Command line:

TOPS program → | train file name → | contains params of training procedure → | new HMM model

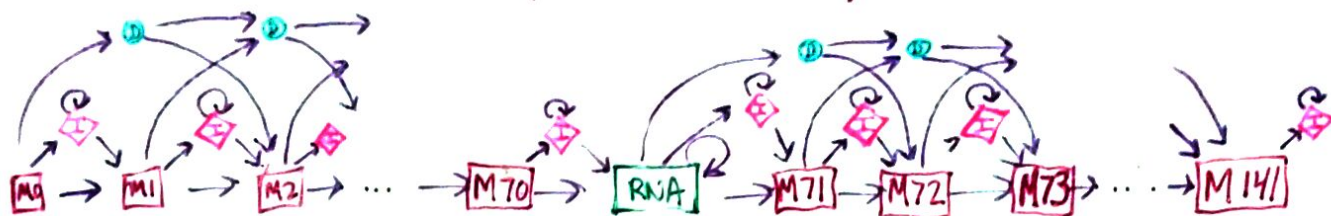train -c   train.txt > model.txt

---
train.txt

```
training_algorithm = "BaumWelchHMM"
training_set = "trainhmm-bw.sequences" ← - format?
initial_specification = "initial-hmm.txt"
maxiter = 300 ← - Max # iterations
```

---
initial-hmm.txt

```
model-name = "HiddenMarkovModel"
state_names =        ("I", "D", "R", "M0", ..."M141"; "Mr0", ..."Mr141")
observation_symbols = ("A", "C", "G", "T")
# transition probabilities
transitions = ("I"  | "M0": 0.05;
               "D"  | "M0": 0.05;
               "M"  | "M0": 0.90;
               "I"  | "M0": 0.5;
               "D"  | "M0": 0.5;
                     ⋮            )
               etc

#emission probabilities
emission-probabilities = ("A" | "R": 0.25;
                          "C" | "R": 0.25;
                          "G" | "R": 0.25;
                          "T" | "R": 0.25;
                              ⋮
                          etc          )

initial_probabilities = ("I": __; "D": __; "R": __; ...etc)
```
probs

One cell line (ex 2_B01):

Legend:
- Delete
- Insert
- RNA insert
- Match states 0 to 141
- → transitions

5'adapter + 5' barcode

3' barcode

assuming adapter is a 5' adapter not a 3' adapter

String	length		
	2B01	3C01	4D01
adapter	45	45	45
5p barcode	25	25	25
3p barcode	71	71	71
adapter + 5pbar	70	70	70
adapter + 3pbar	116	116	116
adap + 3p+5p	141	141	141