

Ευχαριστίες

Ευχαριστώ πρώτα το Θεό που μου έδωσε τη φώτιση και την αντοχή, ώστε να ολοκληρώσω αυτό το πόνημα. Θα επιθυμούσα να ευχαριστήσω θερμά τον επιβλέποντα του διδακτορικού μου, επίκουρο καθηγητή κύριο Στέγγο Δημήτριο, για τον χρόνο που διέθεσε, την ενθάρρυνσή του τις συμβουλές του, την υπομονή, καθώς και την ουσιαστική του συνεισφορά στην παρούσα εργασία. Χωρίς την συμβολή του κυρίου Στέγγου θα ήταν αδύνατη η παρουσίαση αυτής της διατριβής. Είμαι πολύ ευγνώμων για τη γενναιόδωρη υποστήριξή του, τις γόνιμες ιδέες του και την αυστηρή κριτική του που απέτρεψε σφάλματα και με οδήγησε στη σωστή κατεύθυνση.

Θα ήθελα να ευχαριστήσω θερμά επίσης τα μέλη της τριμελούς επιτροπής, καθηγητές κυρίους Μπένο Βασίλειο, και Γεωργακώδη Φώτη για την πολύτιμη βοήθεια, τα πολλαπλά σχόλια και την δημιουργική παρέμβασή τους στην τελική διαμόρφωση της παρούσας έρευνας.

Είμαι πολύ ευγνώμων προς τα μέλη της επιτροπής κρίσης του διδακτορικού μου καθηγητές Αρτίκη Θεόδωρο, Σίσκο Ιωάννη, Χαντζηκωνσταντινίδη Στάθη και Πολίτη Κώστα για την ενθάρρυνσή τους και τον χρόνο που διέθεσαν.

Με τον κίνδυνο να ξεχάσω κάποιους φίλους με τους οποίους είχα ευχάριστη αλληλεπίδραση όλα αυτά τα χρόνια αναφέρομαι με αγάπη στους: Σαχλά Θανάση, Γκίνη Δημήτρη, Παπαιωάννου Αποστόλη, Ψαράκο Γιώργο, και Μπερσίμη Σωτήρη.

Είμαι πολύ ευτυχής να έχω μια υπομονετική και ανεκτική στα όρια της αυτοθυσίας οικογένεια, που χωρίς την υποστήριξή της δεν θα ήταν δυνατή η παρουσίαση της παρούσας εργασίας.

Λίστα – αλγόριθμων

Αλγόριθμος A1: Value-iteration για MDP, ενότητα 1.1.

Αλγόριθμος A2: Policy Iteration για MDP, ενότητα 1.1.

Αλγόριθμος A3: Αλγόριθμος του ενός βήματος, ενότητα 2.2

Αλγόριθμος A4: Αλγόριθμος των ακροτάτων σημείων, ενότητα 3.2

Αλγόριθμος A5: Για επιλογή διδακτικών μεθόδων, ενότητα 9.2

Αλγόριθμος A6: Για επιλογή διδακτικών μεθόδων, ενότητα 9.3

Ακρώνυμα

DP. Dynamic programming

δ.π. διάνυσμα πληροφορίας

f.t. finitely -transient

MDP Markov decision process

POMDP Partially observable Markov decision process

UMDP Unobservable Markov decision process

PWLC piecewise-linear-convex

SI Stochastically-increasing matrix

TP₂ Ολικά θετικός τάξεως 2

ACOE Εξίσωση βελτιστοποίησης για το μέσο κόστος ανά μονάδα χρόνου.

UB Ομοιόμορφο φράγμα

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΚΕΦΑΛΑΙΑ

Σελ.

Πρόλογος-Περίληψη..... 1 - 6

1. Εισαγωγή

Περίληψη.....7

1.1 Μαρκοβιανή διαδικασία αποφάσεων με πεπερασμένο πλήθος αποφάσεων και καταστάσεων8-16

1.2 Μερικά παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων, ιστορική αναδρομή..... 17-19

1.3 Μερικά παρατηρήσιμη Μαρκοβιανή διαδικασία αποφάσεων πεπερασμένου πλήθους καταστάσεων, περιγραφή 20-21

1.4 Μετατροπή μίας POMDP σε πλήρως παρατηρήσιμη MDP..... 22-28

1.5 Πολιτικές και κριτήρια βελτιστοποίησης για την POMDP..... 29-33

Συμπεράσματα..... 34

2.Υπολογισμός της άριστης συνάρτησης τιμών για πεπερασμένο χρονικό ορίζοντα σε μία POMDP με την μέθοδο των Smallwood – Sondik- Lovejoy.

