Speech Processing

Simon King
University of Edinburgh

additional class slides for 2020-21

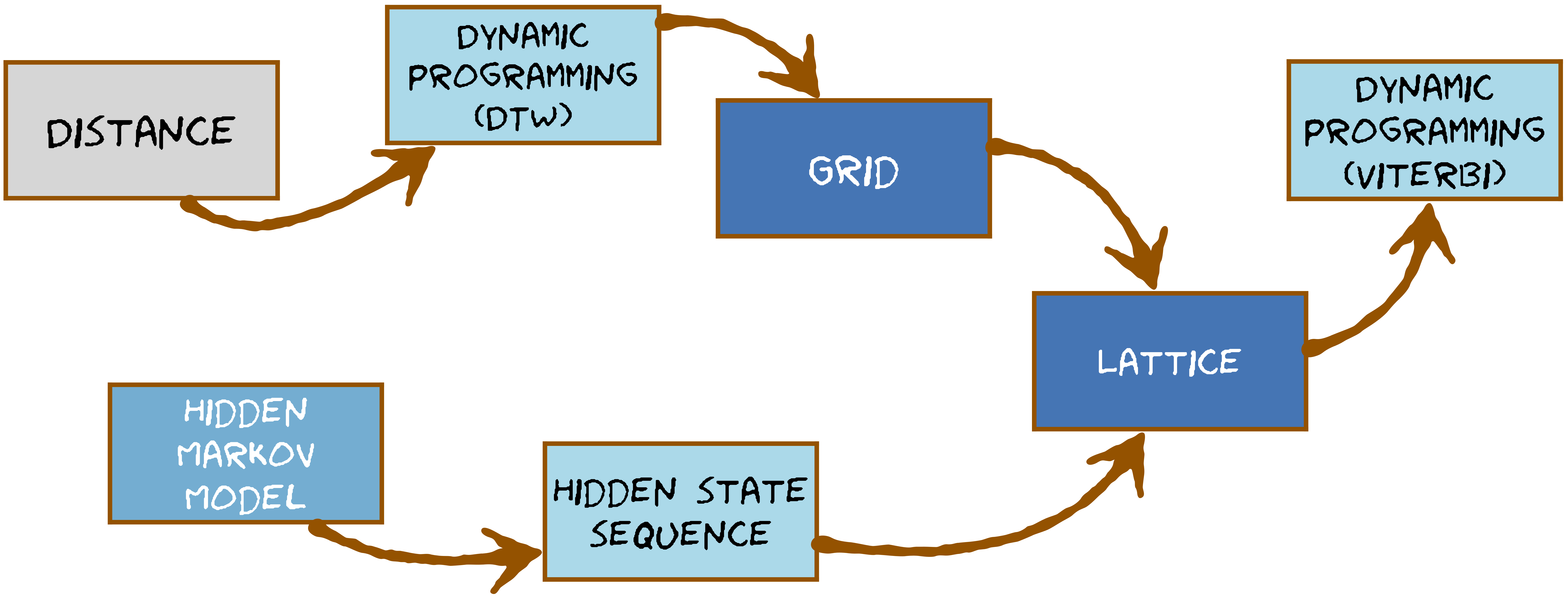
Module 8

The Hidden Markov Model

Orientation

- We've arrived at HMMs
  - Pattern **matching**
  - Extracting **features** from speech
  - Probabilistic **generative** modelling

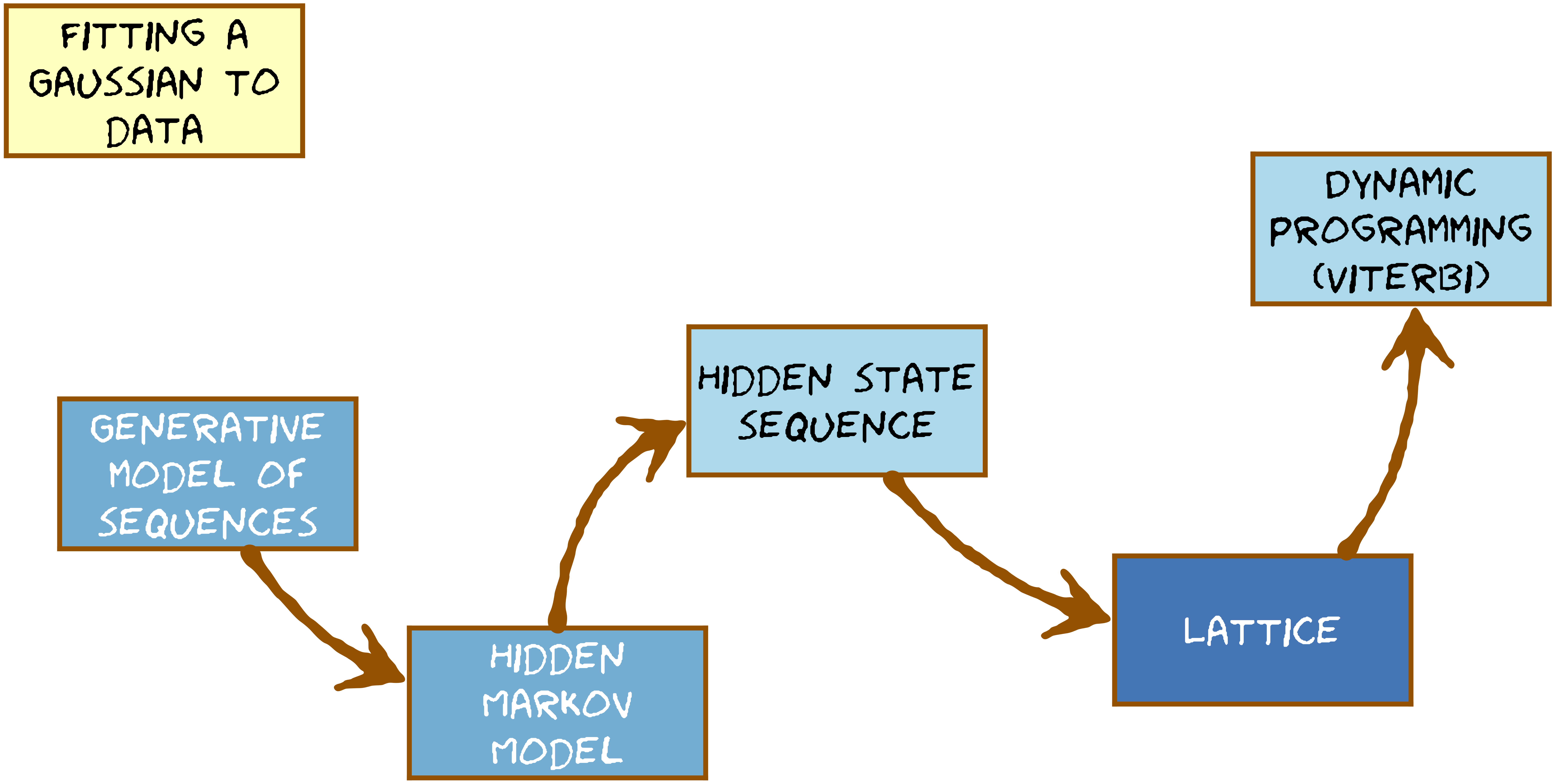
Orientation: from Dynamic Time Warping to the Hidden Markov Model



