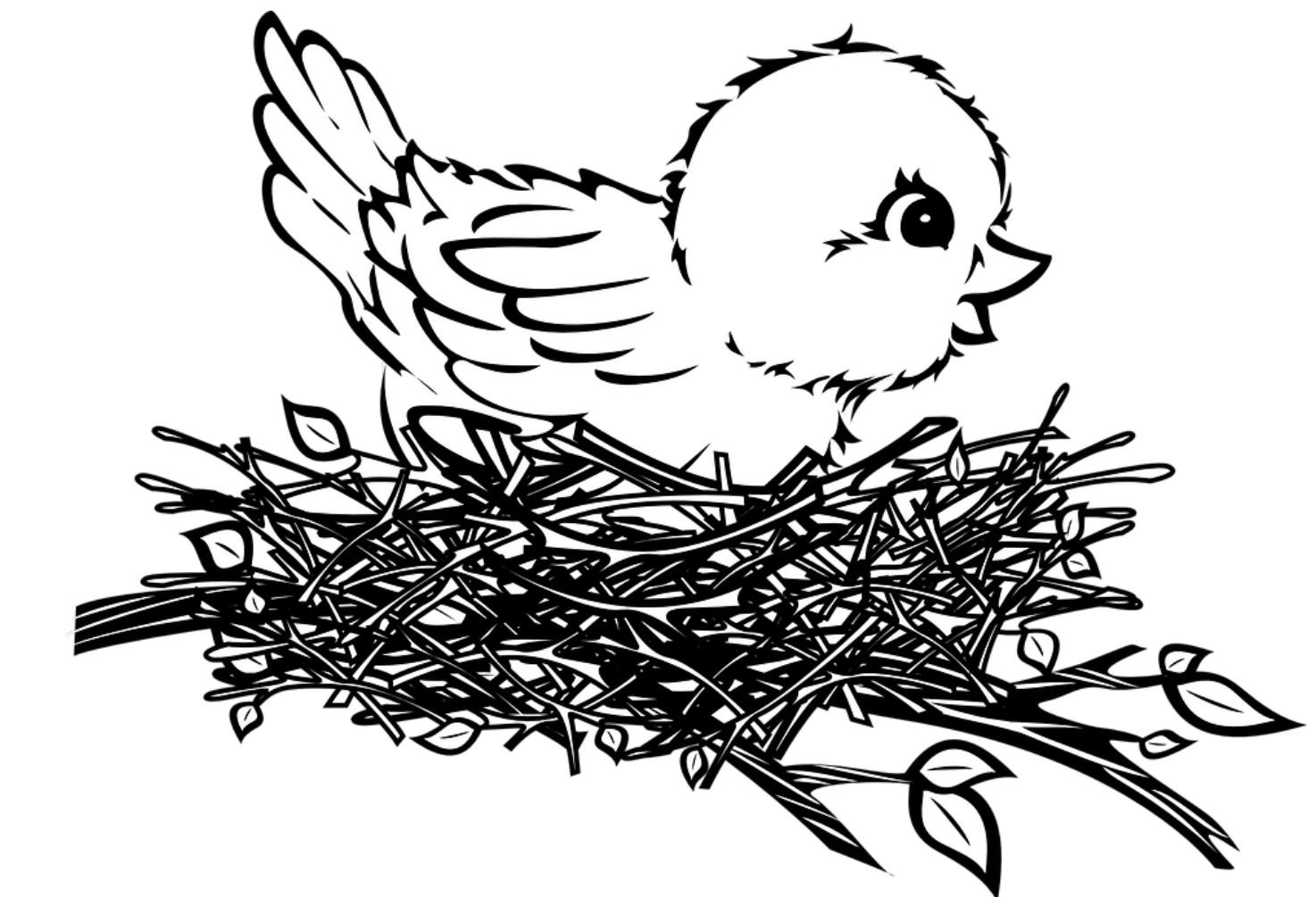
...

