

# **Αυτόματη κατηγοριοποίηση στρατηγικών επίλυσης προβλημάτων από μαθητές με χρήση Δικτύων Bayes.**

Ν. Τσέλιος<sup>1</sup>, Μ. Μαραγκουδάκης<sup>1</sup>, Ν. Αβούρης<sup>1</sup>, Ν. Φακωτάκης<sup>1</sup>, Μ. Κορδάκη<sup>2</sup>

<sup>1</sup> Πανεπιστήμιο Πατρών, Τμήμα Ηλεκτρολόγων Μηχανικών και  
Τεχνολογίας Υπολογιστών, 265 00 Ρίο Πάτρα

<sup>2</sup> Πανεπιστήμιο Πατρών, Τμήμα Μηχανικών  
Ηλ.Υπολογιστών και Πληροφορικής, 265 00 Ρίο Πάτρα

**Υποβάλλον:** Νίκος Αβούρης, Αναπληρωτής Καθηγητής, Τμήμα Ηλεκτρολόγων Μηχανικών και  
Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πατρών, 26500 Ρίο Πάτρα ([N.Avouris@ee.upatras.gr](mailto:N.Avouris@ee.upatras.gr))

**Λέξεις κλειδιά:** έμπειρα συστήματα, αυτόματη καταγραφή στρατηγικών επίλυσης προβλημάτων,  
ανοικτά συστήματα μάθησης με χρήση υπολογιστή, δίκτυα Bayes,

**Επίπεδο εκπαίδευσης:** Η προτεινόμενη τεχνική είναι ανεξάρτητη βαθμίδας εκπαίδευσης, το  
πείραμα που παρουσιάζεται αφορά στο Γυμνάσιο

**Κατηγορία εργασίας:** Εμπειρική πειραματική έρευνα – ανάπτυξη πρωτότυπης τεχνικής/  
συστήματος

# Αυτόματη κατηγοριοποίηση στρατηγικών επίλυσης προβλημάτων από μαθητές με χρήση Δικτύων Bayes.

## Περίληψη

Στα σύγχρονα ανοικτά περιβάλλοντα μάθησης υποστηριζόμενης από υπολογιστή οι στρατηγικές επίλυσης προβλημάτων τις οποίες οι μαθητές μπορούν να ακολουθήσουν είναι πολλαπλές και όχι πάντα εύκολα αναγνωρίσιμες. Η καταγραφή της στρατηγικής επίλυσης που ο κάθε μαθητής ακολουθεί είναι σημαντικό παιδαγωγικό εργαλείο ανάλυσης, το οποίο μπορεί να χρησιμοποιηθεί από εκπαιδευτικούς και ερευνητές για βέλτιστη σχεδίαση διδακτικών εργαλείων και διδακτικών στρατηγικών. Το πρόβλημα της καταγραφής στρατηγικών εμφανίζεται ιδιαίτερα οξύμενο στην περίπτωση μελετών πεδίου όπου καταγράφεται αυτόματα η αλληλεπίδραση του μαθητή με τον υπολογιστή. Η εκ των υστέρων ανάλυση των δεδομένων που συλλέγονται και η καταγραφή των στρατηγικών είναι μια επίπονη και χρονοβόρα διαδικασία. Η αυτοματοποίηση της διαδικασίας αυτής είναι το αντικείμενο της εργασίας αυτής. Στη παρούσα εργασία προτείνεται η εφαρμογή μιας πιθανοτικής μεθόδου κατηγοριοποίησης στρατηγικών, με χρήση δικτύων Bayes, με στόχο τη συσχέτιση των καταγεγραμμένων δράσεων του μαθητή με προκαθορισμένες στρατηγικές. Για την εφαρμογή αυτής της μεθόδου έχει κατασκευαστεί κατάλληλο εργαλείο επεξεργασίας των δεδομένων που διευκολύνει και επιταχύνει την διαδικασία κατασκευής του πιθανοτικού δικτύου. Η μέθοδος αυτή εφαρμόστηκε σε δεδομένα που συλλεχτήκαν από την καταγραφή της αλληλεπίδρασης μαθητών πρώτων τάξεων γυμνασίου με ένα ανοικτό περιβάλλον μάθησης εστιασμένο στη κατανόηση γεωμετρικών εννοιών.

