

Αβεβαιότητα

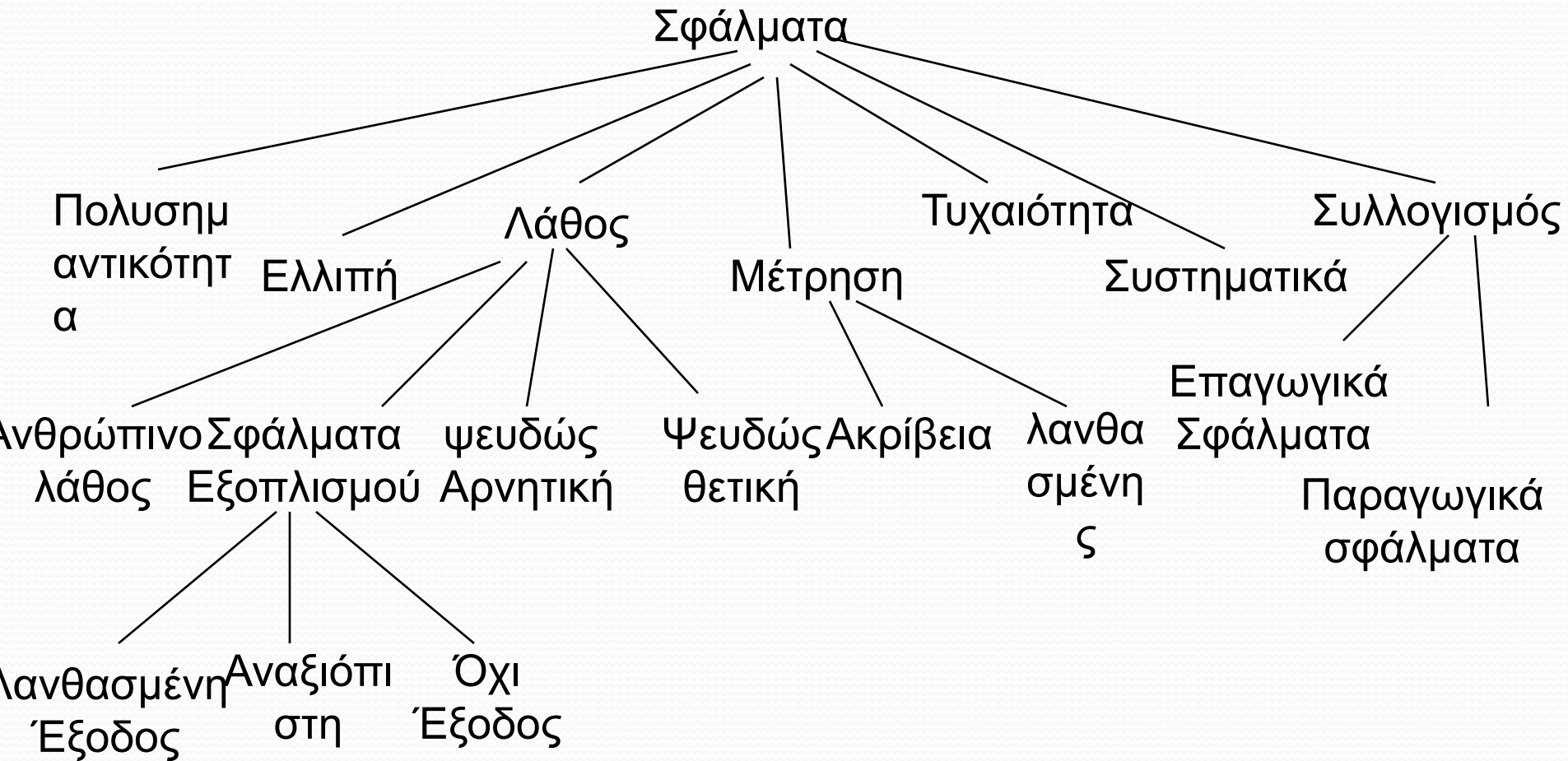
Αβεβαιότητα

- Σε πολλές εφαρμογές θα πρέπει να λυθούν προβλήματα χρησιμοποιώντας όχι πλήρη, λεπτομερή και επακριβή δεδομένα
- Η Αβεβαιότητα μπορεί να θεωρηθεί ως έλλειψη ικανοποιητικής πληροφορίας να φθάσουμε σε μια απόφαση
- Η αβεβαιότητα αυξάνει σε μια βάση γεγονότων ή σε μια βάση κανόνων

Αβέβαιη Γνώση

- Κυριότερες πηγές αβεβαιότητας:
 - **Ανακριβή δεδομένα** (*imprecise data*).
 - **Ελλιπή δεδομένα** (*incomplete data*)
 - **Υποκειμενικότητα** ή/και ελλείψεις στην περιγραφή της γνώσης
 - Κάθε είδους **περιορισμοί** που κάνουν το όλο πλαίσιο λήψης απόφασης ατελές.
- Ανάγκη ύπαρξης "μη ακριβών" μεθόδων συλλογισμού.
 - *Θεωρία Πιθανοτήτων*
 - *Συντελεστές Βεβαιότητας (Certainty Factors)*
 - *Θεωρία Dempster-Shafer*
 - *Ασαφής Λογική (Fuzzy Logic)*.

Τύποι σφαλμάτων



Τύποι σφαλμάτων σε γεγονότα

- Πολυσημαντικότητα (ambiguity)
 - Αποτυχία να διακρίνει μεταξύ των περιπτώσεων και πιθανοτήτων
- Ελλιπής (incompleteness)
 - Αποτυχία να περιγράψει πλήρως τις απαραίτητες πληροφορίες
- Εσφαλμένη (incorrectness)
 - Η πληροφορία είναι εσφαλμένη
- Ψευδώς θετική
 - Δέχεται μια υπόθεση ότι είναι αληθινή ενώ δεν είναι αληθής
- Ψευδώς αρνητική
 - Δέχεται μια υπόθεση ότι είναι ψευδής ενώ είναι αληθής

Τύποι σφαλμάτων σε γεγονότα

- Ανακρίβης (imprecision)
 - Όχι αρκετά κοντά στην αλήθεια
- Όχι ακριβής (inaccuracy)
 - Δεν είναι σωστή
- Αναξιόπιστος (unreliability)
 - Μερικές φορές σωστό μερικές όχι
- Τυχαία λάθη (random error)
 - Τα αίτια μπορεί να προβλέπονται ή να είναι άγνωστα
- Συστηματικά λάθη (systematic error)
 - Ένα λάθος εισάγεται συνέχεια από κάποιο bias

Αβεβαιότητα στους κανόνες

Μια βάση κανόνων εμπεριέχει αβεβαιότητα που οφείλεται σε

- Επαγωγικοί κανόνες
 - Ο επαγωγικός συλλογισμός πηγαίνει από το παράδειγμα σε γενικευμένους κανόνες.
- Ευρεστικούς κανόνες (Heuristic rules)
 - Υπάρχουν κανόνες που είναι βασισμένοι στην εμπειρία
- Όχι μονοτονικούς συλλογισμούς (Non-monotonic reasoning)
 - Όταν επιπρόσθετη πληροφορία δείχνει ότι οι προηγούμενοι συλλογισμοί ήταν εσφαλμένοι, όλα τα συμπεράσματα που είχαν εξαχθεί πριν την εύρεση του συγκεκριμένου λάθους θα πρέπει να εξεταστούν ξανά
- Κανόνες που βασίζονται σε λάθη
 - Εάν η βαλβίδα είναι σε καλή κατάσταση, τότε η έξοδος είναι κανονική
 - Η έξοδος είναι κανονική και επομένως η βαλβίδα είναι σε καλή κατάσταση

Πηγές αβεβαιότητας

- Ατελής γνώση της περιοχής εφαρμογής
 - Η υπάρχουσα θεωρία μπορεί να είναι ασαφής και ελλιπής
- Ατελή δεδομένα για κάθε περίπτωση
 - Τα δεδομένα μπορεί να είναι ανακριβή ή αναξιόπιστα
 - Τα δεδομένα μπορεί να είναι προσεγγιστικά
- Οι ειδικοί χρησιμοποιούν μη λεπτομερείς μεθόδους
 - Όταν οι λεπτομερείς μέθοδοι είναι άγνωστες
 - Όταν οι λεπτομερείς μέθοδοι είναι μη πρακτικά εφαρμόσιμες

Πηγές αβεβαιότητας

- Αξιοπιστία της πληροφορίας
- Ελλιπής πληροφορία
- Πληροφορία από πολλές διαφορετικές πηγές
- Μη ακριβής αναπαράσταση γλωσσών

Αντιμετώπιση της αβεβαιότητας

- Ποιοτικές προσεγγίσεις
 - Αντιμετωπίζουν την ελλιπή πληροφορία
 - Χρησιμοποιούν τη λογική ή κανόνες εξαρτώμενους από τα συμφραζόμενα για τη συμβολική αναπαράσταση
- Ποσοτικές προσεγγίσεις
 - Είναι βασισμένες σε αριθμητικές αναπαραστάσεις και χρησιμοποιούν θεωρία πιθανοτήτων και άλλες μεθόδους λογικής για τη διαχείριση της αβεβαιότητας μέσα από διαδικασία συλλογισμού

Αναπαράσταση αβεβαιότητας

- Προβλήματα

- Είναι δύσκολο να ποσοτικοποιήσουμε τις αλληλεπιδράσεις των πιθανοτήτων
- Ο καθορισμός της πιθανότητας ενός γεγονότος απαιτεί πληροφορία που δεν έχουμε
- Είναι δύσκολο να καθορίσουμε επακριβώς τι σημαίνει πολύ, λίγο, πάρα πολύ ψηλός κ.τ.λ.
- Εφαρμόζοντας θεωρία πιθανοτήτων, απαιτούμε από τους μηχανικούς να περιγράψουν με ακρίβεια ποσότητες που δεν έχουν συγκεκριμένη τιμή.
- Οι πιθανότητες είναι δύσκολο να υπολογίζονται, να βελτιώνονται και μπορεί να απαιτούν πολύ δύσκολους υπολογισμούς

Θεωρίες για την αβεβαιότητα

- Κλασσική θεωρία πιθανοτήτων
- Νόμος Bayes
- Shannon θεωρία βασισμένη σε πιθανότητες
- Υποσυνθήκη Πιθανότητες
- Demster-Shafer θεωρία - συναρτήσεις πεποίθησης
- Ασαφής λογική Zadeh

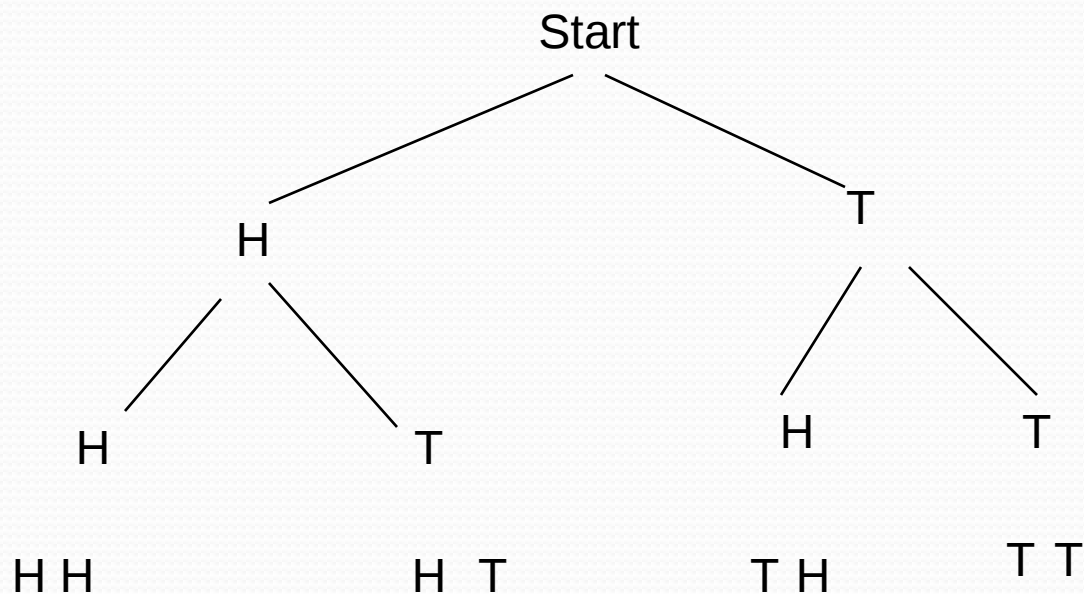
Κλασσική θεωρία πιθανοτήτων

- Αναφέρεται σε ιδανικές καταστάσεις όπως ρίψη ζαριού, νομίσματος και χαρτιών. Όπου υπάρχουν προκαθορισμένες πιθανότητες βασισμένες στην παραδοχή ότι όλοι οι αριθμοί έχουν την ίδια πιθανότητα να συμβούν
- Η πιθανότητα, P , ορίζεται ως ο λόγος των σωστών απαντήσεων προς τον αριθμό των πιθανών αποτελεσμάτων

