

אוגדן בעיות מפורסמות במדעי המחשב
זמן לוגי במערכות מבוזרות
הצגת הבעיה ופתרונה על ידי לזלי למפורט
כתב: מוטי בן-ארי

תיאור הבעיה

אירוע 1: נניח שאתם שולחים הודעת דוא"ל לחבר בארה"ב, ומודיעים לו שאתם מבקשים ממכר שלכם באנגליה לשלוח לו קובץ מעניין. **אחר כך**, אתם שולחים הודעה למכר באנגליה ומבקשים ממנו לשלוח את הקובץ כאשר נוח לו. אולם, האנגלי זריז מאוד, והוא שולח את הקובץ מיידי. נוסף לכך, קווי התקשורת מישראל לאירופה ומאירופה לארה"ב לא עמוסים ולעומת זאת הקו מישראל לארה"ב עמוס מאד, ולכן הקובץ שנשלח מאנגליה לארה"ב מגיע עוד לפני שחברכם בארה"ב קיבל את הודעתכם על הקובץ שהוא עתיד לקבל. החבר בארה"ב עלול לחשוך בקובץ המתקבל כווירוס.

אירוע 2: אתם קובעים עם החבר בארה"ב לקיים שיחת צ'אט בשעה 20:00 (לפי שעון גריניץ). אולם, בשעה שקבעתם חברכם לא מגיב. ביום למחרת, אתם מבררים מה קרה ומתברר שהשעון שלו מפגר ברבע שעה.

שני אירועים אלה מדגימים את אחת הבעיות העיקריות במערכות מבוזרות, דוגמת מערכת האינטרנט, וביתר פירוט: (1) כיצד לקבוע סדר של אירועים כאשר קיימים שינויים משמעותיים בקצב העברת נתונים ועיבודים בין כלל המעבדים ורשתות התקשורת? (2) כיצד לתאם שעונים במערכת מבוזרת? הבעיות קשורות אחת לשנייה כי שעון אינו אלא דרך לספק סדרה של אירועים ("תיק-תיק-תיק").

נתאר כאן את פתרוננו של לזלי למפורט לבעיית קביעת סדר של אירועים במערכת מבוזרת. הצגת הרעיון ופתרוננו היתה נקודת מפנה בהיסטוריה של מדעי המחשב, והוביל לפיתוח של עשרות, אם לא מאות, אלגוריתמים המשמשים בכל מערכת מבוזרת.

שימו לב שיש קווים מקבילים בין הבעיות שתיארנו לבין עקרונות מתורת היחסות של אינשטיין, שטען שאי-אפשר להתעלם ממהירות האור בהגדרת מושגים כגון "סדר" ו-"בו-זמנית".

לזלי למפורט (Leslie Lamport)

לזלי למפורט נולד בניו יורק ב-1941. הוא סיים לימודי תואר ראשון במתמטיקה מוסד היוקרתי MIT, אולם וויתר על לימודי המשך לטובת עבודה כמתכנת וכמורה למתמטיקה במכללה. מאוחר יותר חזר ללימודי והשלים דוקטורט במתמטיקה באוניברסיטת Brandeis בשנת 1972. מאז ועד היום למפורט תמיד עבד במכוני מחקר Digital Equipment, SRI International, Massachusetts Computer Associates, Microsoft Research, Corporation/Compaq. מוסדות אלה עוסקים גם בבעיות יישומיות ומחקריו של למפורט הם בעלי חשיבות עליונה בפיתוח של מערכות מחשב.

לפני שנפרט את שטחי המחקר של לזלי למפורט, נזכיר שהוא מחברו של המערכות לעיבוד מסמכים LaTeX הבנוי על בסיס עבודה קודמת של Donald Knuth. מערכת LaTeX מאפשרת עיבוד תמלילים (בצורה בלתי תלויה במחשב או במערכת הפעלה) למסמכים מסובכים ביותר, ובאמצעותה מתפרסמים כל עיתונות המחקר במתמטיקה ובמדעים מדויקים.

שטח המחקר העיקרי של למפרט מתמקד ביסודות הבסיסיים של מערכות מקביליות ומבוזרות. אלגוריתם "המאפייה" שהוא פתח, פתח ערוץ מחקר פורה לבירור המשמעות המדויקת של ביצוע הוראות על ידי מעבדים מרובים בן-זמנית. בהמשך, עשרות אלגוריתמים פותחו בגלל הצורך באלגוריתמים שונים למערכות בעלות מאפיינים שונים. הנושא השני שלמפרט עסק בו הוא, פיתוח מערכות מבוזרות (כגון מערכת האינטרנט המבוזרת בין מיליוני מחשבים בעולם). הפתרון לבעיה שאנו מציגים כאן הוא מאבני הבסיס לחקר של מערכות אלו. הנושא השלישי בו התמקד למפרט הוא בעיית "הגנרלים הביזנטים" שהיא אב-טיפוס של בעיה במימוש מערכות מהמנות, כאלו שיכולים להמשיך לתפקד גם במקרה של תקלות בחלק מהמעבדים. העבודה נעשתה בחברת SRI במסגרת פיתוח מערכת מחשב מהימנות לחלליות.