## Abstract\*

In computer-supported open learning environments the problem solving strategies used by the students are particularly difficult to detect. This is more an issue in field studies where student interaction with the computing environment is logged. Posteriori analysis of these logging data is a tedious process. Automation of this process is the subject of the research reported here. Problem solving strategies are diagnosed automatically using a Bayesian network that is able to relate student activity to pre-defined strategies. Tools have been built to support this process. A case study involving data from student interaction with a geometric concepts learning environment is also included.

## 1.Εισαγωγή

Τα περιβάλλοντα μάθησης τα οποία σχεδιάζονται με βάση κοινωνικές και εποικοδομιστικές θεωρήσεις για τη γνώση και τη μάθηση είναι κυρίως ανοικτά περιβάλλοντα (Bauersfeld, H. 1988; Confrey, J. 1995). Χαρακτηριστικά αναφέρονται το περιβάλλον της γλώσσας Logo, το περιβάλλον της δυναμικής Γεωμετρίας Cabri-Geometry (Laborde, 1990), όπως και το περιβάλλον ημι-ποσοτικών μοντέλων Models Creator (Dimitracopoulou, et al, 1999). Τα περιβάλλοντα αυτά τα οποία υπό προϋποθέσεις ορίζονται και ως μικρόκοσμοι (Papert, 1980) αποτελούνται: από ένα σύνολο από πρωταρχικά αντικείμενα και βασικές λειτουργίες που επιδρούν σε αυτά όπως και ένα

---

\* *Automatic diagnosis of student problem solving strategies using Bayesian Networks*

σύνολο από κανόνες που διέπουν αυτή την επίδραση, τα οποία σχετίζονται με τη συνήθη δομή ενός τυπικού συστήματος.

Η πλήρης ενσωμάτωση τέτοιων περιβαλλόντων στην εκπαιδευτική διαδικασία παρουσιάζει εγγενείς δυσκολίες από την άποψη του ότι η μαθησιακή πορεία του μαθητή δεν μπορεί να αναγνωρισθεί πλήρως, ιδιαίτερα σε περίπτωση που η αναγνώριση αυτή πρέπει να γίνει εκ των υστέρων από τον εκπαιδευτικό ή τον ερευνητή χωρίς την παρουσία του μαθητή. Στην περίπτωση αυτή η συνήθης πηγή αποτύπωσης της δραστηριότητας του μαθητή είναι ένα αρχείο καταγραφής πληκτρολογήσεων (log file). Η διαδικασία εξαγωγής συμπερασμάτων από τα αρχεία αυτά με μεθόδους όπως η μοντελοποίηση των γνωστικών στόχων του μαθητή μέσω της ανασύνθεσης υψηλότερου ιεραρχικά επιπέδου διεργασιών (Kordaki & Avouris 2000, Tselios et al.2001) αποτελεί μια πρόταση για την αντιμετώπιση του προβλήματος αυτού. Η μέθοδος αυτή όμως απαιτεί συστηματική ανάλυση όλων των αλληλεπιδράσεων καθώς και την εμπλοκή πεπειραμένων αξιολογητών στην παραγωγή των γνωστικών μοντέλων των μαθητών. Κατά συνέπεια θα ήταν επιθυμητή η κατασκευή ενός αυτόματου εργαλείου καταγραφής των ακολουθούμενων στρατηγικών, με σχετικά υψηλό βαθμό ακριβείας. Η ανάλυση τέτοιων αποτελεσμάτων μπορεί να προσφέρει τόσο στο τομέα της σχεδίασης και αξιολόγησης του περιβάλλοντος μάθησης, όσο και στην αξιολόγηση του εκπαιδευτικού αποτελέσματος και κατάλληλης υποστήριξης των μαθητών.

Η έρευνα που περιγράφεται στην εργασία αυτή, εστιάζει στη διαδικασία που ακολουθήθηκε για τη κατασκευή ενός τέτοιου αυτόματου εργαλείου. Το μοντέλο που δημιουργήθηκε και περιγράφεται στη συνέχεια βασίστηκε στη χρήση *Δικτύων Πεποίθησης Bayes* (Bayesian belief networks-BBN). Ως πηγή εκπαίδευσης των δικτύων αυτών χρησιμοποιήθηκαν αρχεία αλληλεπίδρασης (log-files) των μαθητών με το περιβάλλον μάθησης. Τα αρχεία αυτά συσχετίστηκαν με τη στρατηγική επίλυσης, αναλύθηκαν με τη χρήση ενός log file parser ώστε να αποθηκευτούν σε κατάλληλη μορφή για την εφαρμογή αλγορίθμων κατασκευής BBN.