Πιθανότητα γεγονότων

- Σε ένα παιχνίδι υπάρχει ένα σύνολο πιθανών αποτελεσμάτων το οποίο σχηματίζει το χώρο του παραδείγματος. Για ένα ζάρι ο χώρος είναι $\{1,2,3,4,5,6\}$
- Ρίχνοντας ένα ζάρι (προσπάθεια) δημιουργείται ένα αποτέλεσμα το $\{1\}$. Το οποίο ονομάζεται γεγονός
- Εάν ρίξουμε δύο ζάρια θα έχουμε ένα σύνθετο γεγονός $\{2,5\}$

Παράδειγμα δένδρου γεγονότων



	H	T
H	HH	HT
T	TH	TT

Θεωρία πιθανοτήτων

- Τρία αξιώματα για τις πιθανότητες

$$\text{axiom1: } 0 \leq P(E) \leq 1$$

$$\text{axiom2: } \sum_i P(E_i) = 1$$

and its corollary

$$P(E) + P(E^c) = 1$$

$$\text{axiom3: } P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

Πιθανότητες υπό συνθήκη

- Υποδηλώνει την πιθανότητα να ισχύει το υποθετικό συμπέρασμα H δεδομένης της ισχύος μόνο του γεγονότος E .

Συμβολίζεται με $P(H|E)$ και ορίζεται μέσω του πηλίκου της πιθανότητας να συμβούν ταυτόχρονα τα H και E προς την πιθανότητα του E

$$P(H | E) = \frac{P(H \wedge E)}{P(E)}$$

Ιδιότητες

- *Προσθετική Ιδιότητα:* $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
- *Πολ/στική Ιδιότητα για δύο ανεξάρτητα γεγονότα A και B :*
 $P(A \wedge B) = P(A) \bullet P(B)$
- *Πολ/στική Ιδιότητα για δύο μη ανεξάρτητα γεγονότα A και B :*
 $P(A \wedge B) = P(A) \bullet P(B|A)$

Παράδειγμα

- Έστω ότι έχουμε ένα ζάρι:
- $P(A) = P(\text{περιττός αριθμός}) = 3/6 = 0.5$
 - γιατί υπάρχουν 3 δυνατές τιμές (1,3,5) από σύνολο 6 δυνατών τιμών (1,2,3,4,5,6)
- $P(B) = P(\text{αριθμός} \leq 3) = 3/6 = 0.5$
 - γιατί υπάρχουν 3 δυνατές τιμές (1,2,3) από σύνολο 6 δυνατών τιμών (1,2,3,4,5,6)
- $P(B|A) = P(\text{αριθμός} \leq 3 \text{ δεδομένου ότι είναι περιττός}) = 2/3$
 - γιατί υπάρχουν 2 δυνατές τιμές (1,3) από σύνολο 3 δυνατών τιμών (1,3,5)
- $P(A \cap B) = P(\text{περιττός αριθμός και} \leq 3) = P(A) * P(B|A) = 0.5 * 2/3 = 0.33$
- $P(A \cup B) = P(\text{περιττός ή} \leq 3) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.5 - 0.33 = 0.67$
(προσθετική ιδιότητα)

Θεωρίες για την αβεβαιότητα

- Κλασσική θεωρία πιθανοτήτων
- **Νόμος Bayes**
- Shannon θεωρία βασισμένη σε πιθανότητες
- Υποσυνθήκη Πιθανότητες
- Demster-Shafer θεωρία - συναρτήσεις πεποίθησης
- Ασαφής λογική Zadeh

Ο Νόμος του Bayes (*Bayes' rule*)

- Επιτρέπει τον υπολογισμό πιθανοτήτων υπό συνθήκη με χρήση άλλων πιθανοτήτων που είναι ευκολότερο να υπολογιστούν.
- Χρήση εκτιμήσεων αντί συχνοτήτων εμφάνισης γεγονότων.
- Η απλούστερη εκδοχή του νόμου του Bayes:

$$P(H | E) = \frac{P(E | H) \cdot P(H)}{P(E)}$$

- Πιο εύκολο να χρησιμοποιηθεί, συγκριτικά με την σχέση της πιθανότητας υπό συνθήκη.
- Αν H μία ασθένεια και E ένα σύμπτωμα που σχετίζεται με αυτήν, τότε για τον υπολογισμό της πιθανότητας υπό συνθήκη απαιτείται πληροφορία που συνήθως δεν είναι διαθέσιμη:
 - Πόσοι άνθρωποι στον κόσμο πάσχουν από την H και ταυτόχρονα εμφανίζουν το σύμπτωμα E .
 - Πόσοι εμφανίζουν απλά το σύμπτωμα E .
- Στο νόμο του Bayes:
 - Ένας γιατρός μπορεί να δώσει μία εκτίμηση για το πόσοι ασθενείς που έπασχαν από την ασθένεια H εμφάνιζαν το σύμπτωμα E (ποσότητα $P(E|H)$). Αντίθετα, το κλάσμα των ασθενών με σύμπτωμα E που πάσχουν από την ασθένεια H , δηλαδή ο όρος $P(H|E)$, τις περισσότερες φορές είναι αδύνατο να εκτιμηθεί.
 - Το $P(H)$ μπορεί να υπολογιστεί από στατιστικά στοιχεία για τον συνολικό πληθυσμό.
 - Το $P(E)$ από στατιστικά στοιχεία του ίδιου του γιατρού.

Παράδειγμα 1

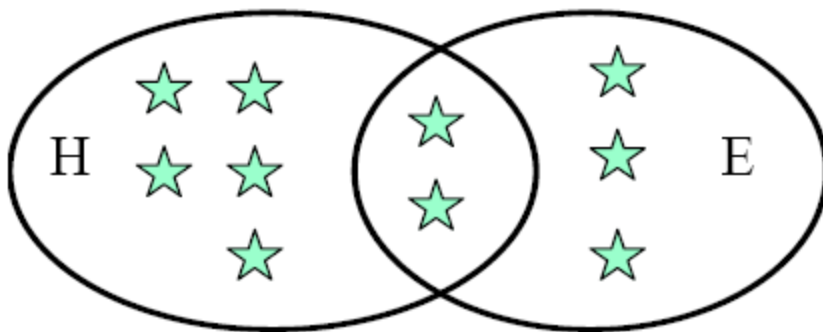
- Έστω τα δύο σύνολα H και E με επτά και πέντε γεγονότα αντίστοιχα από ένα συνολικό πληθυσμό δέκα γεγονότων.
- Το σχήμα μας επιτρέπει να υπολογίσουμε τις απλές (άνευ συνθήκης) και τις υπό συνθήκη πιθανότητες με απλή εφαρμογή του ορισμού τους.

$$P(E) = 5/10 = 0.5$$

$$P(H) = 7/10 = 0.7$$

$$P(H|E) = 2/5 = 0.4$$

$$P(E|H) = 2/7 = 0.2875$$



- Στο παράδειγμα: $P(H|E) * P(E) = P(E|H) * P(H)$
- Γνωρίζοντας τρεις από τις πιθανότητες μπορούμε να υπολογίσουμε την τέταρτη.

Γενική Σχέση του Νόμου του Bayes

- Η πιθανότητα να ισχύει το υποθετικό συμπέρασμα H δεδομένης της ισχύος των γεγονότων $E_1, E_2, \dots, E_k$:

$$P(H | E_1 \wedge E_2 \wedge \dots \wedge E_k) = \frac{P(E_1 \wedge E_2 \wedge \dots \wedge E_k | H) \cdot P(H)}{P(E_1 \wedge E_2 \wedge \dots \wedge E_k)}$$

- **Πρόβλημα χρήσης:** για m πιθανές ασθένειες και n δυνατά συμπτώματα από τα οποία εμφανίζονται τα k , απαιτούνται $(m \cdot n)^k + m + nk$ τιμές πιθανοτήτων, αριθμός υπερβολικά μεγάλος.
- **Απλούστερη περίπτωση:** αν τα διάφορα γεγονότα E θεωρούνται ανεξάρτητα το ένα από το άλλο, τότε απαιτούνται μόνο $m \cdot n + m + n$ τιμές πιθανοτήτων.

Χρήση Θεωρίας Πιθανοτήτων – Συμπεράσματα

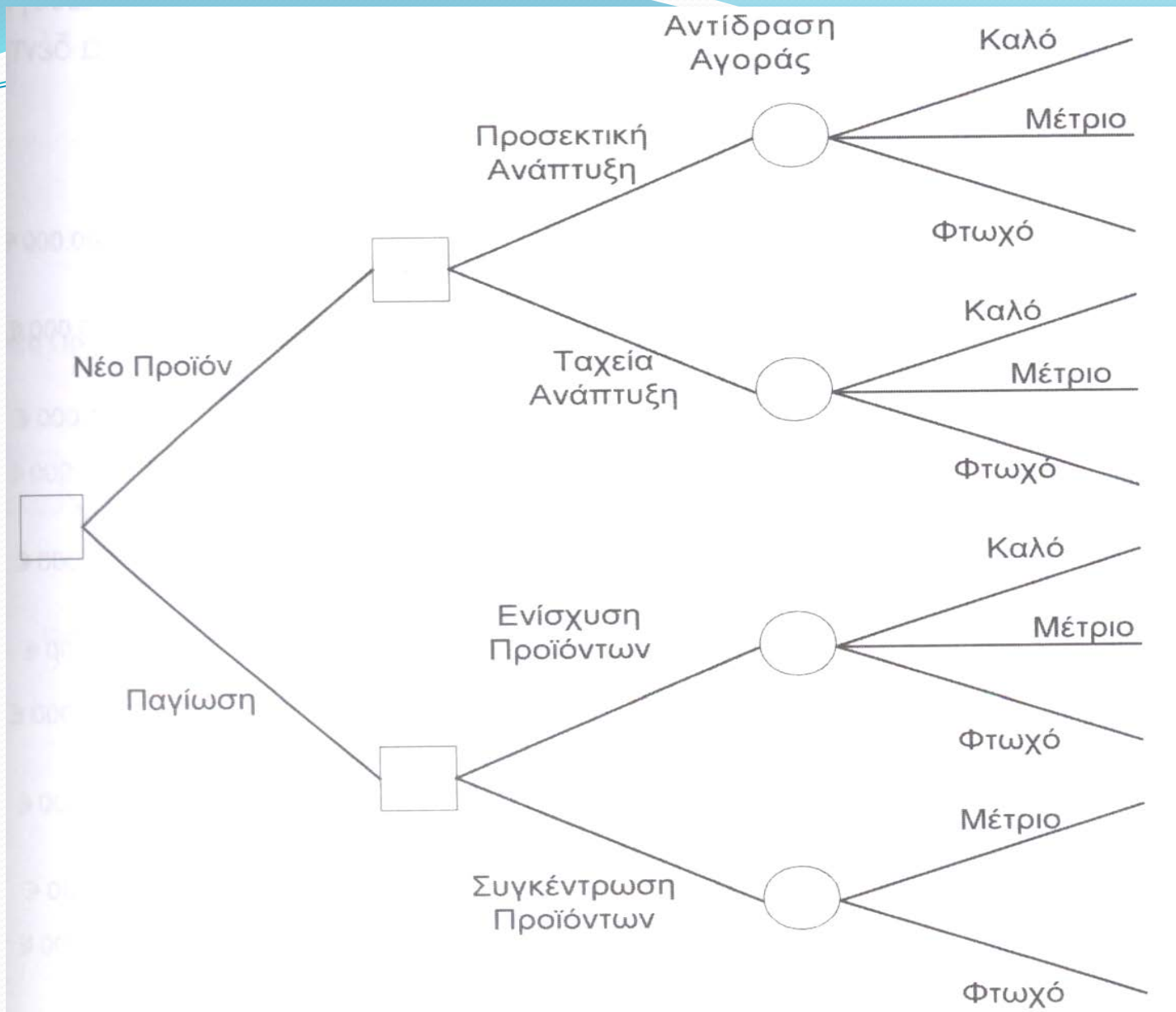
- είτε τα διάφορα γεγονότα θεωρούνται ανεξάρτητα (ευκολότεροι υπολογισμοί σε βάρος της ακρίβειας των συλλογισμών)
- ή καταγράφονται αναλυτικά όλες οι πιθανότητες και οι μεταξύ του συσχετίσεις (ακριβή συμπεράσματα, με υψηλό όμως υπολογιστικό κόστος).
- Εναλλακτική προσέγγιση: Συντελεστές βεβαιότητας.



