

# BiDiTeX – כיווניות למסמכי L<sup>A</sup>T<sub>E</sub>X בעברית

ארתיום בייליס (טונקיך)

11 באוקטובר 2010

## תקציר

מסמך זה תיעוד לשימוש ב-BiDiTeX – תוסף ל-L<sup>A</sup>T<sub>E</sub>X, שמספק תמיכה אוטומטית בכיווניות למסמכים בעברית. הוא תוכנן לעבוד יחד עם חבילת ivritex – החבילה שמוסיפה תמיכה בעברית ל-L<sup>A</sup>T<sub>E</sub>X. מסמך זה נבנה תוך שימוש ב-BiDiTeX ומהווה הדגמה של יכולותיו. הוא מופץ תחת רישיון ייחוס-שיתוף זהה 1.0 ישראל של Creative Commons ו-GFDL.

## תוכן עניינים

1	תמיכה בעברית ב-L <sup>A</sup> T <sub>E</sub> X – מה הבעיה?	1
2	שימוש ב-BiDiTeX	2
2	תרגום התגים	2.1
3	טיפול בתגי L <sup>A</sup> T <sub>E</sub> X	2.2
3	2.2.1 סוגי התגים	
5	2.2.2 מגבלות חשובות	
5	2.3 הפיכת סוגריים	
6	2.4 טיפול במקף	
6	3 התקנה והפעלה	3
6	3.1 התקנה	
6	3.2 הפעלה	
6	3.3 דוגמאות לשימוש	

## רשימת איורים

2	1 דוגמה לקוד של מסמך פשוט	
2	2 שימוש בתג %BIDILTR	
3	3 תגים בעברית	
3	4 תרגום תגים לאנגלית	

## רשימת טבלאות

1	רשימת התגים המתורגמים . . . . .	4
2	רשימת הסביבות המתורגמות . . . . .	4

### 1 תמיכה בעברית ב־ $\text{\LaTeX}$ – מה הבעיה?

חבילת `ivritex` מספקת תמיכה בעברית במסמכי  $\text{\LaTeX}$ . יחד עם זה, היא לא פותרת את אחת הבעיות החשובות ביותר של שימוש בעברית – תמיכה בכיווניות. הסתכלו בפסקה הבאה:

שמי ארתיום (`Artyom`), נולדתי ב־1981 בקייב (אוקראינה).

ללא התייחסות מיוחדת לכיווניות, היא תיראה כך:

שמי ארתיום (`moytrA`), נולדתי ב־1891 בקייב (אוקראינה).

על מנת לגרום לפסקה הזו להיראות כראוי יש להוסיף תגי כיווניות בצורה מפורשת ולהפוך סוגריים.

שמי ארתיום (`\Artyom`), נולדתי ב־`\1981` בקייב (אוקראינה).

כמובן "כתיבה" כזו עלולה להפוך לסיוט בגלל ערבוב של תגים באנגלית ועברית שהופכת את המלל לבלתי קריא לחלוטין. כמו כן, צורת כתיבה כזו לא מאפשרת ייבוא טקסטים רגילים בצורה פשוטה ("ע"י העתק־והדבק). אותה בעיה קיימת בהוספת מילים או משפטים בעברית לטקסטים בשפות שכתובות משמאל לימין. כל משפט חייב להיות מוגדר "ע"י תג מיוחד.

חבילת `BiDiTeX` נועדה לפתור את הבעיה – לבצע ניתוח טקסט לפני מסירתו ל־ $\text{\LaTeX}$  והוספת התגים הנדרשים על־פי אלגוריתם שהוגדר בתקן `UNICODE`.

### 2 שימוש ב־ $\text{BiDiTeX}$ .

השימוש ב־`BiDiTeX` הוא פשוט. מוסיפים חבילה `biditex` "ע"י הוספת הוראה: `\usepackage{biditex}`. כותבים קובץ  $\text{\LaTeX}$  רגיל, מסביב לכל קטע שאמור להיות מטופל "ע"י `BiDiTeX` מוסיפים תגים מיוחדים `%BIDION` להפעלה ותג `%BIDIOFF` לכיבוי – כל אחד בשורה נפרדת. בכל הטקסט שנמצא בין התגים האלו יתווספו תגי כיווניות לפי הצורך. ראה איור 1 לדוגמה.

כאשר עובדים באנגלית ומשלבים בתוכו עברית, ניתן להשתמש בתג `%BIDILTR` שנועד לטפל במקרה זה. ראה איור 2.

בנוסף ניתן להשתמש בתגים בעברית, לדוגמה אפשר לכתוב "`\הדגש{טקסט}`" במקום להשתמש בתג "`\emph{...}`". למרות ששימוש בתגים האלו הוא שקוף מבחינת המשתמש ממולץ להכיר כיצד `BiDiTeX` מטפל במלל ומוסיף תגים.