## **2. Δίκτυα Bayes.**

Τα *Δίκτυα Πεποίθησης Bayes* (Bayesian Belief Networks, BBN) αποτελούν ένα σημαντικό εργαλείο αναπαράστασης γνώσης και εξαγωγής συμπερασμάτων, υπό συνθήκες αβεβαιότητας. Δεδομένου ενός συνόλου μεταβλητών  $D = \langle X_1, X_2 \dots X_N \rangle$ , όπου κάθε μεταβλητή  $X_i$  παίρνει τιμές από ένα σύνολο  $T(X_i)$ , ένα BBN περιγράφει την κατανομή πιθανότητας στο σύνολο αυτό. Με κεφαλαία γράμματα, όπως  $X, Y$  συμβολίζουμε μεταβλητές και πεζά γράμματα όπως  $x, y$  συμβολίζουμε τις τιμές των μεταβλητών αυτών. Τυπικά, ένα BBN είναι ένας κατευθυνόμενος ακυκλικός γράφος (Directed Acyclic Graph- DAG) που αναπαριστά μία συνδυασμένη πιθανοτική κατανομή (joint probability distribution). Ένα δίκτυο  $B$  συμβολίζεται σαν ένα ζεύγος  $B = \langle G, \Theta \rangle$  (Pearl, 1988) όπου  $G$  ένας DAG του οποίου οι κόμβοι αντιστοιχούν στις μεταβλητές του  $D$  και  $\Theta$  το σύνολο των

παραμέτρων που ποσοτικοποιούν το δίκτυο. Ο G ενσωματώνει την ακόλουθη υπόθεση για την ανεξαρτησία των παραμέτρων:

*Κάθε μεταβλητή  $X_i$  είναι ανεξάρτητη από τις μή-απογόνους της, δεδομένου του συνόλου των πατρικών της μεταβλητών.*

Η παράμετρος  $\Theta$  περιλαμβάνει πληροφορίες για την κατανομή πιθανότητας μιας τιμής  $x_i$  μιας μεταβλητής  $X_i$ , δεδομένων των τιμών των αμέσως προηγούμενων μεταβλητών. Η μοναδική συνδυασμένη πιθανοτική κατανομή για το σύνολο  $\langle X_1, X_2, \dots, X_N \rangle$  που περιγράφει ένα δίκτυο  $B$  υπολογίζεται από τον τύπο:

$$P_B(X_1, \dots, X_N) = \prod_{i=1}^N P(x_i | \text{πατρικο } i(X_i)) \quad (1)$$

## 2.1 Εκπαίδευση BBN από δεδομένα.

Για τη διαδικασία εξαγωγής συμπερασμάτων με στόχο τη συσχέτιση των καταγεγραμμένων αλληλεπιδράσεων του μαθητή με τις στρατηγικές που ακολούθησε, χρειάζεται να εκπαιδευτεί το δίκτυο από ήδη διαθέσιμα δεδομένα. Η επιλογή των παραμέτρων χρήζει ιδιαίτερης προσοχής, καθότι ο αριθμός των πιθανών δικτύων που μπορούν να τις περιγράψουν, ισούται με  $2^{\frac{N(N-1)}{2}}$ , όπου  $N$  ο αριθμός των μεταβλητών.

Η εκπαίδευση ενός BBN περιλαμβάνει δύο διαδικασίες, την εκπαίδευση της δομής και την εκπαίδευση των παραμέτρων  $\Theta$  της δομής αυτής. Για τη δεύτερη διαδικασία, χρησιμοποιούμε τις συχνότητες εμφάνισης των τιμών του σώματος εκπαίδευσης (Cooper & Herskovits 1992). Όσον αφορά τη δομή του δικτύου, χρησιμοποιούμε την παρακάτω εξίσωση μαζί με το θεώρημα του Bayes για να εξακριβώσουμε τη σχέση  $r$  μεταξύ δύο υποψηφίων δικτύων  $B_1$  και  $B_2$  αντίστοιχα:

$$r = \frac{P(B_1 | D)}{P(B_2 | D)} \quad (2)$$

$$P(B | D) = \frac{P(D | B)P(B)}{P(D)} \quad (3)$$

όπου:

- $P(B|D)$  η πιθανότητα το δίκτυο  $B$  να είναι το επιθυμητό δεδομένου του συνόλου  $D$ .
- $P(D|B)$  η πιθανότητα που δίνει το δίκτυο  $B$  στα δεδομένα  $D$ .
- $P(D)$  η «γενική» πιθανότητα των δεδομένων.
- $P(B)$  η πιθανότητα του δικτύου  $B$  προτού να δοθούν τα δεδομένα.

