TOWARD AUTISM RECOGNITION USING HIDDEN MARKOV MODELS

by

JOSEPH PAUL LANCASTER JR.

B.S., Kansas State University, 2006

A THESIS

submitted in partial fulfillment of the
requirements for the degree

Master of Science

Department of Computing and Information Sciences
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2008

Approved by:

Major Professor
Dr. David A. Gustafson

Abstract

The use of hidden Markov models in autism recognition and analysis is investigated. More specifically, we would like to be able to determine a person's level of autism (AS, HFA, MFA, LFA) using hidden Markov models trained on observations taken from a subject's behavior in an experiment. A preliminary model is described that includes the three mental states *self-absorbed*, *attentive*, and *join-attentive*. Futhermore, observations are included that are more or less indicative of each of these states. Two experiments are described, the first on a single subject and the second on two subjects. Data was collected from one individual in the second experiment and observations were prepared for input to hidden Markov models and the resulting hidden Markov models were studied. Several questions subsequently arose and tests, written in Java using the JaHMM hidden Markov model toolkit, were conducted to learn more about the hidden Markov models being used as autism recognizers and the training algorithms being used to train them. The tests are described along with the corresponding results and implications. Finally, suggestions are made for future work. It turns out that we aren't yet able to produce hidden Markov models that are indicative of a persons level of autism and the problems encountered are discussed and the suggested future work is intended to further investigate the use of hidden Markov models in autism recognition.

Table of Contents

List of Figures

List of Tables

Chapter 1

Background

Professor David Gustafson, Professor James Teagarden, Professor Marilyn Kaff, Professor Bronwyn Fees, and I have been investigating the use of hidden Markov models for autism recognition and analysis. My literature review indicates that applications of hidden Markov models outside of engineering and biology aren't common. I nevertheless, along with Professor Gustafson, produce a preliminary model and run two experiments to gather data which is used to train hidden Markov models and test their applicability to autism recognition and analysis. It turns out that the hidden Markov models are trained in a way that doesn't allow us to attach meaning to states prior to training. Furthermore, difficulties were encountered when attempting to understand and compare trained hidden Markov models and their parameters. One reason for this is that during training hidden Markov models are only guaranteed to converge to local optima, and the resulting hidden Markov models, although often quite similar, are difficult to measure as being similar by any measure we've examined so far. These problems will be discussed and directions for future work will be proposed. In what follows, I share my progress toward autism recognition using hidden Markov models. I begin with a literature review of autism in Chapter 2 and hidden Markov models in Chapter 3. Then, I elaborate on our preliminary model in Chapter 4, discuss stipulations concerning hidden Markov models and corresponding tests in Chapter 5, discuss the results of those tests in Chapter 6, discuss the implications of the results in Chapter 7, and finally propose future work in Chapter 8.

Chapter 2

Autism Literature Review

Asperger's syndrom (AS), high functioning autism (HFA), medium functioning autism (MFA), and low functioning autism (LFA), collectively known as autism, are the autism spectrum conditions (ASC). They are characterized by social, affective, and communication deficits and in the "lower" forms, also by repetitive and obsessive behavior. Persons with AS tend to have a normal to above average IQ and perform well in mathematics, engineering, and computer science. On the other hand, persons with HFA, MFA, LFA have a progressively lower IQ, more social and communication diffulties, and a much higher degree of repetitive and obsessive behavior. Autism is generally shows up very early on in the infant years. Persons afflicted with ASC disorders may have difficulty developing language skills, relating to others, and my engage in what non-autistic persons might consider unusual behaviors[BC06].

Several, sometimes conflicting autism theories have been proposed. Among them is the Domain Specificity Theory that claims that people have at least four core domains of cognition. They are folk biology, folk physics, folk psychology, and folk mathematics. Folk biology concerns a persons ability to taxonomize objects in the natural world, folk physics concerns the ability of a person to understand the causal and mechanical properties of objects, folk psychology pertains to a persons ability to understand and predict the mental states and intentions of others, and folk mathematics accounts for the ability of persons to count and reason about the probability of events[BCWS+97]. According to Baron-Cohen, autistic persons are lacking in folk psychology and unusually strong in folk physics.

Hyper-systemizing is another theory that suggests that all persons have a tendency to systemize their environment. That is, they need to find order and predictability in their environment. To do so they must deal with change. According to the theory there are two types of change, agentive change and non-agentive change. Agentive change refers to objects in the environment that appear to be self-propelled, and non-agentive refers to objects that do not appear to be self-propelled. Non-agentive change is simpler to reason about since the objects and their corresponding change is highly predictive (sand falling through fingers or a fan blade in motion, for example). Agentive change on the other hand is more difficult to make predictions about. A person's behavior or the behavior of a computer are examples. According to the theory, systemizing involves the identification of laws to predict the probability of an event. Furthermore, there are eight levels of systemizing and each level requires as input systems with increasingly more structured systems (less random). Finally, autistic individuals are in the top 5-8 levels of this system. Persons with AS are at level 5, persons with HFA are at level 6, persons with MFA are at level 7, and persons with LFA are at level 8. Consequently, persons with the lower forms of autism are able to systemize only the most structured of systems and repetitive and obsessive behavior result as the persons try to keep there environment as predictable as possible. Furthermore, language may suffer since human language tends not to be very structured[BC06].

The theory of mind-blindness claims the inability of autistic persons to attribute mental states to themselves or to others[BC90,Fri01]. In other words, they lack a "theory of mind"[FH94]. Central coherence theory maintains that there is an information integration imbalance at the root of autism[FH94], and executive dysfunction theory accounts for the inability of an individual to switch attention flexibly and in a goal-directed way, rather than in a reactive way. Baron-Cohen later introduces the "extreme male brain" theory of autism[BC02] which claims that females tend to empathize more and males tend to systemize more, where systemize has the same meaning as that given above and empathize is what it sounds like. Systemizers try to determine the variables of systems in their environment and determine the laws that govern the systems, where a system is anything that takes inputs, performs some function, and then produces outputs. Sympathizers try to understand the thoughts, emotions, and intentions of others and to predict their behavior. So, according to the theory, autism is an extreme exaggeration of the male brain. This fits the hyper-systemizing theory above.

Then, there are those who claim that autistic individuals have difficulty in recognizing faces, body language, emotion, vocal expressions, and so forth[BC90], and there are those that disagree, claiming that flaws in the conducted experiments are to blame[Cas05]. Further, autistic children have been shown to have difficulty producing certain facial expressions such as smiles, especially facial expressions of positive affect[YKSM89].

Chapter 3

Hidden Markov Model Literature Review

3.1 Hidden Markov Models

Hidden Markov models are models that are similar to Markov models but in which the states are hidden. Hidden Markov models also emit an observation symbol at every time t. While the states of a hidden Markov model are hidden, the observation symbol emissions are not. These output symbols are a function of the hidden underlying state transitions. Defined in this way, hidden Markov models are capable of modeling an unobservable random variable and an observable random variable that is a function of the hidden random variable. Put another way, hidden Markov models generate observations according to the probability distribution functions that are defined for each state in the hidden Markov model. A generated output symbol then depends on the current state and associated probability distribution function. So the probability of a sequence of observations being generated depends on the probability of the underlying state transitions and the observation probability distribution funtions of the states in the transition graph.

In our case, the unobservable random variable is the mental state of a subject and observable random variable is the behavior of the subject; the subject's behavior being a function of the subject's mental state. While we may not be able to talk about the state that a hidden Markov model is in at time t, we can talk about the most probable state sequence of the model. The emitted observation symbols provide clues about the underlying states of the hidden Markov model as the behavior of a subject provides clues about underlying mental state of the subject. So, while it may be difficult to ascertain the mental state of a subject directly due to certain practical limitations such as limitations on our ability to read minds, it is possible that a hidden Markov model can be trained to generate the same observation symbols as a subject, and if so, its transition probability matrix may provide clues about the mental states of the subject such as how long a subject is in a particular state and how the subject's mental states transition from one to the next.

3.1.1 The Hidden Markov Model Definition

A hidden Markov model $\lambda = \langle \pi, A, B \rangle$ with N states and M observation output symbols has an initial state probability vector $\pi = \{\pi_i\}$, where $\pi_i > 0$, $0 \leq i < N$, and $\sum_{i=0}^{N-1} \pi_i = 1$, that describes the probability of being in a state i at time $t = 1$; a transition probability matrix $A = \{a_{ij}\}$, where $a_{ij} > 0$, $0 \leq i, j < N$, and $\sum_{j=0}^{N-1} a_{ij} = 1$, $0 \leq i < N$, that describes the probability of being in state i at time t and state j at time $t+1$; and an observation probability vector $B = \{b_j(k)\}$, where $b_j(k) > 0$, $0 \leq k < M$, and $\sum_{k=0}^{M-1} b_j(k) = 1$, $0 \leq j < N$, describing the probability of the observation symbol k being output in state j.

Furthermore, there are three assumptions concerning hidden Markov models. The first assumption is the Markov assumption, specifically that the current (next) state depends only on the previous (current) state:

$$a_{ij} = p\{\, q_{t+1} = j \,|\, q_t = i \,\} \,. \tag{3.1}$$

Secondly, time t is not dependent on the start time t_0:

$$p\{\, q_{t_1+1} = j \,|\, q_{t_1} = i \,\} = p\{\, q_{t_2+1} = j \,|\, q_{t_2} = i \,\} \,. \tag{3.2}$$

Third and finally, the output o_t at time t depends on the the state q_t at time t and not on any other observations, including any of $o_i, 0 \leq i < t$. If $O = o_1, o_2, \ldots, o_T$ is an observation sequence of length T, then

$$\{\, O \,|\, q_1, q_2, \ldots, q_T, \lambda \,\} = \prod_{t=1}^{T} p(\, o_t \,|\, q_t, \lambda \,) \,. \tag{3.3}$$

Please see Warakagoda's website for a for complete and less cursory overview of hidden Markov models[War96].

3.2 Hidden Markov Model Applications in the Literature

Hidden Markov models have been applied to automatic speech recognition[Rab89], optical character recognition[AsK93], computational biology[KBM+94], signal processing[CNB98], and estimation and control[EMA97].

Chapter 4

Preliminary Modeling and Initial Experiments

Modeling the autistic mind using hidden Markov models could provide autism researchers with a powerful pattern recognition, modeling, and analysis tool. While the application of many other computational learning models could be explored, hidden Markov models offer an attractive temporal component, suggesting that we might understand the behavior of a subject accross time as they move from behavior to behavior and from state to state. As opposed to only considering a limited number of features of an individual for the purposes of classification, for example, although doing might also be useful. As an example, consider how differences in transition probabilities may reflect difference in a subject's level of autism. Hidden Markov models were chosen for their ability to pick up on unobservable temporal patterns given observable data. We discuss our model and our control experiments in this chapter.

4.1 The Model

Three states were chosen to represent the mental states of persons in a way that could possibly allow a hidden Markov model to pick up on behaviors that may distinguish autistic persons from non-autistic persons. These three states were *self-absorbed*, *attentive*, and *joint-attentive*.

When in the self-absorbed state, a person is unattentive with respect to the environment and other persons in it. You might say that the person is occupied with his or her own thoughts and behaviors. An attentive person may be paying attention to or interacting with the surrounding environment or paying some mild attention to persons or other actors in it, but they are not joint-attentive. A person is join-attentive when he or she is engaged with another person with whom attention on objects or other actors might be shared. These states were chosen because it seems that any person (who is not sleeping) must be in one and only one of them at any given time. Furthermore, autistic individuals tend not to be join-attentive[YKSM89]. Even more, they tend to be self-absorbed. Depending on an autistic person's level of autism, when attentive, they tend to repetitive and obsessive behaviors[BC06]. Lining up objects is an example of an obsessive behavior. Coupled with the right observations, these states could be very informative if the observations give us insight into the states the persons are in.

Observations were chosen that could be recorded during experiments. Further, observations were chosen that could be used in experiments that tested an individuals ability to identify and mirror emotions associated with images of faces or an ability to respond to humor. Such tests are potentially informative since autistic persons are thought by some to have difficulty in both interpreting and expressing culturally shared facial expressions, body language, and tone of voice[YKSM89,Cas05,BCWS+97,BC02,Fri01]. Even more, observations were chosen that could provide clues about underlying mental state.

Useful observations include the expression or lack of expression of various facial expressions, eye movements, vocal cues, and body language. Those observations chosen when data from the second experiment was used to train hidden Markov models were smiles and laughter and the degree and duration or the absence thereof. The durations were used because a trigger based system was chosen over the use of strict intervals of time from which observations could be taken and so duration was supplimentary to degree. If durations weren't used, such useful information would be lost since it coudn't show up in any other way. Strict time intervals were rejected since it wouldn't be clear how to choose a time granularity appropriate for gathering observations and then choosing an observation when either nothing occurred or more than one observation occurred in a time interval and in what order and so forth.

4.2 The Initial Experiments

The preliminary experiments were arranged to acquire observation data appropriate for hidden Markov model training that allowed me to test a hidden Markov model's ability to represent a subject's event data. It also gave me a chance to train a few hidden Markov models with real data and to try to analyze and compare them. The goal behind these experiments was to develop one or more tests that would provide the data necessary to learn hidden Markov models that effectively represent autistic and non-autistic persons, respectively. Two separate experiments were conducted.

4.2.1 The First Experiment

In the first experiment, Danva2[Dan] was used to present a 12 year-old male subject with a sequence of images of people in the same age group making happy, sad, angry, and fearful faces of differing intensities. The subject was asked to perform an emotion recognition task (on the entire sequence of images) in three separate phases. In the first phase, the subject was asked to view each image and select one of four buttons labeled "happy", "sad", "angry", and "fearful". In the second phase, the subject was ask to say aloud the selected emotion in addition to selecting a button. The subject might say "He is fearful" or "She is sad", for example. In the third phase, in addition to pressing the appropriate button and saying the selected emotion aloud, the subject was asked to voluntarily mirror the facial expression of the person in each image.

Using a recording of the experiment, observations were taken according to the following criteria:

- Good Labeling Scores
- Involuntary Mirroring
- Involuntary Vocal Responses
- Voluntary Mirroring

If the subject was able to achieve good labeling scores, then he was able to correctly recognize emotion in the faces he had seen, which, modulo any possible memorization of the answers, does not favor an autistic label[YKSM89,Cas05,BCWS+97,BC02,Fri01]. Next, while we would presume that a non-autistic subject would be better able to voluntarily mirror the faces they see in the Danva2 images, in either case, voluntary mirroring was chosen since it might provide discretionary information, if only based upon degree. Finally, involuntary facial and vocal mirroring could indicate that the individual is sharing emotion with the person in an image, as understanding the other persons emotion may involve experiencing it to some degree[YKSM89,Cas05,BCWS+97,BCWS+97,GG98,Sch00,RC04,WWSP01]. So, for example, we might expect a person's voice to indicate a bit of sadness when refering to a sad face. Perhaps their facial expressions would also provide indications of mirroring, if even very mild. Further, non-autistic individuals may be more likely to involuntarily mirror another's emotional state. So again, this may provide our learned model with discretionary information. It may also be interesting to have a subject view images of faces of persons not in their age group, of their mothers, or in other contexts such as in a classroom lecture video, where emotional faces of the teacher might more often be perceived as negative, for example.

Experimental setup included a monitor on a table in front of a seat for the subject. A monitor was also placed behind the subject and the output was the same as that on the monitor on the table before the subject. A clock was placed behind the monitor and faced a camera on the table. The camera faced the clock, the subject, and the monitor behind the subject. The clock and second monitor allowed us to synchronize with the subject as he was tested. This made it possible to know what images each response corresponded to and also provided information about the duration of each response.

No data from the Danva2 experiment was used to train a hidden Markov model. The subject seemed too aware that he was being actively observed and appeared act at least somewhat self-consciously and with reservation. There was concern that this might cause the individual not to involuntarily mirror recognized emotions and perhaps even effect voluntary mirroring if the subject was afraid or embarrased.

4.2.2 The Second Experiment

The second experiment was set up because the first experiment seemed to put the subject in a sufficiently unnatural environment that might prevent the natural behavior necessary to produce the expected behaviors. In this experiment, two subjects, each on a separate occasion, were asked to watch a sequence of six humorous video clips. In between each clip the subjects were asked to say the phrase “I have finished watching this video clip and I am ready to watch the next clip”. This was done in an effort to provide a clean emotional slate before the following clip. As described above, subjects were observed for smiling and laughing. The duration and instensity of the smile or laugh was also recorded. This experiment also included a clock and two monitors. This experiment, on the other hand, did not place a camera on the table in front of the subject, but rather, all video footage was taken from a relatively inconspicuous camera in a globe, on the wall, in front of the subject.

