



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

**Αναγνώριση χειρόγραφων μαθηματικών εκφράσεων
εντός γραμμής (online)**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Φωτεινή Σιμιστήρα

Ιούνιος 2015



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Αναγνώριση χειρόγραφων μαθηματικών εκφράσεων εντός γραμμής (online)

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Φωτεινή Σιμιστήρα

Επταμελής Επιτροπή

Γεώργιος Καραγιάννης
Καθηγητής ΕΜΠ
ΕΠΙΒΛΕΠΩΝ

Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής ΕΜΠ

Στέφανος Κόλλιας
Καθηγητής ΕΜΠ

Τιμολέων Σελλής
Καθηγητής ΕΜΠ

Βασίλειος Κατσούρος
Ερευνητής Α', ΕΚ Αθηνά

Βασίλειος Γάτος
Ερευνητής Α', ΕΚΕΦΕ
Δημόκριτος

Prof. Dr. habil.
Marcus Eichenberger-Liwicki
Lecturer (Maître Assistant)
University of Fribourg
Switzerland



Φωτεινή Σιμιστήρα

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Φωτεινή Σιμιστήρα, 2015.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Αναγνώριση χειρόγραφων μαθηματικών εκφράσεων εντός γραμμής (online)

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Φωτεινή Σιμιστήρα

Τριμελής Συμβουλευτική Επιτροπή

Γ. Καραγιάννης
(Επιβλέπων)

Α. Γ. Σταφυλοπάτης

Στ. Κόλλιας

Περίληψη

Αντικείμενο της διατριβής είναι η αναγνώριση χειρόγραφων μαθηματικών εκφράσεων (ΜΕ). Η κύρια συνεισφορά της παρούσας εργασίας αφορά στο σχεδιασμό και στην υλοποίηση ενός συστήματος αναγνώρισης ΜΕ που λειτουργεί σε τρία στάδια: α) την ομαδοποίηση των κονδυλίων σε σύμβολα και ταυτόχρονη αναγνώριση των συμβόλων, β) την αναγνώριση των σχέσεων μεταξύ δύο μαθηματικών συμβόλων ή υπο-εκφράσεων και γ) την ανασύσταση της μαθηματικής έκφρασης και η αναπαράστασή της με βάση το μορφότυπο MathML (Παράρτημα Β), με δεδομένη την αναγνώριση και την σχέση μεταξύ των συμβόλων που την απαρτίζουν.

Στο πρώτο στάδιο, προτείνεται η κωδικοποίηση της ακολουθίας των συντεταγμένων των συμβόλων σε ακολουθίες κωδικών Freeman-8 επιπέδων συνοδευόμενων με το ποσοστό του επιμέρους μήκους μεταξύ δύο διαδοχικών σημείων. Η σύγκριση των υπό εξέταση συμβόλων με τα πρότυπα αναφοράς γίνεται με την τεχνική της απόστασης του κοντινότερου γείτονα με βάση το ταίριασμα στα δείγματα (template matching) των προτύπων αναφοράς.

Στο δεύτερο στάδιο χρησιμοποιούνται τα γεωμετρικά χαρακτηριστικά των συμβόλων της μαθηματικής έκφρασης (π.χ. κεντροειδές, πλαίσιο οριοθέτησης) για την εύρεση των σχέσεων μεταξύ των συμβόλων (π.χ. εκθέτης, δείκτης) και την κατανομή τους σε επιμέρους υποεκφράσεις (sub-expressions) της ΜΕ.

Στην παρούσα εργασία αναπτύχθηκαν δύο τεχνικές. Στην πρώτη τεχνική χρησιμοποιήθηκε ο γεωμετρικός ταξινομητής για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ΜΕ και την τοποθέτηση αυτών στα αντίστοιχα επίπεδα της ΜΕ. Τα σύμβολα που ανήκουν σε διαφορετικά επίπεδα ενώνονται με σύμβολα άλλων επιπέδων βάση της τοπολογικής σχέσης που υπάρχει μεταξύ τους. Τέλος, αφού όλα τα σύμβολα έχουν τοποθετηθεί στα αντίστοιχα επίπεδα και οι σχέσεις μεταξύ των συμβόλων έχουν οριστεί, παράγεται ένα MathML αρχείο για την αναπαράσταση της ΜΕ. Οι μαθηματικές εκφράσεις που μπορούν να αναγνωριστούν με αυτή την τεχνική είναι μέχρι ενός επιπέδου πάνω (εκθέτης) και ενός επιπέδου κάτω (δείκτης). Στην δεύτερη τεχνική χρησιμοποιήθηκε ένας ταξινομητής βασισμένος σε SVM για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ΜΕ.

Στο τρίτο στάδιο, αναπτύχθηκε μία στοχαστική γραμματική (SCFG) για την περιγραφή της σύνταξης των ΜΕ. Αφού ορίσουμε την γραμματική, μπορούμε πλέον να ψάξουμε για την πιο πιθανή υποψήφια μαθηματική έκφραση λαμβάνοντας υπόψη τις τοπολογικές σχέσεις μεταξύ των συμβόλων και/ή των υπό-εκφράσεων. Για να αναλύσουμε αποτελεσματικότερα μια ΜΕ, πρέπει πρώτα να ορίσουμε ένα αλγόριθμο ανάλυσης (parsing algorithm) της ΜΕ. Σε αυτή την εργασία χρησιμοποιήσαμε τον αλγόριθμο Cocke-Younger-Kasami (CYK) κατάλληλα τροποποιημένο ώστε να μπορεί να αντιμετωπίσει το πρόβλημα της ανάλυσης των μαθηματικών εκφράσεων.

Τέλος, αναπτύχθηκε ένα πρόγραμμα συλλογής χειρόγραφων μαθηματικών εκφράσεων για την δημιουργία βάσης χειρόγραφων μαθηματικών συμβόλων. Η βάση που δημιουργήθηκε αποτελείται από 186 σύμβολα που συλλέχτηκαν από 50 γραφείς. Κάθε γραφέας έγραψε κάθε σύμβολο 5 φορές. Επιπλέον, κάθε γραφέας έγραψε 54 εξισώσεις οι οποίες περιέχουν τουλάχιστον από μία φορά όλα τα σύμβολα που προαναφέρθηκαν.

Για την αξιολόγηση του αλγορίθμου αναγνώρισης των χειρόγραφων μαθηματικών συμβόλων χρησιμοποιήθηκαν δύο βάσεις χειρόγραφων μαθηματικών συμβόλων, η βάση LaViola και η βάση CROHME2014. Το ποσοστό ορθής αναγνώρισης είναι 92% στην βάση LaViola και 77.25% στη βάση CROHME2014. Για την αξιολόγηση του αλγορίθμου εύρεσης των σχέσεων μεταξύ συμβόλων χρησιμοποιήθηκε η βάση χειρόγραφων μαθηματικών εκφράσεων MathBrush, με μέσο ποσοστό λάθους 2.87%. Τέλος, για την αξιολόγηση του αλγορίθμου ανασύστασης της μαθηματικής έκφρασης χρησιμοποιήθηκαν οι βάσεις CROHME2012 και 2013. Οι βάσεις CROHME2012 και 2013 αποτελούνται από 4 μέρη: Μέρος-I, Μέρος-II, Μέρος-III και Μέρος-IV. Το ποσοστό ορθής αναγνώρισης του προτεινόμενου αλγορίθμου είναι 78.70%, 65.78%, 56.37% και 50.22% για τα τέσσερα μέρη των βάσεων CROHME2012 και 2013.

Λέξεις κλειδιά

Αναγνώριση μαθηματικών συμβόλων, αναγνώριση χειρόγραφων μαθηματικών παραστάσεων, αναγνώριση τοπολογικών σχέσεων, SCFG, SVM, CROHME

Abstract

The thesis focuses on the recognition of online handwritten mathematical expressions. The main contribution is the development of a system for the recognition of online handwritten mathematical expressions. This system operates in three stages: a) grouping strokes to symbol hypotheses and recognizing these candidate symbols, b) recognition of the spatial relation between two corresponding mathematical symbols or subexpressions, and c) representation of the mathematical expression using the MathML format (Appendix B).

In the first stage, we use the Freeman-8 chain code followed by the local percentage of the length between two successive points to represent the symbol. The Freeman-8 chain code is used to represent a boundary by a connected sequence of straight-line segments of specified length and direction. The direction of each segment is coded by using an integer number in the range of 0 to 8. For the comparison between the reference symbol and the test symbol, the template matching technique is used.

In the second stage, the spatial relations between the symbols that comprise the mathematical expression are recognized. The geometrical features of the symbols comprising the mathematical expression (ME) (e.g. centroid, bounding box) are used in order to classify the relation between symbols (e.g. superscript, subscript) and finally put the symbols in the right subexpression of the ME.

In the current thesis two different techniques are developed. In the first technique the geometrical classifier for the recognition of the spatial relation between symbols comprising the ME is used. After the spatial relation is recognized, the symbols are positioned in the corresponding levels of the ME. The symbols that belong to different levels are connected with symbols of other levels of the ME based on the spatial relation between them. Finally, when all the symbols are positioned in the corresponding levels and the spatial relations between the symbols are defined, a MathML file is produced for the representation of the ME. The mathematical expression that can be recognized through this technique is till one level up (superscript) and one level down (subscript). In the second technique an SVM based classifier is developed for the recognition of the spatial relation between symbols of the ME.

In the third stage a stochastic context free grammar (SCFG) is used for the description of the syntax of the MEs. When the grammar is defined we can then search for the most likely ME using also the spatial relations between the symbols and/or subexpressions. To analyse effectively the ME, we first need to define a parsing algorithm for the ME. In the current thesis we use a variation of the Cocke–Younger–Kasami (CYK) algorithm in order to tackle the problem of the analysis of the MEs.

Also a program for the acquisition of handwritten mathematical expressions is developed for the creation of a handwritten mathematical symbols dataset. With this system a database 186 symbols selected by 50 different writers was created. Each writer wrote each symbols 5 times. Moreover each writer wrote 54 mathematical equations that contained at least once each symbol of the dataset.

We use two different datasets of mathematical symbols, the Laviola dataset and the CROHME dataset. The accuracy rate for the symbol recognition algorithm is 92% for the LaViola dataset and 77.25% for the CROHME2014 dataset. For the evaluation of the algorithm for the recognition of the spatial relation between objects, we used the MathBrush dataset. The reported error rate in this case is 2.87%. Finally the CROHME 2012 and 2013 datasets were used for the evaluation of the parsing algorithm of the MEs and gave an accuracy result of 78.70%, 65.78%, 56.37% and 50.22% for the four different parts of the dataset.

KeyWords

Mathematical symbol recognition, recognition of handwritten mathematical expressions, recognition of spatial relations, SCFG, SVM, CROHME

Περιεχόμενα

Κεφάλαιο 1. Εισαγωγή	7
Κεφάλαιο 2. Υπάρχοντα συστήματα αναγνώρισης μαθηματικών εκφράσεων	10
2.1 Natural Log System (1999)	10
2.2 Freehand Formula Entry System/ DRACULAE (1999)	11
2.3 PenCalc (2001)	13
2.4 JMathNotes (2003)	14
2.5 InftyEditor (2003)	15
2.6 Math Input (2009)	15
Κεφάλαιο 3. Αναγνώριση χειρόγραφων μαθηματικών συμβόλων	17
3.1 Ανάπτυξη βάσης δεδομένων χειρόγραφων μαθηματικών συμβόλων	18
3.2 Λογισμικό συλλογής χειρόγραφων μαθηματικών συμβόλων	20
3.3 Διαθέσιμες τεχνικές αναγνώρισης συμβόλων	22
3.4 Προτεινόμενες τεχνικές	25
3.4.1 Τεχνική ελαστικού ταιριάσματος	25
3.4.2 Τεχνική ταύτισης απόστασης με υπόδειγμα	29
3.5 Αξιολόγηση	37
3.6 Συμπεράσματα	39
Κεφάλαιο 4. Αναγνώριση τοπολογικής σχέσης μεταξύ δύο συμβόλων της ΜΕ	41
4.1 Βάσεις δεδομένων χειρόγραφων ΜΕ	41
4.2 Δομικά στοιχεία	43
4.3 Προτεινόμενες τεχνικές	44
4.3.1 Ταξινομητής με χρήση ευρυστικών κανόνων	44
4.3.2 Ταξινομητής βασισμένος σε μηχανές διανυσμάτων υποστήριξης	46
4.3.3 Ταξινομητής βασισμένος στην τεχνική των νευρωνικών δικτύων	48
4.4 Αξιολόγηση	48
4.5 Συμπεράσματα	52
Κεφάλαιο 5. Ανασύσταση της δομής χειρόγραφων ΜΕ	53
5.1 Βάση δεδομένων χειρόγραφων ΜΕ	53
5.2 Διαθέσιμες τεχνικές	70
5.3 Προτεινόμενες τεχνικές	73
5.3.1 Σταδιακή ανάλυση ΜΕ με την τοποθέτηση των συμβόλων σε επίπεδα	74
5.3.2 Σταδιακή ανάλυση με χρήση στοχαστικών γραμματικών	77
5.4 Αξιολόγηση	83
5.5 Συμπεράσματα	87
Κεφάλαιο 6. Συμπεράσματα	89
Βιβλιογραφία	91
Παράρτημα Α	95
Παράρτημα Β	98
Παράρτημα Γ	99
Παράρτημα Δ	100
Παράρτημα Ε	101
Παράρτημα Ζ	106

Contents

1.	Introduction	7
2.	Existing systems for recognition of online handwritten mathematical expressions	10
	2.1 Natural Log System (1999)	10
	2.2 Freehand Formula Entry System/ DRACULAE (1999)	11
	2.3 PenCalc (2001)	13
	2.4 JMathNotes (2003)	14
	2.5 InfyEditor (2003)	15
	2.6 Math Input (2009)	15
3.	Recognition of handwritten mathematical symbols	17
	3.1 Development of database of handwritten mathematical symbols	18
	3.2 Software for selection of handwritten mathematical symbols	20
	3.3 Existing techniques for the recognition of handwritten mathematical symbols	22
	3.4 Suggested techniques	25
	3.4.1 Elastic matching technique	25
	3.4.2 Template matching technique	29
	3.5 Evaluation	37
	3.6 Conclusions	39
4.	Recognition of the spatial relation between two symbols of the ME	41
	4.1 Database of handwritten ME	41
	4.2 Structural elements	43
	4.3 Suggested Techniques	44
	4.3.1. Heuristic rule based classifier	44
	4.3.2. SVM based classifier	46
	4.3.3. ANN based classifier	48
	4.4 Evaluation	48
	4.5 Conclusions	52
5.	Reconstruction of handwritten ME	53
	5.1 Databases of handwritten ME	53
	5.2 Existing Techniques	70
	5.3 Suggested Techniques	73
	5.3.1. ME recognition by defining levels	74
	5.3.2. ME recognition by using SCFG and CYK based parser	77
	5.4 Evaluation	83
	5.5 Conclusions	87
6.	Conclusions & Future Work	89
	Bibliography	91
	Appendix A	95
	Appendix B	98
	Appendix Γ	99
	Appendix Δ	100
	Appendix E	101
	Appendix Z	106

Κεφάλαιο 1. Εισαγωγή

Η οπτική αναγνώριση χειρόγραφων μαθηματικών εκφράσεων είναι ένα ερευνητικό θέμα που έχει απασχολήσει την ερευνητική κοινότητα πάνω από τέσσερις δεκαετίες (**Anderson 1968, Tappert 1990, Blostein and Grbavec 1997, Plamondon and Srihari 2000, Chan and Yeung 2000, Tapia and Rojas 2007**). Κατά τη διάρκεια των ερευνητικών προσπαθειών για την δημιουργία κατάλληλων τεχνολογικών εφαρμογών καθιερώθηκαν οι όροι της οπτικής αναγνώρισης χειρογράφων εντός γραμμής (on-line) και εκτός γραμμής (off-line) (**Anderson 1968**). Η διαφορά των δύο διαδικασιών έγκειται στον τρόπο ανάκτησης και αποθήκευσης των δεδομένων που συνιστούν το χειρόγραφο προς αναγνώριση σήμα. Η είσοδος των συστημάτων αναγνώρισης χαρακτήρων εντός γραμμής (on-line) είναι οι συντεταγμένες των ιχνών του ψηφιακού ίχνους/μελανιού ως χρονικές ακολουθίες, ενώ τα συστήματα εκτός γραμμής (off-line) καλούνται να επεξεργαστούν μια δυαδική εικόνα του χειρόγραφου πρωτοτύπου. Οι τεχνολογικές εφαρμογές που έχουν υλοποιηθεί για οπτική αναγνώριση χειρόγραφων γραφημάτων εντός γραμμής είναι διαθέσιμες σε υπολογιστές με γραφίδα (pen-based computers) ή πιο πρόσφατα με τη χρήση ψηφιακών πενών (digital pens) για ταμπλέτες (tablets) ή για ψηφιακό χαρτί (digital paper) και παρουσιάζουν υψηλά ποσοστά ορθής αναγνώρισης (>95%). Η δυσκολία στην υλοποίηση παρόμοιων συστημάτων για μαθηματικές εκφράσεις έγκειται στην ποικιλία του τρόπου γραφής των μαθηματικών εκφράσεων, στην ποικιλία του μεγέθους των μαθηματικών συμβόλων, ακόμα και για ίδια σύμβολα το μέγεθος ποικίλει (π.χ. εκθέτης, δείκτης) και στη δισδιάστατη τοπολογία τους. Η αναγνώριση μαθηματικών εκφράσεων δυσκολεύει ακόμα περισσότερο, καθώς το πλήθος των συμβόλων που ένα τέτοιο σύστημα έχει να επεξεργαστεί μεγαλώνει. Η βάση δεδομένων εκμάθησης ενός τέτοιου συστήματος κυμαίνεται από σύμβολα συμπεριλαμβανομένων των λατινικών και των ελληνικών γραμμάτων και αριθμών, έως και πιο εξειδικευμένα μαθηματικά σύμβολα, όπως το σύμβολο της άθροισης \sum , του γινομένου \prod , του ολοκληρώματος \int κ.λπ. Επιπλέον, υπάρχει μεγάλη διακύμανση στον τρόπο γραφής του κάθε γραφέα. Τέλος, ο δισδιάστατος τρόπος γραφής των ΜΕ σε αντίθεση με τον τρόπο γραφής του κειμένου, κάνει το πρόβλημα της αναγνώρισης ακόμα πιο δύσκολο.

Δομικά στοιχεία μιας μαθηματικής έκφρασης είναι τα σύμβολα. Ένα σύμβολο μπορεί να αποτελείται από μία ή περισσότερες κονδυλιές (strokes) $S = \{S^m, m = 1, \dots, M\}$ όπου M το πλήθος των κονδυλιών. Ως κονδυλιά θεωρείται η χρονική ακολουθία των συντεταγμένων των σημείων από την χρονική στιγμή που ακουμπάμε τη γραφίδα στην ταμπλέτα μέχρι την χρονική στιγμή που σηκώνουμε τη γραφίδα από τη ταμπλέτα.

Το πρόβλημα της αναγνώρισης ΜΕ μπορεί να χωριστεί σε τρία στάδια:

1. Το στάδιο της αναγνώρισης των μαθηματικών συμβόλων, το οποίο προϋποθέτει την σωστή ομαδοποίηση των κονδυλίων της ME σε σύμβολα.
2. Το στάδιο της αναγνώρισης της τοπολογικής σχέσης που διέπει δύο σύμβολα της ME.
3. Το στάδιο της αναγνώρισης της μαθηματικής δομής που περιγράφει την ME.

Αξίζει να σημειωθεί ότι τα τελευταία χρόνια, η αναγνώριση των μαθηματικών συμβόλων (στάδιο 1) έχει φτάσει σε ποσοστό ακρίβειας πάνω από 95% (**Tapia and Rojas 2007**). Ως εκ τούτου, η έρευνα της αναγνώρισης ME έχει επικεντρωθεί στην πιο αποτελεσματική ομαδοποίηση των κονδυλίων που συνθέτουν τα σύμβολα (στάδιο 1) και στην ανάλυση της τοπολογίας των συμβόλων της ME (στάδιο 2), ώστε να γίνει η σωστή ψηφιακή αναπαράστασή της, κατά την σύνθεση της δομής της (στάδιο 3).

Η συγκεκριμένη εργασία επικεντρώνεται και στα τρία στάδια του προβλήματος της αναγνώρισης χειρόγραφων μαθηματικών εκφράσεων. Η δομή της διατριβής είναι η εξής: Στο Κεφάλαιο 2 περιγράφονται υπάρχοντα συστήματα αναγνώρισης μαθηματικών εκφράσεων.

Στο Κεφάλαιο 3 περιγράφεται αναλυτικά το στάδιο της αναγνώρισης χειρόγραφων μαθηματικών συμβόλων. Στο στάδιο αυτό περιλαμβάνεται και η ανάπτυξη μιας βάσης δεδομένων χειρόγραφων μαθηματικών συμβόλων (ILSP) καθώς και το αντίστοιχο λογισμικό συλλογής χειρόγραφων μαθηματικών συμβόλων. Στην παρούσα εργασία μελετήθηκε η αναγνώριση χειρόγραφων μαθηματικών συμβόλων με τις ακόλουθες δύο τεχνικές (**Simistira et al. 2008**): α) την τεχνική ελαστικού ταιριάσματος (elastic matching) και β) την τεχνική ταύτισης απόστασης με υπόδειγμα (string template matching). Οι δύο αυτές τεχνικές αξιολογήθηκαν στις βάσεις LVZE και ILSP. Η τεχνική ταύτισης απόστασης με υπόδειγμα έδωσε ποσοστό ορθής αναγνώρισης για την Top-1 επιλογή ίσο με 92.13% για την βάση LVZE και 92.35% για την βάση ILSP. Τα ποσοστά αναγνώρισης για την μέθοδο του ελαστικού ταιριάσματος κυμαίνονται σε χαμηλότερα ποσοστά: (i) για την βάση ILSP, 81.02 (top1) και (ii) για την βάση LVZE, 82.87 (top1).

Στο Κεφάλαιο 4 περιγράφεται η αναγνώριση της τοπολογικής σχέσης μεταξύ δύο συμβόλων της ME. Αναφέρονται οι βάσεις δεδομένων χειρόγραφων ME που χρησιμοποιήθηκαν για την αξιολόγηση των προτεινόμενων μεθόδων. Στην παρούσα εργασία αναπτύχθηκαν τρεις ταξινομητές για την εύρεση της τοπολογικής σχέσης μεταξύ δύο συμβόλων με βάση τα γεωμετρικά χαρακτηριστικά των πλαισίων οριοθέτησης (bounded boxes) των συμβόλων:

- (α) ένας ταξινομητής με χρήση ευριστικών κανόνων (heuristics) (**Simistira et al. 2012**),
- (β) ένας ταξινομητής βασισμένος στην τεχνική των μηχανών διανυσμάτων υποστήριξης (SVM) (**Simistira et al. 2013**) και

(γ) ένας ταξινομητής βασισμένος στην τεχνική των νευρωνικών δικτύων (ANN) (**Simistira et al. 2013**)

Οι προτεινόμενες μέθοδοι αξιολογήθηκαν στις βάσεις χειρόγραφων ME CROHME2012 και MathBrush. Το μέσο ποσοστό λάθους του ευριστικού ταξινομητή είναι 12.31% στην βάση CROHME2012. Αντίστοιχα το μέσο ποσοστό λάθους του ταξινομητή SVM είναι 2.87%, στην βάση MathBrush και 2.61% στην βάση CROHME2012. Τέλος, το μέσο ποσοστό λάθους του ταξινομητή ANN είναι 3.35% στην βάση MathBrush και 3.73% στην βάση CROHME2012.

Στο Κεφάλαιο 5 περιγράφεται η αναγνώριση της δομής της ME όπου προτάθηκαν δύο τεχνικές. Η πρώτη τεχνική (**Simistira et al. 2012**) χρησιμοποιεί τον ευριστικό ταξινομητή για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ME και στη συνέχεια ξεκινώντας από το αριστερότερο σύμβολο προσπαθεί να ανασυστήσει τη δομή της ME βρίσκοντας τη βασική γραμμή και τοποθετώντας σταδιακά τα σύμβολα σε αντίστοιχα επίπεδα της ME. Στην δεύτερη τεχνική χρησιμοποιήθηκε ο ταξινομητής SVM για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ME (**Simistira et al. 2014**). Παράλληλα αναπτύχθηκε μία γραμματική τύπου SCFG για την περιγραφή της δομής των ME και προτάθηκε ο αλγόριθμος Cocke–Younger–Kasami (CYK) κατάλληλα τροποποιημένος (**Simistira et al. 2015**) ώστε να μπορεί να αντιμετωπίσει το πρόβλημα της ανάλυσης μαθηματικών εκφράσεων.

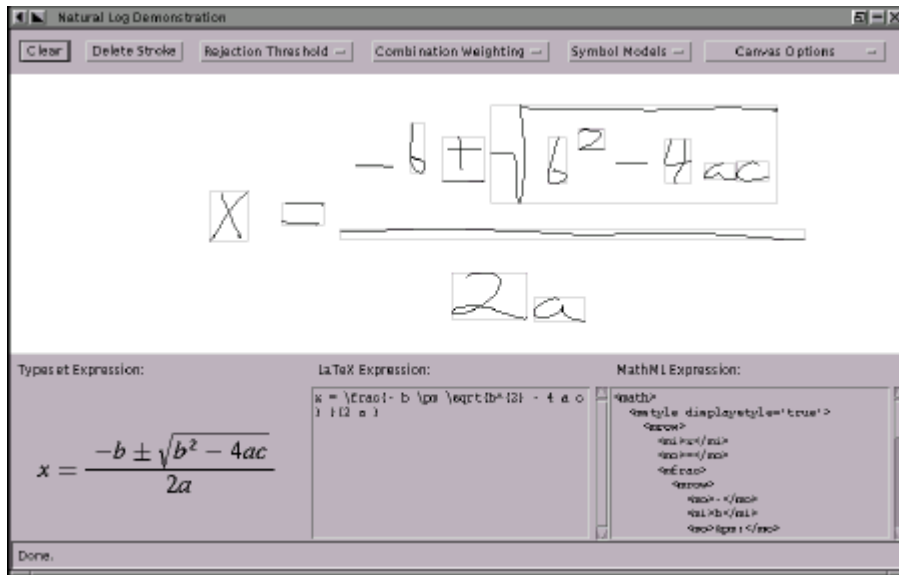
Για την αξιολόγηση του συστήματος, που υλοποιεί την σταδιακή ανάλυση της ME με την τοποθέτηση των συμβόλων σε επίπεδα, χρησιμοποιήθηκε η βάση CROHME 2011 (**Mouchère et al., 2011**).

Για το Μέρος-I της βάσης CROHME2011 το ποσοστό ορθής αναγνώρισης συμβόλων είναι 81.76% και το ποσοστό ορθής αναγνώρισης μαθηματικών εκφράσεων με την πρώτη τεχνική είναι 9.94%. Αντίστοιχα, για το Μέρος-II της βάσης CROHME2011 το ποσοστό ορθής αναγνώρισης συμβόλων είναι 88.90% και το ποσοστό ορθής αναγνώρισης μαθηματικών εκφράσεων είναι 5.45%.

Για την αξιολόγηση της τεχνικής που υλοποιεί την σταδιακή ανάλυση της ME με χρήση στοχαστικών γραμματικών, χρησιμοποιήθηκαν τα σύνολα εξέτασης των βάσεων CROHME2012 και CROHME2013. Τα ποσοστά ορθής αναγνώρισης για τα Μέρη-I,II,III και IV είναι 78.70%, 65.78%, 56.37% και 50.22%. Στην τεχνική αυτή υποθέσαμε ότι όλα τα σύμβολα της ME έχουν αναγνωριστεί σωστά.

Κεφάλαιο 2. Υπάρχοντα συστήματα αναγνώρισης μαθηματικών εκφράσεων

2.1. Natural Log System (1999)



Εικόνα 1. Natural Log System

Το 1999, οι Matsakis et al. ανέπτυξαν ένα applet που έχει την δυνατότητα να αναγνωρίζει απλές μαθηματικές εκφράσεις (**Matsakis 1999**). Το σύστημα έχει την δυνατότητα να αναγνωρίζει σύμβολα όπως: αριθμοί, μικρά γράμματα του λατινικού αλφαβήτου, μερικά γράμματα από το ελληνικό αλφάβητο, σχεσιακούς/δυναδικούς τελεστές, κλάσματα, αθροίσματα, τετραγωνικές ρίζες κ.λπ. Βασικός περιορισμός του συστήματος είναι ότι ο χώρος προτύπων περιέχει πραγματώσεις των συμβόλων από έναν γραφέα και μπορεί να αναγνωρίσει μόνο απλές μαθηματικές εκφράσεις. Επιπλέον, το σύστημα δεν υποστηρίζει ολοκληρώματα και εκθέτες. Η αναγνώριση της μαθηματικής έκφρασης πραγματοποιείται σε 3 στάδια:

1) ταξινόμηση/αναγνώριση συμβόλων: για το στάδιο αυτό έχουν δημιουργηθεί γκαουσιανά μοντέλα που περιγράφουν τα πρότυπα σύμβολα με τα οποία θα συγκριθούν τα υπό εξέταση σύμβολα,

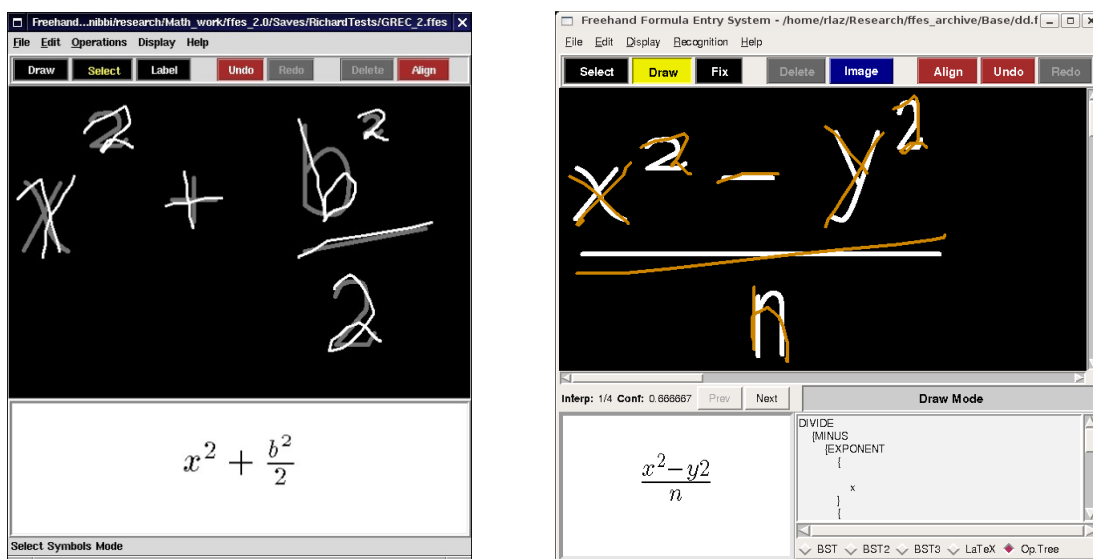
2) κατάτμηση έκφρασης (expression partitioning) σε σύμβολα: οι κωνδυλίες ομαδοποιούνται σε σύμβολα βάση του ταξινομητή συμβόλων. Για κάθε υποψήφιο σύμβολο υπολογίζεται ένα κόστος που ισούται με την λογαριθμική πιθανοφάνεια (log likelihood) πολλαπλασιασμένο με ένα βάρος το οποίο είναι ίσο με το πλήθος των κωνδυλιών που απαρτίζουν το σύμβολο μειωμένο κατά 1. Επιπλέον, υπολογίζεται το ελάχιστο εκτεταμένο δένδρο (minimum spanning tree) μεταξύ των συμβόλων υπολογίζοντας τις αποστάσεις μεταξύ των κεντροειδών τους,

3) τα αναγνωρισμένα σύμβολα τοποθετούνται κατάλληλα ώστε να δημιουργήσουν την δισδιάστατη δομή που περιγράφει την ΜΕ σε Latex και MathML μορφότυπο (Παράρτημα Β). Η δομή της ΜΕ βρίσκεται με χρήση γεωμετρικών χαρακτηριστικών, όπως είναι οι σχέσεις μεταξύ των πλαισίων οριοθέτησης (bounding box) των συμβόλων της έκφρασης. Σαν σύμβολο κλειδί (που ανήκει στην βασική γραμμή της ΜΕ) θεωρείται το μεγαλύτερο σε πλάτος σύμβολο.

Στο στάδιο της προ-επεξεργασίας κάθε κονδυλιά μετατρέπεται σε ένα διάνυσμα χαρακτηριστικών διάστασης 74. Κάθε σύμβολο με N κονδυλιές, επαναδειγματοληπτείται σε $36/N$ ίσα τμήματα χρησιμοποιώντας γραμμική παρεμβολή μεταξύ των σημείων της κονδυλιάς. Τα υπόλοιπα 2 στοιχεία του διανύσματος χαρακτηριστικών είναι το άθροισμα της τετραγωνικής παραγώγου των x και y συντεταγμένων των σημείων της κονδυλιάς ως προς το συνολικό μήκος της κονδυλιάς.

Η βάση συμβόλων που χρησιμοποιήθηκε σ' αυτό το σύστημα αποτελείται από 500 δείγματα τα οποία συλλέχθηκαν από έναν γραφέα με την βοήθεια ενός Wacom tablet και ενός προγράμματος συλλογής δεδομένων γραμμένο σε Java. Το ποσοστό ορθής αναγνώρισης του συστήματος δεν έχει αναφερθεί.

2.2. Freehand Formula Entry System/ DRACULAE (1999)



Εικόνα 2. FFES/DRACULAE

Το σύστημα αναγνώρισης μαθηματικών εκφράσεων FFES (Smithies 1999, Smithies et al. 1999, Smithies et al. 2001, Zanibbi et al. 2001) σχεδιάστηκε και υλοποιήθηκε το 1999 από τους Steve Smithies και Kevin Novins στο πανεπιστήμιο του Otago. Από τότε διάφοροι ερευνητές συνέβαλαν στην βελτίωση του συστήματος.

Η αναγνώριση της ΜΕ πραγματοποιείται σε τρία στάδια: α) ομαδοποίηση κονδυλιών, β) αναγνώριση συμβόλων και γ) ανάλυση τοπολογίας μαθηματικών συμβόλων και μετατροπή της αναγνωρισμένης ΜΕ σε Latex, μέσω του αναλυτή έκφρασης DRACULAE (**Zanibbi et al. 2001, Zanibbi et al. 2002, Zhang et al. 2005**). Επιπλέον δίνει την δυνατότητα στο χρήστη να διορθώσει το αποτέλεσμα της αναγνώρισης. Η ομαδοποίηση κονδυλιών σε σύμβολο γίνεται με την βοήθεια του αναγνωριστή συμβόλων, δίνοντας σε κάθε δυνατό συνδυασμό κονδυλιών ένα βαθμό εμπιστοσύνης (confidence level) το οποίο αντιστοιχεί στο χειρότερα αναγνωρισμένο σύμβολο για κάθε συνδυασμό κονδυλιών. Το σύστημα παράγει όλους τους δυνατούς συνδυασμούς κονδυλιών σε ένα παράθυρο k (βάσει του συνόλου εκμάθησης) γειτονικών κονδυλιών, έχοντας σαν προϋπόθεση ότι ο γραφέας δεν συμπληρώνει σύμβολα που έχει ήδη δημιουργήσει. Το στάδιο της αναγνώρισης των συμβόλων βασίζεται σε έναν ταξινομητή κοντινότερου γείτονα (nearest neighborhood) για ένα 50-διάστατο διάλυμα χαρακτηριστικών (**Smithies et al. 1999**). Το σύστημα εκπαιδεύεται με δεδομένα από τον χρήστη.

Σε επόμενο στάδιο, το πρόγραμμα DRACULAE παίρνει σαν είσοδο την λίστα των αναγνωρισμένων συμβόλων της ΜΕ και τα αντίστοιχα πλαίσια οριοθέτησής τους (bounding boxes) και παράγει την δενδροειδή και στη συνέχεια την Latex αποτύπωση της έκφρασης. Η δενδροειδή αυτή μορφή περιγράφει την ιεραρχική δομή των γραμμών που συνθέτουν την ΜΕ. Το σύστημα αναλύει την ΜΕ σε τρία στάδια:

1) Στο πρώτο στάδιο (Layout Pass) παράγεται, έπειτα από μία σειρά κατάλληλων μετασχηματισμών του αρχικού δέντρου, ένα τελικό BST (Baseline Structure Tree) το οποίο αναπαριστά την δισδιάστατη τοποθέτηση των συμβόλων. Το σύστημα ψάχνει για την βασική γραμμή (baseline) και μετέπειτα χρησιμοποιεί την πληροφορία αυτή για την εύρεση δευτερευουσών γραμμικών δομών. Οι γραμμικές αυτές δομές συνθέτουν μία δενδροειδή αποτύπωση της ΜΕ σε επίπεδα (Baseline Structure Tree). Πιο συγκεκριμένα, στο στάδιο προεπεξεργασίας, για κάθε σύμβολο υπολογίζεται το πλαίσιο οριοθέτησής του (bounding box), το κεντροειδές (centroid) και οι δυνατές περιοχές τοποθέτησης άλλων συμβόλων βάσει της κατηγορίας στην οποία ανήκει. Ορίζονται οι ακόλουθες κατηγορίες συμβόλων: ascender (π.χ. b,d), descender (π.χ. g, j), centered (π.χ. a, e), non-scripted (π.χ. +,>=), open bracket (π.χ. {) , ρίζες και μεταβλητού εύρους σύμβολα (π.χ. Σ, Π).

Έπειτα το σύστημα παράγει την αρχική δομή του δέντρου όπου όλα τα σύμβολα της ΜΕ ανήκουν στην ίδια βασική γραμμή, στοιχισμένα κατά την ελάχιστη x -συντεταγμένη του πλαισίου οριοθέτησής τους. Ακολούθως ψάχνει το σύμβολο εκκίνησης της βασικής γραμμής, το οποίο ορίζεται ως το αριστερότερο σύμβολο της ΜΕ. Έπειτα το σύστημα ψάχνει για τα σύμβολα εκείνα που ανήκουν στην ίδια γραμμή με το σύμβολο εκκίνησης. Στην περίπτωση που δεν ανήκουν όλα τα σύμβολα στη βασική γραμμή, ανατίθενται στην κατάλληλη περιοχή

συμβόλου που ανήκει στην βασική γραμμή. Τα βήματα αυτά γίνονται επαναληπτικά μέχρι να εξετασθούν όλα τα σύμβολα της ΜΕ. Συνοπτικά σε αυτό το στάδιο αναγνωρίζονται οι βασικές γραμμές της ΜΕ και οι περικλειόμενές τους.

