



**מיכאל שטאל**

ארכיטקט בדיקות תוכנה באינטל, ישראל  
 עוסק בעיקר בבדיקות מערכות משובצות מחשב. במסגרת תפקידו, מיכאל מגדיר שיטות בדיקה ומתודולוגיות עבודה, עוסק בהדרכה ולפעמים אפילו מרשים לו לבדוק משהו (שזה הכי כיף).  
 מיכאל מציג תכופות בכנסים בארץ ובחו"ל ומלמד בדיקות תוכנה בפקולטה למדעי המחשב באוניברסיטה העברית. ניתן לראות חלק מהמצגות והמאמרים שלו באתר [www.testprincipia.com](http://www.testprincipia.com)



**כשטוענים שבאג שפתחתי אינו אמיתי, מבחינתי העליבו אותי**

מובן מאליו שבאג לא אמיתי גורם לאובדן זמן בתהליך הפיתוח: הזמן שהבודק משקיע בשחזור הבאג ובדיווח עליו; הזמן שמשקיעים בהבנת הדו"ח ובניתוח הבאג; הזמן שהמפתחים משקיעים בניסיונות שיחזור באג שאינו משתחרר, בתיקון משהו שאינו דורש תיקון ובמחיקת התיקון המיותר.

מעבר לעובדה שהמדד מורה על אי יעילות בעבודה, הוא גם מצביע על רמת המקצועיות של הבודקים. באג שאינו משתחרר בגלל טעות בפרטים שנכתבו בדוח מרמז על כך שהבודק לא ווידא שהבאג משתחרר על פי התיאור שהוא עצמו כתב, או שלא ציין שהבאג משתחרר רק לעיתים רחוקות. דיווח על באג לא אמיתי מראה על חוסר ידע על המוצר. באג כפול נגרם

כשהבודק לא השקיע את המאמץ לעשות חיפוש קצר ברשימת הבאגים הפתוחים על מנת להמנע מלפתוח באג כפול. לחילופין זה מצביע על כך שהכותרת שהמדווח הראשון (בודק, בדרך כלל) נתן לדוח שלו לא הכילה מילות מפתח מתאימות ולכן הבאג לא זוהה בחיפוש מהיר על הכותרות.

אני לא יודע איך אתם, אבל כשטוענים שבאג שפתחתי אינו אמיתי, מבחינתי העליבו אותי; אמרו שאני לא מקצוען. מצד אחד זה גורם לי להילחם על באגים שאני חושב שהם כן נכונים:

- אבדוק שוב אם הבאג משתחרר אצלי; אתקן תיאור לא נכון בדרך השחזור
- אם חוסר ההבנה שלי את המוצר נבע מדרישה שניתנת להבנה בשתי דרכים או מתייעוד שגוי, אתעקש שהבאג יסומן כטעות בדוקומנטציה ושהדרישות או התייעוד יתוקנו.
- אם אני חושב שהתנהגות המוצר, כפי שהוגדרה, אינה נכונה (הרי אני מייצג את הלקוח!) אתווכח עם המפתח והארכיטקט ואנסה לשנות את הדרישות.
- אם נטען שהבאג כפול, אבדוק שלא סגרו לי אותו בטענה של "שורש הבעיה זהה".

מצד שני זה עושה אותי זהיר; אם אני לא לגמרי בטוח שמשעה שראיתי הוא אכן באג, אעשה עבודה נוספת לפני שאדווח. למשל: אשאל אחרים אם אני מבין נכון את הדרישות; אבדוק אם הבאג משתחרר לא רק על ה-setup שלי.

השאיפה של קבוצת הבדיקות היא להוריד את מדד ה-RDR. השאלה היא איזו רמה של RDR היא רמה טובה. על פניו נראה ששאיפה ל-RDR של אפס היא מטרה טובה. אבל האם אכן כך? לדעתי לא. אני חושב שרמה מסוימת של RDR היא בריאה, ו-RDR ששואף לאפס מעלה חשש שקבוצת הבדיקות סובלת מבעיות חמורות לא פחות מ-RDR גבוה.

