Aaron Wu

303890304

Homework 5: Star Trek

Uses: startrek.m, new\_A.m, error\_func.m, transform.m

A list of the rules and assumptions is provided at the end of the document.

The names warp drive vibration index, hailer transponder frequencies, surface reflection colors, and ratios of long to short axes were hard coded into the program as the available input. Each element in the vectors of each corresponded to a different ship and the order of the elements between vectors corresponded with each other. Therefore, each index of all the vectors identified a single ship. These elements were transformed into vectors themselves, governed by a set of rules. These rules determined whether the specific characteristic of this ship fell into the range provided by the rules.

Example (names):

RULE: ships starting with 2 consonants = klingon

RULE: ships starting with A or E = anterean

RULE: ships with the second element as a number = anterean

RULE: ships that are not klingon or anterean are neither

RULE: ships with no name have an equal chance of being klingon, anterean, or neither

These rules were implemented using a graded system of likelihood.

The first thing to notice is that there are only 3 possible outcomes in this situation: Klingon, Anterean, or neither. This produces a vector of likelihoods governed by the rules.

In terms of names, either the name follows the specified rule or doesn’t follow the specified rule. For most of the rules, if the requirement has been met, then the property is true for that rule and is represented by a 1. Likewise, if the requirement has not been met, then the property is not true for that rule and is represented by a 0. However, if there is uncertainty between any of the possible outcomes, then the likelihood is represented as a distribution between the possible outcomes.

For example, the name “Grotz” would follow the first rule, but fails the next four and fulfills the outcome of Klingon. Therefore, the vector of likelihood that this property belongs to a Klingon ship is [1 0 0]. However, if the name was unavailable, then this would follow the fifth rule and the likelihood is evenly distributed among the 3 outcomes is [.3 .3 .3].

It is imperative to note that the number of rules doesn’t entail the number of outcomes; it only influences it. Many rules are actually created to interpret ambiguous information such as “light color” and “dark color” and aren’t a property of their own.

The creation of graded vectors of likelihood is used in order to best avoid conflict between different races on a specified property. In other words, it is to maximize linear independence on a given property in order to maintain a stable linear association. Each property delineates different dimensions towards a different outcome. Though the outcomes themselves might be able to be put on a unidimensional scale, there are far more properties that are unable to be scaled in such as manner (such as name). Furthermore, the graded vector provides a shared scale that would not cause chaos within the error matrix. If we had used the scales given, the hailer transponder frequencies (in the thousands) would diminish the importance of the other properties such as ratio of long to short axes (in the single digits).

A vector of all the input values was created for each ship by concatenating all the graded likelihood vectors. The expected output vectors of each race the ships belonged to was also generated with graded likelihood vectors. There is no uncertainty within these ships and therefore each vector is purely 1’s and 0’s, based on whether the ships are a certain race or not. Since there are 4 races, it is expected that there are only 4 outcomes with 5 ships each. For example, a Klingon ship would yield the vector [1 0 0 0], with the first element representing the likelihood the ship is Klingon.

A linear associator matrix was created as the outer product of one of the input and output vectors. The matrix is then modified, using an error correction technique. The matrix is used for each pair and the error is assessed from every other pair. If unsatisfactory, a new random pair is selected to generate the error matrix which is used to correct the current matrix. The criterion for testing for convergence is set to .01% error of 5 iterations prior.

The matrix is then used to generate the expected output from the corrupted data. The output is transformed into each race and is interpreted as friendly, alert, or hostile. The corrupted data is hard coded into the program. If a user were to introduce more corrupt data into the program, they would be required to code it directly into the program. Because we were supplied with the actual outcomes, the system will also give feedback as to which ships contained an error in learning.

After many trials these are the outcomes:

Trial 1:

Romulon: Alert

Federation: Friendly

Federation: Friendly

Federation: Friendly

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Anterean: Friendly

Klingon: Hostile

Klingon: Hostile

Anterean: Friendly

Anterean: Friendly

Romulon: Alert

Romulon: Alert

Anterean: Friendly

Anterean: Friendly

Federation: Friendly

Federation: Friendly

Trial 2:

Romulon: Alert

Federation: Friendly

Federation: Friendly

Federation: Friendly

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Anterean: Friendly

Klingon: Hostile

Klingon: Hostile

Anterean: Friendly

Anterean: Friendly

Romulon: Alert

Romulon: Alert

Anterean: Friendly

Anterean: Friendly

Federation: Friendly

Federation: Friendly

…..many more trials of the same results

Trial 20

Romulon: Alert

Federation: Friendly

Federation: Friendly

Federation: Friendly

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Anterean: Friendly

Klingon: Hostile

Klingon: Hostile

Anterean: Friendly

Anterean: Friendly

Romulon: Alert

Romulon: Alert

Anterean: Friendly

Anterean: Friendly

Error

Federation: Friendly

19: Federation vs Romulon

In this case, it took 20 trials before coming up with one undesired result.