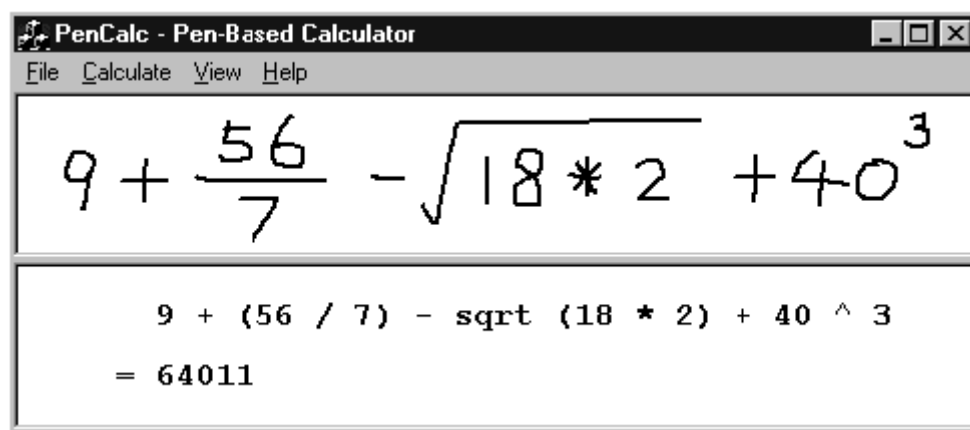
2) Στο δεύτερο στάδιο (Lexical Pass) ομαδοποιούνται και αναγνωρίζονται επιμέρους σύμβολα που αποτελούν υπό-σύμβολα ενός σύνθετου συμβόλου. Π.χ. τα επιμέρους σύμβολα "s", "i" και "n" του σύνθετου συμβόλου "sin", τα σύμβολα < και = του σύνθετου συμβόλου ≤ κ.λπ.

3) Στο τρίτο στάδιο (Expression Analysis Pass) παράγεται ένα δέντρο τελεστών (operator tree) το οποίο προκύπτει από την ανάλυση της δομής της ΜΕ χρησιμοποιώντας κατάλληλη γραμματική.

Για το σύστημα FFES έχει αναφερθεί ποσοστό ακρίβειας αναγνώρισης χαρακτήρων της τάξης του 77%. Το ποσοστό αναγνώρισης του συστήματος μπορεί να φτάσει το 95% αν πρώτα εκπαιδευτεί με τον γραφικό χαρακτήρα του χρήστη/γραφέα.

Το σύστημα FFES/DRACULAE είναι πλέον υπό την εποπτεία του Richard Zanibbi, στο Computer Science Department του Rochester Institute of Technology, USA. Η πιο πρόσφατη έκδοση του συστήματος (v0.4) είναι συμβατή με τα λειτουργικά Linux/MacOSX και αναφέρθηκε τον Αύγουστο του 2007.

2.3. PenCalc (2001)



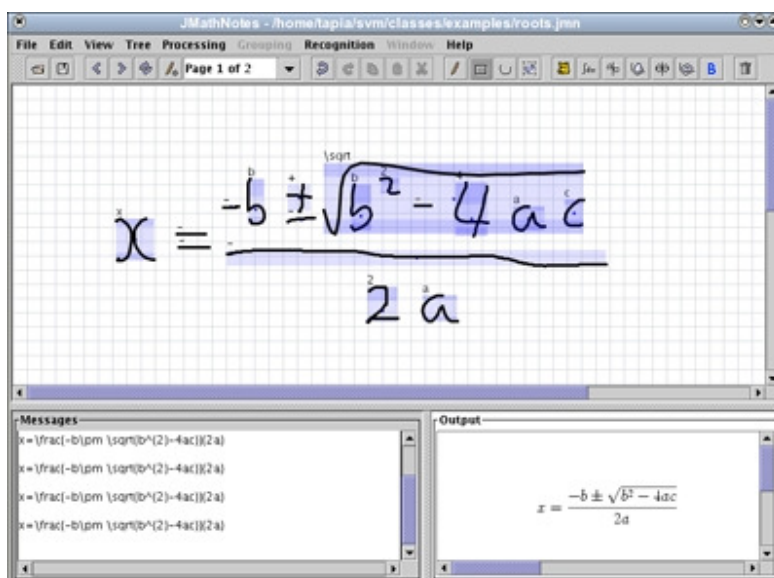
Εικόνα 3. PenCalc

Το **2001**, οι **Chen και Yeung** δημοσίευσαν ένα άρθρο όπου περιέγραφαν την πρώτη αριθμομηχανή που υποστηρίζει γραφίδα (PenCalc). Η αναγνώριση της ΜΕ πραγματοποιείται σε δύο στάδια: α) το στάδιο της αναγνώρισης των συμβόλων της ΜΕ (**Chan and Yeung 1999**), και β) το στάδιο της δομικής ανάλυσης (**Chan and Yeung 2000**) της ΜΕ. Για το στάδιο (α) από τις ακολουθίες συντεταγμένων των σημείων των κονδυλιών του υπό εξέταση συμβόλου εξαγονται τα δομικά αρχέτυπα που περιγράφουν το σύμβολο και κατόπιν

συγκρίνεται με τα σύμβολα της βάσης προτύπων για την ταυτοποίησή τους. Σε περίπτωση που δεν είναι δυνατή η ταυτοποίηση, τότε το προς εξέταση σύμβολο υπόκειται σε μία σειρά παραμορφώσεων ώστε να βρεθεί κάποιο πρότυπο που να ταυτίζεται με το υπό εξέταση σύμβολο. Υπάρχουν 4 επίπεδα παραμόρφωσης: καμία παραμόρφωση, παραμόρφωση αρχικού τύπου (line, up, down, loop, dot), παραμόρφωση κατεύθυνσης (0,1,2,3,4,5,6,7) και συνδυασμός των δύο τελευταίων παραμορφώσεων. Σε περίπτωση αμφισημίας λαμβάνεται υπόψη το ύψος των δομικών αρχέτυπων που συνθέτουν το σύμβολο. Χρησιμοποιήθηκε ένα μέρος της βάσης προτύπων χειρόγραφων του MIT Spoken Language Systems Group (**Kassel 1995**), το οποίο αποτελείται συνολικά από 62 πρότυπα σύμβολα (10 αριθμοί, 26 κεφαλαία γράμματα και 26 μικρά γράμματα) γραμμένα από 150 διαφορετικούς γραφείς (σύνολο 9.300 σύμβολα). Για κάθε κλάση το ποσοστό ορθής αναγνώρισης είναι: 98.6% (αριθμοί), 98.49% (κεφαλαία γράμματα), 98.44% (μικρά γράμματα) και 97.40% (συνολικά για όλες τις κλάσεις).

2.4. JMathNotes (2003)

Το 2003 υλοποιήθηκε από τον Rodriguez (**Rodriguez 2003**) το σύστημα αναγνώρισης χειρόγραφων ME JMathNotes σε γλώσσα Java. Η διεπαφή είναι ιδιαίτερα απλή στην χρήση, δίνοντας επιπλέον την δυνατότητα ρύθμισης των παραμέτρων που αφορούν την μηχανή αναγνώρισης συμβόλων, την έξοδο του συστήματος αναγνώρισης ME κ.ά. Η μέθοδος αναγνώρισης μαθηματικών συμβόλων που χρησιμοποιεί βασίζεται στην τεχνολογία των μηχανών διανυσμάτων υποστήριξης (Support Vector Machines, SVM) και δίνει πολύ χαμηλό σφάλμα αναγνώρισης (1.35%). Η δομική ανάλυση της ME βασίζεται στην παραγωγή ενός ελάχιστου εκτεταμένου δένδρου (minimum spanning tree) και στην ιδέα της περιοχής δράσης των συμβόλων (**Tapia and Rojas 2004**).



Εικόνα 4. JMathNotes

2.5. InftyEditor (2003)

Το σύστημα αναγνώρισης InftyEditor υλοποιήθηκε το 2003 στο Πανεπιστήμιο Kyushu της Ιαπωνίας. Υποστηρίζει δύο τρόπους εισαγωγής μαθηματικών εκφράσεων: α) μέσω του πληκτρολογίου και β) μέσω της γραφίδας ή του ποντικιού. Ανήκει στην ευρύτερη ομάδα προγραμμάτων Infty (**Okamura et al. 1999, Suzuki et al. 2003**) και υποστηρίζει αναγνώριση ME εντός (online) και εκτός γραμμής (offline). Ξεκίνησε το 1995 σαν ερευνητικό πρόγραμμα υποστήριξης ανθρώπων με προβλήματα όρασης.

Τα σύμβολα που υποστηρίζει φαίνονται στον ακόλουθο πίνακα:

Γράμματα λατινικού αλφαβήτου	A-Z, a-z
Αριθμοί	0-9
Γράμματα από το ελληνικό αλφάβητο	Φ, Ψ, Ω, α, β, γ, δ, ε, θ, λ, μ, ν, π, ϕ, ω
Ειδικά σύμβολα	Σ , ∞ , \int , $\sqrt{\quad}$ κλπ.
Διαδικοί τελεστές	+, -, ±, ∓, ×, ÷, !, /, \, ⊗, ∪, ∩
Σχεσιακοί τελεστές	=, ≡, ≠, ∃, ⊂, ⊃, →, ≤, ≥ κλπ.
Συναρτήσεις	lim, log, sin, cos, tan
Παρενθέσεις	(), { }, [],

Για την αντιμετώπιση των αμφισημιών εισάγει κανόνες γραφής και σειρά γραφής των κονδυλιών των συμβόλων.

Για την διάκριση συμβόλων που υπάρχει πιθανότητα να μπερδευτούν, όπως στην περίπτωση των κεφαλαίων/μικρών γραμμάτων, εισάγει κανόνες γραφής. Π.χ. για να διακρίνει το x μικρό από το X κεφάλαιο, το X πρέπει να γραφτεί εισάγοντας 2 οριζόντιες γραμμές στην βάση και την κορυφή του συμβόλου Σ . Επιπλέον ορίζει για κάθε σύμβολο την σειρά γραφής (Εικόνα 5) των κονδυλιών.



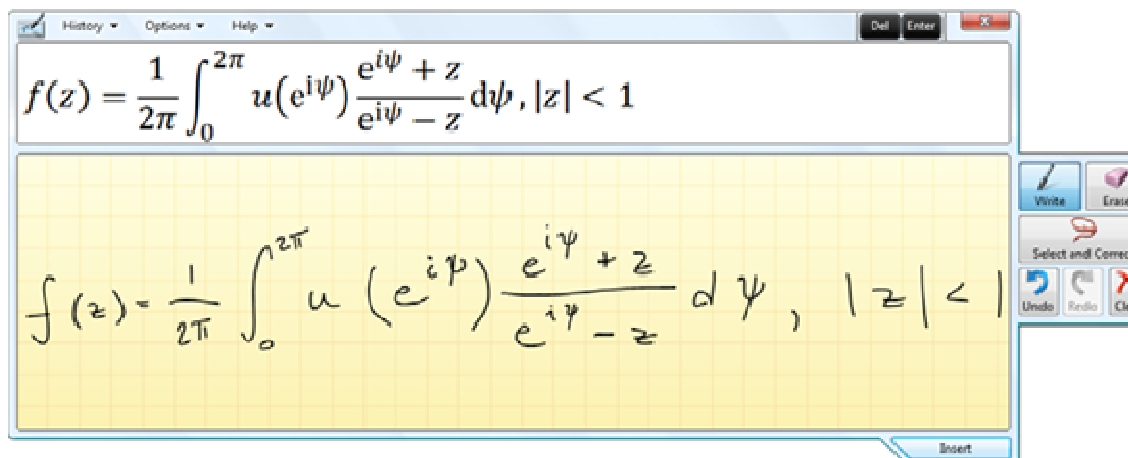
Εικόνα 5. Σειρά γραφής για το σύμβολο A

Το ποσοστό ορθής αναγνώρισης συμβόλων είναι 98.51% και το ποσοστό αναγνώρισης ME είναι 89.6%.

2.6. Math Input (2009)

Το σύστημα αναγνώρισης Math Input (**Panic 2009**) αναπτύχθηκε από το Microsoft Development Center Serbia για το λειτουργικό σύστημα Windows 7. Αποτελείται από δύο

υπό-συστήματα: α) το Math Input Panel και β) το Math Input Control. Καλύπτει μαθηματικές εξισώσεις επιπέδου γυμνασίου/κολλεγίου.

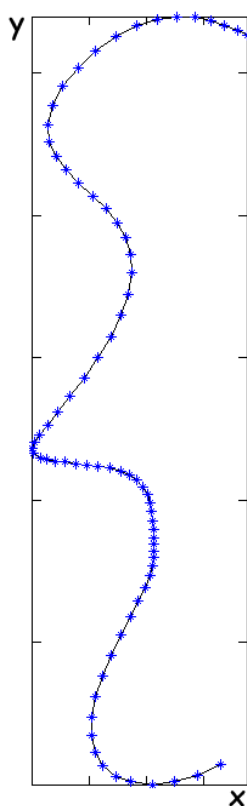


Εικόνα 6. Math Input

Το υποσύστημα Math Input Panel είναι μία αυτόνομη εφαρμογή που επιτρέπει την εισαγωγή χειρόγραφων μαθηματικών παραστάσεων σε windows-based εφαρμογές. Το αποτέλεσμα της αναγνώρισης αποθηκεύεται στο MathML μορφότυπο (Παράρτημα Β). Το υποσύστημα Math Input Control είναι ένα ActiveX UI control που σχεδιάστηκε με σκοπό την καλύτερη ενσωμάτωση σε εφαρμογές Windows (π.χ. επεξεργαστές κειμένου, σημειωματάριο).

Κεφάλαιο 3. Αναγνώριση χειρόγραφων μαθηματικών συμβόλων

Η έξοδος ενός συστήματος γραφίδας-ταμπλέτας είναι το λεγόμενο ψηφιακό μελάνι (digital ink). Η γραφίδα παράγει σημεία που απαρτίζουν το ψηφιακό μελάνι με μια συγκεκριμένη συχνότητα δειγματοληψίας (συνήθως 200 σημεία το δευτερόλεπτο). Το ψηφιακό μελάνι χωρίζεται σε κονδυλιές (strokes). Κάθε κονδυλιά αποτελείται από μια χρονική ακολουθία των συντεταγμένων των σημείων από την χρονική στιγμή που ακουμπάμε τη γραφίδα στην ταμπλέτα μέχρι την χρονική στιγμή που σηκώνουμε τη γραφίδα από τη ταμπλέτα. Υπάρχουν συστήματα γραφίδων που εκτός από την χρονική ακολουθία των συντεταγμένων των σημείων παράγουν επιπλέον και την αντίστοιχη χρονική ακολουθία των πιέσεων ή των γωνιών γραφής. Επιπλέον, υπάρχουν συστήματα που παρέχουν την πληροφορία εκτός από τις συντεταγμένες μιας «εμφανούς» καμπύλης και τις συντεταγμένες μιας «αφανούς» καμπύλης. Αφανής θεωρείται μία καμπύλη που σχηματίζεται από τα σημεία από τη στιγμή που σηκώνουμε τη γραφίδα και μέχρι να ακουμπήσουμε ξανά την γραφίδα στην ταμπλέτα. Στη παρούσα εργασία επικεντρωνόμαστε στην χρήση της χρονικής ακολουθίας των συντεταγμένων των σημείων των εμφανών καμπυλών. Ένα τυπικό παράδειγμα των σημείων μια μονοκονδυλιάς φαίνεται ακολούθως (Εικόνα 7).



Εικόνα 7. Τα σημεία που απαρτίζουν το σύμβολο «{»

Ένα σύμβολο μπορεί να αποτελείται από μία ή περισσότερες κονδυλιές. Για παράδειγμα το σύμβολο της αφαίρεσης (-) αποτελείται από μία κονδυλιά, το σύμβολο της ισότητας (=) από δύο, το σύμβολο του διάφορου (\neq) από τρεις. Ο μέγιστος αριθμός κονδυλιών που μπορεί να αποτελεί ένα σύμβολο εξαρτάται από τον χώρο προτύπων που χρησιμοποιεί το σύστημα αναγνώρισης.

Το πρόβλημα της αναγνώρισης των μαθηματικών συμβόλων που συνθέτουν μία ΜΕ περιλαμβάνει δύο επιμέρους στάδια τα οποία μπορούν να πραγματοποιηθούν είτε ξεχωριστά, είτε ταυτόχρονα. Τα στάδια αυτά είναι (α) η ομαδοποίηση των κονδυλιών ώστε να συνθέσουν ένα σύμβολο και (β) η αναγνώριση του συμβόλου.

3.1. Ανάπτυξη βάσης δεδομένων χειρόγραφων μαθηματικών συμβόλων

Για την ανάπτυξη ενός συστήματος αναγνώρισης χειρόγραφων μαθηματικών παραστάσεων απαιτείται η εκμάθηση ή εκπαίδευση του συστήματος με χειρόγραφα μαθηματικά σύμβολα και παραστάσεις. Έτσι το σύστημα θα μπορεί να αξιολογηθεί με βάση τα αποτελέσματα της αναγνώρισης με ένα διαφορετικό σύνολο χειρόγραφων συμβόλων ή παραστάσεων από τους ίδιους ή διαφορετικούς γραφείς. Βασικό λοιπόν στοιχείο για την υλοποίηση ενός συστήματος αναγνώρισης χειρόγραφων μαθηματικών παραστάσεων είναι η συλλογή χειρόγραφων μαθηματικών συμβόλων και παραστάσεων από την οποία θα προκύψουν τα απαραίτητα πρότυπα για την εκμάθηση του συστήματος καθώς και για την αξιολόγησή του.

Η μεθοδολογία συλλογής της βάσης δεδομένων ILSP, που δημιουργήθηκε σαν βασική υποδομή στο πλαίσιο αυτής της εργασίας, βασίζεται στις ακόλουθες εργασίες:

α) Αρχικά έγινε μια προσπάθεια επιλογής των επιστημονικών περιοχών (κατηγορίες) των μαθηματικών από τις οποίες θα αντληθούν τα σύμβολα. Οι κατηγορίες που επιλέχθηκαν είναι η Άλγεβρα (π.χ. στοιχειώδης άλγεβρα- αριθμοί, αριθμητικές πράξεις, θεωρία αριθμών, γραμμική άλγεβρα, θεωρία πεδίου και πολυώνυμα), Ανάλυση (π.χ. πραγματικές συναρτήσεις, παράγωγοι, ολοκληρώματα, μιγαδικές συναρτήσεις, μερικές παράγωγοι), Γεωμετρία (π.χ. Αλγεβρική Γεωμετρία), Εφαρμοσμένα Μαθηματικά (π.χ. Πιθανότητες και Στατιστική). Η κατηγοριοποίηση των μαθηματικών περιοχών που χρησιμοποιήθηκαν είναι μία εμπειρική κατηγοριοποίηση όπως οργανώνεται από επιστήμονες που ασχολούνται με τους τομείς των μαθηματικών. Σημειώνεται ότι με την κατηγοριοποίηση των μαθηματικών περιοχών έχουν ασχοληθεί επιστήμονες και οργανισμοί διεθνώς. Μία από τις σημαντικότερες προτάσεις για κατηγοριοποίηση των μαθηματικών περιοχών έχει γίνει από την Αμερικάνικη Μαθηματική Κοινότητα (American Mathematical Society) με κύριο στόχο την ταξινόμηση των επιστημονικών άρθρων.

β) Στην συνέχεια έγινε μία προσπάθεια καταγραφής των μαθηματικών συμβόλων που συναντώνται στα ελληνικά βιβλία μαθηματικών της δευτεροβάθμιας εκπαίδευσης (γυμνάσιο, λύκειο). Επίσης, έγινε προσπάθεια καταγραφής των μαθηματικών συμβόλων από ελληνικά και ξένα (κυρίως αμερικάνικα και αγγλικά) βιβλία θεωρητικών ή εφαρμοσμένων μαθηματικών που χρησιμοποιούνται σε θετικές σχολές του Πανεπιστημίου και του Πολυτεχνείου. Τέλος, έγινε καταγραφή μαθηματικών συμβόλων από ένα σύνολο επιστημονικών περιοδικών και βιβλίων αναφοράς μαθηματικών.

Από τα παραπάνω προκύπτουν τα ακόλουθα συμπεράσματα:

- Ο αριθμός των συμβόλων που μπορούν να εμφανισθούν σε μια μαθηματική παράσταση είναι μεγάλος (>400) εξαιρουμένων των λατινικών χαρακτήρων.
- Υπάρχουν περιπτώσεις όπου τα σύμβολα παριστάνονται με περισσότερες από μία μορφές.
- Αρκετά μαθηματικά σύμβολα προκύπτουν από την συνένωση ή υπέρθεση γραφημάτων που αντιστοιχούν σε άλλα μαθηματικά σύμβολα.

Από τον αριθμό των συμβόλων που αναφέρθηκαν επιλέχθηκε ένα σύνολο 186 συμβόλων τα οποία καλύπτουν τις ανάγκες πολλών ομάδων χρηστών. Συνοπτικά οι ομάδες των συμβόλων αυτών είναι οι εξής:

α) τα 24 κεφαλαία γράμματα του ελληνικού αλφαβήτου (Α-Ω),

β) τα 25 μικρά γράμματα του ελληνικού αλφαβήτου (α-ω, ς),

γ) τα 12 κεφαλαία γράμματα του λατινικού αλφαβήτου (μόνο αυτά που διαφοροποιούνται από τα κεφαλαία γράμματα του ελληνικού αλφαβήτου, C, D, F, G, J, L, Q, R, S, U, V, W),

δ) τα 25 μικρά γράμματα του λατινικού αλφαβήτου (εκτός από το ο, το οποίο ταυτίζεται με το ελληνικό γράμμα όμικρον),

ε) τα 10 αριθμητικά ψηφία από 0-9

και ζ) 90 μαθηματικά σύμβολα.

Αναλυτικά στα 90 μαθηματικά σύμβολα περιλαμβάνονται:

- αριθμητικοί τελεστές και σχεσιακά σύμβολα (+, -, <, >, κλπ.),
- λογικοί τελεστές (\cap , \cup κλπ.),
- τελεστές συνόλων,

- σύνολα (N, Z, Q, R, C),
- βέλη (\rightarrow, \uparrow),
- σύμβολα συναρτήσεων (sin, cos, log, κλπ.),
- σύμβολα ορίων (lim, limsup, κ.λπ.),
- υπολογιστικά σύμβολα (\sum, \prod, \int),
- γεωμετρικά σύμβολα (\perp, \circ),
- σημεία στίξης,
- παρενθέσεις, αγκύλες και άγκιστρα,
- και συνοδευτικά σύμβολα (\sim, \wedge , κ.λπ.)

Η βάση που δημιουργήθηκε αποτελείται από 186 σύμβολα που συλλεχτήκαν από 50 γραφείς. Κάθε γραφέας έγραψε κάθε σύμβολο 5 φορές. Επιπλέον, κάθε γραφέας έγραψε 54 εξισώσεις οι οποίες περιέχουν τουλάχιστον από μία φορά όλα τα σύμβολα που προαναφέρθηκαν.

3.2. Λογισμικό συλλογής χειρόγραφων μαθηματικών συμβόλων

Στο πλαίσιο της εργασίας αυτής δημιουργήθηκε μία διεπαφή για την συλλογή χειρόγραφων μαθηματικών συμβόλων. Τα δείγματα αυτά συλλέχθηκαν με ένα tablet PC.

Στη βάση δεδομένων του συστήματος αποθηκεύονται:

- τα στοιχεία του χρήστη (γραφέα)
- οι κωδικοί των συμβόλων
- οι εικόνες των μαθηματικών συμβόλων

Στα στοιχεία του χρήστη περιλαμβάνονται ο κωδικός χρήστη (δίνεται αυτόματα από το σύστημα), το ονοματεπώνυμο, η ηλικία, το φύλλο, πληροφορία για τον αν ο χρήστης είναι αριστερόχειρας ή δεξιόχειρας και η εθνικότητα. Στα στοιχεία που αφορούν την κωδικοποίηση των συμβόλων δίνονται ο κωδικός του συμβόλου με βάση την κωδικοποίηση Unicode (Παράρτημα Γ) και η ονομασία του συμβόλου.

Για κάθε σύμβολο αποθηκεύονται:

- η γραφική απεικόνιση συμβόλου σε μορφότυπο gif ή tiff

- η γραφική απεικόνιση συμβόλου σε μορφότυπο isf (Ink Serialization Format, βλ. Παράρτημα Δ)
- οι συντεταγμένες (x,y) των σημείων ανά stroke και ο χρόνος έναρξης κάθε κονδυλιάς
- η πίεση που αντιστοιχεί σε κάθε σημείο της κονδυλιάς (τιμές από 0 έως 255).

Το λογισμικό συλλογής χειρόγραφων μαθηματικών συμβόλων αποτελείται από 16 καρτέλες συλλογής δεδομένων (Εικόνα 8): 1) User data, 2) Greek Letters – caps, 3) Greek Letters – small, 4) Latin Letters – caps, 5) Latin Letters – small, 6) Numbers, 7) Arrows, 8) Accents, 9) Comparison symbols, 10) Constructs, 11) Functions (1), 12) Functions (2), 13) Functions (3), 14) Operators, 15) Set symbols και 16) Various.

Κάθε καρτέλα συλλογής δεδομένων αποτελείται από ισάριθμα κουτιά εισαγωγής και απεικόνισης του προς γραφή συμβόλου. Πιο συγκεκριμένα στην καρτέλα “Greek caps” υπάρχουν τα 24 γράμματα του ελληνικού αλφαβήτου σε κεφαλαία μορφή και κάτω από κάθε γράμμα η περιοχή εισαγωγής των δεδομένων του χρήστη. Επιπλέον σε κάθε καρτέλα (εκτός από την καρτέλα “user data”), δίνονται οι παρακάτω επιλογές:

Ink: Ενεργοποιώντας αυτή την επιλογή ο χρήστης μπορεί να γράψει. Στην έναρξη του προγράμματος η επιλογή αυτή είναι ενεργοποιημένη.

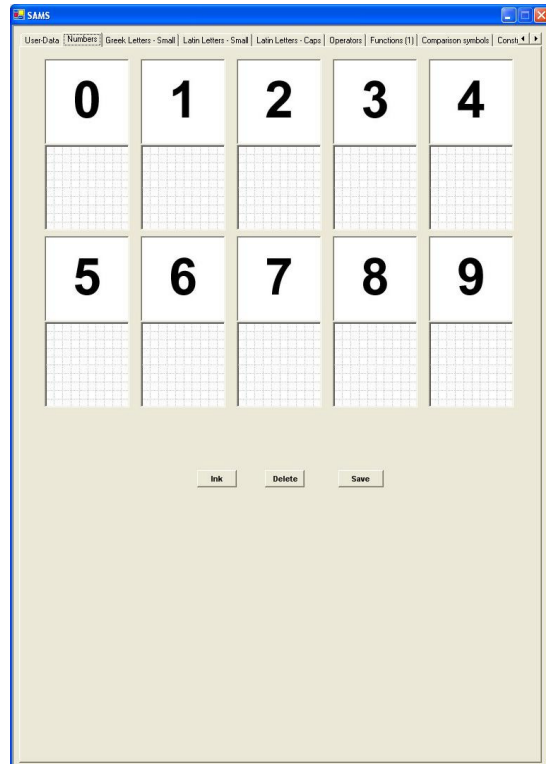
Delete: Ενεργοποιώντας αυτή την επιλογή ο χρήστης μπορεί να διαγράψει. Στην έναρξη του προγράμματος η επιλογή αυτή είναι απενεργοποιημένη.

Save: Ενεργοποιώντας αυτή την επιλογή ο χρήστης μπορεί να αποθηκεύσει τα δεδομένα της επιλεγμένης καρτέλας. Για κάθε σύμβολο αποθηκεύεται:

- η γραφική απεικόνιση συμβόλου σε μορφή gif (Graphics Interchange Format),
- η γραφική απεικόνιση συμβόλου σε μορφή isf (Ink Serialazation Format) ,
- οι συντεταγμένες (x,y) των σημείων των κονδυλιών,
- ο χρόνος αποτύπωσης των σημείων των κονδυλιών και
- η πίεση ανά κονδυλιά (τιμές από 0 έως 255).



(α)



(β)

Εικόνα 8. Καρτέλες του προγράμματος συλλογής χειρόγραφων μαθηματικών συμβόλων α) καρτέλα «Greek Letters – caps» και β) καρτέλα «Numbers »

3.3. Διαθέσιμες τεχνικές αναγνώρισης συμβόλων

Οι πιο διαδεδομένες τεχνικές που έχουν εφαρμοστεί για την αναγνώριση συμβόλου κατηγοριοποιούνται σε τρεις μεγάλες κατηγορίες με βάση τη μεθοδολογία που ακολουθούν. Συγκεκριμένα, διακρίνονται σε αλγορίθμους που εφαρμόζουν στατιστικές τεχνικές όπως HMM και NN (**Rabiner 1989**), ελαστικού ταιριάσματος (**Uchida and Sakoe 2005**) και μεθόδους βασισμένες σε συνδυασμό ταξινομητών.

A) Τεχνικές HMM/NN

Στην ερευνητική εργασία των **Winkler et al. το 1996**, η ομαδοποίηση των κονδυλιών γίνεται βάση του αποτελέσματος που δίνει ο αναγνωριστής. Αρχικά, ο αλγόριθμος παράγει όλους του δυνατούς συνδυασμούς κονδυλιών οι οποίοι κατηγοριοποιούνται βάση ενός HMM ταξινομητή. Τα χαρακτηριστικά που χρησιμοποιούνται βασίζονται στην αναπαράσταση του συμβόλου με online και offline χαρακτηριστικά. Στην έξοδο του ταξινομητή αντιστοιχίζεται μία τιμή πιθανοφάνειας για κάθε συνδυασμό κονδυλιών. Ο συνδυασμός κονδυλιών με την υψηλότερη πιθανοφάνεια επιλέγεται ως η βέλτιστη ομαδοποίηση των κονδυλιών σε σύμβολο. Μειονέκτημα αυτής της μεθόδου είναι η εξάρτησή της από τη σειρά γραφής των κονδυλιών. Με άλλα λόγια, αν τα σύμβολα της ME δεν έχουν γραφτεί με κατεύθυνση από αριστερά προς τα δεξιά και από πάνω προς τα κάτω, το ποσοστό αναγνώρισης μειώνεται δραματικά.

Οι **Hu και Zanibbi το 2011** εφάρμοσαν τεχνικές HMM για την αναγνώριση μεμονωμένων χειρόγραφων μαθηματικών συμβόλων. Η μοντελοποίηση των κλάσεων γίνεται με μίγματα κανονικών κατανομών (Gaussian Mixture Models). Για την αρχικοποίηση των παραμέτρων των GMM μοντέλων χρησιμοποιείται μία παραλλαγή του τμηματικού αλγορίθμου K-μέσων (Segmental-K-means) για Υπονοούμενα Μαρκοβιανά Μοντέλα (HMMs). Για κάθε κλάση συμβόλου χρησιμοποιείται ένα συνεχές αριστερά προς τα δεξιά HMM έξι καταστάσεων. Τέσσερα τοπικά χαρακτηριστικά χρησιμοποιούνται για την περιγραφή κάθε κλάσης: (1) το συνημίτονο της κλίσης μεταξύ δύο διαδοχικών σημείων της καμπύλης της κονδυλιάς, (2) η γ-συντεταγμένη των σημείων των κονδυλιών, (3) η πληροφορία αν το ψηφιακό στυλό ακουμπάει την γραφίδα ή όχι και (4) το πρόσημο της γωνίας μεταξύ δύο σημείων της καμπύλης. Το σύστημα αξιολογήθηκε σε μία βάση δεδομένων 93 διακριτών συμβόλων, αποτελούμενη από 20,281 πρότυπα εκπαίδευσης και 2,202 πρότυπα εξέτασης (**MacLean et al. 2010**). Τα ποσοστά ορθής αναγνώρισης των συμβόλων είναι 82.9% (top-1) και 97.8% (top-5).

Οι τεχνικές HMM δεν χρειάζονται κατατμημένα δεδομένα εισόδου αλλά δεν μπορούν να αξιοποιήσουν την πληροφορία που δίνεται από το περιεχόμενο της ME (π.χ. από την σύνταξη των γειτονικών συμβόλων), ενώ τα νευρωνικά δίκτυα μπορούν να αξιοποιήσουν την πληροφορία περιεχομένου αλλά χρειάζονται κατετμημένα δεδομένα εισόδου για την εκπαίδευση των μοντέλων τους. Για το λόγο αυτό αναπτύχθηκαν υβριδικά συστήματα αναγνώρισης όπου συνδιάζονται HMM και RNN (Recurrent Neural Network). Παρόλα αυτά η αρχιτεκτονική αυτή δεν αξιοποιεί τις δυνατότητες των νευρωνικών δικτύων στο έπαρκο. Επιπλέον τα RNN πάσχουν από το πρόβλημα του vanishing gradient μετά από απόσταση 10 χρονικών βημάτων (**Hochreiter et al. 2001**). Οι **Graves et al. το 2009** χρησιμοποίησαν νευρωνικά δίκτυα τύπου Recurrent Neural Networks (RNN) και ειδικά την αρχιτεκτονική των Long Short-Term Memory για την αναγνώριση χειρόγραφου κειμένου τα οποία μπορούν υπερβούν αυτό τον σκόπελο μέσω των κελιών μνήμης που διαθέτουν. Το **2013**, οι **Alvaro et al.** εφάρμοσαν την ίδια τεχνική για την αναγνώριση μαθηματικών συμβόλων. Οι **Golubitsky and Watt το 2010** και **Keshari and Watt το 2008** χρησιμοποίησαν μηχανές διανυσμάτων υποστήριξης για την αντιμετώπιση του προβλήματος της αναγνώρισης συμβόλων.

B) Ελαστικό ταιρίασμα (elastic matching)

Οι **MacLean και Labahn, το 2010** παρουσίασαν μία παραλλαγή του ελαστικού ταιριάσματος που παρουσιάστηκε στην εργασία του **Tappert το 1982** με σκοπό την βελτίωση του υπολογιστικού κόστους που απαιτεί η τεχνική αυτή. Τα αποτελέσματα που παρουσίασαν είναι παρόμοια με αυτά της πρωτότυπης μεθόδου, απαιτώντας γραμμικό χρόνο για τον υπολογισμό της σύγκρισης.

Οι **Shi, Li και Soong** το **2007** παρουσίασαν μία τεχνική βασισμένη στο ελαστικό ταίριασμα με σκοπό την βελτιστοποίηση της ομαδοποίησης των κονδυλίων σε σύμβολα και της αναγνώρισης των συμβόλων. Ο αλγόριθμος δυναμικού προγραμματισμού εξετάζει όλους τους δυνατούς συνδυασμούς ομαδοποίησης κονδυλίων σε σύμβολα (σε ένα παράθυρο L κονδυλίων ανά σύμβολο) για την βέλτιστη ομαδοποίηση. Σε ένα σύνολο εξέτασης αποτελούμενο από 2.500 εκφράσεις, ο αλγόριθμος μπόρεσε να αναγνωρίσει σωστά το 96.6%.

Οι **Shilman et al.** το **2004** παρουσίασαν μία μέθοδο ομαδοποίησης των κονδυλίων βασισμένη σε έναν γράφο γειτνίασης. Κόμβοι του γράφου είναι οι κονδυλιές. Όποιες κονδυλιές βρίσκονται «κοντά» ομαδοποιούνται. Όταν ο γράφος ολοκληρωθεί, χωρίζεται σε επιμέρους ομάδες ενωμένων κόμβων. Ο μέγιστος αριθμός κόμβων που μπορεί να έχει μία ομάδα είναι ίσος με k. Κάθε μία από αυτές τις ομάδες θεωρείται ως ένα μεμονωμένο σύμβολο το οποίο ταξινομείται μέσω ενός μοντέλου εκπαίδευσης. Σε κάθε ομαδοποίηση δίνεται ένα αρνητικό κόστος πιθανοφάνειας. Για την εύρεση της βέλτιστης ομαδοποίησης χρησιμοποιείται ένας αλγόριθμος δυναμικού προγραμματισμού. Το κυριότερο πλεονέκτημα αυτής της μεθόδου είναι ότι δεν είναι εξαρτώμενη από την σειρά γραφής των κονδυλίων. Η μέθοδος αυτή παρουσιάζει ποσοστό αναγνώρισης ίσο με 94%.

Ο **Tappert** το **1982** χρησιμοποιεί μία μέθοδο ομοιότητας βασισμένη στο ελαστικό ταίριασμα. Ο τρόπος υπολογισμού του βαθμού ομοιότητας βασίζεται στην αντιστοίχιση των σημείων που απαρτίζουν το υπό εξέταση σύμβολο με τα σημεία του πρότυπου συμβόλου χρησιμοποιώντας τον αλγόριθμο δυναμικού ταιριάσματος (DTW - Dynamic Time Warping). Ως βέλτιστο ταίριασμα θεωρείται εκείνο που προκύπτει από την ελαχιστοποίηση του αθροίσματος του κόστους ταιριάσματος των συμβόλων. Ο αλγόριθμος αυτός επιτρέπει την σύγκριση συμβόλων με διαφορετικό πλήθος σημείων.**Γ) Συνδυασμός ταξινομητών**

Οι **Garain και Chaudhuri** το **2004** χρησιμοποιούν μία μέθοδο για την αναγνώριση χειρόγραφων μαθηματικών συμβόλων που βασίζεται στην κωδικοποίηση των σημείων της καμπύλης ενός συμβόλου κατά Freeman και στο τοπικό μήκος μεταξύ δύο σημείων της καμπύλης. Χρησιμοποιούνται δύο ταξινομητές, ένας ταξινομητής HMM και ένας βασισμένος στον κοντινότερο γείτονα. Η μέθοδος αξιολογήθηκε με μία βάση εκφράσεων αποτελούμενη από 98.060 και κατάφερε να αναγνωρίσει σωστά το 93% των συμβόλων.

Οι **LaViola και Zelenik** το **2007** εφάρμοσαν την μέθοδο της προσαρμοστικής ώθησης AdaBoost (**Freund and Schapire 1995**) χρησιμοποιώντας στατιστικά και γεωμετρικά χαρακτηριστικά των κονδυλίων (π.χ. ο αριθμός κονδυλίων συμβόλου, τα σημεία τομής μεταξύ κονδυλίων, η αναλογία ύψους/πλάτους συμβόλου κ.λπ.) και την μηχανή αναγνώρισης χειρόγραφων χαρακτήρων της Microsoft. Η συγκεκριμένη τεχνική εφαρμόστηκε σε σύνολο 6.336 συμβόλων και μπόρεσε να αναγνωρίσει σωστά το 95.8% των συμβόλων. Να σημειώσουμε ότι το σύστημα της Microsoft δεν υποστηρίζει την αναγνώριση ορισμένων

μαθηματικών συμβόλων και για την συγκεκριμένη βάση συμβόλων μπόρεσε να αναγνωρίσει σωστά το 91.3%. Η βάση που χρησιμοποιήθηκε για αξιολόγηση αποτελείται από 48 διαφορετικά σύμβολα (a-z, 0-9, Σ, (,), -, ρίζα, ολοκλήρωμα, αγκύλη, μικρότερο από, μεγαλύτερο από, +, διάφορο και την λέξη else).

Τέλος να σημειώσουμε ότι έχουν αναπτυχθεί τεχνικές αναγνώρισης συμβόλων με χρήση μονο των online χαρακτηριστικών (**MacLean και Labahn 2010, Hu και Zanibbi 2011, Alvaro et al. 2013**), ή και με συνδυασμό online/offline χαρακτηριστικών (**Winkler 1996, Keshari και Watt, 2008, Alvaro et al. 2014**).