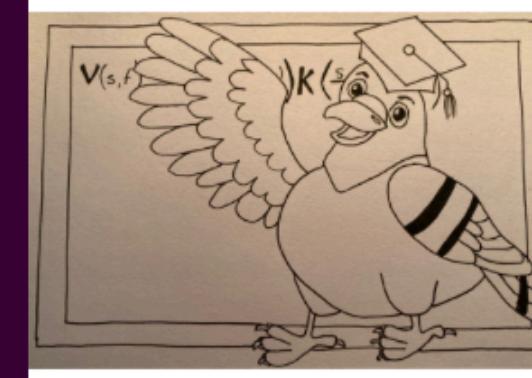
q1_norm: CUSTOM ARG=q1_norm2 FUNC=sqrt(x) PERIODIC=NO ← Length of vector

q1_anorm: CUSTOM ARG1=q1_norm ARG2=q1_denom FUNC=x/y PERIODIC=NO ← Divide by number of bonds from central atom



Documentation = tools to automatically
generate annotations for the inputs in
the PLUMED nest





PLUMED-TUTORIALS

A living textbook for and by the PLUMED consortium

[Home](#)[News](#)[PLUMED](#)[Browse lessons](#)[Going further](#)[Contribute](#)

Contributions work like the
PLUMED nest

Authors acknowledged
and keep control over
content

Browse the lessons

The lessons that have been submitted to the PLUMED-TUTORIALS are listed below. PLUMED-TUTORIAL monitors whether PLUMED input files in these lessons are compatible with the current and development versions of the code and integrates links from these files to the PLUMED manual.

Showing 1 to 10 of 17 entries

Show entries

ID	Name	Instructors	Description
23.001	Developments in PLUMED	Tribello G.	A series of articles were I outline some development work that I have been doing with PLUMED over the last few years
22.013	SASA module - The solvent accessible surface area of proteins as a collective variable, and the application of PLUMED for implicit	Arsiccio A.	An introduction to the SASA module and a description of how PLUMED can be used for implicit



