Homework 10

Theoretical Analysis of Service Stations in Steady State. Priority Queues.

<u>Submit questions</u>: Part I: 2,3,4; Part II: 2,3; Part 3: all; Part4: 6-12.

Part I. 'Anonymous Pizza" Case Study

Adapted from J.A.Fitzsimmons, M.J.Fitzsimmons *"Service Management: Operations, Strategy and Information Technology"*, Irwin McGraw Hill, 1998.

Anonymous Pizza is a delivery-only pizza service promising delivery within 40 minutes of receiving a call for an order, otherwise the customer gets $2 off the price. Anonymous employs a single pizzamaker, who is paid $10 per hour, and who can make, on average, one pizza every 3 minutes. This service time has an exponential distribution. Pizzas are placed in a large oven with a capacity for ten pizzas to bake for approximately 12 minutes. A team of six delivery persons serves the neighboring population. The travel time to deliver a pizza in the market area and then return averages 10 minutes, with an exponential distribution. The one-way travel time is equal to the half of the two-way travel time. During evening times, calls for pizza arrive randomly, on average one every 5 minutes. Pizzas are delivered one at a time to customers by drivers who use their own cars and are paid $8 per hour.

Question 1. (Solved) Draw a process flow diagram, and identify the bottleneck operation.

Briefly summarize the assumptions that you are using in order to construct a mathematical model of Anonymous Pizza activities. Use the model to answer the following questions.

Question 2. Evaluate the delivery guarantee. What is the probability of paying the $2 off on the guarantee?

Question 3. Determine the number of delivery persons required to ensure that the average waiting time for a complete ready-to-send pizza to be picked up for delivery is limited to one minute.

Question 4. Is it reasonable to change the number of pizzamakers (one currently) or drivers (six currently)? Use cost-benefit considerations.

Appendix for Part I

The following formulas can be helpful in your solution.

Sum of two exponential random variables.

Assume that $X \sim \exp(\lambda_1)$, $Y \sim \exp(\lambda_2)$, X and Y are independent and $\lambda_1 \neq \lambda_2$. Then straightforward calculations provide us with the following cumulative distribution function of $X+Y$:

$$F(x) \;=\; 1 - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 x} - \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 x}\,.$$

Sum of three exponential random variables.

Assume that $X \sim \exp(\lambda_1)$, $Y \sim \exp(\lambda_2)$, $Z \sim \exp(\lambda_3)$, X, Y and Z are independent and $\lambda_1 \neq \lambda_2 \neq \lambda_3$. Then the cumulative distribution function of $X+Y+Z$ is given by:

$$F(x) \;=\; 1 - \frac{\lambda_2\lambda_3}{(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)} e^{-\lambda_1 x} - \frac{\lambda_1\lambda_3}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)} e^{-\lambda_2 x} - \frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)} e^{-\lambda_3 x}\,.$$

Part II. M/G/n Queues.

Queueing performance for varying service-time distributions.

A team of company electricians consists of 3 workers. The team receives alarm calls every 15 minutes on average. The arrival of the calls can be modeled by a time-homogeneous Poisson process. It was found that the average service time is 40 minutes. The following four hypotheses concerning the service-time distribution are considered:

a. The service time is exponential.

b. The service time is Pareto distributed with $\alpha = 2.5$.

c. The service time is Pareto distributed with $\alpha = 1.5$.

d. The service time is lognormally distributed with σ=1.

Question 1. (Solved) Calculate the missing parameters of the above distributions (use hours as time-units).

Question 2. For each of the four distributions **a-d** calculate the average waiting time (average time between a call and the start of its service).
Use the Allen-Cunneen approximation for the M/G/3 queue, when applicable. Conjecture what happens when the approximation is not applicable.

Question 3. Assume now that the company has hired enough electricians to handle all alarm calls immediately (an M/G/∞ model). What is the *steady-state* distribution of the number of calls being served?

Appendix for Part II

1. The density and the survival function of the **Pareto** distribution with parameters $\alpha > 0$, $k > 0$, are given as :

$$f(x) = \frac{\alpha k^\alpha}{(k+x)^{\alpha+1}}, \quad x > 0; \qquad S(x) = \left(\frac{k}{k+x}\right)^\alpha, \quad x > 0.$$

For a random variable X that is Pareto distributed,

$$\mathrm{E}X = \frac{k}{\alpha - 1}; \qquad \mathrm{E}X^2 = \frac{2k^2}{(\alpha-2)(\alpha-1)}$$

(If $\alpha \leq 1$ the expectation is infinite. If $\alpha \leq 2$ the variance is infinite.)

