Μοντέλα Markov (MM)
και
κρυφά μοντέλα Markov (HMM)

Χρονικά ή ακολουθιακά δεδομένα

Χρονικά ή ακολουθιακά δεδομένα (π.χ. DNA, κείμενο, ομιλία, οι τιμές μιας μετοχής, καιρός, οι συναισθηματικές καταστάσεις ενός ατόμου.)

Παρατήρηση: Σε πολλά χρονικά ή ακολουθιακά δεδομένα οι τιμές που εμφανίζονται σε κάθε χρονική στιγμή, ή σε κάθε θέση, εξαρτώνται (με κάποιο πιθανοτικό τρόπο) από τις τιμές που εμφανίστηκαν στο παρελθόν ή σε προηγούμενες θέσεις.
(Μια ακολουθία από *ανεξάρτητες* τυχαίες μεταβλητές είναι το ακραίο παράδειγμα όπου δεν υπάρχει καμία εξάρτηση από προηγούμενες τιμές).
Προκειμένου να μοντελοποιήσουμε τέτοια δεδομένα έχουν προταθεί διάφοροι τρόποι (π.χ. νευρωνικά δίκτυα, χρονοσειρές).
Στη συνέχεια θα δούμε δύο απλούς αλλά χρήσιμους τρόπους μοντελοποίησης: τα μοντέλα Markov και τα κρυφά μοντέλα Markov.

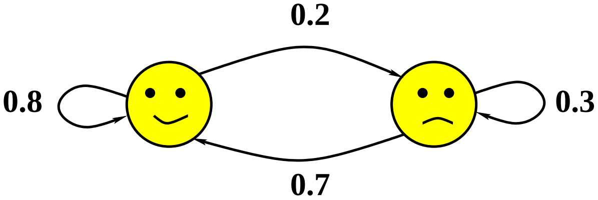
Μοντέλα Markov

Μοντέλο Markov n καταστάσεων (n **state Markov model - MM**) είναι ένα σύστημα τέτοιο ώστε:

- Σε κάθε χρονική στιγμή t βρίσκεται ακριβώς σε μια από τις n καταστάσεις $1, 2, \ldots, n$.
  Με X_t θα συμβολίζουμε την κατάσταση στην οποία βρίσκεται το σύστημα την χρονική στιγμή t.
- Υπάρχει μια μήτρα $A = [a_{ij}]$ τέτοια ώστε, αν την χρονική t το σύστημα βρίσκεται στην κατάσταση i, τότε το στοιχείο a_{ij} της μήτρας A ισούται με την πιθανότητα μετάβασης του συστήματος στην κατάσταση j την χρονική στιγμή $t+1$.
  Ιδιότητα Markov: Η νέα κατάσταση του συστήματος εξαρτάται μόνο από την προηγούμενη κατάσταση και οι πιθανότητες μετάβασης από την κατάσταση i στην κατάσταση j δεν αλλάζουν με τον χρόνο t.
  Η μήτρα A ονομάζεται **μήτρα μετάβασης** του συστήματος.
- Συνήθως τα μοντέλα Markov αναπαρίστανται ως κατευθυνόμενα γραφήματα με ετικέτες.

Μοντέλα Markov

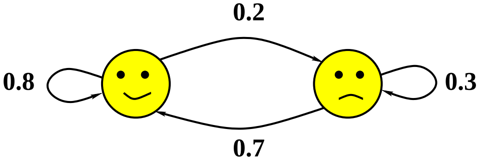
Παράδειγμα 1: Ένα μοντέλο Markov για την διάθεση



$$A = \begin{bmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{bmatrix}$$

Παρατήρηση: Το άθροισμα των πιθανοτήτων στα εξερχόμενα τόξα κάθε κατάστασης ισούται με 1. (Το άθροισμα των στοιχείων κάθε γραμμής της A ισούται με 1)

Μοντέλα Markov

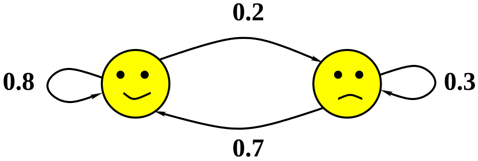


Ερώτηση 1: Έστω ότι σήμερα το άτομο είναι χαρούμενο πόσο πιθανό στις επόμενες 4 μέρες η διάθεση του να είναι

$$H \quad H \quad S \quad S$$

όπου H: Χαρούμενη και S: Στεναχωρημένη.

Μοντέλα Markov



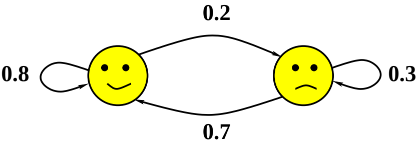
Ερώτηση 1: Έστω ότι σήμερα το άτομο είναι χαρούμενο πόσο πιθανό στις επόμενες 4 μέρες η διάθεση του να είναι

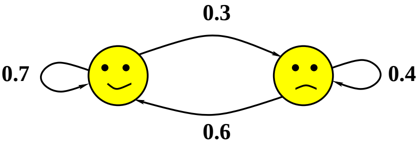
$$H \quad H \quad S \quad S$$

όπου H: Χαρούμενη και S: Στεναχωρημένη.
Απάντηση: Ψάχνουμε την πιθανότητα

$$\begin{aligned} P(HHSS|H) &= P(H|H)P(H|HH)P(S|HHH)P(S|HHHS) \\ &\overset{\text{Από την ιδιότητα Markov}}{=} P(H|H)P(H|H)P(S|H)P(S|S) \\ &= 0.8 \cdot 0.8 \cdot 0.2 \cdot 0.3 = 0.0384 \end{aligned}$$

(Η απάντηση είναι το γινόμενο των πιθανοτήτων στα τόξα)

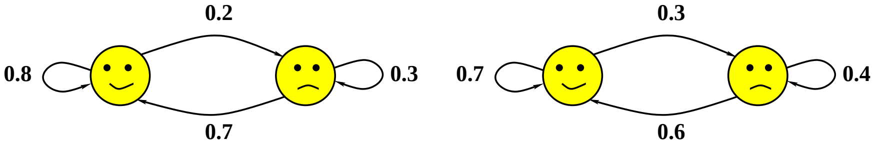




Άτομο A Άτομο B

Ερώτηση 2: Έστω ότι αρχικά τα άτομα A και B είναι χαρούμενα. Ποιο από τα δύο άτομα είναι ποιο πιθανό στις 4 μέρες που ακολούθησαν να είχε την παρακάτω διάθεση: $H \quad H \quad S \quad S$

Μοντέλα Markov



Άτομο A Άτομο B

Ερώτηση 2: Έστω ότι αρχικά τα άτομα A και B είναι χαρούμενα. Ποιο από τα δύο άτομα είναι ποιο πιθανό στις 4 μέρες που ακολούθησαν να είχε την παρακάτω διάθεση: $H \quad H \quad S \quad S$

Απάντηση: Για το άτομο Α:

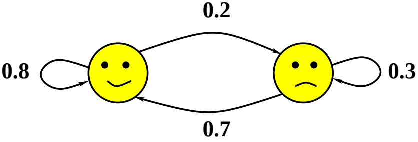
$$P = P(H|H)P(H|H)P(S|H)P(S|S) = 0.8 \cdot 0.8 \cdot 0.2 \cdot 0.3 = 0.0384$$

Για το άτομο Β:

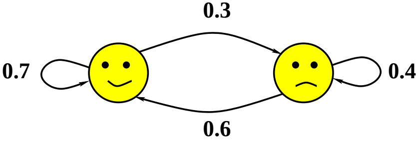
$$P = P(H|H)P(H|H)P(S|H)P(S|S) = 0.7 \cdot 0.7 \cdot 0.3 \cdot 0.4 = 0.0588$$

Επομένως, είναι **πιο πιθανό** η ακολουθία $HHSS$ να περιγράφει την διάθεση του ατόμου B.