Εφαρμόζουμε την εξίσωση (3) στη (2). Θεωρώντας ότι τα υποψήφια δίκτυα είναι ισοπίθανα ( $P(B_1)=P(B_2)$ ) -λαμβάνοντας υπ'όψη ότι δεν έχουμε πρότερη γνώση για την πιθανότητα ενός δικτύου πριν δούμε τα δεδομένα- η σχέση που εκφράζει το  $r$  γίνεται:

$$r = \frac{P(D | B_1)}{P(D | B_2)} \quad (4)$$

Η πιθανότητα που δίνει το δίκτυο στα δεδομένα υπολογίζεται από την εξίσωση των Glymour και Cooper (1999):

$$P(D | B) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(\frac{\Xi}{q_i})}{\Gamma(\frac{\Xi}{q_i} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\frac{\Xi}{r_i q_i} + N_{ijk})}{\Gamma(\frac{\Xi}{r_i q_i})} \quad (5)$$

## 2.2. Naive Bayes

Σε αντίθεση με τα BBN, ο αλγόριθμος ταξινόμησης Naive Bayes βασίζεται στην υπόθεση ότι οι τιμές των μεταβλητών  $\langle X_1, X_2 \dots X_N \rangle$  είναι πιθανοτικά ανεξάρτητες, δεδομένης μιας τιμής ταξινόμησης  $v$  (στρατηγική χρήστη). Έχοντας παράσχει ένα σώμα εκπαίδευσης, όταν εισάγουμε ένα νέο διάνυσμα, το οποίο περιγράφεται από ένα σύνολο τιμών  $\langle x_1, x_2 \dots x_N \rangle$ , ο αλγόριθμος Naive Bayes το ταξινομεί προβλέποντας την τιμή ταξινόμησης  $V_{NB}$  χρησιμοποιώντας την εξίσωση:

$$V_{NB} = \arg \max_{v_j \in V} P(v_j) \prod_{i=1}^n P(x_i | v_j) \quad (6)$$

Όπου  $v_j$  είναι μία πιθανή τιμή ταξινόμησης που ανήκει σε ένα πεπερασμένο σύνολο  $V$ . Οι όροι  $P(v_j)$  και  $P(x_i | v_j)$  εκτιμώνται υπολογίζοντας τη συχνότητα εμφάνισης στο σώμα εκπαίδευσης. Συγκεκριμένα, για τη μελέτη περίπτωση της παρούσας εργασίας οι στρατηγικές είναι 13. Επομένως, ο όρος  $P(v_j)$  υπολογίζεται μετρώντας πόσες φορές εμφανίστηκε η εκάστοτε στρατηγική στο σώμα εκπαίδευσης και διαιρώντας με το συνολικό αριθμό των διανυσμάτων του σώματος εκπαίδευσης. Αντίστοιχα, ο όρος  $P(x_i | v_j)$  είναι ο αριθμός των περιπτώσεων που η τιμή  $x_i$  εμφανίστηκε μαζί με τη τιμή ταξινόμησης  $v_j$ . Για το λόγο αυτό ο χρόνος εκπαίδευσης είναι σημαντικά λιγότερος σε σχέση αυτόν ενός BBN. Εφόσον η εκπαίδευση έχει λάβει χώρα, ο χρόνος απόκρισης για τη διαδικασία εξαγωγής συμπεράσματος για την κλάση ενός νέου διανύσματος είναι ισότιμος. Μάλιστα, όταν έχουμε πολλές μεταβλητές τα BBN υπερτερούν διότι εξετάζουν μόνο εκείνες που επηρεάζουν την κλάση, οπότε στη χειρότερη περίπτωση που την επηρεάζουν όλες, τα BBN αποκρίνονται το πολύ στον ίδιο χρόνο με τον αλγόριθμο Naive Bayes.