2. For the explanations on the Allen-Cunneen approximation, see Hall Section 5.5.2 (pages 153-154). If you don't have a book you can download Section 5 from the site. (Lecture 9, Reading Packet for MJP in "Related Material".)

Part III. Priority Queues.

Problem 1. People arrive at a Xerox machine according to a Poisson process with rate one per minute. The number of copies to be made by each person is uniformly distributed between 1 and 10. Each copy requires 3 sec. Find the average number of people in queue and the average waiting time in queue when:

1. Each person uses the machine on a first-come-first-serve basis.
2. Persons with no more than 3 copies to make are given non-preemptive priority over the others.

Part IV. MJP and Phase-Type Service Times.

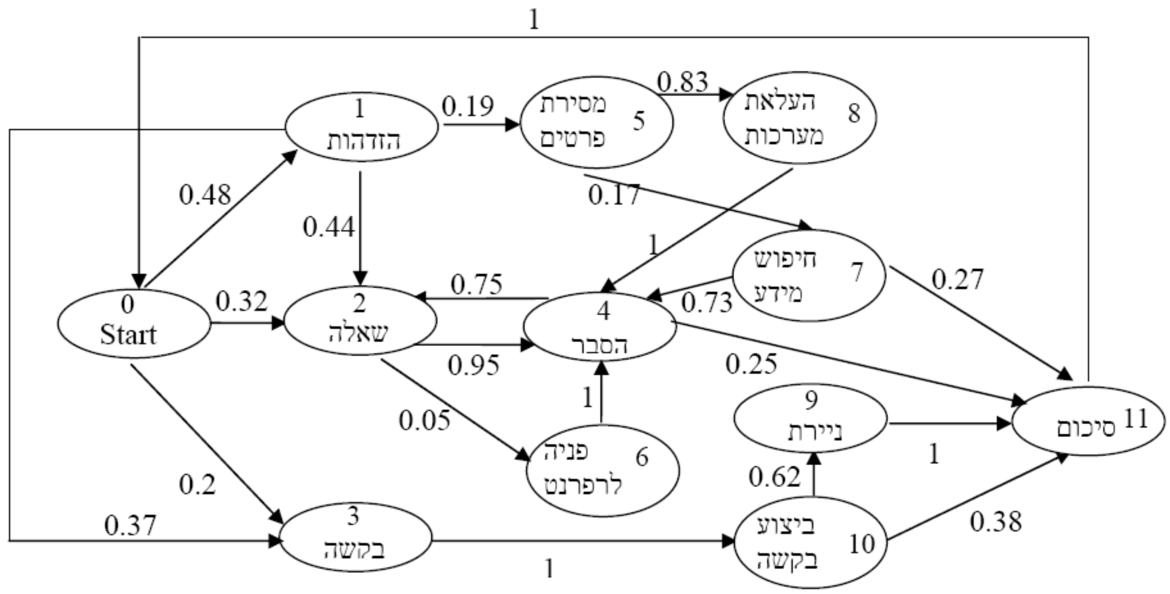
<u>חלק 4 (בנושא MJP והתפלגות Phase Type)</u>

בשנת 2007 בוצע בפקולטה לתעשייה וניהול פרויקט שנתי בשיתוף פעולה עם מוקד טלפוני של בנק ישראלי גדול. אחת ממטרות הפרויקט הייתה: חקר תהליך ביצוע שיחות מסוג שירות מסוים. על מנת לאפיין את תהליך השיחה ביחידת שירות זו, בוצעו מדידות ראשוניות על ידי האזנה למוקדנים. לאחר ניתוח השיחות, נמצא כי ניתן לחלק את השיחה לפאזות עיקריות (אלמנטים). אלמנטים אלו מתאימים לכל השיחות בשירות הנדון.