**רישיון לטעות**

ישנם הרבה מדדים למדידת האפקטיביות של קבוצת בדיקות. בכדי לתת מושג מה זה "הרבה", שווה להציץ ולקבל רעיונות במאמר הבא, שמרכז תיאור מתומצת של למעלה מ-40 מטריקות אפשריות: <https://www.swtestacademy.com/software-testing-metrics-kpi/>

בטור זה אתייחס לאחת המטריקות המוזכרות במאמר: RDR (Rejected Defect Ratio). החישוב הוא פשוט: מספר הבאגים שנסגרו כיוון שלא היו אמיתיים מחולק בסך כל הבאגים שנפתחו. מה נכנס תחת הקטגוריה של RDR? מה הוא "באג לא אמיתי"? יש שלוש קטגוריות של באגים לא אמיתיים:

- 01 באגים שלא משתחררים
- 02 באגים שאינם נכונים
- 03 באגים כפולים

**באגים שלא משתחררים (can't reproduce)**

ישנם שני סוגי באגים לא משתחררים. הסוג הראשון הוא באג שאכן קשה לשחזור. לעיתים באג נובע מאינטראקציה של כמה פרמטרים, שלחלק מהם אנחנו לא מודעים. נניח למשל שהרצנו סדרה של בדיקות ואחת מהן שינתה משתנה קונפיגורציה מערך האתחול שלו A, לערך אחר B, ללא שהיינו מודעים לתוצאת לוואי זו. הבאג קורה רק כאשר משתנה הקונפיגורציה הוא עם ערך B ומשתנה הקלט מקבל ערך C. כשנססה לשחזר את הבאג, יש סיכוי סביר שנרצה להתחיל מנקודת התחלה ידועה, ולכן נאתחל את המערכת. התוצאה: משתנה הקונפיגורציה יהיה A והבאג לא ישתחרר. במקרים אחרים הבודק אכן מצא באג, אבל מידע כלשהו חסר בתיאור הוראות השחזור (חסר צעד; ערכי הקלט שגויים; חוסר הבנה בסדר פעולות מסויים הוא קריטי להופעת הבאג). כתוצאה מכך הבאג לא משתחרר כלל. אבל בשני המקרים שתיארתי אכן יש באג במוצר! הסוג השני הוא כשלוש בשחזור הבאג כיוון שפשוט אין באג. יתכן שהבודק חשב שראה משהו אבל פשוט טעה ולא הבין מה שראה. יתכן שהמערכת עליה הבודק ביצע את הבדיקות לא הייתה תקינה: בשל רכיב חומרה תקול; בגלל דרייבר מגירסה קודמת שאינו מתאים לגירסה החדשה; או שההרשאות לכתיבה וקריאה מספריה מסויימת השתנו; וכו' וכו'. כשמנסים לשחזר באג כזה על מערכת תקינה, הוא (הפלא ופלא) לא משתחרר.

**באגים שאינם נכונים**

("by design" or "not a bug") אלה המקרים בהם הבודק חשב שהמוצר אמור להתנהג באופן A, וכיוון שההתנהגות בפועל הייתה B הבודק דיווח על באג. התעמקות בדרישות תגלה שבצעם ההתנהגות אכן אמורה להיות B; המוצר צודק והבודק, שלא ידע מספיק את הדרישות, טעה.

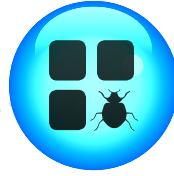


**איזה אחוז של באגים לא אמיתיים הוא טוב?**

**באגים כפולים (duplicate bugs)**

אלה באגים שמישהו כבר דיווח ומישהו אחר דיווח בשנית. **בטור קודם** התייחסתי באריכות לסוג זה של באגים, וחשוב לחזור ולומר: באג הוא כפול רק אם הסימפטומים המדווחים הם זהים. אם שורש הבעיה (root cause) זהה אבל הסימפטומים שונים, זה אינו באג כפול!