masterclass-21-2

Public

Edit Pins ▾

Unwatch 6 ▾

main ▾

1 branch

0 tags

Go to file

Add file ▾

<> Code ▾

Gareth Aneurin Tribello and Gareth Aneurin Tribello

data

figures

notebooks

slides

work

EMBED.yml

NAVIGATION.md

README.md

Added files fo

Added final fi

Fixed notebook

Fixed notebook

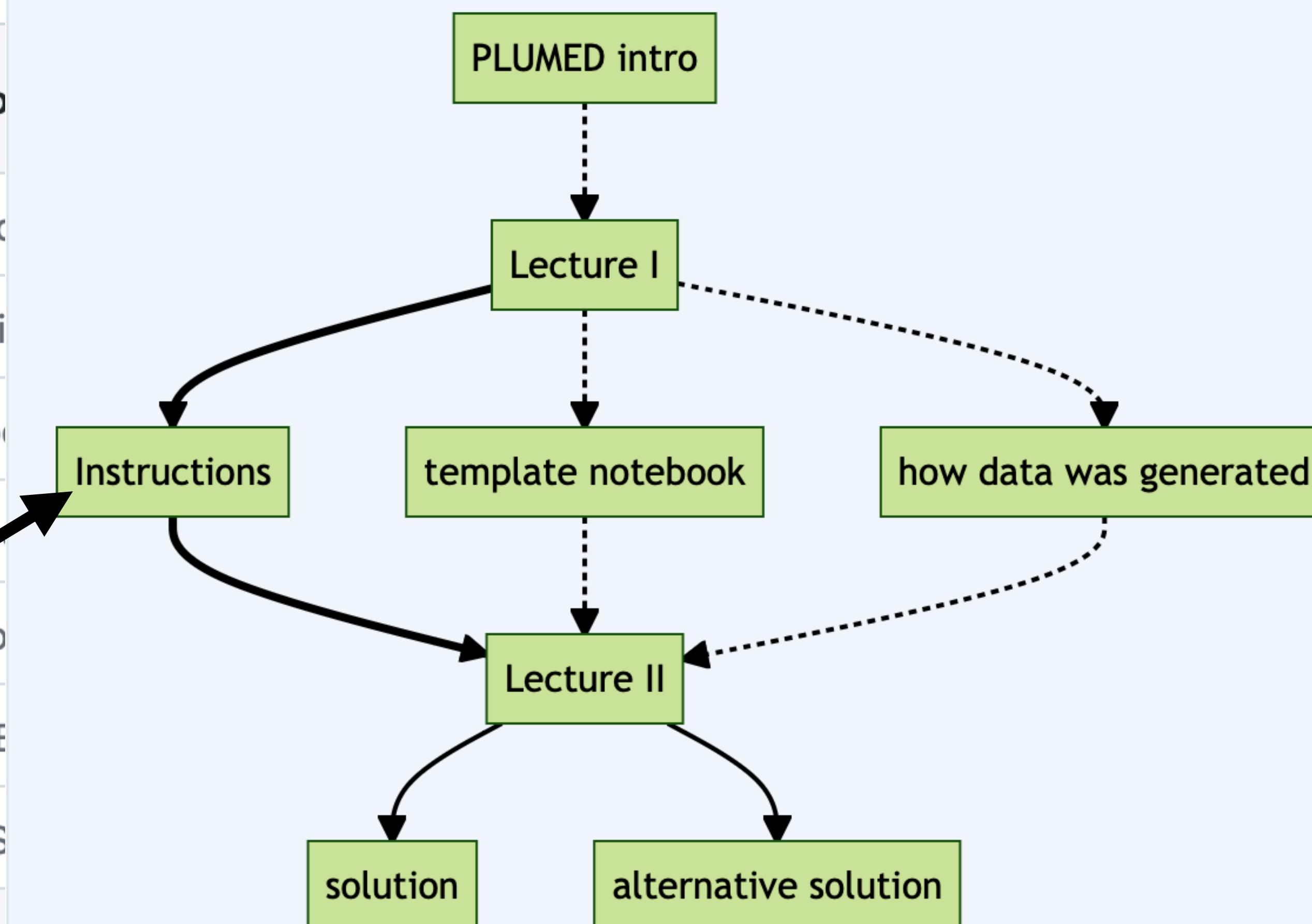
Final steps to

Update EMBED

Renamed INS

Renamed INS

Instructions
for creating
graph



obtain is a sample from a multinomial distribution so we can estimate parameters for the

multinomial by using likelihood max

can then calculate the free energy,

If we estimate $P(s')$ using likelihood maximisation we can thus get an estimate of the free energy surface. To estimate the free energy surface in this way we can use the input file:

Click on the labels of the actions for

```
# We use natural units here so that
UNITS NATURAL
data: READ FILE=__FILL__ VALUES=__FILL__
hhh: HISTOGRAM ARG=__FILL__ STRIDE=1
fes: CONVERT_TO_FES GRID=__FILL__ TEMP=1 # This sets k_B T = 1
DUMPGRID GRID=__FILL__ FILE=fes.dat
```

Copy this input to a file called plumed.dat, fill in the blanks so that you have a grid that runs from -4 to +4 with 100 bins. Run the command:

```
> plumed driver --noatoms
```

The --noatoms flag here is needed since plumed driver is commonly used to analyze a trajectory of atomic coordinates, but here we are directly analyze a collective variable.

You should see that the free energy curve for this data set looks like this:

```
![The free energy as a function of the CV for the data in uncorrelated data](figures/masterclass-21-2-fes1.png)
```

As you can see, there is a single minimum in this free energy surface.

Exercise 3: Calculating the fluctuations for a CV

Physical systems spend the majority of their time fluctuating around minima in the free energy landscape. Let's suppose that this minima is at μ and lets use the Taylor series to write an expression for the free energy at s' as follows:

```
$$
```

You should see that the free energy curve for this data set looks like this:

```
masterclass-21-2 — vi README.md — 149x54

The fluctuations in the average get smaller as this quantity is computed from larger numbers of random variables. We say that the average thus converges to the ensemble average, which is zero for the graph above.

#### Exercise 2: Calculating the free energy

We can estimate the distribution for our CV,  $P(s')$ , by calculating [a histogram](https://www.notion.so/Histogram-2d2527795f0140008b318d3bc958ee4c). The histogram we obtain is a sample from [a multinomial distribution](https://www.notion.so/Multinomial-distribution-f378b34a38564f6999e184cd2df53f5c) so we can estimate parameters for the multinomial by using [likelihood maximisation](https://www.notion.so/Maximum-likelihood-e74a16b691284896bb2bbd5cff502f6d). Once we have the marginal distribution,  $P(s')$  we can then calculate the [free energy](https://www.notion.so/Thermodynamic-potentials-dd0dcbab15cd437b91e803ac7fdb1bf8),  $F(s')$  as a function of  $s(x)$  as  $F(s')$  is related to  $P(s')$  by:

$$
F(s') = - k_B T \ln P(s')
$$

If we estimate  $P(s')$  using likelihood maximisation we can thus get an estimate of the free energy surface. To estimate the free energy surface for the data in `uncorrelated_data` using PLUMED in this way we can use the input file:

```plumed
#SOLUTIONFILE=work/plumed_ex2.dat
We use natural units here so that kBT is set to 1
UNITS NATURAL
data: READ FILE=__FILL__ VALUES=__FILL__
hhh: HISTOGRAM ARG=__FILL__ STRIDE=1 __FILL__=-4.5 __FILL__=4.5 __FILL__=100 __FILL__=DISCRETE
fes: CONVERT_TO_FES GRID=__FILL__ TEMP=1 # This sets k_B T = 1
DUMPGRID GRID=__FILL__ FILE=fes.dat
```
Copy this input to a file called plumed.dat, fill in the blanks so that you have a grid that runs from -4 to +4 with 100 bins. Run the calculation by executing the following command:
```
> plumed driver --noatoms
```

The `--noatoms` flag here is needed since `plumed driver` is commonly used to analyze a trajectory of atomic coordinates, but here we are directly analyzing a collective variable.

You should see that the free energy curve for this data set looks like this:

![The free energy as a function of the CV for the data in uncorrelated data](figures/masterclass-21-2-fes1.png)

As you can see, there is a single minimum in this free energy surface.

#### Exercise 3: Calculating the fluctuations for a CV

Physical systems spend the majority of their time fluctuating around minima in the free energy landscape. Let's suppose that this minima is at  $\mu$  and lets use the Taylor series to write an expression for the free energy at  $s'$  as follows:
```



