

קומפילציה

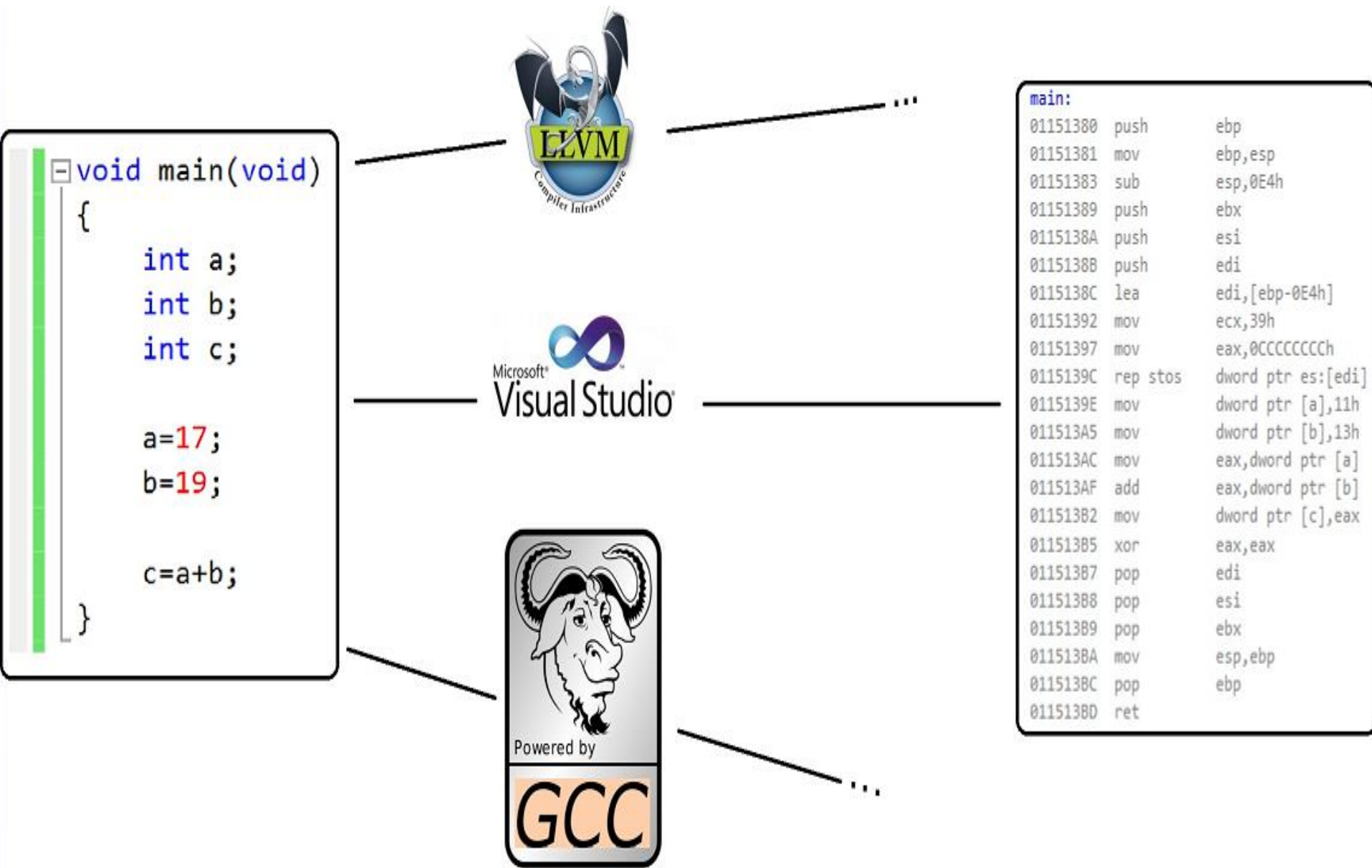
אורן איש שלום

קומפילציה הוא התהליך בו נלקחת שפת
תכנות עילית (נאמר C, ++C, JAVA וכו')
ומתורגמת לשפת מכונה, אסמבלר (x86,
MIPS וכו')

