מילון (Dictionary)

מילון מאחסן אוסף של רְשוּמות מהטיפוס (מפתח, אינפורמציה).

המפתח שונה (בד"כ) מרשומה לרשומה.

אוסף המפתחות האפשריים מסומן ב-U. לדוגמא: מספרים שלמים.

פעולות:

- **אתחול:** יצירת מילון ריק. create(D)
- **חיפוש:** החזר מצביע לרשומה ב- D שמפתחה x או NULL. find(D,x)
- **הוספה:** הוסף ל- D רשומה שמפתחה x. insert(D,x,info)
- **הוצאה:** סלק מ- D רשומה שמפתחה x. delete(D,x)

כללים:

- x שייך לקבוצת המפתחות U.
- כל x מופיע לכל היותר פעם אחת במילון (בדר"כ).

מילון, מבנה חיפוש ועצי חיפוש

פעולות נוספות כאשר מוגדר סדר על U (למשל כאשר מפתח הוא מספר):

מינימום: החזר את המפתח המינימלי ב- D. min(D)

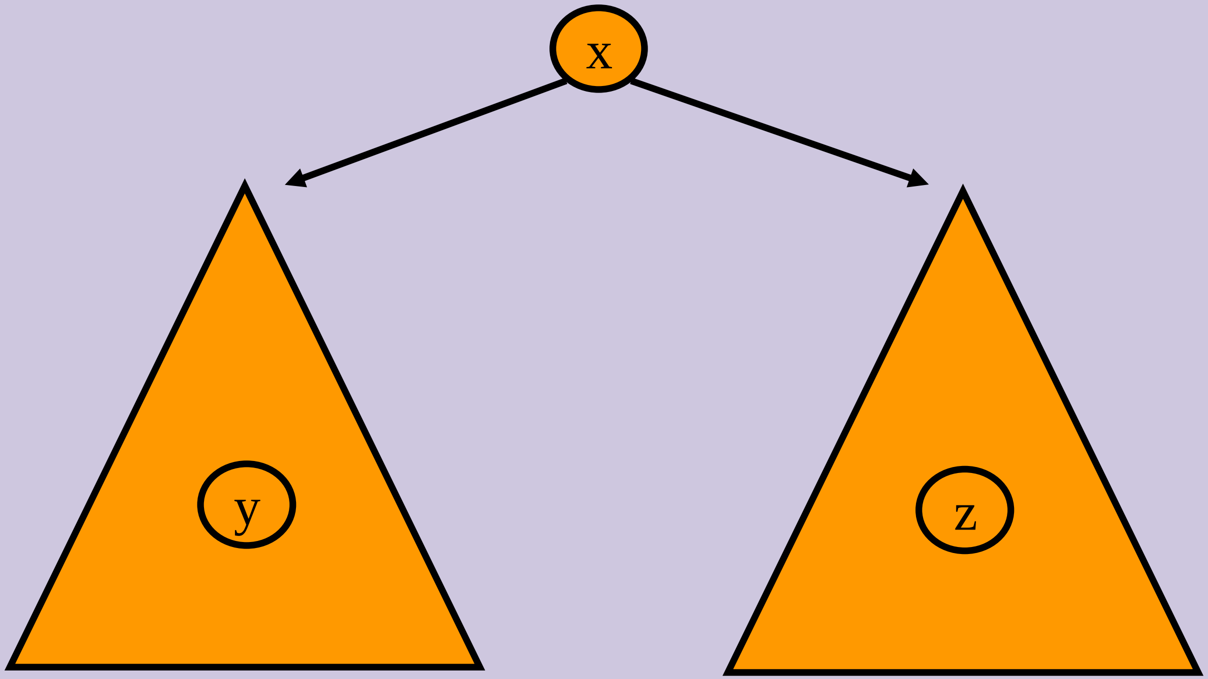
עוקב: החזר מצביע לאיבר במילון D בעל המפתח הקטן ביותר שגדול מ- x. next(D,x)

מילון + מינימום + עוקב נקרא = מבנה חיפוש

מטרה: לבצע את כל הפעולות בזמן O(log n) (במקרה הגרוע ביותר) כאשר n הוא מספר המפתחות הנמצאים במילון בזמן ביצוע הפעולה.

עצי חיפוש: היא משפחה של מימושים למבנה חיפוש.

עץ חיפוש בינרי



$$y < x < z$$

סיבוכיות הפעולות:

פעולה	function	סיבוכיות
אתחול	create(D)	$O(1)$
חיפוש	find(D,x)	$O(h)$
הוספה	insert(D,x,info)	$O(h)$
הוצאה	delete(D,x)	$O(h)$
מינימום	min(D)	$O(h) \rightarrow O(1)$
עוקב	next(D,x)	$O(h) \rightarrow O(1)$

עצים מאוזנים

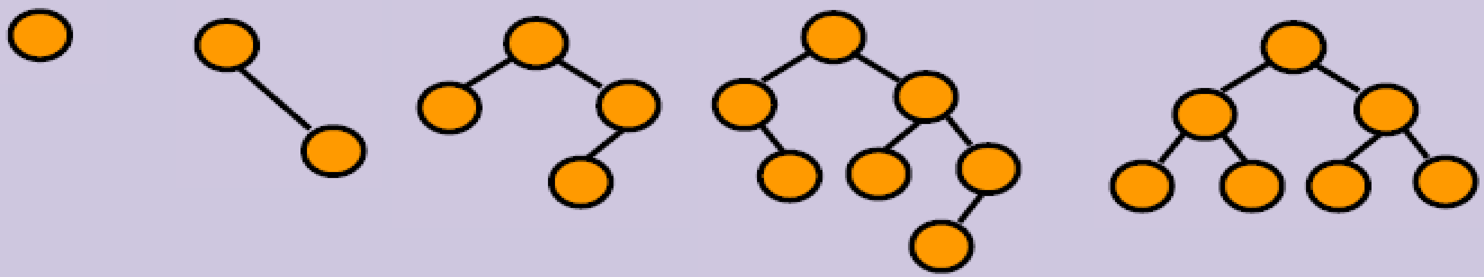
הגדרה: משפחת עצים תקרא מאוזנת אם $h(T) = O(\log n)$.

עצי AVL (Adelson-Velsky, Landis)

הגדרה: עץ AVL הוא עץ חיפוש בינרי שבו לכל צומת v התכונה:

$$| h(v \to left) - h(v \to right) | \leq 1$$

דוגמאות לעצי AVL



עצים מאוזנים

<u>הגדרה</u>: משפחת עצים תקרא <u>מאוזנת </u>אם $h(T) = O(\log n)$.

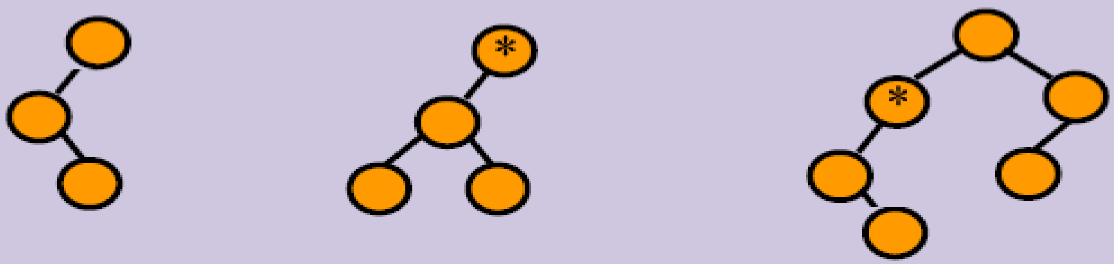
עצי AVL (Adelson-Velsky, Landis)