Μοντέλα Markov

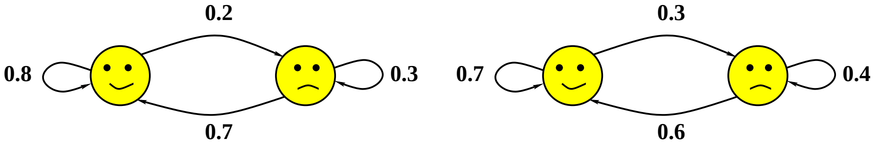


Άτομο A



Άτομο B

Ερώτηση 3: Έστω ότι αρχικά τα άτομα A και B είναι χαρούμενα. Ποιο από τα δύο άτομα έχει περισσότερα ‘‘σκαμπανεβάσματα’’ στην διάθεσή του; Για παράδειγμα έχει τις εναλλαγές: $S \quad H \quad S \quad H$



Άτομο A Άτομο B

Ερώτηση 3: Έστω ότι αρχικά τα άτομα A και B είναι χαρούμενα. Ποιο από τα δύο άτομα έχει περισσότερα ‘‘σκαμπανεβάσματα’’ στην διάθεσή του; Για παράδειγμα έχει τις εναλλαγές: $S \quad H \quad S \quad H$

Απάντηση: Για το άτομο A:

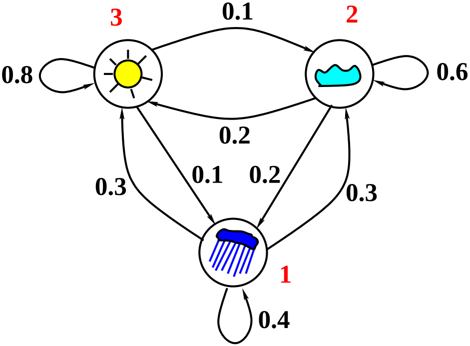
$$P = P(S|H)P(H|S)P(S|H)P(H|S) = 0.2 \cdot 0.7 \cdot 0.2 \cdot 0.7 = 0.0196$$

Για το άτομο B:

$$P = P(S|H)P(H|S)P(S|H)P(H|S) = 0.3 \cdot 0.6 \cdot 0.3 \cdot 0.6 = 0.0324$$

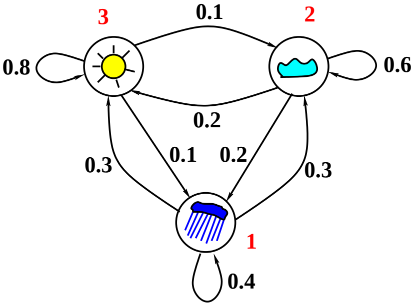
Επομένως, είναι **πιο πιθανό** η ακολουθία $SHSH$ να περιγράφει την διάθεση του ατόμου B.

Παράδειγμα 2: Ένα μοντέλο Markov για τον καιρό.



$$A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Μοντέλα Markov



Ερώτηση 1: Έστω ότι σήμερα ο καιρός είναι βροχερός, ποια είναι η πιθανότητα (σύμφωνα με το μοντέλο του καιρού) στις οκτώ επόμενες μέρες ο καιρός να είναι

Η, Η, Η, Β, Β, Η, Σ, Η

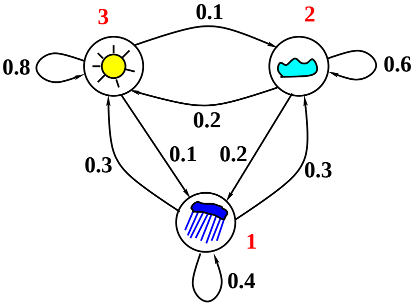
όπου

Η: Ηλιοφάνεια

Β: Βροχερός

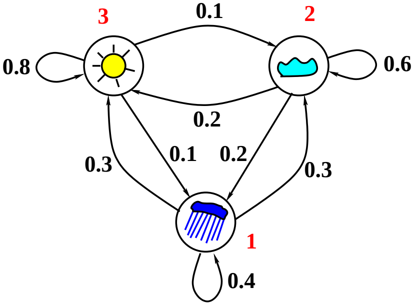
Σ: Συννεφιασμένος

Μοντέλα Markov



Ερώτηση 2: Έστω ότι σήμερα ο καιρός είναι βροχερός, ποια είναι πιθανότητα να ο καιρός να παραμείνει βροχερός για k ημέρες; Το ίδιο ερώτημα για την ηλιοφάνεια και τον συννεφιασμένο καιρό.

Μοντέλα Markov



Ερώτηση 3: Ποιος είναι ο αναμενόμενος αριθμός ημερών που ο καιρός θα είναι βροχερός;

Μοντέλα Markov

Μια απάντηση στο ερώτημα αυτό μπορεί να δοθεί με προσομοίωση του συστήματος καιρού για μεγάλο αριθμό φορών.

```
import numpy as np
import random

#Transition matrix
A = np.array([[0.4,0.3,0.3],[0.2,0.6,0.2],[0.1,0.1,0.8]])

#States
States = {0: 'Rainy', 1: 'Cloudy', 2: 'Sunny'}

#Set number of simulations
N = 1000
#Set initial state
S = 0
TotalCounters = np.array([0,0,0])
```

Μοντέλα Markov

```python
for runs in range(N):

    #Statistics
    Rs, Cs, Ss = 0, 0, 0
    Counters = np.array([Rs, Cs, Ss])
    Weather = []

    #Set number of days for each run
    days = 100
    for i in range(days):
        r = random.uniform(0, 1)
        if(r < A[S,0]):
            S = 0
        else:
            if(r < A[S,0] + A[S,1]):
                S = 1
            else:
                S = 2
        Counters[S]+=1
        Weather.append(States[S])

    print(Weather)
    np.add(TotalCounters, Counters, out=TotalCounters)

#Print results
print("# of runs", N, "# of days in each run", days)
print("# of Rs:", TotalCounters[0], "# of Cs:", TotalCounters[1], "# of Ss:",
     TotalCounters[2])
print("Percentages:")
print("R:", TotalCounters[0]/N, "C:", TotalCounters[1]/N, "S:", TotalCounters[2]/N)
```

Μοντέλα Markov

Μετά από 1000 προσομοιώσεις με 100 μέρες η κάθε μια προκύπτουν τα παρακάτω στατιστικά:

```
# of runs 1000 # of days in each run 100
# of Rs: 18372 # of Cs: 27281 # of Ss: 54347
 Percentages:
R: 18.372 C: 27.281 S: 54.347
```

Με άλλα λόγια,
το 54% των ημερών αναμένουμε ηλιοφάνεια
το 27% των ημερών αναμένουμε συννεφιά
το 18% των ημερών αναμένουμε βροχή

Εναλλακτικά, αντί της προσομοίωσης, μπορούμε να χρησιμοποιήσουμε μοντελοποίηση του συστήματος με μήτρες:
Έστω X_t η τ.μ. η οποία ισούται με την κατάσταση στην οποία βρίσκεται το σύστημα την χρονική στιγμή t.

