

Solving Sudoku with genetic operations that preserve building blocks

Αναστάσιος Βήττας 197

Οι Γενετικοί Αλγόριθμοι (GA) είναι μία μέθοδος υπολογιστικής νοημοσύνης βασισμένη στη διαδικασία της φυσικής επιλογής και της γενετικής εξέλιξης. Αναπτύχθηκαν αρχικά από τον John Holland και βασίζονται στις έννοιες της γενετικής, όπως η αναπαραγωγή, η μετάλλαξη, και η επιλογή. Οι Γενετικοί Αλγόριθμοι χρησιμοποιούνται ευρέως για την επίλυση προβλημάτων βελτιστοποίησης, ιδιαίτερα σε περιπτώσεις που οι παραδοσιακές μέθοδοι αποτυγχάνουν λόγω πολυπλοκότητας ή μεγέθους του χώρου αναζήτησης. Ένα σημαντικό τους πλεονέκτημα είναι η ικανότητά τους να προσαρμόζονται και να βρίσκουν λύσεις χωρίς να απαιτείται πλήρης γνώση του χώρου αναζήτησης.

Στους γενετικούς αλγορίθμους, οι έννοιες Γονίδιο, Χρωμόσωμα και Πληθυσμός προέρχονται από τη βιολογία και αποτελούν τα βασικά δομικά στοιχεία για την αναπαράσταση και επίλυση προβλημάτων.

- **Γονίδιο (Gene):**

Το γονίδιο αναπαριστά τη μικρότερη μονάδα πληροφορίας μέσα στη λύση. Για παράδειγμα, σε ένα πρόβλημα Sudoku το γονίδιο αποτελεί μία γραμμή του, δηλαδή μια λίστα 9 αριθμών από το 1 έως το 9.

- **Χρωμόσωμα (Chromosome):**

Το χρωμόσωμα είναι μια συλλογή γονιδίων που μαζί αποτελούν μια πλήρη λύση στο πρόβλημα. Στο παράδειγμα μας το χρωμόσωμα αντιπροσωπεύει ένα πλήρες πλέγμα Sudoku (9x9). Δηλαδή αποτελείται από 9 γονίδια.

- **Πληθυσμός (Population):**

Ο πληθυσμός είναι ένα σύνολο από πολλά χρωμοσώματα (πλέγματα Sudoku). Κάθε χρωμόσωμα αντιπροσωπεύει μία διαφορετική πιθανή λύση του Sudoku. Μέσα από διαδικασίες που θα αναλύσουμε παρακάτω, όπως η επιλογή (selection), ο συνδυασμός (crossover) και η μετάλλαξη (mutation), ο πληθυσμός εξελίσσεται, βελτιώνοντας τις λύσεις του.

Τρόπος λειτουργίας Γενετικών αλγόριθμων

1. **Πληθυσμός:** Ένα σύνολο πιθανών λύσεων (χρωμοσώματα)
2. **Καταλληλότητα (Fitness):** Συνάρτηση που αξιολογεί την ποιότητα κάθε λύσης
3. **Επιλογή:** Επιλογή των καλύτερων λύσεων για αναπαραγωγή
4. **Διασταύρωση (Crossover):** Ανταλλαγή γενετικού υλικού μεταξύ λύσεων
5. **Μετάλλαξη:** Τυχαίες μικρές αλλαγές στις λύσεις

Αρχικά, δημιουργείται ένας πληθυσμός τυχαίων χρωμοσωμάτων. Κατά τη διάρκεια της διαδικασίας, κάθε χρωμόσωμα αξιολογείται με βάση μια συνάρτηση καταλληλότητας (fitness function), η οποία καθορίζει πόσο "καλή" είναι η λύση που αντιπροσωπεύει.

Η εξέλιξη του πληθυσμού πραγματοποιείται μέσω τριών βασικών λειτουργιών: της επιλογής (selection), της διασταύρωσης (crossover) και της μετάλλαξης (mutation). Το selection διασφαλίζει ότι τα καλύτερα χρωμοσώματα του πληθυσμού έχουν μεγαλύτερη πιθανότητα να χρησιμοποιηθούν για τη δημιουργία της επόμενης γενιάς. Το crossover συνδυάζει δύο χρωμοσώματα-γονείς για να δημιουργήσει απογόνους, μεταφέροντας χαρακτηριστικά και από τα δύο γονικά χρωμοσώματα. Η μέθοδος mutation εισάγει τυχαίες αλλαγές στα χρωμοσώματα, προσφέροντας ποικιλία στον πληθυσμό και αποτρέποντας την πρόωρη σύγκλιση σε τοπικά βέλτιστα.

Solving Sudoku with Genetic Operations that Preserve Building Blocks

Το άρθρο (Yuji Sato & Hazuki Inoue, 2010) προτείνει και αξιολογεί νέες γενετικές τεχνικές για την επίλυση δύσκολων παζλ Sudoku, με έμφαση στη διατήρηση των Building Blocks, που θα αναλύσουμε παρακάτω.

Ένα Sudoku αποτελεί ένα πλέγμα 9x9 κελιών, όπου κάποια κελιά είναι ήδη συμπληρωμένα με αριθμούς από το 1 έως το 9. Ο στόχος είναι να συμπληρωθούν όλα τα κενά κελιά με αριθμούς έτσι ώστε σε κάθε γραμμή, στήλη και 3x3 υποπλέγμα να εμφανίζεται μόνο μία φορά κάθε αριθμός από το 1 έως το 9.