<u>הגדרה</u>: עץ AVL הוא עץ חיפוש בינרי שבו לכל צומת v התכונה:

$$| h(v \to left) \; - \; h(v \to right) | \leq 1$$

<u>דוגמאות לעצים שהם לא AVL</u>

צומת בו מופר האיזון



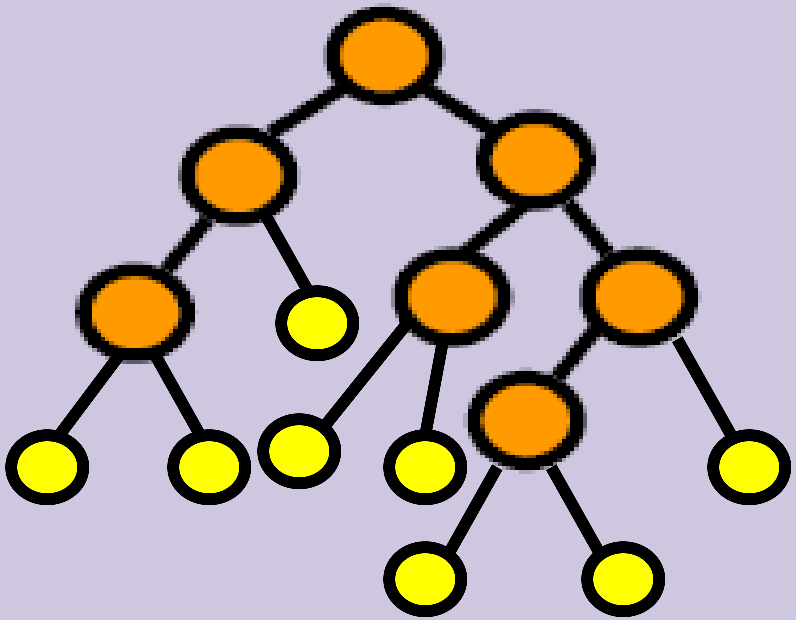
חסם לגובה עץ AVL

משפט:

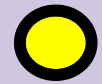
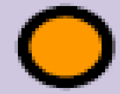
יהי T עץ AVL בן n צמתים וגובה h, אזי $h=O(\log n)$.

הוכחה:

T

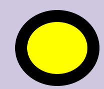
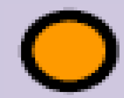


טענה 1:

 = #  +1

הוכחת טענה 1:

באינדוקציה

N(T) = #  n(T) = # 

$$N(T) = N(T \rightarrow left) + N(T \rightarrow right)$$

$$= n(T \rightarrow left)+1 + n(T \rightarrow right)+1$$

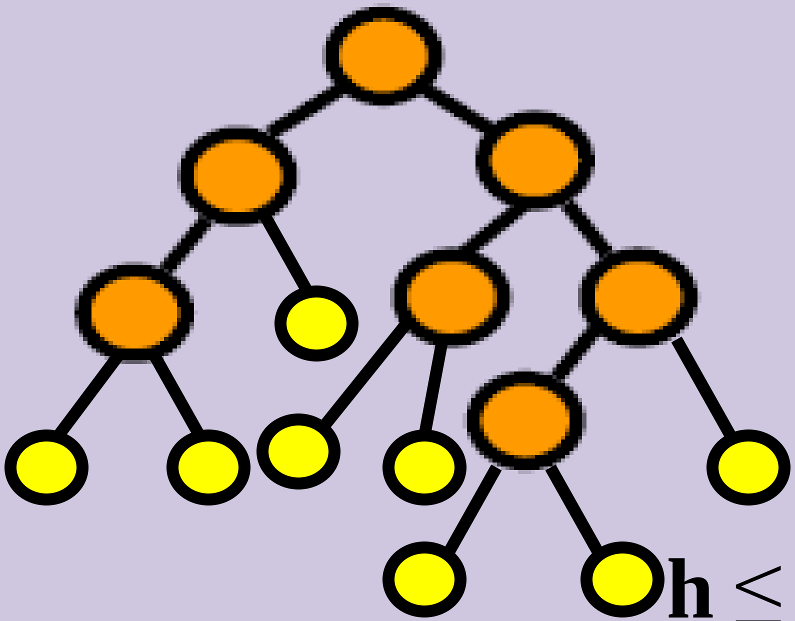
$$= n(T)+1$$

■

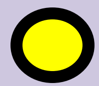
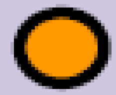
חסם לגובה עץ AVL

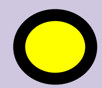
משפט:

יהי T עץ AVL בן n צמתים וגובה h, אזי h=O(log n).



הוכחה:

טענה 1: #  = #  +1

הגדרה: נסמן ב- N_h מספר מינימאלי של  בעץ AVL בגובה $\leq$ **h**

טענה 2: $N_{h-2} \leqslant N_{h-1}$

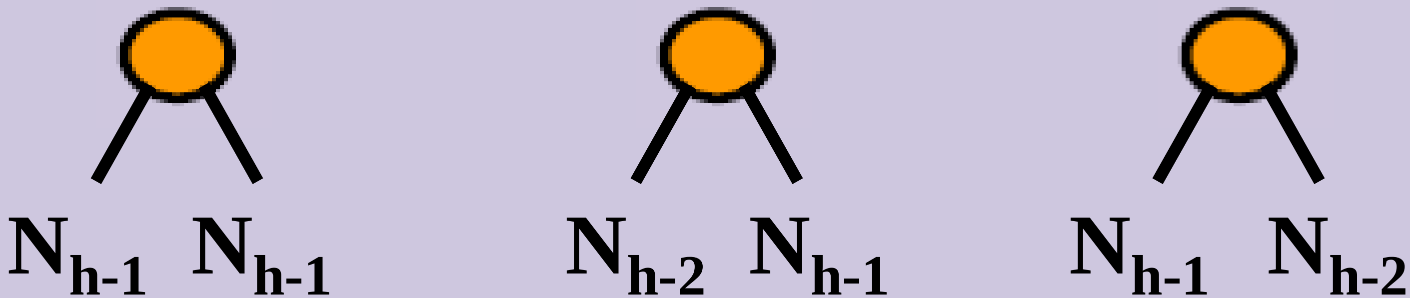
■

<u>הגדרה</u>: נסמן ב- N_h מספר

מינימאלי של בעץ AVL בגובה $\le h$

<u>טענה 3</u>: $N_h=N_{h-1}+N_{h-2}$

<u>הוכחת טענה 3</u>: שלשה מקרים



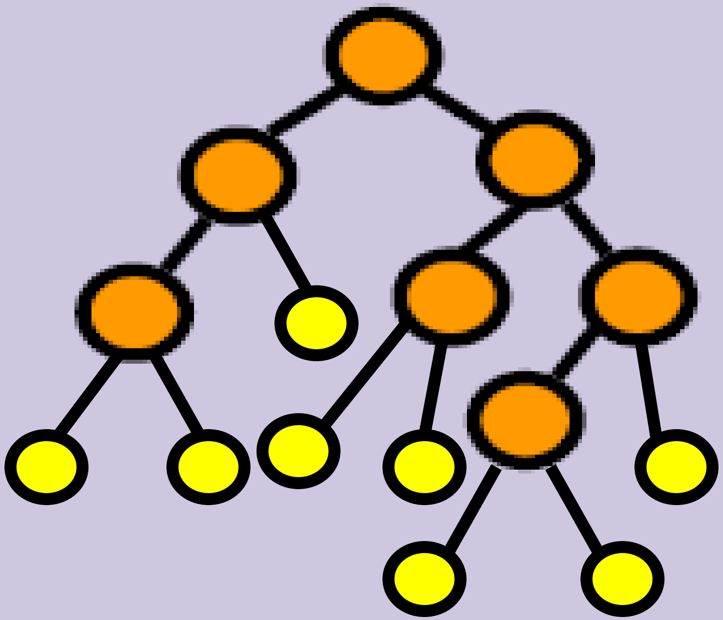
$$N_h=\boldsymbol{min}(N_{h-1}+N_{h-2}\,,\,N_{h-1}+N_{h-1}) = N_{h-1}+N_{h-2}$$

חסם לגובה עץ AVL

משפט:

יהי T עץ AVL בן n צמתים וגובה h, אזי h=O(log n).