Ιδιότητα

Σε κάθε χρονική στιγμή $t \geq 0$ και για κάθε κατάσταση j ισχύει ότι

$$
\begin{aligned}
P(X_{t+1} = j) &= \sum_{i=1}^{n} P(X_t = i) a_{ij} \\
&= P(X_t = 1) a_{1j} + P(X_t = 2) a_{2j} + \cdots + P(X_t = n) a_{nj}
\end{aligned}
$$

Μοντέλα Markov

Έτσι, για κάθε t ισχύουν οι παρακάτω n ισότητες

$$\begin{aligned}
P(X_{t+1} = 1) &= P(X_t = 1)a_{11} + P(X_t = 2)a_{21} + \cdots + P(X_t = n)a_{n1}\\
P(X_{t+1} = 2) &= P(X_t = 1)a_{12} + P(X_t = 2)a_{22} + \cdots + P(X_t = n)a_{n2}\\
&\vdots\\
P(X_{t+1} = n) &= P(X_t = 1)a_{1n} + P(X_t = 2)a_{2n} + \cdots + P(X_t = n)a_{nn}
\end{aligned}$$

Μοντέλα Markov

Αν θεωρήσουμε τον πίνακα-στήλη

$$Y_t = \begin{bmatrix} P(X_t = 1) \\ P(X_t = 2) \\ \vdots \\ P(X_t = n) \end{bmatrix}$$

τότε εύκολα προκύπτει ότι

$$\begin{bmatrix} P(X_{t+1} = 1) \\ P(X_{t+1} = 2) \\ \vdots \\ P(X_{t+1} = n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \ddots & \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} P(X_t = 1) \\ P(X_t = 2) \\ \vdots \\ P(X_t = n) \end{bmatrix}$$

ή, ισοδύναμα

$$Y_{t+1} = A^T Y_t.$$

όπου A^T είναι η ανάστροφη της μήτρας μετάβασης A.

Μοντέλα Markov

Επαναληπτικά προκύπτει ότι

$$
\begin{aligned}
Y_{t+1} &= A^T Y_t \\
&= A^T(A^T Y_{t-1}) = (A^T)^2 Y_{t-1} \\
&= (A^T)^2(AY_{t-2}) = (A^T)^3 Y_{t-2} \\
&= \cdots \\
&= (A^T)^t Y_1
\end{aligned}
$$

Δηλαδή, αν γνωρίζουμε την κατάσταση του μοντέλου την χρονική στιγμή $t = 1$ μπορούμε χρησιμοποιώντας την τελευταία ισότητα για να υπολογίσουμε την πιθανότητα το μοντέλο μας να βρίσκεται σε μια από τις n καταστάσεις την χρονική στιγμή $t + 1$.

Μοντέλα Markov

Για παράδειγμα, στο μοντέλο του καιρού, όπου

$$A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

αν θέσουμε

$$Y_t = \begin{bmatrix} P(X_t = B) \\ P(X_t = \Sigma) \\ P(X_t = H) \end{bmatrix}$$

τότε έχουμε ότι

$$Y_{t+1} = (A^T)^t Y_1 \Leftrightarrow \begin{bmatrix} P(X_{t+1} = B) \\ P(X_{t+1} = \Sigma) \\ P(X_{t+1} = H) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.2 & 0.8 \end{bmatrix}^t \begin{bmatrix} P(X_1 = B) \\ P(X_1 = \Sigma) \\ P(X_1 = H) \end{bmatrix}$$

Μοντέλα Markov

Αν για παράδειγμα ο καιρός την χρονική στιγμή $t = 1$ είναι βροχερός τότε έχουμε

$$Y_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

και

$$\begin{bmatrix} P(X_{t+1} = B) \\ P(X_{t+1} = \Sigma) \\ P(X_{t+1} = H) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.2 & 0.8 \end{bmatrix}^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Για $t = 99$ προκύπτει ότι

$$\begin{bmatrix} P(X_{100} = B) \\ P(X_{100} = \Sigma) \\ P(X_{100} = H) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.2 & 0.8 \end{bmatrix}^{99} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.1818 \\ 0.2727 \\ 0.5454 \end{bmatrix}$$

δηλαδή η πιθανότητα την χρονική στιγμή $t = 100$ ο καιρός να είναι και πάλι βροχερός είναι 0.1818.

Παράδειγμα 3: Μοντέλα Markov για κείμενα
Ένα κείμενο (όπως για παράδειγμα ένα θεατρικό έργο του Shakespeare) μπορεί να ορίσει ένα μοντέλο Markov.

- Κάθε (διαφορετική) λέξη αντιστοιχεί σε μια κατάσταση του μοντέλου.
- Η πιθανότητα μετάβασης από την λέξη1 στην λέξη2 υπολογίζεται από τον αριθμό των φορών που η λέξη1 ακολουθείται από την λέξη2 στο κείμενο (αντί να ακολουθείται από κάποια άλλη λέξη).
- Αν κατασκευασθεί η μήτρα μετάβασης που αντιστοιχεί στο κείμενο, τότε αυτή μπορεί να χρησιμοποιηθεί για να παράγονται τυχαία κείμενα που μοιάζουν με το αρχικό.
- Επίσης, μπορούμε να εξακριβώσουμε αν ένα κείμενο προέρχεται από μια συγκεκριμένη πηγή.

Για παράδειγμα, θεωρώντας ως σύνολο εκπαίδευσης, την επόμενη λίστα με τα γνωστά ονόματα pokemon:

Abomasnow Abra Absol Accelgor Aegislash Aerodactyl Aggron Aipom Alakazam Alomomola Altaria Amaura Ambipom Amoonguss Ampharos Anorith Araquanid Arbok Arcanine Arceus Archen Archeops Ariados Armaldo Aromatisse Aron Articuno Audino Aurorus Avalugg Axew Azelf Azumarill Azurill Bagon Baltoy Banette Barbaracle Barboach Basculin Bastiodon Bayleef Beartic Beautifly Beedrill Beheeyem Beldum Bellossom Bellsprout Bergmite Bewear Bibarel Bidoof Binacle Bisharp Blastoise Blaziken Blissey Blitzle Boldore Bonsly Bouffalant Bounsweet Braixen Braviary Breloom Brionne Bronzong Bronzor Bruxish Budew Buizel Bulbasaur Buneary Bunnelby Burmy Butterfree Buzzwole Cacnea Cacturne Camerupt Carbink Carnivine Carracosta Carvanha Cascoon Castform Caterpie Celebi Celesteela Chandelure Chansey Charizard Charjabug Charmander Charmeleon Chatot Cherrim Cherubi Chesnaught Chespin Chikorita Chimchar Chimecho Chinchou Chingling Cinccino Clamperl Clauncher Clawitzer Claydol Clefable Clefairy Cleffa Cloyster Cobalion Cofagrigus Combee Combusken Comfey Conkeldurr Corphish Corsola Cosmoem Cosmog Cottonee Crabominable Crabrawler Cradily Cranidos Crawdaunt Cresselia Croagunk Crobat Croconaw Crustle Cryogonal Cubchoo Cubone Cutiefly Cyndaquil Darkrai Darmanitan Dartrix Darumaka Decidueye Dedenne Deerling Deino Delcatty Delibird Delphox Deoxys Dewgong Dewott Dewpider Dhelmise Dialga Diancie Diggersby Diglett Ditto Dodrio Doduo Donphan Doublade Dragalge Dragonair Dragonite Drampa Drapion Dratini Drifblim Drifloon Drilbur Drowzee Druddigon Ducklett Dugtrio Dunsparce Duosion Durant Dusclops Dusknoir Duskull Dustox Dwebble Eelektrik Eelektross Eevee Ekans Electabuzz Electivire Electrike Electrode Elekid Elgyem Emboar Emolga Empoleon Entei Escavalier Espeon Espurr Excadrill Exeggcute Exeggutor Exploud Farfetch'd Fearow Feebas Fennekin Feraligatr Ferroseed Ferrothorn Finneon Flaaffy Flab?b? Flareon Fletchinder Fletchling Floatzel Floette Florges Flygon Fomantis Foongus Forretress Fraxure Frillish Froakie Frogadier Froslass Furfrou Furret Gabite Gallade Galvantula Garbodor Garchomp Gardevoir Gastly Gastrodon Genesect Gengar Geodude Gible Gigalith Girafarig Giratina Glaceon Glalie Glameow Gligar Gliscor Gloom Gogoat Golbat Goldeen Golduck Golem Golett Golisopod Golurk Goodra Goomy Gorebyss Gothita Gothitelle Gothorita Gourgeist Granbull Graveler Greninja Grimer Grotle Groudon Grovyle Growlithe Grubbin Grumpig Gulpin Gumshoos Gurdurr Guzzlord Gyarados Hakamo-o Happiny Hariyama Haunter Hawlucha Haxorus Heatmor Heatran Heliolisk Helioptile Heracross Herdier Hippopotas Hippowdon Hitmonchan Hitmonlee Hitmontop Honchkrow Honedge Ho-Oh Hoopa Hoothoot Hoppip Horsea Houndoom Houndour Huntail Hydreigon Hypno Igglybuff Illumise Incineroar Infernape Inkay Ivysaur Jangmo-o Jellicent Jigglypuff Jirachi Jolteon Joltik Jumpluff Jynx Kabuto Kabutops Kadabra Kakuna Kangaskhan Karrablast Kartana Kecleon Keldeo Kingdra Kingler Kirlia Klang Klefki Klink Klinklang Koffing Komala Kommo-o Krabby Kricketot Kricketune Krokorok Krookodile Kyogre Kyurem Lairon Lampent Landorus Lanturn Lapras Larvesta Larvitar Latias Latios Leafeon Leavanny Ledian Ledyba Lickilicky Lickitung Liepard Lileep Lilligant Lillipup Linoone Litleo Litten Litwick Lombre Lopunny Lotad Loudred Lucario Ludicolo Lugia Lumineon Lunala Lunatone Lurantis Luvdisc Luxio Luxray Lycanroc

Μοντέλα Markov

Machamp Machoke Machop Magby Magcargo Magearna Magikarp Magmar Magmortar Magnemite Magneton Magnezone Makuhita Malamar Mamoswine Manaphy Mandibuzz Manectric Mankey Mantine Mantyke Maractus Mareanie Mareep Marill Marowak Marshadow Marshtomp Masquerain Mawile Medicham Meditite Meganium Meloetta Meowstic Meowth Mesprit Metagross Metang Metapod Mew Mewtwo Mienfoo Mienshao Mightyena Milotic Miltank Mime Jr. Mimikyu Minccino Minior Minun Misdreavus Mismagius Moltres Monferno Morelull Mothim Mr. Mime Mudbray Mudkip Mudsdale Muk Munchlax Munna Murkrow Musharna Natu Necrozma Nidoking Nidoqueen Nidoran? Nidorina Nidorino Nihilego Nincada Ninetales Ninjask Noctowl Noibat Noivern Nosepass Numel Nuzleaf Octillery Oddish Omanyte Omastar Onix Oranguru Oricorio Oshawott Pachirisu Palkia Palossand Palpitoad Pancham Pangoro Panpour Pansage Pansear Paras Parasect Passimian Patrat Pawniard Pelipper Persian Petilil Phanpy Phantump Pheromosa Phione Pichu Pidgeot Pidgeotto Pidgey Pidove Pignite Pikachu Pikipek Piloswine Pineco Pinsir Piplup Plusle Politoed Poliwag Poliwhirl Poliwrath Ponyta Poochyena Popplio Porygon Porygon2 Porygon-Z Primarina Primeape Prinplup Probopass Psyduck Pumpkaboo Pupitar Purrloin Purugly Pyroar Pyukumuku Quagsire Quilava Quilladin Qwilfish Raichu Raikou Ralts Rampardos Rapidash Raticate Rattata Rayquaza Regice Regigigas Regirock Registeel Relicanth Remoraid Reshiram Reuniclus Rhydon Rhyhorn Rhyperior Ribombee Riolu Rockruff Roggenrola Roselia Roserade Rotom Rowlet Rufflet Sableye Salamence Salandit Salazzle Samurott Sandile Sandshrew Sandslash Sandygast Sawk Sawsbuck Scatterbug Sceptile Scizor Scolipede Scrafty Scraggy Scyther Seadra Seaking Sealeo Seedot Seel Seismitoad Sentret Serperior Servine Seviper Sewaddle Sharpedo Shaymin Shedinja Shelgon Shellder Shellos Shelmet Shieldon Shiftry Shiinotic Shinx Shroomish Shuckle Shuppet Sigilyph Silcoon Silvally Simipour Simisage Simisear Skarmory Skiddo Skiploom Skitty Skorupi Skrelp Skuntank Slaking Slakoth Sliggoo Slowbro Slowking Slowpoke Slugma Slurpuff Smeargle Smoochum Sneasel Snivy Snorlax Snorunt Snover Snubbull Solgaleo Solosis Solrock Spearow Spewpa Spheal Spinarak Spinda Spiritomb Spoink Spritzee Squirtle Stantler Staraptor Staravia Starly Starmie Staryu Steelix Steenee Stoutland Stufful Stunfisk Stunky Sudowoodo Suicune Sunflora Sunkern Surskit Swablu Swadloon Swalot Swampert Swanna Swellow Swinub Swirlix Swoobat Sylveon Taillow Talonflame Tangela Tangrowth Tapu Bulu Tapu Fini Tapu Koko Tapu Lele Tauros Teddiursa Tentacool Tentacruel Tepig Terrakion Throh Thundurus Timburr Tirtouga Togedemaru Togekiss Togepi Togetic Torchic Torkoal Tornadus Torracat Torterra Totodile Toucannon Toxapex Toxicroak Tranquill Trapinch Treecko Trevenant Tropius Trubbish Trumbeak Tsareena Turtonator Turtwig Tympole Tynamo Type: Null Typhlosion Tyranitar Tyrantrum Tyrogue Tyrunt Umbreon Unfezant Unown Ursaring Uxie Vanillish Vanillite Vanilluxe Vaporeon Venipede Venomoth Venonat Venusaur Vespiquen Vibrava Victini Victreebel Vigoroth Vikavolt Vileplume Virizion Vivillon Volbeat Volcanion Volcarona Voltorb Vullaby Vulpix Wailmer Wailord Walrein Wartortle Watchog Weavile Weedle Weepinbell Weezing Whimsicott Whirlipede Whiscash Whismur Wigglytuff Wimpod Wingull Wishiwashi Wobbuffet Woobat Wooper Wormadam Wurmple Wynaut Xatu Xerneas Xurkitree Yamask Yanma Yanmega Yungoos Yveltal Zangoose Zapdos Zebstrika Zekrom Zigzagoon Zoroark Zorua Zubat Zweilous Zygarde

Μοντέλα Markov

μπορούμε χρησιμοποιώντας ένα μοντέλο Markov με βάση τις συχνότητες μετάβασης από το ένα γράμμα στο άλλο να παράξουμε τυχαία νέα ονόματα που μοιάζουν με τα υπάρχοντα.