During the experiments, the subjects' facial expressions, vocal expressions, eye movement, and body language along with any evidence of involuntary mirroring activity were all captured on video, similar to the first experiment, except that the subjects in this experiment weren't asked to voluntarily mirror. There is difficulty making out eye movement in much of these videos, however, so future experiments should remedy this somehow. As far as natural bahavior goes, one subject, who was also the subject in the first experiment, did seem to act more comfortably in the second experiment. The other failed to produce many emotional cues of any kind and hence, very little useful data could be collected from the subjects behavior.

4.3 Using Hidden Markov Models with the Experimental Data

After collecting data from the second experiment, an observation sequence was produced based on the behavior of the subject from which we were able to elicit useful responses. Hidden Markov models were trained with this data. However, it was necessary to determine whether or not enough data was being colelct to effectively model the subject's behavior using hidden Markov models. This question was approached empirically since the literature review turned up little on uses of hidden Markov models beyond engineering applications such as automatic speech recognition, automatic character recognition, and signal processing. These tests are described in Chapter 5 and the results and implications are discussed in Chapters 6 and 7 respectively.

Chapter 5

Hidden Markov Model Requirements

Before using hidden Markov models to model the mental states of persons, several stipulations must be made and several questions addressed. These questions concern the ability of hidden Markov models to correctly represent the mental states they are intended to represent and our ability to compare them. Futhermore, information extraction must not be prohibitive. While eventual difficulies prevented us from performing any useful information extraction, hidden Markov model generation and comparison were studied and will be discussed here. In what follows, these questions and any corresponding tests are discussed. I begin with a cursory overview of the hidden Markov model tool-kit (Section 5.1) and associated training algorithms (Section 5.2) used in this study's testing implementations. The results of the majority of these tests are presented in Chapter 6. Chapter 7 discusses the implications of those results.

5.1 The JaHMM Hidden Markov Model Took-Kit

The JaHMM hidden Markov model took-kit was used for all test implementations[Fra06]. JaHMM allows hidden Markov models and observation sequences to be written to and read from specially formatted files. Hidden Markov models can also be instantiated directly in code. Once created, JaHMM hidden Markov models can be manipulated using various methods that JaHMM implements. Supported methods include the segemental k-means and Baum-Welch algorithms for training hidden Markov models, methods for determining the probability of the hidden Markov model generating a given obsevation sequence, and methods for generating the most likely state sequence of a hidden Markov model given a sequence of observations. JaHMM also implements the Kullback-Leibler distance measure for hidden Markov models, which will be mentioned in Section 5.6.

5.2 The Segmental K-Means and Baum-Welch Algorithms

Two hidden Markov model training algorithms were used in the following tests. These two algorithms were chosen primarily because of their availability in JaHMM. They are the segmental k-means algorithm[JR90,DD96] and the Baum-Welch algorithm[JR90,DD96,RJ86,War96]. They both use maximum likelyhood critria. Since this likelyhood maximization cannot be solved analytically both algorithms use iterative gradient-based approaches. Despite this, a convergence proof is given for the segmental k-means algorithm by Juang et. al.[JR90]and the Baum-Welch algorithm is guaranteed to converge according to Warakagoda[War96].

A third learning criterion, the maximum mutual information criterion (MMI) was not used in this study. The criterion, a gradient-based method like the segmental k-means and Baum-Welch algorithms, was not used because it was not available in JaHMM. It has not, however, been ruled out for future investigations. It has desirable properties not present in either of the algorithms currently being used in this study[War96,OG02]. Most notably, maximum mutual information hidden Markov models are trained, as Warakagoda puts it, more discriminatively[War96,OG02]. Although it is not presently clear whether or not this will be beneficial to this study.

5.3 Selecting a Training Algorithm

It has not been determined which of the two available algorithms is the most appropriate for the learning task involved or whether or not each has its own niche. What is known is that each algorithm behaves significantly differently than the other in several ways. This was discovered while running the tests that are described in this chapter. The exact differences between these two algorithms are discussed in Section 6.1. As a consequence, both algorithms were used in many of the tests run and their output was compared.

5.4 Understanding Hidden Markov Model Convergence

Over the course of the study several questions arose concerning the convergence of the training algorithms and resulting hidden Markov models with respect to several factors such as the length and the number of training sequences supplied. What's more, the sensitivity of the available training algorithms to small permutations in the data wasn't understood. It is important that these questions be answered before there can be confidence that enough data has been collected and that the learned models have converged. Without confidence in the convergence of the models it is hard to say whether any information that can be gleaned from the model parameters will be useful. Further, without confidence that enough data has been collected it is difficult to determine whether or not it is feasible to collect enough data for the hidden Markov model from human subjects in the first place let alone that enough has actually been collected.

In order to investigate the convergence of the hidden Markov models trained using the available training algorithms a test program was implemented and several program parameters were varied to achieve results for different tests. Each test is designed to compare two or more hidden Markov models that were generated and/or trained under differing cirumstances.

The convergence questions and corresponding tests are discussed in the remainder of this section following an overview of the test program parameters. The actual results are discussed in Section 6.2.

5.4.1 The Convergence Test Program Parameters

In words, the test program proceeds by first generating `NUM_HMMS` hidden Markov models and initializing the training observation sequence to the `empty` sequence. Then, for `ITERATIONS` iterations, the program proceeds by:

1. appending (or prepending if `REVERSE` is `true`) `STEP` (or `max{STEP,MIN_LENGTH}` in the first iteration) observations to the existing training observation sequence,
2. training all `NUM_HMMS` hidden Markov models using the training algorithm specified by `TRAINER` on the newly lengthened sequence,
3. running the comparison algorithm specified by `SIMILARITY` on the set of hidden Markov models specified by `COMPARE`, and finally
4. generating any corresponding data output.

The values that integers appended or prepended to the training sequence can take on are specified by `NUM_OBS`. How the `STEP` successive observations are generated in each iteration is discussed later in this section. The `NUM_HMMS` initial hidden Markov models generated before iteration begins are copied over at each iteration and training is performed on the `NUM_HMMS` copies.

The test program makes use of several important parameters and only the most important will be discussed here. See Table 5.1 for a list of the majority of the available parameters.

Generating Observations

In each iteration, a new training sequence is generated by generating `STEP` observations and appending or prepending those to the existing (previous) training sequence. This is done as follows: If `RANDOM` is `true` then generate a new observation randomly and append or prepend it to the training sequence. Otherwise, let `this_obs` and `counter` be two integer variables initialized to `INIT_OBS` and `0` respectively before the first iteration. Then in each iteration, append or prepend a newly generated observation `STEP` times according to the algorithm given in Figure 5.4.1. If, in the first iteration, the length of the resulting training

Table 5.1: *A list of parameters for the test program in Section 5.4*

`COMPARE`	Takes on the values `CURRENT`, `PREVIOUS`, or `AGGREGATE`.
`RANDOM`	When `true`, successive observations will generated randomly.
`REPEAT`	When `RANDOM` is `false`, determines the number of times an observation will be generated before its successor is generated.
`REVERSE`	When `true`, each new observation will be prepended to the observation sequence. When `false`, each new observation will be appended.
`NUM_HMMS`	The number of hidden Markov models trained in each iteration.
`NUM_OBS`	Representation observations can take on integer values from `0` to `NUM_OBS - 1`. If `RANDOM` is `false`, observations within this range will be generated in order except to begin again at `0` when `NUM_OBS - 1` is reached. Before the successor is generated, however, an observation will be generated `REPEAT` times.
`SIMILARITY`	The similarity metric used.
`ITERATIONS`	The number of times the training sequence will be lengthened, the hidden Markov models will be trained on the new sequence, the similarity metric will be applied to the newly trained hidden Markov models, and any associated data output will be generated.
`STEP`	The number of observations that will be prepended or appended to the training sequence at the beginning of each generation.
`MIN_LENGTH`	The length of the initial observation sequence that is used in the first iteration.
`TRAINER`	The training algorithm used.
`INIT_OBS`	The first observation to be generated and added to the training sequence.

```
if counter < REPEAT then
   counter := counter + 1
else
   if this_obs < NUM_OBS then
      this_obs := this_obs + 1
   else
      this_obs := 0
   end ;
   counter := 0
end ;
if REVERSE then
   prepend_to_sequence ( this_obs )
else
   append_to_sequence ( this_obs )
end
```

Figure 5.1: *Pseudo-code generating a new observation.*

sequence is less than `MIN_LENGTH` then `MIN_LENGTH - STEP` additional observations will be appended or prepended to the sequence before continuing.

If `REVERSE` is `true` then prepend each newly generated observation to the beginning of the previous sequence, othewise append it to the end.

The `COMPARE` Parameter

In each generation a set of hidden Markov models are compared using the `SIMILARITY` similarity measure and output. The `COMPARE` parameter specifies which hidden Markov models are compared to which other hidden Markov models. If `COMPARE = PREVIOUS` then the hidden Markov models in iteration I are compared with the hidden Markov models in iteration $I-1$. If `COMPARE = CURRENT` then all hidden Markov models in iteration I will be compared to eachother. Otherwise, `COMPARE = AGGREGATE` and the hidden Markov models from iterations I_{init} through I are compared.

The `SIMILARITY` and `TRAINER` Parameters

Each time the test program is run a similarity metric must be specified in `SIMILARITY`. The similarity metric is supposed to measure the ”distance” between two hidden Markov models. At the end of every iteration, this measure is applied to each pair of hidden Markov models that are meant to be compared and the average of these ”distances” is output. The three distance metrics tested so far in this study are discussed in Section 5.6.

In addition to a similarity metric a training algorithm must be specified in `TRAINER`. The available training algorithms are discussed in Section 5.2.

5.4.2 The Influence of Training Sequence Length

The necessary length of observation sequences should be known before data collection is begun. Toward this end, the effect of training sequence length on hidden Markov model convergence was tested by comparing hidden Markov models trained on successively longer observation sequences. In each trail, ten hidden Markov models were generated randomly and stored. Also, `COMPARE` was set to `CURRENT` so that in each iteration, all of the newly trained hidden Markov models would be compared with each other, but not to any hidden Markov models from previous generations. `REVERSE` was set to `false` so that we would be appending new observations to the sequence since we concerned only with the length of the sequence and not how it was constructed. `INIT_OBS` was set to `0` and `MIN_LENGTH` was set to `2`. Finally, `NUM_OBS` was set to `3`. The remainder of the parameters were allowed to vary. Varying the these parameters allowed us to test not only the effect of training sequence length on convergence, but also the effects of initial hidden Markov model parameters as the length of an observation observation increases.

In Section 6.2 we discuss the results of running the test program with a `REPEAT` value of `1`, `2`, and `3` when using different combinations of values of `SIMILARITY`and `TRAINER`. `RANDOM` was also allowed to be `true` or `false`.

5.4.3 The Influence of the Number of Training Sequences

The influence of the number of training sequences on convergence wasn't tested. All tests thus far have concentrated on using only one training sequence. This is something that may be investigated in future work. Juang et. al. remark that the segemental k-means algorithm works well with multiple independent observation sequences[JR90] and so the segemental k-means algorithm might be a good place to start.

5.4.4 The Sensitivity of HMMs to Subtle Changes to Input Sequences

Another step taken to ensure the quality of acquired data was to determine how sensitive hidden Markov models are to minor changes in the observation sequences they are trained on. In words, will a subtle change to a training sequence result in significantly different model parameters? Specifically, we were interested in the influence of permuting an observation in a sequence, appending observations to a sequence, and prepending observations to a sequence.

In order to test this influence the test program was run with `COMPARE` set to `PREVIOUS`, `STEP` set to `1`, `NUM_HMMS` set to texttt1, `REPEAT` took on values from `1` to `3`, and finally, `RANDOM` and `REVERSE` were each allowed to be `true` or `false`. This was done for varying values of `SIMILARITY` and `TRAINER` in order to further ascertain the effects of the small obsevation sequence changes on the similarity metrics and training algorithms involved. Using a step size of one with one hidden Markov model per iteration and comparing the hidden Markov model one an iteration with the one in the previous allowed us to determine the sensitivity of a hidden markov model to the addition of a single observation to the beginging or end of a sequence (to add a sequence to the beginning in each iteration `REVERSE` was set to `true`). Permuting two observation values in a training sequence was not done since the addition of an observation the the front or rear of an observation sequence should sufficient. Such a permutation may, however, be considered in future work.

The results of the tests in this section are discussed in Section 6.2.

5.4.5 The Influence of Initial HMM Parameters

It was clear very early on that two hidden Markov models initialized to different random parameters had a good chance of converging to different local optima. However, it wasn't clear whether or not such a pair of hidden Markov models would eventually converge to, at the very least, the same local optima if the observation sequence they were trained on was long enough. To test this, the same tests that were run in Section 5.4.2 were used to observe how each resulting hidden Markov model differed from the others and the length of the training sequence increased. The results of these tests are discussed in Section 6.2.

5.4.6 The Influence of the Training Algorithm

In order to test the influence of the training algorithms on convergence we used the same tests as in Section 5.4.2 and compared the results of using one or the other of the two available algorithms. The results of these tests are discussed in Section 6.2. The number of iterations required by each training algorithm is also discussed.

5.5 Observation Sequence Probability Distribution

To get a sense of how the probabilities of observation sequences are distributed with respect to different hidden Markov model training algorithms, a test program was implemented to collect data on these probabilities. First, all of the possible permutations of an observation sequence of a given length T and a given alphabet size M were generated. Then a random sequence and a randomly generated hidden Markov model were generated and the randomly generated hidden Markov model was trained on the randomly generated sequence. Finally, for each of the possible permutations, the probability of the hidden Markov model generating the permutation is found and stored away. This data along with corresponding statistics was available for analysis. The resulting data is meantioned briefly in Section 6.1 when the differences between the Baum-Welch and the segmental k-means algorithms are discussed.

5.6 Defining an Effective Hidden Markov Model Distance Metric

In order to compare hidden Markov models, particularly two or more each trained on a person that may or may not be autistic, we need to be able to measure or visualize their respective differences. This could be done through the use of one or more metric functions, or through graphing or clustering. While graphing and clustering hidden Markov models representing austic and non-autistic subjects is a major goal of this research, doing so requires us to be able to plot hidden Markov models in a graph and clustering further requires a distance metric. For this reason, research into this problem has so far been concentrated on measuring the distance between two hidden Markov models and determining how one might be plotted.

Three attempts have so far been made to find a distance metric for hidden Markov models. The first was an attempt to use the Kullback-Leibler measure for hidden Markov models, the second attempted to use the expected value of the number of transitions between a state and itself, and the third makes use of the components of the transition probability matrix.

5.6.1 The Kullback-Leibler Measure

The first, the Kullback-Leibler method, is an asymmetric hidden Markov model metric that considers the emission probabilities in addition to the transition probabilities[FRW95]. I used the Kullback-Leibler method implemented in JaHMM. The resulting distances values contained `NaN` values and it's not clear why. Consequently, very little useful data was collected to test the applicability of this measure. The Kullback-Leibler method may be investigated further in future work. It may prove useful for determining the similarity of two hidden Markov models that would be equal given a renaming of one's states since it considers the emission probabilities in addition to the transition probabilities .

5.6.2 Using the Expected Value of the Number of Hops

The second similarity metric, which made use of the expected number of hops between a state and itself, was developed by Professor Gustafson and I. Intuitively this can be used to find the amount of time a hidden Markov model spends in that state since the amount of time spent in a state is the additive inverse of the amount of time spent in all other states. Further, in a step toward the definition of coordinates for hidden Markov model plotting, the additive inverse of the expected value of the number of hops between a state and itself was computed for each state in a hidden Markov model and then used as the coordinate vector components for that hidden Markov model. The corresponing distance metric was simply the cartesian distance between the vectors. Hence, the first definition of the component vector $coord_1(\lambda)$ a for hidden Markov model $\lambda = \{\pi, A, B\}$ with N states is

$$coord_1(\lambda) \ = \ \langle In_\lambda(q_1), \ldots, In_\lambda(q_N) \rangle \tag{5.1}$$

where $In_\lambda(q)$, the time spent in $q \in S$ were S is the set of states in the hidden Markov model λ, is

$$In_\lambda(q) \;=\; E_\lambda(D_q)^{-1} \; . \tag{5.2}$$

$D_{q \in S}$ is a random variable who's value is the length of a sequence of hidden Markov model state transitions starting at q and ending at q, and

$$E_\lambda(D_q) \;=\; \sum_{d \in \delta_q} d \, p_{q,\lambda}(d) \; , \tag{5.3}$$

where d is the length of a path in δ_q, the set of all possible paths from state q to state q, $q \in S$, and $p_{q,\lambda}(d)$ is the probability of such a path.