הוכחה:



טענה 1: $\# = \# + 1$

הגדרה: נסמן ב- N_h מספר

מינימאלי של בעץ AVL בגובה **h**

טענה 2: $N_{h-2} \leqslant N_{h-1}$

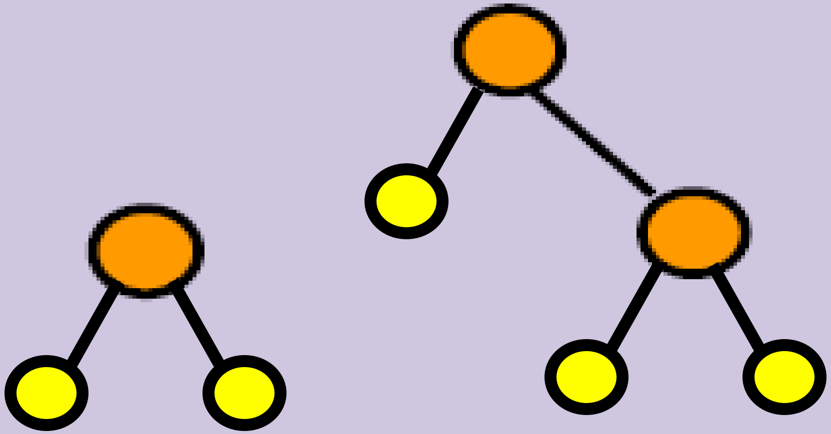
טענה 3: $N_h=N_{h-1}+N_{h-2}$

טענה 4: $N_h> 2^{h/2}$

טענה 4: $N_h > 2^{h/2}$

הוכחת טענה 4:

באינדוקציה



$$N_0 = 2 > 2^{0/2} \quad N_1 = 3 > 2^{1/2}$$

טענה 3

$$N_h = N_{h-1} + N_{h-2} > 2^{(h-1)/2} + 2^{(h-2)/2}$$

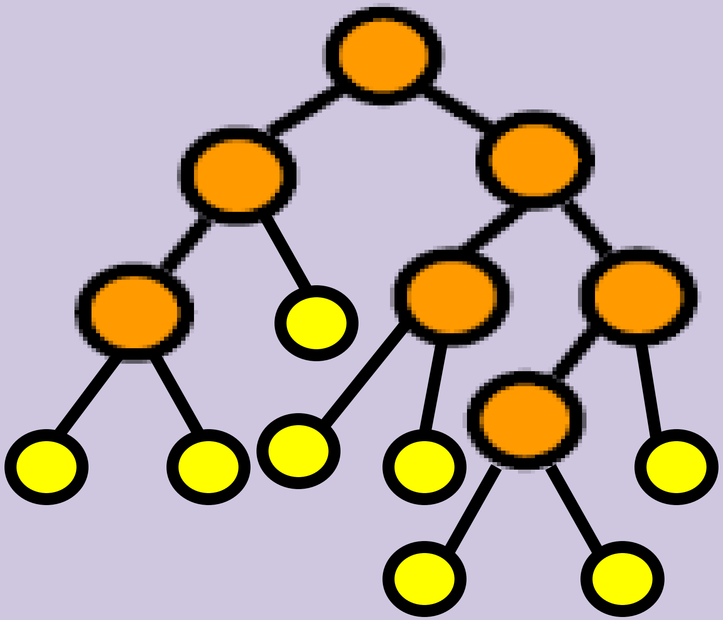
$$N_h > 2^{h/2}\left(\frac{1}{\sqrt{2}} + \frac{1}{2}\right) > 2^{h/2}$$

$\blacksquare$

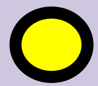
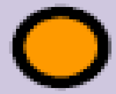
חסם לגובה עץ AVL

משפט:

יהי T עץ AVL בן n צמתים וגובה h, אזי $h=O(\log n)$.



הוכחה:

טענה 1: #  = #  +1

$$N = n+1$$

טענה 4: $N_h > 2^{h/2}$

$$n+1=N \geq N_h > 2^{h/2}$$

$$h=2(h/2) < 2\log(N_h) < 2\log(n+1)$$

■

חסם לגובה עץ AVL

משפט:

יהי T עץ AVL בן n צמתים וגובה h, אזי $h=O(\log n)$.

הוכחנו

יהי T עץ AVL בן n צמתים וגובה h, אזי $h=2\log(n+1)$.

$$N_h > 2^{h/2}$$

$$N_h = N_{h-1} + N_{h-2} \qquad N_0 = 2 \qquad N_1 = 3$$

מספרי פיבונצ'י

$$N_h = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{h+3} + \left(\frac{1-\sqrt{5}}{2}\right)^{h+3}}{\sqrt{5}}$$

$$N_h = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{h+3} + \left(\frac{1-\sqrt{5}}{2}\right)^{h+3}}{\sqrt{5}}$$

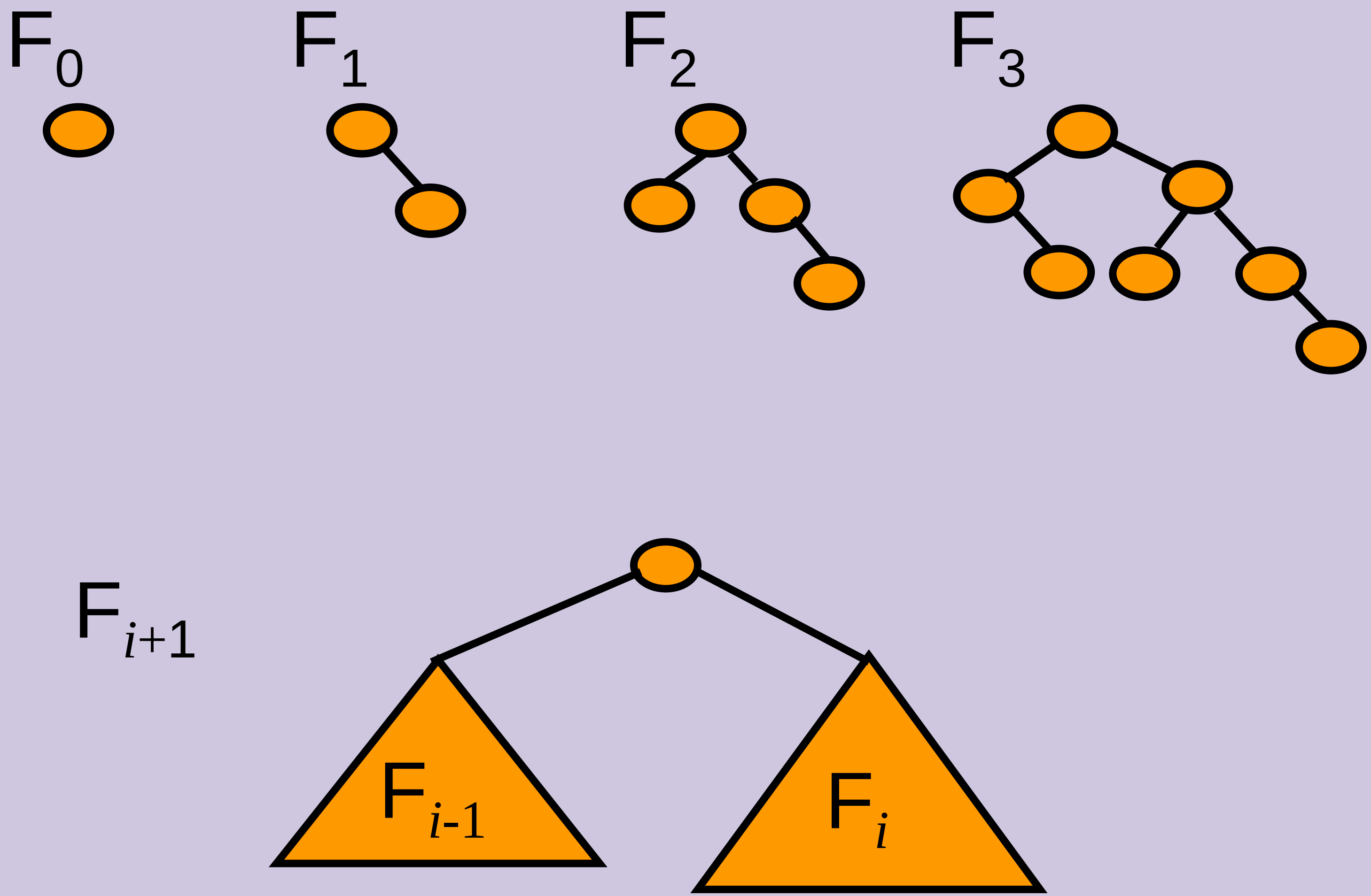
$$N_h \approx \frac{\Phi^{h+3}}{\sqrt{5}}$$

נקרא יחס הזהב $\Phi = \frac{1+\sqrt{5}}{2}$

$$h \approx \log N_n / \log \Phi = 1.44 \log n$$

$\blacksquare$

משפחת **עצי פיבונצ'י** (Fibonacci trees):



ניתוח מספר צמתים בעץ פיבונצ'י

$$n_0 = 0, n_1 = 1, n_2 = 1, n_3 = 2, n_4 = 3, n_5 = 5,$$

טענה: לעץ פיבונאצ'י F_i יש $|F_i| = n_{i+3} - 1$ צמתים כאשר n_i הוא מספר פיבונצ'י ה-i.

