LaTeX בעברית

מסמך זה מתאר כיצד להכין מסמכים עבריים או דו-לשוניים בעזרת LaTeX בעברית. עבור משתמשי PC יש במסמך זה הסבר אודות מעבד התמלילים איינשטיין.

1 LaTeX--XeT

התוכנית LaTeX בעברית מבוססת על התוכנית LaTeX הלועזית, בתוספת פקודות מיוחדות להפיכת כיוון השפה. התוכנית נקראת PUBTeX ב PC ו latex או latex--xet ב Unix. במדריך זה תיקרא התוכנית LaTeX--XeT.

קובץ DVI. שנוצר על ידי TeX--XeT הוא סטנדרטי ואינו דורש כל טיפול מיוחד.

2 hebrew.sty

hebrew.sty הוא קובץ פקודות LaTeX לתמיכה בעברית. יש לכלול אותו בהגדרת ה documentstyle בראש המסמך בין הפרמטרים שבסוגריים המרובעים לפי אחת הדוגמאות המופיעות להלן.

ח ש ו ב — כשעובדים ב PC יש לכתוב hebrew_p.

כשעובדים בעברית 8 bit יש לכתוב hebrew_newcode.

שים לב שיש לשמור את hebrew.sty תמיד מימין להגדרות הגודל וה twoside. הדבר נחוץ כדי שההגדרות שלו יהיו האחרונות בסדר, וכך הן אלו שיהיו בתוקף בזמן קריאת שאר הקובץ.

```
\documentstyle[hebrew]{report}
\documentstyle[12pt,hebrew]{article}
\documentstyle[11pt,twoside,hebrew,bezier]{book}
```

קובץ זה מגדיר את פקודות כיוון השפה שנראה בהמשך. רוב הקובץ דואג להתאמת הפקודות של LaTeX לעברית. פקודות LaTeX--XeT בזמן כתיבת אנגלית אמורות לעבוד בדיוק כפי שמוסבר ב[LAT]. אותן פקודות עם אותו תחביר עובדות בעזרת hebrew.sty גם בעברית, אך הפעולות שהן מבצעות מותאמות לכיוון הכתיבה. הזחת פסקה לדוגמה נעשית מימין בעברית. בנוסף מוגדרות בקובץ זה פקודות המיוחדות לעברית כגון מניה בספרות עבריות, תוכן ענינים דו לשוני ועוד. נסקור פקודות אלה בפרוט בהמשך.

אופציה נוספת שקיימת וקשורה ל hebrew הינה hebcal (hebcal_p כשעובדים ב PUBTeX או hebcal_newcode כשעובדים בעברית 8 bit) . אופציה זו מאפשרת קבלת תאריכים לפי הלוח העברי. ראה פרטים בהמשך.

3 פקודות בסיסיות לתמיכה בעברית בLaTeX--XeT

בסעיף זה נכיר את הפקודות הבסיסיות לתמיכה בעברית בLaTeX--XeT.

כל הפקודות של LaTeX קיימות גם בLaTeX--XeT. פקודות LaTeX--XeT מתחילות בלוכסן הפוך ואחריו סידרה של אותיות. כאשר עובדים ב Einstein הלוכסן נראה לעתים על המסך כאילו הוא נמצא מימין לפקודה.

בתחילת העבודה על הקובץ כיוון הכתיבה העיקרי הוא אנגלי.הפקודה `\sethebrew` משנה את הכיוון העיקרי לעברית. הפקודה `\unsethebrew` מחזירה את הכיוון העיקרי לאנגלית. אסור להשתמש בפקודות אלו בתוך פסקה. הן יכולות להופיע רק בין פסקאות.

להחלפת הכיוון זמנית, למספרים ולטקסט לועזי משתמשים בפקודה

\ L{...string ...}

לצורך החלפת כיוון זמנית והכנסת טקסט בעברית בתוך טקסט לועזי משתמשים בפקודה

\ R{...טקסט בעברית...}

ניתן לשנות את הכיוון בתוך קטע שבו כבר שונה הכיוון ללא הגבלת עומק הקינון לדוגמה:

He said “ תן לי 12 שקל בבקשה ”, so I did.

כאשר משתמשים ב einstein, אזי לכתיבת נוסחאות מתמטיות בתוך פיסקה עברית אין צורך במעבר שפה, המעבר לכתיבה מתמטית מחליף גם את הכיוון. יחד עם זאת, לכתיבת נוסחאות מורכבות בתוך טקסט עברי כדאי מאד לעבור סרגל כי הכיוון הטבעי בנוסחאות הוא משמאל לימין. נוסחאות ממורכזות נכתבות גם הן ללא הפיכת הכיוון, אך עם שינוי סרגל.