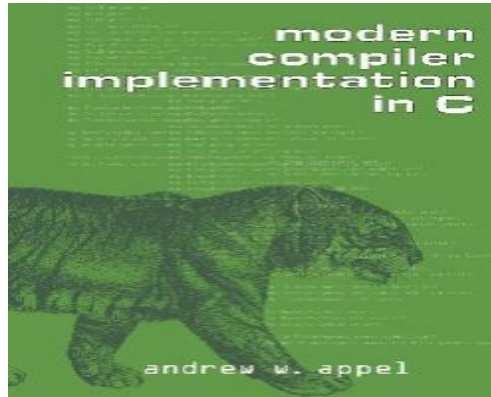
קומפילציה של תוכנית C פשוטה (משמאל) לאסמבלר x86 (מימין)

באיור מופיעים שלושה קומפיילרים נפוצים.

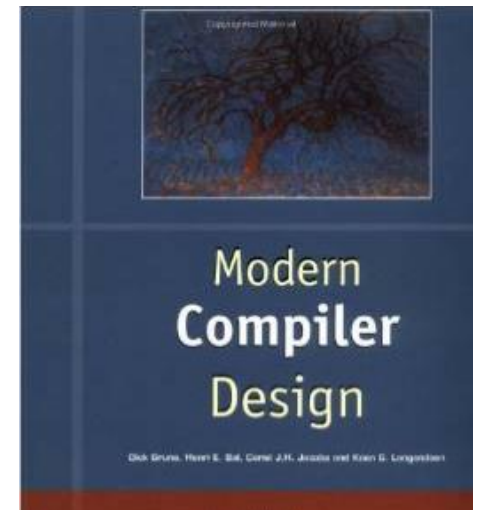
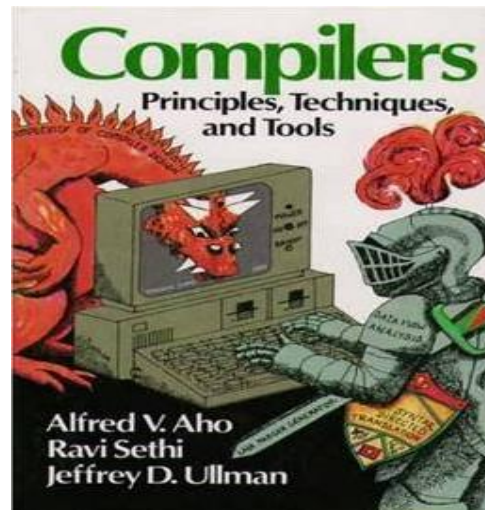
הפלט המוצג הוא של הקומפיילר האמצעי: MSVC



קומפילציה – ספר הקורס ומקורות נוספים



- ספר הקורס (להגיד משהו עליו)
- modern compiler implementation in C / Andrew W Appel
- ספרים נוספים (להגיד משהו עליהם)
- compilers: principles, techniques, and tools / Aho et al.
- modern compiler design / Grune et al.



קומפילציה – מבנה הקורס

- עמוד השדרה של הקורס הוא בניית קומפיילר אמיתי לשפת תכנות מומצאת, שיכלול את כל השלבים שמבצע קומפיילר תעשייתי
- התכנות בקורס יתבצע בשפת C (למה?)
- בנוסף, נלמד ללא מימוש, garbage collection, linkers, אופטימיזציות, ועוד.