To construct these paths, find their lengths, and find the corresponding sequence length probabilities, we begin by defining the set of all possible paths of length l beginning at a state σ and ending at σ. This set is shown in Equation 5.4.

$$V_{l \in \mathbb{N}, \sigma \in S} \;=\; \{\langle q_1, \ldots, q_{l'} \rangle \,|\, q_i \in S, \;\; q_1 = q_{l'} = \sigma, l = l'\} \tag{5.4}$$

Now the set of lengths of all possible state sequences beginning at state σ and ending at state σ, δ_σ, is defined as

$$\begin{aligned} \delta_{q \in S} \;&=\; \{\, length_v \,|\, v \in Paths_\sigma \,\} \; , \\ & \quad\; length_v = l - 1 \text{ such that } v = \langle q_1, \ldots, q_l \rangle, q_i \in S \; , \\ & \quad\; Paths_\sigma = \{\, v \in V_{l,\sigma'} \,|\, l \in \mathbb{N}, \sigma = \sigma' \,\} \; . \end{aligned} \tag{5.5}$$

and the probability of a state sequence beginning at σ and ending at σ having length d, $p_{\sigma,\lambda}(d \in \delta_{\sigma \in S})$, is

$$p_{\sigma,\lambda}(d \in \delta_{\sigma \in S}) = \sum_{v \in V_{d,\sigma}} p_\lambda(v) \tag{5.6}$$

where $p_\lambda(v)$ is defined as

$$p_\lambda(\langle q_1, \ldots, q_l \rangle) = \prod_{k=1}^{l-1} a_{q_k, q_{k+1}} ,$$

where $\lambda = \{\pi, \{a_{ij}\}, B\}$.

To compute $E_\lambda(D_q)$ a depth first search is performed, beginning at q and continuing until either q is reached or until one of the probability or length thresholds, `pthresh` and `lthresh` respectively, is reached. As we perform the search we keep track of the probability of reaching the current state in the graph where the probability of initially leaving q is 1 and when we reach a terminal state, either because it is q or because a threshold has been met, we add the probability of the transition sequence that led to that state to an accumulator. When the search is complete, the value of the accumulator is the summation of the probabilities of all of these paths, and is, hence, an approximation to $E_\lambda(D_q)$. `pthresh` is set to a reasonably low probability so that the estimate of $E_\lambda(D_q)$ is good, but also efficiently computable. By setting `pthresh` low enough, the resulting value of $E_\lambda(D_q)$ is barely effected but infinitely

```
private double e ( Hmm < ObservationInteger > hmm , int state )
{

  double result = 0.0 ;

  s . push ( path ( state , 0.0 , 1.0 ) ) ;

  Pair < Integer , Pair < Double , Double > > path = null ;
  int    r = -1 ;
  double l = -1.0 ;
  double p = -1.0 ;

  while ( ! s . empty () ) {
    path = s . pop () ;
    r = state ( path ) ;
    l = length ( path ) ;
    p = probability ( path ) ;
    if ( r == state && l > 0.0 )
      result += l * p ;
    else if ( p >= pthresh && l <= lthresh )
      for ( int q = 0 ; q < hmm . nbStates () ; q++ )
        s . push (
          path ( q , l + 1.0 , p * hmm . getAij ( r , q ) ) ) ;
  }

  return result ;

}
```

Figure 5.2: *Computation of* $E_\lambda(D_q)$.

long paths are not followed. `lthresh` needs to be set to some reasonable length to prevent infinite recursion when probabilities of `1` in the transition probability matrix lead to paths of infinite length with a probability $p > pthresh$ such that `pthresh` becomes ineffective at preventing the exploration of infinitely long transition sequences. The depth-first search algorithm is given in Figure 5.2.

It turns out that these coordinates and corresponding distance metric are not very useful since we are unable to map states in the world to states in the hidden Markov model such that the mapping holds even after training. This is necessary for states in the resulting hidden Markov models to provide information about the similarity of hidden Markov models with respect to states in the world or states of mind in our case. Specifically, through the use of these coordinates we would like to be able to talk about how often a person is in a particular mental state after training a hidden Markov model on sequences of observations of the person. Thus, a distance metric that can provide this kind of information, or one that is able to correctly compare two hidden Markov models that would be equivalent if one had its states renamed must be found.

5.6.3 Using the Components of the Transition Probability Matrix

The third and final similarity metric that was explored also relies on the cartesian distance between coordinate vectors computed for hidden Markov models. This metric was also the work of Professor Gustafson and I. In this case the coordinates are simply taken directly from the transition probability matrix of a hidden Markov model. So, for a hidden Markov model $\lambda = \langle \pi, \{a_{ij}\}, B \rangle$ with N states, the coordinate vector $coord_2(\lambda)$ is

$$\langle\, a_{ij} \,|\, i \neq j \,\rangle \,. \tag{5.7}$$

Only the transitions from a state to a state other than itself were used since for all i, $0 \leq i < N$, a_{ij} such that $i = j$, is just $1 - \sum_{0 \leq j < N, j \neq i} a_{ij}$.

This similarity metric also had the disadvantage that it only considers the transition matrices and does not take the observation probability vectors into consideration. So, again, two very similar hidden Markov models may yeild a very large distance with the metric if a state renaming is required to make them equal. Also, there is the question of whether any strict subset of a hidden Markov model's parameters can be considered in isolation (without the others) since the training algorithms change all of the parameters to maximize the probability of observation sequence provided regardless of their interpretation before training.

Thus, a suitable distance metric (and coordinate definition) for hidden Markov models has yet to be found. There is the question of whether or not one even exists, yet this is a question that has been left for future work. However, such a similarity metic must be found before we will ever be able to compare, plot, or cluster hidden Markov models. More about the unpredictability of hidden Markov model training and the current inability to deal with this since we don't have such a measure will be discussed in Chapter 6.

5.7 Control over Hidden Markov Model Parameters

Before moving on to Chapter 6 to discuss the results of the tests introduced in this chapter, we return for a moment to a topic that was mentioned briefly and implicitly earlier in this chapter but that didn't receive much attention. The topic is our control over hidden Markov model parameters. It's been suggested that in order to make sense of hidden Markov models, we need to be able to make sense of its parameters, whether in sum or in part. But what limitations exist for us? Specifically, can the values of any hidden Markov model parameters be determined before and held constant during training? Can we compare and intuit the parameters of a hidden Markov model meaningfully? Lastly, can we compare and intuit any parameters of a hidden Markov model in isolation?

At this point we haven't answered the question of whether or not we can gather any meaningful information from the parameters, but what we do know is that the parameters will change during training in order to maximize the probability of the training sequence with respect to those parameters so that we can't assign identities to states before training. Furthermore, the probability of a sequence of observations being generated by a hidden Markov model is a function of all hidden Markov model parameters, so that the initial probabilities, transition probabilities, and emission probabilities don't hold much meaning by themselves. This suggests that we need to find ways of analyzing, evaluating, and comparing hidden Markov models using all of the parameters rather than a subset, which further suggests that we need to find coordinates and a similarity metric that do the same. This seems to rule out $coord_1$ and $coord_2$, which makes plenty of room for the Kullback-Leibler measure, which is a focus of future work.

Chapter 6

A Discussion of the Hidden Markov Model Test Results

The tests arranged in Chapter 5 were run and data was collected. The corresponding results are discussed in this chapter. The results are considered in the same order that the associated tests were discussed in Chapter 5. The chapter begins by discussing the training algorithms, moves on to hidden Markov model convergence and hidden Markov model distance metrics, and then finishes with a discussion of our control over hidden Markov model parameters.

6.1 Training Algorithm Difference Discussion

From what the tests show, both training algorithms seem to converge to local optima very raplidly during training. However, the training algorithms are different in that they tend to give significantly different probability values to observation sequences and to distribute those probabilities much differently, which may account for the difference in probability values. In both cases all observation sequences of a particular length sum to 1, but the segmental k-means algorithm attributes very high probability to very few observation sequences, while the Baum-Welch algorithm seems to attribute relatively low probability to almost all possible observation sequences.

It is difficult to say which of the two algorithms is more suitable for autism modeling without a deeper enquiry into how well each algorithm actually represents the subject being modeled. Also, it may be interesting to ask what the observation sequences being being attributed to a particular probability look like. It may be that the Baum-Welch algorithm is better capable of providing probabilities for a wider variety of sequences, since the hidden Markov models trained using the segmental k-means algorithm attribute 0 to all but a very few. However, at the same time, the segmental k-means algorithm may be better at learning hidden Markov models that generate only the very best sequences. For the moment, the Baum-Welch algorithm is being used, but future work may uncover which of the two algorithms is more appropriate and in what situations.

6.2 Convergence Test Results

The final parameters of a hidden Markov model greatly depend on the initial parameters. Even for long sequences, as the tables in Appendix A suggest, the initial parameters seem to be the deciding factor. Table A.19 through Table A.27 demonstrate how the transition matrices of hidden Markov models initialized in various ways, and using one of two training algorithms, change as the length of the sequences they are trained on become longer.

Of particular importance is the fact that hidden Markov models trained using the Baum-Welch algorithm and initialized in various ways can look very different and at other times very similar although not equal, even for very long sequences. It is very clear that hidden Markov models initialized to uniform probabilities do not tend to converge very well (Table A.3, Table A.12, Table A.21, Table A.30, Table A.39, Table A.48), but for the others it doesn't seem so clear. For example, referring to Appendix A, BW-Random-0 and BW-Random-1 are very similar, but BW-Random-0 and BW-Random-4 aren't so similar. On the other hand, BW-Random-4 is very similar to BW-Random-2. Also, even the hidden Markov models that are similar when the training sequence is long are less similar when it is shorter.

Furthermore, even very similar hidden Markov models can have significant differences. Take BW-Random-3 for example, while it has converged to values that are similar to those of both BW-Random-0 and BW-Random-1, both very close to being equal, it isn't close to being equal to them. However, a simple renaming of states or repositioning of some parameter values would make BW-Random-3 approximately it equal to them. Even BW-Random-0 and BW-Random-1 require some emission probabilities to be moved around before they can be equal.

So, what we have here are hidden Markov models that have reached local optima during training. Now, while in many of our test cases the probabilities of the resulting hidden Markov models generating the sequences they were trained on is no different, differing parameters make comparing them difficult. Comparing them is difficult because we have no way of predicting how they might converge and whether they might be very similar or very different from others trained similarly but intialized differently.

Future work might involve looking for relationships between initial and final parameters, but regardless, until we understand how these resulting hidden Markov models are similar and how they differ, or until we can find relationships between them, we'll be unable to gather useful information from one of them or compare two of them. So far, the distance metrics that we have explored are unable to cope with these issues.

6.3 Hidden Markov Model Distance Metrics

The three similarity metrics that have been tested all had shortcomings. Most of all the Kullback-Leibler measure, which future work can explore more closely, since it has so far only returned useless `NaN` values. $coord_1$ seems to provide more information than $coord_2$ since it does seem to decrease to some degree as the length of the training sequence increases.

However, both measures rely solely upon the transition probabilities, making them sensitive to local optima as described in Section 6.2. Further, $coord_1$ sometimes gives questionable results when a sink is present in the transition graph, most notably, sinks tend to lead to component values that sum to more than 1. For example, it doesn't make sense for a hidden Markov model to be in state 1 half of the time, in state 2 half of the time, and in state 3 all of the time.

The similarity averages taken during the tests described in Section 6.2 are given in Appendix B. Future work will seek a more reliable distance metric, looking more closely at the Kullback Leibler measure along the way. Also, coordinates that reliably represent a hidden Markov model must be found before hidden Markov models can be plotted.

6.4 Control over Hidden Markov Model Parameters

As was mentioned in Chapter 5, we are unable to choose or label hidden Markov model parameters before training. This may be easily seen in the tables of Appendix B where similarly trained hidden Markov models result in different, albeit sometimes similar, parameters. This reflects the fact that the training algorithms fit all parameters to the data regardless of any pre-imposed interpretations. Hence, these interpretations may not hold after training. Even more, it may not be as easy to see which states or observations in a hidden Markov model map to which states or observations in the world when we are considering observation sequences that are much more complicated than the very simple observation sequences used in the tests, at least when attempting to do so manually. For example, when randomly generated observation sequences are used, it is much more difficult to pick out any similarities among the resulting hidden Markov models.

The implications of these results are discussed in Chapter 7.

Chapter 7

The Application of Hidden Markov Models to Autism Research

The implications of the results in Chapter 6 are considered here. I will discuss why more work must be done before hidden Markov models can be trained for the purpose of using the resulting parameters for analysis and comparison.

7.1 Lack of World State to HMM State Mapping

In order for researchers to use the parameters in hidden Markov models effectively there must be a meaningful interpretation of those parameters in the world. Unfortunately this is complicated by the fact that we can't yet map mental states to the integer states of a hidden Markov model prior to training and expect that mapping to hold after training, as was discussed in Chapter 6. Furthermore, we won't be able to glean any information from the parameters of a hidden Markov model until we are able to reason about its parameters, but those parameters are so inextricably connected that reasoning about them in isolation may be impractical or impossible. Finding a relationship between levels of autism and the transition matrix may not be possible, for instance. Therefore, a method of mapping hidden Markov model parameters to world model parameters after training must be found. This will be a focus of future work.

7.2 Understanding the Training Algorithms

Before we can use hidden Markov models to model persons mental processes, it will be necessary to better understand the training algorithms that are being used to learn them. Right now we don't fully understand how either the segmental k-means algorithm or the Baum-Welch algorithm divide up the probability space among observation sequences. For example, it isn't fully understood what sort of sequences are given relatively high probability by either of the algorithms. We need confidence that a trained hidden Markov model will generate sequences similar to the training sequence with high probability. All that is known is that the segmental k-means algorithm learns hidden Markov models that generate very few observation sequences and with fairly high probability, while the Baum-Welch algorithm learns hidden Markov models that generate essentially any possible sequence but with very low absolute probability.

7.3 Graphing and Comparing Hidden Markov Models

Learning the differences between hidden Markov models trained on autistic and non-autistic persons respectively requires us to find one or more similarity measures. Without such a measure we won't be able to compare them. Comparing hidden Markov models may require the ability to graph them, but before we can do so we must determine what their coordinates look like. The coordinates explored so far haven't had successful results. Coordinates of autistic persons may or may not form clusters, but they should definitely not typically reside around non-autistic trained hidden Markov model point clusters, or vice-versa. The similarity measure should make sense intuitively and theoretically and depending on the specific application, may or may not be created to distinguish between autistic and non-autistic individuals, rather than being created intuitively and with the expectation of distinguishing between them. The similarity metric may have to consider all of the hidden Markov model parameters together. Finally, we need to be able to compare hidden Markov models regardless of how they were intialized or trained.

7.4 The Length and Number of Training Sequences

Lastly, we need to know more about the necessary length and number of observation sequences. First of all we need to be sure that we are collecting enough data, but we also need to be sure that the amount of data that needs to be collected in practice is feasible. It may not be feasible to require a subject to provide us with seven sequences of length 100. But again, that depends on how much of the subject's time is required to acquire that amount of data.

Chapter 8

Future Work

Since a working similarity metric has not yet been found, future work may look into the visualization of hidden Markov models, exploring various ways of plotting and graphing them. The Kullback-Leibler measure should be further investigated because it may turn out to be an effective hidden Markov model comparison metric. Recall from Chapter 5 that the Kullback-Leibler measure considers the emission probabilities in addition to the transition probabilities, which is exactly what we expect a working hidden Markov model distance metric to need to do. Alternatively, genetic programming may be used to learn one or more distance metrics. One advantage of using genetic programming in this context is that there is no shortage of data. Data, like that in Appendix A, can be generated on the fly in order to learn distance metrics that appropriately minimize the distance between like hidden Markov models; hidden Markov models which we expect to be similar. In other words, we may be able to construct a distance metric that meets our own hidden Markov model similarity criteria.

Future work should also determine how many and what length of training sequences is enough to effectively learn hidden Markov models that represent the subjects that they are intended to represent. Also, we should determine the situations in which the segmental k-means and Baum-Welch algorithms can best be used. One avenue of exploration may be to investigate further the observation sequences that have higher or low probabilites of being generated by hidden Markov models learned using these algorithms. Given the enormous number of observation sequences in even a short sequence, a concept learning approach might be used to learn the concepts induced by the permutations of observation sequences with a probability of being generated by a hidden Markov model above a certain threshold. Doing so might help us to characterize the observation sequences that a hidden Markov model effectively recognizes.

Once we have learned how to train hidden Markov models to model autistic, non-autistic, and other subjects, we might be able to learn similarities and differences between autism and other related disabilities such as Tourrette's, mental retardation, amygdala disorders, or Alzheimer's. Also, while this research has so far concentrated on building models through observation, it might also be possible to use the hidden Markov models that we build as a therapeutic tool[SMLO06,SMMO06]. In this case the hidden Markov models could be used to predict when a subject is in a specific emotional state based upon sensor readings and modify a plant such as a video game in a closed control loop[EMA97]. Such a tool could be used to encourage children or other patients to express various emotions by intelligently eliciting them from the patients based upon their simultaneously sensed emotional state.

Finally, if hidden Markov models prove incapable of modeling the mental processes of autistic and non-autistic persons, other models, such as Markov models, may be explored.

Bibliography

[AGBOB02] J. Abrisqueta-Gomez, O.F.A. Bueno, M.G.M. Oliveira, and P.H.F. Bertolucci, *Recognition memory for emotional pictures in alzheimer's patients*, Acta Neurologica Scandinavica **105** (2002), no. 1, 51–54.