## 3. Μελέτη περίπτωσης

### 3.1 Χαρακτηριστικά πειράματος

Στην ενότητα αυτή περιγράφεται ένα πείραμα ανάπτυξης ενός εργαλείου αυτόματης καταγραφής στρατηγικών επίλυσης προβλημάτων με χρήση των τεχνικών της προηγούμενης ενότητας. Για το πείραμα αυτό χρησιμοποιήθηκαν δεδομένα από την αλληλεπίδραση μιας ομάδας 30 μαθητών πρώτων τάξεων γυμνασίου σε συνθήκες σχολικού εργαστηρίου με τον μικρόκοσμο C.AR.ME. Το περιβάλλον αυτό είναι ένα ανοικτό, αλληλεπιδραστικό περιβάλλον πολλαπλών αναπαραστάσεων μάθησης των εννοιών που αφορούν στη διατήρηση και στη μέτρηση της

επιφάνειας (Kordaki & Potari, 1998). Το πείραμα περιελάμβανε την επίλυση δύο προβλημάτων, τον μετασχηματισμό ενός μη-κυρτού πολυγώνου σε ένα ισοδύναμο ως προς την επιφάνεια σχήμα και το πρόβλημα της σύγκρισης ενός μη-κυρτού πολυγώνου με ένα τετράγωνο (Κορδάκη, 1999). Οι μαθητές κλήθηκαν να επιλύσουν και τα δύο προβλήματα με όλους τους δυνατούς τρόπους με χρήση των εργαλείων που διατίθενται από τον μικρόκοσμο. Η ανάλυση των στρατηγικών επίλυσης που ακολούθησαν οι μαθητές προκειμένου να αντιμετωπίσουν τα παραπάνω προβλήματα περιγράφονται στην εργασία (Κορδάκη, 1999). Στην παρούσα εργασία χρησιμοποιήθηκαν δεδομένα αλληλεπίδρασης και οι κατηγορίες των στρατηγικών επίλυσης που ακολούθησαν οι μαθητές στο πρόβλημα της σύγκρισης οι οποίες και παρουσιάζονται στον Πίνακα 1

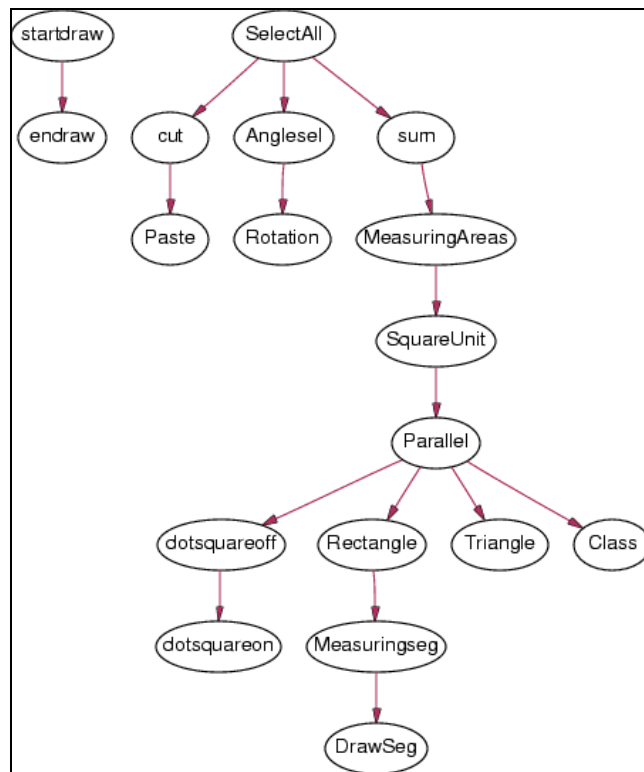
|     | <i>Οι κατηγορίες των συγκρίσεων/στρατηγικών</i>                                                                                                   | <i>Αριθμός<br/>στρατ/κών</i> | <i>%επί του<br/>δείγματος</i> |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|-------------------------------|
| 1.  | Με το 'μάτι'                                                                                                                                      | 0                            | 0                             |
| 2.  | Με την περίμετρο                                                                                                                                  | 1                            | 0,87                          |
| 3.  | Με την αυτόματη λειτουργία της μέτρησης                                                                                                           | 21                           | 18,42                         |
| 4.  | Με τις προσομοιωμένες αισθησιοκινητικές ενέργειες (ακε) των παιδιών                                                                               | 14                           | 12,28                         |
| 5.  | Με τη λειτουργία της μέτρησης                                                                                                                     | 27                           | 23,68                         |
| 6.  | Με τους αυτόματους μετασχηματισμούς & τη λειτουργία της μέτρησης                                                                                  | 11                           | 9,64                          |
| 7.  | Με τους αυτόματους μετασχηματισμούς                                                                                                               | 14                           | 12,28                         |
| 8.  | Με τους αυτόματους μετασχηματισμούς & τις προσομοιωμένες ακε των παιδιών                                                                          | 16                           | 14,03                         |
| 9.  | Με την αυτόματη λειτουργία της μέτρησης & τους αυτόματους μετασχηματισμούς                                                                        | 3                            | 2,63                          |
| 10. | Με την αυτόματη λειτουργία της μέτρησης & τις προσομοιωμένες ακε                                                                                  | 2                            | 1,75                          |
| 11. | Με τους τύπους υπολογισμού & τις προσομοιωμένες ακε                                                                                               | 1                            | 0,87                          |
| 12. | Με τους τύπους υπολογισμού & την αυτόματη λειτουργία της μέτρησης                                                                                 | 1                            | 0,87                          |
| 13. | Με αυτόματους μετασχηματισμούς & προσομοιωμένες ακε & τη λειτουργία της μέτρησης & με τύπους υπολογισμού σε συνδυασμό με τον εγκλεισμό του πολυγ. | 3                            | 2,63                          |
|     | <b>Σύνολο</b>                                                                                                                                     | <b>114</b>                   | <b>100</b>                    |