<u>האלמנטים העיקריים המרכיבים שיחה (אין צורך להתעמק בתיאור האלמנטים):</u>

1. **<u>הזדהות</u>**- אם ללקוח יש קוד סודי והוא מקיש אותו במענה הקולי אזי הזיהוי הוא אוטומאטי. אם הקוד לא הוקש, מתבקש הלקוח להזדהות מול המוקדן ולענות על מספר שאלות לאימות הנתונים.
2. **<u>שאלה</u>**- שאלה של לקוח. לדוגמא הבהרה של מכתב שהתקבל מהבנק, בירורים לגבי אופן התשלום, ריבית ועוד.
3. **<u>בקשה</u>**- בקשה של הלקוח.
4. **<u>הסבר</u>**- מתן הסבר על-ידי הבנקאי לשאלות הלקוח.
5. **<u>מסירת פרטים</u>**- מסירת פרטים על-ידי הלקוח לבנקאי, כדי לאפשר לבנקאי לבצע פעולות מסוימות, לדוגמא משלוח פקס עם פרטי הלקוח.
6. **<u>פנייה לרפרנט</u>**- אם הבנקאי לא מצא תשובה, יש באפשרותו לפנות לבנקאי ותיק יותר (הרפרנט) ולבקש ממנו עזרה. הפנייה לרפרנט **קוטעת** את השיחה.
7. **<u>חיפוש מידע</u>**- מעבר על היסטוריה של הלקוח לצורך מציאת מידע.
8. **<u>העלאת מערכות</u>**- בנקאים עובדים מול שלוש מערכות. על מנת להתחיל את תהליך הטיפול בלקוח, יש צורך לחכות ששלושת המערכות תעלנה על מסך הבנקאי. זמן המתנה זה כלול באלמנט.
9. **<u>ניירת</u>**- כתיבת הוראות במחשב כגון- הוראה למשלוח פקס ועוד. כמעט ואין ביצוע ניירת פיזית כגון שליחת מכתבים.
10. **<u>בצוע בקשה</u>**- ביצוע בקשת הלקוח על-ידי בנקאי, כוללת את כל הפעולות הנדרשות מהבנקאי, לדוגמא שליחת מכתב, ביצוע שינויים ועוד.
11. **<u>סיכום</u>**- תיעוד השיחה במחשב.

בעזרת נתוני ההאזנות, חושבו פרופורציות המעברים (הסתברויות מעבר) בין האלמנטים (המצבים). שרשרת מרקוב, כמתואר להלן, מהווה מודל לתהליך ביצוע האלמנטים על ידי מוקדן כלשהו, ממחלקת השירות שנחקרה בפרויקט.



1.1 מהי ההסתברות שבשיחה כלשהי יהיה אלמנט 7 ("חיפוש מידע")?

$$P(element\ 7\ in\ call) = 0.48 \cdot 0.19 \cdot 0.17$$

1.2 אם מוקדן נמצא עכשיו בשלב של "העלאת מערכות" (אלמנט 8), מהי ההסתברות שלא יבקר באלמנט 6 ("פניה לרפרנט") עד סוף השיחה?

$$P(\ldots) = 0.25 \cdot \left(1 + 0.75 \cdot 0.95 + \left(0.75 \cdot 0.95\right)^2 + \left(0.75 \cdot 0.95\right)^3 + \ldots\right) =$$

$$= 0.25 \cdot \frac{1}{1 - 0.75 \cdot 0.95} = 0.869565$$

נניח שההתפלגות הסטציונרית של השרשרת נתונה על ידי $\pi = (\pi_0, \pi_1, ..., \pi_{11})$. **השתמשו ב- π_iים בתשובתכם לשאלות הבאות:**

ביצוע כל אלמנט מתועד בטבלת נתונים כשורה נפרדת. כלומר אם במשך שיחה מוקדן היה פעמיים במצב "שאלה", פעמיים במצב "הסבר" ופעם אחת במצב "סיכום", אזי בטבלה יהיו 5 שורות עם נתוני שיחה זו.
(שימו לב: בכל שיחה משתתף מצב "סיכום").