Δέντρα Απόφασης

Δέντρο απόφασης

- **Δέντρα απόφασης** πολύ καλά εργαλεία για αποφάσεις με πολλές σύνθετες πληροφορίες. Ένα δέντρο απόφασης δέχεται ως είσοδο ένα αντικείμενο ή μια κατάσταση που περιγράφεται από ένα σύνολο ιδιοτήτων και αποτελέσματα μια απόφαση ναι ή όχι. Οι λειτουργίες με ένα μεγαλύτερο εύρος αποτελεσμάτων μπορούν να αντιπροσωπευθούν με δέντρα αποφάσεων.
- **Αποτελεσματική δομή:** τόσο οι εναλλακτικές αποφάσεις όσο και οι επιπτώσεις τους μπορούν να καθοριστούν και να αξιολογηθούν.
- **Ασφαλή διαμόρφωση,** μιας ισορροπημένης εικόνας των κινδύνων και των ανταμοιβών.



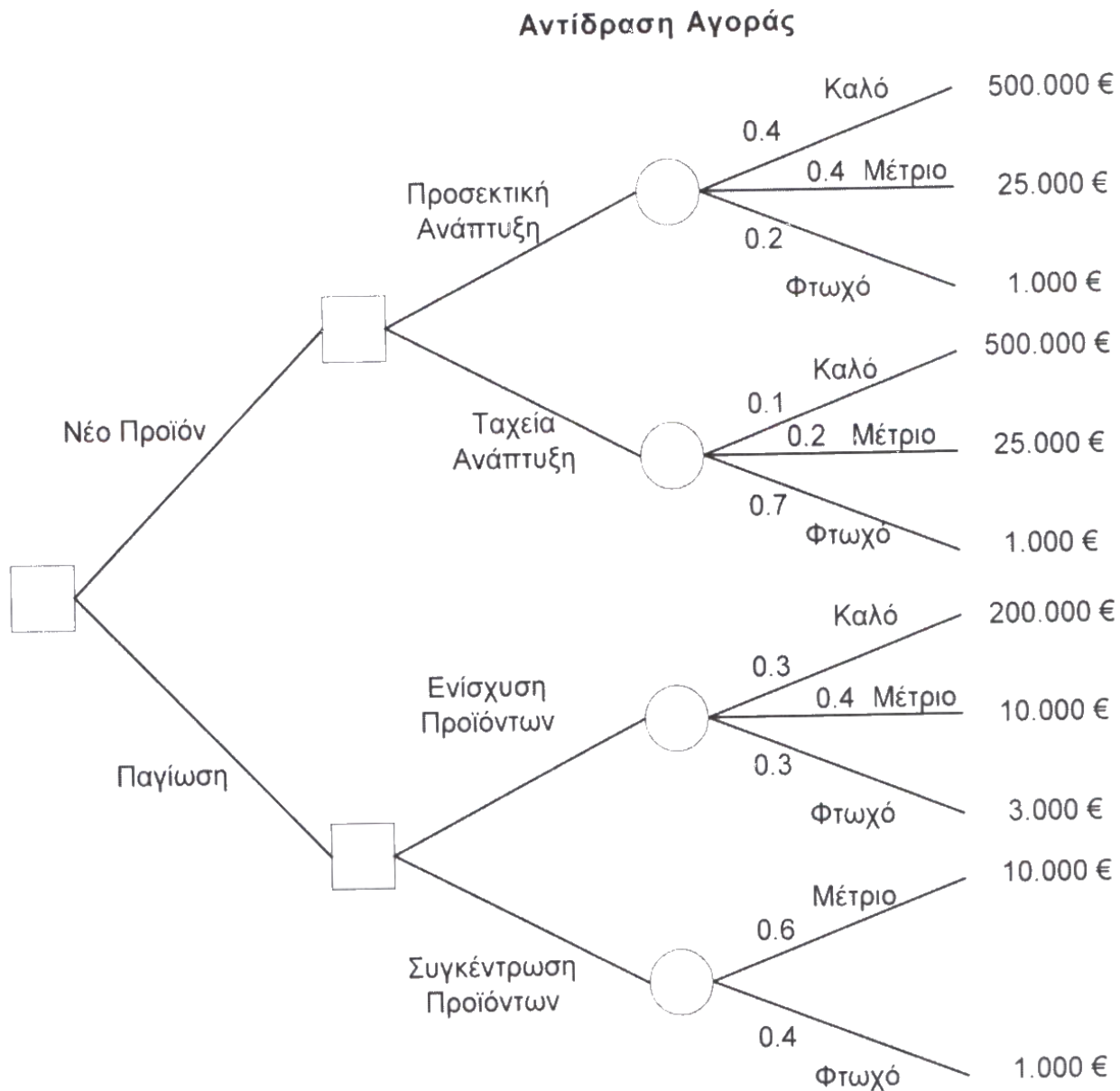
Δέντρο απόφασης για το αν θα πρέπει να αναπτύξουμε νέα προϊόντα ή να εμμένουμε στην υφιστάμενη κατάσταση στην αγορά

Δέντρο απόφασης

- Ένα δέντρο απόφασης ξεκινά με μια απόφαση που πρέπει να ληφθεί.
- Η απόφαση αυτή αναπαριστάται με ένα μικρό τετράγωνο.
- Από το τετράγωνο αυτό τραβάμε γραμμές, μια για κάθε πιθανή λύση και γράφουμε τη λύση αυτή κατά μήκος της γραμμής αυτής.
- Στο τέλος κάθε γραμμής (λύσης), τοποθετούνται τα αποτελέσματα.
- Εάν το αποτέλεσμα της λήψης αυτής της απόφασης είναι αβέβαιο τότε αυτό παρίσταται με έναν μικρό κύκλο.
- Το αποτέλεσμα είναι μια άλλη απόφαση που πρέπει να ληφθεί, αναπαρίσταται με ένα άλλο τετράγωνο.
- Συμβολικά τα τετράγωνα αναπαριστούν αποφάσεις.
- Οι κύκλοι αναπαριστούν αβεβαιότητα ή τυχαίους παράγοντες.
- Εάν στο τέλος μιας γραμμής πρέπει να εμφανισθεί μια λύση, τότε την αφήνουμε κενή.
- Όταν διαμορφώσουμε το δέντρο απόφασης, ελέγχουμε κάθε διαδρομή, τετράγωνο και κύκλο για να δούμε εάν υπάρχουν οποιεσδήποτε λύσεις ή αποτελέσματα που δεν έχουν μελετηθεί. Εάν είναι απαραίτητο, ξανασχεδιάζουμε το δέντρο.

Αξιολόγηση του δέντρου απόφασης

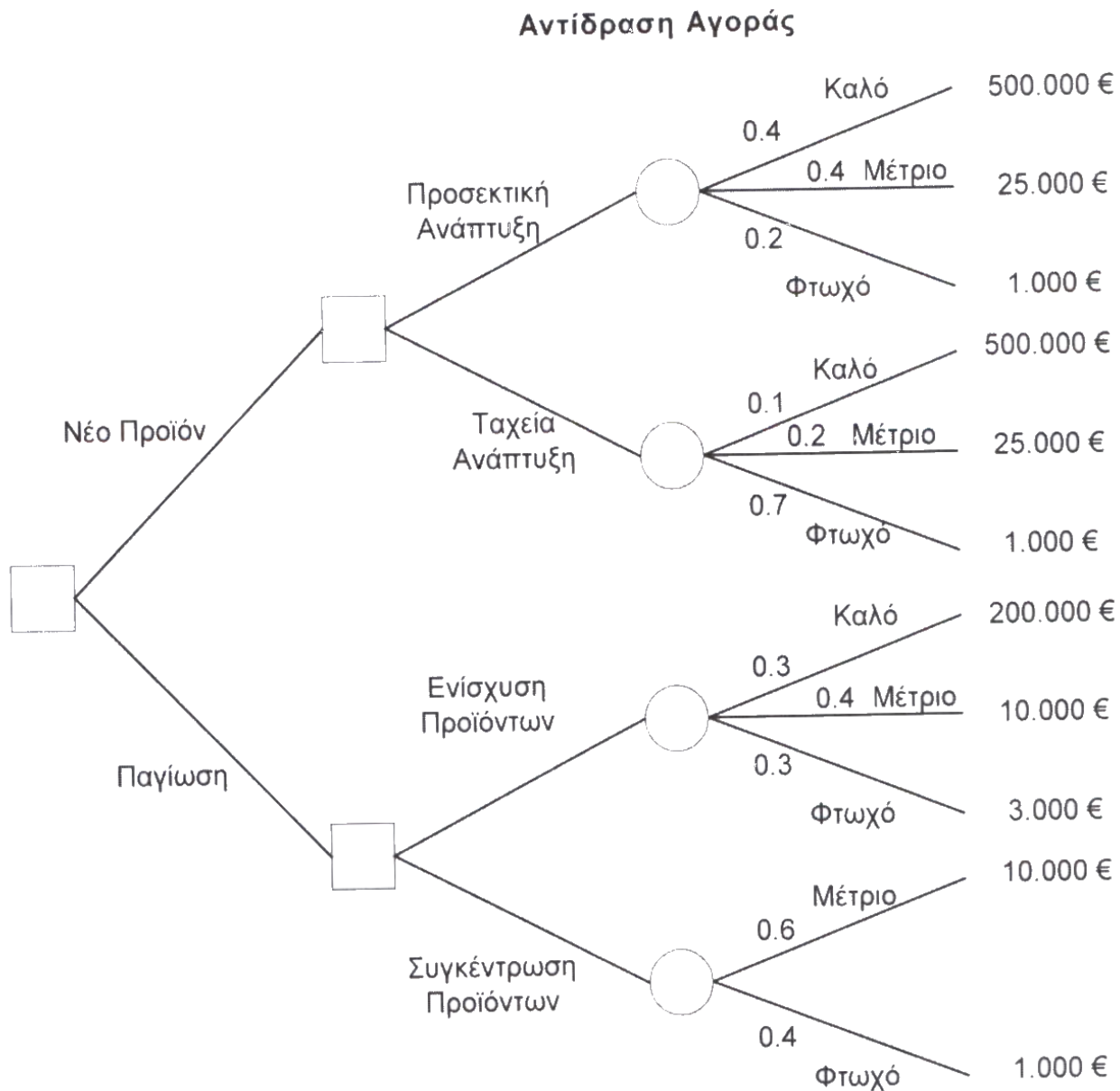
- Υπολογίζουμε την απόφαση που έχει τη μέγιστη αξία για μας. Ανάθεση μετρητών ή μιας αριθμητικής αξίας σε κάθε πιθανό αποτέλεσμα. Πόσο θα άξιζε για μας κάθε πιθανό αποτέλεσμα.
- Εξετάζουμε κάθε κύκλο (σημεία αβεβαιότητας) και υπολογίζουμε την πιθανότητα κάθε αποτελέσματος. Εάν χρησιμοποιήσουμε ποσοστά, τότε τα άθροισμά τους σε κάθε κύκλο θα πρέπει να είναι 100%. Εάν χρησιμοποιείται κλάσματα, τότε αυτά θα πρέπει να αθροίζουν στο 1
- Στοιχεία από παλαιά γεγονότα τότε θα κάνουμε ακριβείς εκτιμήσεις των πιθανοτήτων. Διαφορετικά πρέπει να κάνουμε τις καλύτερες προβλέψεις.