למרות שלזלי למפרט מעולם לא עסק במתמטיקה לשמה לאחר סיום הלימודים, תרומותיו העיקריות מראים שהוא מתמטיקאי בנפשו שלא יסתפק אלא בהוכחות מדויקות ביותר לאלגוריתמים שהוא פיתח. חלק גדול מזמנו הוקדש לחיפוש דרכים לשפר את הדיוק של הוכחות של אלגוריתמים, במיוחד במערכות מקביליות ומבוזרות.

באתר <http://research.microsoft.com/users/lamport/pubs/pubs.html> נמצא רשימה של כל המאמרים של למפרט, ביחד עם התייחסות שלו לחשיבות ולהקשר ההיסטורי של כל מחקר ומחקר.

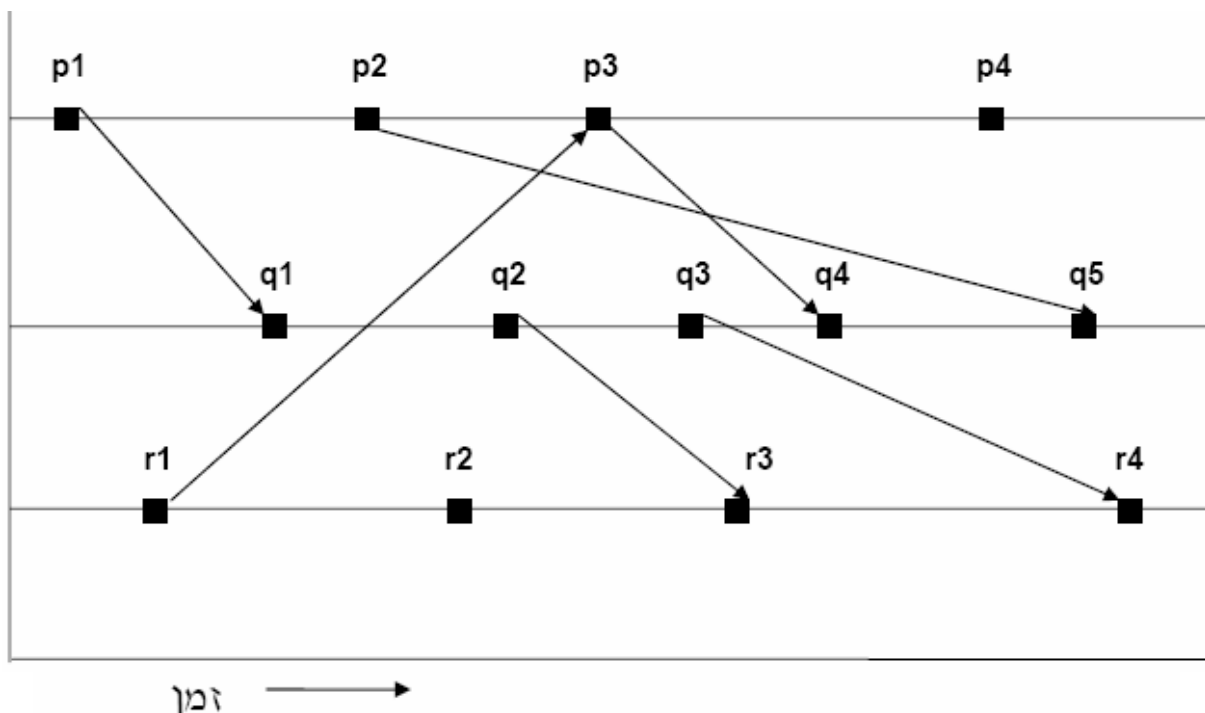
מקור נוסף:

Denis Shasha and Cathy Lazere. Out of Their Minds: The Lives and Discoveries of 15 Great Computer Scientists, NY: Copernicus, 1995, pp. 120-138.

פתרון הבעיה

התנאים לקביעת סדר

התרשים שלהלן מתאר שלושה מחשבים p,q,r. במערכת מבוזרת. בתרשים, הזמן גדל משמאל לימין, המחשבים מבצעים חישובים שונים. סמנו אירועים חשובים בכל חישוב בריבועים שחורים. חלק מהאירועים כוללים שליחת הודעות ממחשב אחד לשני, המסומן בחץ.