איור 1: דוגמה לקוד של מסמך פשוט.

```
\documentclass{article}
\usepackage[cp1255]{inputenc}
\usepackage[english,hebrew]{babel}
\usepackage{biditex}
%BDIDION
\begin{document}
        "שלום עולם" זה Hello World באנגלית.
\end{document}
%BDIDIOFF
```

איור 2: שימוש בתג %BDILTR.

```
%BDILTR
The meaning of word "שלום" is peace.
%BDIOFF
```

## 2.1 תרגום התגים.

BiDiTeX מאפשר להשתמש בתגים בעברית ומתרגם אותם למקביליהם באנגלית, לדוגמה: "הדגש" הופך ל-"emph", "התחל" מתורגם ל-"begin". כל תג בעברית חייב להתחיל כמו כל תג ב- $\LaTeX$  בסימן "\", או אם מדובר בסביבה (כמו enumerate), הוא צריך להופיע בתוך הגדרות בסביבה (begin או end). לדוגמה: הטקסט שמוצג באיור 3 יתורגם לטקסט שמוצג באיור 4.

איור 3: תגים בעברית

```
\סעיף{צבעים}
\התחל{מספור}
\פריט ירוק
\פריט כחול
\סיים{מספור}
\תתסעיף{משהו אחר.}
```

תמיכה בתגים בעברית מקלה על עבודה עם טקסט דו-כיווני ומגדילה קריאות של טקסט. ניתן להוסיף תגים חדשים ע"י שימוש בפקודה %BDIDICTAG עבור תגים ו-%BDIDICENV עבור סביבות. לדוגמה, פקודה mytag התגשלי %BDIDICTAG מגדירה את בתג "\התגשלי" כ-"\mytag". את הרשימה המלאה של התגים והסביבות שמוגדרים כברירת מחדל ניתן למצוא בטבלאות 1 ו-2 בהתאם.

טבלה 1: רשימת התגים המתורגמים

תג באנגלית	תג בעברית
L	ש
R	י
begin	התחל
end	סיים
title	כותרת
tableofcontents	תוכןעניינים
maketitle	צורכותרת
author	מחבר
date	תאריך
part	חלק
chapter	פרק
section	סעיף
subsection	תתסעיף
subsubsection	תתתסעיף
emph	הדגש
item	פריט, נק
label	תג
ref	ראה
footnote	הערתשוליים
cite	צטט
ldots	נקודות

טבלה 2: רשימת הסביבות המתורגמות

תג באנגלית	תג בעברית
itemize	פירוט
description	תיאור
enumerate	מספור
verse	שיר
document	מסמך
abstract	תקציר

## איור 4: תרגום תגים לאנגלית

```
\section{צבעים}
\begin{enumerate}
\item ירוק
\item כחול
\end{enumerate}
\subsection{משהו אחר }
```

## 2.2 טיפול בתגי $\LaTeX$ .

אחת הבעיות ה"קשות" עבור  $\text{BiDi}\TeX$  היא להבחין - מהו הטקסט הרגיל שיש לטפל בו ולהוסיף תגי כיווניות ומתי אסור להכניס תגים אלו. למעשה נדרשת יכולת אבחנה פשוטה בין תוכן המסמן לבין תגי העיצוב שלו. מכיוון שב- $\LaTeX$  קיים מגוון מאוד רחב של תגים שיכולים להתווסף ע"י חבילות שונות, אין זה אפשרי לגרום ל- $\text{BiDi}\TeX$  להכיר את כולם. לכן, נבחרה קבוצה קטנה של סוגי התגים שנמצאים בשימוש הרחב ביותר: `begin`, `end`, `R`, `L`, `Lnum`, `ref`, `label`, `includegraphics`, `bibitem`, `cite`.

### 2.2.1 סוגי התגים.

תגים עם טקסט - תגים נראים בצורה הבאה:

```
\tag[text 1]...{text 2}
\tag*{text 1}...{text 2}
```

כאשר  $\text{BiDi}\TeX$  חייב לטפל בטקסט שנמצא בין הסוגריים ולהוסיף תגי כיווניות במידת הצורך.

תגים עם פקודות - תגים שנראים באותה צורה כמו תגים עם טקסט, רק שאסור לטפל בתוכן שנמצא בין הסוגריים על מנת לא לפגוע בתוכן שלהן, לדוגמה:

```
\begin{tabular}{|l|l|r|}
```

בשלב זה,  $\text{BiDi}\TeX$  מכיר רק את התגים הבאים כפקודות מסוג זה - `begin`, `end`, `L`, `R`. תמיד ניתן להוסיף עוד תגים כאלה ע"י שימוש בפקודה `%BIDITAG`. לדוגמה אפשר לגרום ל- $\text{BiDi}\TeX$  להתעלם מתוכן התג "ref" ע"י הוספת השורה הבאה:

```
%BIDITAG ref
```

נוסחאות - נוסחאות שמשולבות בתוך מלל כמו:  $e = \sum_{i=0}^{\infty} \frac{1}{i!}$  שהוגדרו בין סימני "\$" או בין סימני "[\"-ר\"].  $\text{BiDi}\TeX$  מתעלם מהתוכן שלהן לחלוטין. יש לשים לב, שהנוסחה חייבה להיות כתובה בשורת טקסט אחת.

```
$e=\sum_{i=0}^{\infty}\frac{1}{i!}$
```

ולא פרוסה על מספר שורות:

$$e = \sum_{i=0}^{\infty} \frac{1}{i!}$$

במקרים שזה בלתי אפשרי לכתוב נוסחה בשורה אחת, או הנוסחה מוגדרת בתוך סביבה כמו `equation`, יש לבטל את `BiDiTeX` ע"י שימוש בתג `"%BIDIOFF"`.