Δέντρο απόφασης ανάπτυξης νέων προϊόντων



Αποδίδουμε τις τιμές των αποτελεσμάτων και καθορίζουμε τις πιθανότητες των αποτελεσμάτων της αβεβαιότητας, τότε υπολογίζουμε τις τιμές που μας βοηθούν να πάρουμε την απόφασή μας.

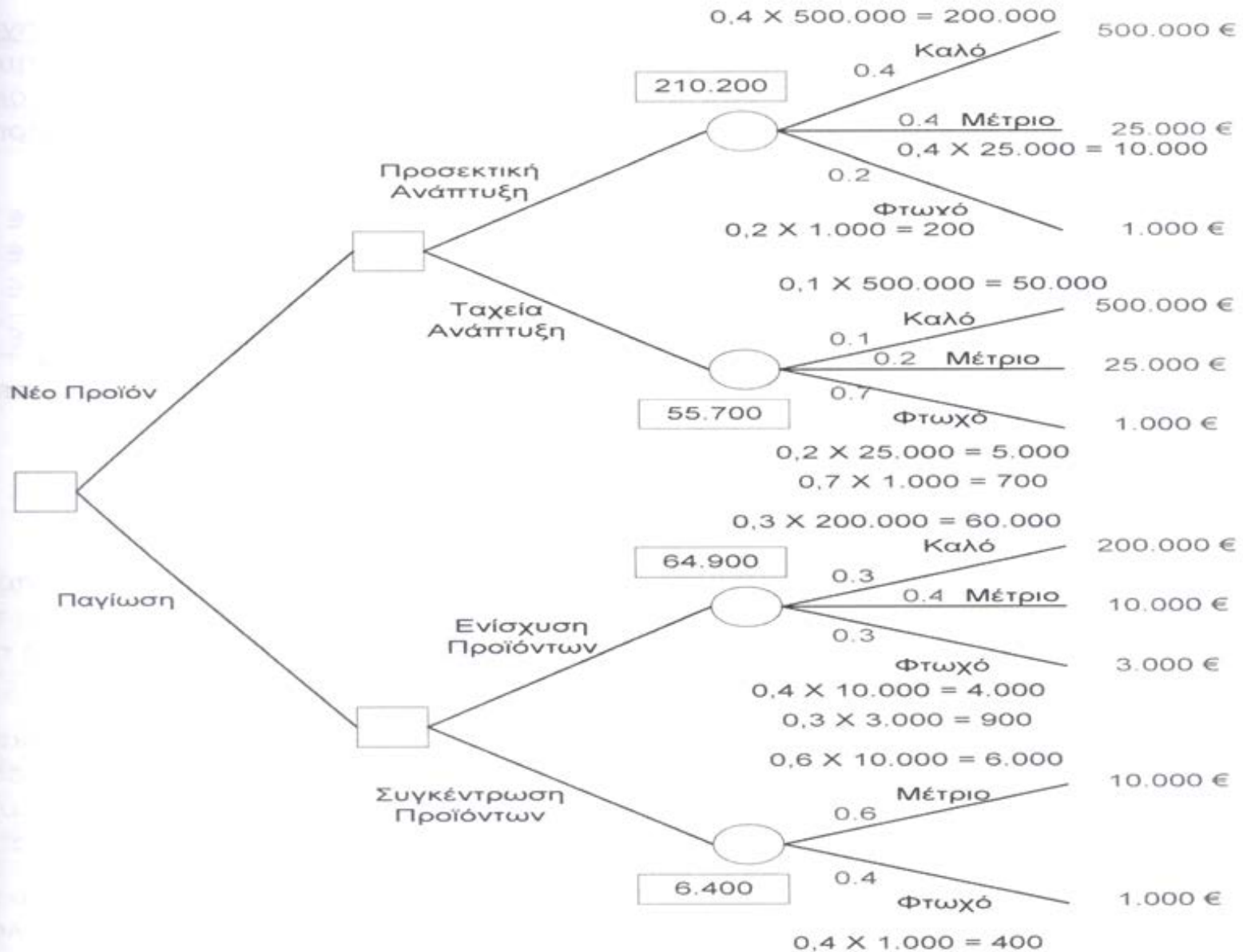


Δέντρο απόφασης ανάπτυξης νέων προϊόντων

Υπολογισμοί στο δέντρο απόφασης

- Δουλεύουμε προς τα πίσω. Μόλις ολοκληρώνουμε ένα σύνολο υπολογισμών, καταγράφουμε το αποτέλεσμα.
- Καθορίζουμε την τιμή των αβέβαιων αποτελεσμάτων (κύκλοι του διαγράμματος), πολλαπλασιάζουμε την τιμή των αποτελεσμάτων με την πιθανότητά τους .
- Η συνολική αξία του κάθε κόμβου του δέντρου προκύπτει με την πρόσθεση όλων αυτών των τιμών.

Αντίδραση Αγοράς

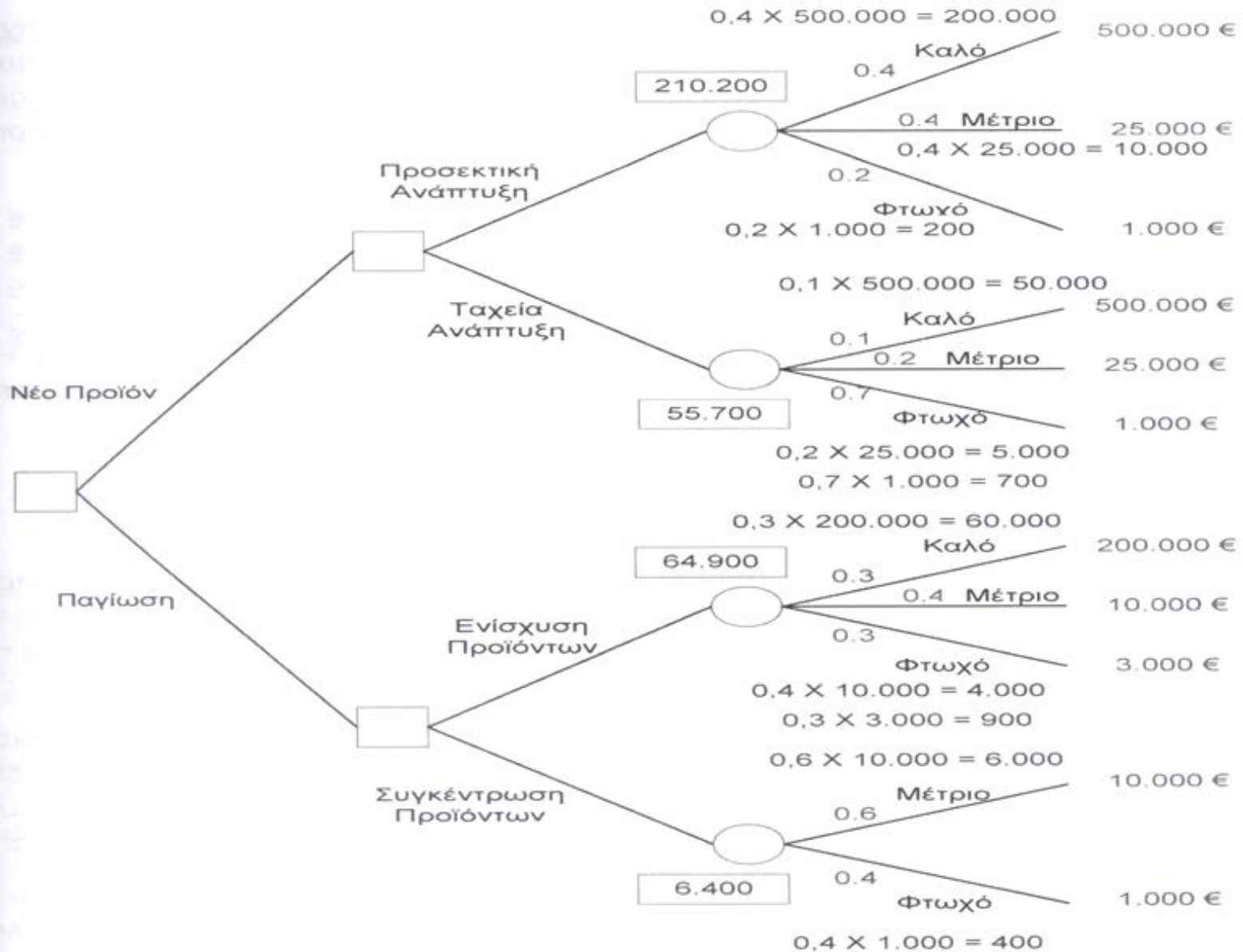


Δέντρο απόφασης ανάπτυξης νέων προϊόντων

Υπολογισμοί στο δέντρο απόφασης

- Η συνολική αξία του κάθε κόμβου του δέντρου προκύπτει με την πρόσθεση όλων αυτών των τιμών.
 - Η τιμή για το “νέο προϊόν – πλήρης ανάπτυξη” υπολογίζεται ως εξής:
 - $0.4 \text{ (πιθανότητα καλού αποτελέσματος)} \times 500,000 \text{ €} = 200,000 \text{ €}$
 - $0.4 \text{ (πιθανότητα μέτριου αποτελέσματος)} \times 25,000 \text{ €} = 10,000 \text{ €}$
 - $0.2 \text{ (πιθανότητα φτωχού αποτελέσματος)} \times 1,000 \text{ €} = 200 \text{ €}$
- 210.200 €