masterclass-21-2

Public

Edit Pins ▾

Unwatch 6 ▾

main ▾

1 branch

0 tags

Go to file

Add file ▾

<> Code ▾

Gareth Aneurin Tribello and Gareth Aneurin Tribello

data

figures

notebooks

slides

work

EMBED.yml

NAVIGATION.md

README.md

Added files fo

Added final fi

Fixed notebook

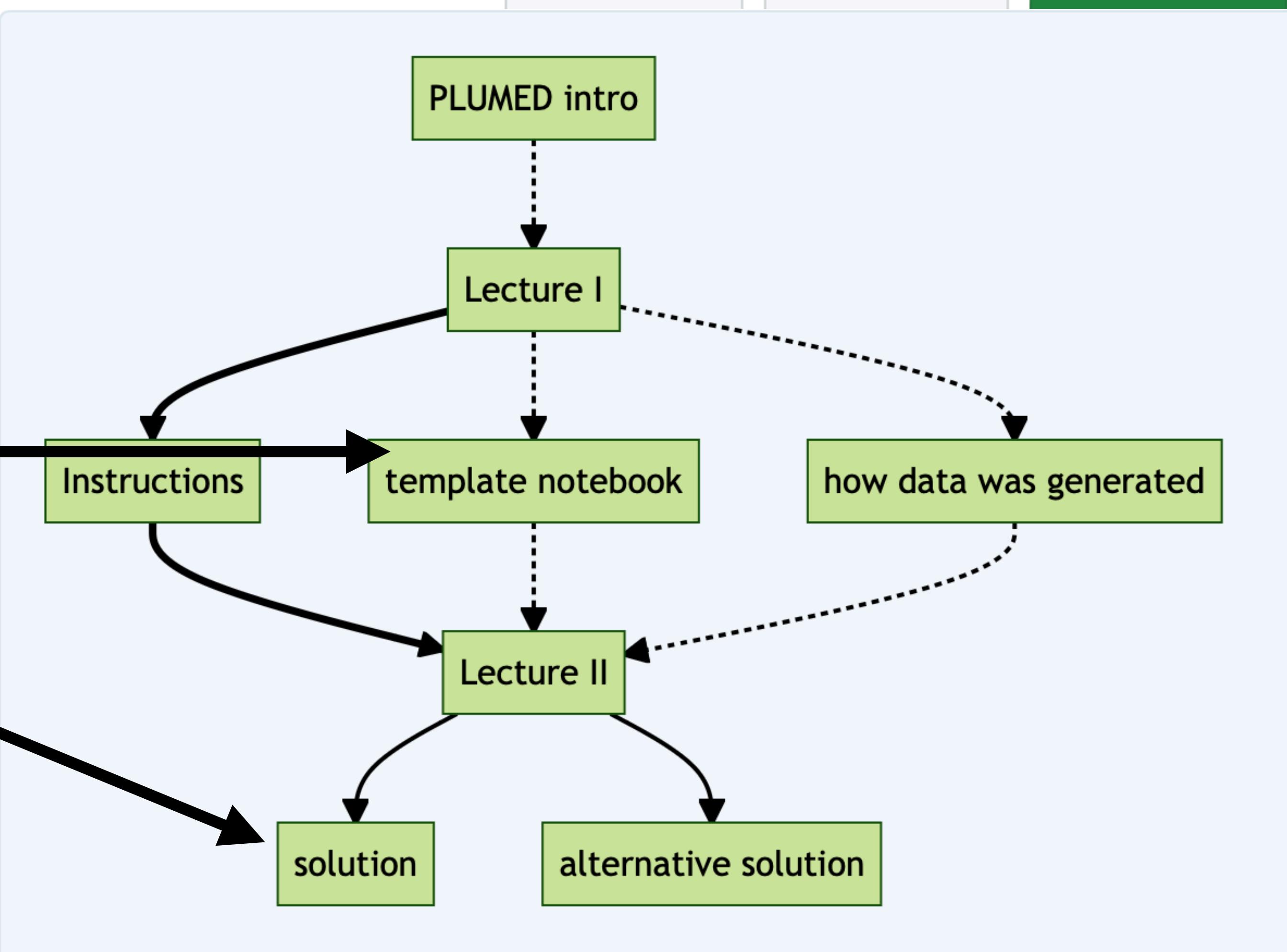
Fixed notebook

Final steps to

Update EMB

Renamed INS

Renamed INS





```
In [1]: # Import some useful python libraries
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
```

PLUMED masterclass 21.2: Statistical errors in MD (template)

I completed all the exercises for the masterclass in a python notebook that was similar to this one. In this template I will explain how I was able to do this.

Running PLUMED

To run PLUMED from the python notebook I first set a variable equal to the string that contains the PLUMED input like this:

```
In [2]: pex0 = '''d1: READ FILE=data VALUE=d1
PRINT ARG=d1 FILE=colvar'''
```

I can now create a directory to run PLUMED in, output the input above to a plumed.dat file in that directory and then run PLUMED as follows.

```
In [ ]: # Make a directory to hold the exercise
!mkdir ../Exercises/Exercise_0
# Output the input file above to the plumed.dat file
f = open("../Exercises/Exercise_0/plumed.dat", "w")
f.write( pex0 )
f.close()
# Now run PLUMED driver
!cd ../Exercises/Exercise_1 && plumed driver --noatoms > /dev/null
```

I kept my notebook with the solutions in the solutions directory of the GitHub repository you downloaded at the start of the exercise. When the exercise was completed I thus had a new directory called Exercises in the repo that contained the solutions to all the exercises in directories called Exercise_1, Exercise_2 and so on.

Reading and plotting the colvar files

You can read in and plot any colvar files that PLUMED produces by using a variant on the command shown below:

```
In [4]: data = np.loadtxt("../data/correlated_data")

plt.plot( data[:,0], data[:,1], 'ko' )
plt.xlabel("Time")
plt.ylabel("Random variable value")
plt.show()
```