טקסט בתוך נוסחאות יש להקיף בקופסה מיוחדת שאומרת לLaTeX--XeT לעבור למצב טקסט, בדומה ל`\mbox{}` של LaTeX. קיימים שני סוגי קופסאות כאלו, אחד לעברית `\hmbox{}` והשני לאנגלית `\embox{}`. בכל מקרה אין להשתמש ב`\mbox{}`.

4 פקודות נוספות

לטובת משתמשים שמעדיפים לכתוב גם את הפקודות של LaTeX--XeT בעברית תורגמו לעברית מספר פקודות, וניתן להשתמש בהן באחת משתי הצורות, עברית או אנגלית. כמו כן הוגדרו מספר פקודות נוספות כדי להתגבר על הקושי שביצירת גרשיים ומקפים יפים בעברית כמו באנגלית. הנה רשימת הפקודות:

`\sethebrew`	Enter hebrew mode
`\unsethebrew`	Go out of Hebrew mode
`\hebrewpar`	Starting one hebrew paragraph with indentation
`\hebrewlines`	Starting one hebrew paragraph without indentation
`\L{...string...}`	English string within a Hebrew environment
`\R{...סידרית...}`	Hebrew string within an English environment
`cr {...}הדגש\ or \hadgesh{...}`	Making bold letters - Poor man's bold
`\מקףגדול or \makafgadol`	Like the regular -- producing – , which doesn't work in Hebrew
`\מקףענק or \makafanak`	Like the regular --- producing — , which doesn't work in Hebrew
`\גרש or \geresh`	Like the regular ‘ , which doesn't work in Hebrew
`\פתחגרש or \opengeresh`	’
`\סגורגרש or \closegeresh`	‘
`\פתחציטוט or \openquote`	Like the regular ”, which doesn't work in Hebrew
`\סגורציטוט or \closequote`	Like the regular “, which doesn't work in Hebrew
`\היום or \hebday`	Today's regular date in Hebrew
`\Hebrewtoday`	Hebrew date in Hebrew (included in hebcal.sty)
`\Hebrewdate`	Converting a Georgian date to an Hebrew (in hebcal.sty)
`\קפיצהגדולה`	Like the regular `\bigskip`
`\קפיצהבינונית`	Like the regular `\medskip`
`\קפיצהקטנה`	Like the regular `\smallskip`
`\שורהימין`	Like the regular `\rightline`
`\שורהשמאל`	Like the regular `\leftline`
`\מרכז`	Like the regular `\centerline`
`\undertext{text טקסט}`	Produces text טקסט

5 מיפוי האותיות העבריות

קיימים מספר מיפויים לאותיות העבריות ועניין זה מסבך את השימוש בתוכנה.
ב Unix קיים Old code שבו האותיות העבריות תופסות את מקומן של האותיות הלועזיות הקטנות (lower case).
ב PC האותיות העבריות נמצאות החל מהמקום ה 128 בטבלת ASCII. במידה ומעבירים קובץ קלט שהוכן ב PC לביצוע על ידי latex ב Unix, צריך להשתמש בתוכנית קונברסיה מתאימה לצורך העברה למיפוי של Unix.
ב Unix קיים גם New code בו האותיות העבריות נמצאות החל מהמקום ה 224 בטבלת ASCII. שימוש במיפוי זה הוא נוח למי שעובד ב emacs וב vi בעברית, משום שאלה מאפשרים לראות על המסך עברית ואנגלית בעת ובעונה אחת.