3.4. Προτεινόμενες τεχνικές

Στο πλαίσιο της παρούσας εργασίας μελετήθηκε η αναγνώριση χειρόγραφων μαθηματικών συμβόλων με τις ακόλουθες δύο τεχνικές:

- α) τεχνική ελαστικού ταιριάσματος (elastic matching)
- β) τεχνική ταύτισης απόστασης με υπόδειγμα (string template matching)

Ακολούθως περιγράφονται οι δύο αυτές τεχνικές.

3.4.1 Τεχνική ελαστικού ταιριάσματος

Για την αναγνώριση μαθηματικών συμβόλων, οι ακολουθίες συντεταγμένων που περιγράφουν το σύμβολο αρχικά υπόκεινται στα ακόλουθα στάδια προ-επεξεργασίας:

- α) απαλοιφή των επαναλαμβανόμενων σημείων
- β) κανονικοποίηση ως προς την ταχύτητα γραφής
- γ) κανονικοποίηση ως προς το μέγεθος

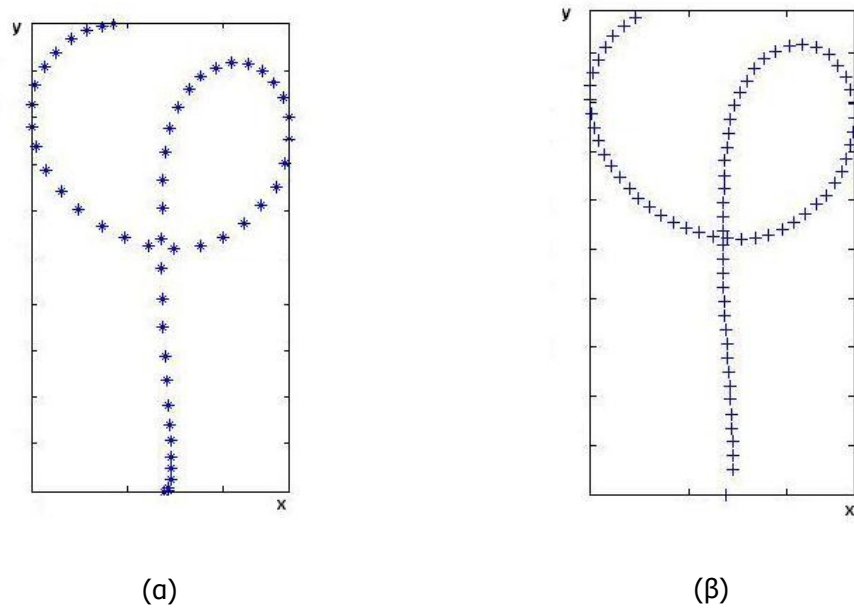
Κατά το πρώτο στάδιο προ-επεξεργασίας απαλείφονται τυχόν επαναλαμβανόμενα σημεία. Ως επαναλαμβανόμενα σημεία θεωρούνται τα σημεία με τις ίδιες ή πολύ κοντινές συντεταγμένες. Τα σημεία αυτά επισημαίνονται και στη συνέχεια αφαιρούνται από τη χρονική ακολουθία των σημείων.

Κατά το στάδιο της κανονικοποίησης ως προς την ταχύτητα γραφής γίνεται επαναδειγματοληψία της χρονικής ακολουθίας των συντεταγμένων ώστε κάθε σημείο να ισαπέχει από το προηγούμενό του. Τα νέα σημεία της ακολουθίας προκύπτουν από γραμμική παρεμβολή ή αποδεκατισμό.

Στο επόμενο στάδιο της προ-επεξεργασίας, δηλ. στην κανονικοποίηση ως προς το μέγεθος, τα νέα σημεία προκύπτουν ώστε να μένει αμετάβλητος ο λόγος πλάτους-ύψους του συμβόλου

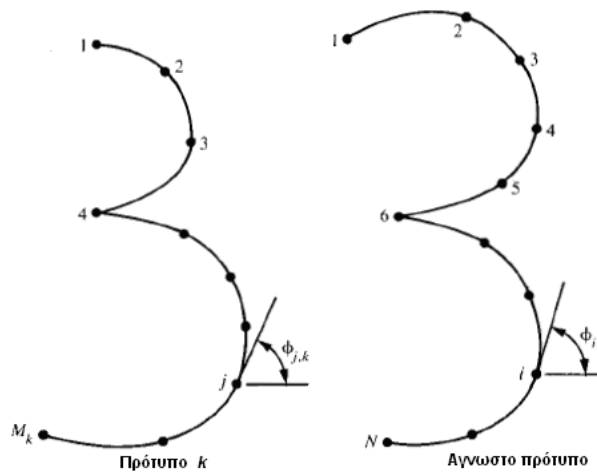
και το πλάτος να μην ξεπερνά ένα συγκεκριμένο κατώφλι, π.χ. το 1. Με αυτόν τον τρόπο το σύμβολο είναι εγγεγραμμένο σε ένα ορθογώνιο του οποίου το μέγιστο πλάτος είναι ίσο με το κατώφλι που επιλέξαμε. Ένα άλλο είδος κανονικοποίησης ως προς το μέγεθος του συμβόλου σχετίζεται με το μήκος της καμπύλης του συμβόλου. Σε αυτή την περίπτωση, τα σημεία κανονικοποιούνται διαιρώντας με το συνολικό μήκος της καμπύλης.

Στην Εικόνα 9 παρουσιάζονται τα αποτελέσματα προ-επεξεργασίας για το σύμβολο 'φ'.



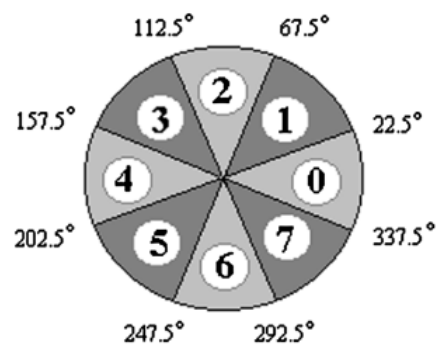
Εικόνα 9. (α) Χρονική ακολουθία σημείων μονοκονδυλιάς και (β) Κανονικοποιημένη ακολουθία σημείων μονοκονδυλιάς.

Κάθε σύμβολο μοντελοποιείται ως μια ακολουθία γωνιών της εφαιπτομένης με την οριζόντια στα σημεία που την απαρτίζουν. Ένα παράδειγμα, τέτοιας μοντελοποίησης παρουσιάζεται στην Εικόνα 10.



Εικόνα 10. Γραφική αναπαράσταση της μοντελοποίησης.

Οι γωνίες που υπολογίζονται με τον παραπάνω εικονιζόμενο τρόπο αντιστοιχίζονται κατά την κωδικοποίηση Freeman (**Freeman 1974**) στους ακεραίους από το 0 έως το 7 (Εικόνα 11).



Εικόνα 11. Κωδικοποίηση Freeman.

Ο κώδικας Freeman χρησιμοποιείται συνήθως σε μια γειτονιά τεσσάρων ή οκτώ κατευθύνσεων. Στην παρούσα εργασία χρησιμοποιούμε την κωδικοποίηση Freeman οκτώ κατευθύνσεων. Βασικές παράμετροι στην κωδικοποίηση Freeman είναι:

- η επιλογή του σημείου έναρξης. Στη συγκεκριμένη εργασία ως σημείο έναρξης θεωρούμε το πρώτο χρονικά σημείο μιας κονδυλιάς.
- ο προσδιορισμός της κατεύθυνσης κατά την οποία πρέπει να κινηθούμε ώστε να βρούμε το επόμενο σημείο. Στη συγκεκριμένη εργασία λαμβάνουμε υπόψη και τις δύο κατευθύνσεις κατά τις οποίες πρέπει να κινηθούμε ώστε να βρούμε το επόμενο σημείο της καμπύλης της κονδυλιάς.

Η μέθοδος σύγκρισης εξαρτάται από την αναπαράσταση του κάθε συμβόλου στο χώρο των χαρακτηριστικών. Στην παρούσα εργασία χρησιμοποιήθηκε η μέθοδος του ελαστικού ταίριασματος που βασίζεται στο δυναμικό προγραμματισμό (Εικόνα 12). Η μέθοδος αυτή, χρησιμοποιείται για απευθείας σύγκριση των επιμέρους συνιστωσών των διανυσμάτων που αναπαριστούν ένα σύμβολο με τη χρήση μιας μετρικής, όπως η Ευκλείδεια απόσταση, για την σύγκριση ζεύγους σημείων και τη χρήση δυναμικού προγραμματισμού για τον συνολικό υπολογισμό της απόστασης μεταξύ δύο προτύπων. Το ελαστικό ταίριασμα πραγματοποιείται και για ακολουθίες διανυσμάτων διαφορετικού μήκους. Η απόσταση ενός σημείου i ενός αγνώστου προτύπου με N σημεία από ένα σημείο j του προτύπου αναφοράς k με M σημεία ορίζεται ως (βλ. Εικόνα 10):

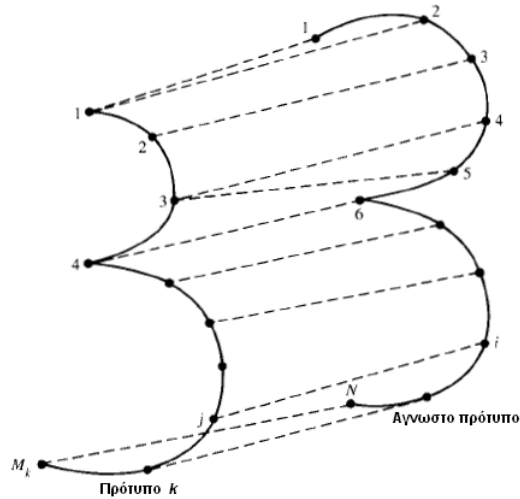
$$d(i,j;k)=|y_i-y_{j,k}|+\min\{|\varphi_i-\varphi_{j,k}|, 2\pi-|\varphi_i-\varphi_{j,k}|\} \quad (1)$$

Όπου φ_i η γωνία του αγνώστου προτύπου στην y_i συντεταγμένη, $\varphi_{j,k}$ η αντίστοιχη γωνία του προτύπου k στην $y_{j,k}$ συντεταγμένη. Παρατηρήστε ότι η απόσταση εξαρτάται από τη διαφορά των υψών (y -συντεταγμένη) και τη διαφορά των γωνιών φ . Σημειώστε, ότι η παραπάνω εξίσωση μπορεί να προσαρμοστεί ανάλογα στην περίπτωση άλλου χώρου χαρακτηριστικών.

Χρησιμοποιώντας την παρακάτω εξίσωση δυναμικού προγραμματισμού:

$$D(i,j;k)=d(i,j;k) + \min\{D(i-1,j;k), D(i-1,j-1;k), D(i-1,j-2;k)\} \quad (2)$$

για $j \geq 2$, η συνολική απόσταση του αγνώστου προτύπου από το πρότυπο αναφοράς δίνεται από την τιμή $D(N,M_k;k)$. Όπως φαίνεται από την παραπάνω εξίσωση δυναμικού προγραμματισμού η συνολική απόσταση του σημείου i του αγνώστου προτύπου από το σημείο j του προτύπου αναφοράς k είναι συνάρτηση των συνολικών αποστάσεων του προηγούμενου σημείου του αγνώστου προτύπου, δηλ. του $i-1$, από τα σημεία j , $j-1$ και $j-2$. Σημειώστε ότι η παραπάνω αναδρομική εξάρτηση μπορεί να επεκταθεί σε μεγαλύτερο παράθυρο προηγούμενων σημείων. Ο κοντινότερος γείτονας από τα πρότυπα αναφοράς αντιστοιχεί στο σύμβολο που αναγνωρίζεται.



Εικόνα 12. Σχηματική αναπαράσταση της τεχνικής του ελαστικού ταιριάσματος.

3.4.2 Τεχνική ταύτισης απόστασης με υπόδειγμα

Στην μέθοδο αυτή το πρόβλημα της αναγνώρισης του συμβόλου έχει αντιμετωπιστεί με ταυτόχρονη ομαδοποίηση κονδυλιών σε σύμβολο και αναγνώριση του σχηματιζόμενου συμβόλου. Σαν βέλτιστη ομαδοποίηση θεωρείται η ομαδοποίηση αυτή που δίνει το μεγαλύτερο ποσοστό αναγνώρισης συγκρινόμενο με ένα από τα πρότυπα αναφοράς.

Αρχικά ομαδοποιούμε τις κονδυλιές της ΜΕ βάση της χρονικής σειράς με την οποία αυτά δημιουργήθηκαν. Σαν βέλτιστη ομαδοποίηση θεωρούμε την ομάδα κονδυλιών που δίνει το μεγαλύτερο ποσοστό αναγνώρισης. Το στάδιο αυτό εκτελείται αναδρομικά για όλες τις κονδυλιές και σε ένα παράθυρο k χρονικά γειτονικών κονδυλιών. Η τιμή του k εξαρτάται από την βάση δεδομένων που χρησιμοποιείται για την αναγνώριση, δηλ. από το μέγιστο αριθμό κονδυλιών ανά σύμβολο. Στην παρούσα εργασία χρησιμοποιήθηκε η βάση δεδομένων χειρόγραφων μαθηματικών εκφράσεων CROHME (βλ. παράγραφο 4.1) και η τιμή του k είναι ίση με 4. Με άλλα λόγια, στη βάση αυτή ο μέγιστος αριθμός κονδυλιών ανά σύμβολο είναι ίσος με 4.

$$\begin{array}{c}
 1 \\
 c^2 \\
 2
 \end{array}
 =
 \begin{array}{c}
 3 \\
 4 \\
 5 \\
 6 \\
 7 \\
 \frac{1}{2} \\
 8 \\
 d^2 \\
 9 \\
 10
 \end{array}
 \quad
 \begin{array}{c}
 1 \\
 c^2 \\
 2 \\
 3,4 \\
 5 \\
 6 \\
 7 \\
 \frac{1}{2} \\
 8,9 \\
 d^2 \\
 10
 \end{array}$$

Εικόνα 13. (α) Χρονική ακολουθία κονδυλιών και (β) Ομαδοποίηση κονδυλιών σε σύμβολα.

Έστω η μαθηματική έκφραση $c^2 = \frac{1}{2}d^2$ (Εικόνα 13). Εξετάζοντας τις κονδυλιές με αριθμό 8,

9 και 10 προκύπτει ότι η βέλτιστη ομαδοποίηση κονδυλιών είναι αυτή που αποτελείται από την 8^η και 9^η κονδυλιά που δίνει και το μεγαλύτερο ποσοστό αναγνώρισης συμβόλου. Η ομαδοποίηση αυτή των κονδυλιών αντιστοιχεί στο σύμβολο «d».

Για την αναγνώριση του υπό εξέταση συμβόλου εφαρμόζεται η ακόλουθη τεχνική ταύτισης απόστασης με υπόδειγμα (template matching).

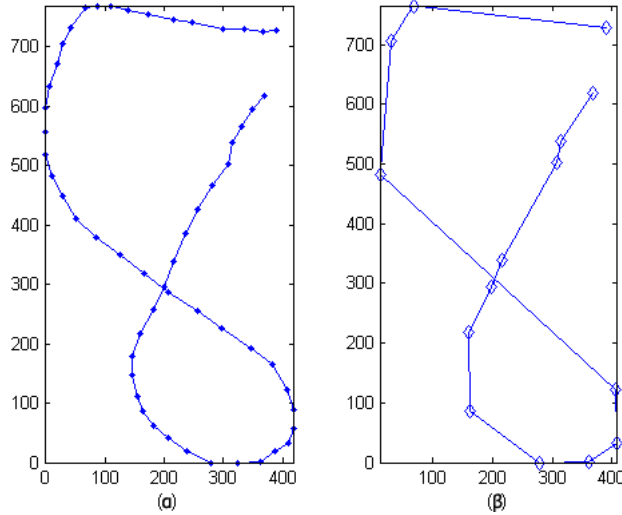
Έστω, $S = \{S^m, m = 1, \dots, M\}$ η χρονική ακολουθία κονδυλιών του υπό εξέταση συμβόλου, όπου S^m η m-ιστή κονδυλιά και M ο αριθμός των κονδυλιών που απαρτίζουν το σύμβολο. Κάθε μία κονδυλιά αναπαρίσταται από την χρονική ακολουθία των συντεταγμένων της ως $S^m = \{p_n^m, n = 1, \dots, N_m\}$, όπου $p_n^m = (x_n^m, y_n^m)$ είναι οι συντεταγμένες του n-οστού σημείου της m κονδυλιάς και N_m ο συνολικός αριθμός των σημείων της m κονδυλιάς. Η κατεύθυνση στο p_n^m σημείο της κονδυλιάς ορίζεται ως η κλίση της γωνίας που σχηματίζεται μεταξύ του ευθύγραμμου τμήματος που ενώνει δύο διαδοχικά σημεία p_n^m και p_{n+1}^m της κονδυλιάς και του οριζόντιου άξονα:

$$\theta_n^m = \tan^{-1} \left(\frac{x_{n+1}^m - x_n^m}{y_{n+1}^m - y_n^m} \right) \quad (3)$$

Η γωνία που προκύπτει από τον προηγούμενο υπολογισμό αντιστοιχίζεται στον κωδικό Freeman 8 επιπέδων και παίρνει μία ακέραια τιμή από το σύνολο $\{0, 1, \dots, 7\}$.

Συνεπώς, η κονδυλιά m μπορεί να περιγραφεί ως μία ακολουθία $S^m = \{c_n^m, n = 1 \dots N_m - 1\}$ από $c_n^m \in \{0, 1, \dots, 7\}$ κωδικούς κατεύθυνσης.

Ορίζουμε ως, τα πιο σημαντικά σημεία της κονδυλιάς εκείνα τα σημεία για τα οποία η διαφορά μεταξύ δύο διαδοχικών κωδικών Freeman που τα περιγράφουν είναι θετικός αριθμός. Η διαφορά αυτή υπολογίζεται μετρώντας τις αλλαγές μεταξύ δύο διαδοχικών στοιχείων της ακολουθίας κωδικών Freeman (**Freeman 1974, Gonzalez and Woods 2002**) που αντιστοιχούν στις κονδυλιές. Έστω $K_m < N_m$ ο αριθμός των πιο σημαντικών σημείων της m κονδυλιάς. Αντί να χρησιμοποιήσουμε τους κωδικούς Freeman των αρχικών σημείων της κονδυλιάς, κρατάμε μόνο τους κωδικούς Freeman των πιο σημαντικών σημείων της κονδυλιάς (Εικόνα 14).

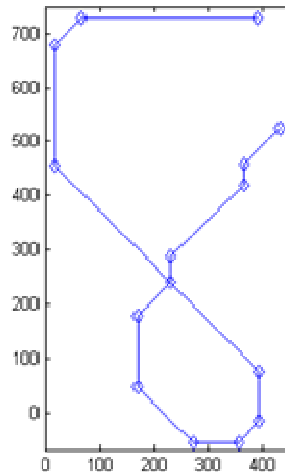


Εικόνα 14. (α) Τα αρχικά σημεία και τα (β) πιο σημαντικά σημεία ενός συμβόλου.

Στο επόμενο στάδιο υπολογίζουμε το ποσοστό του μήκους του τμήματος μεταξύ του k και του $k+1$ πιο σημαντικού σημείου ως προς το συνολικό μήκος της κονδυλιάς m :

$$\lambda_k^m = \frac{\Delta l_k^m}{\sum_{i=1}^{K_m-1} \Delta l_i^m} \quad (4)$$

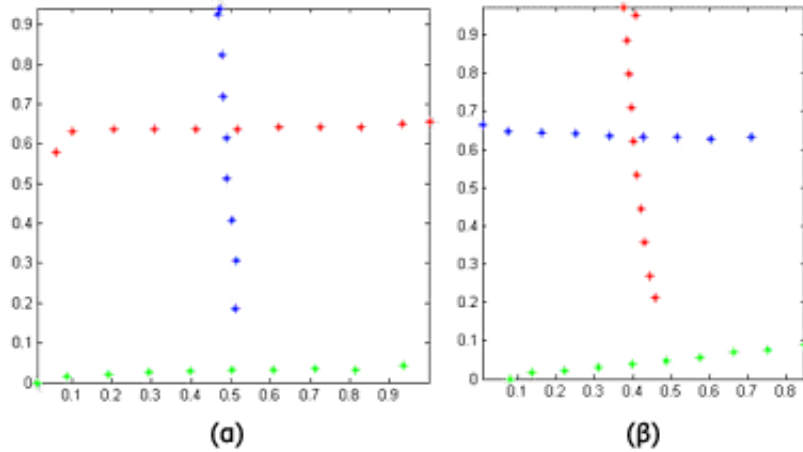
όπου Δl_k είναι η ευκλείδεια απόσταση μεταξύ των k και $k+1$ πιο σημαντικών σημείων. Σαν αποτέλεσμα, η κονδυλιά πλέον μπορεί να περιγραφεί ως μια ακολουθία δύο τιμών: (i) του κωδικού Freeman και (ii) του ποσοστού μήκους μεταξύ δύο συνεχόμενων σημαντικών σημείων της κονδυλιάς. Έτσι, $S^m = \{(c_k^m, \lambda_k^m), n = 1 \dots K_m - 1\}$, όπου $c_k^m \in \{0, 1, \dots, 7\}$ και $\lambda_k^m \in [0, 1]$. Η αναπαράσταση αυτή της κονδυλιάς μπορεί να ερμηνευθεί ως «από κάθε σημαντικό σημείο προς ποια κατεύθυνση και για πόσο ποσοστό μήκους σε σχέση με το συνολικό μήκος της κονδυλιάς πρέπει να κινηθώ ώστε να αναπαράγω την κονδυλιά» (Εικόνα 15).



Εικόνα 15. Κβαντισμένη αναπαράσταση του συμβόλου σύμφωνα με την κωδικοποίηση κατά Freeman 8-επιπέδων.

Η κατηγοριοποίηση του συμβόλου βασίζεται στην τεχνική του κοντινότερου γείτονα (nearest neighbor) της απόστασης μεταξύ του προς εξέταση συμβόλου και του πρότυπου αναφοράς. Έστω $S_T = \{S_T^m, m = 1, \dots, M_T\}$ το υπό εξέταση σύμβολο, όπου M_T ο αριθμός των κονδυλιών και $S_T^m = \{(c_k^m, \lambda_k^m), k = 1 \dots K_T^m - 1\}$ η αναπαράσταση των κονδυλιών σύμφωνα με το ζευγάρι τιμών c_k^m, λ_k^m , όπου c_k^m είναι η κατεύθυνση μεταξύ δύο συνεχόμενων σημαντικών σημείων της κονδυλιάς m και λ_k^m το αντίστοιχο ποσοστό μήκους μεταξύ των δύο συνεχόμενων σημαντικών σημείων της κονδυλιάς.

Όμοια, έστω $S_R = \{S_R^l, l = 1, \dots, M_R\}$ ένα από τα πρότυπα αναφοράς, όπου M_R ο αριθμός των κονδυλιών που απαρτίζει το πρότυπο αναφοράς και $S_R^l = \{(c_k^l, \lambda_k^l), k = 1 \dots K_R^l - 1\}$ η αναπαράσταση των κονδυλιών σύμφωνα με το ζευγάρι τιμών c_k^l, λ_k^l . Από το σύνολο των προτύπων αναφοράς, διαλέγουμε εκείνα που έχουν τον ίδιο αριθμό κονδυλιών με το υπό εξέταση σύμβολο, $M_T = M_R = M$. Παρόλο που ο αριθμός κονδυλιών του προτύπου αναφοράς και του υπό εξέταση συμβόλου είναι ίδιος, η αντιστοίχιση των κονδυλιών δεν συμπίπτει πάντα. Εξαρτάται από την σειρά γραφής των κονδυλιών του συμβόλου. Ας υποθέσουμε μια αντιστοίχιση μία-προς-μία των κονδυλιών που απαρτίζουν το υπό εξέταση σύμβολο και των κονδυλιών που απαρτίζουν το πρότυπο αναφοράς: $\{(m, l), m, l = 1, \dots, M\}$, όπου η m κονδυλιά του υπό εξέταση συμβόλου αντιστοιχίζεται με την l κονδυλιά του προτύπου αναφοράς. Για παράδειγμα στην Εικόνα 16α, η οριζόντια κονδυλιά του συμβόλου της πρόσθεσης (+), έχει γραφεί πριν την διασταυρωμένη κατακόρυφη κονδυλιά, ενώ στην Εικόνα 16β η αντίστοιχη οριζόντια κονδυλιά έχει γραφτεί μετά την διασταυρωμένη κατακόρυφη.



Εικόνα 16. Αντιστοίχιση μεταξύ των κονδυλιών που απαρτίζουν το υπό εξέταση σύμβολο και των κονδυλιών που απαρτίζουν το πρότυπο αναφοράς.

Η σωστή αντιστοίχιση των κονδυλιών επιτυγχάνεται ελαχιστοποιώντας το άθροισμα των αποστάσεων μεταξύ όλων των συνδυασμών που προκύπτουν από τα ζεύγη κονδυλιών του υπό εξέταση συμβόλου και του προτύπου αναφοράς. Κατά συνέπεια, η απόσταση μεταξύ του υπό εξέταση συμβόλου και του προτύπου αναφοράς προκύπτει από τη ακόλουθη σχέση:

$$D(S_T, S_R) = \min_{(m,l)} \left(\sum_{m=1}^M d(S_T^m, S_R^l) \right) \quad (5)$$

όπου $d(S_T^m, S_R^l)$ είναι η απόσταση μεταξύ των δύο κονδυλιών.

Συνήθως οι δύο υπό σύγκριση κονδυλιές έχουν διαφορετικό πλήθος στοιχείων, $K_T^l \neq K_R^m$. Για το λόγο αυτό εφαρμόζουμε μια τεχνική κανονικοποίησης ως προς την ταχύτητα γραφής, ενοποιώντας τις δύο ακολουθίες στοιχείων που περιγράφουν τις υπό σύγκριση κονδυλιές S_T^m , S_R^l .

Χρησιμοποιώντας τα ποσοστά μήκους μεταξύ δύο συνεχόμενων σημαντικών σημείων των κονδυλιών που απαρτίζουν το υπό εξέταση σύμβολο λ_k^m και των κονδυλιών που απαρτίζουν το πρότυπο αναφοράς λ_k^l αντίστοιχα, ορίζουμε δύο νέες ακολουθίες Λ_k^T , Λ_k^R που σαν στοιχεία τους έχουν το αθροιστικό μήκος από το αρχικό σημείο της κονδυλιάς μέχρι το k-οστό σημαντικό σημείο (Πίνακας 1):

$$\left\{ \Lambda_k^T = \sum_{i=1}^{k-1} \lambda_i^m, k = 1, \dots, K_T^m \right\} \quad (6)$$

και

$$\left\{ \Lambda_k^R = \sum_{i=1}^{k-1} \lambda_i^l, k=1, \dots, K_R^l \right\} \quad (7)$$

Σημειώστε ότι και οι δύο αυτές ακολουθίες ξεκινούν από μηδενικό μήκος και τελειώνουν με μοναδιαίο μήκος $\Lambda_{K_T^m}^T = \Lambda_{K_R^l}^R = 1$.

Πρότυπο αναφοράς			Υπό εξέταση σύμβολο		
c_k^R	λ_k^R (%)	Λ_k^R (%)	c_k^T	λ_k^T (%)	Λ_k^T (%)
-	-	0	-	-	0
3	0,63	0,63	4	15,35	15,35
4	13,59	14,22	5	3,32	18,67
5	2,18	16,40	6	10,60	29,27
6	13,50	29,90	7	25,28	54,55
7	14,37	44,27	6	4,31	58,86
0	1,93	46,20	5	2,68	61,54
7	6,77	52,97	4	3,89	65,43
5	3,78	56,75	3	6,84	72,27
4	9,23	65,98	2	6,21	78,48
3	4,54	70,52	1	4,13	82,61
2	8,38	78,90	2	2,22	84,83
1	14,45	93,35	1	8,87	93,70
2	2,97	96,32	2	1,79	95,49
1	1,44	97,76	1	4,51	100,00
1	2,24	100,00			

Πίνακας 1. Ακολουθίες χαρακτηριστικών και αθροιστικά ποσοστιαία τοπικά μήκη για το πρότυπο αναφοράς και το υπό εξέταση σύμβολο.

Στη συνέχεια ομαδοποιούμε τις δύο αυτές ακολουθίες Λ_k^T, Λ_k^R σε μία καινούργια ακολουθία $\{\Lambda_{r(k)}^{T,R}, r=1, \dots, K_T^m + K_R^l - 2\}$, κατά αύξουσα σειρά των στοιχείων που τις απαρτίζουν $\{\Lambda_k^T\}$, $\{\Lambda_k^R\}$. Συνεπώς, η ακολουθία που προκύπτει έχει κατά μέγιστο $K_T^m + K_R^l - 2$ διακριτά στοιχεία και με τα στοιχεία αρχής και τέλους να συμπίπτουν. Με αυτό τον τρόπο δημιουργούνται επιπλέον σημεία στις δύο κονδυλίες m, l στις ίδιες αποστάσεις από τα σημεία εκκίνησης των δύο κονδυλίων $\Lambda_r^{T,R}$.

Οι ακολουθίες των κονδυλίων του υπό εξέταση συμβόλου και του προτύπου αναφοράς περιγράφονται πλέον από τις ακόλουθες σχέσεις:

$$\hat{S}_T^m = \{(z_r^m, \lambda_r^{m,l}), r=1 \dots K_T^m + K_R^l - 3\} \quad (8)$$

$$\hat{S}_R^l = \{(z_r^l, \lambda_r^{m,l}), r=1 \dots K_T^m + K_R^l - 3\} \quad (9)$$

$$\lambda_r^{m,l} = \Lambda_{r+1}^{T,R} - \Lambda_r^{T,R} \quad (10)$$

$$z_r^m = \begin{cases} c_k^m, & \text{if } \Lambda_r^{T,R} \in \{\Lambda_k^T, k = 1, \dots, K_T^m - 1\} \\ c_{k+1}^m, & \text{if } \Lambda_r^{T,R} \in \{\Lambda_k^R, k = 1, \dots, K_R^l - 1\} \end{cases} \quad (11)$$

$$z_r^l = \begin{cases} c_k^l, & \text{if } \Lambda_r^{T,R} \in \{\Lambda_k^R, k = 1, \dots, K_R^l - 1\} \\ c_{k+1}^l, & \text{if } \Lambda_r^{T,R} \in \{\Lambda_k^T, k = 1, \dots, K_T^m - 1\} \end{cases} \quad (12)$$

Οι νέες αυτές ακολουθίες που μόλις ορίσαμε \hat{S}_T^m και \hat{S}_R^l , έχουν ίδιο αριθμό στοιχείων και το ίδιο πλέον ποσοστό μήκους μεταξύ των αρχικών σημείων των κονδυλίων και των σημαντικών τους σημείων $\Lambda_r^{T,R}$ (Πίνακας 2). Τα στοιχεία που προέρχονται από την ακολουθία του πρότυπου αναφοράς φαίνονται με έντονα γράμματα. Παίρνοντας την πρώτη διαφορά της ακολουθίας $\Lambda_r^{T,R}$ προκύπτει από την εξίσωση 10 η ενοποιημένη ακολουθία των ποσοστών μήκους μεταξύ δύο συνεχόμενων σημαντικών σημείων της κονδυλιάς $\lambda_r^{m,l}$.

Τέλος για τις δύο νέες ακολουθίες \hat{S}_T^m και \hat{S}_R^l , υπολογίζουμε την κατεύθυνση μεταξύ δύο συνεχόμενων σημαντικών σημείων της κονδυλιάς του υπό εξέταση συμβόλου και του προτύπου αναφοράς, διατηρώντας τις κατευθύνσεις των αρχικών σημείων τους c_k^T, c_k^R (Πίνακας 1) και βάζοντας στα σημεία που προστίθενται την αντίστοιχη κατεύθυνση ανάλογα με το ποσοστό μήκους $\Lambda_r^{T,R}$ στο οποίο βρίσκεται το προστιθέμενο σημείο (εξισώσεις 11 και 12).

Οι καινούργιες ακολουθίες \hat{S}_T^m και \hat{S}_R^l περιγράφονται στον Πίνακα 2.

$\Lambda_{r(k)}^{T,R}$ (%)	$\lambda_r^{m,l}$ (%)	z_r^l	z_r^m
0,00			
0,63	0,63	3	4
14,22	13,59	4	4
15,35	1,13	5	4
16,40	1,05	5	5
18,67	2,27	6	5
29,27	10,60	6	6
29,90	0,63	6	7
44,27	14,37	7	7
46,20	1,93	0	7
52,97	6,77	7	7
54,55	1,58	5	7

όπου $0 \leq d(s_r^m, s_r^l) \leq 4$, αφού η διαφορά μεταξύ δύο κωδικών Freeman z_r^m, z_r^l δίνεται από τον μαθηματικό τύπο $4 - |4 - |z_r^m - z_r^l||$.

Εφαρμόζοντας την Εξίσωση 13 στα δύο σύμβολα της Εικόνας 17 προκύπτει ότι η απόσταση τους είναι ίση με 0.25.

3.5. Αξιολόγηση

Για την αξιολόγηση της προτεινόμενης μεθόδου χρησιμοποιήθηκαν οι παρακάτω βάσεις χειρόγραφων μαθηματικών συμβόλων: α) η βάση χειρόγραφων μαθηματικών συμβόλων ILSP (βλ. Κεφάλαιο 3.1) και β) η βάση LVZE (**LaViola and Zeleznik 2007**).

Η βάση δεδομένων LVZE αποτελείται από 48 διακριτά σύμβολα γραμμένα από 11 γραφείς. Κάθε γραφέας έγραψε κάθε σύμβολο από 10 φορές για την δημιουργία του συνόλου εκμάθησης (training set) και από 12 φορές για την δημιουργία του συνόλου εξέτασης (test set). Τα δεδομένα συλλέχθηκαν από Tablet PC.

Τα 48 σύμβολα της βάσης δεδομένων LVZE περιλαμβάνουν α) τα 25 γράμματα του λατινικού αλφαβήτου (a-z), β) τα ψηφία από 0 έως 9 και γ) μαθηματικά σύμβολα όπως το σύμβολο της άθροισης, το σύμβολο της τετραγωνικής ρίζας, το σύμβολο του ολοκληρώματος, τα σύμβολα ανισότητας, τους αριθμητικούς τελεστές και μερικά ειδικά σύμβολα όπως οι παρενθέσεις και οι αγκύλες.

Η βάση δεδομένων ILSP αποτελείται συνολικά από 62.100 σύμβολα. Από αυτά, τα 27.900 αποτελούν το σύνολο εκπαίδευσης και τα υπόλοιπα 18.600 το σύνολο εξέτασης. Τα σύμβολα συλλέχθηκαν με ένα Tablet-PC.

Για καθεμία από της προαναφερθείσες βάσεις (ILSP και LVZE) πραγματοποιήθηκαν δύο πειράματα. Στο πρώτο πείραμα εφαρμόσαμε την τεχνική της ταύτισης απόστασης με υπόδειγμα χρησιμοποιώντας όλα τα κβαντισμένα σε κωδικό αλυσίδας σημεία (στη συνέχεια θα αναφέρεται ως Full-Freeman-Chain-Code), ενώ στο δεύτερο πείραμα χρησιμοποιήθηκαν μόνο τα πιο σημαντικά σημεία (dominant points) και οι αντίστοιχοι κωδικοί αλυσίδας κατά Freeman (στη συνέχεια θα αναφέρεται ως Reduced-Freeman-Chain-Code). Τα αποτελέσματα των δύο αυτών πειραμάτων φαίνονται στον Πίνακα 3 για την βάση LVZE και στον Πίνακα 4 για την βάση ILSP. Συγκρίνοντας τις τιμές των αποτελεσμάτων, παρατηρούμε ότι το ποσοστό αναγνώρισης παρουσιάζει μία ελαφρά μείωση όταν χρησιμοποιούνται μόνο τα πιο σημαντικά σημεία (dominant points) της κονδυλιάς. Επιπλέον τα ποσοστά αναγνώρισης και για τις δύο βάσεις κυμαίνονται περίπου στις ίδιες τιμές (**LaViola and Zeleznik 2007**).

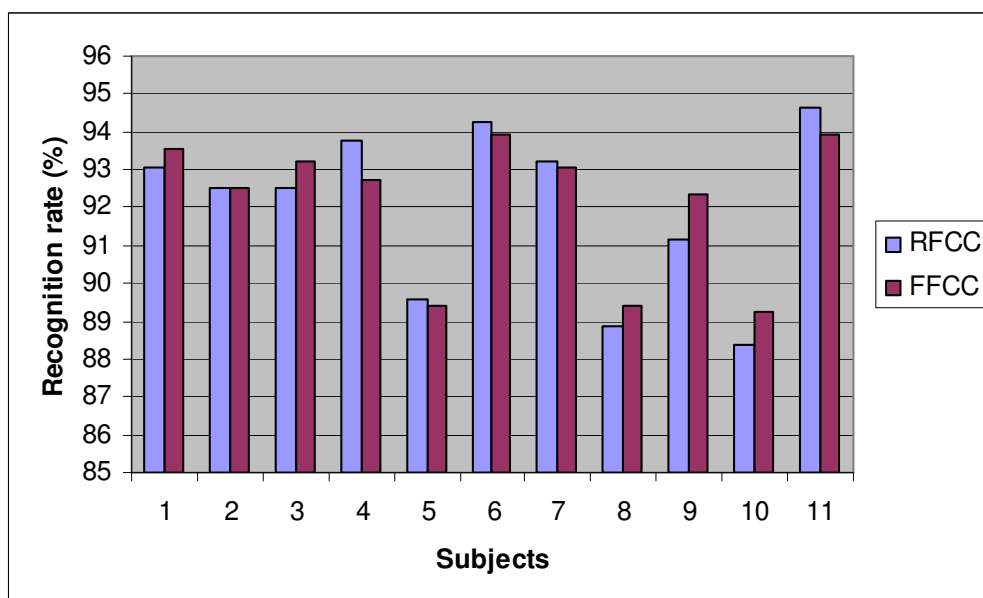
	Top 1	Top 2	Top 3	Top5	Top 10
FFCC	92.13	95.15	96.58	97.91	98.94
RFCC	92.00	95.49	96.85	97.82	98.81

Πίνακας 3. Ποσοστά αναγνώρισης για τις πρώτες N επιλογές με την βάση LVZE.