אנו נגדיר סדר על האירועים השונים $x \rightarrow y$, אולם סדר זה חייב למלא שלושה תנאים:

- אם האירועים x ו- y נמצאים בחישוב של מחשב אחד, ו- x קודם ל- y , אזי $x \rightarrow y$.
- אם האירוע x הוא שליחת הודעה, והאירוע y הוא קבלת אותה הודעה, אזי $x \rightarrow y$.
- אם $x \rightarrow y$ ו- $y \rightarrow z$, אזי $x \rightarrow z$.

תנאים אלה הגיוניים מאוד ובאים להבטיח שאירוע יכול להשפיע על אירוע אחר רק אם זה הגיוני מבחינת המערכת, ולא למשל, מנקודת מבט "מעלי" למערכת שלא קיימת (בדומה לדרישתו של אינשטיין שאין מקום בעולם בו ניתן לקבל מבט "מעלי"). הדרישה הראשונה אומרת שאם x מתרחש לפני y **בתוך מחשב מסוים** אזי x קודמת ל- y ו- x יכול להשפיע על y . דוגמה: p_2 בא לפני p_4 ויכול להשפיע עליו. הדרישה השנייה אומרת שמחשב אחר יכול להשפיע על מחשב אחר, רק על ידי שליחת הודעה, וההשפעה תהיה לכול המוקדם מרגע קבלת ההודעה. דוגמה: r_1 בא לפני p_3 ויכול להשפיע עליו. לבסוף, היחס "בא לפני" או "יכול להשפיע" היא טרנסיטיבית. דוגמה: r_1 בא לפני p_4 ויכול להשפיע עליו, כי $r_1 \rightarrow p_3$ ו- $p_3 \rightarrow p_4$.

אלגוריתם לקביעת סדר חלקי

קיים אלגוריתם פשוט לייחס סדר לאירועים במחשבים בשונים, כך שהסדר יתאים לתנאים שהצבנו. להלן האלגוריתם שמתבצע על ידי כל אחד מהמחשבים:

משתנה מקומי t מאותחל ל- 1
בצע עבור כל אירוע לפי סדר עולה
אין האירוע לא מקבל הודעה אזי
"יחס לאירוע זמן t "
אם האירוע כולל שליחת הודעה, צרף להודעה את הערך של t
הגדל את t באחד
אחרת { האירוע מקבל הודעה }
קבל את ההודעה, הכוללת זמן u
"יחס לאירוע את הגדול מבין u ו- t "
השם ב- t את הגדול מבין u ו- t והגדל באחד

הברירה הראשונה מבטיחה שהתנאי הראשון מתקיים, והברירה השנייה מבטיחה שהתנאי השני מתקיים. התנאי השלישי מתקיים באופן טבעי משמירת המשתנה t מאירוע לאירוע. התוצאה של ביצוע האלגוריתם על האירועים בתרשים הבא:

p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	2	3	4	2	3	4	5	6	1	2	4	5

למשל, אי-אפשר לתת את המספר העוקב 3 לאירוע r_3 כי התקבלה הודעה עם מספר 4. נבדוק למשל, ש- p_1 קודמת ל- q_2 : ואכן, p_1 קודמת ל- q_1 לפי שליחת ההודעה, ו- q_1 קודמת ל- q_2 בגלל שהם באותו מחשב. לפי עיקרון הטרנזיטיביות: $1 < 2 < 3$.

אלגוריתם לקביעת סדר מלא

אם תסתכלו על הטבלה הקודמת, תראו שלמרות שהתנאים מתקיימים, עדיין אין לנו דרך לייחס סדר גלובלי על כל המערכת, כי למשל, לכל מחשב יש אירוע שזמנו הוא 4. קביעת סדר גלובלי היא חשובה כי הוא מאפשר לתאם ביצוע משימות בין המחשבים השונים: ברגע שכל מחשב יודע מהו הסדר הגלובלי, הוא יכול לכוון את חישוביו בהתאם. נתאר כעת אלגוריתם לקביעת סדר גלובלי, כך שהתנאים שקבענו ימשיכו להתקיים. האלגוריתם יתואר ממבט "מעל" לכל המחשבים, אולם ניתן לפתח ממנו אלגוריתם מעשי למערכת מבוזרת.

משתנה t מאותחל ל- 1

בצע עד שאין יותר אירועים הזקוקים לייחוס

בחר את האירוע עם המספר הקטן ביותר

אם יש יותר מאחד, בחר את זה מהחשב המופיע ראשון בסדר האלף-בית

ייחס לאירוע הנבחר את הזמן t

הגדל את t באחד

אמנם אלגוריתם זה נותן עדיפות שרירותית למחשבים עם שמות מסוימים, אולם במערכת אמיתית, עדיפות זו אינה משמעותית.