What you should already know

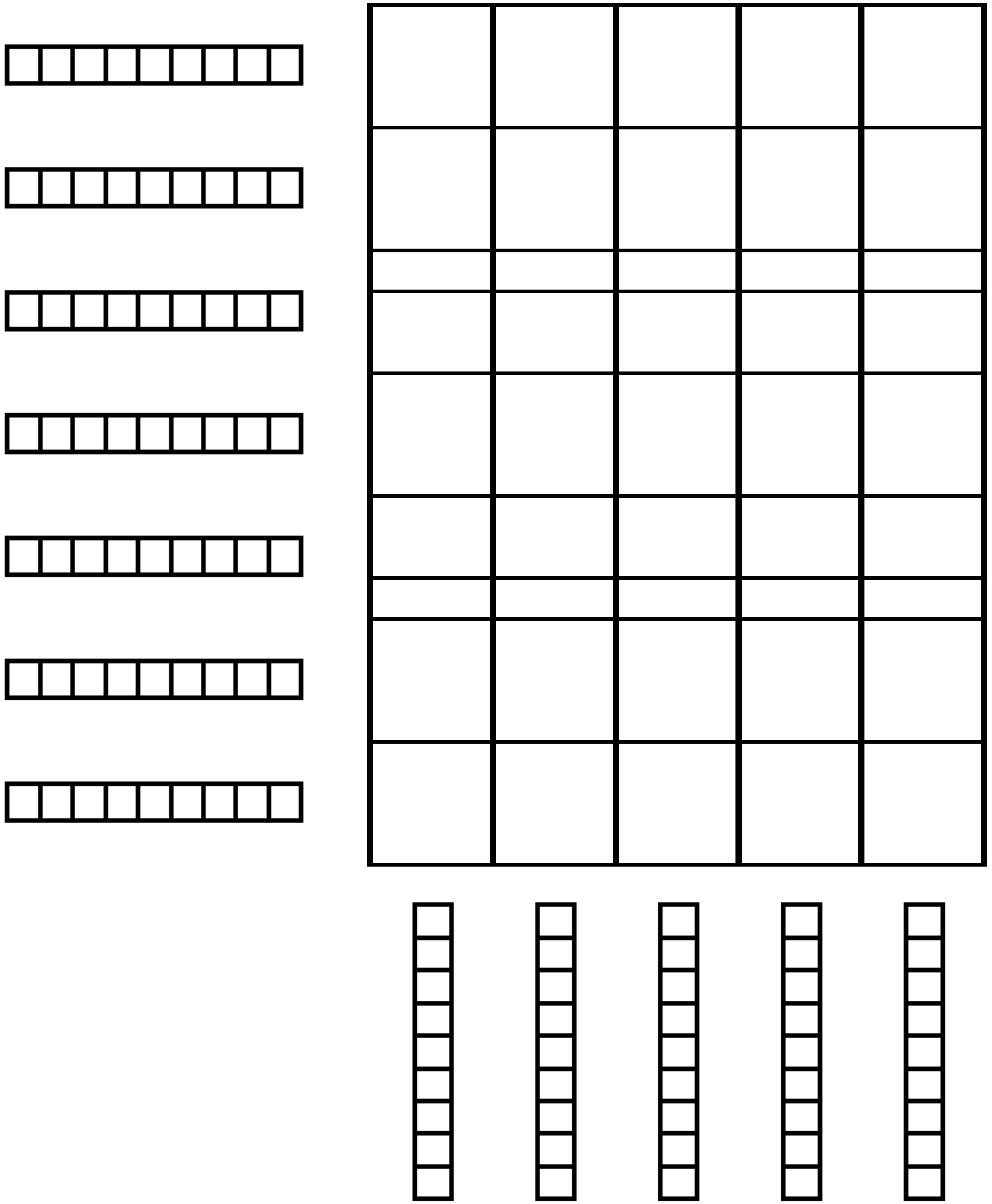
- a single template (as in DTW) cannot capture the natural **variability** of speech
  - *an old-fashioned solution was to store multiple templates*

- a much better solution is to capture variability using **statistics**
  - essentially: mean & variance

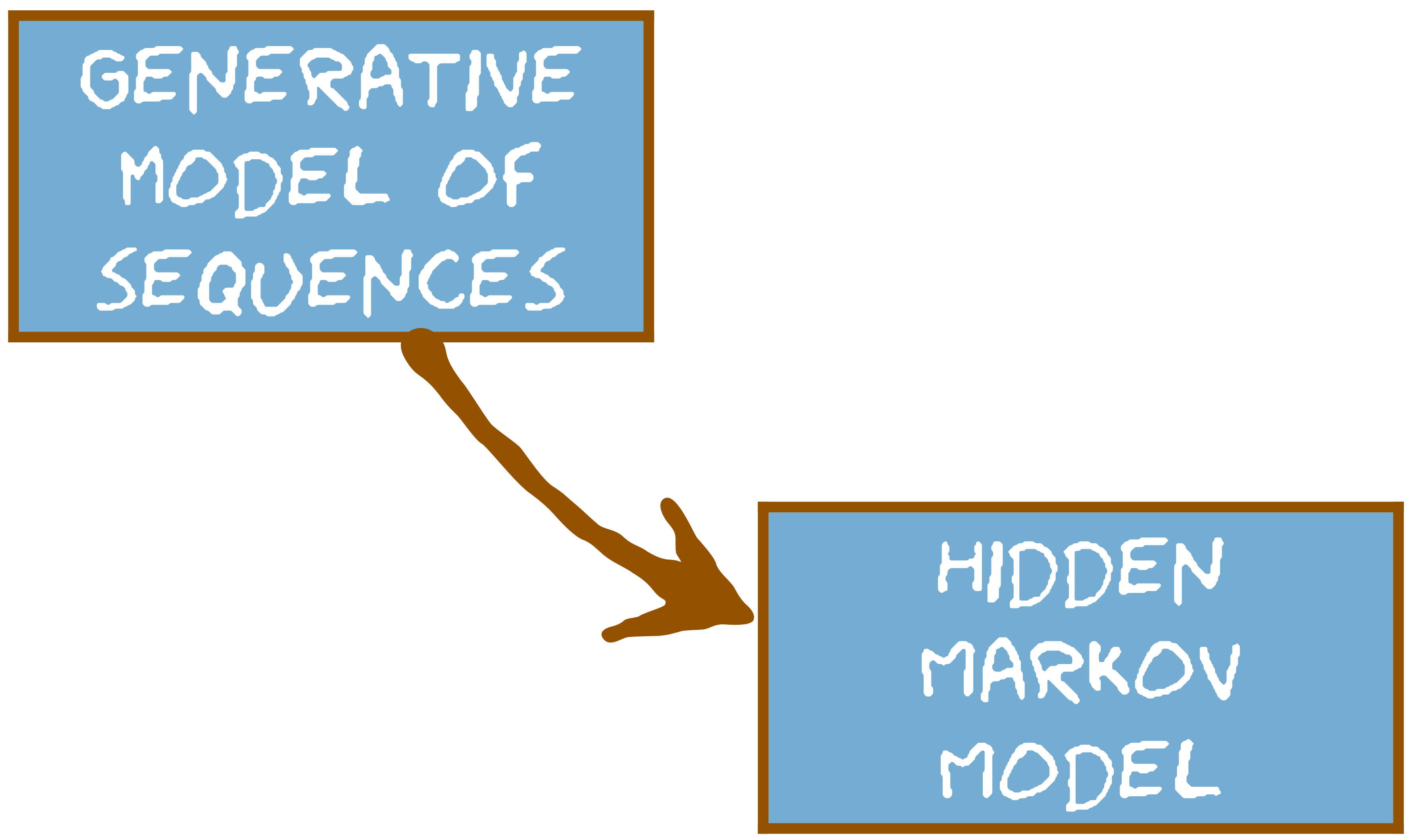

FITTING A GAUSSIAN TO DATA
GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL
HIDDEN STATE SEQUENCE
LATTICE
DYNAMIC PROGRAMMING (VITERBI)