[AsK93] Oscar E. Agazzi and Shyh shiaw Kuo, *Hidden markov model based optical character recognition in the presence of deterministic transformations*, Pattern Recognition **26** (1993), no. 12, 1813–1826.

[BC90] Simon Baron-Cohen, *Autism: A specific cognitive disorder of 'mind-blindness'*, International Review of Psychiatry **2** (1990), no. 1, 81–90.

[BC02] ———, *The extreme mail brain theory of autism*, Trends in Cognitive Sciences **6** (2002), no. 6, 248–254.

[BC06] ———, *Two new theories of autism: Hyper-systemising and assoritive mating*, Archives of Disease in Childhood **91** (2006), no. 1, 2–5.

[BCWS+97] Simon Baron-Cohen, Sally Wheelwright, Carol Scott, Patrick Bolton, and Ian Goodyer, *Is there a link between engineering and autism?*, Autism **1** (1997), no. 1, 101–109.

[Cas05] Fulvia Castelli, *Understanding emotions from standardized facial expressions in autism and normal development*, Autism **9** (2005), no. 4, 428–449.

[CNB98] M.S. Crouse, R.D. Nowak, and R.G. Baraniuk, *Wavelet-based statistical signal processing using hidden markov models*, IEEE Transactions on Signal Processing **46** (1998), no. 4, 886–902.

[Dan] http://www.psychology.emory.edu/clinical/interpersonal/danva.htm.

[DD96] Rakesh Dugad and U.B. Desai, *A tutorial on hidden markov models*, Tech. report, Indian Institute of Technology, May 1996.

[EMA97] Robert James Elliott, John B. Moore, and Lakhdar Aggoun, *Hidden markov models: Estimation and control*, 1 ed., vol. 29, Springer-Verlag New York, LLC, January 1997.

[FH94] Uta Frith and Francesca Happe, *Autism: Beyond "theory of mind."*, Cognition **50** (1994), no. 1-3, 115–132.

[Fra06] Jean-Marc Francois, *Jahmm - hidden markov model (hmm): An implementation in java*, 2006, URL: http://www.run.montefiore.ulg.ac.be/˜ francois/software/jahmm/.

[Fri01] Uta Frith, *Mind blindness and the brain in autism*, Neuron **32** (2001), 969–979.

[FRW95] Markus Falkhausen, Herbert Reininger, and Dietrich Wolf, *Calculation of distance measures between hidden markov models*, In EUROSPEECH-1995 (1995), 1487–1490.

[GG98] Vittorio Gallese and Alvin Goldman, *Mirror neurons and the simulation theory of mind-reading*, Trends in Cognitive Sciences **2** (1998), no. 12, 493–501.

[HMS02] Rita Hargrave, Richard J. Maddock, and Valerie Stone, *Impaired recognition of facial expressions of emotion in alzheimer's disease*, Journal of Neuropsychiatry and Clinical Neurosciences **14** (2002), 64–71.

[JR90] Biing-Hwang Juang and L.R. Rabiner, *The segmental k-means algorithm for estimating parameters of hidden markov models*, IEEE Trans. Acoust. Speech Signal Process. **38** (1990), no. 9, 1639–1641.

[KBM+94] A. Krogh, M. Brown, I.S. Mian, K. Sjolander, and D. Haussler, *Hidden markov models in computational biology. applications to protein modeling.*, Journal of Computational Biology **235** (1994), no. 5, 1501–1531.

[LHK96] M.P. Lawton, K. Van Haitsma, and J. Klapper, *Observed affect in nursing home residents with alzheimer's disease*, Journals of Gerontology Series B: Psychological Sciences and Social Sciences **51** (1996), no. 1, 3–14.

[OG02] Nuria Oliver and Ashutosh Garg, *Mmihmm: Maximum mutual information hidden markov models*, ICML '02: Proceedings of the Nineteenth International Conference on Machine Learning, 2002, pp. 466–473.

[Rab89] L.R. Rabiner, *A tutorial on hidden markov models and selected applications in speech recognition*, vol. 77, 1989, pp. 257–286.

[RC04] G Rizzolatti and L Craighero, *The mirror-neuron system*, Annual Review of Neuroscience **27** (2004), 169–192.

[RJ86] L. Rabiner and B. Juang, *An introduction to hidden markov models*, IEEE ASSP Magazine **3** (1986), no. 1, 4–16.

[Sab04] Mark A. Sabbagh, *Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism*, Brain and Cognition **55** (2004), no. 1, 209–219.

[Sch00] Jay Schulkin, *Theory of mind and mirroring neurons*, Trends in Cognitive Sciences **4** (2000), no. 7, 252–254.

[SMLO06] P. Ravindra De Silva, Ajith P. Madurapperuma, Stephen G. Lambacher, and Minetada Osana, *Therapeutic tool for develop child nonverbal communication skills through interactive game*, cirmca **0** (2006), 33.

[SMMO06] P. Ravindra De Silva, Ajith P. Madurapperuma, Ashu Marasinghe, and Minetada Osano, *A multi-agent based interactive system towards childs emotion performances quantified through affective body gestures*, icpr **1** (2006), 1236–1239.

[SOMM06] P. Ravindra De Silva, Minetada Osano, Ashu Marasinghe, and Ajith Madurapperuma, *Towards recognizing emotion with affective dimensions through body gestures*, fg **0** (2006), 269–274.

[SZW+05] Ye Sun, Jian-Ming Zhang, Liang-Min Wang, Yong-Zhao Zhan, and Shun-Lin Song, *A novel method of recognizing ageing face based on ehmm*, Proceedings of 2005 International Conference on Machine Learning and Cybernetics, vol. 8, 2005, pp. 4599–4604.

[War96] Narada Warakagoda, *http://jedlik.phy.bme.hu/˜ gerjanos/hmm/node2.html*, 1996.

[WRE+98] Paul J. Whalen, Scott L. Rauch, Nancy L. Etcoff, Sean C. McInerney, Michael B. Lee, and Michael A. Jenike, *Masked presentations of emothional facial expressions modulate amygdala activity without explicit knowledge*, 1998, pp. 411–418.

[WWSP01] J.H.G. Williams, A. Whiten, T. Suddendorf, and D.I. Perrett, *Imitation, mirror neurons, autism*, Neuroscience and Biobehavioral Reviews **25** (2001), no. 4, 287–295.

[YKSM89] Nurit Yirmiya, Connie Kasari, Marian Sigman, and Peter Mundy, *Facial expressions of affect in autistic, mentally retarded, and normal children*, Journal of Child Psychology and Psychiatry **30** (1989), no. 55, 725–735.

Appendix A

Convergence Test Data Tables

The following tables record the results of testing the effect of increasingly longer training sequences on hidden Markov models initialized and trained in differing ways. Both the segmental k-means algorithm and the Baum-Welch algorithm were used. While the segmental k-means algorithm as implemented in JaHMM is not parameterized with an initial hidden Markov model, the Baum-Welch algorithm was applied to several initial models. Eight initial hidden Markov models were used. One hidden Markov model was initialized using the segmental k-means algorithm, one using uniform probabilities, and the remaining six with random parameters. Six hidden Markov models were randomly generated and then held constant throughout the test.

One test was performed on the hidden Markov model trained using the segmental k-means algorithm and each of eight the differently initialized hidden Markov models trained using the Baum-Welch algorithm, using `REPEAT = 3` respectively. `STEP` was chosen to be non-constant to keep the tables at a reasonable size. The exact sequences lengths were 2,3,4,5,8,9,16,17,32,33,64,65,128,129. These lengths allowed testing using a wide variation of sequence lengths and also allowed the comparison of hidden Markov models that were one step apart.

Table A.1: *Probability values for the segmental k-means trained HMM.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	2.5000E-1
5	1.4815E-1
8	2.1948E-2
9	2.1948E-2
16	7.1364E-5
17	5.0805E-5
32	3.4828E-9
33	2.4868E-9
64	3.8426E-18
65	2.5912E-18
128	9.9011E-36
129	6.7144E-36

Table A.2: *Probability values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length	*Probability of Training Sequence*
2	?
3	1.0000E0
4	2.5000E-1
5	1.4815E-1
8	2.1948E-2
9	2.1948E-2
16	7.1364E-5
17	5.0805E-5
32	3.4828E-9
33	2.4868E-9
64	3.8426E-18
65	2.5912E-18
128	9.9011E-36
129	6.7144E-36

Table A.3: *Probability values for BW-Uniform-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	1.0547E-1
5	3.4560E-2
8	1.7381E-4
9	5.0805E-5
16	3.0209E-8
17	8.2231E-9
32	6.7434E-16
33	2.3855E-16
64	2.9580E-31
65	1.0318E-31
128	8.7495E-62
129	3.0299E-62

Table A.4: *Probability values for BW-Random-0-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	3.1636E-1
5	2.9627E-1
8	2.1732E-2
9	2.1947E-2
16	7.1349E-5
17	5.0803E-5
32	3.4827E-9
33	2.4868E-9
64	3.8426E-18
65	2.5912E-18
128	9.9011E-36
129	6.7144E-36

Table A.5: *Probability values for BW-Random-1-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	3.1596E-1
5	2.9495E-1
8	2.1440E-2
9	2.1936E-2
16	7.1276E-5
17	5.0793E-5
32	3.4826E-9
33	2.4867E-9
64	3.8422E-18
65	2.5911E-18
128	9.8989E-36
129	6.7131E-36

Table A.6: *Probability values for BW-Random-2-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	1.9056E-1
5	1.5586E-1
8	5.8415E-3
9	3.2505E-3
16	6.5536E-6
17	1.3633E-6
32	4.1598E-12
33	2.0312E-12
64	6.0186E-24
65	4.0586E-24
128	2.4290E-47
129	1.6472E-47

Table A.7: *Probability values for BW-Random-3-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	3.1638E-1
5	2.9627E-1
8	2.1176E-2
9	2.1943E-2
16	7.0744E-5
17	5.0775E-5
32	3.4807E-9
33	2.4823E-9
64	3.6323E-18
65	4.0771E-22
128	9.8845E-36
129	6.6979E-36

Table A.8: *Probability values for BW-Random-4-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	1.8751E-1
5	2.9378E-1
8	1.0097E-2
9	3.9062E-3
16	2.4317E-6
17	4.6109E-7
32	3.1158E-12
33	2.2248E-12
64	4.0586E-24
65	9.5405E-25
128	5.5562E-48
129	3.7679E-48

Table A.9: *Probability values for BW-Random-5-trained.*

Length	*Probability of Training Sequence*
2	1.0000E0
3	1.0000E0
4	2.5000E-1
5	2.5000E-1
8	8.6400E-3
9	3.9062E-3
16	1.6858E-6
17	5.1947E-7
32	1.0311E-12
33	7.4004E-13
64	3.5716E-18
65	2.4397E-18
128	9.4840E-36
129	6.2900E-36

Table A.10: *Initial probability values for the segmental k-means trained HMM.*

Length	π	
2	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
3	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
4	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
5	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
8	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
9	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
16	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
17	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
32	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
33	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
64	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
65	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
128	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
129	0	1.0000E0
	1	0.0000E0
	2	0.0000E0

Table A.11: *Initial probability values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length		π
2	0	?
	1	?
	2	?
3	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
4	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
5	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
8	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
9	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
16	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
17	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
32	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
33	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
64	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
65	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
128	0	1.0000E0
	1	0.0000E0
	2	0.0000E0
129	0	1.0000E0
	1	0.0000E0
	2	0.0000E0

Table A.12: *Initial probability values for BW-Uniform-trained.*

Length	π	
2	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
3	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
4	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
5	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
8	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
9	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
16	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
17	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
32	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
33	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
64	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
65	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
128	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1
129	0	3.3333E-1
	1	3.3333E-1
	2	3.3333E-1

Table A.13: *Initial probability values for BW-Random-0-trained.*

Length	π	
2	0	9.9415E-1
	1	5.5163E-3
	2	3.3672E-4
3	0	9.9446E-1
	1	5.2721E-3
	2	2.7061E-4
4	0	1.0000E0
	1	1.0967E-15
	2	7.3614E-38
5	0	1.0000E0
	1	1.4633E-24
	2	2.6964E-152
8	0	1.0000E0
	1	5.9504E-178
	2	0.0000E0
9	0	1.0000E0
	1	2.7137E-207
	2	0.0000E0
16	0	1.0000E0
	1	5.2448E-183
	2	1.3438E-133
17	0	1.0000E0
	1	2.9502E-194
	2	3.8207E-142
32	0	1.0000E0
	1	7.6598E-168
	2	7.2402E-113
33	0	1.0000E0
	1	5.7988E-171
	2	1.2015E-114
64	0	1.0000E0
	1	7.2712E-164
	2	1.4984E-99
65	0	1.0000E0
	1	2.3067E-172
	2	5.5803E-113
128	0	1.0000E0
	1	4.7239E-174
	2	1.6154E-112
129	0	1.0000E0
	1	2.3069E-175
	2	2.4554E-113

Table A.14: *Initial probability values for BW-Random-1-trained.*

Length		π
2	0	1.4230E-1
	1	7.8972E-1
	2	6.7986E-2
3	0	1.4354E-1
	1	7.8830E-1
	2	6.8154E-2
4	0	7.2493E-17
	1	1.0000E0
	2	3.1413E-10
5	0	6.7032E-74
	1	1.0000E0
	2	1.1408E-24
8	0	0.0000E0
	1	1.0000E0
	2	5.1718E-161
9	0	0.0000E0
	1	1.0000E0
	2	3.6300E-201
16	0	1.8333E-82
	1	1.0000E0
	2	7.7127E-140
17	0	5.8165E-89
	1	1.0000E0
	2	4.0552E-152
32	0	1.6950E-57
	1	1.0000E0
	2	1.5298E-118
33	0	5.8682E-57
	1	1.0000E0
	2	6.7831E-119
64	0	7.7333E-47
	1	1.0000E0
	2	2.7847E-115
65	0	7.9845E-52
	1	1.0000E0
	2	9.5373E-117
128	0	1.8534E-48
	1	1.0000E0
	2	8.4217E-117
129	0	2.7361E-49
	1	1.0000E0
	2	3.3310E-117

Table A.15: *Initial probability values for BW-Random-2-trained.*

Length		π
2	0	1.4903E-3
	1	3.8591E-2
	2	9.5992E-1
3	0	1.6309E-3
	1	3.8282E-2
	2	9.6009E-1
4	0	1.0211E-16
	1	1.2535E-7
	2	1.0000E0
5	0	3.4758E-58
	1	7.8282E-92
	2	1.0000E0
8	0	8.3490E-226
	1	0.0000E0
	2	1.0000E0
9	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
16	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
17	0	4.8465E-239
	1	4.0476E-193
	2	1.0000E0
32	0	0.0000E0
	1	2.4960E-280
	2	1.0000E0
33	0	0.0000E0
	1	1.3089E-286
	2	1.0000E0
64	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
65	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
128	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
129	0	0.0000E0
	1	0.0000E0
	2	1.0000E0

Table A.16: *Initial probability values for BW-Random-3-trained.*

Length	π	
2	0	2.6427E-3
	1	2.2862E-1
	2	7.6873E-1
3	0	2.5158E-3
	1	2.2108E-1
	2	7.7640E-1
4	0	1.7016E-33
	1	3.1856E-9
	2	1.0000E0
5	0	9.4779E-123
	1	1.2432E-11
	2	1.0000E0
8	0	0.0000E0
	1	2.0794E-68
	2	1.0000E0
9	0	0.0000E0
	1	3.7904E-126
	2	1.0000E0
16	0	2.1079E-110
	1	2.1012E-68
	2	1.0000E0
17	0	1.8442E-118
	1	2.8522E-95
	2	1.0000E0
32	0	3.6470E-95
	1	1.0793E-67
	2	1.0000E0
33	0	4.9101E-98
	1	1.4753E-56
	2	1.0000E0
64	0	3.8787E-101
	1	1.0000E0
	2	2.6077E-7
65	0	6.5618E-102
	1	4.1987E-4
	2	9.9958E-1
128	0	9.4496E-118
	1	1.0000E0
	2	3.6727E-22
129	0	1.4924E-116
	1	1.0000E0
	2	1.2515E-20

Table A.17: *Initial probability values for BW-Random-4-trained.*

Length	π	
2	0	2.0790E-1
	1	7.4036E-1
	2	5.1735E-2
3	0	3.3989E-1
	1	6.1274E-1
	2	4.7369E-2
4	0	7.0630E-1
	1	2.9370E-1
	2	2.6886E-11
5	0	3.8996E-7
	1	1.0000E0
	2	8.7697E-48
8	0	2.3848E-9
	1	1.0000E0
	2	8.9069E-158
9	0	1.6809E-14
	1	1.0000E0
	2	7.9362E-156
16	0	3.1351E-2
	1	9.6865E-1
	2	5.9792E-55
17	0	1.8679E-8
	1	1.0000E0
	2	1.8325E-52
32	0	2.3628E-42
	1	1.0000E0
	2	9.7973E-71
33	0	1.4831E-50
	1	1.0000E0
	2	4.2448E-84
64	0	5.1854E-29
	1	1.0000E0
	2	3.2007E-70
65	0	1.4509E-17
	1	1.0000E0
	2	3.9221E-46
128	0	1.6404E-19
	1	1.0000E0
	2	8.0283E-46
129	0	3.0429E-24
	1	1.0000E0
	2	4.1134E-53