	Top 1	Top 2	Top 3	Top5	Top 10
FFCC	92.35	95.15	96.63	97.94	98.97
RFCC	92.21	95.44	96.95	97.89	98.84

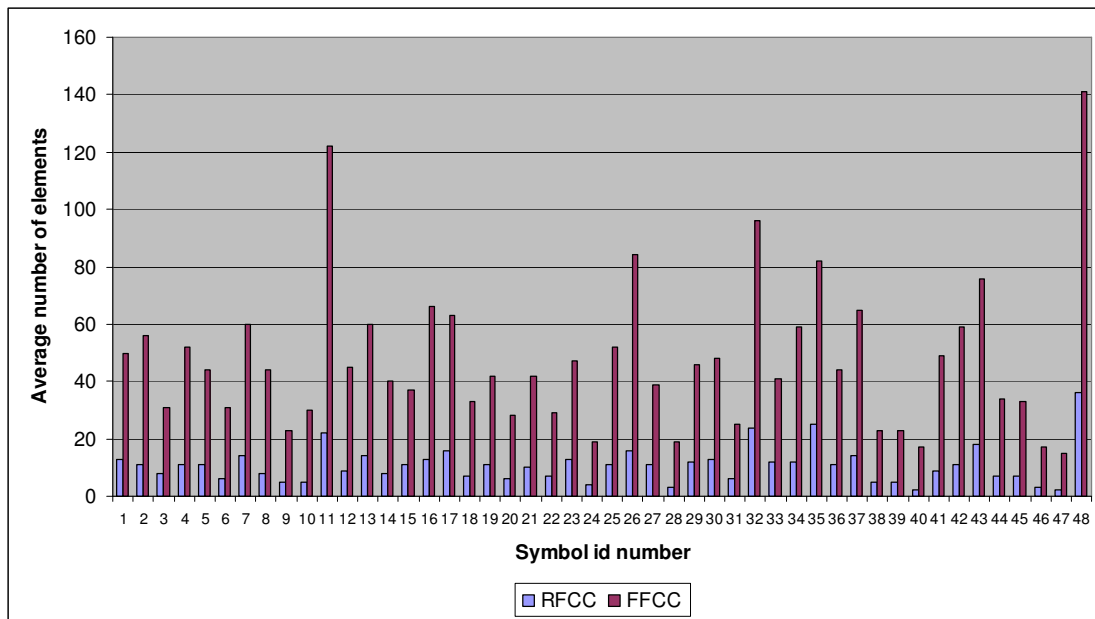
Πίνακας 4. Ποσοστά αναγνώρισης για τις πρώτες N επιλογές με την βάση ILSP.

Η τυπική απόκλιση των ποσοστών αναγνώρισης σε σχέση με τους γραφείς είναι 1.85 για την Top-1 επιλογή του FFCC και 2.18 για την Top-1 επιλογή του RFCC για την βάση LVZE. Στην Εικόνα 18 φαίνονται τα ποσοστά αναγνώρισης ανά γραφέα για την βάση LVZE, χρησιμοποιώντας την FFCC και RFCC κωδικοποίηση.



Εικόνα 18. Ποσοστά αναγνώρισης ανά γραφέα της βάσης LVZE για τις δύο τεχνικές κωδικοποίησης (FFCC και RFCC).

Αξίζει να σημειωθεί ότι, η RFCC κωδικοποίηση έχει πολύ λιγότερα σημεία σε σχέση με την FFCC κωδικοποίηση. Συγκεκριμένα ο μέσος όρος σημείων για ένα σύμβολο της βάσης LVZE με RFCC κωδικοποίηση είναι ίσος με 10.79, ενώ για την κωδικοποίηση FFCC είναι 47.53. Επιπλέον, η μείωση του ποσοστού αναγνώρισης δεν ξεπερνά το 0.15%. Στην Εικόνα 19 φαίνεται ο μέσος αριθμός σημείων με FFCC και RFCC κωδικοποίηση για κάθε σύμβολο της βάσης LVZE.



Εικόνα 19. Μέσος αριθμός σημείων με FFCC και RFCC κωδικοποίηση για τα σύμβολα της βάσης LVZE.

Ακόμη ένα πείραμα πραγματοποιήθηκε χρησιμοποιώντας τα σύμβολα από τις μαθηματικές εκφράσεις της βάσης δεδομένων ILSP. Το ποσοστό αναγνώρισης για την RFCC κωδικοποίηση είναι ίσο με 94.18%. Να σημειωθεί ότι ο αριθμός προτύπων ανά σύμβολο διαφέρει σημαντικά, π.χ. το σύμβολο της ισότητας εμφανίζεται σχεδόν σε όλες τις μαθηματικές εκφράσεις της βάσης, ενώ το σύμβολο του ολοκληρώματος εμφανίζεται μόνο σε δύο.

Τέλος, για την αξιολόγηση της μεθόδου του ελαστικού ταιριάσματος (κεφάλαιο 3.4.1) χρησιμοποιήθηκαν οι 2 βάσεις χειρόγραφων μαθηματικών συμβόλων ILSP και LVZE. Τα ποσοστά αναγνώρισης για την μέθοδο αυτή κυμαίνονται σε χαμηλότερα ποσοστά: (i) για την βάση ILSP, 81.02 (top1), 84.10 (top2), 85.23 (top3) 86.87 (top5) και (ii) για την βάση LVZE, 82.87 (top1), 83.76 (top2), 84.03 (top3) 85.26 (top5).

3.6. Συμπεράσματα

Στην παρούσα εργασία αναπτύχθηκαν δύο τεχνικές αναγνώρισης μαθηματικών συμβόλων: α) η τεχνική ελαστικού ταιριάσματος (elastic matching) και β) η τεχνική ταύτισης απόστασης με υπόδειγμα (string template matching). Οι δύο αυτές τεχνικές αξιολογήθηκαν στην βάση LVZE και στην βάση ILSP. Η τεχνική ταύτισης απόστασης με υπόδειγμα έδωσε ποσοστό ορθής αναγνώρισης για την Top-1 επιλογή ίσο με 92.13% για την βάση LVZE και 92.35% για την βάση ILSP. Τα ποσοστά αναγνώρισης για την μέθοδο του ελαστικού ταιριάσματος κυμαίνονται σε χαμηλότερα ποσοστά: (i) για την βάση ILSP, 81.02 (top-1) και (ii) για την βάση LVZE, 82.87 (top-1). Μπορούμε να συμπεράνουμε ότι τα λάθη στην αναγνώριση των συμβόλων στην πλειονότητά τους οφείλονται σε τρεις λόγους: ο αλγόριθμος αναγνώρισης συμβόλων δεν χειρίζεται σύνθετα σύμβολα (π.χ. log) και δεν λαμβάνει υπόψη του την σχετική θέση των

κονδυλιών του συμβόλου (π.χ. το σύμβολο της ανισότητας μπορεί να μπερδευτεί με το σύμβολο F που συνήθως γράφεται με δύο κονδυλιές). Επιπλέον, υπάρχουν κλάσεις συμβόλων που παρουσιάζουν μεγάλη ομοιότητα, όπως το μικρό γράμμα c με το κεφαλαίο γράμμα C και είναι δύσκολο να διαχωριστούν με τα συγκεκριμένα χαρακτηριστικά.

Κεφάλαιο 4. Αναγνώριση τοπολογικής σχέσης μεταξύ δύο συμβόλων μιας ΜΕ

Η αναγνώριση των σχέσεων μεταξύ των συμβόλων βασίζεται στην δομική ανάλυση της ΜΕ. Μία απλή λύση σε αυτό το πρόβλημα είναι η εισαγωγή περιορισμών που εξετάζουν την σχετική τοπολογική σχέση μεταξύ των συμβόλων. Για το λόγο αυτό εξετάζουμε τις τοπολογικές ιδιότητες του συμβόλου, όπως είναι το κεντροειδές και το πλαίσιο οριοθέτησης του συμβόλου, ώστε να συμπεράνουμε τις τοπολογικές σχέσεις μεταξύ των μαθηματικών συμβόλων. Αξιολογώντας τα γεωμετρικά χαρακτηριστικά που διέπουν τα σύμβολα μιας μαθηματικής έκφρασης αναπτύξαμε γεωμετρικούς κανόνες για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων. Η μέθοδος αυτή υλοποιεί έναν ευριστικό ταξινομητή με βάση γεωμετρικά χαρακτηριστικά.

Επιπλέον, προκειμένου να αναγνωρίσουμε την τοπολογική σχέση που διέπει δύο σύμβολα, προτείνουμε μία καινοτόμο μέθοδο βασισμένη σε έναν πιθανοτικό SVM ταξινομητή (**Platt, 2000, Lin et al. 2007**). Οι μηχανές διανυσματικής υποστήριξης έχουν κατά κύριο λόγο χρησιμοποιηθεί στην αντιμετώπιση του προβλήματος της αναγνώρισης συμβόλων (**Golubitsky and Watt 2010, Keshari and Watt 2008**) και προτάθηκαν για πρώτη φορά στην αναγνώριση της τοπολογικής σχέσης μεταξύ δύο συμβόλων από τη **Simistira et al. 2013**.

Για την αξιολόγηση των δύο μεθόδων αναγνώρισης τοπολογικής σχέσης μεταξύ δύο συμβόλων χρησιμοποιήθηκαν οι ακόλουθες βάσεις χειρόγραφων μαθηματικών εκφράσεων.

4.1. Βάσεις δεδομένων χειρόγραφων ΜΕ

Για να αξιολογήσουμε και να συγκρίνουμε τις τρεις μεθόδους ταξινόμησης σχέσεων μεταξύ συμβόλων, (α) τον ευριστικό ταξινομητή (ILSP-1), (β) τον SVM ταξινομητή και (γ) τον ANN ταξινομητή, δημιουργήσαμε μία βάση τοπολογικών σχέσεων. Τα ζευγάρια συμβόλων προήλθαν από την βάση μαθηματικών εκφράσεων CROHME 2012 η οποία αποτελείται από ένα σύνολο εκπαίδευσης 1.336 μαθηματικών εκφράσεων και από ένα σύνολο αξιολόγησης 488 μαθηματικών εκφράσεων.

Από τις 1.336 μαθηματικές εκφράσεις του συνόλου εκπαίδευσης επιλέξαμε τις 825 με σκοπό (α) την συλλογή επαρκών δειγμάτων για κάθε τοπολογική σχέση και (β) την συλλογή δειγμάτων που μπορεί να είναι ασαφή για τον ταξινομητή. Η βάση δεδομένων που δημιουργήθηκε περιέχει 1906 τοπολογικές σχέσεις μεταξύ συμβόλων. Στον πίνακα 5 παρουσιάζεται το πλήθος των δειγμάτων ανά σχέση. Για κάθε ζευγάρι σχέσεων αποθηκεύσαμε διάφορες πληροφορίες όπως (α) την τοπολογική σχέση, (β) τις συντεταγμένες του πλαισίου οριοθέτησης των συμβόλων, (γ) τις συντεταγμένες του κεντροειδούς για κάθε σύμβολο και

(δ) τις συντεταγμένες των κονδυλιών των συμβόλων. Χρησιμοποιώντας αυτές τις πληροφορίες δημιουργήσαμε ένα διάνυσμα χαρακτηριστικών διάστασης έξι. Ακολουθήθηκε η τεχνική 10-fold cross validation, χρησιμοποιώντας για κάθε επανάληψη (fold) το 90% της βάσης ως σύνολο εκπαίδευσης και το υπόλοιπο 10% ως σύνολο αξιολόγησης.

Τοπολογική σχέση	Πλήθος δειγμάτων
πάνω	176
πάνω-δεξιά	203
δεξιά	838
μέσα	234
κάτω-δεξιά	169
κάτω	286
Σύνολο	1906

Πίνακας 5. Κατανομή τοπολογικών σχέσεων που εξάχθηκαν από την βάση CROHME2012

Επιπλέον, για την εκπαίδευση και αξιολόγηση του πιθανοτικού SVM ταξινομητή χρησιμοποιήθηκε η βάση MathBrush. Η βάση χειρόγραφων μαθηματικών παραστάσεων MathBrush (**MacLean et al. 2010**) περιλαμβάνει 4.655 μαθηματικές εξισώσεις από 20 διαφορετικούς γραφείς, 25.936 σύμβολα και 21.238 τοπολογικές σχέσεις μεταξύ μαθηματικών συμβόλων ή/και υπό-εκφράσεων της ΜΕ. Ορίζονται πέντε τοπολογικές σχέσεις μεταξύ δύο συμβόλων: (i) πάνω-δεξιά, (ii) δεξιά, (iii) κάτω-δεξιά, (iv) κάτω, και (v) μέσα.

Συγκεκριμένα η βάση MathBrush περιλαμβάνει:

- (i) 1.919 τοπολογικές σχέσεις «πάνω-δεξιά»,
- (ii) 14,624 τοπολογικές σχέσεις «δεξιά»,
- (iii) 858 τοπολογικές σχέσεις «μέσα»,
- (iv) 1,259 τοπολογικές σχέσεις «κάτω-δεξιά» και
- (v) 2,578 τοπολογικές σχέσεις «κάτω».

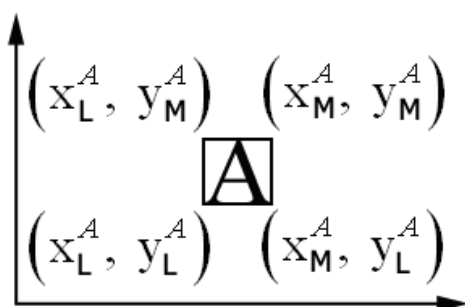
Στον ακόλουθο πίνακα παρουσιάζονται τα 97 σύμβολα που περιέχονται στις ΜΕ:

51 γράμματα του Λατινικού αλφαβήτου	A, a, B, b, C, c, D, d, E, e, F, f, G, g, H, h, I, i, J, j, K, k, L, l, M, m, N, n, O, o, P, p, Q, q, R, r, S, s, T, t, U, u, V, v, W, X, x, Y, y, Z, z
20 γράμματα του ελληνικού αλφαβήτου	Δ, Γ, Ω, Π, Σ, α, β, δ, ε, γ, λ, μ, φ, π, ψ, ρ, σ, τ, θ, ζ
6 σχεσιακοί τελεστές	=, ≠, ≤, <, ≥, >
4 παρενθετικά σύμβολα	(,), [,]
τελεστής άθροισης	+
δομικά σύμβολα	γραμμή κλάσματος, √
τελεστής ολοκληρώματος	∫
ειδικό σύμβολο	∞
σημείο στίξης	.

4.2. Δομικά στοιχεία

Πριν προχωρήσουμε στην τοπολογική ανάλυση των μαθηματικών εκφράσεων, θα κάνουμε μία αναφορά στον ορισμό του πλαισίου οριοθέτησης και στον ορισμό του κεντροειδούς ενός συμβόλου:

Πλαίσιο Οριοθέτησης (bounding box). Το πλαίσιο οριοθέτησης ενός συμβόλου (Εικόνα 20) είναι η ορθογώνια περιοχή που περικλείει το σύμβολο και ορίζεται από τις ακόλουθες συντεταγμένες: $x_L^A = \min(x_i^A)$, $x_M^A = \max(x_i^A)$, $y_L^A = \min(y_i^A)$, $y_M^A = \max(y_i^A)$.



Εικόνα 20. Bounding box ενός συμβόλου.

Κεντροειδές. Μια πολύ διαδεδομένη τεχνική για την εξέταση του αν ένα σύμβολο βρίσκεται σε μία περιοχή είναι εξετάζοντας τις συντεταγμένες του κεντροειδούς του. Στην περίπτωση επεξεργασίας συμβόλων εκτός γραμμής (off-line) το κεντροειδές ορίζεται ως η μέση τιμή των συντεταγμένων των pixels του συμβόλου. Στην περίπτωση επεξεργασίας συμβόλων εντός γραμμής (on-line) η διαδικασία εύρεσης του κεντροειδούς είναι διαφορετική. Ακολουθώντας την προσέγγιση που προτάθηκε στην εργασία των **Zanibbi et al.** το **2002**, μπορούμε να ορίσουμε τις συντεταγμένες του κεντροειδούς ως εξής:

Η x-συντεταγμένη βρίσκεται από τον μαθηματικό τύπο $x_c^A = \frac{w^A}{2}$, όπου $w^A = x_M^A - x_L^A$,

και η y-συντεταγμένη από τον μαθηματικό τύπο $y_c^A = ah^A$, όπου $h^A = y_M^A - y_L^A$.

Η τιμή της παραμέτρου a εξαρτάται από την ταξινόμηση του συμβόλου σε μία από τις ακόλουθες κατηγορίες:

- 1) ascenders ("b", "d", κ.λπ.), δηλ. σύμβολα που εκτείνονται πάνω από την βασική γραμμή
- 2) descenders ("p", "q", κ.λπ.), δηλ. σύμβολα που εκτείνονται κάτω από την βασική γραμμή
- 3) και centered ("+", "=", κ.λπ.) δηλ. σύμβολα που γράφονται επί της βασικής γραμμής

Για καθεμία από τις παραπάνω κατηγορίες η παράμετρος α παίρνει την τιμή 0.2, 0.8 και 0.5, αντίστοιχα.

4.3. Προτεινόμενες τεχνικές

Το στάδιο αυτό προϋποθέτει ότι όλα τα σύμβολα έχουν αναγνωρισθεί. Στόχος τώρα είναι να βρεθεί η βέλτιστη αναπαράσταση της δομικής πληροφορίας που διέπει την ΜΕ. Για το λόγο αυτό, πρέπει να οριστούν οι σχέσεις μεταξύ των συμβόλων της ΜΕ.

Στην παρούσα εργασία αναπτύχθηκαν τρεις ταξινομητές για την εύρεση της σχέσης μεταξύ δύο συμβόλων:

- (α) ένας ταξινομητής με χρήση ευριστικών κανόνων (heuristics),
- (β) ένας ταξινομητής βασισμένος στην τεχνική των μηχανών διανυσμάτων υποστήριξης και
- (γ) ένας ταξινομητής βασισμένος στην τεχνική των νευρωνικών δικτύων

4.3.1. Ταξινομητής με χρήση ευριστικών κανόνων

Δεδομένου ενός συμβόλου, ορίζουμε τις ακόλουθες τοπολογικές σχέσεις: (i) πάνω-δεξιά, (ii) δεξιά, (iii) κάτω-δεξιά, (iv) κάτω, (v) πάνω και (vi) μέσα. Είναι φανερό ότι οι τοπολογικές σχέσεις εξαρτώνται άμεσα από το προς εξέταση σύμβολο. Για παράδειγμα το σύμβολο της ρίζας $\sqrt{\quad}$ έχει σχέση 'μέσα' με κάποιο άλλο σύμβολο της ΜΕ, ενώ για το σύμβολο π μία τέτοια σχέση δεν μπορεί να υπάρξει.

Για να βρούμε ποιά τοπολογική σχέση έχει ένα σύμβολο A με ένα άλλο σύμβολο B, ορίσαμε τις παρακάτω λογικές συναρτήσεις (Εικόνα 21) με τη χρήση ευριστικών κανόνων:



Εικόνα 21. Περιοχές συμβόλου και αντίστοιχες τοπολογικές σχέσεις

πάνω-δεξιά: Ένα σύμβολο A βρίσκεται στην πάνω-δεξιά περιοχή ενός συμβόλου B (π.χ. το σύμβολο A είναι εκθέτης του B), εάν:

$$\tan \frac{\pi}{10} \leq \tan(A, B) < \tan\left(\frac{3\pi}{8}\right) \quad (1)$$

$$y_c^A > y_c^B \quad (2)$$

Όπου (A, B) ορίζει την γωνία που σχηματίζεται μεταξύ της ευθείας γραμμής που ενώνει τα κεντροειδή των συμβόλων A και B και της οριζόντιας γραμμής.

δεξιά: Ένα σύμβολο A βρίσκεται στην δεξιά περιοχή του συμβόλου B, με άλλα λόγια το σύμβολο A είναι στο ίδιο επίπεδο το σύμβολο B, εάν:

$$\left| y_c^A - y_c^B \right| < a_1 \cdot \max(h^A, h^B) \quad (3)$$

$$x_c^A > x_c^B \quad (4)$$

Όλα τα σύμβολα που ανήκουν στο ίδιο επίπεδο, ικανοποιούν αυτή την σχέση.

κάτω-δεξιά: Ένα σύμβολο A βρίσκεται στην κάτω-δεξιά περιοχή ενός συμβόλου B (π.χ. το σύμβολο A είναι δείκτης του B), εάν:

$$\tan \frac{14\pi}{8} \leq \tan(A, B) < \tan\left(\frac{16\pi}{8}\right) \quad (5)$$

$$y_c^A < y_c^B \quad (6)$$

κάτω: Ένα σύμβολο A βρίσκεται στην κάτω περιοχή ενός συμβόλου B (π.χ. το σύμβολο A είναι παρανομαστής ενός κλάσματος), εάν:

$$\left| x_L^A - x_L^B \right| \leq a_2 \cdot \min(w^A, w^B) \quad (7)$$

$$\left| x_M^A - x_M^B \right| \leq a_2 \cdot \min(w^A, w^B) \quad (8)$$

$$y_M^A \leq y_L^B \quad (9)$$

πάνω: Ένα σύμβολο A βρίσκεται στην πάνω περιοχή ενός συμβόλου B (π.χ. το σύμβολο A είναι αριθμητής ενός κλάσματος), εάν:

$$\left| x_L^A - x_L^B \right| \leq a_3 \cdot \min(w^A, w^B) \quad (10)$$

$$\left| x_M^A - x_M^B \right| \leq a_3 \cdot \min(w^A, w^B) \quad (11)$$

$$y_L^A \geq y_M^B \quad (12)$$

μέσα: Ένα σύμβολο A βρίσκεται στην μέσα περιοχή ενός συμβόλου B (π.χ. το σύμβολο A είναι υπόριζο της ρίζας B), εάν:

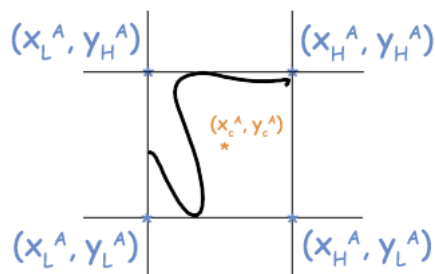
$$x_L^B < x_c^A < x_M^B \quad (13)$$

$$y_L^B < y_c^A < y_M^B \quad (14)$$

4.3.2. Ταξινομητής βασισμένος σε μηχανές διανυσμάτων υποστήριξης

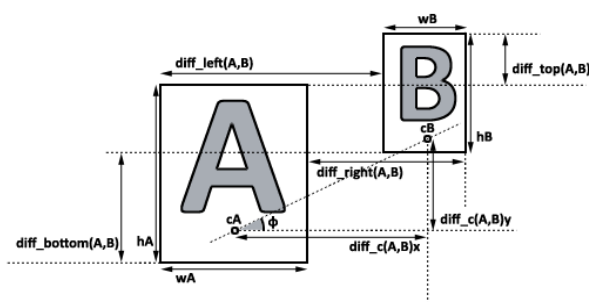
Στην παρούσα εργασία, αναπτύξαμε έναν SVM ταξινομητή παρόμοιο με αυτόν που περιγράφεται στην εργασία **Simistira et al. 2013**. Στην εργασία **Simistira et al. 2013** χρησιμοποιείται είτε το μέγιστο ύψος, είτε το μέγιστο πλάτος των προς σύγκριση συμβόλων. Στην παρούσα εργασία χρησιμοποιούμε ένα διαφορετικό τρόπο κανονικοποίησης των χαρακτηριστικών, καθώς και ένα επιπλέον χαρακτηριστικό, το οποίο ορίζεται από την γωνία που σχηματίζεται από την ευθεία που ενώνει τα κεντροειδή των δύο υπό εξέταση συμβόλων με την οριζόντια ευθεία. Το σύνολο των χαρακτηριστικών περιγράφεται αναλυτικά στην επόμενη παράγραφο. Επιπλέον χρησιμοποιούμε 5 αντί για 6 τοπολογικές σχέσεις, καθώς στην παρούσα εργασία εκπαιδεύσαμε τον SVM ταξινομητή στην βάση MathBrush, όπου δεν υπάρχει η τοπολογική σχέση «πάνω» καθώς θεωρείται συμμετρική με τη σχέση «κάτω».

Ας υποθέσουμε το πλαίσιο οριοθέτησης ενός συμβόλου A, όπως αυτό ορίστηκε στην Ενότητα 4.2. Αρχικά ορίζουμε το πλάτος του πλαισίου οριοθέτησης ενός συμβόλου A ως $w^A = x_H^A - x_L^A$ και το ύψος ως $h^A = y_H^A - y_L^A$ (βλ. Εικόνα 22). Η x-συντεταγμένη του κεντροειδούς του συμβόλου A ορίζεται ως $x_c^A = \frac{w^A}{2}$ και η y-συντεταγμένη του κεντροειδούς του συμβόλου A ορίζεται ως $y_c^A = ah^A$. Η τιμή της παραμέτρου a καθορίζεται όπως στο Κεφάλαιο 4.2 και παίρνει μία από τις τιμές 0.2, 0.8 και 0.5.



Εικόνα 22. Πλαίσιο οριοθέτησης ενός συμβόλου.

Για την αναπαράσταση της τοπολογικής σχέσης μεταξύ δύο αντικειμένων της ΜΕ, ορίσαμε 7 χαρακτηριστικά (βλ. Εικόνα 23). Ως αντικείμενο ορίζεται ένα μαθηματικό σύμβολο ή μία υπο-έκφραση της ΜΕ.



Εικόνα 23. Γεωμετρικά χαρακτηριστικά βασισμένα στο πλαίσιο οριοθέτησης και στα κεντροειδή των αντικειμένων.

Ας υποθέσουμε ότι τα A και B είναι δύο αντικείμενα της ΜΕ. Εάν ορίσουμε ως αντικείμενο αναφοράς το αντικείμενο A, τότε ορίζουμε τα ακόλουθα γεωμετρικά χαρακτηριστικά βασισμένα στα πλαίσια οριοθέτησης και στα κεντροειδή των αντικειμένων:

- I. $diff_left(A,B)$ την απόσταση μεταξύ των δύο αριστερών άκρων των πλαισίων οριοθέτησης των αντικειμένων
- II. $diff_right(A,B)$ την απόσταση μεταξύ των δύο δεξιών άκρων των πλαισίων οριοθέτησης των αντικειμένων
- III. $diff_top(A,B)$ την απόσταση μεταξύ των δύο άνω άκρων των πλαισίων οριοθέτησης των αντικειμένων
- IV. $diff_bottom(A,B)$ την απόσταση μεταξύ των δύο κάτω άκρων των πλαισίων οριοθέτησης των αντικειμένων
- V. $diff_c(A,B)x$ την απόσταση μεταξύ των δύο x-συντεταγμένων των αντικειμένων
- VI. $diff_c(A,B)y$ την απόσταση μεταξύ των δύο y-συντεταγμένων των αντικειμένων
- VII. $\varphi \angle(cA, cB)$ την γωνία μεταξύ των δύο κεντροειδών cA και cB .

Επιπλέον κανονικοποιούμε τα γεωμετρικά αυτά χαρακτηριστικά, διαιρώντας τα I, II και V με το πλάτος του πλαισίου οριοθέτησης που περιλαμβάνει τα δύο αντικείμενα A και B και αντίστοιχα τα III, IV και VI με το ύψος του πλαισίου οριοθέτησης που περιλαμβάνει τα δύο αντικείμενα A και B.

Σε σύγκριση με τους **Álvaro και Zanibbi (2013)**, η προτεινόμενη μέθοδος χρησιμοποιεί 7 αντί για 9 χαρακτηριστικά, από τα οποία τα 6 είναι ταυτόσημα, ενώ το έβδομο που ορίζεται από την γωνία των κεντροειδών είναι καινοτόμο της προτεινόμενης μεθόδου. Άλλη μία διαφορά έγκειται στην διαφορετικές τυπογραφικές κατηγορίες (ascenders, descenders, normal και middle) που προτείνονται από τους **Álvaro και Zanibbi (2013)** καθώς και στο τρόπο υπολογισμού της γ -συντεταγμένης του κεντροειδούς.

Στην συγκεκριμένη εργασία εκπαιδεύσαμε έναν πιθανοτικό SVM ταξινομητή με Gaussian kernel (**Platt 2000, Lin et al. 2007**) χρησιμοποιώντας την βάση MathBrush (**MacLean 2011**). Η βάση MathBrush αποτελείται από 4,654 ME και επισημειωμένες τοπολογικές σχέσεις. Ο ταξινομητής υλοποιήθηκε με την βιβλιοθήκη LibSVM (**Chang και Lin 2011**), αξιολογήθηκε σύμφωνα με την τεχνική 10-cross validation. Τα αντίστοιχα αποτελέσματα παρουσιάζονται στην Ενότητα 4.4.

4.3.3. Ταξινομητής βασισμένος στην τεχνική των νευρωνικών δικτύων

Στην συγκεκριμένη εργασία χρησιμοποιήσαμε τα ίδια χαρακτηριστικά που αναφέρθηκαν στην τεχνική που είναι βασισμένη στις μηχανές διανυσμάτων υποστήριξης (βλ. Ενότητα 4.3.2.). Υλοποιήσαμε 4 διαφορετικούς ταξινομητές χρησιμοποιώντας διαφορετικά σύνολα χαρακτηριστικών και διαφορετικό αριθμό τοπολογικών σχέσεων. Ο αριθμός των τοπολογικών σχέσεων που χρησιμοποιήσαμε ορίστηκε από την αντίστοιχη βάση δεδομένων που χρησιμοποιήθηκε για την αξιολόγηση των ταξινομητών. Για την βάση τοπολογικών σχέσεων CROHME2012 υλοποιήθηκαν οι ταξινομητές ANN-6f-6sp όπου χρησιμοποιούνται 6 γεωμετρικά χαρακτηριστικά και ο ταξινομητής ANN-7f-6sp όπου χρησιμοποιούνται και τα 7 γεωμετρικά χαρακτηριστικά. Αντίστοιχα για την βάση τοπολογικών σχέσεων Mathbrush υλοποιήθηκαν οι ταξινομητές ANN-6f-5sp και ANN-7f-5sp. Οι ANN ταξινομητές που χρησιμοποιήθηκαν αποτελούνται από 7 και 6 μονάδες εισόδου αντίστοιχα. Κάθε μονάδα εισόδου αντιστοιχεί σε ένα από τα 7 γεωμετρικά χαρακτηριστικά που αναφέρθηκαν παραπάνω. Το κρυφό επίπεδο αποτελείται από 10 μονάδες. Τέλος οι μονάδες του επιπέδου της εξόδου αντιστοιχούν σε μία από τις 6 ορισμένες τοπολογικές σχέσεις (πάνω-δεξιά, δεξιά, κάτω-δεξιά, κάτω, πάνω και μέσα). Οι ταξινομητές αυτοί εκπαιδεύτηκαν χρησιμοποιώντας τον αλγόριθμο Levenberg-Marquardt (**Levenberg 1944**) γνωστός και ως damped least-squares (DLS). Η υλοποίηση των ANN ταξινομητών έγινε χρησιμοποιώντας το Neural Network Toolbox του MATLAB (www.mathworks.com/products/matlab).

4.4. Αξιολόγηση

Για τον SVM ταξινομητή χρησιμοποιήσαμε δύο τεχνικές, την τεχνική "one-against-all" και την τεχνική "one-against-one" προκειμένου να διαλέξουμε την τεχνική με την καλύτερη απόδοση στην ταξινόμηση των τοπολογικών σχέσεων μεταξύ συμβόλων. Η πρώτη προσέγγιση χρησιμοποιεί έξι SVM ταξινομητές που εκπαιδεύτηκαν σύμφωνα με την τεχνική "one-against-

all". Στο Πίνακα 6 (στήλη 3) παρουσιάζονται τα αποτελέσματα για κάθε τοπολογική σχέση. Η δευτέρα προσέγγιση χρησιμοποιεί 15 SVM ταξινομητές που εκπαιδεύτηκαν σύμφωνα με την τεχνική "one-against-one". Τα αντίστοιχα αποτελέσματα αξιολόγησης παρουσιάζονται στον Πίνακα 6 (στήλη 4). Στο Πίνακα 7 παρουσιάζονται τα μέσα ποσοστά λάθους (mean error rate) για τους τρεις ταξινομητές.

Τοπολογική σχέση	ILSP-1	SVM (one-against-all)	SVM (one-against-one)
πάνω	1.4	0.5	0
πάνω-δεξιά	4.3	1.9	1.2
δεξιά	6.8	4.8	1.4
μέσα	4.9	2.8	1.2
κάτω-δεξιά	3.4	1.9	0.7
κάτω	2.6	0.9	0.6
Συνολικά	23.4	12.8	5.1

Πίνακας 6. Συγκριτικά αποτελέσματα του μέσου πλήθους λαθών (average number of errors) για τους 3 ταξινομητές (i) ILSP-1, (ii) one-against-all SVM and (iii) one-against-one SVM στην βάση CROHME2012.

Τοπολογική σχέση	ILSP-1	SVM (one-against-all)	SVM (one-against-one)
πάνω	0.73	0.25	0
πάνω-δεξιά	2.28	0.98	0.61
δεξιά	3.62	2.47	0.70
μέσα	2.60	1.43	0.63
κάτω-δεξιά	1.77	0.97	0.36
κάτω	1.30	0.46	0.31
Συνολικά	12.31	6.57	2.61

Πίνακας 7. Συγκριτικά αποτελέσματα του mean error rate για τους 3 ταξινομητές (i) ILSP-1, (ii) one-against-all SVM and (iii) one-against-one SVM στην βάση CROHME2012

Όπως φαίνεται στους Πίνακες 6 και 7 ο ταξινομητής SVM που χρησιμοποιεί την τεχνική «one-against-one» παρουσιάζει καλύτερα αποτελέσματα από τον ταξινομητή SVM που χρησιμοποιεί την τεχνική «one-against-all» καθώς και από τον ILSP-1 ταξινομητή. Συγκεκριμένα το μέσο πλήθος τοπολογικών σχέσεων που αναγνωρίζονται λάθος μειώνεται από 23.4% σε 12.8% όταν ακολουθείται η τεχνική «one-against-all» και μειώνεται ακόμα περισσότερο (5.1%) όταν ακολουθείται η τεχνική «one-against-one». Επιπλέον, όπως φαίνεται από τον Πίνακα 7, το μέσο ποσοστό λάθους των τριών ταξινομητών είναι 12.31% όταν χρησιμοποιείται ο ταξινομητής ILSP-1, 6.57% όταν χρησιμοποιείται ο ταξινομητής SVM "one-against-all" classifier και 2.61% όταν χρησιμοποιείται ο ταξινομητής SVM "one-against-one".

Βασίζόμενοι στο συμπέρασμα ότι ο ταξινομητής SVM "one-against-one" παρουσιάζει καλύτερα αποτελέσματα, δημιουργήσαμε ένα πιθανοτικό ταξινομητή SVM "one-against-one" πολλών

κλάσεων με Γκαουσιανό πυρήνα (Gaussian kernel). Ο ταξινομητής αυτός ονομάστηκε SVM-7f και χρησιμοποιεί τα 7 χαρακτηριστικά που αναφέρθηκαν στην Ενότητα 4.3.2. Για την εκπαίδευση και αξιολόγηση του πιθανοτικού SVM ταξινομητή χρησιμοποιήθηκε η βάση MathBrush. Σε όλα τα πειράματα αξιολόγησης ακολουθήθηκε η τεχνική 10-fold cross validation, χρησιμοποιώντας για κάθε επανάληψη (fold) το 90% της βάσης ως σύνολο εκπαίδευσης και το υπόλοιπο 10% ως σύνολο αξιολόγησης. Επιπλέον, σε κάθε επανάληψη (fold) προσπαθήσαμε να κρατήσουμε παρόμοιο αριθμό δειγμάτων από κάθε κλάση. Οι τιμές των παραμέτρων γ και C του Γκαουσιανού πυρήνα και το soft-margin του ταξινομητή, επιλέχθηκαν (tuned) χρησιμοποιώντας το 20% της βάσης. Στην παρούσα εφαρμογή χρησιμοποιήθηκε η βιβλιοθήκη LibSVM (Chang and Lin 2011) για την εκπαίδευση και αξιολόγηση του ταξινομητή.

Ground Truth	Αναγνώριση τοπολογικών σχέσεων					FN (%)
	δεξιά	κάτω-δεξιά	πάνω-δεξιά	κάτω	μέσα	
δεξιά	14,443	103	73	0	3	1.22
κάτω-δεξιά	261	994	0	4	0	21.05
πάνω-δεξιά	128	0	1,790	1	0	6.72
κάτω	11	18	0	2,547	2	1.20
μέσα	3	0	0	1	854	0.47
FP (%)	2.71	10.85	3.92	0.24	0.58	

Πίνακας 8. Πίνακας ανάλυσης σφαλμάτων για τον ταξινομητή SVM-7f (10-fold cross validation) στην βάση MathBrush.

Παρατηρώντας τον Πίνακα 8, συμπεραίνουμε ότι ο ταξινομητής SVM-7f παράγει περισσότερα λάθη όταν προσπαθεί να αναγνωρίσει τις σχέσεις «πάνω-δεξιά» και «κάτω-δεξιά», από όταν προσπαθεί να αναγνωρίσει τις σχέσεις «κάτω», «μέσα» και «δεξιά». Συγκεκριμένα 128 από τις 1.919 σχέσεις «πάνω-δεξιά» αναγνωρίζονται λανθασμένα ως σχέσεις «δεξιά», ενώ μόνο 1 σχέση από αυτές αναγνωρίζεται ως σχέση «κάτω». Από τις 1.259 σχέσεις «κάτω-δεξιά», 261 αναγνωρίστηκαν λανθασμένα ως σχέσεις «δεξιά» και 4 ως σχέσεις «κάτω».

Τα αποτελέσματα του Πίνακα 8 συνάδουν με τα αντίστοιχα των **Álvaro και Zanibbi (2013)**, όπου και οι δύο μέθοδοι αναδεικνύουν ότι η σχέση «κάτω-δεξιά» έχει το υψηλότερο ποσοστό False Negative. Αυτό έγκειται στο γεγονός της ύπαρξης μεγάλου αριθμού σχέσεων «κάτω-δεξιά» που αναγνωρίζονται λανθασμένα ως σχέσεις «δεξιά». Είναι φανερό λοιπόν, ότι η σχέση «κάτω-δεξιά» παράγει τα περισσότερα λάθη ταξινόμησης. Παρόλα αυτά, το μέσο ποσοστό λάθους του ταξινομητή είναι 2.87%, το οποίο θεωρείται αρκετά χαμηλό και ικανό να αντιμετωπίσει το πρόβλημα της αναγνώρισης των ME, όπως περιγράφεται στο επόμενο κεφάλαιο.

Για την αξιολόγηση των ANN ταξινομητών χρησιμοποιήθηκαν οι βάσεις CROHME2012 και MathBrush. Για την βάση τοπολογικών σχέσεων CROHME2012 υλοποιήθηκαν οι ταξινομητές ANN-6f-6sp που χρησιμοποιεί 6 γεωμετρικά χαρακτηριστικά και ο ταξινομητής ANN-7f-6sp

που χρησιμοποιεί και τα 7 γεωμετρικά χαρακτηριστικά. Αντίστοιχα για την βάση τοπολογικών σχέσεων Mathbrush υλοποιήθηκαν οι ταξινομητές ANN-6f-5sp και ANN-7f-5sp. Ο ταξινομητής ANN-6f-6sp δίνει ορθό ποσοστό αναγνώρισης ίσο με 97.11%, ενώ ο ταξινομητής ANN-7f-6sp δίνει για την ίδια βάση δεδομένων (CROHME2012) ποσοστό ορθής αναγνώρισης ίσο με 96.27%. Αντίστοιχα ο ταξινομητής ANN-6f-5sp δίνει ορθό ποσοστό αναγνώρισης ίσο με 95.91%, ενώ ο ταξινομητής ANN-7f-5sp δίνει για την ίδια βάση δεδομένων (MathBrush) ποσοστό ορθής αναγνώρισης ίσο με 96.65%. Στο πίνακα 9 παρουσιάζονται τα αποτελέσματα των ANN ταξινομητών με 6 ή 7 χαρακτηριστικά στις βάσεις CROHME2012 και MathBrush.