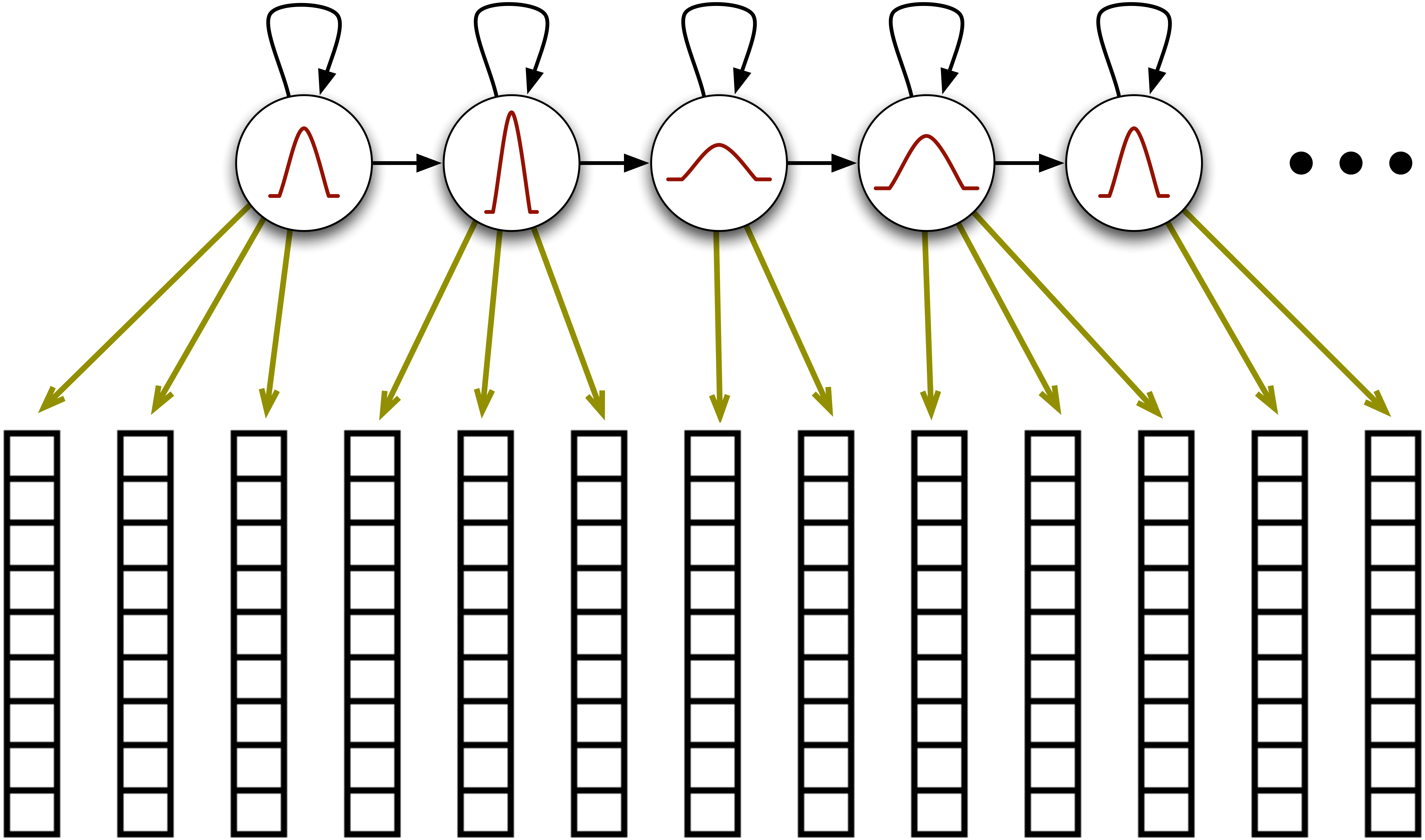
GENERATIVE MODEL OF SEQUENCES

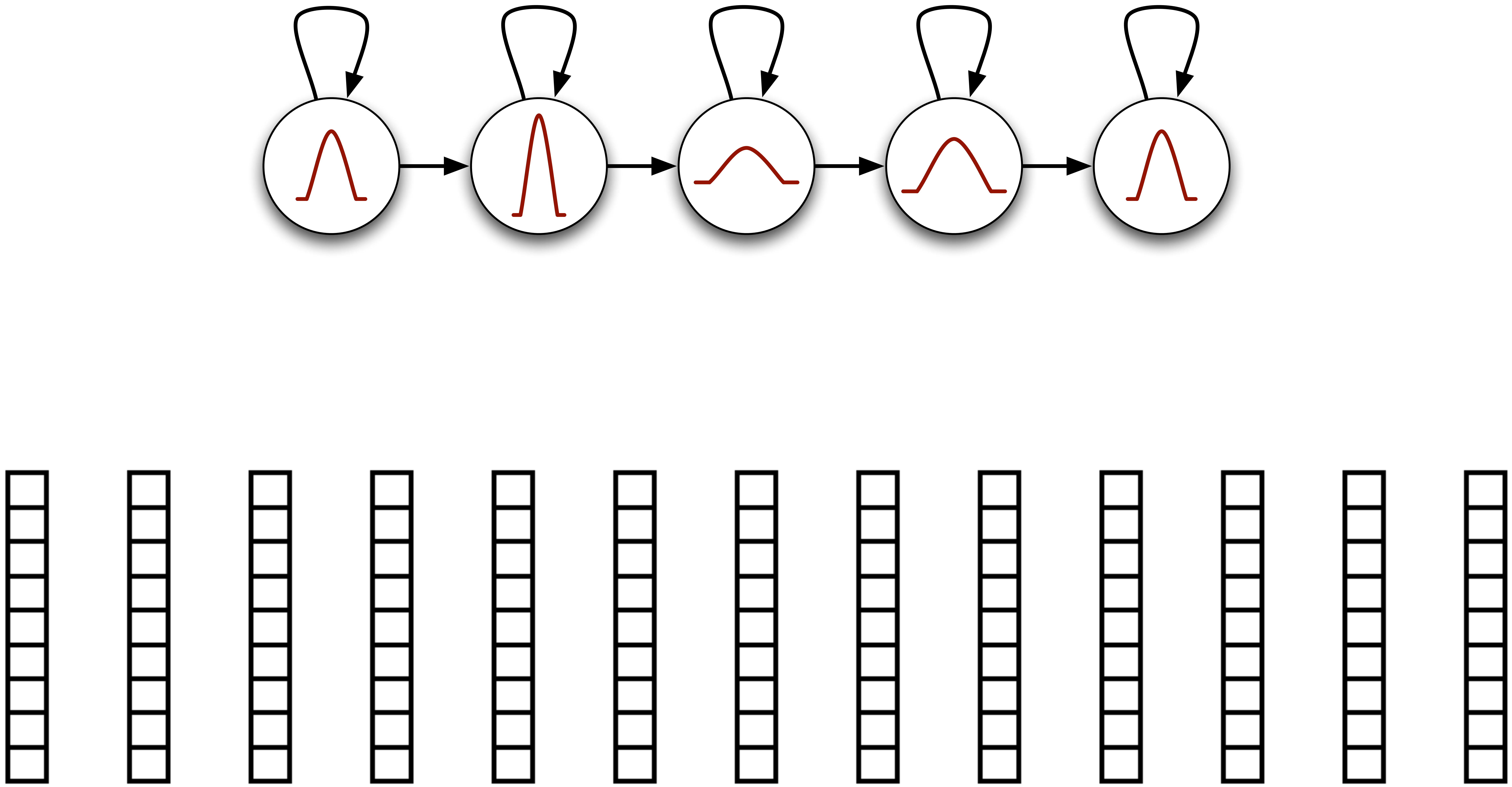
Recap: the multivariate Gaussian as a generative model