Table A.18: *Initial probability values for BW-Random-5-trained.*

Length		π
2	0	2.5671E-1
	1	1.5724E-1
	2	5.8604E-1
3	0	2.0462E-1
	1	1.3141E-1
	2	6.6397E-1
4	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
5	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
8	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
9	0	0.0000E0
	1	0.0000E0
	2	1.0000E0
16	0	4.1879E-74
	1	9.7080E-11
	2	1.0000E0
17	0	3.0946E-71
	1	1.3856E-9
	2	1.0000E0
32	0	3.2317E-51
	1	1.9773E-4
	2	9.9980E-1
33	0	2.2322E-55
	1	1.9271E-4
	2	9.9981E-1
64	0	1.5839E-53
	1	1.2892E-9
	2	1.0000E0
65	0	3.9104E-53
	1	1.8778E-9
	2	1.0000E0
128	0	6.0591E-55
	1	1.4666E-11
	2	1.0000E0
129	0	2.7770E-54
	1	1.6956E-10
	2	1.0000E0

Table A.19: *Transition probability values for the segmental k-means trained HMM.*

Length	*A*			
2		0	1	2
	0	0.0000E0	1.0000E0	0.0000E0
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
3		0	1	2
	0	0.0000E0	1.0000E0	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
4		0	1	2
	0	5.0000E-1	5.0000E-1	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
5		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
8		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	0.0000E0	0.0000E0	1.0000E0
9		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	0.0000E0	0.0000E0	1.0000E0
16		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
17		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	2.5000E-1	0.0000E0	7.5000E-1
32		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	7.0000E-1	3.0000E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
33		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	7.2727E-1	2.7273E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
64		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
65		0	1	2
	0	6.8182E-1	3.1818E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
128		0	1	2
	0	6.7442E-1	3.2558E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
129		0	1	2
	0	6.8182E-1	3.1818E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1

Table A.20: *Transition probability values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length		*A*		
2		0	1	2
	0	?	?	?
	1	?	?	?
	2	?	?	?
3		0	1	2
	0	0.0000E0	1.0000E0	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
4		0	1	2
	0	5.0000E-1	5.0000E-1	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
5		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
8		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	0.0000E0	0.0000E0	1.0000E0
9		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	0.0000E0	0.0000E0	1.0000E0
16		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
17		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	2.5000E-1	0.0000E0	7.5000E-1
32		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	7.0000E-1	3.0000E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
33		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	7.2727E-1	2.7273E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
64		0	1	2
	0	6.6667E-1	3.3333E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
65		0	1	2
	0	6.8182E-1	3.1818E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
128		0	1	2
	0	6.7442E-1	3.2558E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1
129		0	1	2
	0	6.8182E-1	3.1818E-1	0.0000E0
	1	0.0000E0	6.6667E-1	3.3333E-1
	2	3.3333E-1	0.0000E0	6.6667E-1

Table A.21: *Transition probability values for BW-Uniform-trained.*

Length	*A*			
2		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
3		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
4		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
5		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
8		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
9		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
16		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
17		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
32		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
33		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
64		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
65		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
128		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
129		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1

Table A.22: *Transition probability values for BW-Random-0-trained.*

Length		*A*		
2		0	1	2
	0	8.6014E-1	1.3834E-1	1.5231E-3
	1	8.1480E-1	1.1071E-1	7.4487E-2
	2	6.2186E-1	2.3680E-1	1.4134E-1
3		0	1	2
	0	9.0546E-1	9.3632E-2	9.1041E-4
	1	8.2660E-1	1.0433E-1	6.9067E-2
	2	7.1240E-1	1.8675E-1	1.0085E-1
4		0	1	2
	0	2.5807E-1	7.4193E-1	1.0279E-7
	1	7.7417E-35	2.5113E-1	7.4887E-1
	2	1.5151E-84	5.7475E-33	1.0000E0
5		0	1	2
	0	3.3850E-1	6.6150E-1	7.0779E-8
	1	2.2548E-37	3.2807E-1	6.7193E-1
	2	8.6202E-138	7.4740E-33	1.0000E0
8		0	1	2
	0	6.6667E-1	3.3333E-1	2.3067E-32
	1	6.1576E-69	6.6208E-1	3.3792E-1
	2	9.5522E-166	1.1373E-18	1.0000E0
9		0	1	2
	0	6.6667E-1	3.3333E-1	1.7900E-86
	1	1.1818E-78	6.6665E-1	3.3335E-1
	2	7.5790E-194	2.2824E-40	1.0000E0
16		0	1	2
	0	6.6667E-1	3.3333E-1	1.0788E-75
	1	5.3272E-73	6.6664E-1	3.3336E-1
	2	3.3328E-1	1.9831E-32	6.6672E-1
17		0	1	2
	0	6.6667E-1	3.3333E-1	3.9793E-95
	1	5.3814E-76	6.6666E-1	3.3334E-1
	2	2.5000E-1	8.9219E-42	7.5000E-1
32		0	1	2
	0	6.6667E-1	3.3333E-1	1.3352E-88
	1	3.0912E-70	7.0000E-1	3.0000E-1
	2	3.3333E-1	1.0196E-54	6.6667E-1
33		0	1	2
	0	6.6667E-1	3.3333E-1	1.6336E-90
	1	3.8822E-72	7.2727E-1	2.7273E-1
	2	3.3333E-1	6.0485E-53	6.6667E-1
64		0	1	2
	0	6.6667E-1	3.3333E-1	8.8479E-82
	1	7.6430E-77	6.6667E-1	3.3333E-1
	2	3.3333E-1	7.3548E-59	6.6667E-1
65		0	1	2
	0	6.8182E-1	3.1818E-1	1.0097E-84
	1	1.3703E-75	6.6667E-1	3.3333E-1
	2	3.3333E-1	3.9889E-59	6.6667E-1
128		0	1	2
	0	6.7442E-1	3.2558E-1	1.1210E-82
	1	2.2836E-76	6.6667E-1	3.3333E-1
	2	3.3333E-1	2.8927E-59	6.6667E-1
129		0	1	2
	0	6.8182E-1	3.1818E-1	1.5446E-83
	1	7.3191E-77	6.6667E-1	3.3333E-1
	2	3.3333E-1	3.0474E-59	6.6667E-1

Table A.23: *Transition probability values for BW-Random-1-trained.*

Length	A			
2		0	1	2
	0	3.9327E-1	3.5683E-1	2.4990E-1
	1	9.0892E-2	4.9990E-1	4.0921E-1
	2	8.9439E-1	4.8244E-2	5.7369E-2
3		0	1	2
	0	3.9293E-1	3.7785E-1	2.2922E-1
	1	9.1361E-2	5.4489E-1	3.6374E-1
	2	8.9510E-1	4.9546E-2	5.5357E-2
4		0	1	2
	0	1.0000E0	4.5518E-35	2.3415E-13
	1	1.3844E-4	2.7695E-1	7.2292E-1
	2	7.7273E-1	1.2107E-17	2.2727E-1
5		0	1	2
	0	1.0000E0	1.0902E-69	7.1867E-13
	1	3.3728E-5	3.7078E-1	6.2918E-1
	2	7.0868E-1	1.1019E-18	2.9132E-1
8		0	1	2
	0	1.0000E0	1.1777E-115	5.5678E-12
	1	3.1170E-20	6.6667E-1	3.3333E-1
	2	3.4485E-1	5.6461E-57	6.5515E-1
9		0	1	2
	0	1.0000E0	5.5556E-131	2.0143E-25
	1	2.8004E-52	6.6667E-1	3.3333E-1
	2	3.3345E-1	1.6396E-52	6.6655E-1
16		0	1	2
	0	6.6694E-1	3.3306E-1	8.5525E-22
	1	1.2171E-49	6.6667E-1	3.3333E-1
	2	3.3347E-1	1.6078E-48	6.6653E-1
17		0	1	2
	0	7.5003E-1	2.4997E-1	3.6306E-29
	1	5.9895E-62	6.6667E-1	3.3333E-1
	2	3.3336E-1	2.6741E-46	6.6664E-1
32		0	1	2
	0	6.6667E-1	3.3333E-1	1.9669E-40
	1	1.4990E-53	6.6667E-1	3.3333E-1
	2	3.0000E-1	2.4227E-39	7.0000E-1
33		0	1	2
	0	6.6667E-1	3.3333E-1	3.9915E-38
	1	1.3508E-52	6.6667E-1	3.3333E-1
	2	2.7273E-1	5.5349E-38	7.2727E-1
64		0	1	2
	0	6.6667E-1	3.3333E-1	3.7419E-42
	1	4.5635E-45	6.6667E-1	3.3333E-1
	2	3.3333E-1	3.1797E-37	6.6667E-1
65		0	1	2
	0	6.6667E-1	3.3333E-1	4.8317E-42
	1	3.7848E-49	6.8182E-1	3.1818E-1
	2	3.3333E-1	1.7396E-38	6.6667E-1
128		0	1	2
	0	6.6667E-1	3.3333E-1	2.0756E-41
	1	2.9172E-46	6.7442E-1	3.2558E-1
	2	3.3334E-1	2.0988E-35	6.6666E-1
129		0	1	2
	0	6.6667E-1	3.3333E-1	1.9658E-41
	1	5.6341E-47	6.8182E-1	3.1818E-1
	2	3.3334E-1	4.3131E-36	6.6666E-1

Table A.24: *Transition probability values for BW-Random-2-trained.*

Length	A			
2		0	1	2
	0	2.2318E-2	5.3199E-1	4.4569E-1
	1	8.2448E-2	2.0393E-2	8.9716E-1
	2	1.0707E-1	7.8300E-2	8.1463E-1
3		0	1	2
	0	2.2247E-2	5.3274E-1	4.4502E-1
	1	7.9910E-2	2.1566E-2	8.9852E-1
	2	1.0063E-1	8.6109E-2	8.1326E-1
4		0	1	2
	0	2.2745E-1	7.7255E-1	9.4147E-9
	1	9.7801E-1	2.1987E-2	6.6704E-9
	2	2.4444E-1	2.5216E-1	5.0340E-1
5		0	1	2
	0	6.4076E-2	9.3592E-1	2.3759E-19
	1	7.1457E-1	2.8543E-1	9.5117E-72
	2	3.5447E-1	7.5406E-3	6.3799E-1
8		0	1	2
	0	5.3108E-2	9.4689E-1	6.6422E-76
	1	7.0008E-1	2.9992E-1	1.7566E-93
	2	3.3311E-1	2.1974E-4	6.6667E-1
9		0	1	2
	0	2.8991E-5	9.9997E-1	1.0715E-105
	1	9.9964E-1	3.6440E-4	2.0936E-112
	2	3.3333E-1	2.5878E-9	6.6667E-1
16		0	1	2
	0	9.4297E-14	1.0000E0	4.0557E-94
	1	7.5000E-1	2.4825E-20	2.5000E-1
	2	3.3333E-1	1.0474E-18	6.6667E-1
17		0	1	2
	0	2.5618E-5	9.9997E-1	3.4112E-28
	1	7.9985E-1	1.4387E-4	2.0001E-1
	2	3.3333E-1	7.6712E-9	6.6667E-1
32		0	1	2
	0	1.4254E-10	1.0000E0	1.1054E-60
	1	6.6667E-1	2.3931E-58	3.3333E-1
	2	3.3333E-1	1.0319E-11	6.6667E-1
33		0	1	2
	0	2.3188E-11	1.0000E0	4.7620E-63
	1	7.0000E-1	1.5247E-14	3.0000E-1
	2	3.3333E-1	1.9292E-12	6.6667E-1
64		0	1	2
	0	1.0668E-85	1.0000E0	4.7115E-86
	1	6.6667E-1	3.5759E-87	3.3333E-1
	2	3.3333E-1	9.6972E-292	6.6667E-1
65		0	1	2
	0	1.2168E-120	1.0000E0	7.6883E-121
	1	6.6667E-1	4.4139E-122	3.3333E-1
	2	3.1818E-1	0.0000E0	6.8182E-1
128		0	1	2
	0	4.3212E-124	1.0000E0	3.2488E-124
	1	6.6667E-1	2.4337E-125	3.3333E-1
	2	3.2558E-1	0.0000E0	6.7442E-1
129		0	1	2
	0	6.9575E-128	1.0000E0	5.2506E-128
	1	6.6667E-1	3.0599E-129	3.3333E-1
	2	3.1818E-1	0.0000E0	6.8182E-1

Table A.25: *Transition probability values for BW-Random-3-trained.*

Length	*A*			
2		0	1	2
	0	9.1584E-2	5.7224E-1	3.3617E-1
	1	4.0410E-2	6.2943E-1	3.3016E-1
	2	2.0790E-2	5.2352E-1	4.5569E-1
3		0	1	2
	0	8.5839E-2	5.6609E-1	3.4807E-1
	1	3.5029E-2	6.1253E-1	3.5244E-1
	2	1.5114E-2	4.8388E-1	5.0101E-1
4		0	1	2
	0	1.0000E0	1.9081E-41	8.0767E-103
	1	7.4618E-1	2.5382E-1	2.1707E-42
	2	6.2135E-9	7.4603E-1	2.5397E-1
5		0	1	2
	0	1.0000E0	2.3282E-56	7.1214E-169
	1	6.6709E-1	3.3291E-1	6.5536E-58
	2	4.7986E-10	6.6638E-1	3.3362E-1
8		0	1	2
	0	1.0000E0	4.8360E-14	3.3040E-159
	1	3.5175E-1	6.4825E-1	5.6022E-31
	2	2.0842E-16	3.3335E-1	6.6665E-1
9		0	1	2
	0	1.0000E0	1.0930E-32	1.8582E-181
	1	3.3338E-1	6.6662E-1	1.6033E-52
	2	3.6321E-64	3.3333E-1	6.6667E-1
16		0	1	2
	0	6.6858E-1	5.9025E-17	3.3142E-1
	1	3.3429E-1	6.6571E-1	5.7782E-31
	2	1.6468E-30	3.3334E-1	6.6666E-1
17		0	1	2
	0	7.5007E-1	5.6261E-27	2.4993E-1
	1	3.3339E-1	6.6661E-1	5.8922E-41
	2	4.4039E-55	3.3333E-1	6.6667E-1
32		0	1	2
	0	6.6668E-1	2.0880E-29	3.3332E-1
	1	3.0001E-1	6.9999E-1	4.8923E-31
	2	4.4086E-45	3.3334E-1	6.6666E-1
33		0	1	2
	0	6.6673E-1	3.0735E-23	3.3327E-1
	1	2.7277E-1	7.2723E-1	2.3058E-27
	2	1.1046E-37	3.3336E-1	6.6664E-1
64		0	1	2
	0	6.6695E-1	5.5483E-21	3.3305E-1
	1	3.3405E-1	6.6595E-1	1.8720E-17
	2	6.0348E-18	3.3214E-1	6.6786E-1
65		0	1	2
	0	7.2328E-1	2.9400E-8	2.7672E-1
	1	3.8231E-1	6.1769E-1	7.0305E-10
	2	9.9523E-9	3.8552E-1	6.1448E-1
128		0	1	2
	0	6.6668E-1	1.3687E-35	3.3332E-1
	1	3.2559E-1	6.7441E-1	5.1757E-30
	2	7.7652E-30	3.3333E-1	6.6667E-1
129		0	1	2
	0	6.6668E-1	2.6226E-34	3.3332E-1
	1	3.1819E-1	6.8181E-1	2.3504E-28
	2	4.4740E-28	3.3334E-1	6.6666E-1