Araquaza Aron Azuriller Bast Bish Blish Bonsly Bouffa Caterbug Charow Chimsicolos Chimsicor Clefable Cloysteela Cofagronzor Crobopassimise Delp Dewpidash Drath Dratias Dusclops Dustle Emboarkrain Espurr Excadrilluxe Flord Garbodon Garcherros Gastle Genesectabutop Gengas Glissey Gogon-Z Golbat Goliwrat Goth Groth Hippowdon Honchen Hone Hydred Laprasect Ledyba Lill Litwicketunfishi Lucargle Ludin Luxray Manape Marshtomp Metan Mimer Nidos Octill Octivire Omask Oricorit Palkia Palpix Persianchog Pidgeot Poliwratias Puruglytufful Rapione Ratinaras Relia Relibird Rhydon Rhyper Rose Rowlitzelf Sand Scragalga Seadra Shelle Shieldum Slow Snorlax Sprithe Star Starly Stung Swaloss Sylveon Talos Tapu Bulbasaur Toxicrossom Tsareloon Turtoutlangmora Turtwicketune Tyrogadino Ursarillitombee Vigon-Z Vivire Volcariados Volcarow Wobbull Xurkrain Zorokorus

Μοντέλα Markov

Μια ενδιαφέρουσα ιδέα είναι να συνδυάσει κάποιος δύο κείμενα και να παράγει μετά τυχαίο κείμενο που μοιάζει και με τα δύο!
Παράδειγμα τυχαίου κειμένου που προήλθε από μοντέλο Markov για τα έργα Άμλετ και Αλίκη στη χώρα των θαυμάτων:

"Oh, you foolish Alice!" she answered herself.
"How can you learn lessons in the world were now but to follow him thither with modesty enough, and likelihood to lead it, as our statists do, A baseness to write this down on the trumpet, and called out "First witness!" ... HORATIO: Most like. It harrows me with leaping in her hand, watching the setting sun, and thinking of little pebbles came rattling in at the door that led into a small passage, not much larger than a pig, my dear," said Alice (she was so much gentry and good will As to expend your time with us a story!" said the Caterpillar.

Μοντέλα Markov

Κάθε ακολουθία $X_1, X_2, \dots, X_k$ ονομάζεται **ακολουθία παρατηρήσεων** του μοντέλου Markov.
Συχνά, η αντίστοιχη ακολουθία συμβολίζεται με $O = O_1 O_2 \dots O_k$ και τα O_1, O_2, ..., O_k ονομάζονται **παρατηρούμενα σύμβολα**.
Παρατήρηση: Στα (απλά) μοντέλα Markov τα παρατηρούμενα σύμβολα αντιστοιχίζονται ένα προς ένα με τις καταστάσεις του μοντέλου.
Για παράδειγμα, στο μοντέλο του καιρού, μια ακολουθία παρατηρήσεων είναι η

$$H, H, H, B, B, H, \Sigma, H$$

η οποία δείχνει σε ποια κατάσταση βρίσκεται το σύστημα κάθε φορά.

Κρυφά μοντέλα Markov

Υπάρχουν συστήματα όπου τα παρατηρούμενα σύμβολα δεν αντιστοιχούν μονοσήμαντα στις καταστάσεις του συστήματος, οι οποίες είναι **κρυμμένες**.
Μερικές φορές μπορούμε να μοντελοποίησουμε τέτοια συστήματα με τη βοήθεια των κρυφών μοντέλων Markov.

Κρυφά μοντέλα Markov

Κρυφό μοντέλο Markov n **καταστάσεων και** m **συμβόλων (**n **state hidden Markov model with** m **symbols - HMM)** είναι ένα σύστημα τέτοιο ώστε:

- Σε κάθε χρονική στιγμή t βρίσκεται ακριβώς σε μια από τις n καταστάσεις $1, 2, \dots, n$.
- Σε κάθε χρονική στιγμή t το σύστημα εμφανίζει ένα σύμβολο s (από τα m διαθέσιμα) που εξαρτάται από την κατάσταση i στην οποία βρίσκεται το σύστημα.
- Υπάρχει μια $n \times n$ μήτρα $A = [a_{ij}]$ τέτοια ώστε, αν την χρονική t το σύστημα βρίσκεται στην κατάσταση i, τότε το στοιχείο a_{ij} της μήτρας A ισούται με την πιθανότητα μετάβασης του συστήματος στην κατάσταση j την χρονική στιγμή $t+1$.
  Η μήτρα A ονομάζεται **μήτρα μετάβασης** του συστήματος.
- Υπάρχει μια μήτρα $n \times m$ μήτρα $E = [E_{is}] = [E_i(s)]$ τέτοια ώστε το στοιχείο $e_{is} = e_i(s)$ ισούται με την πιθανότητα το σύστημα στην κατάσταση i να εμφανίσει το σύμβολο s.
  Η μήτρα E ονομάζεται **μήτρα εκπομπής** ή **μήτρα εμφάνισης**.

Κρυφά μοντέλα Markov

Παράδειγμα 1: Έστω ότι διαθέτουμε δύο νομίσματα 1 και 2, τα οποία έχουν τις παρακάτω πιθανότητες να φέρουν H (Heads) ή T (Tails).

- Για το νόμισμα 1: $P(H) = 0.5$ και $P(T) = 0.5$.
- Για το νόμισμα 2: $P(H) = 0.7$ και $P(T) = 0.3$.

Κάθε χρονική στιγμή t επιλέγουμε ένα από τα δύο νομίσματα σύμφωνα με τις παράκατω πιθανότητες

- Αν έχουμε επιλέξει το νόμισμα 1,
  - επιλέγουμε ξανά το νόμισμα 1 με πιθανότητα 0.8
  - επιλέγουμε το νόμισμα 2 με πιθανότητα 0.2
- Αν έχουμε επιλέξει το νόμισμα 2,
  - επιλέγουμε το νόμισμα 1 με πιθανότητα 0.5
  - επιλέγουμε ξανά το νόμισμα 2 με πιθανότητα 0.5

και ρίχνοντας το νόμισμα που επιλέξαμε εμφανίζουμε σε κάποιον μόνο το αποτέλεσμα H ή T (χωρίς να φανερώσουμε ποιο νόμισμα επιλέξαμε).

Κρυφά μοντέλα Markov

Έτσι, για παράδειγμα, κάποιος παρατηρεί τα σύμβολα

$$H, H, H, T, H, T, T$$

αλλά δεν γνωρίζει (διότι είναι κρυμμένο) ποιο νόμισμα παράγει κάθε ένα από αυτά σύμβολα.
Τα δύο νομίσματα και η ακολουθία των ρίψεων τους αποτελούν ένα απλό κρυφό μοντέλο Markov.
Εδώ, έχουμε

$$A = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix} \end{array}$$

και

$$E = \begin{array}{c c} & \begin{array}{cc} H & T \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \end{array}$$

Βασικά προβλήματα στα κρυφά μοντέλα Markov (HMM):
Δίδεται ένα HMM και μια ακολουθία παρατηρήσεων $O = O_1 O_2 \ldots O_k$.