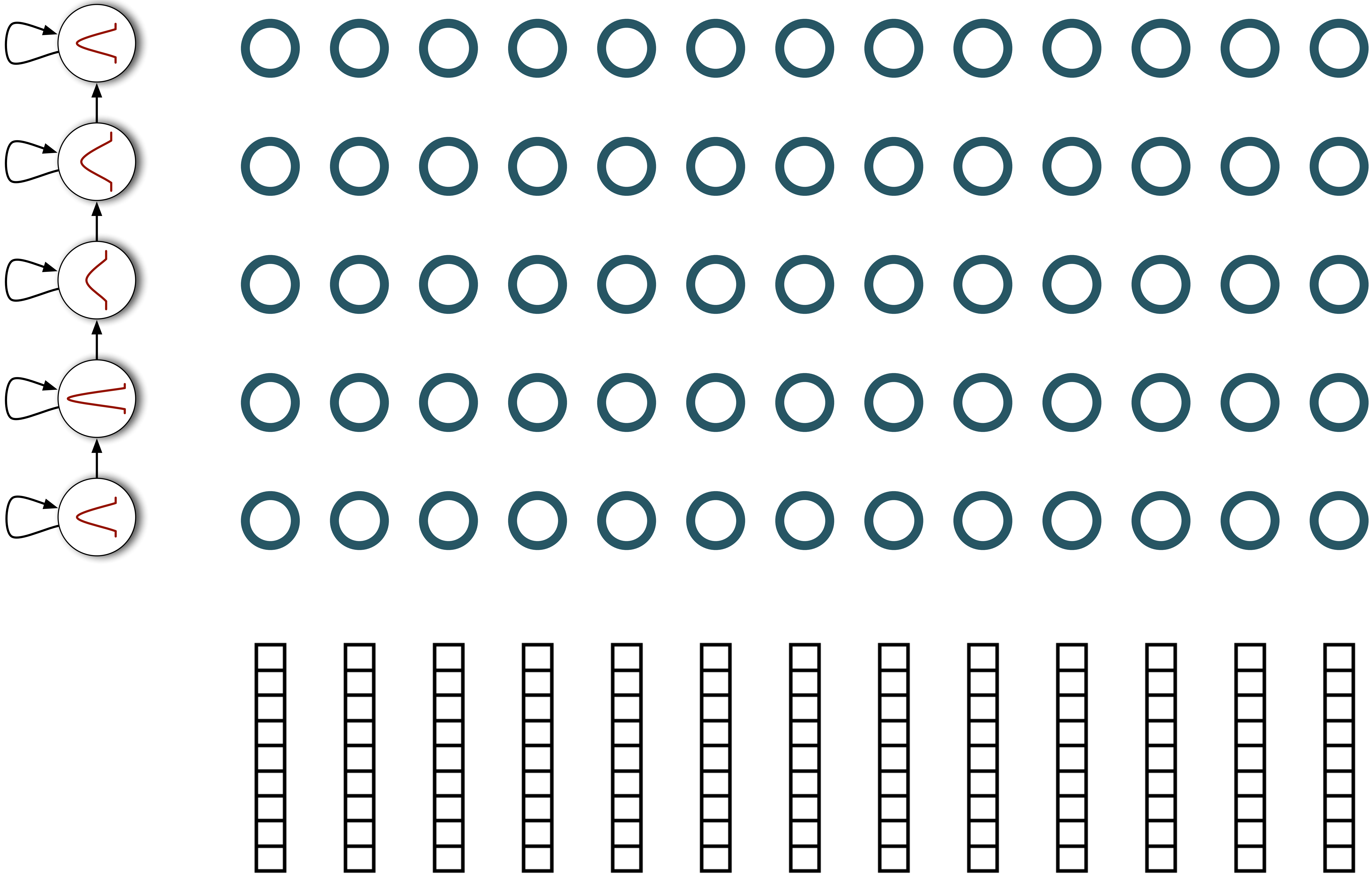


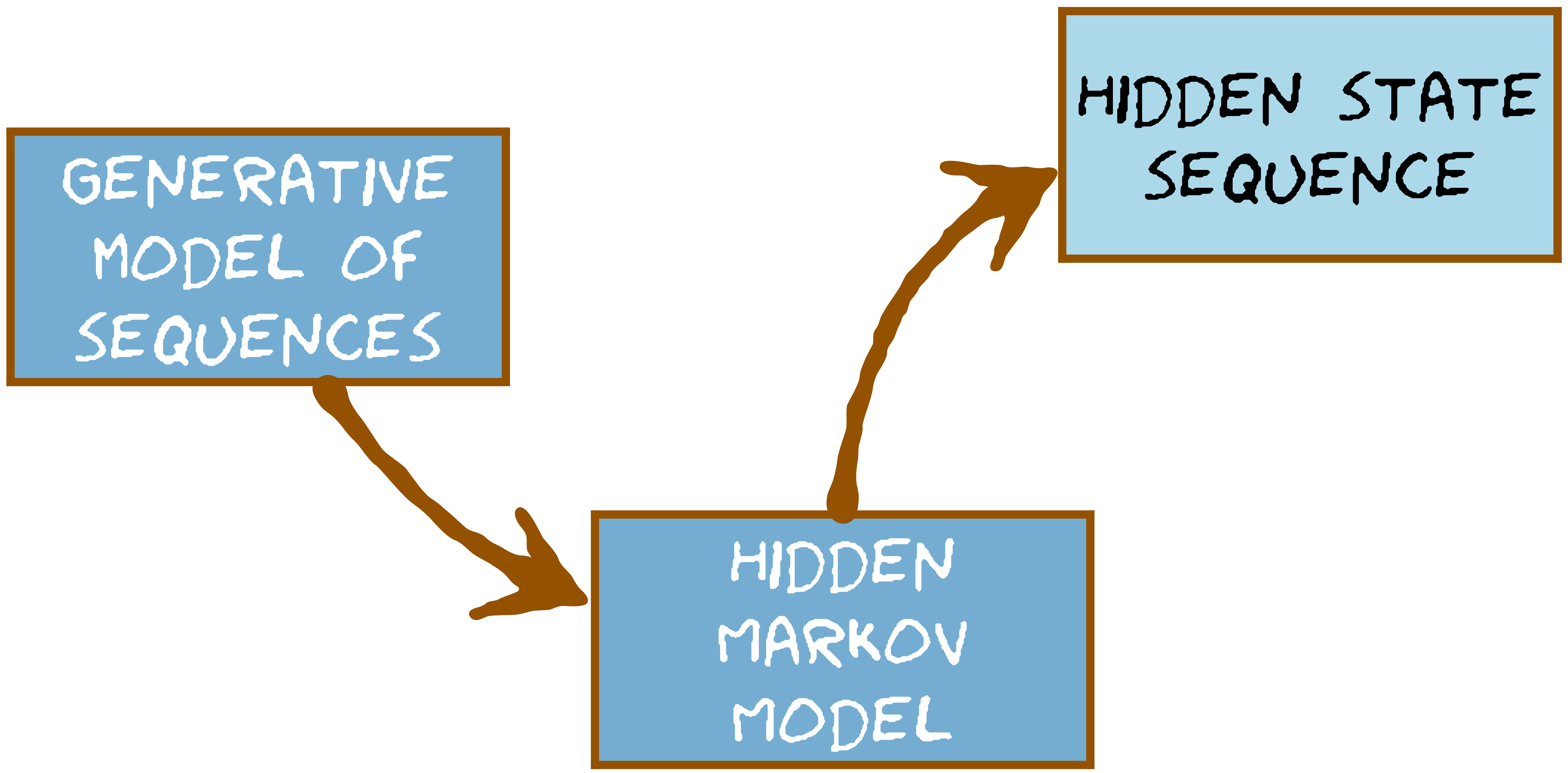
DYNAMIC PROGRAMMING (DTW)