Table A.26: *Transition probability values for BW-Random-4-trained.*

Length	*A*			
		0	1	2
2	0	1.6371E-2	7.0043E-1	2.8320E-1
	1	8.8263E-1	7.3715E-2	4.3658E-2
	2	7.3080E-1	2.4040E-2	2.4516E-1
		0	1	2
3	0	1.2679E-2	7.1977E-1	2.6755E-1
	1	8.1108E-1	1.2571E-1	6.3215E-2
	2	6.9307E-1	2.9368E-2	2.7756E-1
		0	1	2
4	0	1.1136E-3	3.7748E-1	6.2141E-1
	1	7.2568E-1	1.4393E-2	2.5993E-1
	2	6.2315E-6	3.3741E-12	9.9999E-1
		0	1	2
5	0	2.9487E-1	7.4391E-13	7.0513E-1
	1	6.4123E-1	3.5854E-1	2.3019E-4
	2	1.4801E-29	7.4489E-48	1.0000E0
		0	1	2
8	0	2.6438E-1	1.9268E-9	7.3562E-1
	1	6.0486E-1	3.9434E-1	7.9798E-4
	2	5.6883E-20	3.3972E-83	1.0000E0
		0	1	2
9	0	7.2755E-17	1.0000E0	1.8555E-14
	1	5.0000E-1	3.2483E-14	5.0000E-1
	2	2.5125E-52	2.6188E-55	1.0000E0
		0	1	2
16	0	1.4169E-2	9.6603E-1	1.9799E-2
	1	5.8604E-2	5.6160E-1	3.7980E-1
	2	1.3746E-1	3.9184E-6	8.6253E-1
		0	1	2
17	0	9.8028E-4	9.9854E-1	4.7898E-4
	1	5.2121E-1	1.7582E-3	4.7704E-1
	2	1.4412E-1	5.7099E-5	8.5583E-1
		0	1	2
32	0	1.9436E-31	1.0000E0	1.4675E-30
	1	6.3636E-1	2.0796E-29	3.6364E-1
	2	3.0000E-1	1.5969E-44	7.0000E-1
		0	1	2
33	0	1.4400E-35	1.0000E0	1.1478E-34
	1	6.3636E-1	1.7671E-33	3.6364E-1
	2	2.7273E-1	8.3827E-53	7.2727E-1
		0	1	2
64	0	4.2968E-11	1.0000E0	3.9131E-30
	1	6.6667E-1	6.9996E-12	3.3333E-1
	2	3.3333E-1	1.0601E-12	6.6667E-1
		0	1	2
65	0	2.3668E-8	1.0000E0	9.0326E-19
	1	6.8181E-1	3.6026E-6	3.1818E-1
	2	3.3333E-1	6.5547E-10	6.6667E-1
		0	1	2
128	0	9.4438E-10	1.0000E0	8.7408E-21
	1	6.7442E-1	2.0824E-7	3.2558E-1
	2	3.3333E-1	3.7958E-11	6.6667E-1
		0	1	2
129	0	2.3903E-11	1.0000E0	1.3088E-25
	1	6.7442E-1	1.1188E-9	3.2558E-1
	2	3.3333E-1	1.7646E-13	6.6667E-1

Table A.27: *Transition probability values for BW-Random-5-trained.*

Length	*A*			
2		0	1	2
	0	1.0259E-1	3.2004E-1	5.7737E-1
	1	8.6086E-2	3.9881E-1	5.1510E-1
	2	3.4994E-1	6.2348E-1	2.6580E-2
3		0	1	2
	0	1.2123E-1	3.8077E-1	4.9800E-1
	1	9.1792E-2	4.2657E-1	4.8164E-1
	2	3.5162E-1	6.3345E-1	1.4932E-2
4		0	1	2
	0	0.0000E0	0.0000E0	1.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	7.3105E-1	2.6895E-1	0.0000E0
5		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	5.0000E-1	5.0000E-1	0.0000E0
8		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	5.0000E-1	5.0000E-1	0.0000E0
9		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	0.0000E0	1.0000E0
	2	5.0000E-1	5.0000E-1	0.0000E0
16		0	1	2
	0	8.5406E-1	1.4594E-1	3.8729E-7
	1	2.7064E-5	3.0012E-2	9.6996E-1
	2	5.0463E-1	4.9537E-1	3.5591E-7
17		0	1	2
	0	8.7618E-1	1.2382E-1	5.8042E-6
	1	3.7178E-4	4.4171E-2	9.5546E-1
	2	5.0904E-1	4.9095E-1	7.2835E-6
32		0	1	2
	0	7.0006E-1	2.9994E-1	2.1287E-27
	1	3.2445E-6	6.2533E-1	3.7467E-1
	2	6.1311E-1	3.7135E-1	1.5535E-2
33		0	1	2
	0	7.2729E-1	2.7271E-1	1.5334E-28
	1	2.2807E-6	6.2688E-1	3.7311E-1
	2	6.1330E-1	3.6822E-1	1.8472E-2
64		0	1	2
	0	6.6668E-1	3.3332E-1	1.2201E-34
	1	1.9392E-18	6.6942E-1	3.3058E-1
	2	3.3475E-1	7.0287E-10	6.6525E-1
65		0	1	2
	0	6.6668E-1	3.3332E-1	2.8503E-33
	1	1.2337E-17	6.6805E-1	3.3195E-1
	2	3.1946E-1	1.7490E-9	6.8054E-1
128		0	1	2
	0	6.6668E-1	3.3332E-1	6.1724E-35
	1	1.3169E-21	6.6716E-1	3.3284E-1
	2	3.2606E-1	1.8512E-12	6.7394E-1
129		0	1	2
	0	6.6668E-1	3.3332E-1	4.2214E-34
	1	1.0620E-19	6.6741E-1	3.3259E-1
	2	3.1887E-1	4.3217E-11	6.8113E-1

Table A.28: *Emission probability values for the segmental k-means trained HMM.*

Length	*Opdf*			
		0	1	2
2	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	3.3333E-1	3.3333E-1	3.3333E-1
		0	1	2
3	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
		0	1	2
4	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	0.0000E0	1.0000E0	0.0000E0
		0	1	2
5	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	1.0000E0	0.0000E0
		0	1	2
8	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
9	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
16	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
17	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
32	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
33	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
64	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
65	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
128	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
129	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0

Table A.29: *Emission probability values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length	*Opdf*			
		0	1	2
2	0	?	0.0000E0	0.0000E0
	1	?	0.0000E0	0.0000E0
	2	?	0.0000E0	0.0000E0
		0	1	2
3	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
		0	1	2
4	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	0.0000E0	1.0000E0	0.0000E0
		0	1	2
5	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	1.0000E0	0.0000E0
		0	1	2
8	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
9	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
16	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
17	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
32	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
33	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
64	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
65	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
128	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0
		0	1	2
129	0	1.0000E0	0.0000E0	0.0000E0
	1	0.0000E0	1.0000E0	0.0000E0
	2	0.0000E0	0.0000E0	1.0000E0

Table A.30: *Emission probability values for BW-Uniform-trained.*

Length	*Opdf*			
2		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4		0	1	2
	0	7.5000E-1	2.5000E-1	0.0000E0
	1	7.5000E-1	2.5000E-1	0.0000E0
	2	7.5000E-1	2.5000E-1	0.0000E0
5		0	1	2
	0	6.0000E-1	4.0000E-1	0.0000E0
	1	6.0000E-1	4.0000E-1	0.0000E0
	2	6.0000E-1	4.0000E-1	0.0000E0
8		0	1	2
	0	3.7500E-1	3.7500E-1	2.5000E-1
	1	3.7500E-1	3.7500E-1	2.5000E-1
	2	3.7500E-1	3.7500E-1	2.5000E-1
9		0	1	2
	0	3.3333E-1	3.3333E-1	3.3333E-1
	1	3.3333E-1	3.3333E-1	3.3333E-1
	2	3.3333E-1	3.3333E-1	3.3333E-1
16		0	1	2
	0	3.7500E-1	3.7500E-1	2.5000E-1
	1	3.7500E-1	3.7500E-1	2.5000E-1
	2	3.7500E-1	3.7500E-1	2.5000E-1
17		0	1	2
	0	3.5294E-1	3.5294E-1	2.9412E-1
	1	3.5294E-1	3.5294E-1	2.9412E-1
	2	3.5294E-1	3.5294E-1	2.9412E-1
32		0	1	2
	0	3.7500E-1	3.4375E-1	2.8125E-1
	1	3.7500E-1	3.4375E-1	2.8125E-1
	2	3.7500E-1	3.4375E-1	2.8125E-1
33		0	1	2
	0	3.6364E-1	3.6364E-1	2.7273E-1
	1	3.6364E-1	3.6364E-1	2.7273E-1
	2	3.6364E-1	3.6364E-1	2.7273E-1
64		0	1	2
	0	3.4375E-1	3.2813E-1	3.2813E-1
	1	3.4375E-1	3.2813E-1	3.2813E-1
	2	3.4375E-1	3.2813E-1	3.2813E-1
65		0	1	2
	0	3.5385E-1	3.2308E-1	3.2308E-1
	1	3.5385E-1	3.2308E-1	3.2308E-1
	2	3.5385E-1	3.2308E-1	3.2308E-1
128		0	1	2
	0	3.4375E-1	3.2812E-1	3.2812E-1
	1	3.4375E-1	3.2812E-1	3.2812E-1
	2	3.4375E-1	3.2812E-1	3.2812E-1
129		0	1	2
	0	3.4884E-1	3.2558E-1	3.2558E-1
	1	3.4884E-1	3.2558E-1	3.2558E-1
	2	3.4884E-1	3.2558E-1	3.2558E-1

Table A.31: *Emission probability values for BW-Random-0-trained.*

Length		0	1	2
2	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4	0	1.0000E0	8.2578E-17	0.0000E0
	1	1.0000E0	3.6974E-6	0.0000E0
	2	2.4060E-1	7.5940E-1	0.0000E0
5	0	1.0000E0	7.0770E-49	0.0000E0
	1	1.0000E0	3.9379E-6	0.0000E0
	2	1.5087E-5	9.9998E-1	0.0000E0
8	0	1.0000E0	7.6200E-10	0.0000E0
	1	3.1190E-8	1.0000E0	7.4017E-15
	2	2.2344E-98	1.9968E-2	9.8003E-1
9	0	1.0000E0	1.3159E-9	0.0000E0
	1	4.7584E-9	1.0000E0	5.7012E-14
	2	9.5863E-166	3.6924E-5	9.9996E-1
16	0	1.0000E0	1.3703E-9	7.4640E-14
	1	1.1104E-8	1.0000E0	2.2540E-9
	2	2.1579E-13	1.0970E-4	9.9989E-1
17	0	1.0000E0	1.5914E-9	3.0300E-13
	1	6.3574E-9	1.0000E0	5.5646E-12
	2	3.2525E-14	1.2741E-5	9.9999E-1
32	0	1.0000E0	1.0162E-9	1.8932E-11
	1	3.8867E-8	1.0000E0	4.0403E-12
	2	4.3669E-11	5.4202E-7	1.0000E0
33	0	1.0000E0	7.0758E-10	2.5692E-11
	1	2.0852E-8	1.0000E0	2.0290E-12
	2	3.3846E-11	4.3683E-7	1.0000E0
64	0	1.0000E0	1.0451E-9	1.0730E-9
	1	5.1381E-9	1.0000E0	1.0802E-10
	2	9.6373E-10	1.0829E-7	1.0000E0
65	0	1.0000E0	6.7289E-10	6.9792E-10
	1	1.0665E-8	1.0000E0	9.3998E-11
	2	3.6594E-11	1.0825E-7	1.0000E0
128	0	1.0000E0	1.2579E-9	1.2976E-9
	1	6.9313E-9	1.0000E0	9.9074E-11
	2	4.2428E-11	1.0578E-7	1.0000E0
129	0	1.0000E0	1.0105E-9	1.0385E-9
	1	5.8181E-9	1.0000E0	9.7225E-11
	2	3.4865E-11	1.0622E-7	1.0000E0

Table A.32: *Emission probability values for BW-Random-1-trained.*

Length	*Opdf*			
2		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4		0	1	2
	0	2.4440E-1	7.5560E-1	0.0000E0
	1	1.0000E0	1.6712E-12	0.0000E0
	2	9.9984E-1	1.6238E-4	0.0000E0
5		0	1	2
	0	4.2042E-4	9.9958E-1	0.0000E0
	1	1.0000E0	2.8502E-25	0.0000E0
	2	9.9921E-1	7.8517E-4	0.0000E0
8		0	1	2
	0	4.5281E-75	4.7703E-2	9.5230E-1
	1	1.0000E0	8.7782E-8	5.0724E-217
	2	9.9316E-8	1.0000E0	6.4544E-10
9		0	1	2
	0	2.1359E-124	3.4759E-4	9.9965E-1
	1	1.0000E0	3.7425E-7	1.4023E-197
	2	4.3532E-9	1.0000E0	1.1808E-9
16		0	1	2
	0	1.4028E-10	6.1782E-4	9.9938E-1
	1	1.0000E0	5.1338E-7	2.3066E-12
	2	2.5854E-8	1.0000E0	2.0306E-6
17		0	1	2
	0	2.9669E-11	8.4636E-5	9.9992E-1
	1	1.0000E0	8.4594E-7	3.4449E-12
	2	7.3301E-9	1.0000E0	1.8815E-8
32		0	1	2
	0	7.1751E-8	3.8995E-6	1.0000E0
	1	1.0000E0	1.7237E-6	7.6396E-11
	2	3.2141E-8	1.0000E0	2.4605E-8
33		0	1	2
	0	7.8246E-8	4.1361E-6	1.0000E0
	1	1.0000E0	1.6473E-6	8.3500E-11
	2	1.7924E-8	1.0000E0	1.7832E-8
64		0	1	2
	0	1.2643E-6	7.5033E-7	1.0000E0
	1	1.0000E0	3.6852E-6	2.1092E-9
	2	2.7291E-9	1.0000E0	4.9223E-7
65		0	1	2
	0	1.7378E-7	8.0929E-7	1.0000E0
	1	1.0000E0	2.5419E-6	1.7423E-9
	2	7.2802E-9	1.0000E0	4.6278E-7
128		0	1	2
	0	3.3741E-7	6.3408E-7	1.0000E0
	1	9.9999E-1	5.9304E-6	2.7887E-9
	2	2.8943E-9	1.0000E0	7.7468E-7
129		0	1	2
	0	2.6862E-7	6.7343E-7	1.0000E0
	1	1.0000E0	4.8510E-6	2.4121E-9
	2	2.9650E-9	1.0000E0	7.2832E-7

Table A.33: *Emission probability values for BW-Random-2-trained.*

Length	*Opdf*			
2		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4		0	1	2
	0	3.2461E-1	6.7539E-1	0.0000E0
	1	7.4374E-1	2.5626E-1	0.0000E0
	2	1.0000E0	1.9204E-9	0.0000E0
5		0	1	2
	0	1.9583E-1	8.0417E-1	0.0000E0
	1	2.8779E-8	1.0000E0	0.0000E0
	2	1.0000E0	2.9940E-14	0.0000E0
8		0	1	2
	0	1.7422E-8	7.4130E-1	2.5870E-1
	1	1.3261E-33	4.5975E-1	5.4025E-1
	2	1.0000E0	1.9511E-10	2.3278E-238
9		0	1	2
	0	1.8542E-14	6.6660E-1	3.3340E-1
	1	1.1932E-62	3.3339E-1	6.6661E-1
	2	1.0000E0	2.1999E-13	8.6712E-286
16		0	1	2
	0	3.9002E-22	8.0000E-1	2.0000E-1
	1	3.8357E-100	4.0000E-1	6.0000E-1
	2	1.0000E0	7.5022E-21	2.2885E-94
17		0	1	2
	0	2.0749E-13	6.6672E-1	3.3328E-1
	1	9.4106E-34	3.9998E-1	6.0002E-1
	2	1.0000E0	1.2560E-11	7.8842E-13
32		0	1	2
	0	3.7876E-15	7.0000E-1	3.0000E-1
	1	1.7899E-64	4.0000E-1	6.0000E-1
	2	1.0000E0	3.8998E-14	1.3168E-60
33		0	1	2
	0	1.6232E-16	7.2727E-1	2.7273E-1
	1	1.8059E-67	4.0000E-1	6.0000E-1
	2	1.0000E0	2.4300E-13	3.8188E-63
64		0	1	2
	0	4.0376E-15	6.6667E-1	3.3333E-1
	1	2.2257E-90	3.3333E-1	6.6667E-1
	2	1.0000E0	7.6757E-96	5.0863E-89
65		0	1	2
	0	2.9925E-22	6.6667E-1	3.3333E-1
	1	1.4861E-125	3.3333E-1	6.6667E-1
	2	1.0000E0	4.8635E-131	4.2251E-124
128		0	1	2
	0	1.5563E-24	6.6667E-1	3.3333E-1
	1	1.2361E-129	3.3333E-1	6.6667E-1
	2	1.0000E0	1.5892E-133	1.2794E-126
129		0	1	2
	0	6.4293E-25	6.6667E-1	3.3333E-1
	1	2.5689E-133	3.3333E-1	6.6667E-1
	2	1.0000E0	6.1413E-138	6.4134E-131