<i>Βάση δεδομένων</i>	<i>Ταξινομητής</i>	<i>Ποσοστό λανθασμένης αναγνώρισης (%)</i>
CROHME-2012	ANN-7f-6sp	3.73
	ANN-6f-6sp	2.89
MathBrush	ANN-7f-5sp	3.35
	ANN-6f-5sp	4.09

Πίνακας 9. Ποσοστά λανθασμένης αναγνώρισης για τους ANN ταξινομητές.

Παρατηρώντας τους Πίνακες 10 και 11 μπορούμε να συμπεράνουμε ότι οι ANN-7f-6sp και ANN-7f-5sp ταξινομητές παρουσιάζουν δυσκολία στην αναγνώριση των σχέσεων «πάνω-δεξιά» και «κάτω-δεξιά». Συγκεκριμένα όταν χρησιμοποιείται ο ταξινομητής ANN-7f-6sp, 8 τοπολογικές σχέσεις πάνω δεξιά ταξινομούνται ως «δεξιά». Αντίστοιχα από τις 169 τοπολογικές σχέσεις «κάτω-δεξιά» 6 ταξινομούνται ως «δεξιά», 7 ταξινομούνται ως «μέσα» και 1 ως «κάτω». Επιπλέον 7 τοπολογικές σχέσεις «δεξιά» ταξινομούνται ως «πάνω-δεξιά» και 6 τοπολογικές σχέσεις «δεξιά» ταξινομούνται ως «κάτω-δεξιά».

Ground Truth	Αναγνώριση τοπολογικών σχέσεων						FN (%)
	πάνω	πάνω-δεξιά	δεξιά	μέσα	κάτω-δεξιά	κάτω	
πάνω	171	4	1	0	0	0	2.84
πάνω-δεξιά	0	195	8	0	0	0	3.94
δεξιά	2	7	817	5	6	1	2.51
μέσα	2	0	6	218	6	2	6.84
κάτω-δεξιά	0	0	6	7	155	1	8.28
κάτω	0	0	1	4	2	279	2.45
FP (%)	2.29	5.34	2.62	6.84	8.28	1.41	

Πίνακας 10. Πίνακας ανάλυσης σφαλμάτων για τον ταξινομητή ANN-7f-6sp στην βάση δεδομένων CROHME2012.

Αντίστοιχα όταν χρησιμοποιείται ο ταξινομητής ANN-7f-5sp, 138 τοπολογικές σχέσεις πάνω δεξιά ταξινομούνται ως «δεξιά» και 12 τοπολογικές σχέσεις «κάτω-δεξιά» ταξινομούνται ως «δεξιά». Επιπλέον 81 τοπολογικές σχέσεις «δεξιά» ταξινομούνται ως «πάνω-δεξιά» και 152 τοπολογικές σχέσεις «δεξιά» ταξινομούνται ως «κάτω-δεξιά».

Ground Truth	Αναγνώριση τοπολογικών σχέσεων					FN (%)
	πάνω-δεξιά	δεξιά	μέσα	κάτω-δεξιά	κάτω	
πάνω-δεξιά	1780	138	0	0	1	7.24
Δεξιά	81	14380	3	152	8	1.67

Μέσα	0	9	846	2	1	1.40
κάτω-δεξιά	0	268	0	979	12	22.24
Κάτω	0	9	6	22	2541	1.44
FP (%)	4.35	2.86	1.05	15.24	0.86	

Πίνακας 11. Πίνακας ανάλυσης σφαλμάτων για τον ταξινομητή ANN-7f-5sp στην βάση δεδομένων MathBrush.

4.5. Συμπεράσματα

Στην παρούσα εργασία υλοποιήθηκαν τρεις διαφορετικοί ταξινομητές (ILSP-1, SVM και ANN) και χρησιμοποιήθηκαν τρία διαφορετικά σύνολα γεωμετρικών χαρακτηριστικών για την αντιμετώπιση του προβλήματος της αναγνώρισης τοπολογικής σχέσης μεταξύ συμβόλων. Οι ταξινομητές αξιολογήθηκαν σε δύο βάσεις ME με επισημειωμένες τοπολογικές σχέσεις: (α) ένα υποσύνολο της βάσης χειρόγραφων ME CROHME-2012 και (β) την βάση χειρόγραφων ME MathBrush. Για την υλοποίηση του SVM ταξινομητή χρησιμοποιήθηκαν δύο τεχνικές: (α) η τεχνική «one-against-one» και (β) η τεχνική «one-against-all». Τα πειραματικά αποτελέσματα έδειξαν ότι ο SVM ταξινομητής που υλοποιήθηκε σύμφωνα με την τεχνική «one-against-one» παρουσιάζει τα καλύτερα αποτελέσματα ορθής αναγνώρισης.

Παρατηρήθηκε ότι χρησιμοποιώντας ένα επιπλέον χαρακτηριστικό στον SVM ταξινομητή, το ποσοστό σφάλματος μειώνεται από 3.20% σε 2.87% για την βάση χειρόγραφων ME MathBrush. Επιπλέον, το ποσοστό λάθους 2.87% είναι άμεσα συγκρίσιμο με το ποσοστό λάθους άλλων δημοσιευμένων τεχνικών που χρησιμοποιούν την ίδια βάση δεδομένων για την αξιολόγηση τους (**Álvaro και Zanibbi 2013**). Αντίστοιχα το ποσοστό λάθους του ANN ταξινομητή μειώνεται από 4.09% σε 3.35% όταν χρησιμοποιείται ένα επιπλέον χαρακτηριστικό στην βάση MathBrush. Παρόλα αυτά είναι μεγαλύτερο από το αντίστοιχο ποσοστό λάθους του SVM ταξινομητή. Είναι πιθανό αυτό να οφείλεται στην απλή αρχιτεκτονική του προτεινόμενου ANN ταξινομητή (περιέχει μόνο ένα κρυφό επίπεδο). Αναμένεται ότι η χρήση περισσότερων κρυφών επιπέδων στην αρχιτεκτονική του ANN θα μειώσει το ποσοστό λάθους του ταξινομητή.

Αξίζει να σημειωθεί ότι υπάρχει ένας σημαντικός περιορισμός στην ορθή αναγνώριση της τοπολογικής σχέσης μεταξύ δύο μεμονωμένων συμβόλων, καθώς είναι μάλλον απίθανο να υπάρξει τέλειο ποσοστό αναγνώρισης χωρίς την «γνώση» του περιεχομένου της ME που περιλαμβάνει τα υπό εξέταση σύμβολα. Μία πιθανή επέκταση της προτεινόμενης τεχνικής μπορεί να είναι η ανάπτυξη διαφορετικών μεθόδων που να αξιοποιούν την πληροφορία του «περιεχομένου» της ME (**Awal et al. 2010**).

Κεφάλαιο 5. Ανασύσταση της δομής χειρόγραφων ΜΕ

Το τελευταίο στάδιο της διαδικασίας αναγνώρισης μαθηματικών παραστάσεων είναι η ανακατασκευή της παράστασης στην αρχική της μορφή από τους μεμονωμένους χαρακτήρες που έχουμε αναγνωρίσει. Το σημείο αυτό είναι και το πιο σημαντικό στο οποίο η αναγνώριση μαθηματικών παραστάσεων διαφοροποιείται από την αναγνώριση απλών κειμένων. Στην περίπτωση του χειρόγραφου κειμένου οι χαρακτήρες είναι διατεταγμένοι στην ίδια ευθεία, από αριστερά προς τα δεξιά, δηλαδή σε μία διάσταση. Αντίθετα, σε μία μαθηματική έκφραση, τα σύμβολα που την αποτελούν δεν τοποθετούνται όλα στην ίδια ευθεία, αλλά σε διάφορα σημεία του επιπέδου. Πρόκειται δηλαδή για μία διάταξη στις δύο διαστάσεις. Οι βασικοί λόγοι που δίνουν στις μαθηματικές εκφράσεις αυτό το δισδιάστατο χαρακτήρα είναι οι εξής: (1) Ορισμένα σύμβολα ακολουθούνται από άλλα που δεν τοποθετούνται στη βασική γραμμή (baseline) της έκφρασης αλλά ψηλότερα ή χαμηλότερα από αυτό, ώστε να λειτουργούν ως εκθέτες ή δείκτες αντίστοιχα. (2) Πολλά μαθηματικά σύμβολα (όπως το σύμβολο του αθροίσματος, το ολοκλήρωμα, η γραμμή κλάσματος, το σύμβολο της ρίζας) συνοδεύονται από ολόκληρες παραστάσεις που τα προσδιορίζουν περαιτέρω. Για μερικά από αυτά οι εν λόγω παραστάσεις τοποθετούνται εκτός της βασικής γραμμής της έκφρασης. Επιπλέον, η ιδιαίτερη θέση του κάθε συμβόλου στο χώρο δίνει στον αναγνώστη πληροφορίες σχετικά με τον ρόλο του στην παράσταση. Συνεπώς μία μαθηματική έκφραση αποτελείται από την βασική γραμμή και/ή από υπό-παραστάσεις που βρίσκονται τοποθετημένες εκτός της βασικής γραμμής, σε άλλα επίπεδα της ΜΕ. Οι υπό-παραστάσεις με την σειρά τους μπορεί να περιέχουν επιμέρους επίπεδα κ.ο.κ. Το φαινόμενο μπορεί να επαναλαμβάνεται, δίνοντας στη συνολική έκφραση αυθαίρετη μορφή και πολυπλοκότητα.

Το στάδιο αυτό στοχεύει στην εύρεση της τοπολογίας των συμβόλων που απαρτίζουν την μαθηματική έκφραση, στην ανασύσταση και στην αναπαράστασή της σε μορφή MathML (Παράρτημα Β).

5.1. Βάση δεδομένων χειρόγραφων ΜΕ

Η βάση χειρόγραφων μαθηματικών παραστάσεων CROHME2011 (Competition on Recognition of Online Handwritten Mathematical Expressions) (**Mouchère et al. 2011**) περιλαμβάνει 1450 μαθηματικές εξισώσεις, από τις οποίες οι 921 ανήκουν στο σύνολο εκπαίδευσης και οι 529 στο σύνολο εξέτασης. Τα σύμβολα που περιέχονται στις ΜΕ φαίνονται στον Πίνακα 12. Όλες οι ΜΕ είναι αποθηκευμένες σε μορφότυπο InkML (Παράρτημα Α) (<http://www.w3.org/TR/InkML/>). Το σύνολο εκμάθησης είναι χωρισμένο σε δύο μέρη, το Μέρος-I που αποτελείται από 296 ΜΕ και το Μέρος-II που αποτελείται από 921 ΜΕ (το σύνολο των ΜΕ του Μέρους-I και 625 διαφορετικές ΜΕ). Τα 37 σύμβολα που περιέχονται στις ΜΕ του Μέρους-I είναι κοινά και για το Μέρος-II (Πίνακας 12). Οι ΜΕ του Μέρους-II

περιέχουν επιπλέον άλλα 20 σύμβολα που παρουσιάζονται στον Πίνακα 13. Συνολικά το Μέρος-II περιλαμβάνει 87 σύμβολα. Αντίστοιχα το σύνολο εξέτασης αποτελείται από το Μέρος-I με 181 ΜΕ και το Μέρος-II με 348 ΜΕ. Ο τρόπος που χωρίζονται οι ΜΕ σε δύο μέρη είναι ανάλογα με την πολυπλοκότητα της ΜΕ και το πλήθος των συμβόλων που τις απαρτίζουν. Συνεπώς, οι ΜΕ που ανήκουν στο Μέρος-I είναι απλούστερες από αυτές που ανήκουν στο Μέρος-II καθώς περιλαμβάνουν λιγότερα σύμβολα και η τοπολογία των συμβόλων είναι απλούστερη.

Κοινά σύμβολα για Μέρος-I και Μέρος-II	
a, b, c, d, e, i, k, η, x, y, z	11
0, 1, 2, 3, 4, 5, 6, 7, 8, 9	10
φ, π, θ	3
+ , - , \pm	3
sin, cos	2
\dots	1
$\neq, \leq, <$	3
=	1
(,)	2
$\sqrt{\quad}$	1
	37

Πίνακας 12. Κοινά σύμβολα που περιέχονται στις ΜΕ του Μέρους-I και Μέρους-II.

Επιπλέον σύμβολα για το Μέρος-II	
A, B, C, F, j	5
α, β, γ	3
∞, \dots	2
\div, \times	2
\sum, \int	2
log, tan	2
\geq	1
lim, !, \longrightarrow	3
	20

Πίνακας 13. Επιπλέον σύμβολα που περιέχονται στις ΜΕ του Μέρους-II.

f, g, h, l, m, o, p, q, r, s, t, u, v, w, E, G, H, I, L, M, N, P, R, S, T, V, X, Y	28
λ, μ, σ	3
{, }, [,]	4
, ...	2
prime, " , ' , " "	4
\exists, \forall, \in	3
	44

Πίνακας 14. Επιπλέον σύμβολα που περιέχονται στις ΜΕ του Μέρους-III και Μέρους-IV.

Η βάση χειρόγραφων μαθηματικών παραστάσεων CROHME2012 (**Mouchère et al. 2012**) είναι χωρισμένη σε 3 μέρη (Μέρος-I, Μέρος-II και Μέρος-III) ανάλογα με την πολυπλοκότητα της ΜΕ και το πλήθος των συμβόλων που τις απαρτίζουν, ενώ η βάση χειρόγραφων μαθηματικών παραστάσεων CROHME2013 (**Mouchère et al. 2013**) αποτελείται μόνο από ένα μέρος (Μέρος-IV). Συνεπώς, οι ΜΕ που ανήκουν στο Μέρος-I είναι απλούστερες από

αυτές που ανήκουν στο Μέρος-IV καθώς περιλαμβάνουν λιγότερα σύμβολα και η τοπολογία των συμβόλων είναι απλούστερη. Τα τέσσερα μέρη του συνόλου εκμάθησης αποτελούνται από 296, 921, 1.336 και 8.836 μαθηματικές εκφράσεις αντίστοιχα. Τα αντίστοιχα τέσσερα μέρη του συνόλου εξέτασης αποτελούνται από 108, 301, 486 και 671 μαθηματικές εκφράσεις. Κάθε μέρος περιγράφεται από την δική του γραμματική. Συγκεκριμένα το Μέρος-IV περιέχει όλα τα 101 μαθηματικά σύμβολα που περιέχονται στα υπόλοιπα 3 μέρη (Μέρος-I, Μέρος-II και Μέρος-III).

Η δομή των αποθηκευμένων InkML αρχείων (Παράρτημα Α) που περιγράφουν τις ΜΕ περιλαμβάνει τα ακόλουθα είδη πληροφορίας: α) τα στοιχεία του γραφέα, β) την αποτύπωση της ΜΕ σε μορφότυπο Latex, γ) τον κωδικό ταυτοποίησης του ψηφιακού μελανιού, δ) τις συντεταγμένες x , y του ψηφιακού μελανιού, ε) τα αληθή δεδομένα (ground truth) των συμβόλων και στ) τα αληθή δεδομένα (ground truth) της ΜΕ (δηλ. τη MathML δομή της ΜΕ βλ. Παράρτημα Β). Στα στοιχεία του γραφέα περιλαμβάνονται πληροφορίες όπως ο κωδικός ταυτοποίησης του χρήστη, το ονοματεπώνυμο, η ηλικία, το φύλλο, πληροφορία για τον αν ο χρήστης είναι αριστερόχειρας ή δεξιόχειρας κ.λπ.

Οι μαθηματικές εκφράσεις που περιέχονται στο Μέρος-I επιτρέπουν ένα σύμβολο στην θέση εκθέτη/δείκτη (Εικόνα 24 (α), (β)) και επαναληπτικές ρίζες (Εικόνα 24 (γ)). Δεν επιτρέπονται εκφράσεις με επαναληπτικά κλάσματα. Παρόλα αυτά επιτρέπονται αθροίσματα κλασμάτων (Εικόνα 24 (ε)) και κλάσματα με αθροίσματα (Εικόνα 24 (δ)).

(a) $y = ax^2 + b$

(β) $\phi_k = \pi$

(γ) $\sqrt{1 + \sqrt{2 + \sqrt{3 + \sqrt{4}}}}$

(δ) $\frac{\sin B + \sin C}{\cos B + \cos C}$

(ε) $\frac{c}{d} - \frac{a}{b} = \frac{(bc - ad)}{bd}$

Εικόνα 24. Παραδείγματα χειρόγραφων μαθηματικές εκφράσεων από το Μέρος-I της βάσης

δεδομένων CROHME (α) $y = ax^2 + b$, (β) $\phi_k = \pi$, (γ) $\sqrt{1 + \sqrt{2 + \sqrt{3 + \sqrt{4}}}}$,
 (δ) $\frac{\sin B + \sin C}{\cos B + \cos C}$, (ε) $\frac{c}{d} - \frac{a}{b} = \frac{(bc - ad)}{bd}$

Η μαθηματική έκφραση $y = ax^2 + b$ που παρουσιάζεται στην Εικόνα 24 (α) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mi>y</mi>
4 <mrow>
5 <mo>=</mo>
6 <mrow>
7 <mi>a</mi>
8 <mrow>
9 <msup>
10 <mi>x</mi>
11 <mn xml:id="2_1">2</mn>
12 </msup>
13 <mrow>
14 <mo>+</mo>
15 <mi>b</mi>
16 </mrow>
17 </mrow>
18 </mrow>
19 </mrow>
20 </mrow>
21 </math>

```

Στην γραμμή-1 ορίζεται το μορφότυπο MathML. Το μορφότυπο MathML χρησιμοποιείται συνήθως για την αποτύπωση των μαθηματικών εκφράσεων στο διαδίκτυο. Η πληροφορία που περιέχει είναι δομημένη και αποτελείται από το περιεχόμενο, τις ετικέτες που δείχνουν τι είδος είναι το περιεχόμενο (mi για τους identifiers, mo για τους τελεστές και mn για τους αριθμούς) και τις ετικέτες δομής του περιεχομένου (π.χ. η ετικέτα mrow ορίζει μία οριζόντια γραμμή με αντικείμενα, η ετικέτα mfrac για τα κλάσματα, η ετικέτα msqrt για την τετραγωνική ρίζα, η ετικέτα mroot ορίζει μία ν-ρίζα, η ετικέτα msup ορίζει μία εκθετική έκφραση κ.λπ.). Για παράδειγμα η γραμμή 3 δείχνει ότι το περιεχόμενο y είναι mi, και η γραμμή 2 δείχνει ότι το y βρίσκεται σε μία οριζόντια γραμμή mrow. Στις γραμμές 9-12 ορίζεται η βάση x του εκθέτη 2, δηλ. η υποέκφραση x^2 .

Ένας εναλλακτικός τρόπος γραφής της μαθηματικής έκφρασης $y = ax^2 + b$ είναι ο ακόλουθος:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mi xml:id="y_1" style="">y</mi>
4 <mo xml:id="=_1" style="">=</mo>
5 <mrow>
6 <mi xml:id="a_1" style="">a</mi>
7 <msup>
8 <mi xml:id="x_1">x</mi>
9 <mn xml:id="2_1">2</mn>
10 </msup>
11 </mrow>
12 <mo xml:id="+_1">+</mo>
13 <mi xml:id="b_1">b</mi>
14 </mrow>
15 </math>

```


Συνεπώς μία μαθηματική έκφραση μπορεί να γραφτεί με περισσότερους από ένα τρόπους σύμφωνα με το μορφότυπο MathML. Η διαφορετικότητα έγκειται στον ορισμό των οριζόντιων γραμμών της ΜΕ. Για παράδειγμα στην πρώτη αποτύπωση της ΜΕ $y = ax^2 + b$, ο ορισμός των οριζόντιων γραμμών μπορεί να αποτυπωθεί ως: $(y=(a(x^2+b))))$ ενώ στην δεύτερη αποτύπωση της ίδιας ΜΕ, ο ορισμός των οριζόντιων γραμμών μπορεί να αποτυπωθεί ως: $(y=(ax^2)+b)$.

Αντίστοιχα η μαθηματική έκφραση $\phi_k = \pi$ που παρουσιάζεται στην Εικόνα 24 (β) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <msub>
4 <mi xml:id="phi_1" style="">&phi;</mi>
5 <mi xml:id="k_1">k</mi>
6 </msub>
7 <mrow>
8 <mo xml:id="=_1">=</mo>
9 <mi xml:id="pi_1">&pi;</mi>
10 </mrow>
11 </mrow>
12 </math>

```

Η μαθηματική έκφραση $\sqrt{1+\sqrt{2+\sqrt{3+\sqrt{4}}}}$ που παρουσιάζεται στην Εικόνα 24 (γ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <msqrt xml:id="_1" style="">
3 <mn xml:id="1_1">1</mn>
4 <mrow>
5 <mo xml:id="+_1">+</mo>
6 <msqrt xml:id="_2" style="">
7 <mn xml:id="2_1" style="">2</mn>
8 <mrow>
9 <mo xml:id="+_2" style="">+</mo>
10 <msqrt xml:id="_3" style="">
11 <mn xml:id="3_1" style="">3</mn>
12 <mrow>
13 <mo xml:id="+_3" style="">+</mo>
14 <msqrt xml:id="_4" style="">
15 <mn xml:id="4_1">4</mn>
16 </msqrt>
17 </mrow>
18 </msqrt>
19 </mrow>
20 </msqrt>
21 </mrow>
22 </msqrt>
23 </math>

```

Η μαθηματική έκφραση $\frac{\sin B + \sin C}{\cos B + \cos C}$ που παρουσιάζεται στην Εικόνα 24 (δ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mfrac xml:id="-_1" style="">
3 <mrow>
4 <mo xml:id="\sin_1">sin</mo>
5 <mrow>
6 <mi xml:id="B_1">B</mi>
7 <mrow>
8 <mo xml:id="+_1">+</mo>
9 <mrow>
10 <mo xml:id="\sin_2">sin</mo>
11 <mi xml:id="C_1" style="">C</mi>
12 </mrow>
13 </mrow>
14 </mrow>
15 </mrow>
16 <mrow>
17 <mo xml:id="\cos_1">cos</mo>
18 <mrow>
19 <mi xml:id="B_2">B</mi>
20 <mrow>
21 <mo xml:id="+_2" style="">+</mo>
22 <mrow>
23 <mo xml:id="\cos_2" style="">cos</mo>
24 <mi xml:id="C_2">C</mi>
25 </mrow>
26 </mrow>
27 </mrow>
28 </mrow>
29 </mfrac>
30 </math>

```

Η μαθηματική έκφραση $\frac{c}{d} - \frac{a}{b} = \frac{bc - ad}{bd}$ που παρουσιάζεται στην Εικόνα 24 (ε) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mfrac xml:id="_1">
4 <mi xml:id="c_1">c</mi>
5 <mi xml:id="d_1">d</mi>
6 </mfrac>
7 <mrow>
8 <mo xml:id="-_1">-</mo>
9 <mrow>
10 <mfrac xml:id="_2" style="">
11 <mi xml:id="a_1">a</mi>
12 <mi xml:id="b_1" style="">b</mi>
13 </mfrac>
14 <mrow>
15 <mo xml:id="=_1">=</mo>
16 <mfrac xml:id="_3">
17 <mrow>
18 <mo xml:id="(_1" style="">(</mo>
19 <mrow>
20 <mi xml:id="b_2">b</mi>
21 <mrow>
22 <mi xml:id="c_2">c</mi>

```

```

23 <mrow>
24 <mo xml:id="-_2" style="">-</mo>
25 <mrow>
26 <mi xml:id="a_2">a</mi>
27 <mrow>
28 <mi xml:id="d_2">d</mi>
29 <mo xml:id=")_1" style="">></mo>
30 </mrow>
31 </mrow>
32 </mrow>
33 </mrow>
34 </mrow>
35 </mrow>
36 <mrow>
37 <mi xml:id="b_3">b</mi>
38 <mi xml:id="d_3">d</mi>
39 </mrow>
40 </mfrac>
41 </mrow>
42 </mrow>
43 </mrow>
44 </mrow>
45 </math>

```

Οι περιορισμοί που διέπουν τις μαθηματικές εκφράσεις του Μέρους-II είναι λιγότερες (βλ. Εικόνα 25). Επιπλέον, δεν υπάρχει περιορισμός στην επαναληπτικότητα των τελεστών (sum, product, function call, fraction, root, sub/superscript) όπως συμβαίνει στις ΜΕ του Μέρους-I.

$$c_{e_i+1}$$

(α)

$$2^{e^k} + 1$$

(β)

$$\lim_{x \rightarrow 0} \frac{(1 - \cos x)(1 + \cos x)}{x^2 (1 + \cos x)}$$

(γ)

$$\sum_{k=1}^n k = \frac{1}{2}(n^2 + n)$$

(δ)

$$\int_1^2 \left(\frac{x^2 - 1}{x^2} \right) e^{x + \frac{1}{x}} dx$$

(ε)

Εικόνα 25. Παραδείγματα χειρόγραφων μαθηματικών εκφράσεων από το Μέρος-II της βάσης

δεδομένων CROHME (α) c_{e_i+1} , (β) $2^{e^k} + 1$, (γ) $\lim_{x \rightarrow 0} \frac{(1 - \cos x)(1 + \cos x)}{x^2 (1 + \cos x)}$,

(δ) $\sum_n^{k=1} k = \frac{1}{2}(n^2 + n)$, (ε) $\int_1^2 \left(\frac{x^2 - 1}{x^2} \right) e^{x + \frac{1}{x}} dx$

Η μαθηματική έκφραση c_{e_i+1} που παρουσιάζεται στην Εικόνα 25 (α) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <msub>
3 <mi xml:id="c_1" style="">c</mi>
4 <mrow>
5 <msub>
6 <mi xml:id="e_1" style="">e</mi>
7 <mi xml:id="i_1" style="">i</mi>
8 </msub>
9 <mrow>
10 <mo xml:id="+_1" style="">+</mo>
11 <mn xml:id="1_1" style="">1</mn>
12 </mrow>
13 </mrow>
14 </msub>
15 </math>

```

Η μαθηματική έκφραση $2^{2^k} + 1$ που παρουσιάζεται στην Εικόνα 25 (β) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <msup>
4 <mn xml:id="2_1">2</mn>
5 <msup>
6 <mn xml:id="2_2" style="">2</mn>
7 <mi xml:id="k_1" style="">k</mi>
8 </msup>
9 </msup>
10 <mrow>
11 <mo xml:id="+_1" style="">+</mo>
12 <mn xml:id="1_1" style="">1</mn>
13 </mrow>
14 </mrow>
15 </math>

```

Η μαθηματική έκφραση $\lim_{x \rightarrow 0} \frac{(1 - \cos x)(1 + \cos x)}{x^2(1 + \cos x)}$ που παρουσιάζεται στην Εικόνα 25 (γ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <munder>
4 <mo xml:id="\lim_1">lim</mo>
5 <mrow>
6 <mi xml:id="x_1">x</mi>
7 <mrow>
8 <mo xml:id="\rightarrow_1">&rarr;</mo>
9 <mn xml:id="0_1">0</mn>
10 </mrow>
11 </mrow>
12 </munder>
13 <mfrac xml:id="-_1" style="">
14 <mrow>

```

```

15 <mo xml:id="(_1">(</mo>
16 <mrow>
17 <mn xml:id="1_1">1</mn>
18 <mrow>
19 <mo xml:id="-_2">-</mo>
20 <mrow>
21 <mo xml:id="\cos_1">cos</mo>
22 <mrow>
23 <mi xml:id="x_2">x</mi>
24 <mrow>
25 <mo xml:id=")_1" style="">)</mo>
26 <mrow>
27 <mo xml:id="(_2" style="">(</mo>
28 <mrow>
29 <mn xml:id="1_2">1</mn>
30 <mrow>
31 <mo xml:id="+_1" style="">+</mo>
32 <mrow>
33 <mo xml:id="\cos_2" style="">cos</mo>
34 <mrow>
35 <mi xml:id="x_3" style="">x</mi>
36 <mo xml:id=")_2" style="">)</mo>
37 </mrow>
38 </mrow>
39 </mrow>
40 </mrow>
41 </mrow>
42 </mrow>
43 </mrow>
44 </mrow>
45 </mrow>
46 </mrow>
47 </mrow>
48 <mrow>
49 <msup>
50 <mi xml:id="x_4" style="">x</mi>
51 <mn xml:id="2_1">2</mn>
52 </msup>
53 <mrow>
54 <mo xml:id="(_3">(</mo>
55 <mrow>
56 <mn xml:id="1_3">1</mn>
57 <mrow>
58 <mo xml:id="+_2">+</mo>
59 <mrow>
60 <mo xml:id="\cos_3">cos</mo>
61 <mrow>
62 <mi xml:id="x_5">x</mi>
63 <mo xml:id=")_3">)</mo>
64 </mrow>
65 </mrow>
66 </mrow>
67 </mrow>
68 </mrow>
69 </mrow>
70 </mfrac>
71 </mrow>

```

72 </math>

Η μαθηματική έκφραση $\sum_n^{k=1} k = \frac{1}{2}(n^2 + n)$ που παρουσιάζεται στην Εικόνα 25 (δ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```
1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <munderover>
4 <mo xml:id="\sum_1" style="">&Sigma;</mo>
5 <mrow>
6 <mi xml:id="k_1" style="">k</mi>
7 <mrow>
8 <mo xml:id="=_1" style="">=</mo>
9 <mn xml:id="1_1">1</mn>
10 </mrow>
11 </mrow>
12 <mi xml:id="n_1">n</mi>
13 </munderover>
14 <mrow>
15 <mi xml:id="k_2" style="">k</mi>
16 <mrow>
17 <mo xml:id="=_2" style="">=</mo>
18 <mrow>
19 <mfrac xml:id="-_1" style="">
20 <mn xml:id="1_2" style="">1</mn>
21 <mn xml:id="2_1">2</mn>
22 </mfrac>
23 <mrow>
24 <mo xml:id="(_1" style="">(</mo>
25 <mrow>
26 <msup>
27 <mi xml:id="n_2" style="">n</mi>
28 <mn xml:id="2_2" style="">2</mn>
29 </msup>
30 <mrow>
31 <mo xml:id="+_1">+</mo>
32 <mrow>
33 <mi xml:id="n_3" style="">n</mi>
34 <mo xml:id=")_1" style="">)</mo>
35 </mrow>
36 </mrow>
37 </mrow>
38 </mrow>
39 </mrow>
40 </mrow>
41 </mrow>
42 </mrow>
43 </math>
```

Η μαθηματική έκφραση $\int_1^2 \left(\frac{x^2-1}{x^2}\right) e^{x+\frac{1}{x}} dx$ που παρουσιάζεται στην Εικόνα 25 (ε) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```
1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
```

```

3 <msubsup>
4 <mo xml:id="\int_1">&int;</mo>
5 <mn xml:id="1_1">1</mn>
6 <mn xml:id="2_1">2</mn>
7 </msubsup>
8 <mrow>
9 <mo xml:id="(_1" style=""></mo>
10 <mrow>
11 <mfrac xml:id="-_1" style="">
12 <mrow>
13 <msup>
14 <mi xml:id="x_1">x</mi>
15 <mn xml:id="2_2">2</mn>
16 </msup>
17 <mrow>
18 <mo xml:id="-_2" style="">-</mo>
19 <mn xml:id="1_2" style="">1</mn>
20 </mrow>
21 </mrow>
22 <msup>
23 <mi xml:id="x_2">x</mi>
24 <mn xml:id="2_3">2</mn>
25 </msup>
26 </mfrac>
27 <mrow>
28 <mo xml:id=")_1" style=""></mo>
29 <mrow>
30 <msup>
31 <mi xml:id="e_1">e</mi>
32 <mrow>
33 <mi xml:id="x_3">x</mi>
34 <mrow>
35 <mo xml:id="+_1">+</mo>
36 <mfrac xml:id="-_3">
37 <mn xml:id="1_3">1</mn>
38 <mi xml:id="x_4">x</mi>
39 </mfrac>
40 </mrow>
41 </mrow>
42 </msup>
43 <mrow>
44 <mi xml:id="d_1">d</mi>
45 <mi xml:id="x_5" style="">x</mi>
46 </mrow>
47 </mrow>
48 </mrow>
49 </mrow>
50 </mrow>
51 </mrow>
52 </math>

```

Οι μαθηματικές εκφράσεις του Μέρους-III και Μέρους-IV είναι ακόμα πιο πολύπλοκες από αυτές των Μέρους-I και II. Μερικές αντιπροσωπευτικές μαθηματικές εκφράσεις φαίνονται στην Εικόνα 26. Περιλαμβάνουν πολλά καινούργια σύμβολα (101 κλάσεις), όπως επιπρόσθετα γράμματα του λατινικού και ελληνικού αλφαβήτου, σύμβολα όπως το σύμβολο της πρώτης

παραγώγου (βλ. Εικόνα 26 (δ)), το σύμβολο \int κλπ. Επίσης επιτρέπονται ν-ρίζες (βλ. Εικόνα 26 (ζ)).

$$d_{\theta,0}^2 = \frac{3 \cos^2 \theta - 1}{2}$$

(α)

$$\left[\frac{2}{3} \left(1 + \frac{9}{4} x \right)^{\frac{3}{2}} \frac{4}{9} \right]^4$$

(β)

$$\exists \alpha, \forall x \in [a, b]$$

(γ)

$$b' = \sum_{ij} \mu_{ij} a_{ij}$$

(δ)

$$\|f\|^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n^2 + b_n^2)$$

(ε)

$$9^{9^{9^{9^9}}}$$

(στ)

$$\sqrt[x]{y}$$

(ζ)

$$A = \{a_0, \dots, a_{|A|}\}$$

(η)

Εικόνα 26. Παραδείγματα χειρόγραφων μαθηματικών εκφράσεων από το Μέρος-III και Μέρος-IV της βάσης δεδομένων CROHME.

(α) $d_{\theta,0}^2 = \frac{3 \cos^2 \theta - 1}{2}$, (β) $\left[\frac{2}{3} \left(1 + \frac{9}{4} x \right)^{\frac{3}{2}} \frac{4}{9} \right]^4$, (γ) $\exists \alpha, \forall x \in [a, b]$, (δ) $b' = \sum_{ij} \mu_{ij} a_{ij}$, (ε)

$\|f\|^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n^2 + b_n^2)$, (στ) $9^{9^{9^{9^9}}}$, (ζ) $\sqrt[x]{y}$, (η) $A = \{a_0, \dots, a_{|A|}\}$.

Η μαθηματική έκφραση $d_{\theta,0}^2 = \frac{3 \cos^2 \theta - 1}{2}$ που παρουσιάζεται στην Εικόνα 26 (α) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

- 1 <math xmlns="http://www.w3.org/1998/Math/MathML">
- 2 <mrow>
- 3 <munderover>
- 4 <mi xml:id="d_1">d</mi>
- 5 <mrow>


```

6 <mn xml:id="0_1">0</mn>
7 <mrow>
8 <mo xml:id=","_1" style="">,</mo>
9 <mn xml:id="0_2" style="">0</mn>
10 </mrow>
11 </mrow>
12 <mn xml:id="2_1">2</mn>
13 </munderover>
14 <mrow>
15 <mo xml:id="=_1" style="">=</mo>
16 <mfrac xml:id=" "_1" style="">
17 <mrow>
18 <mn xml:id="3_1" style="">3</mn>
19 <mrow>
20 <msup>
21 <mi xml:id="\cos_1" style="">cos</mi>
22 <mn xml:id="2_2" style="">2</mn>
23 </msup>
24 <mrow>
25 <mi xml:id="\theta_1" style="">&theta;</mi>
26 <mrow>
27 <mo xml:id="-_1">-</mo>
28 <mn xml:id="1_1">1</mn>
29 </mrow>
30 </mrow>
31 </mrow>
32 </mrow>
33 <mn xml:id="2_3" style="">2</mn>
34 </mfrac>
35 </mrow>
36 </mrow>
37 </math>

```

Η μαθηματική έκφραση $[\frac{2}{3}(1 + \frac{9}{4}x)^2 \frac{4}{9}]_0^4$ που παρουσιάζεται στην Εικόνα 26 (β) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mo xml:id="["_1" style="">[</mo>
4 <mrow>
5 <mfrac xml:id="-_1">
6 <mn xml:id="2_1">2</mn>
7 <mn xml:id="3_1">3</mn>
8 </mfrac>
9 <mrow>
10 <mo xml:id="(_1">(</mo>
11 <mrow>
12 <mn xml:id="1_1">1</mn>
13 <mrow>
14 <mo xml:id="+_1">+</mo>
15 <mrow>
16 <mfrac xml:id="-_2">
17 <mn xml:id="9_1">9</mn>
18 <mn xml:id="4_1" style="">4</mn>
19 </mfrac>
20 <mrow>

```

```

21 <mi xml:id="x_1">x</mi>
22 <mrow>
23 <msup>
24 <mo xml:id=")_1" style="">></mo>
25 <mfrac xml:id="-_3" style="">
26 <mn xml:id="3_2" style="">3</mn>
27 <mn xml:id="2_2">2</mn>
28 </mfrac>
29 </msup>
30 <mrow>
31 <mfrac xml:id="-_4">
32 <mn xml:id="4_2">4</mn>
33 <mn xml:id="9_2" style="">9</mn>
34 </mfrac>
35 <msubsup>
36 <mo xml:id="]_1">]</mo>
37 <mn xml:id="0_1" style="">0</mn>
38 <mn xml:id="4_3" style="">4</mn>
39 </msubsup>
40 </mrow>
41 </mrow>
42 </mrow>
43 </mrow>
44 </mrow>
45 </mrow>
46 </mrow>
47 </mrow>
48 </mrow>
49 </math>

```

Η μαθηματική έκφραση $\exists \alpha, \forall x \in [a, b]$ που παρουσιάζεται στην Εικόνα 26 (γ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mo xml:id="\exists_1" style="">&exist;</mo>
4 <mrow>
5 <mi xml:id="\alpha_1" style="">&alpha;</mi>
6 <mrow>
7 <mo xml:id=",_1">,</mo>
8 <mrow>
9 <mo xml:id="\forall_1" style="">&forall;</mo>
10 <mrow>
11 <mi xml:id="x_1">x</mi>
12 <mrow>
13 <mo xml:id="\in_1">&isin;</mo>
14 <mrow>
15 <mo xml:id="[_1" style="">[</mo>
16 <mrow>
17 <mi xml:id="a_1" style="">a</mi>
18 <mrow>
19 <mo xml:id=",_2">,</mo>
20 <mrow>
21 <mi xml:id="b_1" style="">b</mi>
22 <mo xml:id="]_1" style="">]</mo>
23 </mrow>
24 </mrow>

```

25 </mrow>
 26 </mrow>
 27 </mrow>
 28 </mrow>
 29 </mrow>
 30 </mrow>
 31 </mrow>
 32 </mrow>
 33 </math>

Η μαθηματική έκφραση $b' = \sum_{ij} \mu_{ij} a_{ij}$ που παρουσιάζεται στην Εικόνα 26 (δ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <msup>
4 <mi xml:id="b_1" style="">b</mi>
5 <mo xml:id="\prime_1">&prime;</mo>
6 </msup>
7 <mrow>
8 <mo xml:id="=_1">=</mo>
9 <mrow>
10 <munder>
11 <mo xml:id="\sum_1" style="">&Sigma;</mo>
12 <mrow>
13 <mi xml:id="i_1">i</mi>
14 <mi xml:id="j_1">j</mi>
15 </mrow>
16 </munder>
17 <mrow>
18 <msub>
19 <mi xml:id="\mu_1">&mu;</mi>
20 <mrow>
21 <mi xml:id="i_2" style="">i</mi>
22 <mi xml:id="j_2" style="">j</mi>
23 </mrow>
24 </msub>
25 <msub>
26 <mi xml:id="a_1" style="">a</mi>
27 <mrow>
28 <mi xml:id="i_3">i</mi>
29 <mi xml:id="j_3">j</mi>
30 </mrow>
31 </msub>
32 </mrow>
33 </mrow>
34 </mrow>
35 </mrow>
36 </math>
```