……

Trial 55

Error

Federation: Friendly

Federation: Friendly

Federation: Friendly

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Romulon: Alert

Klingon: Hostile

Anterean: Friendly

Klingon: Hostile

Klingon: Hostile

Anterean: Friendly

Anterean: Friendly

Romulon: Alert

Romulon: Alert

Anterean: Friendly

Anterean: Friendly

Error

Federation: Friendly

1: Romulon vs Federation

19: Federation vs Anterean

There can be up to 2 undesired results, but this is exceedingly rare.

The pairs that are most often mixed up are numbers 1 and 19, with the word “often” being a relative term towards the other numbers. In fact, errors in the other numbers have not yet been witnessed, but may occur. The errors in the numbers are mostly Federation and Romulan. The 2 races have slight crosses in their ranges when it comes to surface reflection color (Light Gray) and ratio for long to short axes (2.1). When the corrupted data is mostly limited to these ranges for only these properties, the system will have trouble correctly identifying the correct race.

List of Rules:

names:

RULE: ships starting with 2 consonants = klingon

RULE: ships starting with A or E = anterean

RULE: ships with the second element as a number = anterean

RULE: ships with no name have an equal chance of being klingon, anterean or neither

RULE: ships that are not klingon or anterean are neither

warp drive vibration index:

PREMISE: if warp drive is 0, the data for the warp drive is not available

RULE: if warp drive data does not exist, then the ship has an equally likely chance of being any of the warp drive ranges

RULE: if a warp drive falls into any one of these ranges, then it cannot

fall into the other ranges

ASSUMPTION: if warp drive is 6.4, the ship is likely federation

ASSUMPTION: if warp drive is 6.5- 6.8, the ship is likely federation or anterean

ASSUMPTION: if warp drive is 6.9- 7.1, the ship is likely klingon

ASSUMPTION: if warp drive is 7.2- 7.3, the ship is likely romulan or klingon

ASSUMPTION: if warp drive is greater than 7.4, the ship is likely romulan

hailer transponder frequencies:

PREMISE: if hailer frequency is 10000, then the hailer frequency is considered greater than 1000

RULE: if hailer frequency is greater than 1000, then the ship has an equally likely chance of being any of the hailer frequency ranges greater than 1000

PREMISE: if hailer frequency is -10000, then the hailer frequency is considered less than 1000

RULE: if hailer frequency is less than 1000, then the ship has an equally likely chance of being any of the hailer frequency ranges less than 1000

PREMISE: if hailer frequency is 0, the data for the hailer frequency is not available

RULE: if hailer frequency data does not exist, then the ship has an equally likely chance of being any of the hailer frequency ranges

ASSUMPTION: hailer frequencies up to 978.0 are considered romulan

ASSUMPTION: hailer frequencies from 978.0 to 980.4 are considered romulan or klingon

ASSUMPTION: hailer frequencies from 980.4 to 1008.7 are considered klingon

ASSUMPTION: hailer frequencies from 1008.7 to 1044.9 are considered mostly anterean

ASSUMPTION: hailer frequencies from above 1050.0 are considered mostly federation

color:

ASSUMPTION: black and dark gray ships are likely klingon

ASSUMPTION: light gray ships are likely romulan or federation

ASSUMPTION: dark blue or green ships are likely romulan

ASSUMPTION: light blue or blue ships are likely romulan or anterean

ASSUMPTION: pink, orange, or yellow ships are likely anterean

ASSUMPTION: white ships are likely federation

RULE: if color is dark, then the ship has an equally likely chance of being any of the dark color ranges

RULE: if color is light, then the ship has an equally likely chance of being any of the light color ranges

PREMISE: if color is \*, the data for the color is not available

RULE: if color data does not exist, then the ship has an equally likely chance of being any of the color ranges

ratio of long to short axes:

PREMISE: if axis ratio is 0, the data for the axis ratio is not available

RULE: if axis ratio data does not exist, then the ship has an equally likely chance of being any of the axis ratio ranges

ASSUMPTION: axis ratios from 1- 1.5 are likely anterean

ASSUMPTION: axis ratios from 1.5- 1.8 are likely romulan

ASSUMPTION: axis ratios from 1.8- 2.2 are likely romulan or federation

ASSUMPTION: axis ratios from 1- 1.5 are likely federation

ASSUMPTION: axis ratios greater than 2.7 are likely klingon