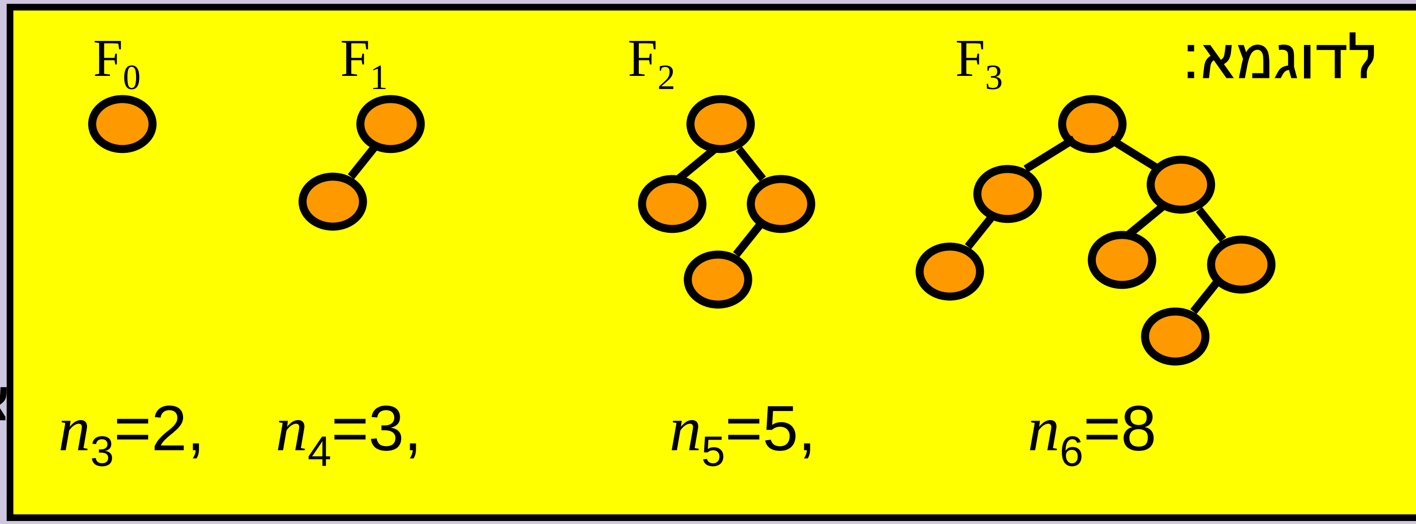
הוכחה: ב

בסיס האי

הנחת הא

צעד האינ

צ'י,



$$|F_i| = |F_{i-1}| + |F_{i-2}| + 1 = (n_{i+2} - 1) + (n_{i+1} - 1) + 1$$

$$= n_{i+3} - 1$$

מספרי פיבונצ'י (המשך)

הוכחת הטענה : נתונה המשוואה הבאה. $n_{i+1} = n_i + n_{i-1}$

נניח פתרון מהצורה: $n_i = a \cdot x^i$

נציב במשוואה ונקבל:

$$a \cdot x^{i+1} = a \cdot x^i + a \cdot x^{i-1}$$

$$x^2 = x + 1$$

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad x_{1,2} = \frac{1 \pm \sqrt{5}}{2} \qquad \Phi = \frac{1+\sqrt{5}}{2} \qquad \overline{\Phi} = \frac{1-\sqrt{5}}{2}$$

משוואת ההפרשים ליניארית ולכן כל צרוף ליניארי של פתרונות מהווה פתרון:

$$n_i = a \cdot \Phi^i + b \cdot \overline{\Phi}^i$$

שימוש בתנאי השפה מוביל למציאת הקבועים:

$$n_0 = 0 \quad \Rightarrow \quad a \cdot \Phi^0 + b \cdot \overline{\Phi}^0 = a + b = 0 \quad \Rightarrow \quad b = -a$$

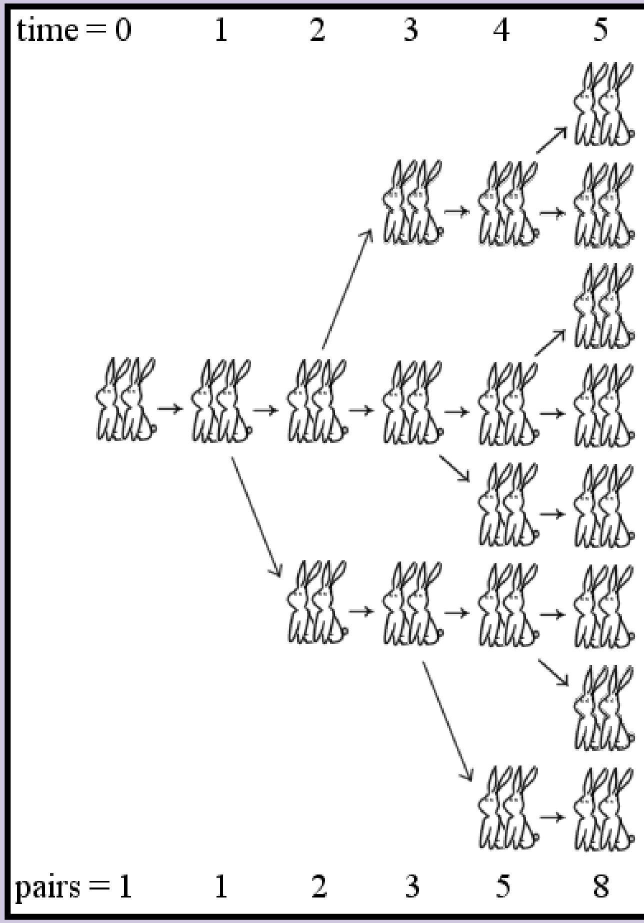
$$n_1 = 1 \quad \Rightarrow \quad a \cdot \left(\frac{1+\sqrt{5}}{2}\right) - a \cdot \left(\frac{1-\sqrt{5}}{2}\right) = 1 \quad \Rightarrow \quad a = \frac{1}{\sqrt{5}}$$

לפיכך פתרון המשוואה (כולל תנאי ההתחלה) הוא: $n_i = \dfrac{\Phi^i - \overline{\Phi}^i}{\sqrt{5}}$

על מספרי פיבונצ'י

פיבונאצ'י

סיפור הארנבים (תרגיל בספרו משנת 1202):



יש זוג ארנבים צעירים

כעבור חודש מתבגרים

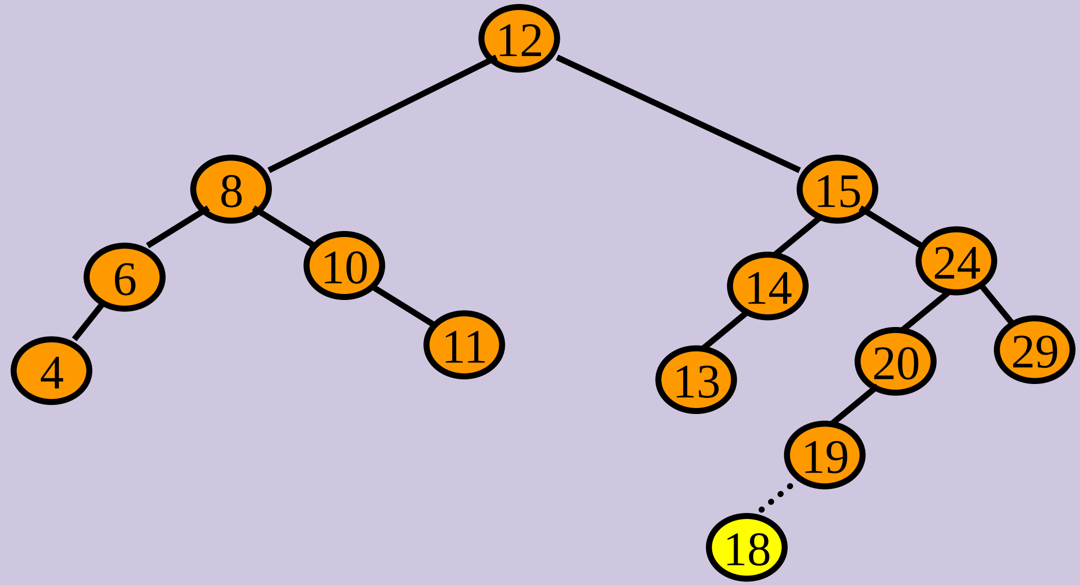
כעבור חודש מַמְלִיטִים זוג ארנבים

וכו'

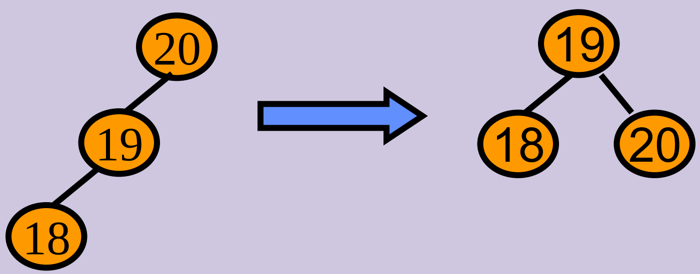
כמה ארנבים יהיו בחודש i?