Η μαθηματική έκφραση $\|f\|^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n^2 + b_n^2)$ που παρουσιάζεται στην Εικόνα 26 (ε) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
```

```

3 <mo xml:id="|_1">|</mo>
4 <mrow>
5 <mo xml:id="|_2">|</mo>
6 <mrow>
7 <mi xml:id="f_1">f</mi>
8 <mrow>
9 <mo xml:id="|_3" style="">|</mo>
10 <mrow>
11 <msup>
12 <mo xml:id="|_4">|</mo>
13 <mn xml:id="2_1">2</mn>
14 </msup>
15 <mrow>
16 <mo xml:id="=_1" style="">=</mo>
17 <mrow>
18 <mfrac xml:id="_1">
19 <mn xml:id="1_1" style="">1</mn>
20 <mn xml:id="4_1">4</mn>
21 </mfrac>
22 <mrow>
23 <msubsup>
24 <mi xml:id="a_1" style="">a</mi>
25 <mn xml:id="0_1" style="">0</mn>
26 <mn xml:id="2_2">2</mn>
27 </msubsup>
28 <mrow>
29 <mo xml:id="+_1">+</mo>
30 <mrow>
31 <mfrac xml:id="_2">
32 <mn xml:id="1_2">1</mn>
33 <mn xml:id="2_3" style="">2</mn>
34 </mfrac>
35 <mrow>
36 <msubsup>
37 <mo xml:id="sum_1" style="">&Sigma;</mo>
38 <mrow>
39 <mi xml:id="n_1">n</mi>
40 <mrow>
41 <mo xml:id="=_2">=</mo>
42 <mn xml:id="1_3">1</mn>
43 </mrow>
44 </mrow>
45 <mrow>
46 <mo xml:id="+_2">+</mo>
47 <mi xml:id="\infty_1">&infin;</mi>
48 </mrow>
49 </msubsup>
50 <mrow>
51 <mo xml:id="(_1">(</mo>
52 <mrow>
53 <msubsup>
54 <mi xml:id="a_2">a</mi>
55 <mi xml:id="n_2">n</mi>
56 <mn xml:id="2_4">2</mn>
57 </msubsup>
58 <mrow>
59 <mo xml:id="+_3">+</mo>

```

```

60 <mrow>
61 <msubsup>
62 <mi xml:id="b_1">b</mi>
63 <mi xml:id="n_3">n</mi>
64 <mn xml:id="2_5">2</mn>
65 </msubsup>
66 <mo xml:id=")__1">)</mo>
67 </mrow>
68 </mrow>
69 </mrow>
70 </mrow>
71 </mrow>
72 </mrow>
73 </mrow>
74 </mrow>
75 </mrow>
76 </mrow>
77 </mrow>
78 </mrow>
79 </mrow>
80 </mrow>
81 </mrow>
82 </math>

```

Η μαθηματική έκφραση $9^{9^{9^{9^9}}}$ που παρουσιάζεται στην Εικόνα 26 (στ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <msup>
3 <mn>9</mn>
4 <msup>
5 <mn>9</mn>
6 <msup>
7 <mn>9</mn>
8 <msup>
9 <mn>9</mn>
10 <msup>
11 <mn>9</mn>
12 <mn>9</mn>
13 </msup>
14 </msup>
15 </msup>
16 </msup>
17 </msup>
18 </math>

```

Η μαθηματική έκφραση $\sqrt[x]{y}$ που παρουσιάζεται στην Εικόνα 26 (ζ) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mroot xml:id="_1" style="">
3 <mi xml:id="y_1" style="">y</mi>
4 <mi xml:id="x_1" style="">x</mi>
5 </mroot>
6 </math>

```

Η μαθηματική έκφραση $A = \{a_0, \dots, a_{|A|}\}$ που παρουσιάζεται στην Εικόνα 26 (η) μπορεί να γραφτεί σύμφωνα με το μορφότυπο MathML ως:

```

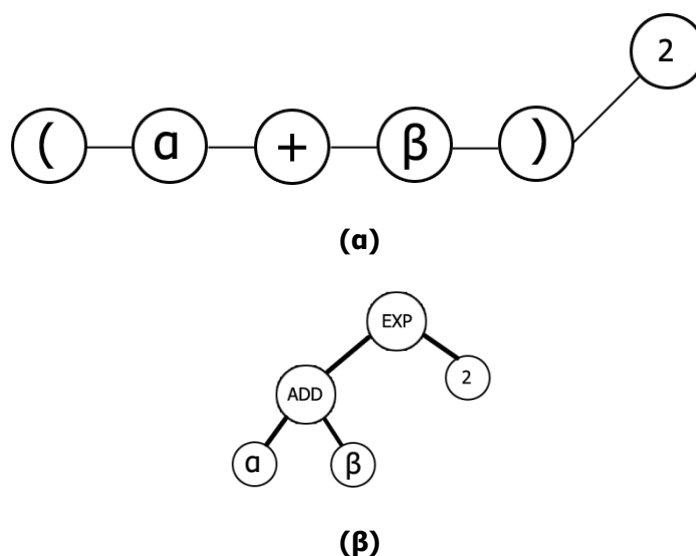
1 <math xmlns="http://www.w3.org/1998/Math/MathML">
2 <mrow>
3 <mi xml:id="A_1">A</mi>
4 <mrow>
5 <mo xml:id="=_1">=</mo>
6 <mrow>
7 <mo xml:id="\{ _1" style="">\{</mo>
8 <mrow>
9 <msub>
10 <mi xml:id="a_1">a</mi>
11 <mn xml:id="0_1" style="">0</mn>
12 </msub>
13 <mrow>
14 <mo xml:id=",_1">,</mo>
15 <mrow>
16 <mo xml:id="\ldots_1" style="">&hellip;</mo>
17 <mrow>
18 <mo xml:id=",_2">,</mo>
19 <mrow>
20 <msub>
21 <mi xml:id="a_2" style="">a</mi>
22 <mrow>
23 <mo xml:id="|_1">|</mo>
24 <mrow>
25 <mi xml:id="A_2" style="">A</mi>
26 <mo xml:id="|_2">|</mo>
27 </mrow>
28 </mrow>
29 </msub>
30 <mo xml:id="\}_1">\}</mo>
31 </mrow>
32 </mrow>
33 </mrow>
34 </mrow>
35 </mrow>
36 </mrow>
37 </mrow>
38 </mrow>
39 </math>

```

5.2. Διαθέσιμες τεχνικές

Η αναπαράσταση μιας μαθηματικής έκφρασης συνήθως γίνεται είτε μέσω ενός σχεσιακού δέντρου συμβόλων (symbol layout tree), είτε μέσω ενός σχεσιακού δέντρου τελεστών (operator tree) (Εικόνα 27). Τα δέντρα συμβόλων αναπαριστούν πληροφορία όμοια με εκείνη της Latex/MathML αναπαράστασης της ΜΕ. Προσδιορίζουν ποιες ομάδες συμβόλων βρίσκονται στην ίδια γραμμή, ποια είναι εκθέτες, ποια δείκτες, ποιά είναι αριθμητές/παρονομαστές ενός κλάσματος κ.λπ.

Οι τεχνικές που έχουν εφαρμοστεί στο τομέα της τοπολογικής ανάλυσης της ΜΕ κατηγοριοποιούνται σε πέντε μεγάλες κατηγορίες με βάση τη κύρια μεθοδολογία που ακολουθούν. Συγκεκριμένα, διακρίνονται σε αλγορίθμους που είναι βασισμένοι στην ιδέα της περιοχής κυριότητας των τελεστών (operator dominance) της ΜΕ (**Chang 1970**), που αναλύουν σε υπό-περιοχές της ΜΕ βάση των προβολών (projections ή projection profiles) των συμβόλων που τις απαρτίζουν (**Okamoto and Miao 1991**), στην εύρεση της βασικής γραμμής της έκφρασης (**Tapia and Rojas 2004**), στην ελαχιστοποίηση του κόστους του γράφου που περιγράφει την ΜΕ (**Eto and Suzuki 2001**) και στην χρήση γραμματικών κανόνων για την απεικόνιση της ΜΕ. Στη συνέχεια του κεφαλαίου παρουσιάζονται συνοπτικά τεχνικές που είτε έχουν κριθεί ως πολύ αποτελεσματικές στα αντίστοιχα σύνολα εξέτασης, είτε αποτελούν τμήματα ολοκληρωμένων συστημάτων.



Εικόνα 27. (α) Σχεσιακό δέντρο συμβόλων (symbol layout tree), και (β) σχεσιακό δέντρο τελεστών (operator tree) για την ΜΕ $(a+b)^2$.

Στην πρώτη κατηγορία των σχετικών αλγορίθμων ανήκουν αυτοί που βασίζονται στην εύρεση της περιοχής δράσης των τελεστών (operator dominance) της ΜΕ. Ο αλγόριθμος προσπαθεί επαναληπτικά να βρει τον τελεστή εκείνο που έχει την πλειοψηφία των υπολοίπων συμβόλων της ΜΕ υπό την κυριαρχία του (**Chang 1970**). Τα σύμβολα αυτά μοιράζονται στις κατάλληλες περιοχές των τελεστών (**Chan and Yeung 2001a, Chang 1970**) και αποτελούν πλέον τα σύμβολα στα οποία δρα ο τελεστής. Οι αλγόριθμοι αυτής της κατηγορίας δημιουργούν ένα σχεσιακό δέντρο τελεστών (Εικόνα 27β) το οποίο προκύπτει από την ανάλυση των συμβόλων της ΜΕ. Ένας τέτοιος αλγόριθμος περιγράφεται στην εργασία των Chan και Yeung το 2001 (**Chan and Yeung 2001b**). Οι Lee και Wang περιγράφουν στο (**Wang and Lee 1997**) μία παρόμοια προσέγγιση εφαρμόζοντας την περιοχή δράσης των τελεστών πρώτα κατά την κάθετη και έπειτα κατά την οριζόντια κατεύθυνση.

Η δεύτερη κατηγορία περιλαμβάνει αλγορίθμους που βασίζονται στην εύρεση υπό-περιοχών

της ME βάση των προβολών (projections ή projection profiles) των συμβόλων που τις απαρτίζουν. Η τεχνική αυτή είναι παρόμοια αυτής που εφαρμόζεται στην επεξεργασία εικόνων (**Nagy 2000**). Η προβολή μιας εικόνας A διαστάσεων $M \times N$, ορίζεται με την ακόλουθη σχέση, ως ένα σήμα μήκους M (ύψος εικόνας) που κάθε τιμή του ισούται με το άθροισμα των τιμών των pixels που έχουν την ίδια τετμημένη:

$$PR(i) = \sum_{j=1}^N A(i, j), \quad i=1, \dots, M \quad (1)$$

Είναι προφανές ότι οι προβολές παρουσιάζουν σημαντικούς λοβούς στις περιοχές εντός συμβόλων και εξίσου σημαντικές κοιλάδες μεταξύ των συμβόλων (πιθανές θέσεις διαχωριστικών). Από τις προβολές αυτές υπολογίζονται τα ιστογράμματα των pixels της ME κατά την οριζόντια και κάθετη κατεύθυνση. Η πρώτη διαχωριστική γραμμή υπολογίζεται στην κατακόρυφη κατεύθυνση, ώστε να διαχωρίσει τυχόν υπό-εκφράσεις που βρίσκονται στον οριζόντιο άξονα. Έπειτα υπολογίζεται η διαχωριστική γραμμή κατά την κάθετη κατεύθυνση. Η διαδικασία αυτή ακολουθείται επαναληπτικά μέχρι να μην μπορεί να υπάρξει άλλο διαχωριστικό και για τις δύο κατευθύνσεις.

Μία βελτίωση αυτής της μεθόδου αναφέρεται στην εργασία των **Raja et al. το 2006** σύμφωνα με την οποία πριν τον υπολογισμό των κάθετων και οριζόντιων διαχωριστικών της ME, εξάγονται από την ME οι κονδυλίες που τέμνονται. Κατόπιν αφαιρείται το μεγαλύτερο σύμβολο (π.χ. το σύμβολο της ρίζας) και επαναλαμβάνεται για τα υπόλοιπα σύμβολα ο υπολογισμός των διαχωριστικών γραμμών.

Στην τρίτη κατηγορία ανήκουν αλγόριθμοι βασισμένοι στην εύρεση της βασικής γραμμής της ME. Ο αλγόριθμος λειτουργεί επαναληπτικά εντοπίζοντας τα σύμβολα εκείνα που είναι τοποθετημένα από αριστερά προς τα δεξιά, συνεχόμενα στη βασική γραμμή της ME. Έπειτα διαχωρίζει τα εναπομείναντα σύμβολα σε περιοχές που βασίζονται στα σύμβολα της βασικής γραμμής (**Zanibbi et al. 2001, 2002**). Η πληροφορία που προκύπτει από την περιοχή δράσης των τελεστών χρησιμοποιείται ώστε να μην χρειάζεται τα σύμβολα να είναι τοποθετημένα σε αυστηρά οριζόντια διάταξη. Η μέθοδος αυτή έχει χρησιμοποιηθεί σε μία πληθώρα pen-based συστημάτων αναγνώρισης ME (**Tapia and Rojas 2004, Anthony et al. 2007, Smirnova and Watt 2008**). Βελτιώσεις αυτής της μεθόδου αναφέρονται στο (**Tapia and Rojas 2004**), χρησιμοποιώντας πολλαπλά σημεία στο πλαίσιο οριοθέτησης (bounding box) του συμβόλου για τον προσδιορισμό των τοπολογικών σχέσεων μεταξύ δύο συμβόλων και χρησιμοποιώντας την έννοια του ελάχιστου εκτεταμένου δένδρου (minimum spanning tree) στο στάδιο του διαχωρισμού των συμβόλων. Τέλος στο (**Zhang et al. 2005**) τεχνικές βασισμένες σε ασαφή λογική (fuzzy logic) χρησιμοποιούνται για την διάκριση αμφισημιών των τοπολογικών σχέσεων (π.χ. εκθέτης, δείκτης).

Στην τέταρτη κατηγορία ανήκουν μέθοδοι που ακολουθούν μία πιο σφαιρική προσέγγιση για την τοπολογική ανάλυση της ΜΕ, βασισμένοι σε τεχνικές ελαχιστοποίησης του κόστους ενός γράφου. Στην προσέγγιση αυτή προτού οριστεί ένα κριτήριο ελαχιστοποίησης κόστους, ορίζονται οι υποψήφιες σχέσεις μεταξύ των συμβόλων. Οι **Eto, Suzuki et al.** χρησιμοποιούν δίκτυα τύπου Virtual Link Networks για την αναπαράσταση του κόστους των σχέσεων μεταξύ των συμβόλων. Έπειτα εξάγεται από τον γράφο, το ελάχιστο εκτεταμένο δένδρο (minimum spanning tree) και υπολογίζονται οι σχέσεις μεταξύ των συμβόλων (**Eto and Suzuki 2001**). Ορίζονται πέντε τοπολογικές σχέσεις μεταξύ δύο συμβόλων: (1) πάνω, (2) κάτω, (3) σε ευθεία, (4) εκθέτης και (5) δείκτης. Οι υποψήφιες σχέσεις και τα αντίστοιχα κόσθη βασίζονται στα πλαίσια οριοθέτησης (bounding box) των συμβόλων και στο κεντρικό σημείο των πλαισίων οριοθέτησης (**Eto and Suzuki 2001, Aly et al. 2008**). Η συγκεκριμένη τεχνική εφαρμόστηκε σε 158,308 συνεχόμενα ζεύγη συμβόλων της βάσης δεδομένων και μπόρεσε να υπολογίσει σωστά την τοπολογική σχέση για 99.57% από τα ζεύγη συμβόλων.

Στην πέμπτη κατηγορία ανήκουν αλγόριθμοι που χρησιμοποιούν γραμματικές τύπου Context Free Grammars (CFG) (**Chou 1989, Yamamoto et al. 2006, Pruša and Hlavác 2007, MacLean and Labahn 2013, Álvaro et al. 2014**). Οι γραμματικοί κανόνες που περιγράφουν την σύνταξη των μαθηματικών εκφράσεων μπορούν να διαβαστούν με κατάλληλους αλγορίθμους όπως είναι ο αλγόριθμος Cocke–Younger–Kasami (CYK). Στην εργασία τους οι **Yamamoto et al. (2006)** χρησιμοποιούν ένα μοντέλο βασισμένο σε στοχαστικές γραμματικές τύπου CFG, τις λεγόμενες SCFG (stochastic CFG), για την μοντελοποίηση των μαθηματικών εκφράσεων και προτείνουν ένα αλγόριθμο βασισμένο στον CYK για την εύρεση της πιο πιθανής έκφρασης ξεκινώντας από την χρονική σειρά των κονδυλίων της ΜΕ. Η εύρεση των τοπολογικών σχέσεων μεταξύ των συμβόλων βασίζεται σε πιθανότητες που προκύπτουν από ευριστικούς κανόνες. Οι **Pruša και Hlavác (2007)** ορίζουν μία δισδιάστατη γραμματική χρησιμοποιώντας μία συνάρτηση κόστους για να εκφράσουν την αξιοπιστία της αναγνωρισμένης μαθηματικής έκφρασης. Ο αλγόριθμος τους βασίζεται στην χρονική ακολουθία των κονδυλίων των συμβόλων που απαρτίζουν την ΜΕ. Οι **MacLean και Labahn (2013)** προτείνουν για την χρήση σχεσιακών γραμματικών σε συνδυασμό με fuzzy σύνολα για την αναγνώριση των ΜΕ. Ο αλγόριθμός τους βασίζεται στην κατάτμηση της ΜΕ σε επιμέρους περιοχές. Επιπλέον χρησιμοποιούν έναν μηχανισμό επαλήθευσης της ΜΕ από τον χρήστη με σκοπό την διόρθωση τυχόν λαθών κατά την αναγνώριση. Οι **Álvaro et al. (2014)** προτείνουν μία 2D επέκταση του αλγορίθμου SCFG και του CYK για την αναγνώριση των ΜΕ.

5.3. Προτεινόμενες τεχνικές

Στο στάδιο της τοπολογικής ανάλυσης αναπτύχθηκαν δύο τεχνικές. Στην πρώτη τεχνική χρησιμοποιήθηκε ο γεωμετρικός ταξινομητής ILSP-1 για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ΜΕ και την τοποθέτηση αυτών στα αντίστοιχα

επίπεδα της ΜΕ. Τα σύμβολα που ανήκουν σε διαφορετικά επίπεδα ενώνονται με σύμβολα άλλων επιπέδων βάση της τοπολογικής σχέσης που υπάρχει μεταξύ τους. Τέλος, αφού όλα τα σύμβολα έχουν τοποθετηθεί στα αντίστοιχα επίπεδα και οι σχέσεις μεταξύ των συμβόλων έχουν οριστεί, παράγεται ένα MathML αρχείο (Παράρτημα Β) για την αναπαράσταση της ΜΕ.

Στην δεύτερη τεχνική χρησιμοποιήθηκε ο ταξινομητής SVM για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ΜΕ. Παράλληλα αναπτύχθηκε μία γραμματική τύπου SCFG για την περιγραφή της σύνταξης των ΜΕ. Αφού ορίσουμε την γραμματική, μπορούμε πλέον να ψάξουμε για την πιο πιθανή υποψήφια μαθηματική έκφραση λαμβάνοντας υπόψη τις τοπολογικές σχέσεις μεταξύ των συμβόλων και/ή των υπό-εκφράσεων. Για να αναλύσουμε αποτελεσματικότερα μιας ΜΕ, πρέπει πρώτα να ορίσουμε ένα αλγόριθμο ανάλυσης (parsing algorithm) της ΜΕ. Σε αυτή την εργασία χρησιμοποιήσαμε τον αλγόριθμο Cocke–Younger–Kasami (CYK) κατάλληλα τροποποιημένο ώστε να μπορεί να αντιμετωπίσει το πρόβλημα της ανάλυσης μαθηματικών εκφράσεων.

5.3.1. Σταδιακή ανάλυση ΜΕ με την τοποθέτηση των συμβόλων σε επίπεδα

Στην τεχνική αυτή χρησιμοποιήθηκε ο γεωμετρικός ταξινομητής ILSP-1 (βλ. Κεφάλαιο 4.3.1.) για την εύρεση της τοπολογικής σχέσης μεταξύ των συμβόλων που απαρτίζουν την ΜΕ.

Αρχικά προσπαθούμε να βρούμε την βασική γραμμή (επίπεδο_0) της ΜΕ. Για το λόγο αυτό πρέπει να εντοπίσουμε το σύμβολο εκκίνησης της ΜΕ (start symbol). Για παράδειγμα στην ΜΕ που φαίνεται στην Εικόνα 28α, σύμβολο εκκίνησης είναι το σύμβολο 'c'. Δεδομένου του συμβόλου εκκίνησης, η τοπολογική σχέση 'δεξιά' εξετάζεται μεταξύ του κεντροειδούς του συμβόλου εκκίνησης και των κεντροειδών των λοιπών συμβόλων της ΜΕ. Με αυτό τον τρόπο βρίσκουμε όλα τα σύμβολα που ορίζουν την βασική γραμμή (level_0) της ΜΕ. Η διαδικασία αυτή επαναλαμβάνεται για όλα τα σύμβολα της ΜΕ. Τα σύμβολα που δεν ανήκουν στην βασική γραμμή, τοποθετούνται σε επίπεδα γύρω από την βασική γραμμή, ανάλογα με την απόστασή τους από τα 'κοντινότερα' σύμβολα της βασικής γραμμής. Με αυτόν τον τρόπο ορίζονται όλα τα επίπεδα της ΜΕ, τα οποία και ταξινομούνται σε σχέση με την βασική γραμμή. Κατόπιν αριθμούμε τα επίπεδα χρησιμοποιώντας θετικό πρόσημο για εκείνα που βρίσκονται πάνω από την βασική γραμμή και αρνητικό πρόσημο για εκείνα που βρίσκονται κάτω από την βασική γραμμή (Εικόνα 28δ).

Ακολούθως εισάγουμε την έννοια του συνδέσμου (connector) μεταξύ των συμβόλων της ΜΕ. Ας υποθέσουμε ότι έχουμε ένα σύμβολο που ανήκει στο επίπεδο_i. Ο δεσμός του συμβόλου αυτού ορίζεται βάση του κοντινότερου σε απόσταση συμβόλου που ανήκει σε κάποιο επίπεδο_j, όπου $i \neq j$, και $j > 0$ εάν $i > 0$ ή $j < 0$ εάν $i < 0$. Η κατηγορία του δεσμού μεταξύ δύο συμβόλων εξαρτάται από την τοπολογική σχέση που διέπει τα δύο αυτά σύμβολα. Έτσι έχουμε τις ακόλουθες κατηγορίες δεσμών σε σχέση με την αντίστοιχη τοπολογική σχέση των

συμβόλων: α) πάνω-δεξιά, β) δεξιά, γ) κάτω-δεξιά, δ) πάνω, ε) κάτω και στ) μέσα. Κατά συνέπεια τα σύμβολα της ΜΕ που ανήκουν στην βασική γραμμή συνδέονται μεταξύ τους με δεσμό κατηγορίας 'δεξιά' (Εικόνα 28ε). Ακολούθως βρίσκουμε τους δεσμούς όλων των συμβόλων της ΜΕ. Με αυτό τον τρόπο, όλες οι δυνατές σχέσεις μεταξύ συμβόλων έχουν ορισθεί και όλα τα σύμβολα έχουν τοποθετηθεί σε υπό-εκφράσεις των επιπέδων της ΜΕ. Η συγκεκριμένη μέθοδος αναγνωρίζει ΜΕ μέχρι ενός επιπέδου εκθέτη ή/και δείκτη.

$$c^2 = \frac{1}{2} d^2$$

(α)

(β)

{1} : c_1 {2} : 2_1

{3, 4} : =_1 {5} : 1_1

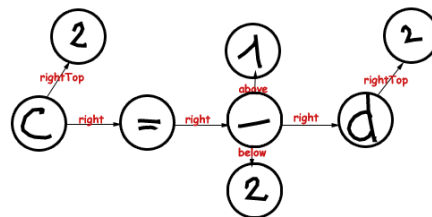
{6} : -_1 {7} : 2_2

{8, 9} : d_1 {10} : 2_3

(γ)



(δ)



(ε)

Εικόνα 28. Στάδια αναγνώρισης της ΜΕ. (α) οι κονδυλιές της ΜΕ αριθμημένες χρονικά, (β) ομαδοποίηση των κονδυλιών σε σύμβολα, (γ) αναγνώριση των συμβόλων, (δ) τοποθέτηση των συμβόλων σε επίπεδα, (ε) σχεσιακό δέντρο συμβόλων της ΜΕ.

Για παράδειγμα, ας υποθέσουμε την ΜΕ της Εικόνας 28α. Όπως περιγράψαμε, πρώτα ορίζουμε το επίπεδο_0 και μετά όλα τα άλλα επίπεδα της ΜΕ. Έτσι χωρίζουμε σε επίπεδα τα σύμβολα της ΜΕ (Εικόνα 28δ). Στην ΜΕ που φαίνεται στην Εικόνα 28γ, η έξοδος του αναγνωριστή συμβόλων είναι {c_1, 2_1, =_1, 1_1, -_1, 2_2, d_1, 2_3}. Η ακολουθία αυτή αποτυπώνει το σύμβολο που αναγνωρίστηκε και την απαρίθμηση του συγκεκριμένου συμβόλου στην ΜΕ, π.χ. το στοιχείο 2_1 αντιστοιχεί στο πρώτο χρονικά γραμμένο σύμβολο 2,

το στοιχείο 2_2 αντιστοιχεί στο δεύτερο χρονικά γραμμένο σύμβολο 2 κ.λπ. Το σύμβολο εκκίνησης της ME είναι το σύμβολο c_1, δηλ. στο πρώτο χρονικά γραμμένο σύμβολο c. Ακολούθως ψάχνουμε για εκείνα τα σύμβολα που έχουν σχέση 'δεξιά' με το σύμβολο εκκίνησης. Προκύπτει ότι η βασική γραμμή (επίπεδο_0) αποτελείται από τα ακόλουθα σύμβολα επίπεδο_0={c_1, =_1, -_1, d_1}. Έπειτα ελέγχουμε τα σύμβολα εκείνα που δεν έχουν ανατεθεί ακόμα σε κάποιο επίπεδο, ξεκινώντας κάθε φορά από το επόμενο σύμβολο της χρονικά ταξινομημένης λίστας συμβόλων. Έτσι προκύπτουν τα ακόλουθα επίπεδα: επίπεδο_+1={1_1}, επίπεδο_+2={2_1, 2_3} και επίπεδο_-1={2_2}.

Στο επόμενο βήμα ψάχνουμε για δεσμούς μεταξύ των επιπέδων που μόλις ορίσαμε (επίπεδο_+1, επίπεδο_+2, επίπεδο_-1), ώστε να ορίσουμε όλες τις δυνατές υπό-εκφράσεις (Εικόνα 28ε). Έτσι για το συγκεκριμένο παράδειγμα (Εικόνα 28δ), βρίσκουμε στο επίπεδο_+1 το σύμβολο {1_1} το οποίο είναι συνδεδεμένο με το σύμβολο {-_1} του επιπέδου_0 με σχέση 'πάνω', το σύμβολο {2_2} του επιπέδου_-1 που είναι συνδεδεμένο με το ίδιο σύμβολο {-_1} με σχέση 'κάτω', το σύμβολο {2_1} του επιπέδου_+2 που είναι συνδεδεμένο με σχέση πάνω-δεξιά με το σύμβολο {c_1} του επιπέδου_0 και το σύμβολο {2_3} του επιπέδου_+2 που είναι συνδεδεμένο με σχέση πάνω-δεξιά με το σύμβολο {d_1} του επιπέδου_0. Αυτή η πληροφορία αποτυπώνεται στο αρχείο MathML της ME (Εικόνα 28ε).