בניית קומפיילר – היבטים תכנותיים

- הפרוייקט שלנו הוא פלטפורמה של משימת תכנות מהחיים האמתיים, וכולל: תכנון, כתיבה, התממשקות עם קוד קיים, שימוש במשאבים זמינים באינטרנט (להגיד על זה משהו), ובדיקות!
- מה עוד טוב במימוש קומפיילר? אפשר להשתמש ב references קיימים כמו gcc ו clang שהקוד שלהם פתוח, ו visual studio של מיקרוסופט (שהקוד שלו לא פתוח, אבל אפשר ללמוד הרבה מהפלט שלו)

קומפילציה אצלנו – תיאום ציפיות

- הקורס בנוי על תרגילים דו שבועיים, שמממשים קומפייילר מראשיתו ועד סופו. עבודה בצוותים.
- סביבת העבודה הרשמית בקורס היא לינוקס, אבל יש תמיכה בפיתוח על גבי ווינדוס.
- בהתאם לכישורי התכנות, התרגילים הראשונים אמורים לקחת שעות בודדות, והתרגילים האחרונים בערך יומיים.
- הפרוייקט המסכם אמור לקחת כמה ימי תכנות
- נוכחות בשיעור, השתתפות פעילה מומלצת מאוד, ותקצר משמעותית את הזמן למימוש תרגילי הבית

קומפילציה, השלבים הראשונים – Front End

- התהליך הינו רב שלבי, ודומה בשלביו הראשונים לתהליך תרגום רגיל משפה אחת לאחרת (נאמר תרגום ספר מספרדית לעברית):
- **ניתוח לקסיקלי** – וידוא שכל המילים בשפת המקור הן מילים חוקיות: מי שמצפלב אינו סרפפ
- **ניתוח סינטקטי** – וידוא שתחביר המשפט נכון: הדלי קיר בחזרה (משפט ללא פועל)
- **ניתוח סמנטי** – וידוא שהמשפט הוא בעל משמעות: הים התיכון קיבל 100 בבגרות באנגלית, או: יואב אורזת מזוודות גדולות
- אחר כך מתחיל מעבר ליצירת הטקסט בשפת היעד.

קומפילציה, השלבים הראשונים – Front End

- שימו לב, בשלושת השלבים הראשונים הנ"ל, אין התייחסות לשפת היעד – אלא רק לשפת המקור!
- אם נסתכל נניח על תרגום מעברית לאנגלית – איזה שלב מהשלושה יהיה הקל ביותר ליישום? ניתוח לקסיקלי? ניתוח סינטקטי? ניתוח סמנטי?
- מה אפשר להגיד למשל על המשפט: אישה נעלה נעלה נעלה, נעלה את הדלת בפני בעלה!
- ומה אפשר להגיד על המשפט: הייתי עלול לזכות בלוטו

קומפילציה – Back End

- אחרי שנבדקה תקינותו של הטקסט בשפת המקור, מתחילים ליצור את הקוד בשפת היעד.
- השלב הראשון בחלק זה הוא **יצירת קוד ביניים**. יצירת קוד המורכב מאבני בניין פשוטות ביותר, שמהן יהיה נוח לעבור לכל סוג של אסמבלר שהוא.
- משם ממשיכים ל**אסמבלר ללא הגבלת רגיסטרים**, כלומר, פקודות אסמבלר אמתיות, אבל ביחס לסט וירטואלי של אינסוף רגיסטרים
- בשלב האחרון, מנותח ה **flow** של התוכנית: איזו שורה יכולה להתבצע אחרי איזו שורה, ומתבצעת **הקצאת הרגיסטרים**

קומפילציה – שבעת השלבים

- ניתוח לקסיקלי (וידוא שכל המילים אכן שייכות לשפה)
- ניתוח סינטקטי (וידוא שתחביר המשפט נכון)
- ניתוח סמנטי (וידוא שהמשפט הוא בעל משמעות)
- יצירת קוד ביניים
- מעבר לאסמבלר עם אינסוף רגיסטרים
- ניתוח flow – איזו פקודה יכולה להתבצע אחרי איזו פקודה
- הקצאת רגיסטרים