איזון בעץ AVL

נובע שזמן החיפוש בעץ AVL הוא O(log n). נצטרך לדאוג שלאחר הכנסה או הוצאה, העץ הנותר יהיה עץ AVL.



לאחר הוספת האיבר 18 נתקבל עץ שאינו עץ AVL.
אבל נתן לשנות את תת העץ שבו הופר האיזון בצורה הבאה:



תיקון כזה נקרא <u>גלגול</u>.

בזמן הוצאה קיימת הפרת איזון דומה. למשל בהוצאת 29.

איזון בעץ AVL (המשך)

עבור צומת v בעץ בינרי נסמן:

$h_L(v)$ גובה תת העץ השמאלי של v.

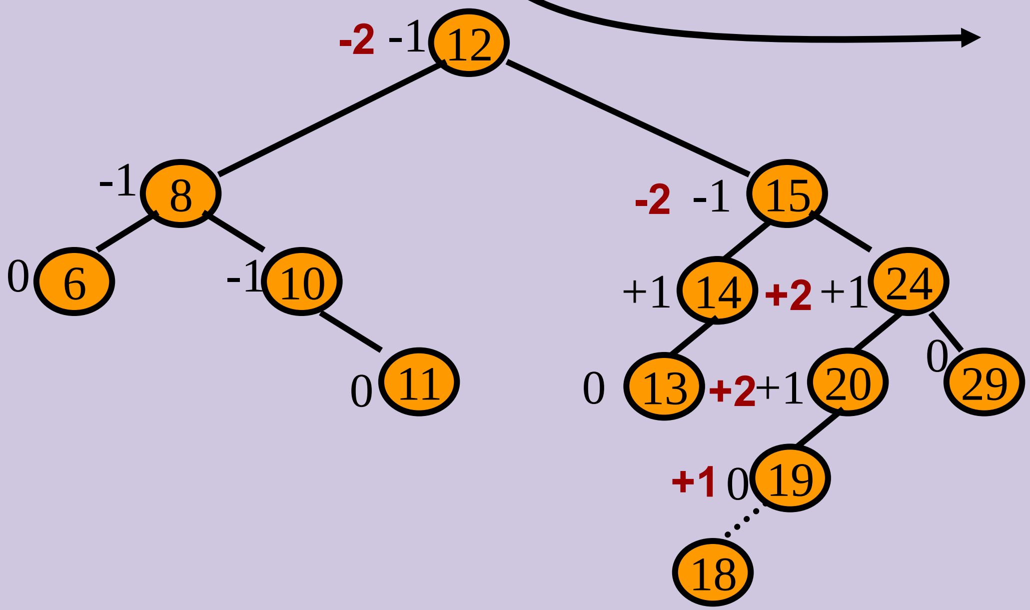
$h_R(v)$ גובה תת העץ הימני של v.

גורם האיזון (Balance Factor) מחושב כהפרש הגבהים: $BF(v) = h_L(v) - h_R(v)$

בעץ AVL תקין: $|BF(v)| \le 1$

לדוגמא:

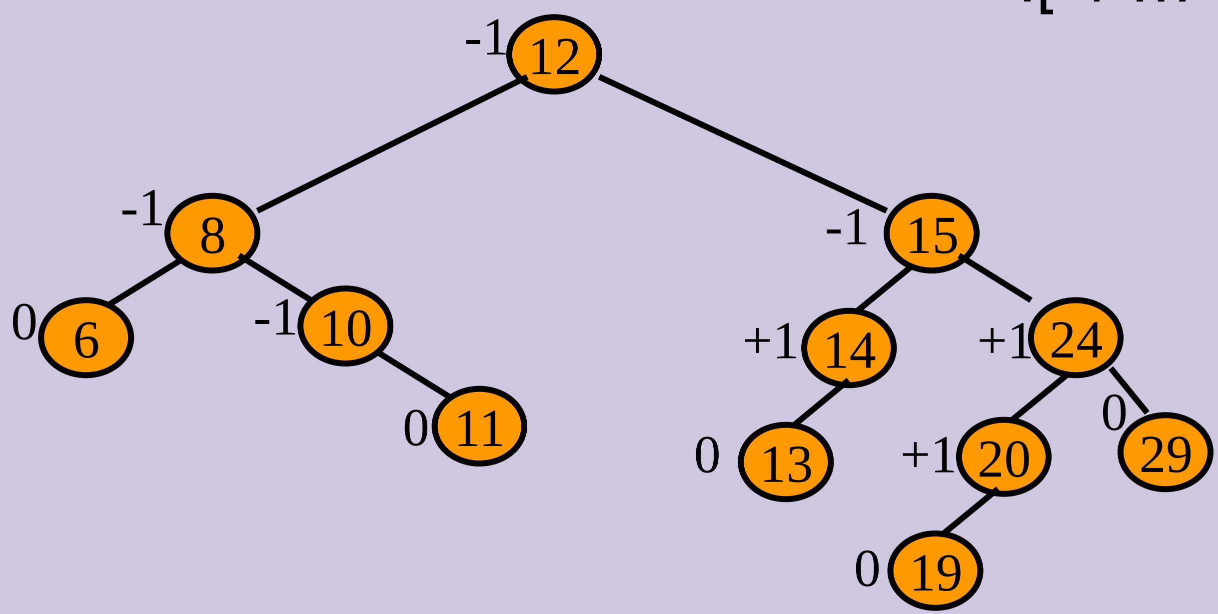
מצד שמאל של כל צומת מסומן גורם האיזון.



אחרי ההכנסה של 18 גורם האיזון מופר על מסלול ההכנסה.

אבחנות

1. הצמתים היחידים שאולי הופר בהם האיזון הם הצמתים לאורך מסלול הכנסה/הוצאה.
2. אם עבור צומת v במסלול הנ"ל גובה העץ ששורשו v לא השתנה אזי גורמי האיזון בצמתים שמעליו במסלול לא השתנו.
3. אם גורם האיזון הופך ל-2 או ל 2-, אזי יש לבצע גלגול על מנת שהעץ יחזור להיות עץ AVL.
4. גורם האיזון לא יכול להיות גדול מ-2 בערכו המוחלט כי בכל הכנסה/הוצאה הוא משתנה ב-1 לכל היותר.
5. **גלגול**– פעולה המתבצעת על צומת שהופר בו האיזון על מנת להחזירו לתחום המותר [-1 … 1].