```

<annotationXML type="truth" encoding="Content-MathML">
  <math xmlns='http://www.w3.org/1998/Math/MathML'>
    <math display="block">
      \begin{array}{c}
        \begin{array}{c}
          c_1^2 \\
          = \\
          \frac{1}{2} \\
          \begin{array}{c}
            d_1^2 \\
            \frac{2}{3}
          \end{array}
        \end{array}
      \end{array}
    \end{math}
  </math>
</annotationXML>

```

Εικόνα 29. Δομή MathML της ME που φαίνεται στην Εικόνα 28α.

Αξίζει να σημειώσουμε ότι ο διαχωρισμός μεταξύ του συμβόλου της γραμμής κλάσματος και του συμβόλου της αφαίρεσης, πραγματοποιείται κατά την φάση της τοπολογικής ανάλυσης

της ME και όχι κατά την φάση της αναγνώρισης συμβόλων. Ο διαχωρισμός ακολουθεί τον παρακάτω απλό αλγόριθμο: Εάν το σύμβολο '·' έχει σχέση τύπου 'πάνω' και 'κάτω' τότε το σύμβολο αυτό είναι η γραμμή κλάσματος. Σε διαφορετική περίπτωση πρόκειται για το σύμβολο της αφαίρεσης.

Τέλος, το αποτέλεσμα της τοπολογικής ανάλυσης της ME οδηγείται σε έναν parser ώστε να παραχθεί το αντίστοιχο MathML αρχείο που περιγράφει την ME (Εικόνα 29).

5.3.2. Σταδιακή ανάλυση με χρήση στοχαστικών γραμματικών

Για να αντιμετωπίσουμε το πρόβλημα της αναγνώρισης των χειρόγραφων ME αναπτύξαμε μια γραμματική που περιγράφει την σύνταξη των ME που πρόκειται να αναγνωρίσουμε και στην συνέχεια αναπτύξαμε έναν αλγόριθμο ανάλυσης της ME.

Μία δισδιάστατη στοχαστική γραμματική χωρίς συμφραζόμενα SCF συμβολίζεται

$$G=(N,T,R,S,P) \quad (2)$$

όπου, N είναι ένα πεπερασμένο σύνολο από μη τερματικά σύμβολα, T είναι ένα πεπερασμένο σύνολο από τερματικά σύμβολα, R είναι ένα πεπερασμένο σύνολο από κανόνες παραγωγής, S είναι ένα αρχικό σύμβολο και P είναι η πιθανότητα που αντιστοιχεί στους κανόνες παραγωγής.

Οι κανόνες παραγωγής των τερματικών και μη-τερματικών συμβόλων είναι της μορφής Chomsky normal form (CNF) :

$$A \rightarrow t_A \quad (3)$$

$$C \xrightarrow{rel} AB \quad (4)$$

Η Εξίσωση 3 αντιπροσωπεύει την γενική μορφή ενός κανόνα παραγωγής τερματικού συμβόλου και η Εξίσωση 4 αντιπροσωπεύει την γενική μορφή ενός κανόνα παραγωγής μη τερματικού συμβόλου. Συγκεκριμένα η Εξίσωση 3 υποδηλώνει ότι το τερματικό σύμβολο t_A μπορεί να παραχθεί από το μη-τερματικό σύμβολο A χωρίς την ύπαρξη κάποιας τοπολογικής σχέσης. Ο μη-τερματικός κανόνας στην Εξίσωση 4 υποδηλώνει ότι το μη-τερματικό σύμβολο B έχει τοπολογική σχέση *rel* με το σύμβολο αναφοράς A, και ο συνδυασμός τους παράγει το μη-τερματικό σύμβολο C. Η πιθανότητα παραγωγής του μη-τερματικού συμβόλου C δίνεται από την ακόλουθη εξίσωση:

$$P(C)=P(C \xrightarrow{rel} AB) \cdot P(A) \cdot P(B) \quad (5)$$

Ας υποθέσουμε την ύπαρξη της ακόλουθης γραμματικής G , ώστε να παρουσιάσουμε τον αλγόριθμο αναγνώρισης μαθηματικών εκφράσεων:

$$N = \{ \text{Sym}, \text{Exp}, \text{OprExp}, \text{Opr} \} \quad (6)$$

$$T = \{ \text{a-z}, \text{0-9}, \text{+}, \text{-} \} \quad (7)$$

$$R = \{ R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8 \} \quad (8)$$

$$R_1 : \text{Sym} \rightarrow \{ \text{0-9}, \text{a-z} \} \quad (9)$$

$$R_2 : \text{Opr} \rightarrow \{ \text{+}, \text{-} \} \quad (10)$$

$$R_3 : \text{OprExp} \xrightarrow{\text{right}} \text{Opr Sym} \quad (11)$$

$$R_4 : \text{Exp} \xrightarrow{\text{r-Above}} \text{Sym Sym} \quad (12)$$

$$R_5 : \text{Exp} \xrightarrow{\text{right}} \text{Exp OprExp} \quad (13)$$

$$R_6 : \text{Exp} \xrightarrow{\text{right}} \text{Sym Sym} \quad (14)$$

$$R_7 : \text{Exp} \xrightarrow{\text{right}} \text{Sym Exp} \quad (15)$$

$$R_8 : \text{Exp} \xrightarrow{\text{right}} \text{Sym OprExp} \quad (16)$$

Η παραπάνω γραμματική G μπορεί να παράγει αλγεβρικές εκφράσεις που περιέχουν εκθέτες, όπως π.χ. η $3y^2 + 3y + 1$, εφαρμόζοντας επαναληπτικά τους κανόνες παραγωγής και έχοντας σαν αρχικό μη-τερματικό σύμβολο το σύμβολο Exp .

Στην παρούσα εργασία χρησιμοποιούμε μία δισδιάστατη στοχαστική γραμματική χωρίς συμπραζόμενα (SCFG) για την παραγωγή των γραμματικών κανόνων των μαθηματικών εκφράσεων. Για να αντιμετωπίσουμε την σύνταξη των ΜΕ που περιέχονται στα τέσσερα διαφορετικά μέρη των βάσεων CROHME 2012 και 2013, δημιουργήσαμε τρεις διαφορετικές γραμματικές, Gram-I , Gram-II και Gram-III (βλ. Παράρτημα Ε). Οι γραμματικές αυτές περιέχουν 38, 59 και 64 γραμματικούς κανόνες αντίστοιχα. Κάθε γραμματική είναι υποσύνολο της άλλης, δηλαδή $\text{Gram-III} \supset \text{Gram-II} \supset \text{Gram-I}$. Να σημειωθεί ότι για το Μέρος-IV χρησιμοποιούμε την ίδια γραμματική με αυτή του Μέρους-III.

Αφού ορίσουμε την γραμματική, μπορούμε πλέον να ψάξουμε για την πιο πιθανή υποψήφια μαθηματική έκφραση λαμβάνοντας υπόψη τις τοπολογικές σχέσεις μεταξύ των συμβόλων και/ή των υπό-εκφράσεων. Για να αναλύσουμε αποτελεσματικότερα μια ΜΕ, πρέπει πρώτα να ορίσουμε ένα αλγόριθμο ανάλυσης (parsing algorithm) της ΜΕ. Σε αυτή την εργασία χρησιμοποιήσαμε τον αλγόριθμο Cocke–Younger–Kasami (CYK) κατάλληλα τροποποιημένο ώστε να μπορεί να αντιμετωπίσει το πρόβλημα της ανάλυσης μαθηματικών εκφράσεων. Η πολυπλοκότητα του αλγορίθμου αυτού είναι της τάξης $O(n^3 |G|)$, όπου n το πλήθος των συμβόλων της ΜΕ και $|G|$ το πλήθος των γραμματικών κανόνων που περιγράφουν την γραμματική G . Ο αλγόριθμος υπολογίζει έναν κάτω διαγώνιο πίνακα T , διάστασης $n \times n$, όπου η πιο πιθανή υποψήφια ΜΕ βρίσκεται στο κελί $T(n,1)$. Η αρίθμηση των γραμμών του πίνακα ξεκινά από την βάση του πίνακα.

Σε κάθε κελί του πίνακα T είναι αποθηκευμένες 3 τιμές: (i) η απεικόνιση της ΜΕ σε μορφότυπο LaTeX/MathML, (ii) η πιθανότητα του παραγόμενου αντικειμένου/αποτελέσματος και (iii) η ετικέτα του παραγόμενου αντικειμένου/αποτελέσματος. Για χάρη απλότητας, στον Πίνακα 14 παρουσιάζεται μόνο η Latex απεικόνιση της ΜΕ.

8	$3y^2+3y+1$ 0.7402 (0.9562) R_7 , Exp	$3y^2+3y+1$ 0.8650 (0.9697) R_5 , Exp							
7						
6						
5	$3y^2+3$ 0.5996 (0.9510) R_7 , Exp	$3y^2+3$ 0.9235 (0.9510) R_7 , Exp	y^2+3y 0.8011 (0.9250) R_7 , Exp	y^2+3y 0.9563 (0.9893) R_5 , Exp		$+ 3y+ 1$ 0.8498 (0.9841) R_9 , OprExp			
4			y^2+3 0.6305 (0.7352) R_7 , Exp	y^2+3 0.9710 (0.9868) R_5 , Exp	$2+3y$ 0.8660 (0.8951) R_8 , Exp		$3y+1$ 0.8635 (0.8848) R_6 , Exp		
3	$3y^2$ 0.9453 (0.9463) R_7 , Exp				$2+3$ 0.8576 (0.8707) R_8 , Exp	$+3y$ 0.9675 (0.9892) R_9 , OprExp		$y+1$ 0.8708 (0.8728) R_8 , Exp	
2	$3y$ 0.9696 (0.9696) R_6 , Exp	y^2 0.9990 (0.9990) R_4 , Exp				$+3$ 0.9849 (0.9849) R_3 , OprExp	$3y$ 0.9781 (0.9781) R_6 , Exp		$+1$ 0.9978 (0.9978) R_3 , OprExp
1	3 1 R_1 , Sym	y 1 R_1 , Sym	2 1 R_1 , Sym	+	1 R_2 , Opr	3 1 R_1 , Sym	y 1 R_1 , Sym	+	1 1 R_2 , Opr
	1	2	3	4	5	6	7	8	

Πίνακας 14. CYK πίνακας της μαθηματικής έκφρασης $3y^2+3y+1$. Σε κάθε κελί του πίνακα T είναι αποθηκευμένες 3 τιμές: (i) η απεικόνιση της ΜΕ σε μορφότυπο LaTeX/MathML, (ii) η πιθανότητα του παραγόμενου αντικειμένου και (iii) η ετικέτα του παραγόμενου αντικειμένου.

Κατά την αρχικοποίηση, τα κελιά της πρώτης γραμμής του CYK πίνακα T γειμίζουν με τις κατάλληλες τιμές. Στην δικής μας προσέγγιση χρησιμοποιούμε τα αναγνωρισμένα σύμβολα που απαρτίζουν την ME σύμφωνα με την χρονολογική σειρά εμφάνισής τους.

Για κάθε σύμβολο αναθέτουμε τιμή πιθανότητας ίση με 1, αφού προϋποθέτουμε ότι όλα τα σύμβολα που απαρτίζουν την ME έχουν αναγνωριστεί σωστά. Οι πιθανότητες αυτές θα ήταν ίσες με την πιθανότητα του αναγνωρισμένου συμβόλου, εάν κατά την διάρκεια της τοπολογικής ανάλυσης κάναμε ταυτόχρονη αναγνώριση των συμβόλων.

Για να υπολογίσουμε τις τιμές των υπόλοιπων κελιών του πίνακα T, ακολουθούμε τα παρακάτω βήματα και κάθε $T(row,col)$ κελί, λαμβάνοντας υπόψη όλους τους δυνατούς συνδυασμούς των ζευγών αντικειμένων που βρίσκονται στα κελιά $T(row-m,col)$ και $T(m,row-m+col)$, για $m = 1 \dots row-1$, ξεκινώντας από $row=2$:

I. Αξιολόγηση των πιθανοτήτων των τοπολογικών σχέσεων χρησιμοποιώντας τον πιθανοτικό ταξινομητή SVM

$$P(EXP_{row,col} \xrightarrow{rel} EXP_{row-m,col} EXP_{m,col-m+row})$$

II. Διατήρησε μόνο τις τοπολογικές σχέσεις για τις οποίες η πιθανότητα είναι μεγαλύτερη από την τιμή κατωφλίσωσης th

$$P(EXP_{row,col} \xrightarrow{rel} EXP_{row-m,col} EXP_{m,col-m+row}) > th$$

III. Εφαρμογή των κανόνων της γραμματικής G ώστε να διατηρήσουμε μόνο τις ME που είναι έγκυρες $\{EXP_{row,col}^G \xrightarrow{rel} EXP_{row-m,col}^G EXP_{m,col-m+row}^G\}$

IV. Αξιολόγηση των αντίστοιχων αθροιστικών πιθανοτήτων βάση της Εξ. (4)

$$P(EXP_{row,col}^G) = P(EXP_{row,col}^G \xrightarrow{rel} EXP_{row-m,col}^G EXP_{m,col-m+row}^G) \cdot P(EXP_{row-m,col}^G) \cdot P(EXP_{m,col-m+row}^G)$$

Ας υπολογίσουμε για παράδειγμα τις τιμές για το κελί $T(2,1)$ για την ανάλυση της ME $3y^2 + 3y + 1$. Υπολογίζουμε πρώτα τις πιθανότητες που αντιστοιχούν στις πιθανές τοπολογικές σχέσεις μεταξύ των συμβόλων «3» και «y» που βρίσκονται στα κελιά $T(1,1)$ και $T(1,2)$. Ο πιθανοτικός SVM ταξινομητής δίνει τις ακόλουθες τιμές πιθανότητας:

Rel	δεξιά	πάνω-δεξιά	κάτω-δεξιά	μέσα	κάτω
P	0.9696	0.0153	0.0086	0.0063	2.9525e-04

Εφαρμόζοντας την τιμή κατωφλίσωσης th στα παραπάνω αποτελέσματα, συμπεραίνουμε ότι μόνο μία από τις υποψήφιες τοπολογικές σχέσεις ικανοποιεί αυτό τον κανόνα. Η τοπολογική

σχέση που ικανοποιεί αυτόν τον κανόνα είναι η σχέση «δεξιά» με τιμή πιθανότητας ίση με 0.9696. Η τιμή αυτή αποθηκεύεται στο κελί T(2,1). Όπως φαίνεται παρακάτω (Πίνακας 15) για τιμή κατωφλίσωσης $th=0.03$, παίρνουμε τα καλύτερα αποτελέσματα όσον αφορά την μετρική *Exp_Rec*. Πρέπει να σημειωθεί ότι οι πιθανότητες των τοπολογικών σχέσεων εμφανίζονται σε κάθε κελί του πίνακα T μέσα σε παρενθέσεις, δίπλα στις πιθανότητες του παραγόμενου αντικειμένου.

Στο επόμενο βήμα, διερευνούμε αν υπάρχει κάποιος κατάλληλος κανόνας R που να παράγει αυτή την σχέση. Αν υπάρχει κάποιος τέτοιος κανόνας, τότε το αντίστοιχο αποτέλεσμα αποθηκεύεται στο κελί T(row,col).

Για να ελαχιστοποιήσουμε περαιτέρω την πολυπλοκότητα του αλγορίθμου, διαγράφουμε πιθανές επαναλήψεις των υποψήφιων εκφράσεων. Οι επαναλήψεις αυτές δημιουργούνται όταν υπάρχουν περισσότεροι από έναν τρόποι για την παραγωγή της ίδιας υπό-έκφρασης. Για παράδειγμα, στο κελί T(4,5) του Πίνακα 14, έχει παραχθεί δύο φορές η ίδια υπό-έκφραση « $3y+1$ » με τιμές πιθανότητας 0.8653 και 0.8545 αντίστοιχα. Κατά τον ίδιο τρόπο, συνδυάζοντας τα κελιά T(1,5) και T(3,6) μπορεί να παραχθεί η ίδια υπο-έκφραση αλλά με μικρότερη τιμή πιθανότητας 0.8545. Έτσι ελαχιστοποιούμε τον αριθμό των υποψήφιων εκφράσεων, κρατώντας τις εκφράσεις με την μεγαλύτερη πιθανότητα (στο συγκεκριμένο παράδειγμα το αποτέλεσμα του συνδυασμού των κελιών T(2,5) και T(2,7)). Εφαρμόζοντας αυτή την διαδικασία επαναληπτικά, καταλήγουμε στις τιμές που φαίνονται στον Πίνακα 14. Σημειώστε ότι τα κενά κελιά του πίνακα T οφείλονται στους γραμματικούς κανόνες. Για παράδειγμα το κελί T(2,6) είναι κενό γιατί στην συγκεκριμένη γραμματική δεν υπάρχει ο κανόνας που συνδυάζει ένα σύμβολο τύπου "Sym" με ένα σύμβολο τύπου "Or".

Στο τελευταίο κελί του πίνακα CYK, δηλαδή στο κελί T(n,1), αποθηκεύονται όλες οι πιθανές ME που παράγονται με την αντίστοιχη γραμματική. Ταξινομούμε αυτές τις εκφράσεις σύμφωνα με την τιμή της αθροιστικής τους πιθανότητας. Η πιο πιθανή έκφραση είναι εκείνη που έχει την υψηλότερη πιθανότητα. Στο συγκεκριμένο παράδειγμα συμπεραίνουμε από τον Πίνακα 14 ότι η πιο πιθανή ME βρίσκεται στο κελί T(8,1) και είναι η $3y^2 + 3y + 1$ με πιθανότητα $P=0.8650$.

Όπως αναφέρθηκε παραπάνω, η πολυπλοκότητα του CYK αλγορίθμου είναι της τάξης $O(n^3 |G|)$, όπου n το πλήθος των συμβόλων της ME και |G| το πλήθος των γραμματικών κανόνων που περιγράφουν την γραμματική G. Η μείωση της πολυπλοκότητας του αλγορίθμου, πραγματοποιείται σε δύο βήματα. Για την ρύθμιση των παραμέτρων των δύο αυτών βημάτων, διαλέξαμε από κάθε μέρος της βάσης CROHME ένα τυχαίο δείγμα αποτελούμενο από το 20% των ME σαν σύνολο επικύρωσης (validation set).

th	Μέρος-I	Μέρος -II	Μέρος -III	Μέρος -IV
0.01	78.70%	65.78%	56.37%	50.22%
0.02	78.70%	65.78%	56.37%	50.22%
0.03	78.70%	65.78%	56.37%	50.22%
0.04	78.70%	65.12%	55.97%	50.07%
0.05	78.70%	65.12%	55.97%	49.93%
0.06	78.70%	64.45%	55.56%	49.78%
0.07	78.70%	64.45%	55.56%	49.63%
0.08	78.70%	64.12%	55.14%	49.63%
0.09	78.70%	64.12%	54.94%	49.48%
0.10	78.70%	64.12%	54.94%	49.48%

Πίνακας 15. Αποτελέσματα της μετρικής EXP_Rec results στο σύνολο αξιολόγησης του CROHME 2012 και 2013 για διαφορετικές τιμές του κατωφλίου *th*.

Στο πρώτο βήμα, ταξινομούνται οι πιθανότητες των τοπολογικών σχέσεων μεταξύ των αντικειμένων κατά φθίνουσα σειρά και κρατάμε τις τρεις πρώτες τοπολογικές σχέσεις των οποίων η πιθανότητα είναι μεγαλύτερη από μία ορισμένη τιμή κατωφλίου, $P > th$.

Έχουμε ορίσει πειραματικά την τιμή κατωφλίου ως $th=0.03$, έτσι ώστε σε όλες τις πιθανές περιπτώσεις να μεταφέρεται η σωστή τοπολογική σχέση στο επόμενο επίπεδο του αλγορίθμου CYK. Εξετάσαμε το ποσοστό ακρίβειας (accuracy rate) της μετρικής EXP_Rec για διάφορες τιμές κατωφλίου *th* από 0.01 μέχρι 0.10, με βήμα 0.01. Παρουσιάζουμε τα αντίστοιχα αποτελέσματα στον Πίνακα 15. Τα πειραματικά αποτελέσματα δείχνουν ότι, αυξάνοντας την τιμή κατωφλίου *th*, η ακρίβεια της μετρικής EXP_Rec μειώνεται. Να σημειωθεί ότι με την τιμή κατωφλίου ίση με $th=0.03$, επιτυγχάνεται το μέγιστο ποσοστό ακρίβεια για την μετρική EXP_Rec.

Στο δεύτερο βήμα, ταξινομούμε τις αθροιστικές πιθανότητες των παραγόμενων ΜΕ σε φθίνουσα σειρά και κρατάμε τις 10 πρώτες. Η βέλτιστη τιμή του πλήθους των επιτρεπόμενων ΜΕ έχει υπολογιστεί πειραματικά, χρησιμοποιώντας ένα υποσύνολο ΜΕ των βάσεων CROHME 2012 and 2013, σαν σύνολο επικύρωσης (validation set). Στον Πίνακα 16 παρουσιάζονται τα αποτελέσματα ακρίβειας για την μετρική EXP_Rec, για διαφορετικό πλήθος επιτρεπόμενων ΜΕ.

		1	5	10
Gram-I	Μέρος-I	78.70%	78.70%	78.70%
Gram-II	Μέρος -II	56.15%	65.45%	65.78%
Gram-III	Μέρος -III	48.26%	55.72%	56.37%
Gram-III	Μέρος -IV	49.18%	50.07%	50.22%

Πίνακας 16. Αποτελέσματα για την μετρική EXP_Rec στο σύνολο επικύρωσης (validation set) των βάσεων CROHME 2012 και 2013 για διαφορετικό πλήθος επιτρεπόμενων ΜΕ.

Παρατηρώντας τον Πίνακα 16 συμπεραίνουμε ότι το πλήθος των επιτρεπόμενων ΜΕ πρέπει να αυξηθεί όσο αυξάνεται η πολυπλοκότητα της γραμματικής που περιγράφει τις ΜΕ. Αξίζει να σημειωθεί ότι χρησιμοποιώντας ΜΕ από το Μέρος-I επιτυγχάνεται το ίδιο ποσοστό ακρίβειας ανεξάρτητα από το πλήθος των επιτρεπόμενων ΜΕ, ενώ το μέγιστο ποσοστό ακρίβειας για τα

υπόλοιπα μέρη (Μέρος-II, Μέρος-III και Μέρος-IV) επιτυγχάνεται χρησιμοποιώντας τα 10 πιθανότερα σενάρια. Αυτό σημαίνει ότι το Μέρος-I περιέχει απλούστερες ΜΕ σε σχέση με τα υπόλοιπα μέρη, για τις οποίες μόνο η μεταφορά της πιο πιθανής ΜΕ είναι αρκετή ώστε να επιτευχθεί το βέλτιστο αποτέλεσμα.





5.4. Αξιολόγηση

Η αξιολόγηση των συστημάτων αναγνώρισης χειρόγραφων ΜΕ βασίζεται στις ακόλουθες τέσσερις μετρικές:

- i. STrec: ποσοστό ορθής αναγνώρισης των κονδυλιών,
- ii. SYMseg: ποσοστό ορθής ομαδοποίησης των συμβόλων,
- iii. SYMrec: ποσοστό ορθής αναγνώρισης των συμβόλων,
- iv. STRUCT: ποσοστό ορθής τοπολογικής αναγνώρισης της έκφρασης και
- v. EXPrec: ποσοστό ορθής αναγνώρισης της μαθηματικής έκφρασης.

Η μετρική STRUCT υπολογίζει το ποσοστό των εκφράσεων που έχουν το σωστό δέντρο τοπολογικής αναπαράστασης, ανεξάρτητα από το αν τα σύμβολα της έκφρασης έχουν αναγνωρισθεί σωστά. Επίσης, αξίζει να σημειωθεί ότι η μετρική SYMrec υπολογίζεται μόνο για εκείνα τα σύμβολα που έχουν ομαδοποιηθεί σωστά. Η τελική βαθμολογία του συστήματος βασίστηκε στην σωστή αναγνώριση της ΜΕ. Θα πρέπει να επισημάνουμε ότι η βαθμολογία για μία ΜΕ μπορεί να πάρει την τιμή 1 στην περίπτωση που είναι ακριβώς ίδια με τα πραγματικά δεδομένα (ground truth) της έκφρασης. Σε κάθε άλλη περίπτωση η βαθμολογία της ΜΕ ισούται με 0. Αυτό σημαίνει ότι αν υπάρχει έστω και ένα λάθος στην αναγνώριση κάποιου συμβόλου, ή στην ομαδοποίηση των κονδυλιών, ή εάν η αναγνώριση της ΜΕ είναι τμηματικά σωστή, τότε η βαθμολογία της συγκεκριμένης ΜΕ μηδενίζεται.

Για παράδειγμα, ας υποθέσουμε ότι θέλουμε να αναγνωρίσουμε την μαθηματική έκφραση $n = 2k$ η οποία αποτελείται από 4 σύμβολα και 6 κονδυλιές. Στον Πίνακα 17 φαίνεται η ορθή αποτύπωση της μαθηματικής έκφρασης σαν αποτέλεσμα της Α αξιολόγησης. Ακολούθως παρουσιάζονται άλλα 3 παραδείγματα λανθασμένης αναγνώρισης (Β, Γ και Δ). Το αποτέλεσμα της Β αξιολόγησης περιλαμβάνει ένα λάθος στην αναγνώριση συμβόλου. Το αποτέλεσμα της Γ αξιολόγησης περιλαμβάνει ένα λάθος στην κατάτμηση/ομαδοποίηση των κονδυλιών. Το αποτέλεσμα της Δ αξιολόγησης περιλαμβάνει ένα λάθος στην τοπολογία της μαθηματικής έκφρασης.

Αποτέλεσμα αξιολόγησης Α: $n = 2k$	Ground truth
	<pre> 1 <mrow> 2 <mi xml:id="n_1" style="">n</mi> 3 <mrow> 4 <mo xml:id="=_1" style="">=</mo> 5 <mrow> 6 <mn xml:id="2_1" style="">2</mn> 7 <mi xml:id="k_1" style="">k</mi> 8 </mrow> 9 </mrow> 10 </mrow> </pre>
Αποτέλεσμα αξιολόγησης Β: $u = 2k$	ένα λάθος στην αναγνώριση συμβόλου
	<pre> 1 <mrow> 2 <mi xml:id="n_1" style="">u</mi> 3 <mrow> 4 <mo xml:id="=_1" style="">=</mo> 5 <mrow> 6 <mn xml:id="2_1" style="">2</mn> 7 <mi xml:id="k_1" style="">k</mi> 8 </mrow> 9 </mrow> 10 </mrow> </pre>
Αποτέλεσμα αξιολόγησης Γ: $n = 2ru$	ένα λάθος στην κατάτμηση
	<pre> 1 <mrow> 2 <mi xml:id="n_1" style="">n</mi> 3 <mrow> 4 <mo xml:id="=_1" style="">=</mo> 5 <mrow> 6 <mn xml:id="2_1" style="">2</mn> 7 <mi xml:id="r_1" style="">r</mi> 8 <mi xml:id="u_1" style="">u</mi> 9 </mrow> 10 </mrow> 11 </mrow> </pre>
Αποτέλεσμα αξιολόγησης Δ: $n = 2_k$	ένα λάθος στην τοπολογία της έκφρασης
	<pre> 1 <mrow> 2 <mi xml:id="n_1" style="">n</mi> 3 <mrow> 4 <mo xml:id="=_1" style="">=</mo> 5 <msub> 6 <mn xml:id="2_1" style="">2</mn> 7 <mi xml:id="k_1" style="">k</mi> 8 </msub> 9 </mrow> 10 </mrow> </pre>

Πίνακας 17. Παραδείγματα αποτελέσματος αναγνώρισης για την μαθηματική έκφραση $n = 2k$, (Α) ορθή αναγνώριση, (Β) ένα λάθος στην αναγνώριση συμβόλου, (Γ) ένα λάθος στην κατάτμηση/ομαδοποίηση κονδυλίων και (Δ) ένα λάθος στην τοπολογία της έκφρασης.

Στον Πίνακα 18 παρουσιάζονται τα αποτελέσματα της αξιολόγησης για την μαθηματική έκφραση $n = 2k$ για τις 5 μετρικές STrec, SYMseg, SYMrec, STRUCT και EXPrec για τα 5 παραδείγματα αξιολόγησης που παρουσιάζονται στον Πίνακα 17.

	STrec	SYMseg	SYMrec	STRUCT	EXPrec
A: $n = 2k$	100,00% (6/6)	100,00% (4/4)	100,00% (4/4)	100,00%	100,00%
B: $u = 2k$	80,00% (5/6)	100,00% (4/4)	75,00% (3/4)	100,00%	0,00%
Γ: $n = 2ru$	60,00% (4/6)	75,00% (3/4)	75,00% (3/4)	0,00%	0,00%
Δ: $n = 2_k$	100,00% (6/6)	100,00% (4/4)	100,00% (4/4)	0,00%	0,00%
Συνολικά	87,50%(21/24)	93,75%(15/16)	93,33%(14/16)	50,00%(2/4)	25,00% (1/4)

Πίνακας 18. Αποτελέσματα της αξιολόγησης για την μαθηματική έκφραση $n = 2k$, για τις 5 μετρικές STrec, SYMseg, SYMrec, STRUCT και EXPrec.

Όπως φαίνεται στον Πίνακα 18 στο αποτέλεσμα B της αξιολόγησης 5 από τις 6 κονδυλίες της ME αναγνωρίστηκαν σωστά (STrec=80%), οι κονδυλίες όλων των συμβόλων της ME ομαδοποιήθηκαν σωστά (SYMseg=100%), 3 από τα 4 σύμβολα της ME αναγνωρίστηκαν σωστά (SYMrec=75%). Η τοπολογική αναγνώριση της έκφρασης ανεξάρτητα από την ορθή αναγνώριση των συμβόλων της είναι σωστή (STRUCT=100%), ενώ μόνο μία από τις 4 ME έχει αναγνωρισθεί σωστά (EXPrec=25%).

Για την αξιολόγηση του συστήματος, που υλοποιεί την σταδιακή ανάλυση της ME με την τοποθέτηση των συμβόλων σε επίπεδα, χρησιμοποιήθηκε η βάση CROHME 2011 (**Mouchère et al. 2011**). Στον Πίνακα 19 και 20 παρουσιάζονται τα αποτελέσματα της προτεινόμενης μεθόδου για το Μέρος-I και για το Μέρος-II της βάσης χειρόγραφων μαθηματικών εκφράσεων CHROME 2011. Το ποσοστό αναγνώρισης συμβόλου για το Μέρος-I είναι 81.76% και για το Μέρος-II είναι 88.90%. Τα αποτελέσματα αυτά είναι πολύ κοντά στο ποσοστό αναγνώρισης που αναφέρθηκε στο (**Simistira et al. 2008**).

Μπορούμε να συμπεράνουμε ότι τα λάθη στην αναγνώριση των συμβόλων στην πλειονότητά τους οφείλονται σε τρεις λόγους Πρώτον ο αλγόριθμος αναγνώρισης συμβόλων δεν χειρίζεται σύνθετα σύμβολα (π.χ. log) και δεν λαμβάνει υπόψη του την σχετική θέση των κονδυλιών του συμβόλου (π.χ. το σύμβολο της ανισότητας μπορεί να μπερδευτεί με το σύμβολο F που αποτελείται από δύο κονδυλίες Γ και -). Επιπλέον, υπάρχουν κλάσεις συμβόλων που παρουσιάζουν μεγάλη ομοιότητα, όπως το μικρό γράμμα c με το κεφαλαίο γράμμα C.

Συστήματα	ST_{rec}	SYM_{seg}	SYM_{rec}	EXP_{rec}
I	53.23	59.06	88.78	4.42
II	22.39	27.98	82.11	0.55
III	78.73	88.07	92.22	29.28
IV	37.41	55.15	81.71	0.00
Προτεινόμενο	75.90	86.67	81.76	9.94

Πίνακας 19. Αποτελέσματα για το Μέρος-I.

Συστήματα	ST_{rec}	SYM_{seg}	SYM_{rec}	EXP_{rec}
I	51.58	56.50	91.29	2.59
II	22.11	28.25	83.76	0.29
III	78.38	87.82	92.56	19.83
IV	52.28	78.77	78.67	0.00
προτεινόμενο	63.80	79.86	88.90	5.45

Πίνακας 20. Αποτελέσματα για το Μέρος-II.

Τα λάθη στην αναγνώριση της τοπολογίας των συμβόλων της ΜΕ οφείλονται στο ότι ο αλγόριθμος μπορεί να αναγνωρίσει εκφράσεις μέχρι ένα επίπεδο πάνω και ένα επίπεδο κάτω από την βασική γραμμή της ΜΕ. Κατά συνέπεια, αφού το Μέρος-II περιλαμβάνει πιο σύνθετες ΜΕ, το ποσοστό αναγνώρισης της ΜΕ φθίνει σε σχέση με το ποσοστό αναγνώρισης των μαθηματικών εκφράσεων που ανήκουν στο Μέρος-I. Επιπλέον η λάθος επιλογή του συμβόλου εκκίνησης της ΜΕ μπορεί να οδηγήσει σε λανθασμένη τοπολογική ανάλυσή της.

Για την αξιολόγηση της τεχνικής που υλοποιεί την σταδιακή ανάλυση της ΜΕ με χρήση στοχαστικών γραμματικών, αναπτύχθηκε μία δισδιάστατη στοχαστική γραμματική χωρίς συμπραζόμενα (2D-SCFG) με σκοπό την περιγραφή της σύνταξης των μαθηματικών εκφράσεων. Οι πιθανότητες των τοπολογικών σχέσεων μεταξύ συμβόλων της ΜΕ υπολογίζονται από ένα πιθανοτικό SVM ταξινομητή. Για να αξιολογηθεί η ακρίβεια της αναγνώρισης των ΜΕ, χρησιμοποιήθηκαν τα σύνολα εξέτασης των βάσεων CROHME2012 και CROHME2013.

Επτά ερευνητικές ομάδες πήραν μέρος στον διαγωνισμό του CROHME 2012. Το σύστημά μας, όπως αυτό περιγράφηκε στο **Simistira et al. (2012)**, πέτυχε ποσοστό ακρίβειας για την μετρική STRUCT ίση με 21.30%, 12.29%, 9.84% για τα Μέρος-I, Μέρος -II και Μέρος -III αντίστοιχα, ενώ τα αντίστοιχα ποσοστά ακρίβειας για την μετρική EXP_Rec είναι 8.33%, 6.64% και 3.69% αντίστοιχα.

Έξι ερευνητικές ομάδες πήρα μέρος στο διαγωνισμό CROHME 2013. Για το Μέρος-IV, δεν έχουμε προηγούμενα αποτελέσματα γιατί δεν είχαμε πάρει μέρος στο διαγωνισμό αυτό.

Στον Πίνακα 21 φαίνονται τα αποτελέσματα του προτεινόμενου αλγορίθμου για την αναγνώριση ΜΕ στα σύνολα εξέτασης των διαγωνισμών CROHME 2012 και CROHME 2013.

Μετρική	Μέρος-I	Μέρος -II	Μέρος -III	Μέρος -IV
EXP_Rec	78.70%	65.78%	56.37%	50.22%
STRUCT	78.70%	66.44%	57.41%	50.52%

Πίνακας 21. Αποτελέσματα για τις μετρικές EXP_Rec και STRUCT στα σύνολα εξέτασης των διαγωνισμών CROHME 2012 και 2013.

Τα τρία πρώτα συστήματα στο διαγωνισμό του CROHME 2012 έδωσαν ποσοστό τοπολογικής αναγνώρισης (μετρική STRUCT) ίσο με 91.67%, 78.70% και 70.37% για το Μέρος-I, 88.37%, 69.44% και 50.17% για το Μέρος -II και 80.33%, 61.27% και 42.21% για το Μέρος -III. Για το Μέρος-IV υπάρχουν αναφορές μόνο της μετρικής EXP_Rec. Συγκεκριμένα τα τρία πρώτα συστήματα του διαγωνισμού CROHME 2013 έδωσαν ποσοστό αναγνώρισης (μετρική EXP_Rec) ίσο με 60.36%, 23.40% και 19.79% (**Mouchère et al. 2012, 2013**).

Για να αξιολογήσουμε την χρήση διαφορετικών γραμματικών, χρησιμοποιήσαμε το σύνολο εξέτασης του CROHME 2012 και 2013. Στο Πίνακα 22 παρουσιάζονται τα αποτελέσματα της

μετρικής EXP_Rec στα σύνολα εξέτασης των Μέρους-I, Μέρους -II, Μέρους -III και Μέρους-IV χρησιμοποιώντας τις γραμματικές Gram-I, Gram-II και Gram-III. Πειραματικά αποτελέσματα έδειξαν ότι το μέγιστο ποσοστό ακρίβειας επιτυγχάνεται χρησιμοποιώντας την πιο κατάλληλη γραμματική για κάθε μέρος. Συγκεκριμένα, το ποσοστό ακρίβειας για το Μέρους-I στην πραγματικότητα δεν επηρεάζεται από την χρήση πολυπλοκότερης γραμματικής. Ενώ το ποσοστό ακρίβειας για τα Μέρους-II, Μέρους-III και Μέρους-IV μειώνεται όταν χρησιμοποιούνται απλούστερες γραμματικές.

	Gram1	Gram2	Gram3
Μέρους-I	78.70%	77.78%	77.78%
Μέρους -II	23.59%	65.78%	65.45%
Μέρους -III	26.95%	55.56%	56.37%
Μέρους -IV	42.32%	50.07%	50.22%

Πίνακας 22. Αποτελέσματα της μετρικής EXP_Rec για τα σύνολα εξέτασης του CROHME 2012 και 2013 χρησιμοποιώντας διαφορετικές γραμματικές.

Εξετάζοντας προσεχτικά τα αποτελέσματα που παρουσιάζονται στους Πίνακες 21 και 22, συμπεραίνουμε ότι η προτεινόμενη μέθοδος βελτιώνει τα ποσοστά ακρίβειας σε σχέση με την προηγούμενη μέθοδο. Πρέπει να σημειωθεί επίσης ότι τα ποσοστά ακρίβειας των δύο μετρικών EXP_Rec και STRUCT είναι σχεδόν ταυτόσημα.

Συγκεκριμένα τα ποσοστά ακρίβειας των μετρικών EXP_Rec και STRUCT στο σύνολο εξέτασης του Μέρους-I αυξάνεται από το 8.33% στο 78.70% και από το 21.30% στο 78.70% αντίστοιχα. Επίσης στο σύνολο εξέτασης του Μέρους-II αυξάνεται από 6.64% σε 65.78% και από 12.29% σε 66,44% αντίστοιχα. Τέλος στο σύνολο εξέτασης του Μέρους-III, το ποσοστό ακρίβειας αυξάνεται από 3.69% σε 56.37% και από 9.84% σε 57.41% αντίστοιχα.

Στο σύνολο εξέτασης του Μέρους-IV η απόδοση του προτεινόμενου αλγόριθμου δίνει ποσοστό ακρίβειας ίσο με 50.52% για την μετρική STRUCT και 50.22% για την μετρική EXP_Rec. Να σημειωθεί ότι για το Μέρους-IV δεν έχουμε προηγούμενα αποτελέσματα για σύγκριση. Επιπλέον φαίνεται ότι η απόδοση της προτεινόμενης μεθόδου είναι ανταγωνιστική των κορυφαίων συστημάτων του διαγωνισμού CROHME. Τέλος να σημειωθεί ότι τα αποτελέσματα των συστημάτων των συμμετεχόντων του διαγωνισμού θα μπορούσαν να βελτιωθούν περαιτέρω αν είχαν επίσης προϋποθέσει την σωστή ομαδοποίηση των κονδυλίων σε σύμβολα και την σωστή αναγνώριση των συμβόλων αυτών.

5.5. Συμπεράσματα

Στην παρούσα εργασία προτείνεται μία μέθοδος για την αναγνώριση χειρόγραφων ΜΕ. Συγκεκριμένα εκπαιδεύσαμε ένα πιθανοτικό SVM ταξινομητή για να αναγνωρίσουμε την τοπολογική σχέση μεταξύ συμβόλων της ΜΕ. Η εκπαίδευση και αξιολόγηση της μεθόδου έγινε με την βάση χειρόγραφων ΜΕ MathBrush. Τα πειραματικά αποτελέσματα έδωσαν ποσοστό λάθους ίσο με 2.8% στην βάση MathBrush. Για την αναγνώριση των χειρόγραφων ΜΕ

αναπτύχθηκε μία γραμματική SCFG για την μοντελοποίηση της σύνταξης των ΜΕ. Έπειτα χρησιμοποιήθηκε ένα αλγόριθμος βασισμένος στον αλγόριθμο CYK για την ανάλυση των ΜΕ και την εξαγωγή του MathML μορφότυπου που αποτυπώνει το αποτέλεσμα της αναγνώρισης της ΜΕ. Για την αξιολόγηση των αποτελεσμάτων χρησιμοποιήθηκαν τα σύνολα εξέτασης των βάσεων CROHME2012 και CROHME2013, καθώς και τα εργαλεία αξιολόγησης που παρείχαν οι διοργανωτές του διαγωνισμού CROHME. Να σημειωθεί ότι στην προτεινόμενη τεχνική προϋποθέτουμε ότι όλα τα σύμβολα της ΜΕ έχουν αναγνωρισθεί σωστά. Τα πειραματικά αποτελέσματα έδωσαν ποσοστά ορθής αναγνώρισης (EXP_Rec μετρική) 78.70%, 65.78%, 56.37% και 50.22% για τα Μέρος-I, Μέρος -II, Μέρος -III και Μέρος -IV αντίστοιχα. Επίσης αξίζει να σημειωθεί ότι ο SVM ταξινομητής που χρησιμοποιήθηκε για την αναγνώριση των ΜΕ του συνόλου εξέτασης CROHME, εκπαιδεύτηκε στην βάση MathBrush.

Οι περισσότερες από τις λανθασμένα αναγνωρισμένες ΜΕ οφείλονται στην διαφορετική σειρά γραφής των συμβόλων που απαρτίζουν την ΜΕ. Για παράδειγμα, στο σύνολο εξέτασης υπάρχουν ΜΕ που περιέχουν κλάσματα τα οποία είναι γραμμένα ακολουθώντας διαφορετική σειρά γραφής των συμβόλων: (i) αριθμητής/ γραμμή_κλάσματος/παρονομαστής, (ii) γραμμή_κλάσματος/παρονομαστής/αριθμητής, (iii) παρονομαστής /γραμμή_κλάσματος/αριθμητής και (iv) παρονομαστής/αριθμητής/ γραμμή_κλάσματος. Για την αντιμετώπιση αυτού το προβλήματος θα πρέπει μελλοντικά να αναπτυχθεί ένας αλγόριθμος που ορίζει την τοπολογική σειρά των συμβόλων. Επιπρόσθετα η προτεινόμενη τεχνική αναγνώρισης συχνά ταξινομεί λανθασμένα τις τοπολογικές σχέσεις εκθέτη/δείκτη της ΜΕ. Πιθανότατα αυτά τα λάθη οφείλονται στους διαφορετικούς τρόπους γραφής των συμβόλων που πολλές φορές οδηγούν σε λανθασμένη κατηγοριοποίησή τους σαν ascender/descender/centered. Για παράδειγμα το μικρό γράμμα β μπορεί να ταξινομηθεί είτε σαν ascender, είτε σαν descender, ανάλογα με τον τρόπο γραφής του εκάστοτε γραφέα. Το πρόβλημα αυτό θα μπορούσε να αντιμετωπισθεί μελλοντικά με την χρήση αυτόματης κατηγοριοποίησης των συμβόλων στις τρεις αυτές κατηγορίες ανάλογα με τον τρόπο γραφής.

Κεφάλαιο 6. Συμπεράσματα

Στην παρούσα εργασία εξετάστηκε το πρόβλημα της αναγνώρισης χειρόγραφων μαθηματικών εξισώσεων. Το πρόβλημα της αναγνώρισης χειρόγραφων μαθηματικών εξισώσεων έγκειται στην γενική κατηγορία της επεξεργασίας εγγράφου, στην οποία περιλαμβάνεται επίσης και το πρόβλημα της αναγνώρισης κειμένου (τυπωμένου ή χειρόγραφου). Σε αντίθεση με την αναγνώριση κειμένου, η διαδικασία αναγνώρισης χειρόγραφων μαθηματικών εξισώσεων είναι μία πιο πολύπλοκη διαδικασία λόγω της δισδιάστατης γραφής των ΜΕ, του μεγάλου αριθμού μαθηματικών συμβόλων και των διαφορετικών τρόπων γραφής.

Το πρόβλημα της αναγνώρισης των ΜΕ μπορεί να χωριστεί στα ακόλουθα επιμέρους προβλήματα: (α) το πρόβλημα της αναγνώρισης των συμβόλων, (β) το πρόβλημα αναγνώρισης της τοπολογικής σχέσης μεταξύ δύο συμβόλων και (γ) το πρόβλημα της ανασύνθεσης της ΜΕ. Στο πρόβλημα της αναγνώρισης των μαθηματικών συμβόλων δημιουργήθηκε αρχικά ένα πρόγραμμα συλλογής δεδομένων με σκοπό της δημιουργίας βάσης προτύπων για την εκπαίδευση και αξιολόγηση του συστήματος αναγνώρισης μαθηματικών συμβόλων. Ακολούθως υλοποιήθηκαν δύο τεχνικές αναγνώρισης: (α) η τεχνική ελαστικού ταιριάσματος (elastic matching) και (β) η τεχνική ταύτισης απόστασης με υπόδειγμα (string template matching). Η τεχνική ελαστικού ταιριάσματος αποδείχθηκε χρονοβόρα και με χαμηλότερα ποσοστά αναγνώρισης συγκριτικά με την τεχνική της ταύτισης απόστασης με υπόδειγμα. Η μέθοδος αξιολογήθηκε στις βάσεις LVZE και ILSP και παρουσίασε ποσοστό ορθής αναγνώρισης ίσο με 92%.

Στο πρόβλημα της αναγνώρισης της τοπολογικής σχέσης μεταξύ δύο συμβόλων ορίστηκαν 5 τοπολογικές σχέσεις: (α) πάνω-δεξιά, (β) δεξιά, (γ) κάτω-δεξιά, (δ) κάτω και (ε) μέσα. Ακολούθως προτάθηκε μία καινοτόμος τεχνική βασισμένη στις μηχανές διανυσμάτων υποστήριξης (SVM). Η μέθοδος αυτή χρησιμοποιεί 7 γεωμετρικά χαρακτηριστικά βασισμένα στο πλαίσιο οριοθέτησης και στο κεντροειδές των συμβόλων. Αξιολογήθηκε στην βάση MathBrush και παρουσίασε ποσοστό λανθασμένης αναγνώρισης ίσο με 2.71%.

Τέλος στο πρόβλημα της ανασύνθεσης της ΜΕ υλοποιήθηκαν στοχαστικές γραμματικές τύπου CFG και ένας CYK based parser. Η μέθοδος αξιολογήθηκε στην βάση CROHME και παρουσίασε ποσοστό ορθής αναγνώρισης ίσο με 78.70% για το Μέρος-I, 65.78% για το Μέρος-II, 56.37% για το Μέρος-III και 50.22% για το Μέρος-IV.

Σε μελλοντικές εργασίες θα προσπαθήσουμε να παραλληλοποιήσουμε τα στάδια αναγνώρισης συμβόλου και ανασύνθεσης ΜΕ με σκοπό την αξιοποίηση της γνώσης του περιεχομένου της ΜΕ. Παράλληλα θα εξεταστεί η χρήση νευρωνικών δικτύων τύπου LSTM (Long Short Term

Memory) για την αναγνώριση των συμβόλων. Τα νευρωνικά δίκτυα τύπου LSTM έχουν παρουσιάσει υψηλό ποσοστό αναγνώρισης αρχικά στην αντιμετώπιση της αναγνώρισης φωνής και τελευταία στην αναγνώριση κειμένου.

Βιβλιογραφία

- Alvaro, F., Sanchez, J.-A. and Benedi, J.-M., 2013.** Classification of Online Mathematical Symbols with Hybrid Features and Recurrent Neural Networks. International Conference on Document Analysis and Recognition, pp. 1012–1016.
- Alvaro, F., Sanchez, J.-A. and Benedi, J.-M., 2014.** Recognition of on-line handwritten mathematical expressions using 2D stochastic context-free grammars and hidden Markov models. Pattern Recognition Letters, vol. 35, no. 0, pp. 58 – 67.
- Aly, W., Uchida, S., Suzuki, M., 2008.** Identifying subscripts and superscripts in mathematical documents. Math. Comput. Sci. 2(2), pp. 195–209.
- Anderson, R.H., 1968.** Syntax-Directed Recognition of Hand-Printed Two-Dimensional Equations. PhD thesis, Harvard University.
- Anthony, L., Yang, J., Koedinger, K.R., 2007.** Adapting handwriting recognition for applications in algebra learning. In: Proceedings of the ACM Workshop on Educational Multimedia and Multimedia Education, pp. 47–56. Augsburg
- Awal, A., Mouchère, H. and Viard-Gaudin, C., 2010.** Improving online handwritten mathematical expressions recognition with contextual modeling. 12th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 427-432.
- Blostein, D. and Grbavec, A., 1997.** Recognition of mathematical notation. Handbook of Character Recognition and Document Image Analysis, 557–582.
- Chan, K. and Yeung, D., 2000.** Mathematics expression recognition: A Survey. Journal on Document Analysis and Recognition, 3(1):3-15, August 2000.
- Chan, K.-F. and Yeung, D.-Y., 2001a.** Error detection, error correction and performance evaluation in on-line mathematical expression recognition. Pattern Recognition, 34(8):1671–1684.
- Chan, K.-F. and Yeung, D.-Y., 2001b.** Pencilc: A novel application of on-line mathematical expression recognition technology. In Proc. Int'l Conf. Document Analysis and Recognition, pp. 774–778, Seattle, USA.
- Chan, K.-F., Yeung, D.-Y., 1999.** Recognizing on-line handwritten alphanumeric characters through flexible structural matching. Pattern Recognition, 32(7):1099–1114.
- Chan, K.-F., Yeung, D.-Y., 2000.** An efficient syntactic approach to structural analysis of on-line handwritten mathematical expressions. Pattern Recognition, 33(3):375–384.
- Chan, K.-F., Yeung, D.-Y., 2001.** Pencilc: a novel application of on-line mathematical expression recognition technology. In: Proceedings of the International Conference on Document Analysis and Recognition, Seattle, pp. 774–778.
- Chang, Chih-Chung and Lin, Chih-Jen, 2011.** LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, pp.1-27.
- Chang, S. K., 1970.** A method for the structural analysis of two-dimensional mathematical expressions. Information Sciences, 2:253-272.
- Eto, Y. and Suzuki, M., 2001.** Mathematical formula recognition using virtual link network. Int'l Conf. Document Analysis and Recognition, 430–437.
- Freeman, H., 1974.** Computer processing of line drawing images, ACM Comput. Surveys 6 (1) 57-98.
- Freund, Y. and Schapire, R.E., 1995.** A decision-theoretic generalization of on-line learning and an application to boosting. J. Computer and System Sciences, 55(1):119–139.

- Garain, U., Chaudhuri, B.B., 2004.** Recognition of Online Handwritten Mathematical Expressions. *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, Vol. 34, No. 6, pp. 2366-2376.
- Golubitsky, O., Watt, S. M., 2010.** Improved classification through runoff elections. In: *International Workshop on Document Analysis Systems*, 59–64.
- Gonzalez, R.C. and Woods, R.E., 2002.** *Digital Image Processing*. Prentice Hall.
- Graves, A., Liwicki, M., Fernandez, S., Bertolami, R., Bunke, H. and Schmidhuber, J., 2009.** A novel connectionist system for unconstrained handwriting recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 5, pp. 855–868.
- <http://www.w3.org/TR/InkML/>**
- Hu, L. and Zanibbi, R., 2011.** HMM-Based Recognition of Online Handwritten Mathematical Symbols Using Segmental K-Means Initialization and a Modified Pen-Up/Down Feature. *International Conference on Document Analysis and Recognition*, pp. 457–462.
- Kassel, R., 1995.** A comparison of approaches to on-line handwritten character recognition. Ph.D. Thesis, MIT Department of Electrical Engineering and Computer Science.
- Keshari, B., Watt, S.M., 2008.** Online mathematical symbol recognition using svms with features from functional approximation. *Mathematical User-Interfaces Workshop*, 27 July 2008, Birmingham, UK.
- LaViola, J. J., and Zeleznik, R.C., 2007.** A practical Approach for Writer-Depended Symbol Recognition Using a Writer-Independent Symbol Recognizer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, pp. 1917-1926.
- Levenberg, K., 1944.** A Method for the Solution of Certain Non-Linear Problems in Least Squares. *Quarterly of Applied Mathematics* 2: 164–168
- MacLean, S. and Labahn, G., 2010.** Elastic matching in linear time and constant space. *IAPR Workshop on Document Analysis Systems*.
- MacLean, S., Labahn, G., Lank, E., Marzouk, M., and Tausky, D., 2010.** Grammar-based techniques for creating groundtruthed sketch corpora. *International Journal on Document Analysis and Recognition*, pp. 1–21.
- Matsakis, N., 1999.** Recognition of handwritten mathematical expressions. Master's thesis, MIT, Cambridge, MA.
- Mouchère H., Viard-Gaudin C., R. Zanibbi, Garain U., Kim D.H., Kim J. H., 2013.** ICDAR 2013 CROHME: Third International Competition on Recognition of Online Handwritten Mathematical Expressions. In: *Proc. 12th International Conference on Document Analysis and Recognition*.
- Mouchère, H., Viard-Gaudin, C., Garain, U., Kim, D. H., and Kim, J. H., 2011.** CROHME 2011: Competition on Recognition of Online Handwritten Mathematical Expression. *11th International Conference on Document Analysis and Recognition (ICDAR)*.
- Mouchère, H., Viard-Gaudin, C., Kim, D.H., Kim, J.H., and Garain, U., 2012.** ICFHR 2012 – Competition on recognition of on-line mathematical expressions (CROHME 2012). In: *Proc. 13th International Conference on Frontiers in Handwriting Recognition*, pp. 807-812.
- Nagy, G., 2000.** Twenty years of document image analysis in PAMI. *IEEE Trans on Pattern Anal. Mach. Intell.* 22(1), 38–62.
- Okamoto, M. and Miao, B., 1991.** Recognition of mathematical expressions by using the layout structures of symbols. *Int'l Conf. Document Analysis and Recognition*, 1:242-250.
- Okamura, H., Kanahori, T., Suzuki, M., Fakuda, H., Cong, R. and Tamari, F., 1999.** *Handwriting Interface for Computer Algebra Systems*. In *Proceedings of Graphics Interface*.
- Panic, M., 2009.** Math handwriting recognition in Windows 7 and its benefits. In: *Intelligent Computer Mathematics*, vol. 5625 of LNCS, Springer, pp. 29–30.

- Plamondon, R. and Srihari, S.N. 2000.** On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intell*, 22(1): 63–84.
- Rabiner, L.R., 1989.** A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proc. IEEE*, 77(2):257–286.
- Raja, A., Rayner, M., Sexton, A.P. and Sorge, V., 2006.** Towards a parser for mathematical formula recognition. In *Mathematical Knowledge Management*, V 4108 of LNAI, pp. 139–151.
- Rodrguez, E. T., 2003.** JMathNotes. <<http://page.mi.fu-berlin.de/tapia/JMathNotes/>>.
- Shi, Y., Li, H.Y., and Soong, F.K., 2007.** A unified framework for symbol segmentation and recognition of handwritten mathematical expressions. *Int'l Conf. Document Analysis and Recognition*, volume 2, pp. 854–858, Curitiba, Brazil.
- Shilman, M., Viola, P., and Chellapilla, K., 2004.** Recognition and grouping of handwritten text in diagrams and equations. In *Ninth International Workshop on Frontiers in Handwriting Recognition*, pp. 569–574.
- Simistira, F., Katsouros, V., and Carayannis, G., 2008.** A Template Matching Distance for Recognition of On-Line Mathematical Symbols. *11th International Conference on Frontiers in Handwriting Recognition (ICFHR)*.
- Smirnova, E., Watt, S.M., 2008.** Communicating mathematics via penbased computer interfaces. In: *Proceedings of the International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, pp. 9–18.
- Smithies, S., 1999.** Freehand Formula Entry System. Master's thesis, University of Otago, Dunedin, New Zealand.
- Smithies, S., Novins, K., Arvo, J., 1999.** A Handwriting-Based Equation Editor, In *Proc. Graphics Interface*, Kingston, Ontario, Canada.
- Smithies, S., Novins, K., Arvo, J., 2001.** Equation Entry and Editing via Handwriting and Gesture Recognition. *Behaviour and Information Technology*, 20(1):53-67.
- Suzuki, M., Tamari, F., Fukuda, R., Uchida, S., Kanahori, T., 2003.** Infty- an integrated OCR system for mathematical documents, *Proceedings of ACM Symposium on Document Engineering 2003*, Grenoble, Ed.C.Vanoirbeek, C.Roisin, E. Munson, pp.95-104.
- Tapia, E. and Rojas, R. 2004.** Recognition of on-line handwritten mathematical expressions using a minimum spanning tree construction and symbol dominance. *Graphics Recognition: Recent Advances and Perspectives*, 3088:329–340.
- Tapia, E. and Rojas, R., 2007.** A survey on recognition of on-line handwritten mathematical notation. *Technical Report B-07-01*, Free University of Berlin.
- Tappert, C. C., 1982.** Cursive Script Recognition by Elastic Matching. *IBM Journal of Research and Development*, 26(6):765–771.
- Tappert, C., Suen, C. Y. and Wakahara, T., 1990.** The State of the Art in On-Line Handwriting Recognition. *IEEE Transactions on Pattern Analysis and Machine Intell*, 12(8):787-808.
- Tausky, D., Labahn, G., Lank, E., Marzouk, M., 2007.** Managing ambiguity in mathematical matrices. In: *Proceedings of the Eurographics Workshop on Sketch-Based Interfaces and Modeling*, pp. 115–122.
- Uchida, S., Sakoe, H., 2005.** A survey of Elastic Matching Techniques for Handwritten Character Recognition. *IEICE Transactions on Information and Systems*, Vol E88-D, No. 8, pp. 1781-1790.
- Wang, H.-J., and Lee, J.-S., 1997.** Design of a mathematical expression understanding system. *Pattern Recognition Letters*, 18(3):289–298.

Winkler, H.-J., 1996. HMM-based handwritten symbol recognition using on-line and off-line features. In Proc. IEEE Int'l Conf. Acoustics Speech and Signal Processing, pp. 3438–3441, Atlanta, GA.

Zanibbi, R., Blostein, D., and Cordy, J.R., 2001. Baseline Structure Analysis of Handwritten Mathematics Notation, In Proceedings of the Sixth International Conference on Document Analysis and Recognition, Seattle, Washington, pp. 768-773.

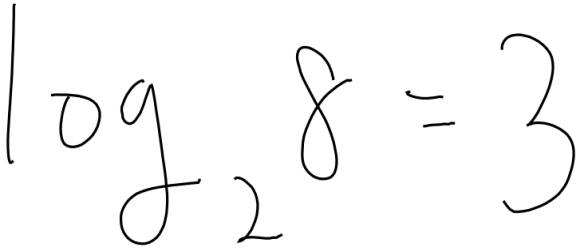
Zanibbi, R., Blostein, D., and Cordy, J.R., 2002. Recognizing Mathematical Expressions Using Tree Transformation. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 11, pp. 1455-1467.

Zanibbi, R., Novins, K., Arvo, J. and Zanibbi, K., 2001. Aiding Manipulation of Handwritten Mathematical Expressions through Style-Preserving Morphs, In Proceedings of Graphics Interface 2001, Ottawa, Canada, pp. 127-134.

Zhang, L., Blostein, D. and Zanibbi, R., 2005. Using Fuzzy Logic to Analyze Superscript and Subscript Relations in Handwritten Mathematical Expressions, in Proc. Int'l Conf. Document Analysis and Recognition, pp. 972-976, Seoul, Korea.

Παράρτημα Α

Αναπαράσταση της χειρόγραφης μαθηματικής έκφρασης $\log_2 8 = 3$ σε μορφότυπο inkml.

A handwritten mathematical expression in black ink on a white background. The expression is $\log_2 8 = 3$. The 'log' is written in a cursive style, the subscript '2' is written below the 'g', the '8' is written to the right of the subscript, and the equals sign and '3' are written to the right of the '8'.