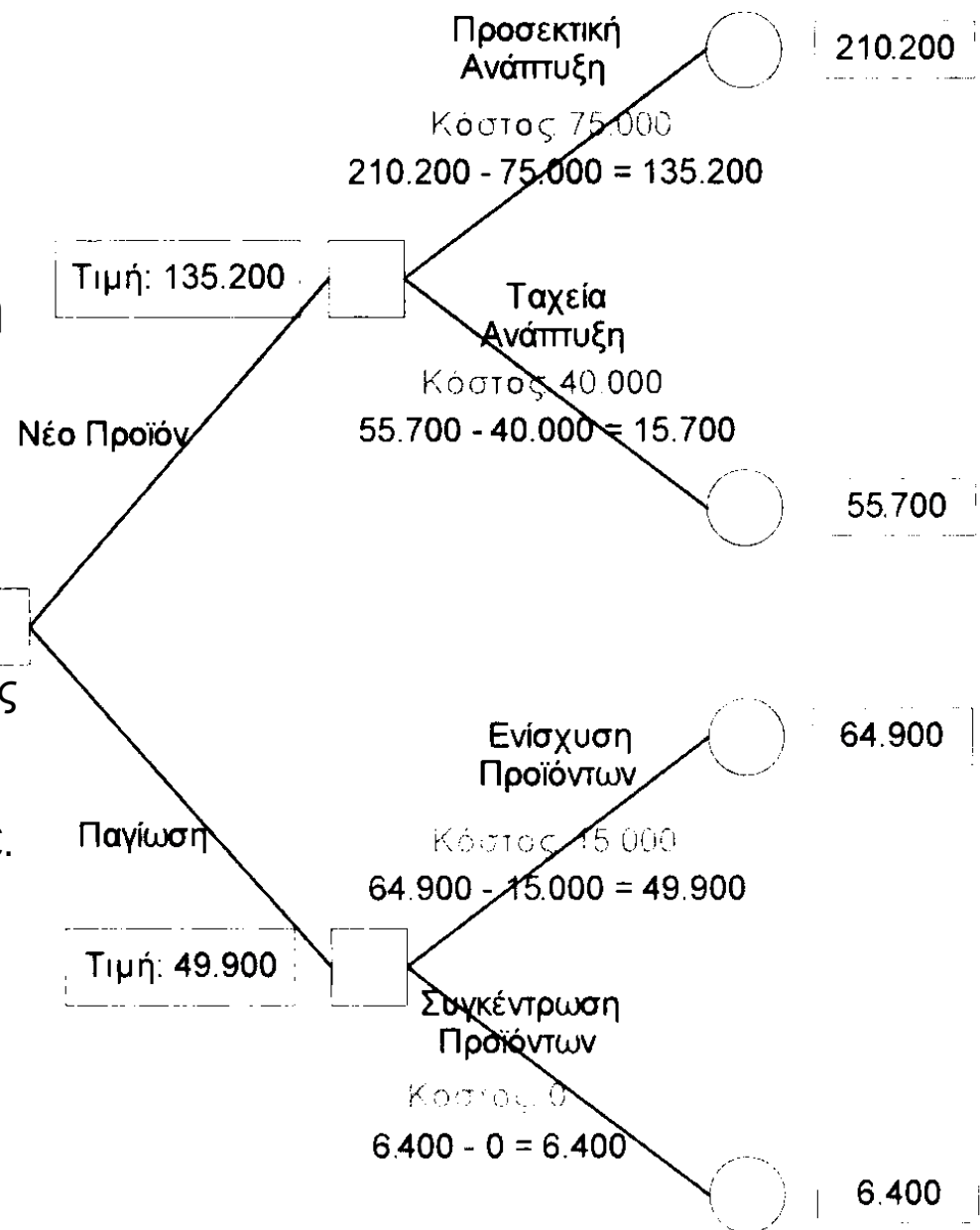
Αντίδραση Αγοράς



Δέντρο απόφασης ανάπτυξης νέων προϊόντων

Αντίδραση Αγοράς

- Όταν αξιολογούμε έναν κόμβο απόφασης, σημειώνουμε σε κάθε γραμμή απόφασης το κόστος της κάθε επιλογής. Αφαιρούμε το κόστος από την τιμή του κάθε αποτελέσματος που έχουμε ήδη υπολογίσει. Αυτό θα μας δώσει μια τιμή η οποία μας δίνει το όφελος αυτής της απόφασης.
- Το κέρδος που υπολογίσαμε για το “νέο προϊόν – προσεκτική ανάπτυξη” ήταν 210,200 €. Υπολογίζουμε το κόστος αυτής της προσέγγισης σε 75.000 € και παίρνουμε ένα καθαρό κέρδος 135.200 €. Το κέρδος του “νέου προϊόντος – ταχεία ανάπτυξη” ήταν 55.700 €. Σε αυτόν τον κλάδο επιλέγουμε την καλύτερη επιλογή, “νέο προϊόν – προσεκτική ανάπτυξη” και αποδίδουμε αυτήν την τιμή στον κόμβο



- 
- Η ανάπτυξη ενός νέου προϊόντος αξίζει τον κόπο και το χρόνο μας για να πάρουμε ένα καλό προϊόν (προσεκτική ανάπτυξη) από το να «σπρώξουμε» (ταχεία ανάπτυξη) το προϊόν στην αγορά.
 - Βελτιώνουμε τα υπάρχοντα προϊόντα μας (ενίσχυση προϊόντων) από το να αναπτύξουμε πρόχειρα (ταχεία ανάπτυξη) ένα νέο προϊόν, ακόμη και αν αυτό έχει μικρότερο κόστος για εμάς.

Δέντρα απόφασης

Τα δέντρα απόφασης παρέχουν μια αποτελεσματική μέθοδο λήψης απόφασης επειδή αυτά:

- παρουσιάζουν σαφώς το πρόβλημα έτσι ώστε όλες οι επιλογές να μπορούν να εμφανισθούν, να συζητηθούν και να συγκριθούν,
- παρέχουν ένα πλαίσιο που ποσοτικοποιεί τις τιμές των αποτελεσμάτων και των πιθανοτήτων επίτευξης των, και
- μας βοηθούν για να πάρουμε τις καλύτερες αποφάσεις επί τη βάση των υπάρχουσών πληροφοριών και των καλύτερων εκτιμήσεών μας.

Η ανάλυση δέντρων απόφασης θα πρέπει να χρησιμοποιηθεί από κοινού με την κοινή αίσθηση-γνώση. Τα δέντρα απόφασης είναι ένα μόνο σημαντικό εργαλείο λήψης αποφάσεων που έχουμε στη διάθεσή μας.



Δέντρα απόφασης

Bayesian Decision Making

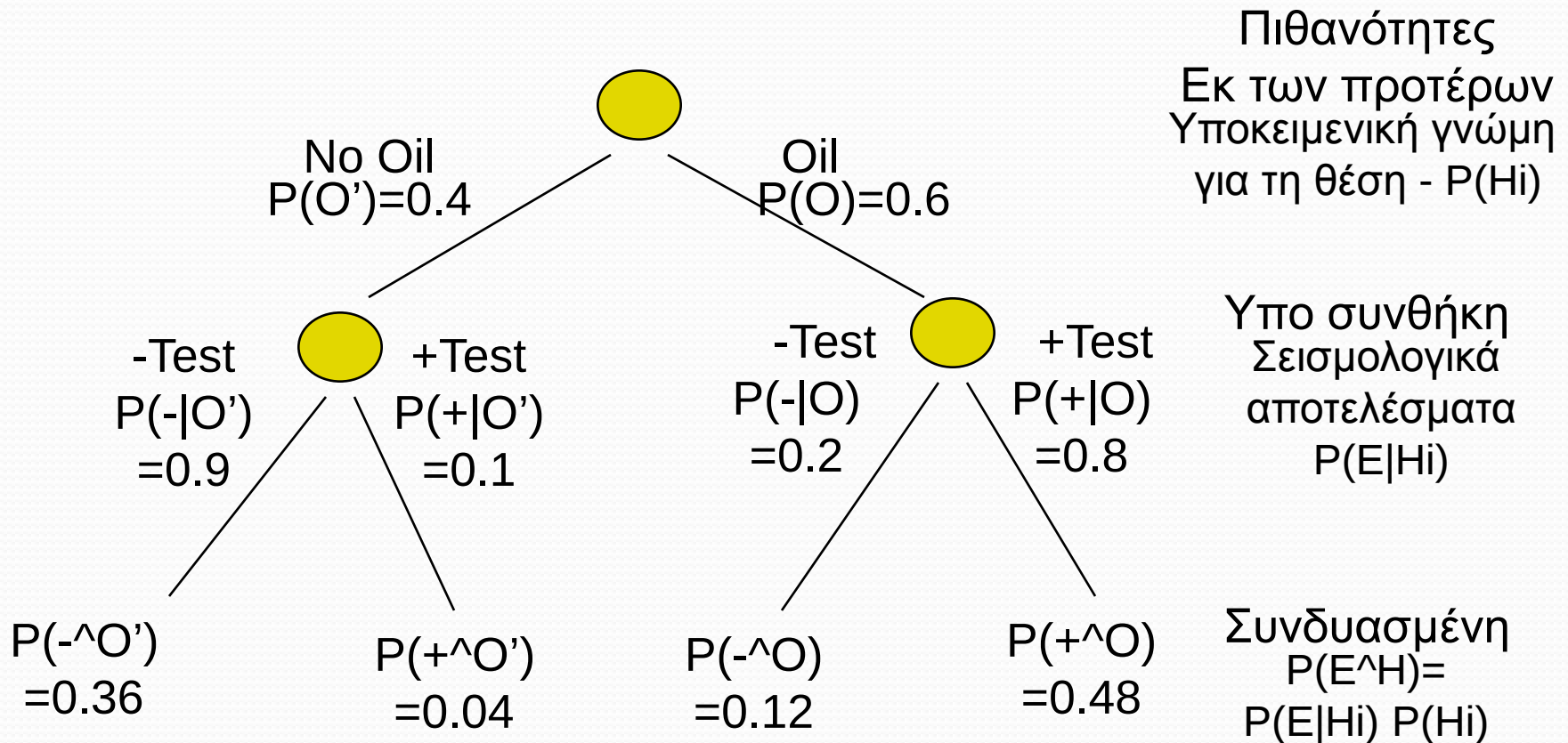
Παράδειγμα Bayesian Decision Making

- Το δένδρο απόφασης είναι ένα παράδειγμα υποθετικού συλλογισμού ή καταστάσεων “Τι Εάν”.
- Περιγράφει τη διαδρομή που οδηγεί στη βέλτιστη στρατηγική
- Περιγράφει πως η αβεβαιότητα αναπαράγεται διαμέσου ενός δένδρου και πως η πιθανότητα σε έναν κόμβο επηρεάζεται από τις πιθανότητες στους άλλους κόμβους

Χρήση του νόμου Bayesian για λήψη αποφάσεων

- Χρησιμοποιείται για ανάλυση δένδρου αποφάσεων στις κοινωνικές επιστήμες και στις επιχειρήσεις
- Το PROSPECTOR σύστημα χρησιμοποιεί Bayesian decision making.
 - Χρησιμοποιήθηκε για να καθορίσει τις κατάλληλες περιοχές για εξόρυξη