הצמת האחרון שהופר בו האיזון

<u>טענה</u> גורם האיזון החדש b

לא יכול להיות 0

$$BF(v) = h_L(v) - h_R(v)$$

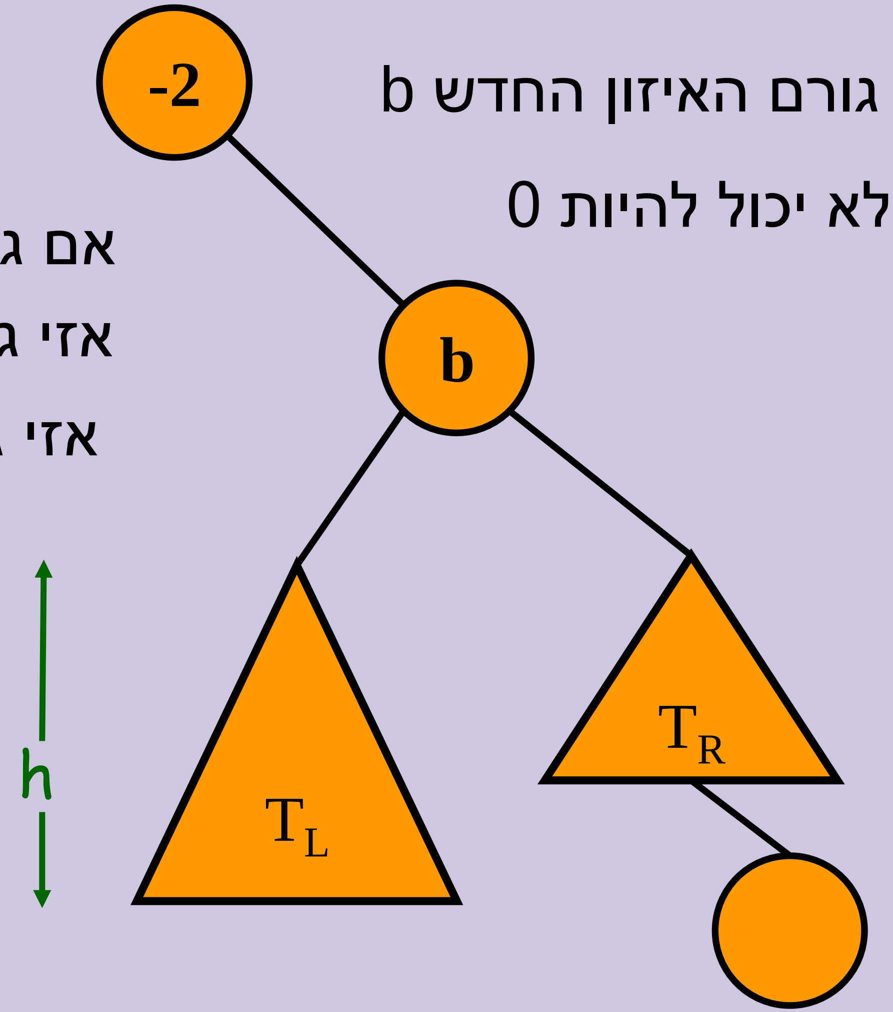
אם גורם האיזון החדש הוא 0

אזי גורם האיזון הישן היה 1+

אזי גובה תת העץ של- b לא השתנה

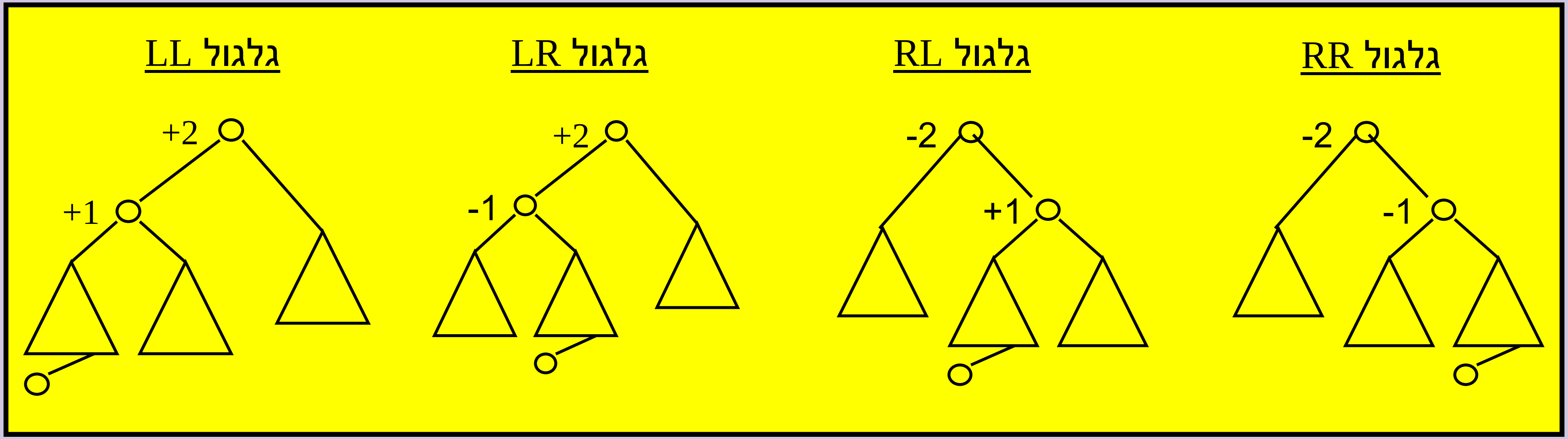
סתירה

לכן b=+1 או b=-1



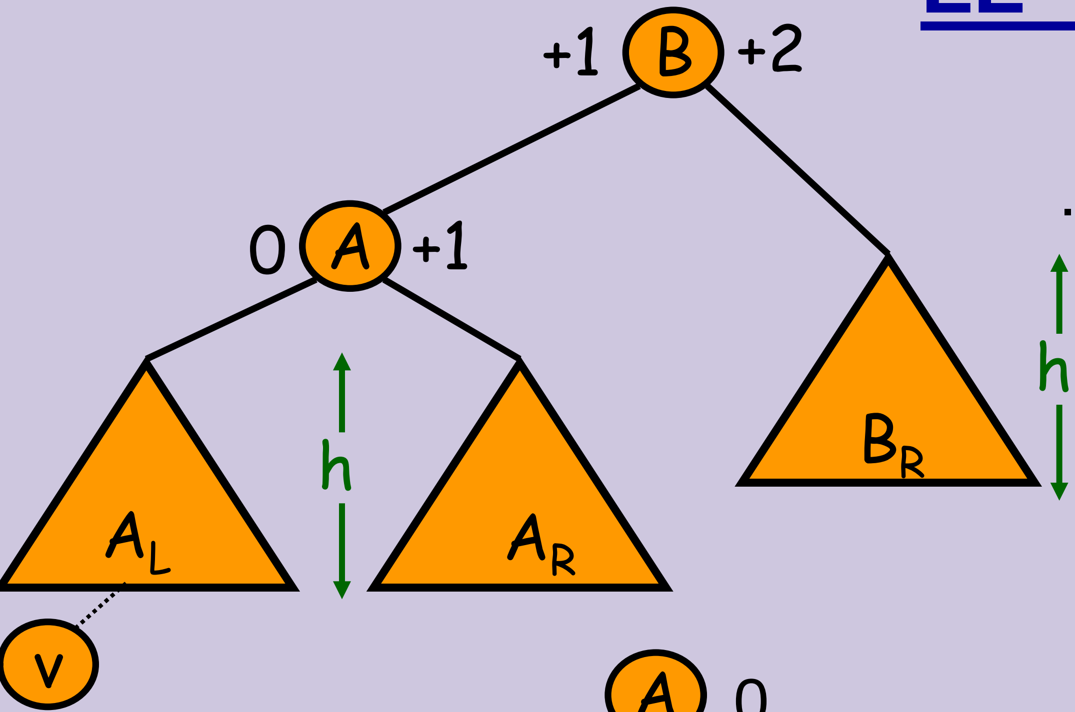
סוגי הגלגולים

סוג הגלגול, כלומר הדרך לתקן חוסר איזון בצומת, תלוי בצורה בה האיזון מוּפָר.
נתן לסווג חוסר איזון בארבע קטגוריות שונות המכסות את כל המקרים.



הגלגול המתאים	בבן הימני v_R	בבן השמאלי v_L	בשורש v
LL		$BF(v_L) = 1$	$BF(v)=2$
LR		$BF(v_L) = -1$	$BF(v)=2$
RR	$BF(v_R) = -1$		$BF(v)= -2$
RL	$BF(v_R) = 1$		$BF(v)= -2$