אני לא באמת חושב שיש הסבר טוב למספר הזה מעבר להרגשת בטן. הוא בטח לא מדעי. לא אלך ללויכוח גדול עם מישהו שישאף ל-10%. אבל גם הסכמה לסבול 30% וגם שאיפה ל-5% ופחות, נראות לי בעייתיות.

בסופו של דבר זו החלטה מקומית של כל קבוצת בדיקות. כל מה שאני ממליץ הוא לבדוק את ה-RDR של הקבוצה, ולחשוב אם צריך לעשות משהו. אם הוא נראה לכם גבוה, נסו לשפר אותו, אבל תזהרו של להגזים לכיוון האפס. זה עלול להביא לתופעות לוואי שאתם ודאי לא מעוניינים בהן.

על מנת להגיע ל-RDR נמוך במיוחד צריך להשקיע. כל באג לא אמיתי יחייב ניתוח וכל בודק שיפתח באג לא אמיתי יצטרך להסביר מה קרה ואיך למנוע תקלה זו בעתיד. התוצאה? בודקים ידווחו רק על באגים שהם ממש ממש בטוחים בהם. אם המוצר עושה משהו מוזר שלדעתם לא ממש נוח ללקוח, הם יעדיפו לקבל את המוזרות ולא לעמוד למשפט למה פתחו באג על מה שהוגדר כדרישה. אם יש בידיהם הוכחה שבאג אכן ארע, אבל הם לא מצליחים לשחזר אותו, הם יעדיפו לא לדווח – למה להם להכנס לכל הסיפור זה. יתכן שהם יעדיפו שלא לדווח על באגים ברמת חומרה נמוכה; הרי גם ככה באגים אלה הם בסיכוי נמוך לתיקון, ולמה להסתבך עם משהו שאולי עוד יוכרז כלא אמיתי? בקיצור: שאיפה ל-RDR נמוך מאוד מייצרת מתח לא בריא בקבוצת הבדיקות ומעלה את הסיכון שחלק מהבאגים לא ידווחו.

אנו רוצים בודקים שמצד אחד מדווחים היטב ובמדויק על באגים שראו, אבל גם מרימים דגל בכל מצב שנראה להם חשוד. בודקים שלא משקיעים שעה בלחפש האם הבאג שחוו כבר דווח מחשש שינפו אם יתברר שפתחו באג כפול, אלא מייחסים יותר חשיבות לכך שאף באג לא יברח, גם אם פה ושם יחטאו בדוח כפול. בודקים שמעדיפים לחשוב שהמילה הכתובה בארכיטקטורה או בעיצוב אינה קדושה ומותר להתנוכח עליה, גם במחיר של רמה מסויימת של באגים שבסוף יסגרו כלא אמיתיים.

אנחנו רוצים בודקים שלא מפחדים לטעות מידי פעם.

כלומר, צריך שיווי משקל כלשהו. מהו אם כן ערך RDR שכדאי לשאוף אליו? חוק האצבע שלי הוא RDR של 15%. הערך הזה מבוסס על נסיון שצברתי עם קבוצת בדיקות שנחשבה לאפקטיבית וטובה. זו הייתה רמת ה-RDR שלנו לאורך כמה פרויקטים. בקבוצה שעבדה במקביל לנו על אותו פרויקט ונחשבה לפחות אפקטיבית ופחות מכירה את המוצר, היה ערך RDR של 30%.

ITCB®  
Israeli Testing  
Certification Board  
למען  
הקהילה

כחלק מעידוד מחקרים מדעיים בתחום בדיקות התוכנה עמותת ITCB® הציעה מלגות לימודים לחוקרים מאוניברסיטת בן גוריון, מחקרים יפורסם בקרוב.

