

לימוד במערכות מורכבות (049004)**גיליון תרגילים 2 – תכנות דינמי (ב)**

הגשה: 4.4.11

השאלות הבאות מתייחסות לבעיית התכנות הדינמי עם מרחב מצב ופעולה סופיים, אופק זמן אינסופי, הסתברויות מעבר $p(s'|s, a)$ ופונקצית פרס $r(s, a)$.

1. הוכיחו כי האופרטור T_γ^* הוא אופרטור כיווץ עבור $\gamma < 1$ (משפט 1 בעמוד 3.3 ברשימות).

רמז: מצאו חסם עליון וחסם תחתון עבור $(T_\gamma^*V_1)(s) - (T_\gamma^*V_2)(s)$ (ללא ערך מוחלט).

2. נתבונן בבעיית התכנות הדינמי עם מקדם היוון 1. ידוע כי קיים "מצב סופג", s_o , נטול פעולות, כך שמתקיים: $r(s_o) = 0$, $p(s_o | s_o) = 1$, וכן $p(s_o | s, a) = \beta_s > 0$ לכל מצב

$s \neq s_o$ ופעולה a .

א. רשמו את משוואת האופטימאליות עבור מודל זה. השוו למשוואה המתקבלת במודל

הרגיל (ללא מצב s_o) עם מקדם היוון קטן מ-1.

ב. הוכיחו כי למשוואה זו קיים פתרון יחיד.

3. חישוב נומרי של פונקצית ערך ומדיניות אופטימאלית: נתבונן במודל בעל שני מצבים ושתי פעולות. בחרו הסתברויות מעבר כרצונכם (שונות מאפס), פונקצית פרס "מעניינת" (לא קבועה), ומקדם היוון 0.9. יש לפרט חישובים ולבצעם ידנית במידת האפשר.

א. שרטטו דיאגרמת מעברי-מצב (state transition diagram) עבור מודל זה.

ב. עבור המדיניות הבוחרת בפעולה הראשונה בכל מצב: רשמו באופן מפורש את משוואת התכנות הדינמי, וחשבו את פונקציית הערך למדיניות זו.

ג. באמצעות אלגוריתם איטרציית-הערך, חשבו את פונקציית הערך האופטימלי. רמת הדיוק הנדרשת הינה אחוז אחד מטווח פונקציית הפרס. חשבו את המדיניות החמדנית ביחס לפונקצית הערך המקורבת שמצאתם, ובדקו באם היא אופטימאלית ע"י הצבה במשוואת האופטימאליות.

ד. בעזרת אלגוריתם איטרציית-הערך, מצא מדיניות אופטימאלית ואת פונקציית הערך האופטימאלי. יש לפרט כל שלב בחישוב.

4. תזמון שרות (scheduling): חוק $c\mu$. נתבונן ב- N עבודות המיועדות לטיפול (הרצה) על

ידי שרת יחיד. השרת רשאי לבחור בכל זמן $t = 0, 1, 2, \dots$ באחת מהעבודות שהטיפול בהן

טרם הסתיים. באם נבחרה עבודה i , אזי בהסתברות $\mu_i > 0$ הטיפול בעבודה זו יסתיים

והעבודה תעזוב את המערכת; אחרת העבודה נשארת במערכת. תהליך השרות הוא אם כן חסר זיכרון – ההסתברות לסיום הטיפול בעבודה מסוימת אינו תלוי בהיסטוריית השרות שלה. לכל עבודה i קיים מחיר המתנה $c_i > 0$, שהוא המחיר המשולם לכל יחידת זמן שבו העבודה מבלה במערכת. מטרת השרת היא למזער את תוחלת המחיר הכולל.

א. תארו את הבעייה כתהליך החלטה מרקובי. האם בעייה זו הינה מאחד מהסוגים שנדונו בכיתה? רשמו את משוואת האופטימאליות (משוואת בלמן) עבור בעייה זו.
 ב. הראו כי המדיניות האופטימאלית הינה לבחור בעבודה i (מתוך אילו שעדיין נמצאות במערכת) שעבורה $c_i \mu_i$ הינה מכסימאלית. (הדרכה: חשבו באופן מפורש את פונקציית הערך למדיניות המוצעת ע"י חישוב ישיר, והראו כי היא מקיימת את משוואת האופטימאליות).

הערה: חוק $c\mu$ הינו מהתוצאות הבסיסיות והידועות בתחום תזמון-השרות, והוא תקף בתנאים כלליים יותר מאלו שתוארו. בפרט, ניתן להכלילו למערכת עם הגעות חיצוניות של עבודות (מערכת תורים).

5. בעיית המזכירה: פתרו את הבעייה הבאה, מתוך הספר של Bertsekas.

4.18 (Optimal Termination of Sampling) www

This is a classical problem, which when appropriately paraphrased, is known as the job selection, or as the secretary selection, or as the spouse selection problem. A collection of $N \geq 2$ objects is observed randomly and sequentially one at a time. The observer may either select the current object observed, in which case the selection process is terminated, or reject the object and proceed to observe the next. The observer can rank each object relative to those already observed, and the objective is to maximize the probability of selecting the “best” object according to some criterion. It is assumed that no two objects can be judged to be equal. Let r^* be the smallest positive integer r such that

$$\frac{1}{N-1} + \frac{1}{N-2} + \dots + \frac{1}{r} \leq 1.$$

Show that an optimal policy requires that the first r^* objects be observed. If the r^* th object has rank 1 relative to the others already observed, it should be selected; otherwise, the observation process should be continued until an object of rank 1 relative to those already observed is found. *Hint*: We assume that, if the r th object has rank 1 relative to the previous $(r-1)$ objects, then the probability that it is best is r/N . For $k \geq r^*$, let $J_k(0)$ be the maximal probability of finding the best object assuming k objects have been selected and the k th object is not best relative to the previous $(k-1)$ objects. Show that

$$J_k(0) = \frac{k}{N} \left(\frac{1}{N-1} + \dots + \frac{1}{k} \right).$$