גלגול LL

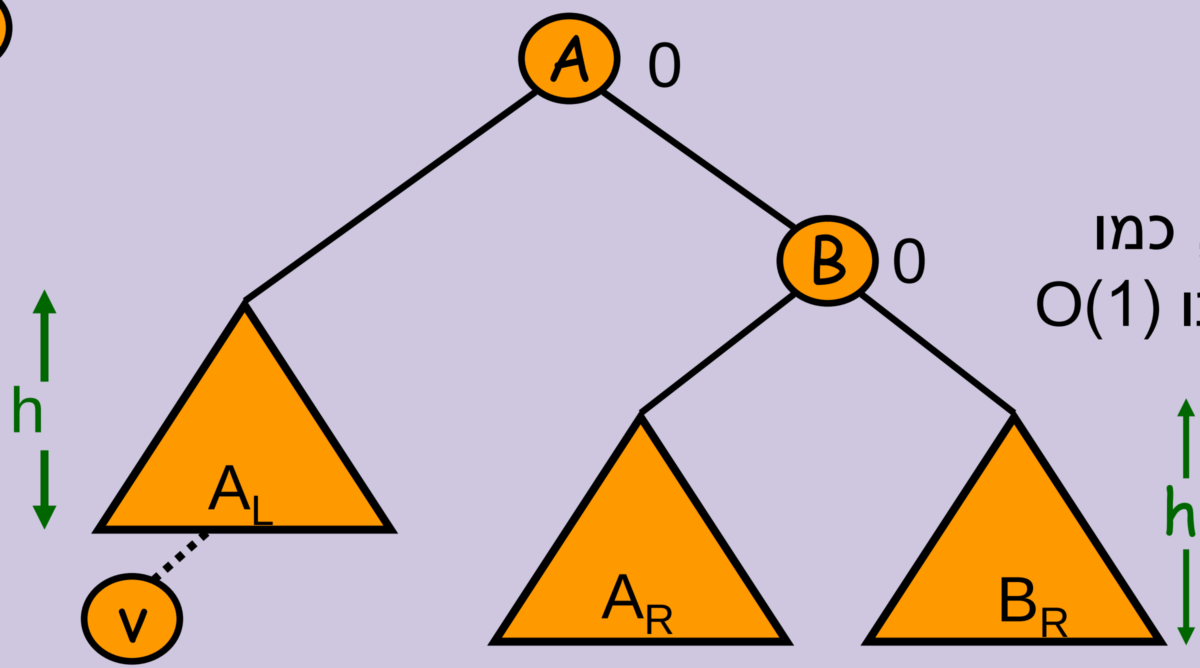


<u>לפני הכנסת v</u>: גובה העץ הוא h +2.

הוכנס צומת v שהגדיל את גובה A_L ל-1+ h.

מצד ימין של הצמתים מסומנים גורמי האיזון שהשתנו.

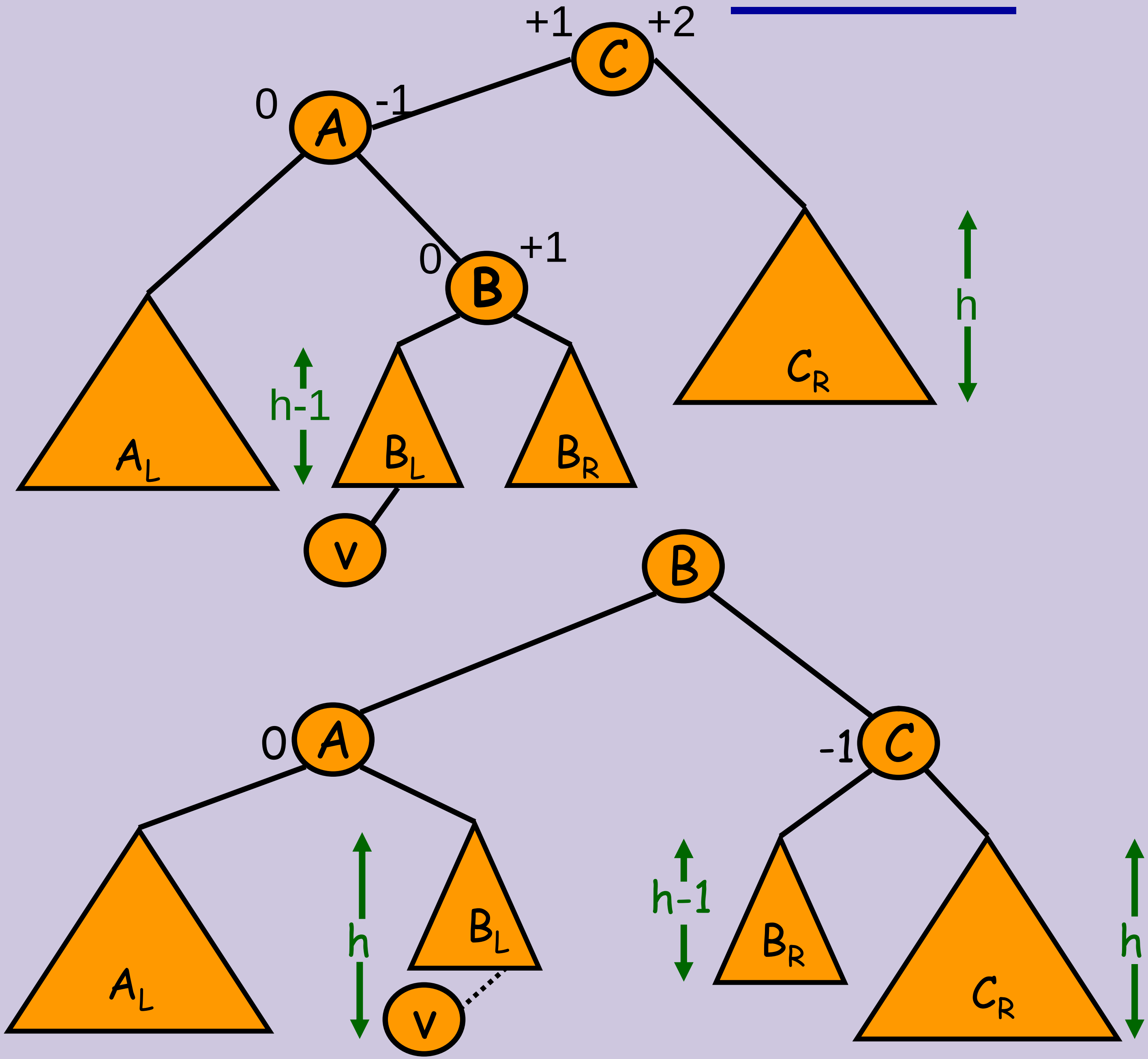
<u>גלגול LL</u> : יעביר את A לשורש



<u>אחרי הגלגול</u>:

גובה העץ לאחר הגלגול הוא h +2, כמו לפני ההכנסה. השורש מאוזן. שינינו O(1) מצביעים ולכן זמן הגלגול O(1).

גלגול LR



<u>לפני הכנסת איבר v:</u>

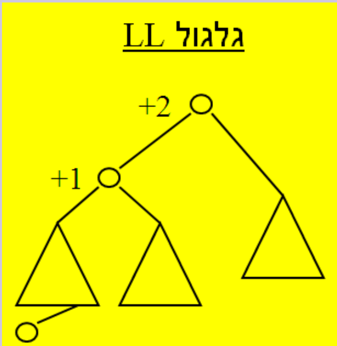
הוכנס איבר ל-B_L שגרם לו להעלות את גובהו ל-h.

<u>גלגול LR:</u>

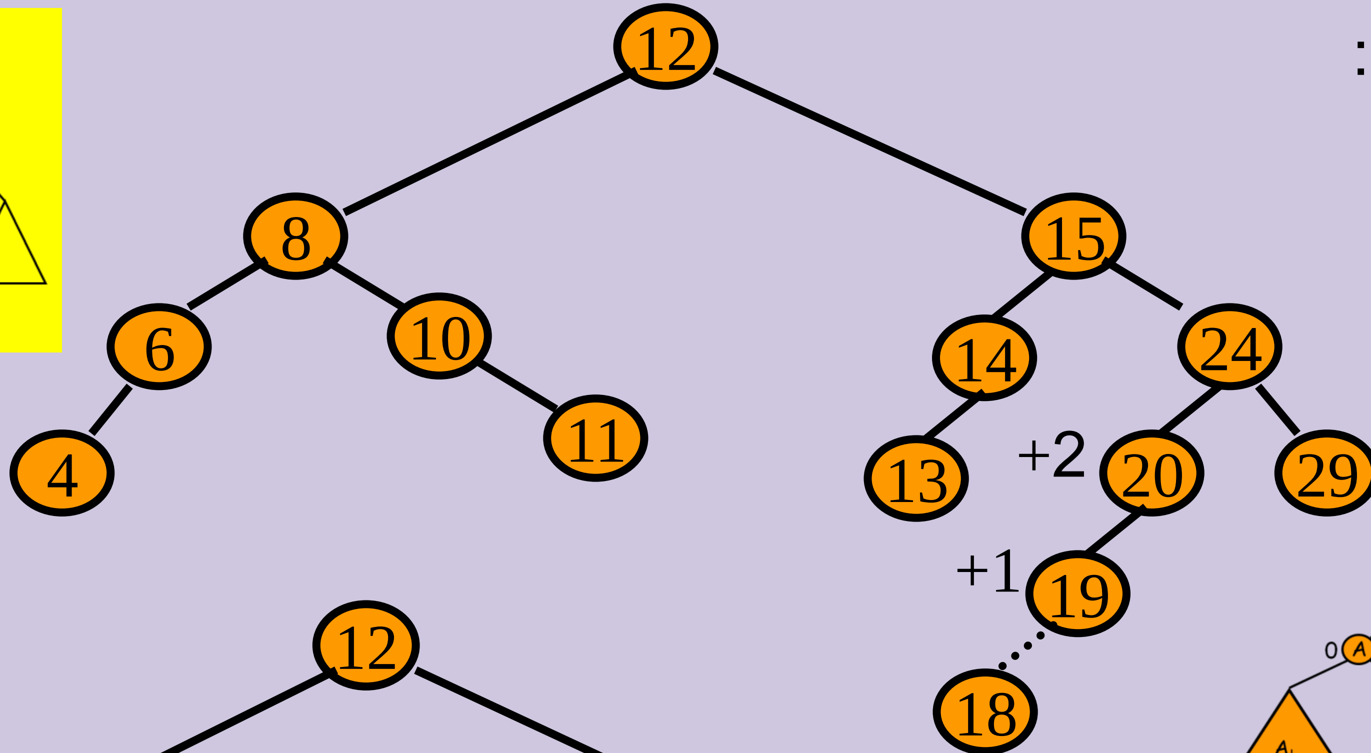
גובה העץ אחרי הגלגול הוא h +2, כמו לפני ההכנסה.