- (**Πρόβλημα πιθανοφάνειας**) Ποια είναι η πιθανότητα η ακολουθία O να έχει παραχθεί από το HMM;
- (**Πρόβλημα αποκωδικοποίησης**) Ποια είναι η πιο πιθανή ακολουθία καταστάσεων του HMM που παρήγαγε αυτή την ακολουθία;
- (**Πρόβλημα εκτίμησης παραμέτρων**) Πως μπορούμε να κατασκευάσουμε ένα HMM για ένα σύστημα;

Κρυφά μοντέλα Markov (HMM)

Τα κρυφά μοντέλα Markov είναι κατάλληλα για την ταξινόμηση σειρών δεδομένων επειδή μπορούν να μοντελοποιήσουν μια ακολουθία δεδομένων και κατόπιν να εκτιμήσουν την πιθανότητα μια άλλη ακολουθία να προέρχεται από το ίδιο μοντέλο.
Μια από τις πρώτες εφαρμογές τους είναι η αναγνώριση ομιλίας, όπου οι λέξεις μοντελοποιούνται ως ακολουθίες ήχων.
Επίσης, χρησιμοποιούνται στην υπολογιστική βιολογία, στην σύνθεση ομιλίας, στην αναγνώριση ανθρώπινου περπατήματος, στην αυτόματη αναγνώριση εκφράσεων προσώπου, στην κρυπτανάλυση, και σε πολλά άλλα πεδία.

Δίδεται ένα HMM και μια ακολουθία παρατηρήσεων $O = O_1 O_2 \ldots O_k$.
Ποια είναι η πιθανότητα η ακολουθία O να έχει παραχθεί από το HMM;

Ένας απλός τρόπος να λύσουμε το πρόβλημα της πιθανοφάνειας είναι χρησιμοποιήσουμε τον **αλγόριθμο forward**.
Για το σκοπό αυτό θα ορίσουμε την βοηθητική συνάρτηση

$$f_j(t) = \begin{array}{r} \text{Η πιθανότητα τις χρονικές στιγμές } 1, 2, \ldots, t \text{ το σύστημα} \\ \text{να έχει εμφανίσει την ακολουθία } O_1, O_2, \ldots, O_t \\ \text{και να έχει καταλήξει στην κατάσταση } j \end{array}$$

για κάθε $t = 1, 2, \ldots, k$ και $j = 1, 2, \ldots, n$.
Το ζητούμενο είναι η τιμή του αθροίσματος

$$f_1(k) + f_2(k) + \cdots + f_n(k)$$

(αφού υπάρχουν ακριβώς n διαφορετικές καταστάσεις στις οποίες μπορεί να βρίσκεται το σύστημα όταν εμφανίσει την ακολουθία O.)

Κρυφά μοντέλα Markov

Η συνάρτηση $f_j(t)$ ικανοποιεί τις παρακάτω ιδιότητες:

- $$f_j(1) = P(X_1 = j)e_j(O_1), \text{ για κάθε } j = 1, 2, \ldots, n.$$

Δηλαδή, η πιθανότητα την χρονική στιγμή $t = 1$ το σύστημα να εμφανίσει το σύμβολο O_1 και να βρίσκεται στην κατάσταση j ισούται με $P(X_1 = j)e_j(O_1)$.
(Πράγματι, η πιθανότητα να βρίσκεται στην κατάσταση i την χρονική στιγμή $t = 1$ ισούται με $P(X_1 = j)$ και η πιθανότητα να εμφανίσει το σύμβολο O_1 ισούται με $e_j(O_1)$.)

Κρυφά μοντέλα Markov

Η συνάρτηση $f_j(t)$ ικανοποιεί τις παρακάτω ιδιότητες:

- $$\begin{aligned} f_j(t+1) &= \left(\sum_{i=1}^{n} f_i(t) a_{ij} \right) e_j(O_{t+1}) \\ &= \left(f_1(t)a_{1j} + f_2(t)a_{2j} + \cdots + f_n(t)a_{nj} \right) e_j(O_{t+1}) \end{aligned}$$

  για κάθε $t \geq 1$ και $j = 1, 2, \ldots, n$.
  Δηλαδή, η πιθανότητα την χρονική στιγμή $t+1$ το σύστημα να εμφανίσει το σύμβολο O_{i+1} δεδομένου τις χρονικές στιγμές $1, 2, \ldots, t$ έχει εμφανίσει τα σύμβολα $O_1, O_2, \ldots, O_t$ και τώρα να βρίσκεται στην κατάσταση j ισούται με $\left(\sum_{i=1}^{n} f_i(t) a_{ij} \right) e_j(O_{t+1})$.

Κρυφά μοντέλα Markov

Παράδειγμα: Να βρεθεί η πιθανότητα το HMM που αποτελείται από τα δύο νομίσματα 1 και 2 να παράγει την ακολουθία

$$\begin{array}{cccc} O_1 & O_2 & O_3 & O_4 \\ H & H & T & H \end{array}$$

Υπενθύμιση:

$$A = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix} \end{array} \qquad E = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} H & T \end{array} \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \end{array}$$

Υπόθεση: Έστω ότι την χρονική στιγμή $t = 1$ έχουμε ίση πιθανότητα να επιλέξουμε το νόμισμα 1 ή το νόμισμα 2, δηλαδή

$$P(X_1 = 1) = P(X_1 = 2) = 0.5.$$

Εδώ, ψάχνουμε την τιμή του αθροίσματος

$$f_1(4) + f_2(4).$$

Θα υπολογίσουμε διαδοχικά τις τιμές της $f_i(4)$ για $t = 1, 2, 3, 4$.

- $t = 1,\ O_1 = H$.
  $f_1(1) = P(X_1 = 1)e_1(H) = 0.5 \cdot 0.5 = \frac{1}{4} = 0.25$.
  $f_2(1) = P(X_1 = 2)e_2(H) = 0.5 \cdot 0.7 = \frac{7}{20} = 0.35$.
  Δηλαδή, η πιθανότητα να εμφανισθεί H ισούται με
  $f_1(1) + f_2(1) = \frac{3}{5} = 0.6$.
- $t = 2,\ O_2 = H$.
  $f_1(2) = (f_1(1) \cdot a_{11} + f_2(1) \cdot a_{21}) \cdot e_1(H) = \left(\frac{1}{4} \cdot 0.8 + \frac{7}{20} \cdot 0.5\right) \cdot 0.5 = \frac{3}{16}$.
  $f_2(2) = (f_1(1) \cdot a_{12} + f_2(1) \cdot a_{22}) \cdot e_2(H) = \left(\frac{1}{4} \cdot 0.2 + \frac{7}{20} \cdot 0.5\right) \cdot 0.7 = \frac{63}{400}$.
  Δηλαδή, η πιθανότητα να εμφανισθεί HH ισούται με
  $f_1(2) + f_2(2) = \frac{69}{200} = 0.345$.