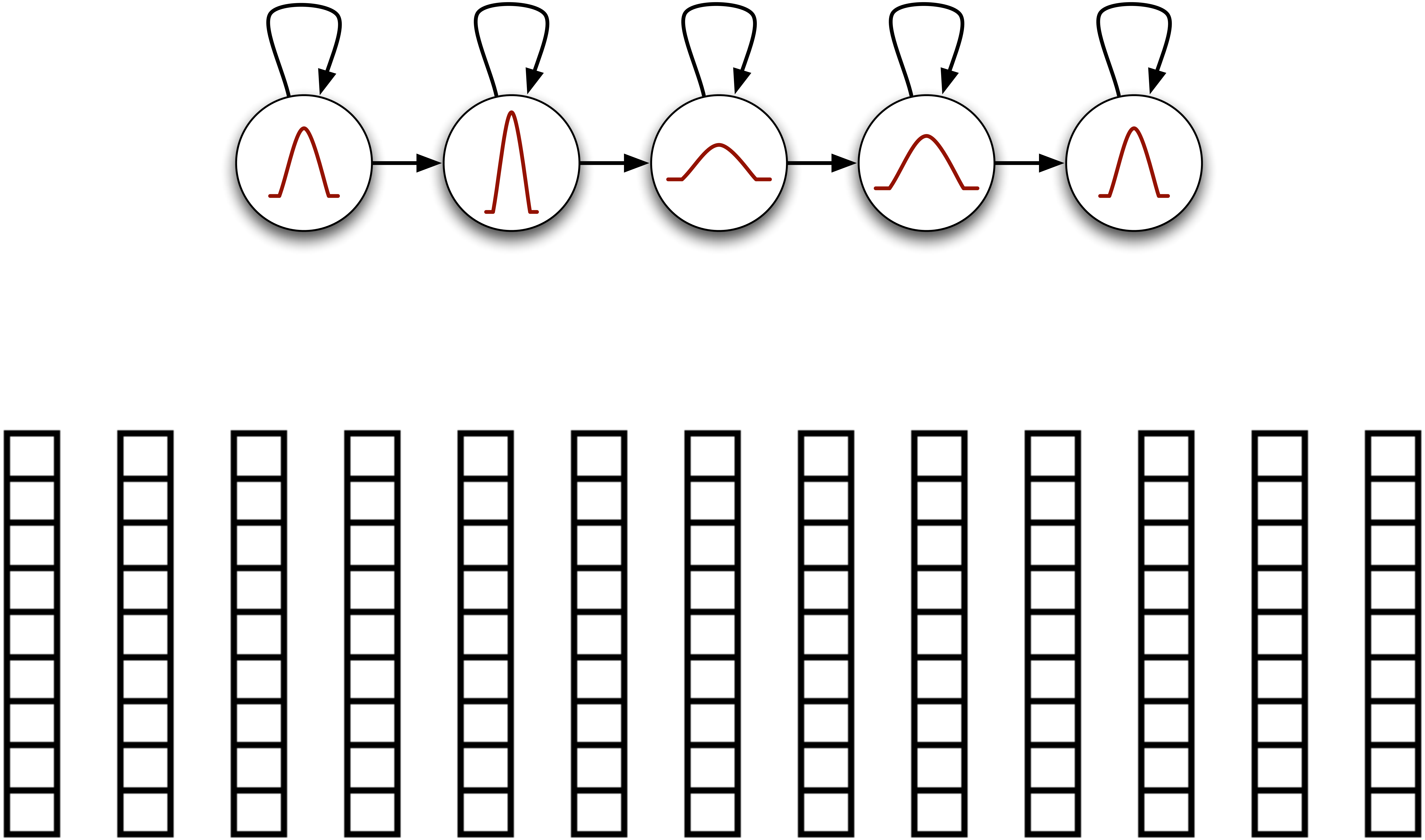

GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL







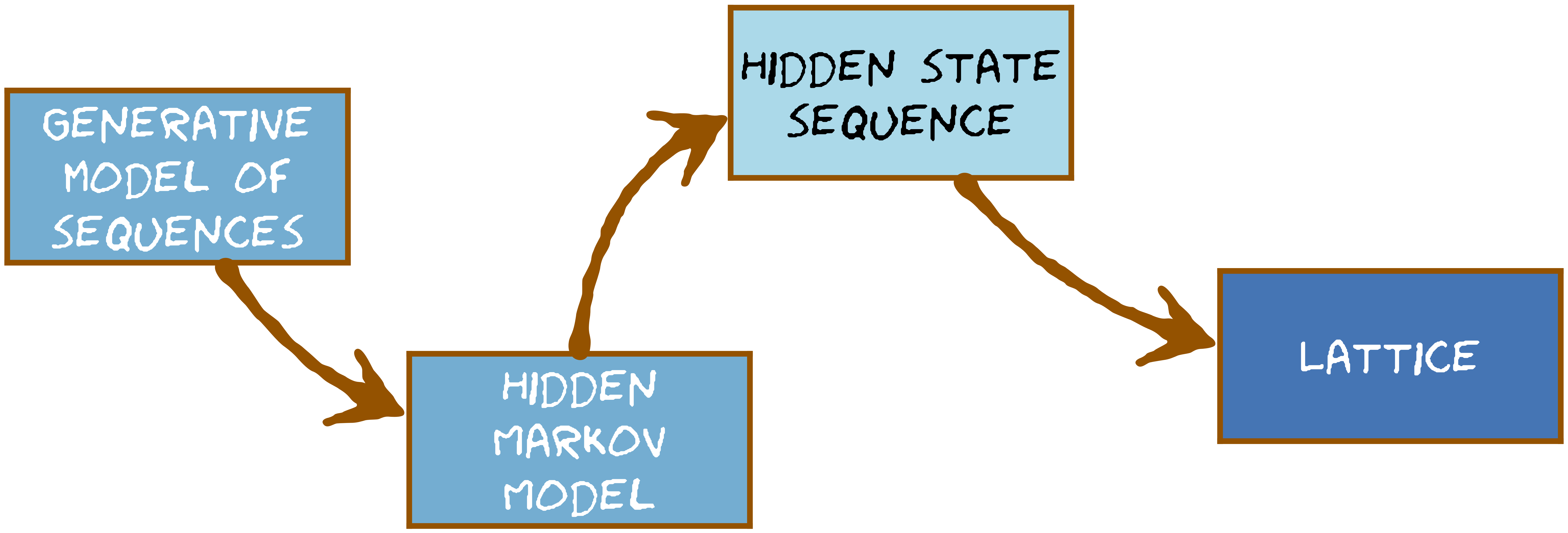

GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL
HIDDEN STATE SEQUENCE

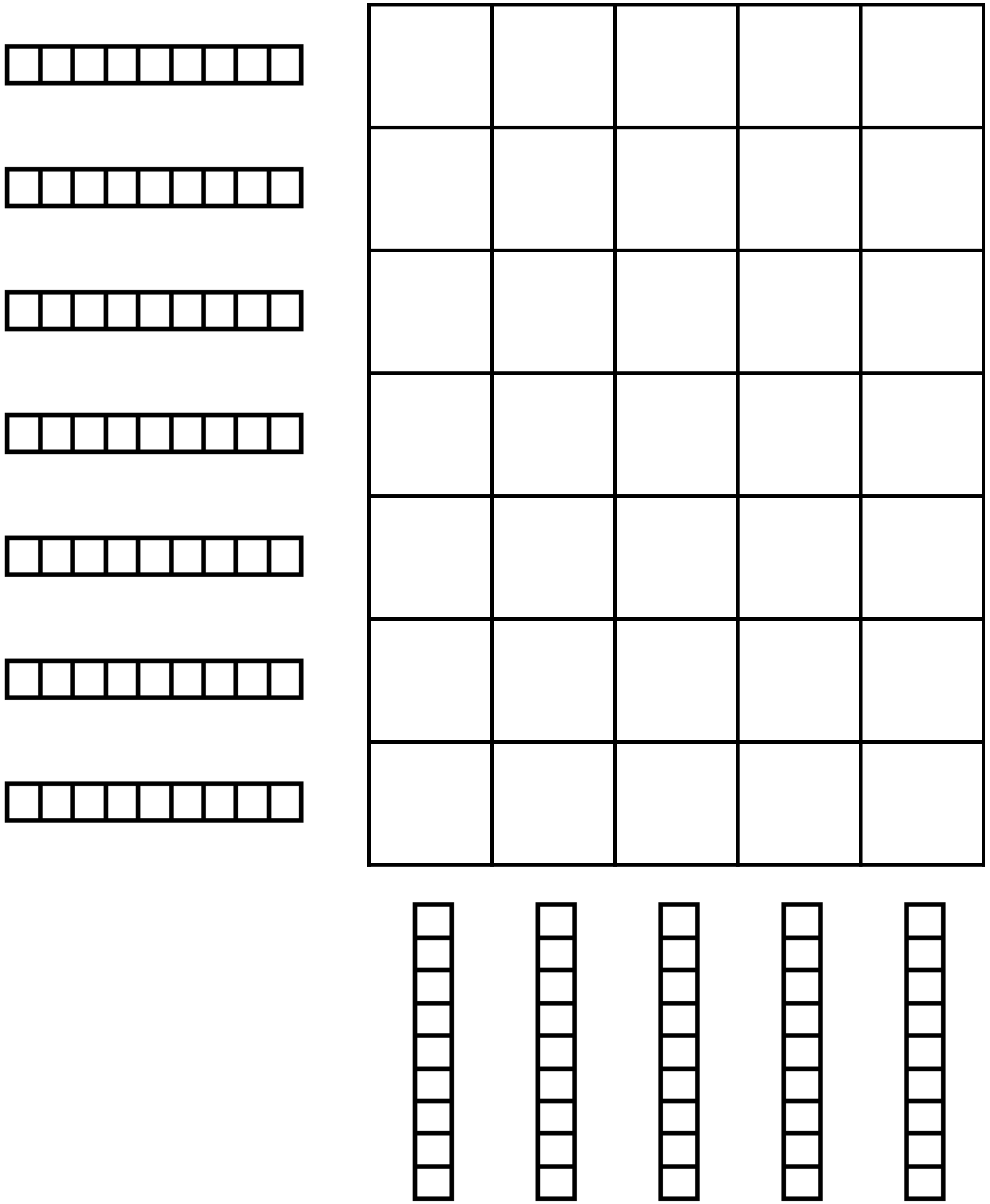