6 מיקוף ושורות ארוכות מדי

TEX פותר את רוב הבעיות בחלוקת הפיסקה לשורות על ידי מיקוף. הוא מוצא לרוב את כל המקומות בהם מותר לשבור מילה בסוף שורה על ידי מקף, ואינו מבצע מיקוף במקום לא חוקי. לרוב השפות הלטיניות יש קבצי מיקוף מיוחדים, בהם כלולות הוראות לתוכנה היכן רצוי לשבור מלים. לעברית עדיין אין. בנוסף, הבעיה של מיקוף דו לשוני עדיין לא נפתרה. כלומר אין אפשרות לעבוד עם יותר מקובץ מיקוף אחד לכל גירסה של TEX. ב LATEX מותקן מיקוף אנגלי כשהשפה העיקרית היא אנגלית. כשהשפה העיקרית היא עברית hebrew.sty מבוטל המיקוף כליל.
העובדה שאין מיקוף עברי יוצרת הרבה יותר מקרים בהם LATEX--XeT לא מוצא צורה טובה לחלוקת הפיסקה לשורות, כמו בגירסאות הלטיניות. LATEX--XeT משאיר במקרה זה שורות ארוכות מדי במסמך התוצאה, אך מתריע עליהן בזמן הריצה ובקובץ הlog (התראות בסגנון Overful hbox...). בניגוד למצופה, שורות אלה בולטות ימינה בדף כמו באנגלית. כפי שמומלץ ב[LAT] כדאי להשאיר את תיקון השורות האלה לגירסה האחרונה של המסמך. פרטים כיצד לתקן אותן ניתן למצוא באותו ספר.
בקצרה, הפתרונות המקובלים הם שבירת מילה עם מקף, שינוי הרגישות של LATEX--XeT למרווחים בין מילים (הפקודה `\sloppy`), סיום השורה לפני השוליים או ניסוח מחדש של הכתוב. השורה הבאה היא דוגמה לשורה ארוכה מדי:
טרנזאקציות קוסמופוליטיות אינדוקטביות ניתנות למניפולציה אגריסיבית קונסטיטוציונית.

במקרים רבים, הן בעברית והן באנגלית, השורות הארוכות נוצרות בגלל כתיבת נוסחה מתמטית בתוך שורות הפיסקה. פתרון מקובל במקרה כזה הוא להפוך את הנוסחה לנוסחה ממורכזת.

7 תוכן עניינים ורשימת תמונות

LaTeX יודע לייצר תוכן עניינים ורשימת תמונות בצורה אוטומטית. בזמן עיבוד הקבצים LaTeX מכין רשימה של הסעיפים והתמונות בקבצים נפרדים בשם toc. ו lof. בהתאמה. בריצה הבאה הוא קורא את הקבצים האלה ויוצר את הטבלאות. פירוש הדבר שבהרצה הסופית של המסמך יש להריץ את LaTeX--XeT פעמיים כדי לוודא שהרשימות מעודכנות. הפקודות הרגילות של LaTeX ליצירת הכותרת של הסעיפים הן הדואגות לעדכן את הרשימה, והפקודות

```
\starttoc{toc}
\starttoc{lof}
```

מכניסות את תוכן העניינים ורשימת הטבלאות בהתאמה אל המסמך. עד כאן אין חדש. התוספות הנדרשות לתמיכה בכתיבה דו כוונית מתוארות בהמשך.

הפקודות ליצירת כותרת סעיפים המופיעות ב[LAT] קיימות גם כאן. הכיוון של הטקסט שבתוכן הוא השפה הראשית. לכתיבת טקסט בשפה המשנית בתוך כותרות יש להוסיף protect\ לפקודת הפיכת הכיוון, לדוגמה

`\section{\protect\L{22} מלכוד}`

הכותרת תיכתב בגוף המסמך בכיוון השפה הראשית באותו מקום. לעומת זאת תוכן העניינים יכתב בכיוון השפה הראשית במקום בו מופיעה הפקודה starttoc\. אם כל המסמך נכתב באותה שפה אין צורך בהסבר נוסף. לעומת זאת אם חלק מהמסמך נכתב בכיוון ראשי שונה יש לדאוג להתאמת השפה הראשית בין פקודות הכותרת שבגוף המסמך לתוכן העניינים בעזרת התפקודות שלהלן.

לכל פקודות הסעיפים הרגילות יש גם גירסה דו כוונית. הפקודה הדו כוונית מתחילה באות b ויש לה שני פרמטרים, האחד בעברית והשני באנגלית, לדוגמה:

`\bsection{טקסט עברי}{English text}`

הכותרת בגוף המסמך תילקח לפי השפה הראשית שם. הכותרת בתוכן העניינים תילקח לפי הכיוון הראשי בזמן כתיבת תוכן העניינים. הדבר מאפשר יצירת שתי גרסאות של תוכן העניינים במסמך, אחת אנגלית שמופיעה באיזור בעל שפה ראשית אנגלית ומכילה את הגירסה האנגלית של הכותרות, והאחרת עברית המופיעה באיזור בו השפה הראשית עברית ומכילה את הגירסה העברית של הכותרות.