Table A.34: *Emission probability values for BW-Random-3-trained.*

Length		*Opdf*		
2		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4		0	1	2
	0	2.4209E-1	7.5791E-1	0.0000E0
	1	1.0000E0	7.7263E-7	0.0000E0
	2	1.0000E0	6.5821E-22	0.0000E0
5		0	1	2
	0	1.4756E-4	9.9985E-1	0.0000E0
	1	1.0000E0	4.3369E-8	0.0000E0
	2	1.0000E0	1.3535E-75	0.0000E0
8		0	1	2
	0	4.6088E-45	7.2859E-2	9.2714E-1
	1	4.1052E-5	9.9996E-1	2.8434E-17
	2	1.0000E0	8.8579E-13	5.6795E-234
9		0	1	2
	0	4.0321E-116	1.4836E-4	9.9985E-1
	1	3.3210E-7	1.0000E0	2.6781E-13
	2	1.0000E0	1.8738E-11	0.0000E0
16		0	1	2
	0	7.2408E-12	4.3067E-3	9.9569E-1
	1	2.4313E-5	9.9998E-1	2.8974E-11
	2	1.0000E0	3.2388E-12	2.0984E-10
17		0	1	2
	0	9.3725E-13	2.1529E-4	9.9978E-1
	1	2.4803E-6	1.0000E0	1.1615E-11
	2	1.0000E0	1.4119E-11	7.4701E-11
32		0	1	2
	0	5.0332E-10	5.9772E-5	9.9994E-1
	1	3.7026E-5	9.9996E-1	6.5241E-12
	2	1.0000E0	7.3557E-12	3.8633E-8
33		0	1	2
	0	3.1531E-10	2.0661E-4	9.9979E-1
	1	7.6719E-5	9.9992E-1	6.2683E-13
	2	1.0000E0	4.1546E-12	2.3758E-7
64		0	1	2
	0	5.0770E-10	9.9900E-1	1.0034E-3
	1	9.9984E-1	1.5514E-4	2.1078E-10
	2	3.0620E-3	6.2423E-14	9.9694E-1
65		0	1	2
	0	6.8181E-11	7.9823E-1	2.0177E-1
	1	9.5788E-1	4.2117E-2	1.0776E-12
	2	2.2577E-1	9.2974E-17	7.7423E-1
128		0	1	2
	0	4.2551E-10	9.9997E-1	3.0242E-5
	1	1.0000E0	4.1450E-6	1.3247E-9
	2	2.5419E-5	1.1481E-12	9.9997E-1
129		0	1	2
	0	3.8975E-10	9.9995E-1	4.7405E-5
	1	1.0000E0	4.9260E-6	8.8535E-10
	2	3.4444E-5	6.2073E-13	9.9997E-1

Table A.35: *Emission probability values for BW-Random-4-trained.*

Length	*Opdf*			
		0	1	2
2	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
		0	1	2
3	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
		0	1	2
4	0	1.0000E0	9.5761E-9	0.0000E0
	1	9.9995E-1	5.0525E-5	0.0000E0
	2	4.8252E-1	5.1748E-1	0.0000E0
		0	1	2
5	0	1.0000E0	3.0869E-7	0.0000E0
	1	1.0000E0	2.8206E-28	0.0000E0
	2	1.1565E-2	9.8844E-1	0.0000E0
		0	1	2
8	0	9.9358E-1	6.4156E-3	6.0387E-63
	1	1.0000E0	1.0273E-18	3.9007E-104
	2	1.8452E-8	5.9930E-1	4.0070E-1
		0	1	2
9	0	1.0000E0	2.5512E-18	1.6967E-90
	1	1.0000E0	8.1981E-20	3.6931E-106
	2	5.4193E-22	5.0000E-1	5.0000E-1
		0	1	2
16	0	5.1582E-1	2.0314E-9	4.8418E-1
	1	9.9998E-1	1.3672E-7	1.5673E-5
	2	1.1311E-6	6.4744E-1	3.5256E-1
		0	1	2
17	0	5.8497E-1	1.8717E-9	4.1503E-1
	1	9.4199E-1	6.8250E-10	5.8011E-2
	2	6.7392E-8	6.4285E-1	3.5715E-1
		0	1	2
32	0	4.0000E-1	2.2703E-11	6.0000E-1
	1	7.2727E-1	1.1201E-34	2.7273E-1
	2	7.4588E-33	1.0000E0	2.4897E-34
		0	1	2
33	0	4.0000E-1	2.9567E-12	6.0000E-1
	1	7.2727E-1	3.0383E-39	2.7273E-1
	2	1.6113E-37	1.0000E0	3.5939E-39
		0	1	2
64	0	3.3333E-1	1.0326E-14	6.6667E-1
	1	6.8182E-1	3.1362E-34	3.1818E-1
	2	1.8086E-30	1.0000E0	1.2772E-12
		0	1	2
65	0	3.6363E-1	3.5690E-12	6.3637E-1
	1	6.8182E-1	5.9628E-23	3.1818E-1
	2	4.8548E-10	1.0000E0	8.5593E-10
		0	1	2
128	0	3.4884E-1	3.3273E-13	6.5116E-1
	1	6.7442E-1	9.4803E-25	3.2558E-1
	2	5.0570E-11	1.0000E0	2.5297E-11
		0	1	2
129	0	3.4884E-1	1.5022E-15	6.5116E-1
	1	6.8182E-1	1.3775E-29	3.1818E-1
	2	1.3389E-14	1.0000E0	1.2958E-13

Table A.36: *Emission probability values for BW-Random-5-trained.*

Length	*Opdf*			
2		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
3		0	1	2
	0	1.0000E0	0.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
4		0	1	2
	0	3.7590E-1	6.2410E-1	0.0000E0
	1	8.3733E-1	1.6267E-1	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
5		0	1	2
	0	0.0000E0	1.0000E0	0.0000E0
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
8		0	1	2
	0	0.0000E0	6.0000E-1	4.0000E-1
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
9		0	1	2
	0	0.0000E0	5.0000E-1	5.0000E-1
	1	1.0000E0	0.0000E0	0.0000E0
	2	1.0000E0	0.0000E0	0.0000E0
16		0	1	2
	0	1.8259E-9	6.8050E-1	3.1950E-1
	1	6.3263E-1	1.7059E-21	3.6737E-1
	2	9.9998E-1	9.0085E-11	2.3154E-5
17		0	1	2
	0	1.0571E-8	6.0643E-1	3.9357E-1
	1	6.5209E-1	2.0824E-22	3.4791E-1
	2	1.0000E0	3.2498E-9	2.7215E-6
32		0	1	2
	0	1.0314E-6	9.9981E-1	1.9115E-4
	1	3.7835E-1	5.2457E-11	6.2165E-1
	2	9.9996E-1	2.0601E-9	4.4864E-5
33		0	1	2
	0	4.2438E-7	9.9995E-1	5.0164E-5
	1	3.7840E-1	2.5232E-11	6.2160E-1
	2	9.9996E-1	9.6024E-10	3.5659E-5
64		0	1	2
	0	2.8491E-6	9.9997E-1	3.0704E-5
	1	5.5561E-3	4.7353E-10	9.9444E-1
	2	1.0000E0	4.6721E-9	1.6807E-6
65		0	1	2
	0	4.0419E-6	9.9995E-1	5.0329E-5
	1	4.1866E-3	3.9093E-10	9.9581E-1
	2	1.0000E0	5.9427E-9	2.6210E-6
128		0	1	2
	0	3.2709E-6	9.9997E-1	2.3427E-5
	1	1.4983E-3	7.5760E-10	9.9850E-1
	2	1.0000E0	1.4185E-8	1.8116E-6
129		0	1	2
	0	3.2450E-6	9.9996E-1	3.4961E-5
	1	2.2619E-3	5.5172E-10	9.9774E-1
	2	1.0000E0	1.3664E-8	2.2639E-6

Table A.37: *coord*1 *values for the segmental k-means trained HMM.*

Length	*Coordinates*		
2	x1	x2	x3
	2.6133E-1	5.0004E-1	2.6575E-1
3	x1	x2	x3
	1.7422E-1	3.3335E-1	5.0000E-1
4	x1	x2	x3
	3.0139E-1	2.8583E-1	4.2861E-1
5	x1	x2	x3
	3.9288E-1	2.5013E-1	3.7512E-1
8	x1	x2	x3
	1.5000E0	1.5000E0	1.0000E0
9	x1	x2	x3
	1.5000E0	1.5000E0	1.0000E0
16	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
17	x1	x2	x3
	3.0070E-1	3.0088E-1	4.0087E-1
32	x1	x2	x3
	3.2219E-1	3.5798E-1	3.2213E-1
33	x1	x2	x3
	3.1118E-1	3.8017E-1	3.1105E-1
64	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
65	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1
128	x1	x2	x3
	3.3940E-1	3.3137E-1	3.3139E-1
129	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1

Table A.38: *coord1 values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length	*Coordinates*		
2	x1	x2	x3
	?	?	?
3	x1	x2	x3
	1.7422E-1	3.3335E-1	5.0000E-1
4	x1	x2	x3
	3.0139E-1	2.8583E-1	4.2861E-1
5	x1	x2	x3
	3.9288E-1	2.5013E-1	3.7512E-1
8	x1	x2	x3
	1.5000E0	1.5000E0	1.0000E0
9	x1	x2	x3
	1.5000E0	1.5000E0	1.0000E0
16	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
17	x1	x2	x3
	3.0070E-1	3.0088E-1	4.0087E-1
32	x1	x2	x3
	3.2219E-1	3.5798E-1	3.2213E-1
33	x1	x2	x3
	3.1118E-1	3.8017E-1	3.1105E-1
64	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
65	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1
128	x1	x2	x3
	3.3940E-1	3.3137E-1	3.3139E-1
129	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1

Table A.39: *coord*1 *values for BW-Uniform-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
3	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
4	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
5	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
8	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
9	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
16	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
17	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
32	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
33	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
64	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
65	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
128	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1
129	x1	x2	x3
	3.5234E-1	3.5234E-1	3.5234E-1

Table A.40: *coord1 values for BW-Random-0-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	8.5159E-1	1.3655E-1	3.0244E-1
3	x1	x2	x3
	8.9684E-1	9.5924E-2	2.7408E-1
4	x1	x2	x3
	3.8749E0	3.9821E0	1.0000E0
5	x1	x2	x3
	2.9542E0	3.0481E0	1.0000E0
8	x1	x2	x3
	1.5000E0	1.5104E0	1.0000E0
9	x1	x2	x3
	1.5000E0	1.5000E0	1.0000E0
16	x1	x2	x3
	3.3413E-1	3.3411E-1	3.3418E-1
17	x1	x2	x3
	3.0070E-1	3.0088E-1	4.0088E-1
32	x1	x2	x3
	3.2219E-1	3.5798E-1	3.2213E-1
33	x1	x2	x3
	3.1118E-1	3.8017E-1	3.1105E-1
64	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
65	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1
128	x1	x2	x3
	3.3940E-1	3.3137E-1	3.3139E-1
129	x1	x2	x3
	3.4457E-1	3.2881E-1	3.2885E-1

Table A.41: *coord*1 *values for BW-Random-1-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	4.2393E-1	3.4530E-1	2.5978E-1
3	x1	x2	x3
	4.0553E-1	3.7801E-1	2.4481E-1
4	x1	x2	x3
	1.0000E0	3.6108E0	4.4000E0
5	x1	x2	x3
	1.0000E0	2.6970E0	3.4327E0
8	x1	x2	x3
	1.0000E0	1.5000E0	1.5264E0
9	x1	x2	x3
	1.0000E0	1.5000E0	1.5003E0
16	x1	x2	x3
	3.3437E-1	3.3410E-1	3.3396E-1
17	x1	x2	x3
	4.0091E-1	3.0070E-1	3.0086E-1
32	x1	x2	x3
	3.2213E-1	3.2219E-1	3.5798E-1
33	x1	x2	x3
	3.1105E-1	3.1118E-1	3.8017E-1
64	x1	x2	x3
	3.3414E-1	3.3414E-1	3.3414E-1
65	x1	x2	x3
	3.2885E-1	3.4457E-1	3.2881E-1
128	x1	x2	x3
	3.3139E-1	3.3940E-1	3.3137E-1
129	x1	x2	x3
	3.2885E-1	3.4457E-1	3.2881E-1

Table A.42: *coord*1 *values for BW-Random-2-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	1.1693E-1	1.2855E-1	7.8914E-1
3	x1	x2	x3
	1.1544E-1	1.3092E-1	7.9034E-1
4	x1	x2	x3
	5.5870E-1	4.4133E-1	1.9865E0
5	x1	x2	x3
	4.3295E-1	5.6706E-1	1.5674E0
8	x1	x2	x3
	4.2508E-1	5.7494E-1	1.5000E0
9	x1	x2	x3
	4.9992E-1	5.0008E-1	1.5000E0
16	x1	x2	x3
	3.6368E-1	3.6372E-1	2.7278E-1
17	x1	x2	x3
	3.8463E-1	3.8474E-1	2.3105E-1
32	x1	x2	x3
	3.3336E-1	3.3343E-1	3.3339E-1
33	x1	x2	x3
	3.4486E-1	3.4492E-1	3.1039E-1
64	x1	x2	x3
	3.3336E-1	3.3343E-1	3.3339E-1
65	x1	x2	x3
	3.2815E-1	3.2820E-1	3.4381E-1
128	x1	x2	x3
	3.3074E-1	3.3079E-1	3.3864E-1
129	x1	x2	x3
	3.2815E-1	3.2820E-1	3.4381E-1

Table A.43: *coord1 values for BW-Random-3-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	2.9215E-1	5.8889E-1	3.8167E-1
3	x1	x2	x3
	3.3982E-1	5.5919E-1	4.1703E-1
4	x1	x2	x3
	1.0000E0	3.9398E0	3.9375E0
5	x1	x2	x3
	1.0000E0	3.0038E0	2.9974E0
8	x1	x2	x3
	1.0000E0	1.5426E0	1.5000E0
9	x1	x2	x3
	1.0000E0	1.5001E0	1.5000E0
16	x1	x2	x3
	3.3573E-1	3.3270E-1	3.3369E-1
17	x1	x2	x3
	4.0096E-1	3.0082E-1	3.0068E-1
32	x1	x2	x3
	3.2215E-1	3.5797E-1	3.2218E-1
33	x1	x2	x3
	3.1111E-1	3.8012E-1	3.1116E-1
64	x1	x2	x3
	3.3408E-1	3.3296E-1	3.3508E-1
65	x1	x2	x3
	4.1019E-1	2.9710E-1	2.9450E-1
128	x1	x2	x3
	3.3138E-1	3.3939E-1	3.3139E-1
129	x1	x2	x3
	3.2882E-1	3.4456E-1	3.2884E-1

Table A.44: *coord1 values for BW-Random-4-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	4.5757E-1	3.5464E-1	1.9955E-1
3	x1	x2	x3
	4.3859E-1	3.7138E-1	2.0667E-1
4	x1	x2	x3
	3.5654E-1	1.7757E0	1.0000E0
5	x1	x2	x3
	3.3913E0	2.7891E0	1.0000E0
8	x1	x2	x3
	3.7824E0	2.5359E0	1.0000E0
9	x1	x2	x3
	1.0000E0	1.0000E0	1.0000E0
16	x1	x2	x3
	1.0610E-1	2.3761E-1	6.6639E-1
17	x1	x2	x3
	1.8830E-1	1.8865E-1	6.2441E-1
32	x1	x2	x3
	3.1135E-1	3.1140E-1	3.7743E-1
33	x1	x2	x3
	3.0003E-1	3.0011E-1	4.0007E-1
64	x1	x2	x3
	3.3336E-1	3.3343E-1	3.3339E-1
65	x1	x2	x3
	3.3849E-1	3.3855E-1	3.2313E-1
128	x1	x2	x3
	3.3597E-1	3.3603E-1	3.2818E-1
129	x1	x2	x3
	3.3597E-1	3.3603E-1	3.2818E-1

Table A.45: *coord1 values for BW-Random-5-trained.*

Length	*Coordinates*		
2	x1	x2	x3
	2.0994E-1	4.6623E-1	3.5554E-1
3	x1	x2	x3
	2.1028E-1	4.8866E-1	3.3302E-1
4	x1	x2	x3
	3.6554E-1	1.3449E-1	5.0000E-1
5	x1	x2	x3
	1.0000E0	1.0000E0	1.0000E0
8	x1	x2	x3
	1.0000E0	1.0000E0	1.0000E0
9	x1	x2	x3
	1.0000E0	1.0000E0	1.0000E0
16	x1	x2	x3
	6.3171E-1	1.8785E-1	1.8247E-1
17	x1	x2	x3
	6.7008E-1	1.6993E-1	1.6278E-1
32	x1	x2	x3
	3.9768E-1	4.6340E-1	1.7656E-1
33	x1	x2	x3
	4.2142E-1	4.4757E-1	1.7034E-1
64	x1	x2	x3
	3.3358E-1	3.3644E-1	3.3209E-1
65	x1	x2	x3
	3.2882E-1	3.3022E-1	3.4321E-1
128	x1	x2	x3
	3.3137E-1	3.3188E-1	3.3891E-1
129	x1	x2	x3
	3.2882E-1	3.2958E-1	3.4384E-1