Hidden state sequence

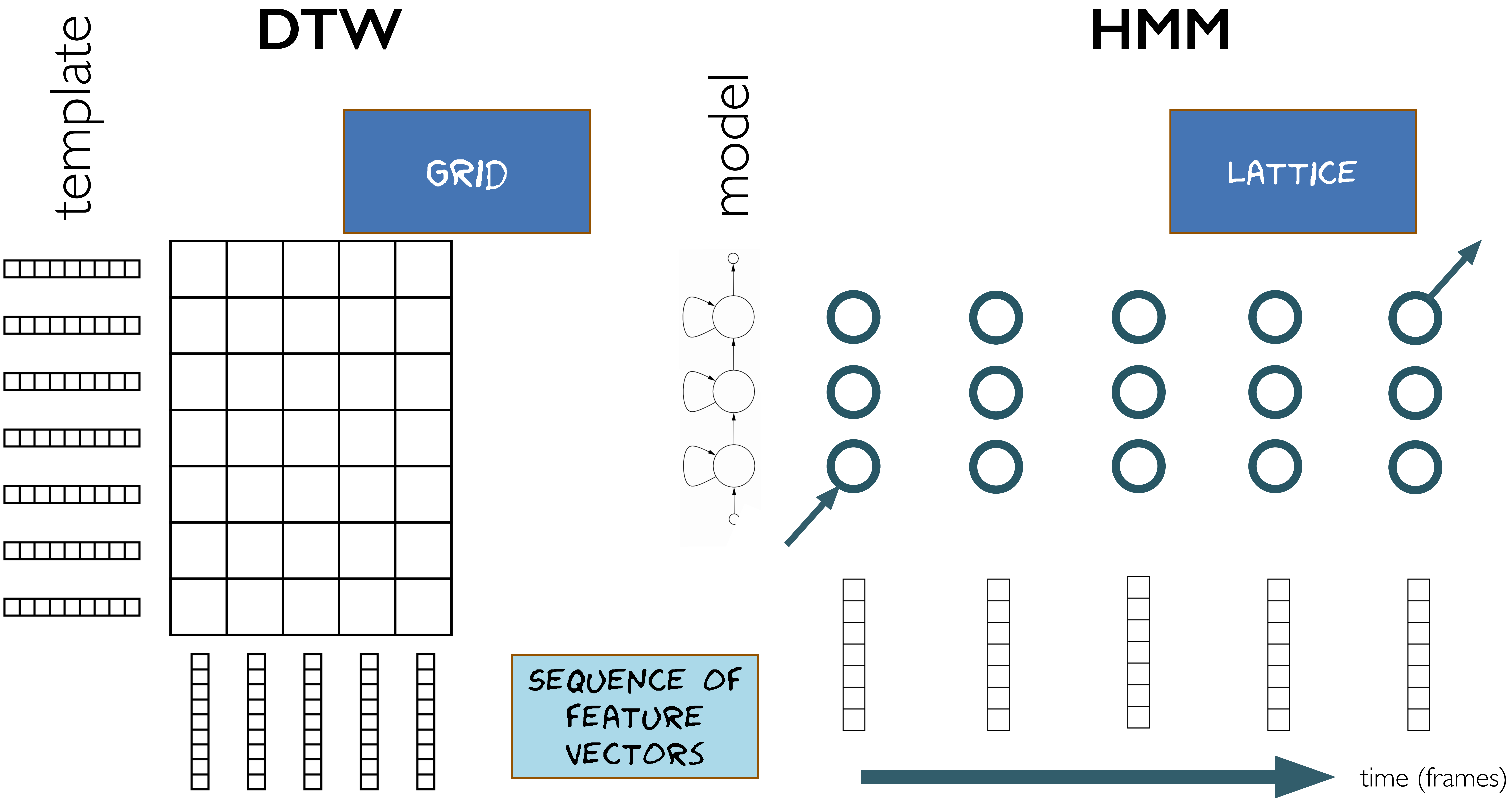
$$P(\mathbf{O} \,|\, \text{model})$$

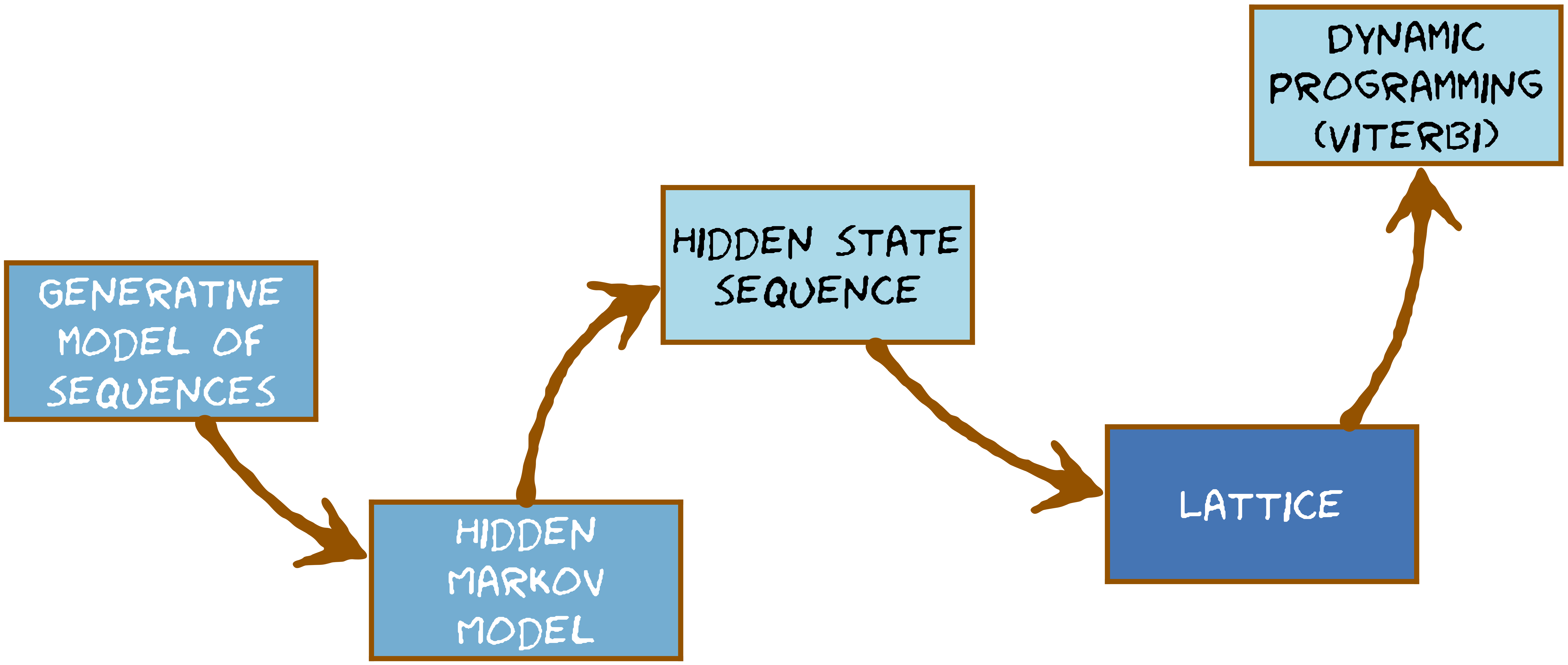
$$P(\mathbf{O}, Q \,|\, \text{model})$$


GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL
HIDDEN STATE SEQUENCE
LATTICE

