עקרונות שפות תכנות, סמסטר ב' – 2016

תרגול 6 : Lazy-Lists, Continuation Passing Style (CPS)

**חלק א' – Lazy-Lists**

תזכורת: רשימות עצלות הן מבני נתונים סדרתיים המאפשרים דחייה של חישוב ושמירה של איברים מתוכם.

היתרונות בשימוש בהן:

* אין צורך לאחסן בזיכרון את כל איברי הרשימה. בעזרת רשימות עצלות ניתן לייצג אף סדרות אינסופיות.
* דחיית חישוב איברים ברשימה לזמן בו נדקק להם – ייתכן שלא נזדקק לכל איברי הרשימה.

נגדיר את ערכי הטיפוס של רשימות עצלות באופן רקורסיבי:

ADT עבור רשימות עצלות (המימוש מופיע בספר הקורס):



## שאלה 1 – השערת קולץ

נגדיר את הפונקציה:

השערת קולץ גורסת כי לכל מתקיים כי הסדרה תמיד מתכנסת ב-1. למשל עבור n=563:

563 -> 1690 -> 845 -> 2536 -> 1268 -> 634 -> 317 -> 952 -> 476 -> 238 -> 119 -> 358 -> 179 -> 538 -> 269 -> 808 -> 404 -> 202 -> 101 -> 304 -> 152 -> 76 -> 38 -> 19 -> 58 -> 29 -> 88 -> 44 -> 22 -> 11 -> 34 -> 17 -> 52 -> 26 -> 13 -> 40 -> 20 -> 10 -> 5 -> 16 -> 8 -> 4 -> 2 -> 1

נרצה להגדיר את הרשימה העצלה המכילה את סדרת קולץ עבור n כלשהו.



מה הטיפוס של lzl-collatz? האם הפרוצדורה היא well-typed? האם (head (lzl-collatz n)) עבור n טבעי כלשהי **תמיד** תצליח?

## שאלה 2 – סדרת ההפעלות העצמיות

בהינתן פונקציה וערך , נגדיר את סדרת ההרכבות העצמיות להיות . הסדרה יכולה להיות סופית, או אינסופית. נרצה לייצר את הרשימה העצלה שאיבריה הם איברי הסדרה. הרשימה תיגמר כאשר הגיעה לנקודת שבת, כלומר כאשר הערך המוחזר מהפעלת הוא אותו הערך שעליו הופעלה.



**חלק ב' – Continuation Passing Style (CPS)**

Continuation Passing Style: פרדיגמה תכנותית בה לכל פרוצדורת משתמש ניתן פרמטר נוסף שהינו פרוצדורה, אשר מהווה את המשך החישוב שיש לבצע בתום פעולת פרוצדורת המשתמש (ועל כן נקרא הפרמטר continuation). חלק מן המוטיבציה לשימוש בפרדיגמת תכנות זו היא השליטה שהיא מאפשרת בבקרה של התוכנית:

1. במהלך כתיבת פרוצדורות משתמש בשיטת CPS, כל קריאה רקורסיבית נכתבת כך שהיא בעמדת זנב. לכן, מעצם השימוש בשיטה, הפרוצדורה תייצר תהליך איטרטיבי.
2. CPS מאפשר לשלוט בסדר לפיו התוכנית תחשב ביטויים. למשל, אם נממש חיפוש בעץ, נוכל לבחור אם לחפש קודם בענף הימני או בשמאלי (לעומת זאת, ללא CPS נהייה תלויים בסדר הערכת הביטויים המובנה ב-interpreter).
3. השליטה בבקרת התוכנית מאפשרת לנקוט בפעולה שונה במקרה של הצלחה / כישלון או יציאה ישירה (exception) במהלך רקורסיה (או איטרציה) וכן מאפשרת להחזיר כמה ערכים ביחד.

## שאלה 1 – אופן פעולת פרוצדורה בגרסת CPS והוכחת נכונות

ניזכר במימוש של הפרוצדורה fact ללא שימוש ב-CPS:



כעת, נממש את fact בשיטת CPS ונמחיש את אופן הפעולה שלה:



**דיון**: מדוע התוצאות אינן זהות? כיצד נרצה לקבוע בדרך כלל את פרוצדורת ה-continuation שנעביר לקריאה הראשונה?

**הוכחת נכונות:** לאחר המרת פרוצדורה לגרסת ה-CPS שלה, נדרשת הוכחת נכונות. הנכונות נשענת על הצגת שקילות בין שתי הגרסאות, במובן הבא: בהפעלת האלגוריתם applicative-eval על קריאה מתאימה לכל אחת מן הגרסאות, החישוב יסתיים ויוחזר ערך זהה. לשם כך, נזדקק להגדרת השקילות הבאה:

**הגדרה**: פרוצדורה f$ שקולה-CPS לפרוצדורה f אם לכל ערכי קלט x1,x2,…,x3 ולכל continuation המסומן cont, מתקיים:



**טענה**: הפרוצדורה fact$ שקולה-CPS לפרוצדורה fact. כלומר, לכל מספר טבעי, n, ולכל continuation, אשר יסומן על ידי c, מתקיים: 

**הוכחה:** כיוון שהפרוצדורה fact היא רקורסיבית, ההוכחה מתבצעת על ידי שימוש באינדוקציה.

**בסיס האינדוקציה**:



**הנחת האינדוקציה**: עבור הטענה מתקיימת לכל . כלומר



**צעד האינדוקציה**: יהא , אזי:



מהנחת האינדוקציה, נקבל:



**כללי אצבע להמרת פרוצדורה רקורסיבית לפרוצדורה שקולה-CPS:**

בהינתן פרוצדורה רקורסיבית, f, עם הפרמטרים הפורמאליים x1,…,xn, פרוצדורה שקולה-CPS לה תסומן f$ ותופעל על ארגומנט נוסף, c, ה-continuation. העקרונות שינחו אותנו בכתיבת f$ הם כדלקמן:

1. נפעיל את c על ערכים מוחזרים (כחלק ממקרה הבסיס, או בתוך continuation הנוצר במהלך הריצה).
2. נקפיד לכתוב כל קריאה לפרוצדורת משתמש בתוך f$ כך שהקריאה תהיה בעמדת זנב.
3. כתיבת ה-continuation:
   1. כשנבצע בתוך f$ קריאה רקורסיבית כלשהי עם ארגומנטים x1,…,xn, נתבסס על הנחת האינדוקציה (המקטינה את הבעיה). כלומר, נניח כי התשובה עבור הבעיה המוקטנת מתקבלת כארגומנט של  
      ה-continuation.
   2. בכתיבת ה-continuation, נקיים את צעד האינדוקציה ונפעל על הפתרון עבור הבעיה המוקטנת כדי לקבל את הפתרון לבעיה הגדולה יותר.
4. בכתיבת פרוצדורה בשיטת CPS, נקפיד שכל פרוצדורת משתמש (לא פרימיטיבית) המשמשת את f$ תהיה גם היא כתובה בשיטת CPS.
5. בדרך כלל, לא נרצה לבצע מניפולציה על הערך המוחזר של הפונקציה ולכן נשתמש בפונקציית הזהות id  
   כ-continuation ההתחלתי.

**שאלה 2 – פונקציות CPS מסדר גבוה**

נמיר את הפונקציה foldr ל-CPS. תזכורת:





**נשים לב**: מדוע לא יכולנו להפעיל את c על תוצאת op$ במקום להעביר לה את c כארגומנט? ראשית, לו עשינו כך, הקריאה לא הייתה בעמדת זנב. בנוסף, הבעיה המהותית יותר היא שבמקרה כזה היינו מאבדים את השליטה על תהליך החישוב המגולמת בתוך ה-continuation.

**שאלה 3 – שימוש במספר continuations**

לצורך הפשטת העבודה מול רשימות, נשתמש ב-ADT עבור עצים (המימוש מופיע בספר הקורס). כפי שכבר ראינו בעבר עבור רשימות ורשימות עצלות, גם עץ מוגדר בצורה רקורסיבית:

1. Empty-Tree הוא עץ
2. אם t עץ, אזי לכל e: הערך של (add-subtree (make-leaf e) t) הוא עץ

ממשק ה-ADT:



הפונקציה replace-leaves$ מקבלת עץ הומוגני, פרדיקט (שאינו כתוב בגרסת CPS), ביטוי new מטיפוס עלי העץ ושני continuations: אחד עבור הצלחה והשני עבור כישלון. המטרה היא להחליף את כל העלים אשר מקיימים את הפרדיקט, בביטוי new. נשים לב שניתן להכליל את הקוד ע"י העברת פרוצדורה בתור new ולשנות את העלה בצורה דינמית.



**הערה:** יכולנו במקום succ ו-fail להעביר פרוצדורות שמבצעות מניפולציות אחרות על התוצאה – כל זאת מבלי לשנות שורה אחת בקוד הקיים של replace-leaves$. לדוגמה:



**שאלה 4 – continuation עם מספר ארגומנטים**

פרוצדורות ה-CPS שאנו יוצרים, למעשה אינן בעצמן מחשבות את התוצאה, אלא מייצרות (על ידי יצירת continuations) פרוצדורה המהווה את תהליך החישוב. אם נקפיד לייצר continuations אשר מקבלים מספר ארגומנטים ובאופן עקבי נפעיל אותם על אותו מספר של ארגומנטים, נוכל לומר שתהליך החישוב שנייצר "מחזיר" מספר ערכים.