1.3 כמה שורות בממוצע יכתבו בין שתי שורות המתעדות שני ביקורים סמוכים של המוקדן באלמנט "בקשה" (לא כולל השורות של האלמנט "בקשה")?

$$\frac{1}{\pi_3} - 1$$

1.4 מהו ממוצע מספר השורות המתעדות ביקור באלמנט "פניה לרפרנט" (6), מתוך 100,000 שורות שנבחרו באקראי?

$$100{,}000 \cdot \pi_6$$

1.5 אם מוקדן נמצא עכשיו בשלב של "שאלה" (2), כמה שורות (כולל השורה של המצב הנוכחי) בטבלת הנתונים יתועדו עד לסיום השיחה?

v_i - תוחלת מספר הצעדים (מספר השורות) עד לסיום השיחה אם מוקדן נמצא עכשיו במצב i.

$$v_2 = 9.2 \quad \Leftarrow \quad \begin{cases} v_2 = 1 + 0.05 v_6 + 0.95 v_4 \\ v_6 = 1 + v_4 \\ v_4 = 1 + 0.75 v_2 + 0.25 v_{11} \\ v_{11} = 1 \end{cases}$$

לאחר הגדרת האלמנטים (הפאזות) בשיחה, בוצע חקר זמנים לצורך אמידת ממוצע משכי האלמנטים והתפלגותם.

עבור כל אלמנט התקבלו הממוצעים הבאים :

מספר	תיאור אלמנט	ממוצע משך האלמנט (שניות)
1	הזדהות	46
2	שאלה	120
3	בקשה	44
4	הסבר	180
5	מסירת פרטים	64
6	פניה לרפרנט	260
7	חיפוש מידע	20
8	העלאת מערכות	46
9	ניירת	61
10	ביצוע בקשה	20
11	סיכום	10

לצורך פתרון השאלות הבאות, ניתן להניח שהתפלגויות משך כל אלמנט הן מעריכיות עם תוחלות הנתונות בטבלה (המציאות לא רחוקה מהנחה זו). בנוסף, התברר ששיחות למוקדן מגיעות לפי תהליך פואסון, בקצב של 12 שיחות לשעה. אם מוקדן עסוק, השיחות עוברות למוקדן אחר.

לאחר חקר הזמנים, משתתפי הפרויקט הגיעו למסקנה שאפשר לתאר את תהליך השיחות שמבצע מוקדן על-ידי תהליך קפיצה מרקובי $X = \{X_t, t \geq 0\}$, כאשר X_t מתאר את המצב (אלמנט, פאזה) של שיחה בזמן t .

1.6 עבור X, נסמן ב-P את מטריצה הסתברויות המעבר שלו וב- q את וקטור קצבי עזיבה. השלימו את הערכים הבאים :

$p_{12} =$ $\qquad$ $q_0 =$

$p_{11} =$ $\qquad$ $q_2 =$

$p_{26} =$ $\qquad$ $q_4 =$

$p_{24} =$ $\qquad$ $q_{11} =$

1.7 כיצד מתפלג משך שיחה? תארו במדויק ככל שתוכלו.

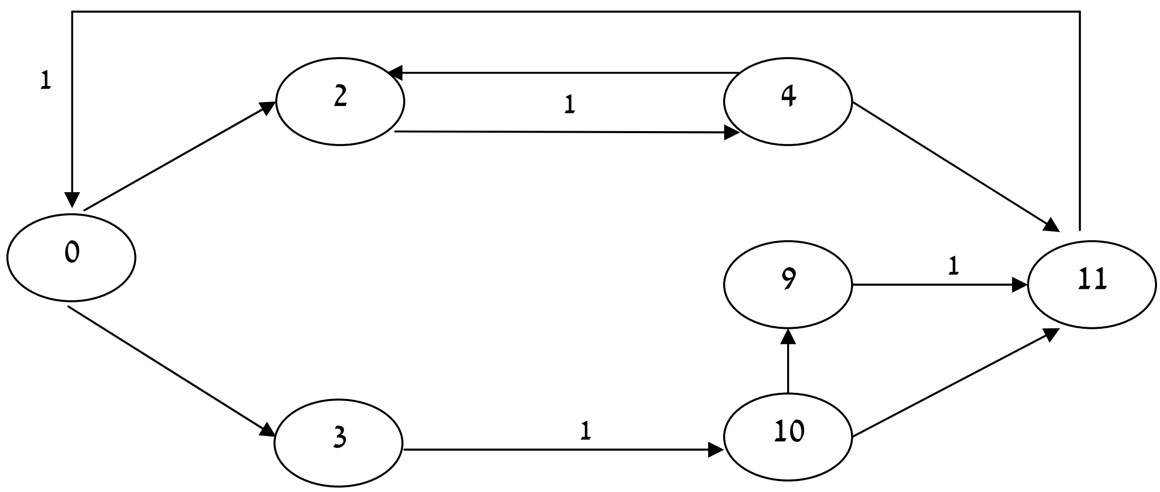
1.8 מהי תוחלת משך השיחה של לקוח שמתקשר עם "בקשה"?