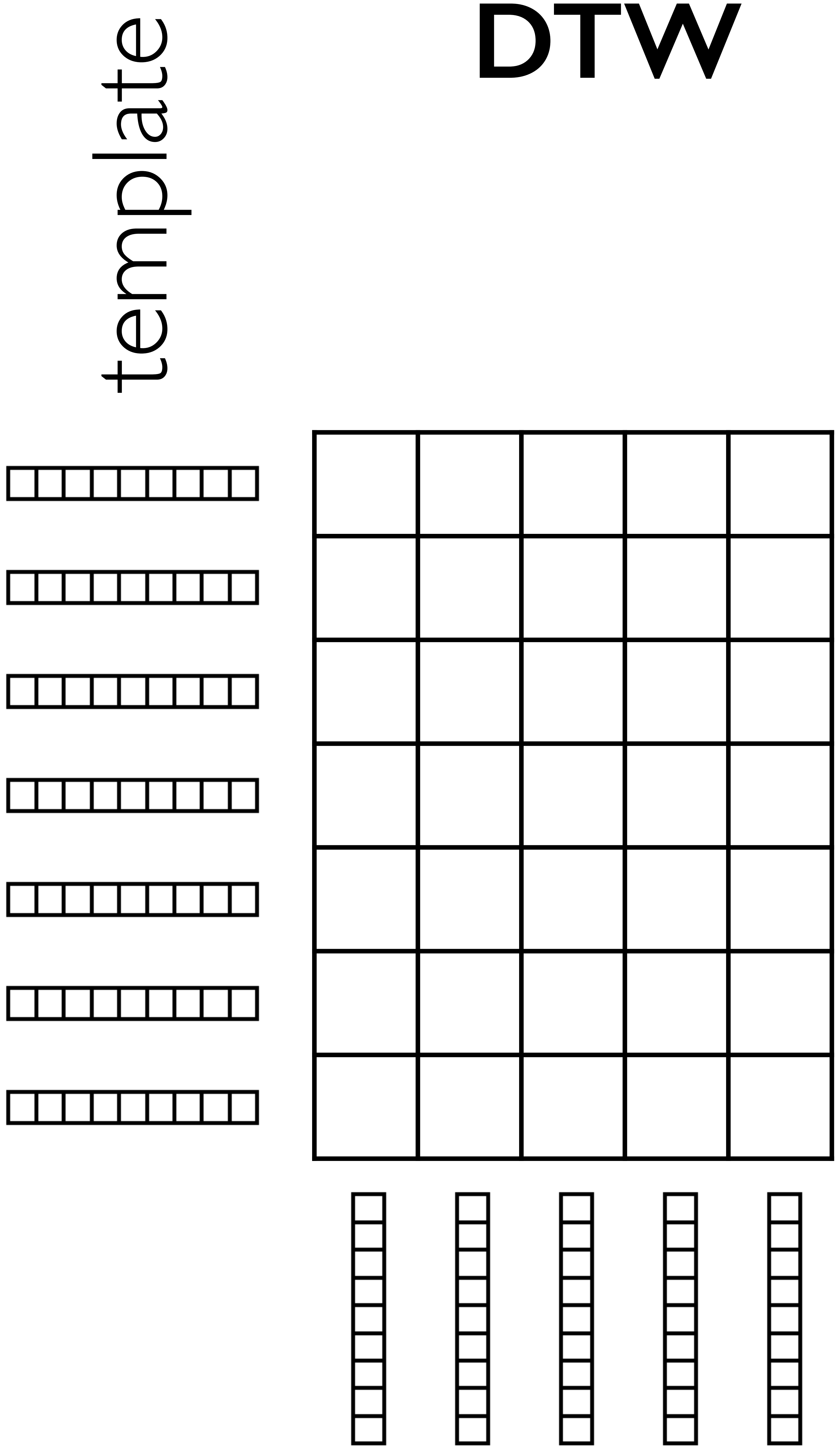


DYNAMIC PROGRAMMING (DTW)

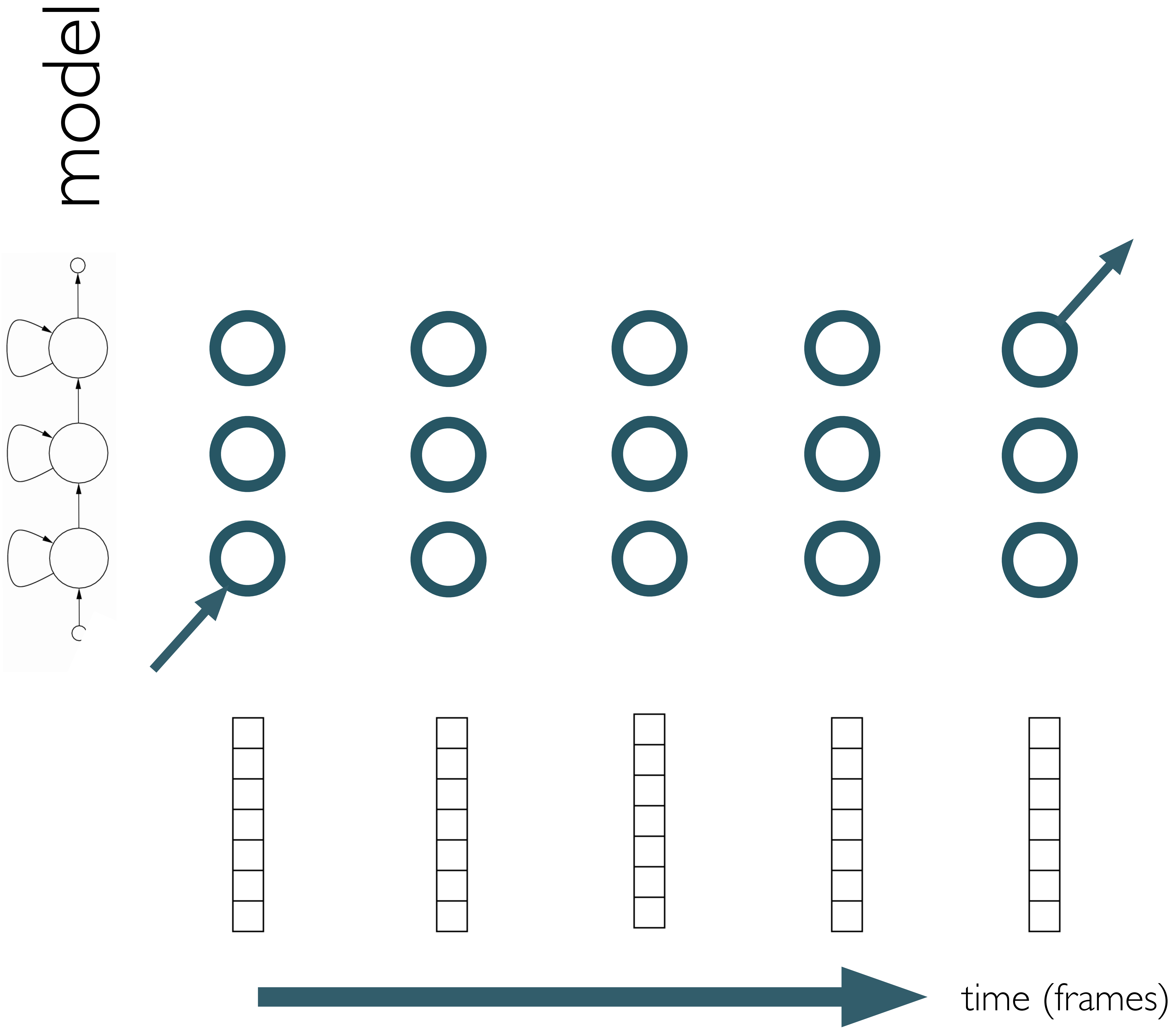

DTW
HMM
template
GRID
model
LATTICE
SEQUENCE OF FEATURE VECTORS
time (frames)


GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL
HIDDEN STATE SEQUENCE
LATTICE
DYNAMIC PROGRAMMING (VITERBI)

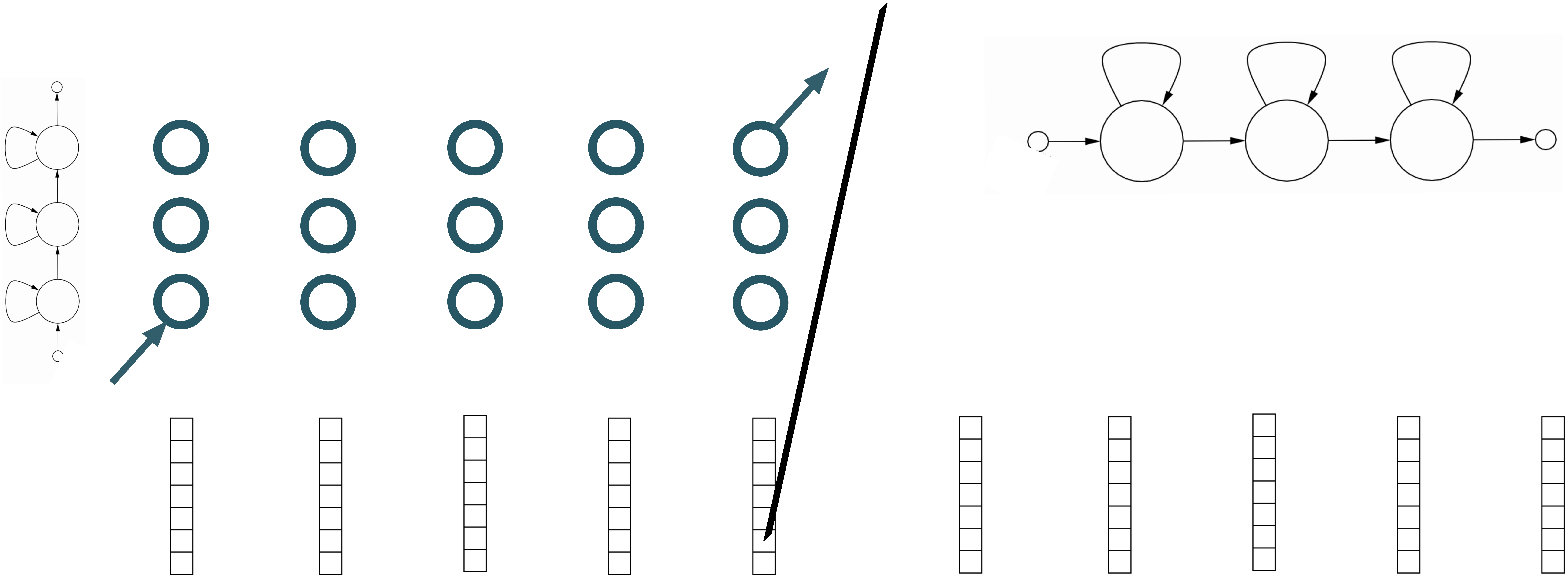
DTW



HMM



Two ways to implement the Viterbi algorithm



What next?