בזמן עיבוד הפקודה starttoc\ בנוסף ליצירת הטבלה LaTeX מוחק את הגירסה הקודמת של הקבצים, כך שבזמן המשך הריצה הקבצים נוצרים מחדש על פי הכותרות הנוכחיות. פרוש הדבר שאי אפשר להשתמש פעמיים באותו מסמך באותם קבצי עזר toc. ו lof.. לכך קיימים שני פתרונות. האחד הוא הכנסת הפקודה nofiles\ בכותרת המסמך (לפני {begin{document\ לפני הגירסה הסופית, והשני הוא יצירת עותק נוסף מהקבצים האלו בשם אחר ושימוש בו בגירסה השניה של תוכן העניינים. הפקודות הבאות (ב PC) מייצרות עותק נוסף של תוכן העניינים שניתן להדפיסו עם הפקודה {starttoc{eoc\:

```
latex myname
copy myname.toc myname.eoc
latex myname
```

8 האופציה hebcal

אופציה זו מאפשרת קבלת תאריכים לפי הלוח העברי.

אם תשתמש באופציה זו, כלומר תכתוב \documentstyle[hebrew,hebcal...]... אזי לרשותך שתי פקודות נוספות:

1. \Hebrewtoday – מדפיסה את התאריך העברי של היום.

2. \Hebrewdate – הופכת תאריך גרגוריאני לעברי. השימוש הוא כלהלן:

```
\newcount\day   \newcount\month   \newcount\year
\day=12   \month=10   \year=1884
\Hebrewdate{\day}{\month}{\year}
```

התוצאה:
כג' תשרי , תרמ'ה

ללא האופציה [hebrew] מתקבל התאריך העברי בתעתיק לועזי –
Tishrei 23, 5645

ח ש ו ב —

ב PUBTeX כתוב \documentstyle[hebrew_p,hebcal_p...]...

בעברית 8 bit כתוב \documentstyle[hebrew_newcode,hebcal_newcode...]...

9 נספח: הכנת המסמך בעזרת IBM PC

9.1 איינשטיין

איינשטיין הוא מעבד תמלילים דו-לשוני (עברי אנגלי) נפוץ. בחרנו בו מכל שאר מעבדי התמלילים הקיימים כיום משום שהוא היחידי ששומר את קובץ העבודה שלו בצורה בה LaTeX רוצה לראות את הקלט. קיימים שני סטנדרטים לשמירת אותיות של שפה משנית בכיוון כתובה הפוך בתוך קובץ. השיטה הנפוצה היא לשמור אותם לפי סדר ההופעה על המסך או על הדף. בשיטה זו האותיות בעברית יופיעו בקובץ הפוך מסדר הקריאה. לדוגמה אם נכתוב "דני" בקובץ יופיע קודם "י" ואחריו "נ" ולבסוף "ד". בציוד המחשבים האנגלי הרגיל, אם נציג את הקובץ על המסך או נדפיסו במדפסת נקבל כתב עברי קריא. השיטה השניה שומרת את כל האותיות הסדר הכתיבה והקריאה. אות שנקראת קודם תופיע קודם בקובץ. "דני" יכתב "ד" "נ" "י". בקובץ כזה העברית תראה כתובה הפוך אם נדפיס את הקובץ, אך זו צורה קלה יותר לטיפול בתוכנה. זו הצורה הדרושה לנו. כל מעבד תמלילים אחר ששומר את האותיות בצורה זו יתאים למטרתינו.

9.2 T.EXE

איינשטיין הוא יותר מעורך טקסטים פשוט. הוא מכיל פקודות ואופציות שממקמות אותו ברמה של מעבד תמלילים. פקודות כגון מירכוז, הדגשה, אינדכסים וכו. קיימות

בתוך איינשטיין, אך אלו פרימיטיביות יחסית לכלים המקבילים בTEX. לצורך ממוש פקודות אלה אישטיין משאיר בתוך הקובץ סימנים מיוחדים שמסמנים עבורו מתי ואיך להפעיל את הפקודות האלה. לפני מסירת הקובץ שנוצר בעזרת איינשטיין לTEX, יש לנפות מתוכו את הסימנים האלה.

ניתן לעשות זאת על ידי הפעלת התוכנית T.EXE המצויה על הדיסקט המקורי של einstein.

רשימת מקורות

[LAT] Leslie Lamport. LATEX *A Document Preperation system* Addison-Wesley, 1986. ISBN 0-201-15790-X