להלן סדרה של טבלאות המתארות את ייחוס המספרים לפי האלגוריתם:

p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1									2			
p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	3			4					2	5		
p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	3	6		4	7				2	5		
p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	3	6	8	4	7	9			2	5	10	
p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	3	6	8	4	7	9	11		2	5	10	12
p1	p2	p3	p4	q1	q2	q3	q4	q5	r1	r2	r3	r4
1	3	6	8	4	7	9	11	13	2	5	10	12

בדוק שהתנאים לסדר עדיין מתקיימים!!

לזלי למפורט (Leslie Lamport)

לזלי למפורט נולד בניו יורק ב-1941. הוא סיים לימודי תואר ראשון במתמטיקה מוסד היוקרתי MIT, אולם וויתר על לימודי המשך לטובת עבודה כמתכנת וכמורה למתמטיקה במכללה. מאוחר יותר חזר ללימודי והשלים דוקטורט במתמטיקה באוניברסיטת Brandeis בשנת 1972. מאז ועד היום למפורט תמיד עבד במכוני מחקר SRI International, Massachusetts Computer Associates, Digital Equipment Corporation/Microsoft Research, Corporation/Compaq. מוסדות אלה עוסקים גם בבעיות יישומיות ומחקריו של למפורט הם בעלי חשיבות עליונה בפיתוח של מערכות מחשב.

לפני שנפרט את שטחי המחקר של לזלי למפורט, נזכיר שהוא מחברו של המערכות לעיבוד מסמכים LaTeX הבנוי על בסיס עבודה קודמת של Donald Knuth. מערכת LaTeX מאפשרת עיבוד תמלילים (בצורה בלתי תלויה במחשב או במערכת הפעלה) למסמכים מסובכים ביותר, ובאמצעותה מתפרסמים כל עיתונות המחקר במתמטיקה ובמדעים מדויקים.

שטח המחקר העיקרי של למפורט מתמקד ביסודות הבסיסיים של מערכות מקבילות ומבוזרות. אלגוריתם "המאפיייה" שהוא פתח, פתח ערוץ מחקר פורה לבירור המשמעות המדויקת של ביצוע הוראות על ידי מעבדים מרובים בן-זמנית. בהמשך, עשרות אלגוריתמים פותחו בגלל הצורך באלגוריתמים שונים למערכות בעלות מאפיינים שונים. הנושא השני של למפורט עסק בו הוא, פיתוח מערכות מבוזרות (כגון מערכת האינטרנט המבוזרת בין מיליוני מחשבים בעולם). הפתרון לבעיה שאנו מציגים כאן הוא מאבני הבסיס לחקר של מערכות אלו. הנושא השלישי בו התמקד למפורט הוא בעיית "הגנרלים הביזנטים" שהיא אב-טיפוס של בעיה במימוש מערכות מהמנות, כאלו שיכולים להמשיך לתפקד גם במקרה של תקלות בחלק מהמעבדים. העבודה נעשתה בחברת SRI במסגרת פיתוח מערכת מחשב מהימנות לחלליות.

למרות שלזלי למפורט מעולם לא עסק במתמטיקה לשמה לאחר סיום הלימודים, תרומותיו העיקריות מראים שהוא מתמטיקאי בנפשו שלא יסתפק אלא בהוכחות מדויקות ביותר לאלגוריתמים שהוא פיתח. חלק גדול מזמנו הוקדש לחיפוש דרכים לשפר את הדיוק של הוכחות של אלגוריתמים, במיוחד במערכות מקבילות ומבוזרות.

מקורות

במאמר שלהלן מופיע האלגוריתמים שתיארנו ביחד עם אלגוריתמים אחרים המבוססים עליהם. מאמר זה הוא המאמר המצוטט ביותר של לזלי למפורט, ועל חילוקי הדעות על המאמר הוא מעיר: יש כאלה שחושבים שהוא טריביאלי ויש שחושבים שהוא מזהיר! לדעתי, לעיתים קרובות, הפתרונות המזהירים הם הפשוטים ביותר.

Time, Clocks and the Ordering of Events in a Distributed System

Communications of the ACM 21, 7 (July 1978), 558-565.

Denis Shasha and Cathy Lazere. Out of Their Minds: The Lives and Discoveries of 15 Great Computer Scientists, NY: Copernicus, 1995, pp. 120-138.

מאמר זה ורשימה של כל המאמרים של למפורט, ביחד עם התייחסות שלו לחשיבות ולהקשר ההיסטורי של כל מחקר ומחקר, נמצאים באתר:

<http://research.microsoft.com/users/lamport/pubs/pubs.html>