```
1 <ink xmlns="http://www.w3.org/2003/InkML">
2 <traceFormat>
3 <channel name="X" type="decimal" />
4 <channel name="Y" type="decimal" />
5 </traceFormat>
6 <annotation type="truth">\log_{2} 8 = 3</annotation>
7 <annotation type="writer">fr11Sog/zxB5kFbt2NBZV8Evv9s=</annotation>
8 <annotation type="gender">Male</annotation>
9 <annotation type="age">25</annotation>
10 <annotation type="hand">right-handed</annotation>
11 <annotation type="UI">2011_KAIST_Data1_0_sub_1</annotation>
12 <annotation type="copyright">KAIST</annotation>
13 <annotationXML type="truth" encoding="Content-MathML">
14 <math xmlns="http://www.w3.org/1998/Math/MathML">
15 <mrow>
16 <msub>
17 <mo xml:id="\log_1">\log</mo>
18 <mn xml:id="2_1">2</mn>
19 </msub>
20 <mrow>
21 <mn xml:id="8_1">8</mn>
22 <mrow>
23 <mo xml:id="=_1">=</mo>
24 <mn xml:id="3_1">3</mn>
25 </mrow>
26 </mrow>
27 </mrow>
28 </math>
29 </annotationXML>
30 <trace id="0">10350 2248, 10338 2224, 10334 2207, 10334 2207, 10323 2184,
10323 2184, 10323 2184, 10323 2184, 10321 2221, 10321 2276, 10315 2353, 10306
2457, 10287 2581, 10269 2729, 10245 2885, 10220 3044, 10196 3193, 10175 3327,
10157 3442, 10145 3532, 10139 3603, 10135 3647, 10136 3668, 10136 3668, 10159
3656</trace>
31 <trace id="1">10676 2848, 10652 2857, 10629 2866, 10614 2888, 10606 2919,
10603 2959, 10612 3006, 10628 3063, 10658 3119, 10689 3169, 10741 3211, 10795
3238, 10860 3249, 10922 3240, 10982 3216, 11030 3173, 11062 3118, 11077 3054,
11068 2992, 11035 2934, 10985 2882, 10923 2842, 10854 2809, 10793 2792, 10745
2792, 10709 2793, 10709 2793, 10707 2812</trace>
32 <trace id="2">11778 2860, 11773 2833, 11755 2804, 11730 2783, 11696 2775,
11653 2775, 11606 2783, 11556 2806, 11506 2842, 11461 2884, 11431 2932, 11415
2980, 11415 3029, 11428 3074, 11458 3118, 11494 3142, 11541 3157, 11591 3161,
11639 3157, 11681 3143, 11718 3122, 11745 3095, 11764 3066, 11776 3041, 11784
```

3017, 11785 2997, 11785 2982, 11785 2982, 11785 2982, 11785 2982, 11785 2982, 11785 2982, 11790 3008, 11796 3032, 11802 3066, 11811 3115, 11817 3178, 11825 3255, 11829 3344, 11835 3439, 11838 3544, 11841 3654, 11849 3761, 11849 3867, 11849 3966, 11846 4060, 11840 4143, 11823 4220, 11800 4284, 11767 4334, 11722 4372, 11671 4388, 11615 4384, 11555 4355, 11497 4307, 11446 4244, 11402 4164, 11374 4085, 11360 4004, 11368 3924, 11393 3846, 11437 3775, 11497 3710, 11564 3650, 11645 3599, 11731 3544, 11817 3498, 11889 3455, 11957 3424, 12015 3410, 12052 3401, 12091 3406</trace>

33 <trace id="3">12453 3859, 12453 3859, 12461 3841, 12461 3841, 12476 3826, 12476 3826, 12500 3813, 12515 3804, 12538 3804, 12559 3804, 12587 3810, 12615 3823, 12643 3841, 12666 3865, 12684 3895, 12693 3929, 12696 3966, 12690 4013, 12673 4054, 12649 4094, 12619 4131, 12584 4164, 12548 4189, 12512 4209, 12474 4223, 12446 4227, 12423 4226, 12408 4220, 12408 4220, 12405 4194, 12422 4180, 12446 4165, 12485 4152, 12526 4137, 12575 4125, 12624 4113, 12678 4107, 12729 4101, 12776 4096, 12815 4094, 12847 4094, 12872 4096, 12886 4096, 12904 4104, 12904 4104, 12904 4104, 12924 4116, 12924 4116</trace>

34 <trace id="4">13432 3158, 13445 3161, 13445 3161, 13466 3146, 13491 3139, 13515 3122, 13549 3106, 13581 3078, 13620 3044, 13652 3003, 13687 2965, 13711 2919, 13729 2872, 13735 2822, 13733 2772, 13723 2726, 13703 2689, 13678 2655, 13644 2623, 13605 2602, 13561 2591, 13516 2590, 13469 2598, 13424 2614, 13382 2638, 13344 2670, 13317 2706, 13296 2750, 13287 2800, 13292 2852, 13308 2908, 13334 2967, 13368 3029, 13412 3090, 13457 3157, 13509 3222, 13563 3286, 13613 3351, 13658 3415, 13703 3476, 13741 3535, 13769 3592, 13792 3644, 13803 3689, 13809 3728, 13807 3764, 13801 3799, 13786 3828, 13765 3853, 13738 3876, 13703 3894, 13665 3909, 13622 3920, 13573 3920, 13521 3911, 13466 3897, 13412 3874, 13362 3843, 13319 3804, 13281 3757, 13249 3709, 13228 3654, 13215 3602, 13213 3543, 13224 3484, 13246 3427, 13283 3372, 13331 3323, 13385 3276, 13444 3237, 13506 3207, 13563 3184, 13611 3167, 13656 3169, 13687 3175, 13703 3181, 13720 3190, 13720 3190</trace>

35 <trace id="5">14695 3005, 14695 3005, 14682 3005, 14682 3005, 14682 3005, 14682 3005, 14718 3006, 14748 3002, 14795 3003, 14852 2999, 14926 2997, 15009 2991, 15099 2985, 15188 2977, 15273 2971, 15346 2964, 15410 2955, 15461 2952, 15491 2944, 15491 2944, 15511 2940, 15511 2940</trace>

36 <trace id="6">14817 3553, 14817 3553, 14817 3553, 14817 3553, 14847 3540, 14872 3528, 14920 3520, 14977 3504, 15060 3490, 15154 3469, 15262 3454, 15377 3437, 15490 3422, 15592 3409, 15684 3398, 15754 3390, 15800 3378, 15832 3368, 15832 3368, 15832 3368, 15832 3353</trace>

37 <trace id="7">16563 2634, 16553 2611, 16559 2594, 16575 2584, 16598 2572, 16639 2564, 16687 2552, 16753 2549, 16827 2543, 16916 2548, 17002 2552, 17089 2569, 17169 2585, 17240 2611, 17297 2643, 17338 2679, 17359 2715, 17358 2757, 17335 2800, 17290 2843, 17230 2888, 17157 2932, 17080 2973, 17005 3012, 16933 3048, 16871 3077, 16822 3103, 16795 3122, 16785 3139, 16792 3154, 16815 3167, 16857 3181, 16916 3196, 16993 3214, 17079 3237, 17174 3267, 17269 3303, 17361 3345, 17442 3390, 17511 3443, 17561 3496, 17591 3555, 17599 3612, 17587 3665, 17551 3718, 17493 3772, 17413 3817, 17314 3856, 17198 3886, 17070 3905, 16945 3906, 16822 3905, 16715 3894, 16623 3874, 16559 3853, 16516 3834, 16495 3810, 16501 3787</trace>

38 <traceGroup xml:id="8">

39 <annotation type="truth">Segmentation</annotation>

40 <traceGroup xml:id="9">

41 <annotation type="truth">\log</annotation>

42 <annotationXML href="\log_1" />

43 <traceView traceDataRef="0" />

44 <traceView traceDataRef="1" />

45 <traceView traceDataRef="2" />

46 </traceGroup>

47 <traceGroup xml:id="10">


```

48 <annotation type="truth">2</annotation>
49 <annotationXML href="2_1" />
50 <traceView traceDataRef="3" />
51 </traceGroup>
52 <traceGroup xml:id="11">
53 <annotation type="truth">8</annotation>
54 <annotationXML href="8_1" />
55 <traceView traceDataRef="4" />
56 </traceGroup>
57 <traceGroup xml:id="12">
58 <annotation type="truth">=</annotation>
59 <annotationXML href="=_1" />
60 <traceView traceDataRef="5" />
61 <traceView traceDataRef="6" />
62 </traceGroup>
63 <traceGroup xml:id="13">
64 <annotation type="truth">3</annotation>
65 <annotationXML href="3_1" />
66 <traceView traceDataRef="7" />
67 </traceGroup>
68 </traceGroup>
69 </ink>

```

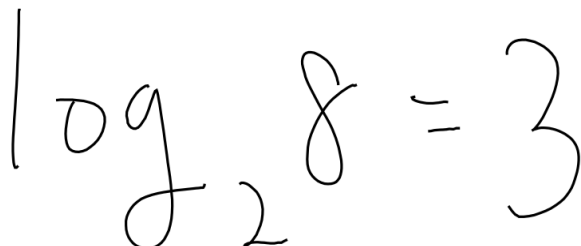
Στις γραμμές 3,4 ορίζεται ο τύπος της πληροφορίας που αποθηκεύεται, δηλ. στο συγκεκριμένο inkml αρχείο αποθηκεύονται οι x,y συντεταγμένες της ME. Στην γραμμή-6 αποθηκεύεται η LaTeX αποτύπωση της ME: $\log_{2} 8 = 3$. Στις γραμμές 7-10 αποθηκεύονται διάφορες πληροφορίες σχετικά με τον γραφέα της ME (κωδικός γραφέα, φύλο, ηλικία κλπ). Στις γραμμές 11 αποθηκεύεται ο μοναδικός κωδικός UI του ψηφιακού μελανιού και στην γραμμή 12 αποθηκεύεται το copyright του ψηφιακού μελανιού. Στις γραμμές 14-29 αποθηκεύεται η αντίστοιχη MathML αποτύπωση της ME. Στις γραμμές 30-37 αποθηκεύονται οι x,y συντεταγμένες για κάθε κονδυλιά της ME. Στην συγκεκριμένη ME περιλαμβάνονται 8 κονδυλιές (trace id="0" - "8"). Στις γραμμές 38-68 παρουσιάζεται η ομαδοποίηση των κονδυλιών σε σύμβολα και η αντίστοιχη ετικέτα του συμβόλου. Π.χ. το σύμβολο "log" αποτελείται από τις κονδυλιές με id 0,1,2, η αντίστοιχη ετικέτα του συμβόλου είναι \log και αντιστοιχίζεται με την 17 μέσω του identifier xml:id=" \log_1 ".

Παράρτημα Β

Η γλώσσα MathML (Mathematical Markup Language) δημιουργήθηκε το 1995 από το World Wide Web Consortium. Είναι μια εφαρμογή της XML που χρησιμοποιείται για την αναπαράσταση των μαθηματικών εκφράσεων κυρίως στο διαδίκτυο.

Π.χ. η χειρόγραφη μαθηματική έκφραση $\log_2 8 = 3$, μπορεί να γραφτεί σε γλώσσα MathML

ως:

A handwritten mathematical expression in black ink on a white background. The expression is $\log_2 8 = 3$. The 'log' is written in a cursive style, the subscript '2' is below the 'log', '8' is to the right of the subscript, and '=' is followed by '3'.

```
<annotationXML type="truth" encoding="Content-MathML">
<math xmlns="http://www.w3.org/1998/Math/MathML">
  <mrow>
    <msub>
      <mo xml:id="\log_1">\log</mo>
      <mn xml:id="2_1">2</mn>
    </msub>
    <mrow>
      <mn xml:id="8_1">8</mn>
    </mrow>
    <mo xml:id="=_1">=</mo>
    <mn xml:id="3_1">3</mn>
  </mrow>
</math>
</annotationXML>
```

Παράρτημα Γ

Για την απεικόνιση των μαθηματικών συμβόλων και παραστάσεων χρησιμοποιήθηκε το Unicode Standard, Version 5.0

Συγκεκριμένα:

1. Mathematical Operators (Range: 2200 – 22FF)
<http://www.unicode.org/charts/PDF/U2200.pdf>
2. Arrows (Range: 2190 – 21FF)
<http://www.unicode.org/charts/PDF/U2190.pdf>
3. General Punctuation (Range: 2000-2026F)
<http://www.unicode.org/charts/PDF/U2000.pdf>
4. C1 Controls and Latin-1 Supplement (Range: 0080-00FF)
<http://www.unicode.org/charts/PDF/U0080.pdf>
5. Miscellaneous Mathematical Symbols-A (Range: 27C0-27EF)
<http://www.unicode.org/charts/PDF/U27C0.pdf>
6. C0 Controls and Basic Latin (Range: 0000-007F)
<http://www.unicode.org/charts/PDF/U0000.pdf>

Παράρτημα Δ

Το μορφότυπο isf είναι ένα από τα βασικά μορφότυπα αποθήκευσης ψηφιακού μελανιού. Αποθηκεύει τις συντεταγμένες της μονάδας του ψηφιοποιητή (digitizer) σε συμπιεσμένο δυαδικό μορφότυπο (binary format). Επιτρέπει την δημιουργία πρόσθετων παραμέτρων οριζόμενων από τον χρήστη (π.χ. χρόνος δημιουργίας μελανιού) και μπορεί να αποθηκεύσει οποιαδήποτε μορφή πακέτων δεδομένων εξαγόμενων από τον ψηφιοποιητή. Το ISF είναι μορφότυπο της Microsoft®.

Παράρτημα Ε

Στο Παράρτημα Ε παρουσιάζονται οι γραμματικές που δημιουργήθηκαν για την περιγραφή των ΜΕ που περιέχει η βάση χειρόγραφων μαθηματικών εκφράσεων CROHME.

Η Context Free γραμματική που χρησιμοποιήθηκε αποτελείται από:

N είναι ένα σύνολο από μη τερματικά σύμβολα:

Exp
Sym
ExpOp
OpUn
ROpUn
OBExp
OpBin
OverExp
Over
OverSym
LeftPar
RightPar
RPExp
SSExp
BigOpExp
BigOp
Sqrt
Func
2Let
Let
SupSym

S είναι ένα αρχικό σύμβολο:

Exp
Sym

Για το **Μέρος-I** χρησιμοποιήθηκαν οι ακόλουθοι κανόνες παραγωγής (σύνολο R):

P	Σχέση	S	A	B	Latex	MathML
1.0	H	Exp	Exp	Exp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Exp	Sym	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Sym	Exp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Sym	Sym	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	OpUn	Exp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	OpUn	Sym	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Exp	ROpUn	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Sym	ROpUn	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Exp	OBExp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	Sym	OBExp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	OBExp	OpBin	Exp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	OBExp	OpBin	Sym	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	Sup	Exp	Sym	Sym	$\{&1\}^{\{&2\}}$	$\langle msup \rangle &1&2 \langle /msup \rangle$
1.0	Sub	Exp	Sym	Sym	$\{&1\}_{\{&2\}}$	$\langle msub \rangle &1&2 \langle /msub \rangle$
1.0	V	Exp	Exp	OverExp	$\frac{\{&1\}}{\{&2\}}$	$\langle mfrac \rangle &1&2 \langle /mfrac \rangle$
1.0	V	Exp	Sym	OverExp	$\frac{\{&1\}}{\{&2\}}$	$\langle mfrac \rangle &1&2 \langle /mfrac \rangle$
1.0	V	OverExp	Over	Exp	$&2$	$\langle mrow \rangle &2 \langle /mrow \rangle$
1.0	V	OverExp	Over	Sym	$&2$	$\langle mrow \rangle &2 \langle /mrow \rangle$
1.0	H	Exp	LeftPar	RPExp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	ExpOp	LeftPar	RPExp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	RPExp	Exp	RightPar	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	RPExp	Sym	RightPar	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	RPExp	OBExp	RightPar	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$
1.0	H	Exp	ExpOp	ExpOp	$&1&2$	$\langle mrow \rangle &1&2 \langle /mrow \rangle$

1.0	H	Exp	BigOpExp	Exp	&1&2	<mrow>&1&2</mrow>
1.0	H	Exp	BigOpExp	Sym	&1&2	<mrow>&1&2</mrow>
1.0	Ins	Exp	Sqrt	Exp	sqrt{&2}	<msqrt&1>&2</msqrt>
1.0	Ins	Exp	Sqrt	Sym	sqrt{&2}	<msqrt&1>&2</msqrt>
1.0	Ins	Exp	Sqrt	OBEExp	sqrt{&2}	<msqrt&1>&2</msqrt>
1.0	H	Exp	Func	Exp	&1&2	<mrow>&1&2</mrow>
1.0	H	Exp	Func	Sym	&1&2	<mrow>&1&2</mrow>
1.0	H	Func	Let	2Let	mathop{&1&2}	<mrow>&1&2</mrow>
1.0	H	Func	2Let	2Let	mathop{&1&2}	<mrow>&1&2</mrow>
1.0	H	2Let	Let	Let	&1&2	<mrow>&1&2</mrow>
1.0	V	LimExp	Lim	Exp	&1{&2}	<munder>&1&2</munder>
1.0	H	Exp	LimExp	Exp	&1{&2}	<mrow>&1&2</mrow>

Για το **Μέρος-II** χρησιμοποιήθηκαν οι ακόλουθοι κανόνες παραγωγής (σύνολο R):

P	Σχέση	S	A	B	Latex	MathML
1.0	Sup	Exp	ExpOp	SupSym	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	Sup	Exp	Sym	SupSym	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	H	Exp	Exp	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Exp	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Sym	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Sym	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	OpUn	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	OpUn	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Exp	ROpUn	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Sym	ROpUn	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Exp	OBExp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Sym	OBExp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	Sup	Exp	Sym	OBExpMinus	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	H	OBExpMinus	OpBinMinus	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	OBExpMinus	OpBinMinus	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	OBExp	OpBin	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	OBExp	OpBin	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	Sup	Exp	ExpOp	Exp	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	Sup	Exp	ExpOp	Sym	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	Sup	Exp	Sym	Exp	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	Sup	Exp	Sym	Sym	$\{&1\}^{\{&2\}}$	$\langle \text{msup} \rangle \&1\&2 \langle \text{msup} \rangle$
1.0	Sub	Exp	ExpOp	Exp	$\{&1\}_{\{&2\}}$	$\langle \text{msub} \rangle \&1\&2 \langle \text{msub} \rangle$
1.0	Sub	Exp	Sym	Exp	$\{&1\}_{\{&2\}}$	$\langle \text{msub} \rangle \&1\&2 \langle \text{msub} \rangle$
1.0	Sub	Exp	ExpOp	Sym	$\{&1\}_{\{&2\}}$	$\langle \text{msub} \rangle \&1\&2 \langle \text{msub} \rangle$
1.0	Sub	Exp	Sym	Sym	$\{&1\}_{\{&2\}}$	$\langle \text{msub} \rangle \&1\&2 \langle \text{msub} \rangle$
1.0	V	Exp	Exp	OverExp	$\text{frac}\{&1\}\{&2\}$	$\langle \text{mfrac} \rangle \&1\&2 \langle \text{mfrac} \rangle$
1.0	V	Exp	Sym	OverExp	$\text{frac}\{&1\}\{&2\}$	$\langle \text{mfrac} \rangle \&1\&2 \langle \text{mfrac} \rangle$
1.0	V	OverExp	Over	Exp	$\&2$	$\langle \text{mrow} \rangle \&2 \langle \text{mrow} \rangle$
1.0	V	OverExp	Over	Sym	$\&2$	$\langle \text{mrow} \rangle \&2 \langle \text{mrow} \rangle$
1.0	H	Exp	LeftPar	RPExp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	ExpOp	LeftPar	RPExp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	RPExp	Exp	RightPar	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	RPExp	Sym	RightPar	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	RPExp	OBExp	RightPar	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Sym	ExpOp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Exp	ExpOp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	ExpOp	ExpOp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	V	Exp	Sym	BigOpExp	$\&2^{\{&1\}}$	$\langle \text{munderover} \rangle \&2^{\{&1\}} \langle \text{munderover} \rangle$
1.0	H	Exp	BigOpExp	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	BigOpExp	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	Ins	Exp	Sqrt	Exp	$\text{sqrt}\{&2\}$	$\langle \text{msqrt} \rangle \&1\> \&2 \langle \text{msqrt} \rangle$
1.0	Ins	Exp	Sqrt	Sym	$\text{sqrt}\{&2\}$	$\langle \text{msqrt} \rangle \&1\> \&2 \langle \text{msqrt} \rangle$
1.0	Ins	Exp	Sqrt	OBExp	$\text{sqrt}\{&2\}$	$\langle \text{msqrt} \rangle \&1\> \&2 \langle \text{msqrt} \rangle$
1.0	H	Exp	Func	Exp	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$
1.0	H	Exp	Func	Sym	$\&1\&2$	$\langle \text{mrow} \rangle \&1\&2 \langle \text{mrow} \rangle$

1.0	H	Func	Let	2Let	$\mathop{\&1&2}$	$\langle\mathop{\&1&2}\rangle$
1.0	H	Func	2Let	2Let	$\mathop{\&1&2}$	$\langle\mathop{\&1&2}\rangle$
1.0	H	2Let	Let	Let	$\&1&2$	$\langle\mathop{\&1&2}\rangle$
1.0	V	LimExp	Lim	Exp	$\&1\{\&2\}$	$\langle\mathop{\&1\{\&2\}}\rangle$
1.0	H	Exp	LimExp	Exp	$\&1\{\&2\}$	$\langle\mathop{\&1\{\&2\}}\rangle$
1.0	Sup	Exp	Big2OpExp	Sym	$\&1^{\&2}$	$\langle\mathop{\&1^{\&2}}\rangle$
1.0	H	Exp	Big2OpExp	Exp	$\&1&2$	$\langle\mathop{\&1&2}\rangle$
1.0	Sup	FExp	Func	Sym	$\{\&1\}^{\&2}$	$\langle\mathop{\{\&1\}^{\&2}}\rangle$
1.0	H	Exp	FExp	Sym	$\{\&1\}^{\&2}$	$\langle\mathop{\{\&1\}^{\&2}}\rangle$
1.0	H	Exp	FExp	Exp	$\{\&1\}^{\&2}$	$\langle\mathop{\{\&1\}^{\&2}}\rangle$

Για το **Μέρος-III** και **Μέρος-IV** χρησιμοποιήθηκαν οι ακόλουθοι κανόνες παραγωγής (σύνολο R):

P	Σχέση	S	A	B	Latex	MathML
1.0	Sup	Exp	ExpOp	SupSym	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	Sup	Exp	Sym	SupSym	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	H	Exp	Exp	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Exp	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Sym	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Sym	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	OpUn	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	OpUn	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Exp	ROpUn	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Sym	ROpUn	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Exp	OBExp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Sym	OBExp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	Sup	Exp	Sym	OBExpMinus	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	H	OBExpMinus	OpBinMinus	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	OBExpMinus	OpBinMinus	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	OBExp	OpBin	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	OBExp	OpBin	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	Sup	Exp	ExpOp	Exp	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	Sup	Exp	ExpOp	Sym	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	Sup	Exp	Sym	Exp	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	Sup	Exp	Sym	Sym	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	Sub	Exp	ExpOp	Exp	$\{&1\}_{\{&2\}}$	<code><msub>&1&2</msub></code>
1.0	Sub	Exp	Sym	Exp	$\{&1\}_{\{&2\}}$	<code><msub>&1&2</msub></code>
1.0	Sub	Exp	ExpOp	Sym	$\{&1\}_{\{&2\}}$	<code><msub>&1&2</msub></code>
1.0	Sub	Exp	Sym	Sym	$\{&1\}_{\{&2\}}$	<code><msub>&1&2</msub></code>
1.0	V	Exp	Exp	OverExp	$\frac{\{&1\}}{\{&2\}}$	<code><mfrac>&1&2</mfrac></code>
1.0	V	Exp	Sym	OverExp	$\frac{\{&1\}}{\{&2\}}$	<code><mfrac>&1&2</mfrac></code>
1.0	V	OverExp	Over	Exp	$&2$	<code><mrow>&2</mrow></code>
1.0	V	OverExp	Over	Sym	$&2$	<code><mrow>&2</mrow></code>
1.0	H	Exp	LeftPar	RPExp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	ExpOp	LeftPar	RPExp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	RPExp	Exp	RightPar	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	RPExp	Sym	RightPar	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	RPExp	OBExp	RightPar	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Sym	ExpOp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Exp	ExpOp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	ExpOp	ExpOp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	V	Exp	Sym	BigOpExp	$&2^{\{&1\}}$	<code><munderover>&2&1</munderover></code>
1.0	H	Exp	BigOpExp	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	BigOpExp	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	Ins	Exp	Sqrt	Exp	$\sqrt{\{&2\}}$	<code><msqrt>&1>&2</msqrt></code>
1.0	Ins	Exp	Sqrt	Sym	$\sqrt{\{&2\}}$	<code><msqrt>&1>&2</msqrt></code>
1.0	Ins	Exp	Sqrt	OBExp	$\sqrt{\{&2\}}$	<code><msqrt>&1>&2</msqrt></code>
1.0	H	Exp	Func	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	Func	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Func	Let	2Let	$\mathop{\{&1&2\}}$	<code><mrow>&1&2</mrow></code>
1.0	H	Func	2Let	2Let	$\mathop{\{&1&2\}}$	<code><mrow>&1&2</mrow></code>
1.0	H	2Let	Let	Let	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	V	Exp	Func	Exp	$&1_{\{&2\}}$	<code><></code>
1.0	V	Exp	Func	Sym	$&1_{\{&2\}}$	<code><></code>
1.0	V	LimExp	Lim	Exp	$&1\{&2\}$	<code><munder>&1&2</munder></code>
1.0	H	Exp	LimExp	Exp	$&1\{&2\}$	<code><mrow>&1&2</mrow></code>
1.0	Sup	Exp	Big2OpExp	Sym	$&1^{\{&2\}}$	<code><msubsup>&1&2</msubsup></code>
1.0	H	Exp	Big2OpExp	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	Sup	FExp	Func	Sym	$\{&1\}^{\{&2\}}$	<code><msup>&1&2</msup></code>
1.0	H	Exp	FExp	Sym	$\{&1\}^{\{&2\}}$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	FExp	Exp	$\{&1\}^{\{&2\}}$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	OpLog	Exp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	OpLog	Sym	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	Exp	OpLog	OpBin	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	H	OBExp	OBExp	OBExp	$&1&2$	<code><mrow>&1&2</mrow></code>
1.0	Sup	RightPar	Big3ParExp	Sym	$&1^{\{&2\}}$	<code><msubsup>&1&2</msubsup></code>
1.0	Sub	Big3ParExp	BigRPar	Sym	$&1_{\{&2\}}$	<code>&1&2</code>
1.0	Sup	Exp	Big2OpExp	Exp	$&1^{\{&2\}}$	<code><msubsup>&1&2</msubsup></code>

Παράρτημα Ζ: Δημοσιεύσεις

Περιοδικά

Simistira, F., Katsouros, V. & Carayannis, G. (2015). Recognition of online handwritten mathematical formulas using probabilistic SVMs and stochastic context free grammars, *Pattern Recognition Letters*, V. 53, No 1, pages 85-92.

Συνέδρια

Simistira, F., Papavassiliou, V., Katsouros, V. & Carayannis, G. (2014). Recognition of spatial relations in mathematical formulas. In 14th International Conference on Frontiers in Handwriting Recognition, pages 164-168. Heraklion, Greece.

Simistira, F., Papavassiliou, V., Katsouros, V. & Carayannis, G. (2013). Structural analysis of online handwritten mathematical symbols based on support vector machines. In Document Recognition and Retrieval XX. San Francisco, USA.

Simistira, F., Papavassiliou, V., Katsouros, V. & Carayannis, G. (2012). A System for Recognition of On-Line Handwritten Mathematical Expressions. In 13th International Conference on Frontiers in Handwriting Recognition, pages 193-198. Bari, Italy.

Simistira, F., Katsouros, V. & Carayannis, G. (2008). A Template Matching Distance for Recognition of On-Line Mathematical Symbols. In Fotini Simistira, Vassilis Katsouros & George Carayannis (editors), 11th International Conference on Frontiers in Handwriting Recognition (ICFHR 2008). Quebec, Canada.