Περίληψη..... 35

2.1 Εισαγωγή..... 36-37

2.2 Ο αλγόριθμος υπολογισμού του ενός βήματος..... 38

2.3 Προσδιορισμός του συνόλου Γ_H των λειτουργικών gradient vectors για την συνάρτηση H_u 39-41

2.4 Προσδιορισμός της υποστηρίζουσας περιοχής $R(\gamma^*, \Gamma_H)$	42-44
Συμπεράσματα.....	45

3. Ο αλγόριθμος των ακρότατων σημείων

Περίληψη.....	46
3.1 Οι διευρυμένες περιοχές.....	47-55
3.2 Η περιγραφή τού αλγορίθμου.....	56-58
3.3 Εφαρμογή τού αλγορίθμου.....	59-60
3.4 Υπολογισμός του συσσωρευμένου σφάλματος της προσέγγισης για την συνάρτηση του μέγιστου αναμενόμενου ολικού οφέλους για δοσμένο ορίζοντα.	61-62
3.5 Εφαρμογή του παραπάνω αλγορίθμου στον υπολογισμό της συνάρτησης ελαχίστου κόστους για πεπερασμένο χρονικό ορίζοντα.....	63-64
Συμπεράσματα	65

4. Αλγόριθμοι για το πρόβλημα των μερικά παρατηρήσιμων Μαρκοβιανών διαδικασιών απόφασης σε άπειρο χρονικό ορίζοντα .

Περίληψη.....	66
4.1 Προσέγγιση της άριστης συνάρτησης τιμών για άπειρο χρονικό ορίζοντα και εύρεση σχεδόν άριστων πολιτικών στα πλαίσια της επαναληπτικής μεθόδου τιμών (Value-iteration).....	67-79
4.2 Κατασκευή φραγμάτων για την βέλτιστη συνάρτηση τιμών.....	80-84
4.3 Προσεγγίσεις της άριστης συνάρτησης τιμών για άπειρο χρονικό ορίζοντα και προσδιορισμός σχεδόν άριστων πολιτικών μέσω φραγμάτων.....	85-93
Συμπεράσματα	94

5. Επαναληπτική μέθοδος πολιτικής για προβλήματα POMDP σε άπειρο χρονικό ορίζοντα

Περίληψη.....	95
---------------	----

5.1	Εισαγωγή.....	96-100
5.2	Μαρκοβιανή διαμέριση και πεπερασμένα μεταβατικές πολιτικές.....	101-108
5.3	Περιοδικές πολιτικές.....	109-115
	Συμπεράσματα	116
6. Πρόβλημα POMDP για την άριστη πολιτική αντικατάστασης συστήματος σε άπειρο χρονικό ορίζοντα στα πλαίσια της διάταξης του λόγου πιθανοφανειών \leq_L.		
	Περίληψη.....	117
6.1	Περιγραφή και υποθέσεις.....	118-120
6.2	Στοχαστικές διατάξεις στον χώρο Π	121-126
6.3	Δομικές ιδιότητες της βέλτιστης πολιτικής αντικατάστασης.....	127-128
6.4	Γεωμετρική ερμηνεία της μερικής διάταξης $\leq_L, (N=3)$	129-137
	Συμπεράσματα.....	138
7. Διερεύνηση control-limit πολιτικών σε προβλήματα αντικατάστασης συστήματος με δύο καταστάσεις δύο μηνύματα και δύο αποφάσεις.		
	Περίληψη.....	139
7.1	Πεπερασμένα μεταβατικές control-limit πολιτικές.....	140-164
7.2	Περιοδικές control-limit πολιτικές (παραδείγματα).....	165-178
	Συμπεράσματα.....	179
8. Το πρόβλημα της αντικατάστασης με βάση το κριτήριο του μέσου κόστους ανά μονάδα χρόνου.		
	Περίληψη.....	180
8.1	Το κριτήριο του μέσου κόστους ανά μονάδα χρόνου.....	181-184
8.2	Τό πρόβλημα της αντικατάστασης συστήματος στα πλαίσια της στοχαστικής διάταξης του λόγου πιθανοφανειών με το κριτήριο του μέσου κόστους ανά μονάδα χρόνου.....	185-186

8.3 Το πρόβλημα με δύο καταστάσεις.....	187-199
Συμπεράσματα.....	200

9. Εφαρμογές των POMDP σε βέλτιστες πολιτικές διδασκαλίας και μάθησης.

Περίληψη.....	201
9.1 Περιγραφή του προβλήματος επιλογής διδακτικών μεθόδων.....	202-206
9.2 Ένα μοντέλο μάθησης με δύο καταστάσεις, δύο μηνύματα και δύο αποφάσεις.....	207-215
9.3 Ένα πρόβλημα με πλήρη αβεβαιότητα στην μια διδακτική μέθοδο και ατελή πληροφόρηση στην άλλη.....	216-230
Συμπεράσματα.....	231

ΠΑΡΑΡΤΗΜΑΤΑ

A. Τα σταθερά σημεία μιας κοίλης ή κυρτής συνάρτησης.....	232-234
---	---------

ΒΙΒΛΙΟΓΡΑΦΙΑ	235-246.
---------------------------	----------