Κρυφά μοντέλα Markov

- $t = 3,\ O_3 = T$.
  $f_1(3) = (f_1(2) \cdot a_{11} + f_2(2) \cdot a_{21}) \cdot e_1(T) = \left(\frac{3}{16} \cdot 0.8 + \frac{63}{400} \cdot 0.5\right) \cdot 0.5 = \frac{183}{1600} = 0.114$.
  $f_2(3) = (f_1(2) \cdot a_{12} + f_2(2) \cdot a_{22}) \cdot e_2(T) = \left(\frac{3}{16} \cdot 0.2 + \frac{63}{400} \cdot 0.5\right) \cdot 0.3 = \frac{279}{8000} = 0.034$.
  Δηλαδή, η πιθανότητα να εμφανισθεί HHT ισούται με
  $f_1(3) + f_2(3) = \frac{597}{4000} = 0.14925$.
- $t = 4,\ O_4 = H$.
  $f_1(4) = (f_1(3) \cdot a_{11} + f_2(3) \cdot a_{21}) \cdot e_1(H) = \left(\frac{183}{1600} \cdot 0.8 + \frac{279}{8000} \cdot 0.5\right) \cdot 0.5 = \frac{1743}{32000} = 0.054$.
  $f_2(4) = (f_1(3) \cdot a_{12} + f_2(3) \cdot a_{22}) \cdot e_2(H) = \left(\frac{183}{1600} \cdot 0.2 + \frac{279}{8000} \cdot 0.5\right) \cdot 0.7 = \frac{903}{32000} = 0.028$.
  Άρα, η πιθανότητα να εμφανισθεί $HHTH$ ισούται με
  $f_1(4) + f_2(4) = \frac{1323}{16000} = 0.082$.

Κρυφά μοντέλα Markov

Ένα δίκαιο νόμισμα θα παρήγαγε οποιαδήποτε ακολουθία με 4 ρίψεις με πιθανότητα

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^4} = 0.0625$$

Το νόμισμα που ευνοεί την ένδειξη H παρήγαγε την ακολουθία $HHTH$ με πιθανότητα

$$0.7 \cdot 0.7 \cdot 0.3 \cdot 0.7 = 0.1029$$

Η πιθανότητα που υπολογίσαμε είναι 0.082 και είναι αρκετά μεγάλη οπότε η ακολουθία $HHTH$ είναι αρκετά πιθανό ότι παράχθηκε από το συγκεκριμένο μοντέλο.

Παρατήρηση: Μπορούμε να συστηματοποιήσουμε τους υπολογισμούς των τιμών $f_t(i)$ χρησιμοποιώντας πίνακες.
Αν θεωρήσουμε τον πίνακα-στήλη

$$F_t = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \cdots \\ f_n(t) \end{bmatrix}$$

και τους διαγώνιους πίνακες

$$D(S) = \begin{bmatrix} e_1(S) & 0 & \cdots & 0 \\ 0 & e_2(S) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & e_n(S) \end{bmatrix}$$

που ορίζονται για κάθε σύμβολο S του συστήματος, τότε, εύκολα προκύπτει ότι

$$F_{t+1} = D(O_{t+1})A^T F_t$$

οπότε για την ακολουθία παρατηρήσεων $O_1, O_2, \ldots, O_{k-1}, O_k$ προκύπτει ότι

$$F_k = (D(O_k) \cdot A^T)(D(O_{k-1}) \cdot A^T) \cdots (D(O_2) \cdot A^T)(D(O_2) \cdot A^T)F_1.$$

Επειδή κάποια σύμβολα S μπορεί να εμφανίζονται παραπάνω από μια φορά, συμφέρει να υπολογίσουμε αρχικά τους πίνακες

$$P(S) = D(S)A^T, \text{ για κάθε σύμβολο } S.$$

Τότε η ζητούμενη απάντηση είναι το άθροισμα των στοιχείων της μήτρας

$$F_k = P(O_k) \cdot P(O_{k-1}) \cdots P(O_2) \cdot F_1.$$

Για το παράδειγμα με τα 2 νομίσματα έχουμε

$$F_1 = \begin{bmatrix} f_1(1) \\ f_2(1) \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{7}{20} \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.35 \end{bmatrix} \text{ αν το πρώτο σύμβολο είναι } H$$

και

$$F_1 = \begin{bmatrix} f_1(1) \\ f_2(1) \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{20} \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.15 \end{bmatrix} \text{ αν το πρώτο σύμβολο είναι } T$$

$$P(H) = \begin{bmatrix} e_1(H) & 0 \\ 0 & e_2(H) \end{bmatrix} A^T = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.7 \end{bmatrix} \begin{bmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.25 \\ 0.14 & 0.35 \end{bmatrix}$$

$$P(T) = \begin{bmatrix} e_1(T) & 0 \\ 0 & e_2(T) \end{bmatrix} A^T = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.3 \end{bmatrix} \begin{bmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.25 \\ 0.06 & 0.15 \end{bmatrix}$$

και

$$F_k = \begin{bmatrix} f_1(k) \\ f_2(k) \end{bmatrix} = P(O_k) \cdot P(O_{k-1}) \cdots P(O_2) F_1$$

Οπότε, για την ακολουθία H, H, T, H έχουμε

$$F_1 = \begin{bmatrix} 0.25 \\ 0.35 \end{bmatrix}$$

$$F_4 = \begin{bmatrix} f_1(4) \\ f_2(4) \end{bmatrix} = P(H) \cdot P(T) \cdot P(H) \cdot F_1 = \begin{bmatrix} \frac{1743}{32000} \\ \frac{903}{32000} \end{bmatrix}$$

ενώ, για την ακολουθία $T, T, T, T, T, T, T, T, T, T$ έχουμε

$$F_1 = \begin{bmatrix} 0.25 \\ 0.15 \end{bmatrix}$$

$$F_{10} = P(T)^9 F_1 = \begin{bmatrix} 0.000243219 \\ 0.0000486438 \end{bmatrix}$$

δηλαδή υπάρχει $0.000243219 + 0.0000486438 = 0.000291863$ πιθανότητα να εμφανισθεί αυτή η ακολουθία κατά την ρίψη των 2 νομισμάτων.

Δίδεται ένα HMM και μια ακολουθία παρατηρήσεων $O = O_1 O_2 \ldots O_k$. Ποια είναι η πιο πιθανή ακολουθία καταστάσεων του HMM που παρήγαγε αυτή την ακολουθία;

Ένας απλός αλγόριθμος για το πρόβλημα αυτό είναι ο **αλγόριθμος του Viterbi**. (Ανάλογος με τον αλγόριθμο forward.)

Κρυφά μοντέλα Markov

Ο αλγόριθμος του Viterbi χρησιμοποιεί δύο βοηθητικές συναρτήσεις:
Την συνάρτηση

$$\delta_j(t) = \begin{array}{r} \text{Η μέγιστη πιθανότητα να εμφανισθεί η ακολουθία} \\ O_1, O_2, \ldots, O_t \text{ τις χρονικές στιγμές } 1, 2, \ldots, t \\ \text{και το σύστημα να καταλήξει στην κατάσταση } i \\ \text{για κάθε δυνατή ακολουθία καταστάσεων που τελειώνει στην } j \end{array}$$

για κάθε $t = 1, 2, \ldots, k$ και $j = 1, 2, \ldots, n$.
Το ζητούμενο είναι η μέγιστη από τις τιμές

$$\delta_1(k), \delta_2(k), \ldots, \delta_n(k).$$

Κρυφά μοντέλα Markov

Εύκολα προκύπτει ότι η συνάρτηση $\delta_j(t)$ ικανοποιεί τις ιδιότητες

- $$\delta_j(1) = P(X_1 = j)e_j(O_1) \text{ για κάθε } j = 1, 2, \dots, n.$$

- $$\begin{aligned}\delta_j(t+1) &= \left(\max_{\text{για κάθε } i}\{\delta_i(t)a_{ij}\}\right) e_j(O_{t+1})\\ &= \left(\max\{\delta_1(t)a_{1j}, \delta_2(t)a_{2j}, \dots, \delta_n(t)a_{nj}\}\right) e_j(O_{t+1})\end{aligned}$$

  για κάθε $t \geq 1$ και $j = 1, 2, \dots, n$.