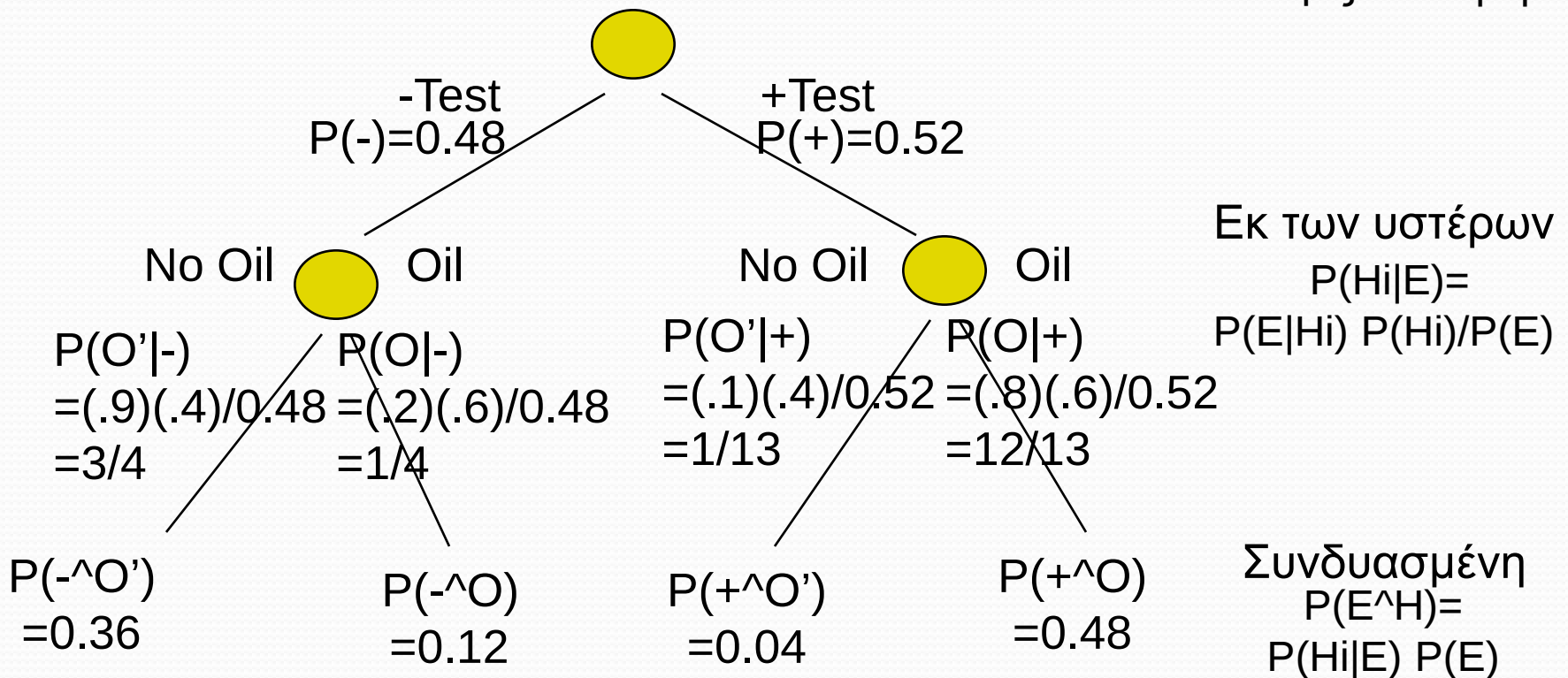
Παράδειγμα Bayesian για λήψη αποφάσεων



Αρχικό δένδρο πιθανοτήτων για εύρεση πετρελαίου

Παράδειγμα Bayesian για λήψη αποφάσεων

Πιθανότητες
Χωρίς συνθήκη

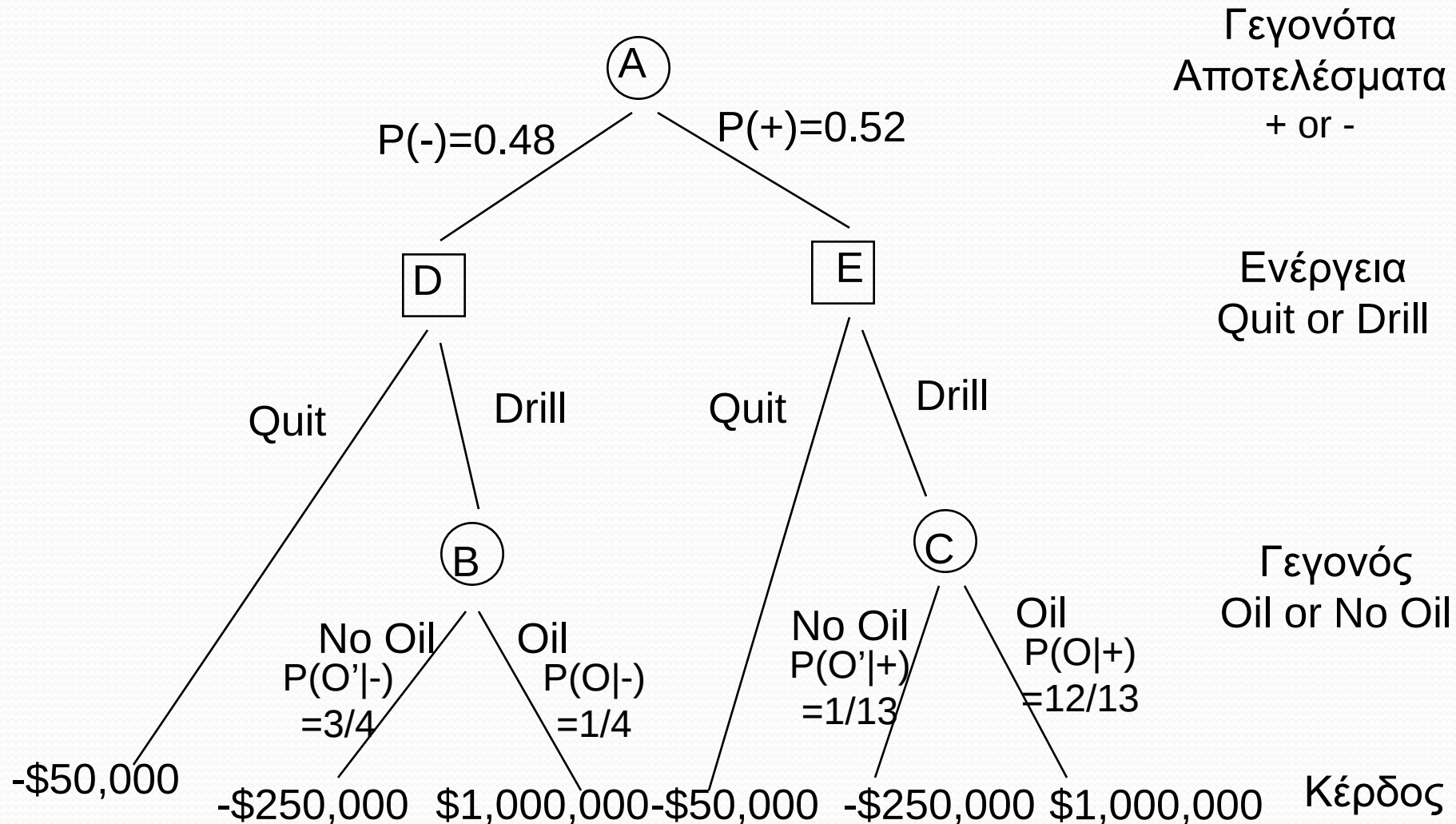


Αναδιατεταγμένο δένδρο πιθανοτήτων για εξερεύνηση πετρελαίου

Παράδειγμα Bayesian για λήψη αποφάσεων

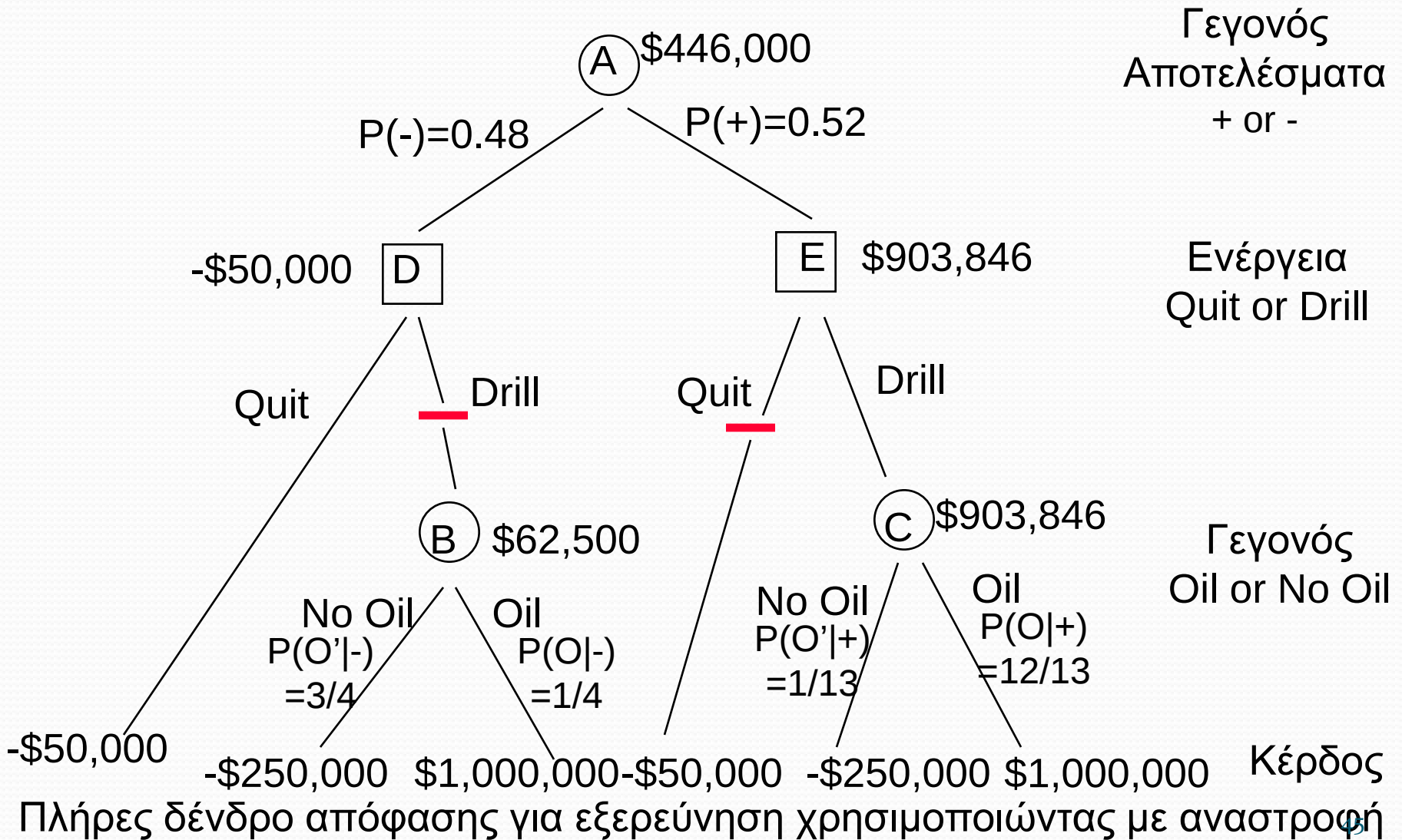
- Εξερεύνηση πετρελαίου
 - πετρέλαιο, επιτυχής \$1,250,000
 - έξοδα γεώτρησης -\$200,000
 - Σεισμολογική έρευνα -\$50,000
- Το αναμενόμενο κέρδος είναι το ποσό που προκύπτει ακολουθώντας την ενδεικνυόμενη ενέργεια
- Εφαρμόζεται ανάστροφη επαγωγή (backward induction).
- Το αναμενόμενο κέρδος από ένα γεγονός-κόμβο είναι το άθροισμα των κερδών επί τις αντίστοιχες πιθανότητες

Παράδειγμα Bayesian Decision Making



Αρχικό δένδρο απόφασης για εξερεύνηση πετρελαίου

Παράδειγμα Bayesian Decision Making





Δίκτυα Πιθανοτήτων **(*Bayesian Probability*** ***Networks*)**

Δίκτυα Πιθανοτήτων (*Bayesian Probability Networks*)

❖ Αντιμετωπίζουν το πρόβλημα της αλληλεπίδρασης των πιθανοτήτων που εμφανίζεται όταν ο χειρισμός της αβεβαιότητας γίνεται αυστηρά κατά Bayes.

❖ Βασίζονται στην παρατήρηση ότι στον πραγματικό κόσμο τα διάφορα γεγονότα δεν αλληλεπιδρούν όλα το ένα με το άλλο αλλά μερικώς. Δηλαδή μπορεί να οριστούν ομάδες από γεγονότα που αλληλεπιδρούν, δηλαδή δεν είναι απαραίτητο να υπολογίζονται οι πιθανότητες όλων των συνδυασμών γεγονότων.