SUDOKU								
1		7			6	4	5	
	2	5	3	4				8
	6				1		7	
	5	3					2	9
6	1				9	8		
			6		2			7
		1		9	3	2		
		8						
4				7	8	5	9	1

ANSWER								
1	3	7	9	8	6	4	5	2
9	2	5	3	4	7	1	6	8
8	6	4	5	2	1	9	7	3
7	5	3	8	1	4	6	2	9
6	1	2	7	3	9	8	4	5
4	8	9	6	5	2	3	1	7
5	7	1	4	9	3	2	8	6
2	9	8	1	6	5	7	3	4
3	4	6	2	7	8	5	9	1

Για την επίλυση του Sudoku οι πιο συμβατικές μέθοδοι επίλυσης είναι το backtracking ή τα metaheuristics, αλλά σε πιο σύνθετα και δύσκολα Sudoku, αυτές οι μέθοδοι μπορεί να μην είναι αποτελεσματικές. Για αυτόν κυρίως τον λόγο οι ερευνητές προτείνουν τη χρήση Γενετικών Αλγορίθμων ως εναλλακτική λύση.

Μέθοδος Επίλυσης του Προβλήματος

Το βασικό πρόβλημα των Γενετικών στο Sudoku είναι η τεράστια διάσταση του χώρου αναζήτησης και η ύπαρξη πολλών τοπικών βέλτιστων λύσεων. Σύμφωνα με τους συγγραφείς υπάρχει και άλλο ένα πρόβλημα το οποίο επηρεάζει αρνητικά την ταχύτητα και την ποιότητα αποτελεσμάτων των Γενετικών Αλγορίθμων, ότι οι βασικές λειτουργίες (όπως το crossover και το mutation) καταστρέφουν τα building blocks, γεγονός που οδηγεί σε λιγότερο αποτελεσματική αναζήτηση λύσεων. Τα building blocks είναι μικρές ομάδες γονιδίων μέσα σε ένα χρωμόσωμα που έχουν υψηλή καταλληλότητα (fitness) και αποτελούν βασικές δομικές μονάδες για τη σύνθεση καλών λύσεων. Σκοπός είναι οι διατήρηση τους εφόσον αποτελούν καλές λύσεις και όχι η καταστροφή τους για να παραχθούν καινούργιες στις νέες "γενιές".

1	3	7
9		5
8	6	4

Όπως βλέπουμε στο παραπάνω σχήμα οι αριθμοί 1,3,7,9,_,5,8,6,4 αποτελούν ένα σχεδόν ολοκληρωμένο υπόπλεγμα που τηρεί τους κανόνες του Sudoku. Αν η μέθοδος crossover ή η μέθοδος mutation αλλάξουν αυτό το υπόπλεγμα τότε "καταστρέφεται" ένα building block, χάνεται μια ήδη καλή δομή.

Crossover: Εστιάζει στη διατήρηση των καλύτερων σειρών ή στηλών εντός των υποπλεγμάτων, αποφεύγοντας τη διάσπαση των building blocks. Κατά τη δημιουργία παιδιών, υπολογίζεται το fitness κάθε γραμμής ή στήλης, και τα παιδιά κληρονομούν αυτές με τις υψηλότερες βαθμολογίες (Ελιτισμός).

Mutation: Εκτελείται μόνο μέσα στα υποπλέγματα 3x3 και περιλαμβάνει την ανταλλαγή θέσεων μεταξύ δύο αριθμών που δεν είναι προκαθορισμένοι από την εκκίνηση.

Επιπλέον, οι συγγραφείς συνδιάζουν τον Γενετικό Αλγόριθμο και με την μέθοδο της Τοπικής Αναζήτησης όπου εισάγεται μια προσέγγιση και κάθε γονέας δημιουργεί πολλούς απογόνους, και επιλέγονται οι καλύτεροι από αυτούς. Η μέθοδος δηλαδή δίνει έμφαση στη διατήρηση στοιχείων υψηλής καταλληλότητας, μειώνοντας την πιθανότητα "καταστροφής" καλών τμημάτων της λύσης.

Crossover Operation: Rows and Columns Preservation

Parent 1	Parent 2	Crossover Offspring																																				
<table> <tr><td>1</td><td>2</td><td>3</td></tr> <tr><td>4</td><td>5</td><td>6</td></tr> <tr><td>7</td><td>8</td><td>9</td></tr> </table>	1	2	3	4	5	6	7	8	9	<table> <tr><td>9</td><td>8</td><td>7</td></tr> <tr><td>6</td><td>5</td><td>4</td></tr> <tr><td>3</td><td>2</td><td>1</td></tr> </table>	9	8	7	6	5	4	3	2	1	<table> <tr><td>1</td><td>2</td><td>3</td></tr> <tr><td>6</td><td>5</td><td>4</td></tr> <tr><td>7</td><td>8</td><td>9</td></tr> <tr><td>9</td><td>8</td><td>7</td></tr> <tr><td>4</td><td>5</td><td>6</td></tr> <tr><td>3</td><td>2</td><td>1</td></tr> </table>	1	2	3	6	5	4	7	8	9	9	8	7	4	5	6	3	2	1
1	2	3																																				
4	5	6																																				
7	8	9																																				
9	8	7																																				
6	5	4																																				
3	2	1																																				
1	2	3																																				
6	5	4																																				
7	8	9																																				
9	8	7																																				
4	5	6																																				
3	2	1																																				

Mutation Operation: Swap within Sub-Block
Before Mutation After Mutation

4	1	2
3	5	6
7	9	8

4	1	2
3	6	5
7	9	8