נראה להלן פרוצדורה בה תכונה זו שימושית. הפרוצדורה מפצלת רשימה נתונה לשתי רשימות לפי פרדיקט פרימיטיבי, pred. ה-continuation מקבל שני ארגומנטים: רשימת האיברים מן הבעיה המוקטנת אשר קיימו את הפרדיקט ורשימת אלו שלא. על שתי הרשימות יתבצע המשך החישוב.



**דוגמאות נוספות לעיון**

# שאלה 1 – "7 בום"

נייצר את הרשימה העצלה שאיבריה מתאימים לחוקי המשחק 7 בום:



**שאלה 2 – רקורסיה הדדית עם CPS**



כאשר נרצה להמיר פונקציות רקורסיביות הדדיות ל-CPS, נמיר את כולן ל-CPS.



**נשים לב:** על שתי הפונקציות להיות כתובות בצורת CPS. אם היינו משתמשים ב- even?$יחד עם odd שאינה כתובה כ-CPS (ובהנחה ש-odd שוב קוראת ל-even?$), היינו "זורקים" את ה-continuation שהצטבר בכל פעם שהיינו קוראים ל-even?$ מתוך odd (כיוון ש-odd אינה כתובה בצורת CPS, היא אינה מקבלת continuation כארגומנט ולכן גם לא מעבירה continuation).

**שאלה 3 – גרסת CPS עבור פרוצדורה למציאת מחלק משותף מקסימאלי (gcd)**

נתחיל בבחינת הפרוצדורה בה תהליך החישוב הוא איטרטיבי:



כתיבת גרסת ה-CPS היא מידית: ה-continuation יוכל להישאר id (כלומר ללא שינוי) לאורך כל החישוב. זאת, כיוון ש-n משמש כ-"צובר" לתשובה במקרה זה. נבחין כי למעשה איננו מרחיבים את הפונקציה "הנשלחת" כ - continuation במהלך ריצת gcd$:



**שאלה 4 – גרסתCPS עבור פרוצדורה לחישוב אורך רשימה**

נבחן תחילה את המימוש של length, המחשבת את אורכה של רשימה:



ובגרסת CPS:



הרעיון הוא להניח כי אל המשתנה cdr-length "יוכנס" האורך של זנב הרשימה (זאת כיוון שקראנו רקורסיבית לפונקציה length$ עם (cdr lst)). כלומר אנו מניחים כי cdr-length חושב וכל שנותר הוא להפעיל את ה-continuation על התוצאה שהיא אורך הרשימה כולה (אורך זנב הרשימה + 1). נבחין כי בכל שלב באיטרציה נוצרת פונקציית continuation חדשה (מורחבת) פרט לשלב האחרון (מקרה הבסיס).

נשים לב: השימוש ב-CPS אינו חוסך באופן כללי את הזיכרון הנדרש, אלא רק חוסך את השימוש במקום על המחסנית. למעשה, העברנו את השימוש בזיכרון בכל איטרציה מן המחסנית אל ה-heap (ה-continuations הם closures והם מיוצרים על ה-heap).

**שאלה 5 – גרסת CPS עבור פרוצדורות מסדר גבוה (map$$)**



אם אנו מניחים כי פרוצדורות המיפוי f המתקבלת כארגומנט היא פרימיטיבית (ואינה כתובה בצורת CPS):



כיוון שהפרוצדורה f אינה כתובה בגרסת CPS, הפעלנו את c על (cons (f (car lst)) mapped-cdr). כעת, נניח כי f אינה פרימיטיבית, ולכן הומרה לגרסת ה CPS שלה, f$. נראה כיצד לכתוב את גרסת ה-CPS של map, כאשר הפרוצדורה אותה מפעילים על כל אחד מאיברי הרשימה, f$, נתונה בעצמה בגרסת CPS:



נתבונן ב-continuation שנבנה לאורך החישוב:







map$$ מפעילה את f$ על אברי הרשימה מהסוף להתחלה (בדוגמה 5 → 4 → 3 → 2 → 1). הפעלת f והקריאה ל-map אינן תלויות (אינן מקוננות זו בזו) ולכן ניתן לבחור מה יתרחש קודם בגרסת ה CPS. כך נקבע את הסדר בו תופעל f$ על אברי הרשימה (מההתחלה לסוף או להיפך). הגרסה הבאה מפעילה את f$ על אברי הרשימה לפי הסדר הרגיל של מההתחלה לסוף:



**שאלה 6 – עצירת החישוב**

הפונקציה mul-list$ מקבלת רשימה של רשימות ומחזירה מכפלה של כל המספרים ברשימה. אם יש מופע אטומי שאינו מספר אז יש לעצור את החישוב.