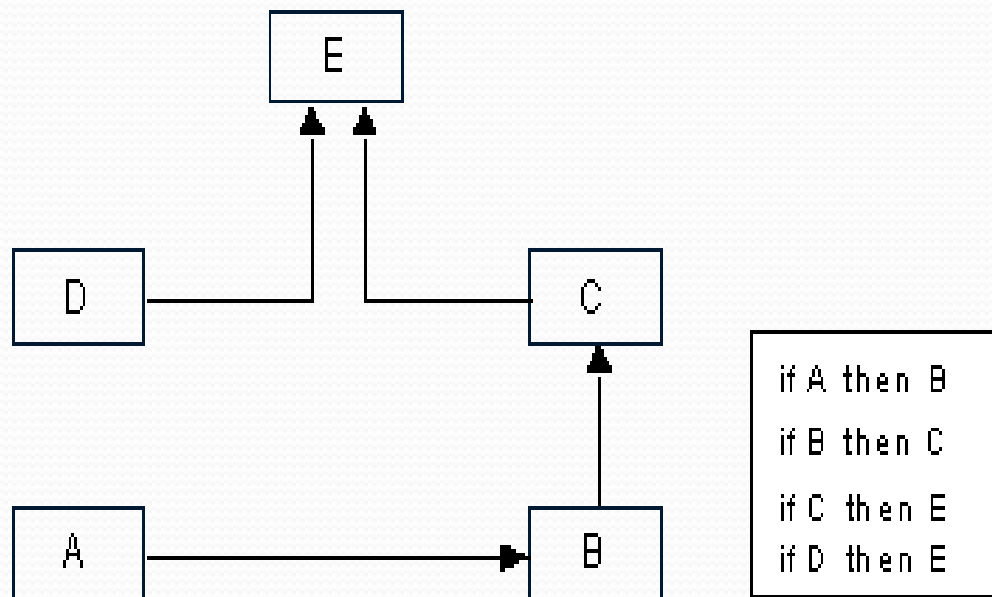
Κανόνες στα δίκτυα πιθανοτήτων: if γεγονός then υποθετικό συμπέρασμα

❑ Διασυνδέονται μεταξύ τους με το υποθετικό συμπέρασμα του ενός να αποτελεί το γεγονός κάποιου άλλου.

❑ Δύο ή περισσότεροι κανόνες μπορεί να καταλήγουν στο ίδιο υποθετικό συμπέρασμα κάτω όμως από διαφορετικές παραδοχές.

❑ Απαγορεύεται η ύπαρξη βρόχων μέσα στο δίκτυο.

Δίκτυα Πιθανοτήτων (*Bayesian Probability Networks*)



Ειδικός μηχανισμός επιτρέπει την αλλαγή της ροής της πληροφορίας (πιθανοτήτων) στο δίκτυο των κανόνων, χωρίς όμως να γίνεται ταυτόχρονη χρήση κανόνων που αντιστρέφουν την σχέση αιτίας-αποτελέσματος.

Δίκτυα Συμπερασμού (Inference Networks)

Παραλλαγή των δικτύων πιθανοτήτων.

Οι κανόνες υποδηλώνουν μία "χαλαρή" συνεπαγωγή:

if γεγονός then υποθετικό συμπέρασμα με βαθμό ισχύος S

Μία περίπτωση υλοποίησης Δικτύων Συμπερασμού είναι με τη χρήση των μεγεθών της λογικής επάρκειας και της λογικής αναγκαιότητας σε συνδυασμό με την ποσότητα εύνοια γεγονότος.

□ **Εύνοια Γεγονότος** (Odds - O):

Εκφράζει τη σύγκριση μεταξύ πιθανοτήτων να συμβεί και να μην συμβεί ένα γεγονός.

$$O(E) = \frac{P(E)}{1 - P(E)}$$

Δίκτυα Συμπερασμού (Inference Networks)

❑ **Λογική Επάρκεια** (Logical Sufficiency – LS):

Εκφράζει πόσο πιθανότερο είναι να συνδεθεί ένα γεγονός E με την αλήθεια ενός υποθετικού συμπεράσματος H , παρά με την άρνηση του H (συμβολίζεται $\neg H$)

$$LS = \frac{P(E | H)}{P(E | \neg H)} = \frac{O(H | E)}{O(H)}$$

❑ **Λογική Αναγκαιότητα** (Logical Necessity – LN)

Εκφράζει το πόσο πιθανότερο είναι να συνδεθεί η απουσία ενός γεγονότος E με την αλήθεια του υποθετικού συμπεράσματος H παρά με την άρνηση του H .

$$LN = \frac{P(\neg E | H)}{P(\neg E | \neg H)} = \frac{O(H | \neg E)}{O(H)}$$

Δίκτυα Συμπερασμού (Inference Networks)

Στα δίκτυα συμπερασμού οι κανόνες περιέχουν και αρχικές τιμές πιθανότητας. Η γενική μορφή τους είναι:

$$E \xrightarrow{(LS, LN)} \begin{matrix} H \\ P_o(H) \end{matrix}$$

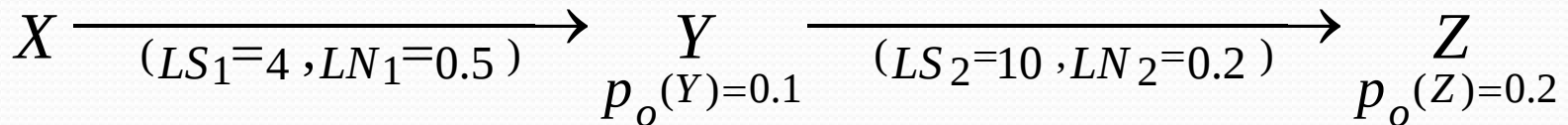
➤ $P_o(H)$ είναι η πιθανότητα να ισχύει το υποθετικό συμπέρασμα H όταν απουσιάζει οποιασδήποτε ένδειξη για την ισχύ του γεγονότος E .

➤ Αν κατά τη διάρκεια του συμπερασμού γίνει γνωστή η ύπαρξη του γεγονότος E τότε και η πιθανότητα του H αλλάζει σε $P(H|E)$.

➤ Ο βαθμός αλλαγής καθορίζεται από το βαθμό ισχύος του κανόνα, δηλαδή τις τιμές LS , LN .

➤ Όλες οι αλλαγές προωθούνται μέσα στο δίκτυο των κανόνων από κόμβο σε κόμβο, ανάλογα και με τις συσχετίσεις που υπάρχουν.

Παράδειγμα Δικτύου Συμπερασμού (Inference Networks)



- $P_o(Y)$ και $P_o(Z)$ είναι οι αρχικές τιμές πιθανότητας για τα Y και Z αντίστοιχα, χωρίς να υπάρχει οποιασδήποτε γνώση για το X .

- Έστω ότι η τιμή του X γίνεται γνωστή. Μετά από πράξεις, η τελική μορφή του δικτύου:

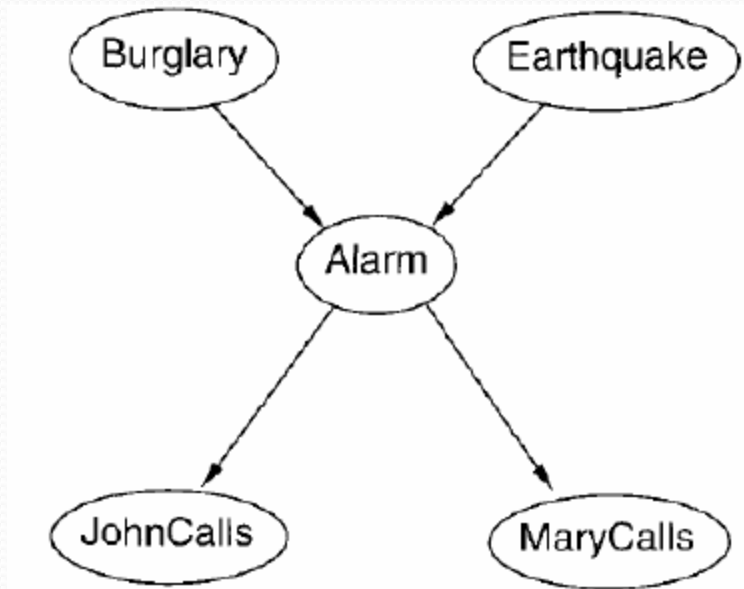


$$O(Y) = \frac{P(Y)}{1 - P(Y)} = \frac{0.1}{1 - 0.1} = 0.111$$

$$O(Y | X) = LS_1 * O(Y) = 4 * 0.111 = 0.444$$

$$P(Y/X) = \frac{O(Y/X)}{1 + O(Y/X)} = \frac{0.444}{1 + 0.444} = 0.307$$

Δίκτυα Πιθανοτήτων Bayes



- You have a new burglar alarm installed at home. It is fairly reliable at detecting a burglary, but also responds on occasion to minor earthquakes.
- You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm. John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary, on the other hand, likes rather loud music and sometimes misses the alarm altogether.

Δίκτυα Πιθανοτήτων Bayes ^{1/2}

- Once we have specified the topology, we need to specify the conditional probability table or CPT for each node. Each row in the table contains the conditional probability of each node value for a conditioning case. A conditioning case is just a possible combination of values for the parent nodes (a miniature atomic event, if you like). For example, the conditional probability table for the random variable Alarm might look like this:

Burglary	Earthquake	$P(\text{Alarm} \text{Burglary}, \text{Earthquake})$	
		<i>True</i>	<i>False</i>
<i>True</i>	<i>True</i>	0.950	0.050
<i>True</i>	<i>False</i>	0.950	0.050
<i>False</i>	<i>True</i>	0.290	0.710
<i>False</i>	<i>False</i>	0.001	0.999

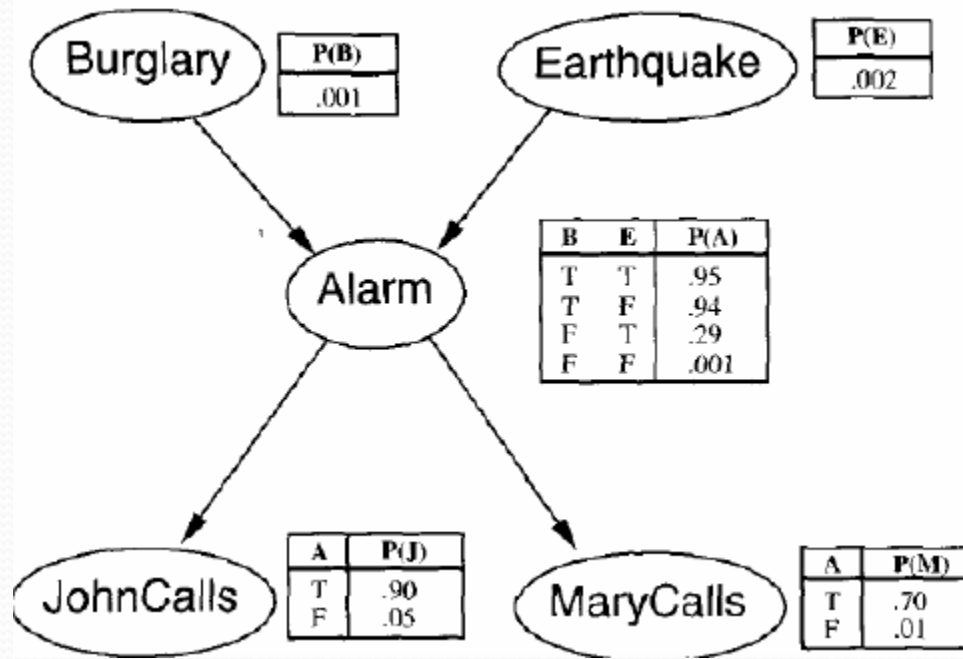
- (Each row in a conditional probability table must sum to 1, because the entries represent an exhaustive set of cases for the variable. Hence only one of the two numbers in each row shown above is independently specifiable. In general, a table for a Boolean variable with n Boolean parents contains 2^n independently specifiable probabilities. A node with no parents has only one row, representing the prior probabilities of each possible value of the variable.)

Δίκτυα Πιθανοτήτων Bayes ^{1/2}

- You have a new burglar alarm installed at home. It is fairly reliable at detecting a burglary, but also responds on occasion to minor earthquakes.
- You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm. John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary, on the other hand, likes rather loud music and sometimes misses the alarm altogether.