1.9 מנהל משמרת מגיע למוקדן עם שאלה חשובה אך המוקדן עסוק בטיפול בלקוח. מהי תוחלת הזמן שיצטרך לחכות מנהל המשמרת עד שהמוקדן יסיים את השיחה?

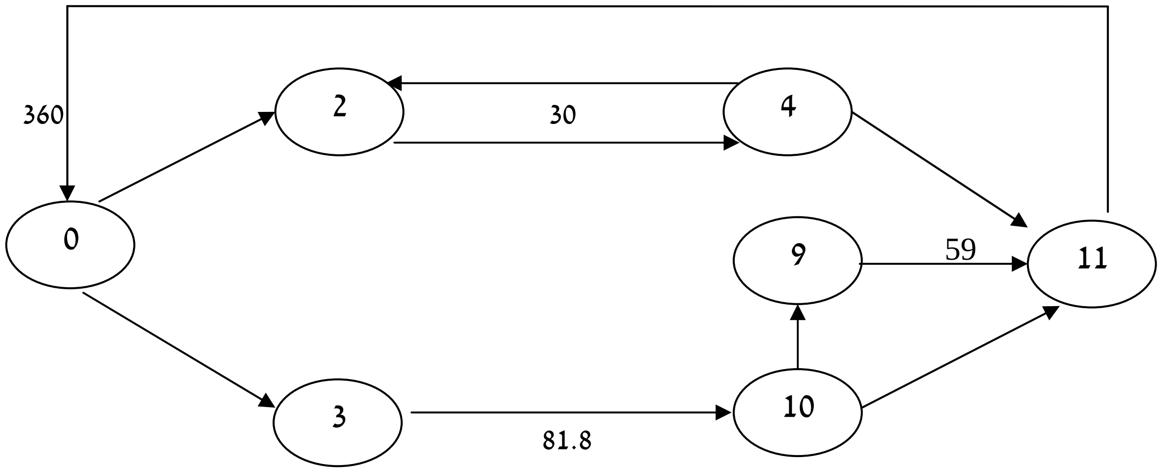
ניתוח מראה שהאלמנטים "הזדהות" (אלמנט 1) ו"פניה לרפרנט" (אלמנט 6) גוזלים זמן משמעותי. כדי למזער משך זמן זה נעשו השינויים הבאים:

- הותקנה תוכנה חדשה המאפשרת זיהוי של לקוח במענה הקולי. כתוצאה מכך הפאזה של "הזדהות" (אלמנט 1) וכל הפאזות הקשורות אליה, כגון: "מסירת פרטים" (אלמנט 5), "העלאת מערכת" (אלמנט 8) ו"חיפוש מידע" (אלמנט 7), "נעלמו" מהשיחות.
- נערכה סדנה שבה למדו המוקדנים כיצד לקבל החלטות ללא פניה לרפרנט. כתוצאה מהסדנה "נעלם" אלמנט 6 ("פניה לרפרנט") מהשיחה. בנוסף הסתבר שמספר השאלות של לקוח לאחר קבלת הסבר קטן. לכן, ההסתברות לעבור מאלמנט 4 ("הסבר") לאלמנט 2 ("שאלה") שווה עכשיו ל-0.25.
- נוספה אפשרות לקבל שירות באתר האינטרנט של הבנק. כתוצאה מכך קצב הגעת הלקוחות קטן בשליש.

1.10 השלימו את דיאגראמת המצבים של שרשרת מרקוב המתארת מעברים בין הפאזות של שיחה לפי התרחיש החדש?



1.11 השלימו את דיאגראמת קצבי המעבר של תק״מ המתאר את תהליך השיחות לפי התרחיש החדש?



1.12 איך מתפלג T – משך הזמן עד לסיום השיחה של לקוח שנמצא עכשיו בשלב של ״שאלה״ (אלמנט 2)?
מהי תוחלת ושונות הזמן הזה? הסבר.