## 2.2.2 מגבלות חשובות

- כל טיפול בתגים וכיווניות מתבצע שורה-שורה. לכן, יש סיכון, שהפקודה או תג `LaTeX` שיתפרס על יותר משורה אחת, יטופל בצורה לא נכונה.
- שילוב של תגי כיווניות `\L{...}` ו-`\R{...}` עלול לפגוע בסדר הסוגריים והתגים שהוצבו. לכן, במקרים מסובכים של פקודות מכוננות שהכיווניות עלולה להשתנות באחד הצדדים יש לטפל בבעיית כיווניות באופן מפורש. ייתכן ובעתיד הבעיה תטופל בצורה שלא תפגע בדקדוק של `LaTeX`.

## 2.3 הפיכת סוגריים.

`BiDiTeX` תומך בהפיכת סוגריים על-פי הדרישות של תקן `UNICODE`. הוא מכיר את כל סוגי הסוגריים שנמצאים בקידוד `ASCII` - `<`, `>`, `()`, `[]`, `{}` (שמיוצגים ע"י `\{` ו-`\}`). אין תמיכה בהפיכת סוגריים אחרים, שאמורים להיות מטופלים לפי התקן, למשל - סימן שייכות `"∈"`.

## 2.4 טיפול במקף.

חבילת `ivritex` מחליפה סימן מינוס `"-"` בסימן מקף-עליון `"-"`. אבל זה מתבצע גם עבור סידרה של שנים או שלושה מינוסים<sup>1</sup> `"--"` או `"---` שלפי כללי `LaTeX` אמורים להפוך למקף `"-"` או `"--"` (קצר או ארוך יותר). זה מקשה על הוספת סימנים אלו לטקסט. למרות זאת, `BiDiTeX` מטפל בשלב זה במקרים האלו והופך את הסימון `"-"` ל-`"-"` ואת הסימון `"--"` ל-`"-"`; כאשר סימן מינוס רגיל `"-"` עדיין הופך למקף-עליון `"-"`. מכיוון שהפתרון הזה אינו פתרון אמתי אלא מעקף, החלפת סימני מינוס במקף לא מתבצעת כברירת מחדל אלא מופעלת ע"י מסירת פרמטר `"-m"` ל-`biditex`. מכיוון שגופן ברירת מחדל של `LaTeX` לא נראה מתאים להצגת קרמפריד בתוך הטקסט העברי, ניתן להגדיר גופן אחר עבור הצגת `"קרמפריד"` או `"טווח-מספרים"` ע"י הוספת פקודה:

$$\backslash\text{bididashfont}\{\text{fontname}\}$$

<sup>1</sup>הבעיה לא קיימת כשמשתמשים בחבילת ניקוד - `nikud.berlios.de` ומשלבים סגנון/חבילה `culmus`. יש לשים לב, שהפתרון המוצע הוא חלקי, בגלל שהגופן המשמש להצגת סימן `"-"` אינו עברי, לכן המראה של הסימן לא טבעי כשמשתמשים בחבילת `ivritex` כפי-שהיא.

## 3 התקנה והפעלה.

### 3.1 התקנה.

הוראות התקנה ניתן למצוא בקובץ README שבא יחד עם קוד המקור. הקוד נבדק על שתי פלטפורמות: Debian Etch ו־Ubuntu Dapper יחד עם ivritex 1.2.1-1 ו־nikud, אך צפוי לעבוד בכל פלטפורמה שניתן להשתמש בה ב־libfribidi.

### 3.2 הפעלה.

כפי שהוסבר לעיל, יש להריץ BiDiTeX לפני הפעלת L<sup>A</sup>T<sub>E</sub>X. לדוגמה:

```
biditex input.tex -o tmp.tex
```

רק אז אפשר להפעיל את L<sup>A</sup>T<sub>E</sub>X על קובץ tmp.tex. פירוט מלא של פרמטרים ניתן למצוא בתיעוד man.

### 3.3 דוגמאות לשימוש.

מסמך זה, יחד עם המקור שלו מהווים הדגמה מקיפה לשימוש ב־BiDiTeX, יחס עם זו, הוא מאוד מסובך להתחלה מכיוון שמשתמש בהרבה כלים שקיימים ב־L<sup>A</sup>T<sub>E</sub>X ומשלב כמות גדולה של טקסטים בעברית ואנגלית. התקנה של BiDiTeX מספקת מסמך (example.tex) שמדגים שימוש בסיסי ב־BiDiTeX.