Burglary	Earthquake	$P(\text{Alarm} \text{Burglary}, \text{Earthquake})$	
		<i>True</i>	<i>False</i>
<i>True</i>	<i>True</i>	0.950	0.050
<i>True</i>	<i>False</i>	0.950	0.050
<i>False</i>	<i>True</i>	0.290	0.710
<i>False</i>	<i>False</i>	0.001	0.999

Δίκτυα Πιθανοτήτων Bayes 2/2



$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Parents}(X_i))$$

- We can calculate the probability of the event that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both John and Mary call

$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = P(J | A)P(M | A)P(A | \neg B \wedge \neg E)P(\neg B)P(\neg E)$$

$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.00062$$

Προσέγγιση Dempster-Shafer

Βασίζεται σε λογισμό με αριθμητικές τιμές *πεποίθησης* (*belief*), δηλαδή πίστης για την ισχύ κάποιου υποθετικού συμπεράσματος για το οποίο υπάρχουν κάποιες ενδείξεις (γεγονότα).

Δεν απαιτείται η συλλογή όλων των απλών και των υπό συνθήκη πιθανοτήτων.

Προσέγγιση Dempster-Shafer

❖ *Πλαίσιο διάκρισης (frame of discernment)*: το σύνολο U των διακριτών και αμοιβαία αποκλειόμενων προτάσεων ενός τομέα γνώσης. π.χ. αν εξετάζεται η ασθένεια κάποιου, το U αντιπροσωπεύει όλες τις πιθανές διαγνώσεις, δηλαδή τα υποθετικά συμπεράσματα.

❖ $\text{Pow}(U)$: Το σύνολο των υποσυνόλων του U (δυναμοσύνολο)

Αν $U = \{A, B, C\}$ είναι το σύνολο των πιθανών ασθενειών τότε το σύνολο:

$$\text{Pow}(U) = \{ \{\}, \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\} \}$$

υποδηλώνει τις πιθανές διαγνώσεις για μια περίπτωση ασθένειας.

❑ Κάθε στοιχείο του $\text{Pow}(U)$ αντιστοιχεί σε διαζευγμένες προτάσεις. Π.χ. $\{A, B\}$ σημαίνει "ασθένεια A ή B".

❑ Στοιχεία του U που δεν ανήκουν σε ένα στοιχείο του $\text{Pow}(U)$, (π.χ. η ασθένεια C στο $\{A, B\}$), κάνουν σαφή την άρνηση του αντίστοιχου υποθετικού συμπεράσματος.

❑ Το κενό υποσύνολο $\{\}$ αντιστοιχεί στην περίπτωση που όλα τα υποθετικά συμπεράσματα είναι ψευδή (*null hypothesis*).

Προσέγγιση Dempster-Shafer (D-S) _{1/2}

- Δεν απαιτείται η συλλογή όλων των απλών και των υπό συνθήκη πιθανοτήτων.
- Λογισμός με αριθμητικές τιμές πεποίθησης (*belief*)
- Πλαίσιο διάκρισης (*frame of discernment*).
- Αν $U = \{A, B, C\}$ πιθανές ασθένειες τότε Pow είναι το σύνολο των υποσυνόλων του U :
 $Pow(U) = \{ \{\}, \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\} \}$ - πιθανές διαγνώσεις.
- Διαζευγμένες Προτάσεις: $\{A, B\}$ σημαίνει "ασθένεια A ή B".
- Στοιχεία του U που δεν ανήκουν σε ένα στοιχείο του $Pow(U)$, (π.χ. η ασθένεια C στο $\{A, B\}$), κάνουν σαφή την άρνηση του αντίστοιχου υποθετικού συμπεράσματος.
- $\{\}$: *null hypothesis*

Προσέγγιση Dempster-Shafer ^{2/2}

- Η βασική κατανομή πιθανότητας (*basic probability assignment - bpa*) είναι μία απεικόνιση: $m: \text{Pow}(U) \rightarrow [0,1]$ η οποία αναθέτει μία τιμή πεποίθησης για κάθε στοιχείο του $\text{Pow}(U)$
- Είναι δηλαδή το μέτρο της πεποίθησης που υπάρχει για το κατά πόσο ισχύει το υποθετικό συμπέρασμα που εκφράζεται με το συγκεκριμένο στοιχείο του U .
 - η πεποίθηση $m(\{A, B\})=0.3$, δε μοιράζεται στα $\{A\}$ και $\{B\}$ αλλά αφορά το $\{A, B\}$.
 - ισχύει $m(\{\})=0$

$$\sum_{X \in \text{Pow}(U)} m(X) = 1$$

- Επειδή θεωρούμε ότι το υποθετικό συμπέρασμα βρίσκεται κάπου μέσα στα στοιχεία του $\text{Pow}(U)$
- Η συνολική πεποίθηση (*belief*) ότι ένα στοιχείο του U ανήκει στο X καθώς και στα τυχόν υποσύνολα του X συμβολίζεται με $Bel(X)$

$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

Dempster-Shafer vs. Bayes

- **Bayes:** η απουσία άλλων ενδείξεων για τις δυνατές εκδοχές τις καθιστά ισοπίθανες.
- **Dempster-Shafer:** η απουσία κάποιων ενδείξεων θέτει την πιθανότητα (*likelihood*) κάθε εκδοχής κάπου στο διάστημα $[0, 1]$.

Κανόνας Dempster-Shafer

- Αν m_1 και m_2 δύο ανεξάρτητες εκτιμήσεις (βασικές κατανομές πιθανότητας) που αποδίδουν κάποιο βαθμό πεποίθησης στα στοιχεία του $\text{Pow}(U)$, τότε αυτές συνδυάζονται σε μία τρίτη εκτίμηση $m_3 = m_1 \oplus m_2$ με τρόπο που ορίζεται με τον κανόνα D-S:

$$m_3(A) = m_1 \oplus m_2(A) = \frac{\sum_{X,Y \in U: X \cap Y = A} m_1(X) \cdot m_2(Y)}{1 - \sum_{X,Y \in U: X \cap Y = \emptyset} m_1(X) \cdot m_2(Y)}$$

$A \in \text{Pow}(U)$

Παράδειγμα: Διάγνωση Ασθένειας

- Έστω $U = \{A, B, C\}$ το σύνολο των δυνατών ασθενειών που μπορεί να διαγνωσθούν.
- Πιθανές Διαγνώσεις $Pow(U) = \{\{\}, \{A\}, \{B\}, \{C\}, \{A,B\}, \{A,C\}, \{B,C\}, \{A,B,C\}\}$
- $m(\{\{\}, \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\}\}) = 1$
 - υποδηλώνει τη βεβαιότητα ότι η διάγνωση βρίσκεται κάπου στα στοιχεία του $Pow(U)$ αλλά ελλείπει άλλων ενδείξεων δεν είναι δυνατό να δοθεί ιδιαίτερη βαρύτητα σε κάποιο
 - Bayes: θα έπρεπε κάθε στοιχείο του $Pow(U)$ να θεωρηθεί ισοπίθανο
- Έστω ότι γίνεται διαθέσιμη επιπλέον πληροφορία, (π.χ. πραγματοποιούνται ιατρικές εξετάσεις) και προκύπτει ότι η ασθένεια είναι μία από τις A ή B με βαθμό πίστης 0.7
 - $m1(\{A, B\}) = 0.7$
 - $m1(\{\{\}, \{A\}, \{B\}, \{C\}, \{A, C\}, \{B, C\}, \{A, B, C\}\}) = 0.3$
 - Δηλαδή, η έλλειψη πίστης σε ένα από τα υποθετικά συμπεράσματα του $Pow(U)$, ισοδυναμεί αυτόματα με ισόποσο βαθμό πίστης στα υπόλοιπα στοιχεία του $Pow(U)$, χωρίς όμως να δίνεται ιδιαίτερη προτίμηση σε κάποιο από αυτά.
 - Bayes: απαιτείται ο υπολογισμός μεγάλου αριθμού υπό συνθήκη πιθανοτήτων, κάτι που είναι υπολογιστικά ακριβό και πολλές φορές αδύνατο.
- Πώς μπορεί να συνδυαστούν δύο ανεξάρτητες εκτιμήσεις (π.χ. δύο ιατρών) σε μία;

Παράδειγμα: Συνδυασμός Διαγνώσεων

- Εστω ότι δύο γιατροί εξετάζουν ανεξάρτητα τον ασθενή και δίνουν την εκτίμησή τους m_1 και m_2 αντίστοιχα, για την αρρώστια από την οποία αυτός πάσχει.

	Γιατρός 1		Γιατρός 2	
Δυνατές περιπτώσεις διάγνωσης	m_1	Bel_1	m_2	Bel_2
{A}	0.05	0.05	0.15	0.15
{B}	0	0	0	0
{C}	0.05	0.05	0.05	0.05
{A, B}	0.15	0.2	0.05	0.2
{A, C}	0.1	0.2	0.2	0.4
{B, C}	0.05	0.1	0.05	0.1
{A, B, C}	0.6	1	0.5	1

$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

- π.χ. $Bel_1(\{A,B\}) = m_1(\{A,B\}) + m_1(\{A\}) + m_1(\{B\}) = 0.015 + 0.05 + 0 = 0.2$
- Οι δύο ανεξάρτητες εκτιμήσεις m_1 και m_2 μπορεί να συνδυαστούν στην m_3 χρησιμοποιώντας τον κανόνα Dempster-Shafer:

$$m_3(A) = m_1 \oplus m_2(A) = \frac{\sum_{X,Y \in Pow(U): X \cap Y = A} m_1(X) \cdot m_2(Y)}{1 - \sum_{X,Y \in Pow(U): X \cap Y = \emptyset} m_1(X) \cdot m_2(Y)}$$

Παράδειγμα: Συνδυασμός Εκτιμήσεων _{1/2}

$m_3 = m_1 \oplus m_2$		m_1						
		{A}	{B}	{C}	{A,B}	{A,C}	{B,C}	{A,B,C}
m_2		0.05	0	0.05	0.15	0.1	0.05	0.6
{A}	0.15	{A} .0075	{ } 0	{ } .0075	{A} .0225	{A} .015	{ } .0075	{A} .09
{B}	0	{ } .0	{B} 0	{ } .0	{B} .0	{ } .0	{B} .0	{B} .0
{C}	0.05	{ } .0025	{ } 0	{C} .0025	{ } .0075	{C} .005	{C} .0025	{C} .03
{A,B}	0.05	{A} .0025	{B} 0	{ } .0025	{A,B} .0075	{A} .005	{B} .0025	{A,B} .03
{A,C}	0.2	{A} .01	{ } 0	{C} .01	{A} .03	{A,C} .02	{C} .01	{A,C} .012
{B,C}	0.05	{ } .0025	{B} 0	{C} .0025	{B} .0075	{C} .005	{B,C} .0025	{B,C} .03
{A,B,C}	0.5	{A} .025	{B} 0	{C} .025	{A,B} .075	{A,C} .05	{B,C} .025	{A,B,C} .3

Παράδειγμα: Συνδυασμός Εκτιμήσεων _{2/2}

Δυνατές περιπτώσεις διάγνωσης	m_3	Bel_3
{A}	0.21	0.21
{B}	0.01	0.01
{C}	0.09	0.09
{A, B}	0.12	0.34
{A, C}	0.20	0.50
{B, C}	0.06	0.16
{A, B, C}	0.31	1.00

$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

- Η αρχική εκτίμηση ότι η ασθένεια είναι μία από τις A ή B αποδυναμώθηκε.
 - η διάγνωση βρίσκεται μάλλον στο σύνολο $\{A, C\}$
 - επειδή $Bel_3(\{A\}) > Bel_3(\{C\})$, αρχίζει να διαφαίνεται ότι η τελική διάγνωση είναι η A
- Η παραπάνω συνδυασμένη εκτίμηση μπορεί να συνδυαστεί εκ νέου με μια άλλη εκτίμηση (π.χ. 3^{ου} ιατρού).