*Πίνακας 1. Οι δυνατές στρατηγικές που μπορεί να χρησιμοποιήσει ο χρήστης.*

### 3.2 Προετοιμασία δεδομένων

Το δείγμα καταγεγραμμένων αρχείων πληκτρολογήσεων από τη διεξαγωγή του πειράματος χρήσης του C.AR.ME. στο οποίο στηρίχθηκε η έρευνα είναι 114, τα οποία στην συνέχεια αναλύθηκαν. Στα περισσότερα από αυτά είχαν καταχωρηθεί σχόλια που περιέγραφαν την στρατηγική που ακολούθησε ο μαθητής, ενώ παράλληλα είχε καταγραφεί και ένα στιγμιότυπο της διεπιφάνειας εργασίας για επαλήθευση του σεναρίου αλληλεπίδρασης που ακολουθήθηκε. Τα δεδομένα αυτά αναλύθηκαν με την βοήθεια εργαλείου που αναπτύχθηκε για υποβοήθηση της διαδικασίας αξιολόγησης εκπαιδευτικού λογισμικού αντίστοιχου τύπου, του Usability Analyzer (Tselios et al. 2001). Στο εργαλείο αυτό ο αξιολογητής έχει τη δυνατότητα να επισκοπήσει αναλυτικά την αλληλεπίδραση του μαθητή με το περιβάλλον μάθησης, μέσω της ταυτόχρονης παρουσίας του αρχείου πληκτρολογήσεων, των σχετικών σχολίων που καταγράφηκαν στη φάση διεξαγωγής του πειράματος καθώς επίσης και αποθηκευμένων στιγμιότυπων από το χώρο εργασίας του μαθητή.

Με τη χρήση της δυνατότητας αυτής καταγράφηκε σε σχεσιακή βάση δεδομένων η συχνότητα εμφάνισης κάθε γεγονότος της αλληλεπίδρασης του μαθητή με το σύστημα. Συνολικά καταγράφηκαν 20 διαφορετικά γεγονότα όπως αναγνωρίστηκαν από τα σειριακά αρχεία πληκτρολόγησης, ενώ δύο επιπλέον μεταβλητές που προστέθηκαν αφορούσαν στο σύνολο των αλληλεπιδράσεων σε κάθε συνεδρία χρήσης (μεταβλητή *sum*) και στον τύπο της στρατηγικής που ακολουθήθηκε όπως αυτή αναγνωρίστηκε στη φάση διεξαγωγής του πειράματος (μεταβλητή *class*).



Σχήμα 1. Το Bayesian δίκτυο πεποίθησης που κατασκευάστηκε για τον πιθανοτικό συμπερασμό της ακολουθούμενης στρατηγικής από τα αρχεία πληκτρολογήσεων.

### 3.3 Κατασκευή BBN-Σύγκριση με άλλες μεθόδους.

Το σύνολο των δεδομένων αυτών τροφοδότησε τη διαδικασία κατασκευής του BBN, διαγραμματική αναπαράσταση του οποίου φαίνεται στο σχήμα 1. Γενικά, ένα BBN είναι ικανό να υπολογίζει την κατανομή πιθανότητας για οποιοδήποτε υποσύνολο μεταβλητών, δεδομένων των τιμών οποιουδήποτε υποσυνόλου των υπόλοιπων μεταβλητών.