Table A.46: *coord2 values for the segmental k-means trained HMM.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	1.0000E0	0.0000E0	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
3	x1	x2	x3	x4	x5	x6
	1.0000E0	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
4	x1	x2	x3	x4	x5	x6
	5.0000E-1	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
5	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
8	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	0.0000E0	0.0000E0
9	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	0.0000E0	0.0000E0
16	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
17	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	2.5000E-1	0.0000E0
32	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.0000E-1	3.3333E-1	0.0000E0
33	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	2.7273E-1	3.3333E-1	0.0000E0
64	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
65	x1	x2	x3	x4	x5	x6
	3.1818E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
128	x1	x2	x3	x4	x5	x6
	3.2558E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
129	x1	x2	x3	x4	x5	x6
	3.1818E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0

Table A.47: *coord2 values for the Baum-Welch trained HMM initialized using segmental k-means.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	?	?	?	?	?	?
3	x1	x2	x3	x4	x5	x6
	1.0000E0	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
4	x1	x2	x3	x4	x5	x6
	5.0000E-1	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
5	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	1.0000E0	3.3333E-1	3.3333E-1
8	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	0.0000E0	0.0000E0
9	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	0.0000E0	0.0000E0
16	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
17	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	2.5000E-1	0.0000E0
32	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.0000E-1	3.3333E-1	0.0000E0
33	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	2.7273E-1	3.3333E-1	0.0000E0
64	x1	x2	x3	x4	x5	x6
	3.3333E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
65	x1	x2	x3	x4	x5	x6
	3.1818E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
128	x1	x2	x3	x4	x5	x6
	3.2558E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0
129	x1	x2	x3	x4	x5	x6
	3.1818E-1	0.0000E0	0.0000E0	3.3333E-1	3.3333E-1	0.0000E0

Table A.48: *coord2 values for BW-Uniform-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
3	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
4	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
5	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
8	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
9	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
16	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
17	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
32	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
33	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
64	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
65	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
128	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1
129	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1	3.3333E-1

Table A.49: *coord2 values for BW-Random-0-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	1.3834E-1	1.5231E-3	8.1480E-1	7.4487E-2	6.2186E-1	2.3680E-1
3	x1	x2	x3	x4	x5	x6
	9.3632E-2	9.1041E-4	8.2660E-1	6.9067E-2	7.1240E-1	1.8675E-1
4	x1	x2	x3	x4	x5	x6
	7.4193E-1	1.0279E-7	7.7417E-35	7.4887E-1	1.5151E-84	5.7475E-33
5	x1	x2	x3	x4	x5	x6
	6.6150E-1	7.0779E-8	2.2548E-37	6.7193E-1	8.6202E-138	7.4740E-33
8	x1	x2	x3	x4	x5	x6
	3.3333E-1	2.3067E-32	6.1576E-69	3.3792E-1	9.5522E-166	1.1373E-18
9	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.7900E-86	1.1818E-78	3.3335E-1	7.5790E-194	2.2824E-40
16	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.0788E-75	5.3272E-73	3.3336E-1	3.3328E-1	1.9831E-32
17	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.9793E-95	5.3814E-76	3.3334E-1	2.5000E-1	8.9219E-42
32	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.3352E-88	3.0912E-70	3.0000E-1	3.3333E-1	1.0196E-54
33	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.6336E-90	3.8822E-72	2.7273E-1	3.3333E-1	6.0485E-53
64	x1	x2	x3	x4	x5	x6
	3.3333E-1	8.8479E-82	7.6430E-77	3.3333E-1	3.3333E-1	7.3548E-59
65	x1	x2	x3	x4	x5	x6
	3.1818E-1	1.0097E-84	1.3703E-75	3.3333E-1	3.3333E-1	3.9889E-59
128	x1	x2	x3	x4	x5	x6
	3.2558E-1	1.1210E-82	2.2836E-76	3.3333E-1	3.3333E-1	2.8927E-59
129	x1	x2	x3	x4	x5	x6
	3.1818E-1	1.5446E-83	7.3191E-77	3.3333E-1	3.3333E-1	3.0474E-59

Table A.50: *coord2 values for BW-Random-1-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	3.5683E-1	2.4990E-1	9.0892E-2	4.0921E-1	8.9439E-1	4.8244E-2
3	x1	x2	x3	x4	x5	x6
	3.7785E-1	2.2922E-1	9.1361E-2	3.6374E-1	8.9510E-1	4.9546E-2
4	x1	x2	x3	x4	x5	x6
	4.5518E-35	2.3415E-13	1.3844E-4	7.2292E-1	7.7273E-1	1.2107E-17
5	x1	x2	x3	x4	x5	x6
	1.0902E-69	7.1867E-13	3.3728E-5	6.2918E-1	7.0868E-1	1.1019E-18
8	x1	x2	x3	x4	x5	x6
	1.1777E-115	5.5678E-12	3.1170E-20	3.3333E-1	3.4485E-1	5.6461E-57
9	x1	x2	x3	x4	x5	x6
	5.5556E-131	2.0143E-25	2.8004E-52	3.3333E-1	3.3345E-1	1.6396E-52
16	x1	x2	x3	x4	x5	x6
	3.3306E-1	8.5525E-22	1.2171E-49	3.3333E-1	3.3347E-1	1.6078E-48
17	x1	x2	x3	x4	x5	x6
	2.4997E-1	3.6306E-29	5.9895E-62	3.3333E-1	3.3336E-1	2.6741E-46
32	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.9669E-40	1.4990E-53	3.3333E-1	3.0000E-1	2.4227E-39
33	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.9915E-38	1.3508E-52	3.3333E-1	2.7273E-1	5.5349E-38
64	x1	x2	x3	x4	x5	x6
	3.3333E-1	3.7419E-42	4.5635E-45	3.3333E-1	3.3333E-1	3.1797E-37
65	x1	x2	x3	x4	x5	x6
	3.3333E-1	4.8317E-42	3.7848E-49	3.1818E-1	3.3333E-1	1.7396E-38
128	x1	x2	x3	x4	x5	x6
	3.3333E-1	2.0756E-41	2.9172E-46	3.2558E-1	3.3334E-1	2.0988E-35
129	x1	x2	x3	x4	x5	x6
	3.3333E-1	1.9658E-41	5.6341E-47	3.1818E-1	3.3334E-1	4.3131E-36

Table A.51: *coord2 values for BW-Random-2-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	5.3199E-1	4.4569E-1	8.2448E-2	8.9716E-1	1.0707E-1	7.8300E-2
3	x1	x2	x3	x4	x5	x6
	5.3274E-1	4.4502E-1	7.9910E-2	8.9852E-1	1.0063E-1	8.6109E-2
4	x1	x2	x3	x4	x5	x6
	7.7255E-1	9.4147E-9	9.7801E-1	6.6704E-9	2.4444E-1	2.5216E-1
5	x1	x2	x3	x4	x5	x6
	9.3592E-1	2.3759E-19	7.1457E-1	9.5117E-72	3.5447E-1	7.5406E-3
8	x1	x2	x3	x4	x5	x6
	9.4689E-1	6.6422E-76	7.0008E-1	1.7566E-93	3.3311E-1	2.1974E-4
9	x1	x2	x3	x4	x5	x6
	9.9997E-1	1.0715E-105	9.9964E-1	2.0936E-112	3.3333E-1	2.5878E-9
16	x1	x2	x3	x4	x5	x6
	1.0000E0	4.0557E-94	7.5000E-1	2.5000E-1	3.3333E-1	1.0474E-18
17	x1	x2	x3	x4	x5	x6
	9.9997E-1	3.4112E-28	7.9985E-1	2.0001E-1	3.3333E-1	7.6712E-9
32	x1	x2	x3	x4	x5	x6
	1.0000E0	1.1054E-60	6.6667E-1	3.3333E-1	3.3333E-1	1.0319E-11
33	x1	x2	x3	x4	x5	x6
	1.0000E0	4.7620E-63	7.0000E-1	3.0000E-1	3.3333E-1	1.9292E-12
64	x1	x2	x3	x4	x5	x6
	1.0000E0	4.7115E-86	6.6667E-1	3.3333E-1	3.3333E-1	9.6972E-292
65	x1	x2	x3	x4	x5	x6
	1.0000E0	7.6883E-121	6.6667E-1	3.3333E-1	3.1818E-1	0.0000E0
128	x1	x2	x3	x4	x5	x6
	1.0000E0	3.2488E-124	6.6667E-1	3.3333E-1	3.2558E-1	0.0000E0
129	x1	x2	x3	x4	x5	x6
	1.0000E0	5.2506E-128	6.6667E-1	3.3333E-1	3.1818E-1	0.0000E0

Table A.52: *coord2 values for BW-Random-3-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	5.7224E-1	3.3617E-1	4.0410E-2	3.3016E-1	2.0790E-2	5.2352E-1
3	x1	x2	x3	x4	x5	x6
	5.6609E-1	3.4807E-1	3.5029E-2	3.5244E-1	1.5114E-2	4.8388E-1
4	x1	x2	x3	x4	x5	x6
	1.9081E-41	8.0767E-103	7.4618E-1	2.1707E-42	6.2135E-9	7.4603E-1
5	x1	x2	x3	x4	x5	x6
	2.3282E-56	7.1214E-169	6.6709E-1	6.5536E-58	4.7986E-10	6.6638E-1
8	x1	x2	x3	x4	x5	x6
	4.8360E-14	3.3040E-159	3.5175E-1	5.6022E-31	2.0842E-16	3.3335E-1
9	x1	x2	x3	x4	x5	x6
	1.0930E-32	1.8582E-181	3.3338E-1	1.6033E-52	3.6321E-64	3.3333E-1
16	x1	x2	x3	x4	x5	x6
	5.9025E-17	3.3142E-1	3.3429E-1	5.7782E-31	1.6468E-30	3.3334E-1
17	x1	x2	x3	x4	x5	x6
	5.6261E-27	2.4993E-1	3.3339E-1	5.8922E-41	4.4039E-55	3.3333E-1
32	x1	x2	x3	x4	x5	x6
	2.0880E-29	3.3332E-1	3.0001E-1	4.8923E-31	4.4086E-45	3.3334E-1
33	x1	x2	x3	x4	x5	x6
	3.0735E-23	3.3327E-1	2.7277E-1	2.3058E-27	1.1046E-37	3.3336E-1
64	x1	x2	x3	x4	x5	x6
	5.5483E-21	3.3305E-1	3.3405E-1	1.8720E-17	6.0348E-18	3.3214E-1
65	x1	x2	x3	x4	x5	x6
	2.9400E-8	2.7672E-1	3.8231E-1	7.0305E-10	9.9523E-9	3.8552E-1
128	x1	x2	x3	x4	x5	x6
	1.3687E-35	3.3332E-1	3.2559E-1	5.1757E-30	7.7652E-30	3.3333E-1
129	x1	x2	x3	x4	x5	x6
	2.6226E-34	3.3332E-1	3.1819E-1	2.3504E-28	4.4740E-28	3.3334E-1

Table A.53: *coord2 values for BW-Random-4-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	7.0043E-1	2.8320E-1	8.8263E-1	4.3658E-2	7.3080E-1	2.4040E-2
3	x1	x2	x3	x4	x5	x6
	7.1977E-1	2.6755E-1	8.1108E-1	6.3215E-2	6.9307E-1	2.9368E-2
4	x1	x2	x3	x4	x5	x6
	3.7748E-1	6.2141E-1	7.2568E-1	2.5993E-1	6.2315E-6	3.3741E-12
5	x1	x2	x3	x4	x5	x6
	7.4391E-13	7.0513E-1	6.4123E-1	2.3019E-4	1.4801E-29	7.4489E-48
8	x1	x2	x3	x4	x5	x6
	1.9268E-9	7.3562E-1	6.0486E-1	7.9798E-4	5.6883E-20	3.3972E-83
9	x1	x2	x3	x4	x5	x6
	1.0000E0	1.8555E-14	5.0000E-1	5.0000E-1	2.5125E-52	2.6188E-55
16	x1	x2	x3	x4	x5	x6
	9.6603E-1	1.9799E-2	5.8604E-2	3.7980E-1	1.3746E-1	3.9184E-6
17	x1	x2	x3	x4	x5	x6
	9.9854E-1	4.7898E-4	5.2121E-1	4.7704E-1	1.4412E-1	5.7099E-5
32	x1	x2	x3	x4	x5	x6
	1.0000E0	1.4675E-30	6.3636E-1	3.6364E-1	3.0000E-1	1.5969E-44
33	x1	x2	x3	x4	x5	x6
	1.0000E0	1.1478E-34	6.3636E-1	3.6364E-1	2.7273E-1	8.3827E-53
64	x1	x2	x3	x4	x5	x6
	1.0000E0	3.9131E-30	6.6667E-1	3.3333E-1	3.3333E-1	1.0601E-12
65	x1	x2	x3	x4	x5	x6
	1.0000E0	9.0326E-19	6.8181E-1	3.1818E-1	3.3333E-1	6.5547E-10
128	x1	x2	x3	x4	x5	x6
	1.0000E0	8.7408E-21	6.7442E-1	3.2558E-1	3.3333E-1	3.7958E-11
129	x1	x2	x3	x4	x5	x6
	1.0000E0	1.3088E-25	6.7442E-1	3.2558E-1	3.3333E-1	1.7646E-13

Table A.54: *coord2 values for BW-Random-5-trained.*

Length	*Coordinates*					
2	x1	x2	x3	x4	x5	x6
	3.2004E-1	5.7737E-1	8.6086E-2	5.1510E-1	3.4994E-1	6.2348E-1
3	x1	x2	x3	x4	x5	x6
	3.8077E-1	4.9800E-1	9.1792E-2	4.8164E-1	3.5162E-1	6.3345E-1
4	x1	x2	x3	x4	x5	x6
	0.0000E0	1.0000E0	0.0000E0	1.0000E0	7.3105E-1	2.6895E-1
5	x1	x2	x3	x4	x5	x6
	0.0000E0	0.0000E0	0.0000E0	1.0000E0	5.0000E-1	5.0000E-1
8	x1	x2	x3	x4	x5	x6
	0.0000E0	0.0000E0	0.0000E0	1.0000E0	5.0000E-1	5.0000E-1
9	x1	x2	x3	x4	x5	x6
	0.0000E0	0.0000E0	0.0000E0	1.0000E0	5.0000E-1	5.0000E-1
16	x1	x2	x3	x4	x5	x6
	1.4594E-1	3.8729E-7	2.7064E-5	9.6996E-1	5.0463E-1	4.9537E-1
17	x1	x2	x3	x4	x5	x6
	1.2382E-1	5.8042E-6	3.7178E-4	9.5546E-1	5.0904E-1	4.9095E-1
32	x1	x2	x3	x4	x5	x6
	2.9994E-1	2.1287E-27	3.2445E-6	3.7467E-1	6.1311E-1	3.7135E-1
33	x1	x2	x3	x4	x5	x6
	2.7271E-1	1.5334E-28	2.2807E-6	3.7311E-1	6.1330E-1	3.6822E-1
64	x1	x2	x3	x4	x5	x6
	3.3332E-1	1.2201E-34	1.9392E-18	3.3058E-1	3.3475E-1	7.0287E-10
65	x1	x2	x3	x4	x5	x6
	3.3332E-1	2.8503E-33	1.2337E-17	3.3195E-1	3.1946E-1	1.7490E-9
128	x1	x2	x3	x4	x5	x6
	3.3332E-1	6.1724E-35	1.3169E-21	3.3284E-1	3.2606E-1	1.8512E-12
129	x1	x2	x3	x4	x5	x6
	3.3332E-1	4.2214E-34	1.0620E-19	3.3259E-1	3.1887E-1	4.3217E-11

Appendix B

Similarity Test Data Tables

The following tables present the similarity test data taken during the tests that generated the data in the tables in Appendix A. For each recorded length, the trained hidden Markov models were compaired pairwise using each of the three similarity metrics mentioned in Section 5.6 and for each similarity the resulting values were averaged. For more information on how the tests were performed, see Appendix A.

Table B.1: *Similarity averages computed using* $Coord1 - Similarity.$

Length	*Average Similarity*
2	?
3	4.0601E-1
4	3.0942E0
5	2.5897E0
8	1.4277E0
9	8.8697E-1
16	1.8974E-1
17	2.2795E-1
32	7.6272E-2
33	9.6770E-2
64	8.3137E-3
65	3.6460E-2
128	1.3865E-2
129	2.0022E-2

Table B.2: *Similarity averages computed using* $Coord2 - Similarity$.

Length	*Average Similarity*
2	?
3	9.5425E-1
4	1.1284E0
5	1.0311E0
8	8.4061E-1
9	8.6067E-1
16	7.1799E-1
17	7.4970E-1
32	6.5192E-1
33	6.6075E-1
64	5.9006E-1
65	5.9992E-1
128	5.9323E-1
129	5.9555E-1

Table B.3: *Similarity averages computed using Kullback – Leibler – Similarity.*

Length	*Average Similarity*
2	?
3	0.0000E0
4	?
5	?
8	?
9	?
16	?
17	?
32	?
33	?
64	?
65	?
128	?
129	?