this extension will turn out to be quite easy

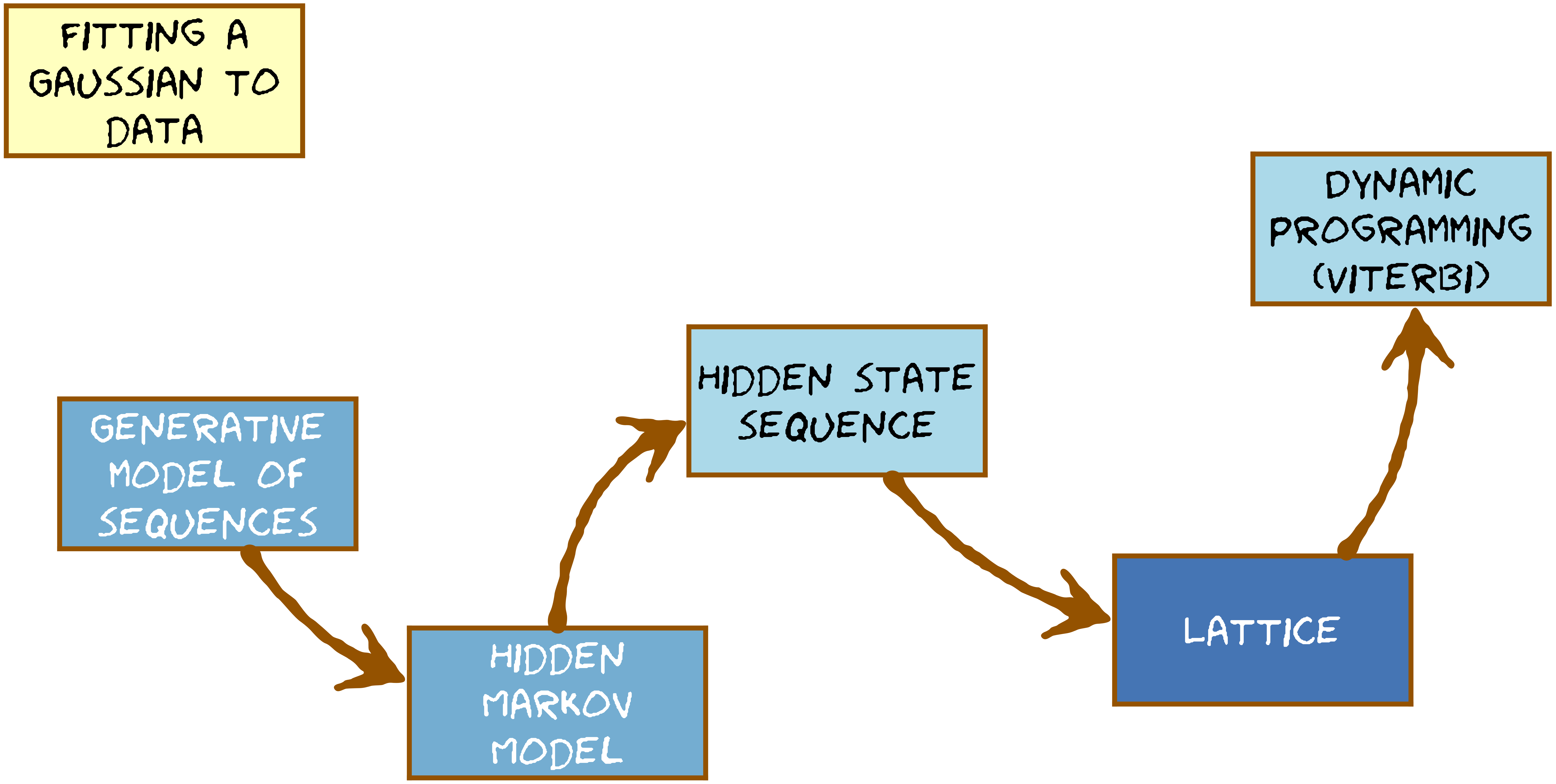
- connected speech

we just need to add a language model

computationally, this is straightforward

- training the model from data

but you may find it conceptually challenging

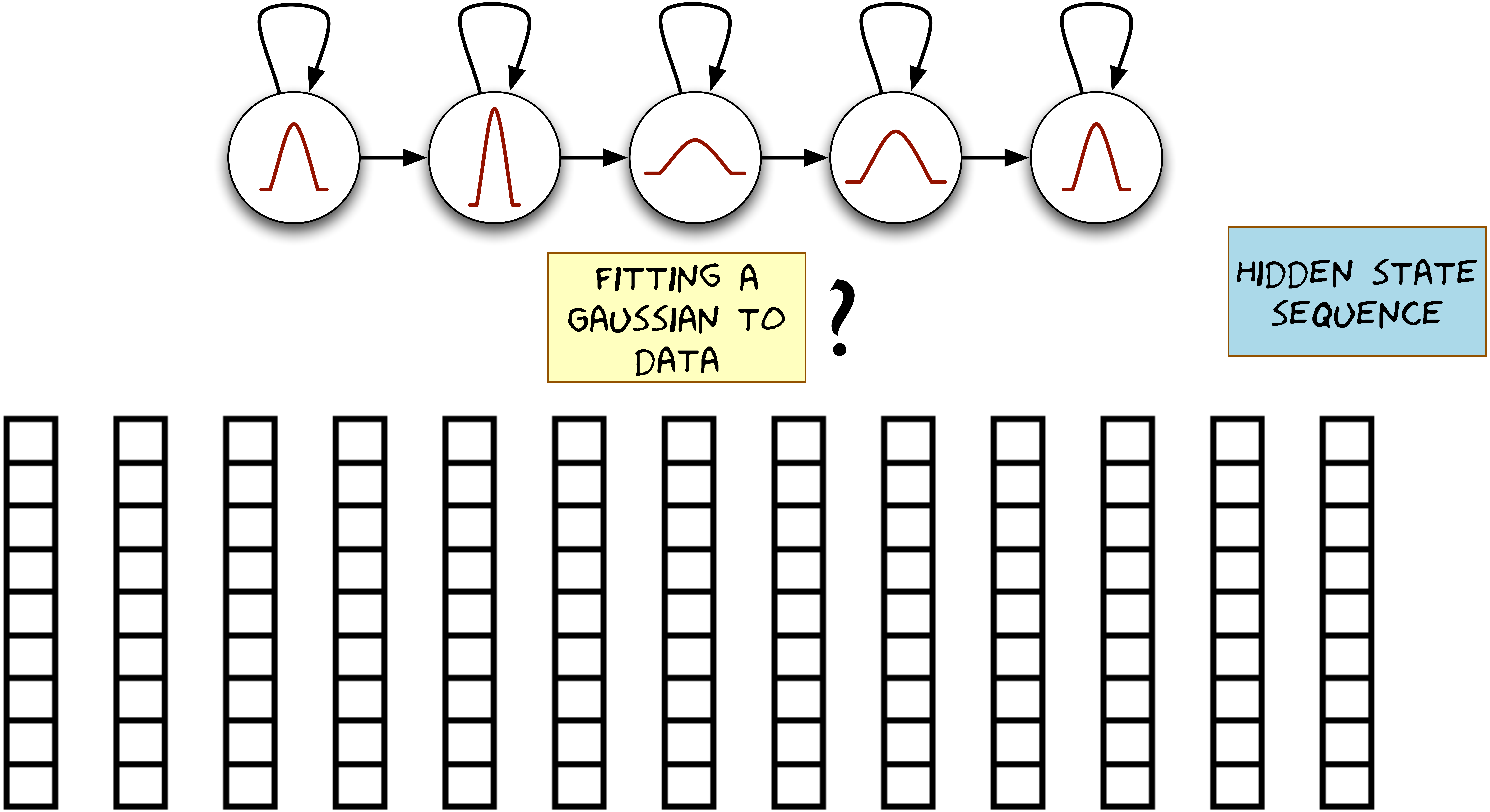

FITTING A GAUSSIAN TO DATA
GENERATIVE MODEL OF SEQUENCES
HIDDEN MARKOV MODEL
HIDDEN STATE SEQUENCE
LATTICE
DYNAMIC PROGRAMMING (VITERBI)

Fitting the Gaussian to data

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \hat{\mu})^2$$


FITTING A GAUSSIAN TO DATA
?
HIDDEN STATE SEQUENCE