Από τη δομή του δικτύου βλέπουμε ότι οι μεταβλητές *startdraw* και *enddraw* που είναι τα γεγονότα που δημιουργούνται όταν ο μαθητής ξεκινά ή τελειώνει αντίστοιχα τη σχεδίαση ενός γεωμετρικού σχήματος, εξαρτώνται μεταξύ τους όχι όμως και με το κύριο σώμα του δικτύου. Επίσης, έχουν καταγραφεί ακόμα τέσσερις (4) μεταβλητές στα αρχεία πληκτρολογήσεων, οι *UnitIteration*, *Erase*, *Clear* και *Symmetry* που προκύπτουν ανεξάρτητες από οποιαδήποτε άλλη μεταβλητή και για το λόγο αυτό δεν συμπεριλαμβάνονται στον κόμβο. Για να εκτιμηθεί η αποτελεσματικότητα του κατασκευασθέντος δικτύου στην ορθή εκτίμηση της ακολουθούμενης

στρατηγικής με βάση τα αρχεία πληκτρολογήσεων, χρησιμοποιήθηκε η μεθοδολογία ten fold cross validation. Με τη μέθοδο αυτή ο πληθυσμός των δεδομένων μας διαμερισματοποιείται σε σύνολα εκπαίδευσης και ελέγχου, με αναλογία 9 προς 1, με συνολικά 10 τυχαίες τέτοιες διαμερισματοποιήσεις. Το τελικό αποτέλεσμα με τη χρήση της μεθόδου αυτής ήταν 84,07% (+/- 6,75% στο  $p < 0.05$ ) ορθών προβλέψεων από τα σύνολα ελέγχου.

Ακολούθως, το πείραμα επανελήφθη με χρήση των διαδεδομένων μηχανικής μάθησης decision tables, J48 PART και Naive Bayes, οι οποίες έχουν ενσωματωθεί στο περιβάλλον αυτόματης εκμάθησης των Holmes et al. (1994). Στόχος του πειράματος αυτού ήταν η συγκριτική αποτίμηση της αποτελεσματικότητας της μεθόδου BBN σε σχέση και με τις τεχνικές αυτές. Τα αποτελέσματα που προέκυψαν συνοψίζονται στον πίνακα 2.

| Μέθοδος              | Bayesian networks | Naive Bayes | Decision Tables | J48 PART |
|----------------------|-------------------|-------------|-----------------|----------|
| % επιτυχία           | 84,07%            | 72,56%      | 63,71%          | 64,60%   |
| Δείκτης <sup>1</sup> | 10                | 8,63        | 7,57            | 7,68     |

Πίνακας 2. Σύγκριση απόδοσης τεχνικών κατηγοριοποίησης.

Από τον πίνακα 2 η προτεινόμενη μέθοδος BBN φαίνεται να υπερτερεί των άλλων μεθόδων αν και όπως περιγράφηκε στην ενότητα 3, άλλες τεχνικές παρουσιάζουν πλεονεκτήματα όσον αφορά την πολυπλοκότητα εφαρμογής τους.

#### 4. Συμπεράσματα.

Η προσέγγιση του προβλήματος της αυτόματης καταγραφής των στρατηγικών επίλυσης προβλημάτων ενός μαθητή με τη χρήση BBN φαίνεται να οδηγεί σε συμπαγή και πολυπλεύρως αξιοποιήσιμα αποτελέσματα. Ένα πλεονέκτημα της τεχνικής BBN, σε σχέση με άλλους αλγόριθμους μηχανικής μάθησης, είναι η δυνατότητα εισαγωγής πρότερης γνώσης που αφορά στις ακολουθίες ενεργειών των μαθητών όταν αυτές είναι γνωστές στον εκπαιδευτικό και στον σχεδιαστή του συστήματος. Το ποσοστό επιτυχίας της τεχνικής αυτής συγκρινόμενο ενδεικτικά και με άλλες μεθόδους κατηγοριοποίησης είναι ιδιαίτερα υψηλό (Πίνακας 2), δεδομένου ότι το πλήθος των μετρήσεων του πειράματος που περιγράφηκε είναι ιδιαίτερα μικρό. Βέβαια, πιθανόν το δίκτυο να αντιμετωπίζει δυσκολία στην αναγνώριση στρατηγικών που εμφανίστηκαν με πολύ μικρή συχνότητα στο δείγμα αναφοράς, κάτι στο οποίο η μελέτη δεν εστίασε. Πιθανά περαιτέρω εμπλουτισμός με περισσότερα πειραματικά δεδομένα θα συμβάλλει στην ευρωστία και μη πόλωση των αποτελεσμάτων.