masterclass-21-2

Public

Edit Pins ▾

Unwatch 6 ▾

main ▾

1 branch

0 tags

Go to file

Add file ▾

<> Code ▾

Gareth Aneurin Tribello and Gareth Aneurin Tribello

data

Added files fo

figures

Added final fi

notebooks

Fixed notebook

slides

Fixed notebook

work

Final steps to

EMBED.yml

Update EMBED

NAVIGATION.md

Renamed INS

README.md

Other PLUMED
tutorials

External
resources

Instructions

PLUMED intro

Lecture I

template notebook

how data was generated

Lecture II

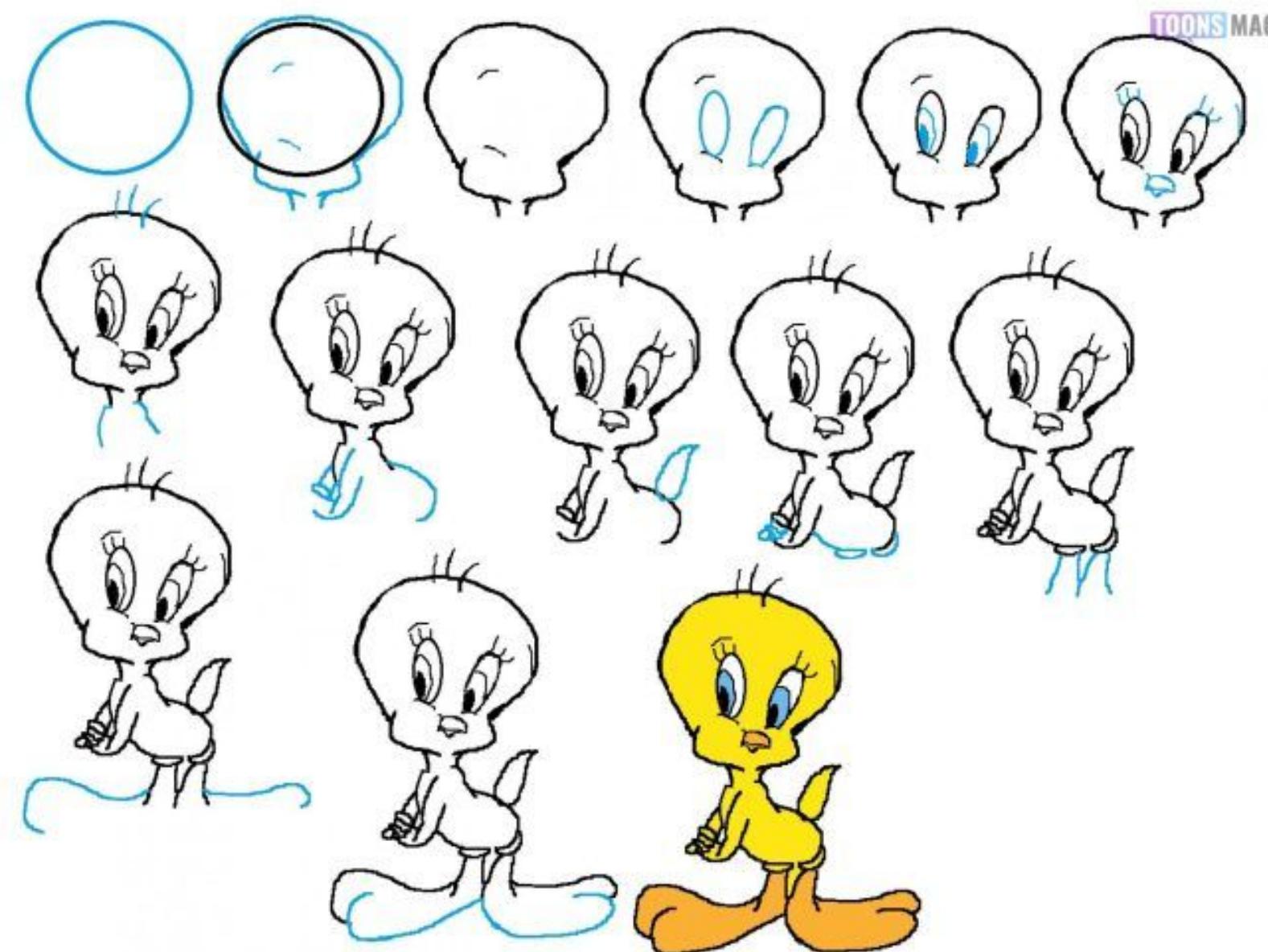
solution

alternative solution



```
video1:  
    title: Lecture I  
    location: https://www.youtube.com/embed/dJKajNwbJ74  
video2:  
    title: Lecture II  
    location: https://www.youtube.com/embed/XKqis1C2GYA  
ref1:  
    location: 21/001  
    type: internal
```

Documentation design principles



Easy to follow, step by step instructions on completing specific, well-defined tasks that connect to the background theory.

Not a lot of work to produce

