**יבש 2 – NLP**

מגישים:

אמיר בלדר – 204179659

רועי גנץ – 204506349

**שאלה 1**

אם נבחן את המשפט ואת סט החוקים שקיבלנו, נשים לב כי ישנן רק 2 אפשרויות ל-parse trees נכונים. שמנו לב לכך תוך כדי הבנייה של העצים bottom-up, כשבעצם התחלנו מהעלים. סט החוקים הנתון הינו כזה שלא מאפשר בחירות מרובות. אנחנו מניחים שהסיבה לכך היא הכמות הלא כל כך גדולה של חוקים במקרה הזה, והעובדה שהמשפט עצמו הינו מאוד קצר. לפי הגדרת השאלה עלינו להסתכל על עצים שההסתברות לקבלם היא C כלשהו, חיובי ממש, ועל כן, בדקנו את שתי האפשרויות שלנו.

העץ הראשון:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S | | | | | | | | | |
| NP | VP | | | | | | | | |
| Vt | NP | | | | | | | |
| NP | | | | | PP | | |
| NP | | CC | NP | | IN | NP | |
| DT | NN | DT | NN | DT | NN |
| John | saw | The | cat | And | the | dog | with | the | Mouse |

נחשב את ההסתברות של העץ הזה:

העץ השני:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S | | | | | | | | | |
| NP | VP | | | | | | | | |
| Vt | NP | | | | | | | |
| NP | | CC | NP | | | | |
| NP | | PP | | |
| IN | NP | |
| DT | NN | DT | NN | DT | NN |
| John | saw | The | cat | and | the | dog | with | The | Mouse |

עתה נחשב את ההסתברות של העץ הזה:

אם כן, ישנם רק שני עצים אפשריים ולשניהם יש את אותה ההסתברות, . מש"ל.

**שאלה 2**

נסתכל על הנתון במקרה הזה, לפיו האורך של כל ענף ימני בעץ הינו בהכרח 1, כפי שהוגדר בשאלה, "directly into a terminal symbol". על כן, אם נסתכל על משפט, ועל המילה ה-j במשפט.  
הנוסחא המקורית הינה:



עתה נשתמש בנתון של השאלה, ונבין כי הביטוי שלנו בעצם ישתנה. אנו יודעים תמיד כי הענף הימני עתיד להיות terminal symbol. אם כן, אם קודם הסתכלנו על ביטוי כללי , עתה אנו יודעים בהכרח שמה שיבחר הוא terminal symbol עבור הענף הימני, לכן בעצם נוכל להסתכל על אוטומטית. אם נסמן את המילה ה-j במשפט ב- , אז בעצם מדובר . עתה נציב זאת בתוך המשוואה המקורית ונקבל:

אם כן, זוהי בעצם הנוסחא הרקורסיבית החדשה שלנו, כאשר המעבר במקרה הזה בין השורה הראשונה לשנייה, ובין השורה השנייה לשלישית, נובע ישירות מההסבר לעיל.