---

<sup>1</sup> Δείκτης: κανονικοποιημένη αποτελεσματικότητα εκτίμησης στρατηγικής σε σχέση με το ποσοστό επιτυχίας του καλύτερου αλγόριθμου.

Σημαντικό επίσης πλεονέκτημα της προτεινόμενης μεθόδου αποτελεί ο παραγόμενος γράφος (σχήμα 1), ο οποίος παρέχει μια αναπαράσταση των μεταβλητών που καθορίζουν την φύση της αλληλεπίδρασης χρήστη-συστήματος. Η αναπαράσταση αυτή αποτελεί επίσης μια εύληπτη και κατανοητή μορφή περιγραφής της αλληλεπίδρασης αυτής. Η μελέτη του γράφου μπορεί να μας οδηγήσει σε βελτίωση του περιβάλλοντος. Αναγνωρίζοντας ανά πάσα στιγμή το σημείο στο οποίο βρίσκεται ο χρήστης είναι δυνατή π.χ. η παρουσίαση επίκαιρης βοήθειας με προτεραιότητα σε πιο χρήσιμες συμβουλές ή η παρουσίαση των πλέον πιθανών να ακολουθήσουν επιλογών που σχετίζονται με την τρέχουσα κατάσταση του διαλόγου.

Από παιδαγωγική σκοπιά, η αυτόματη καταγραφή στρατηγικών σε ανοιχτά περιβάλλοντα μάθησης υποστηρίζει και εμπλουτίζει την εκπαιδευτική διαδικασία καθώς οδηγεί στην ταχεία αναγνώριση της υιοθέτησης ή όχι από τους μαθητές συγκεκριμένων στρατηγικών που μεγιστοποιούν την μάθηση, ενώ η εφαρμογή τους σε πραγματικό χρόνο ιδίως σε συστήματα από απόσταση εκπαίδευσης μπορεί να είναι μεγάλη για την διάγνωση και αξιολόγηση του μαθησιακού αποτελέσματος.

## 5. Αναφορές

- Bauersfeld, H. (1988). Interaction, Construction and Knowledge: Alternative perspectives for Mathematics Education. In D. A. Grows, et al (Eds), *Effective Mathematics Teaching* (pp.27-46, Lawrence Erlbaum.
- Confrey, J. (1995). How Compatible are Radical Constructivism, Sociocultural Approaches, and Social Constructivism? In L.P. Steffe & J. Gale (Eds), *Constructivism in Education*, pp. 185-226, Lawrence Erlbaum.
- Cooper J. and Herskovits E. 1992. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9, pp.309-347.
- Dimitracopoulou, A., Komis, V., Aposotolopoulos, P., & Politis, P., (1999). Design principles of a new modelling environment for young students, *AI-ED Proc, 9<sup>th</sup> Int. Conf. on AI in Education*, Le Mans, France.
- Glymour C. and Cooper G. (eds.). 1999. *Computation, Causation & Discovery*. AAAI Press/The MIT Press
- Holmes, G., Donkin, A., and Witten, I.H. (1994) *Weka: a machine learning workbench*. Proceedings of the 2nd Australian and N. Zealand Conference on Intelligent Information Systems, Brisbane, Australia, pp. 357-361.
- Kordaki, M., & Avouris, N., (2000, submitted ). Modeling in Design and Evaluation of Open Learning Environments, *Computers & Education*.
- Kordaki, M., & Potari, D. (1998). A learning environment for the conservation of area and its measurement: a computer microworld. *Computers and Education*, 31, 405-422.
- Κορδάκη, Μ. (1999). *Οι έννοιες της διατήρησης και της μέτρησης της επιφάνειας μέσα από το σχεδιασμό την υλοποίηση και την αξιολόγηση εκπαιδευτικού λογισμικού*. Δημοσίευτη Διδακτορική διατριβή, Πάτρα, Μάιος 1999.
- Laborde, J-M. (1990). *Cabri-Geometry* [Software]. France: Universite de Grenoble.
- Papert, S., (1980). *Mindstorms: Pupils, Computers, and Powerful Ideas*. New York: Basic Books.
- Pearl J. 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. San Mateo, CA: M.K.
- Tselios N., Avouris N., Kordaki M., (2001, submitted) Task Modelling to support design and evaluation of open problem solving environments, *Journal of Research and Practice in Information Technology*.