Κρυφά μοντέλα Markov

Επίσης, θα χρησιμοποιηθεί η βοηθητική συνάρτηση

$$\psi_j(t) = \begin{array}{l} \text{Ο δείκτης } i \text{ για τον οποίο προκύπτει} \\ \qquad\qquad \text{το μέγιστο } \max\limits_{\text{για κάθε } i} \delta_i(t) a_{ij} \end{array}$$

οι τιμές της οποίες αποθηκεύονται σε μια μήτρα προκειμένου να μπορούμε να ανακτήσουμε την ζητούμενη ακολουθία των καταστάσεων.

Αρχικά, θέτουμε $\psi_i(1) = 0$, **για κάθε** $i = 1, 2, \ldots, n$.

Κρυφά μοντέλα Markov

Παράδειγμα: Να βρεθεί η πιθανότερη ακολουθία καταστάσεων του HMM που αποτελείται από τα δύο νομίσματα 1 και 2 που παράγει την ακολουθία

$$\begin{array}{cccccc} O_1 & O_2 & O_3 & O_4 & O_5 & O_6 \\ H & H & T & H & H & H \end{array}$$

Υπενθύμιση:

$$A = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix} \end{array} \qquad E = \begin{array}{c c} & \begin{array}{cc} H & T \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \end{array}$$

Υπόθεση: Έστω ότι την χρονική στιγμή $t = 1$ έχουμε ίση πιθανότητα να επιλέξουμε το νόμισμα 1 ή το νόμισμα 2, δηλαδή

$$P(X_1 = 1) = P(X_1 = 2) = 0.5.$$

Κρυφά μοντέλα Markov

Εδώ, ψάχνουμε το μέγιστο των τιμών

$$\delta_1(6),\ \delta_2(6).$$

Θα υπολογίσουμε διαδοχικά τις τιμές της $\delta_i(t)$ για $t = 1, 2, 3, 4, 5, 6$ (και αντίστοιχες τιμές $\psi_i(t)$).

- $t = 1,\ O_1 = H$.
  $\delta_1(1) = P(X_1 = 1)e_1(H) = 0.5 \cdot 0.5 = \frac{1}{4} = 0.25$.
  $\psi_1(1) = 0$.
  $\delta_2(1) = P(X_1 = 2)e_2(H) = 0.5 \cdot 0.7 = \frac{7}{20} = 0.35$.
  $\psi_2(1) = 0$.
  Άρα, η πιθανότερη ακολουθία καταταστάσεων για να εμφανισθεί H είναι η ακολουθία: 2.

- $t = 2,\ O_2 = H$.
  $\delta_1(2) = \max\{\delta_1(1) \cdot a_{11}, \delta_2(1) \cdot a_{21}\} \cdot e_1(H) =$
  $\max\{\frac{1}{4} \cdot 0.8, \frac{7}{20} \cdot 0.5\} \cdot 0.5 = \frac{1}{10} = 0.1$.
  $\psi_1(2) = 1$.
  $\delta_2(2) = \max\{\delta_1(1) \cdot a_{12}, \delta_2(1) \cdot a_{22}\} \cdot e_2(H) =$
  $\max\{\frac{1}{4} \cdot 0.2, \frac{7}{20} \cdot 0.5\} \cdot 0.7 = \frac{49}{400} = 0.1225$.
  $\psi_2(2) = 2$.
  Άρα, η πιθανότερη ακολουθία καταστάσεων για να εμφανισθεί H, H είναι η ακολουθία: $2, 2$.

- $t = 3,\ O_3 = T$.
  $\delta_1(3) = \max\{\delta_1(2) \cdot a_{11}, \delta_2(2) \cdot a_{21}\} \cdot e_1(T) =$
  $\max\{\frac{1}{10} \cdot 0.8, \frac{49}{400} \cdot 0.5\} \cdot 0.5 = \frac{1}{25} = 0.04$.
  $\psi_1(3) = 1$.
  $\delta_2(3) = \max\{\delta_1(2) \cdot a_{12}, \delta_2(2) \cdot a_{22}\} \cdot e_2(T) =$
  $\max\{\frac{1}{10} \cdot 0.2, \frac{49}{400} \cdot 0.5\} \cdot 0.3 = \frac{147}{8000} = 0.018$.
  $\psi_2(3) = 2$.
  Άρα, η πιθανότερη ακολουθία καταστάσεων για να εμφανισθεί H, H, T είναι η ακολουθία: $2, 2, 1$.

Κρυφά μοντέλα Markov

- $t = 4,\ O_4 = H$.
  $\delta_1(4) = \max\{\delta_1(3)\cdot a_{11}, \delta_2(3)\cdot a_{21}\}\cdot e_1(H) =$
  $\max\{\frac{1}{25}\cdot 0.8, \frac{147}{8000}\cdot 0.5\}\cdot 0.5 = \frac{2}{125} = 0.016$.
  $\psi_1(4) = 1$.
  $\delta_1(4) = \max\{\delta_1(3)\cdot a_{12}, \delta_2(3)\cdot a_{22}\}\cdot e_2(H) =$
  $\max\{\frac{1}{25}\cdot 0.2, \frac{147}{8000}\cdot 0.5\}\cdot 0.7 = \frac{1029}{160000} = 0.0064$.
  $\psi_2(4) = 2$.
  Άρα, η πιθανότερη ακολουθία καταστάσεων για να εμφανισθεί H, H, T, H είναι η ακολουθία: $2, 2, 1, 1$.

Κρυφά μοντέλα Markov

- $t = 5,\ O_5 = H.$
  $\delta_1(5) = \max\{\delta_1(4) \cdot a_{11}, \delta_2(4) \cdot a_{21}\} \cdot e_1(H) =$
  $\max\{\frac{2}{125} \cdot 0.8, \frac{1029}{160000} \cdot 0.5\} \cdot 0.5 = \frac{4}{625} = 0.0064.$
  $\psi_1(5) = 1.$
  $\delta_2(5) = \max\{\delta_1(4) \cdot a_{12}, \delta_2(4) \cdot a_{22}\} \cdot e_2(H) =$
  $\max\{\frac{2}{125} \cdot 0.2, \frac{1029}{160000} \cdot 0.5\} \cdot 0.7 = \frac{7203}{3200000} = 0.0022.$
  $\psi_2(5) = 2.$
  Άρα, η πιθανότερη ακολουθία καταστάσεων για να εμφανισθεί H, H, T, H, H είναι η ακολουθία: $2, 2, 1, 1, 1$.

Κρυφά μοντέλα Markov

- $t = 6, \ O_6 = H$.
  $\delta_1(6) = \max\{\delta_1(5) \cdot a_{11}, \delta_2(5) \cdot a_{21}\} \cdot e_1(H) =$
  $\max\{\frac{4}{625} \cdot 0.8, \frac{7203}{3200000} \cdot 0.5\} \cdot 0.5 = \frac{8}{3125} = 0.00256$.
  $\psi_1(6) = 1$.
  $\delta_2(6) = \max\{\delta_1(5) \cdot a_{12}, \delta_2(5) \cdot a_{22}\} \cdot e_2(H) =$
  $\max\{\frac{4}{625} \cdot 0.2, \frac{7203}{3200000} \cdot 0.5\} \cdot 0.7 = \frac{14}{15625} = 0.000896$.
  $\psi_2(6) = 1$.
  Άρα, η πιθανότερη ακολουθία καταστάσεων για να εμφανισθεί H, H, T, H, H, H είναι η ακολουθία: $2, 2, 1, 1, 1, 1$.