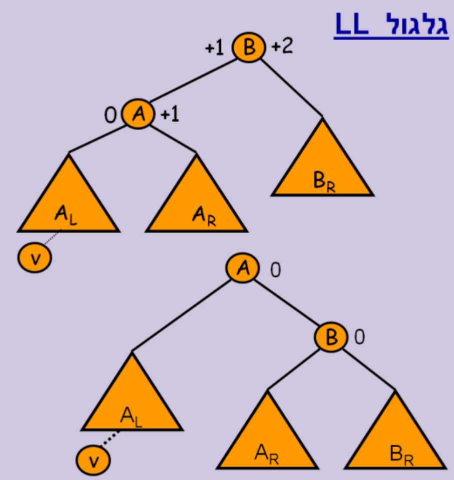
שינינו O(1) מצביעים ולכן זמן הגלגול O(1).

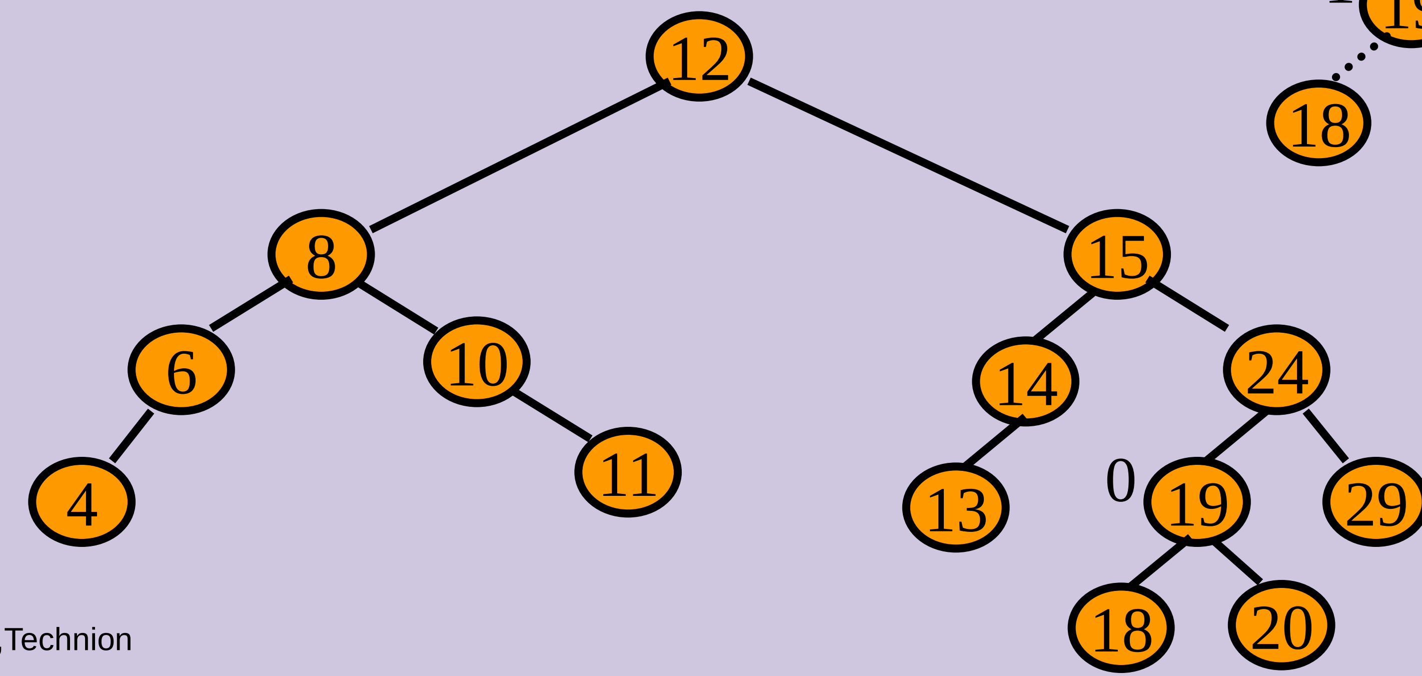
דוגמא להכנסת ערך x לעץ AVL



הוסף 18:

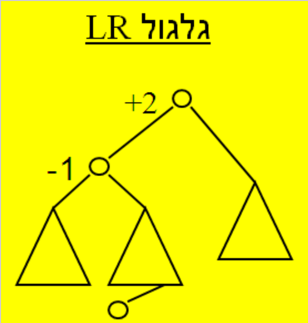




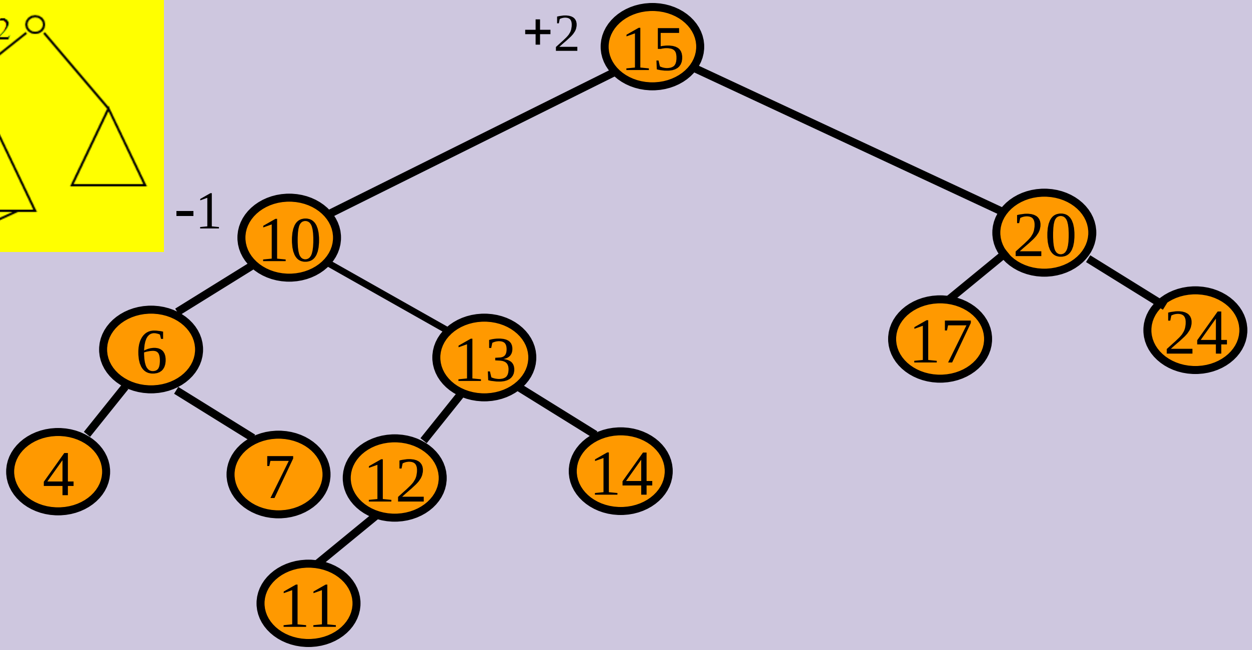


לאחר גלגול LL:

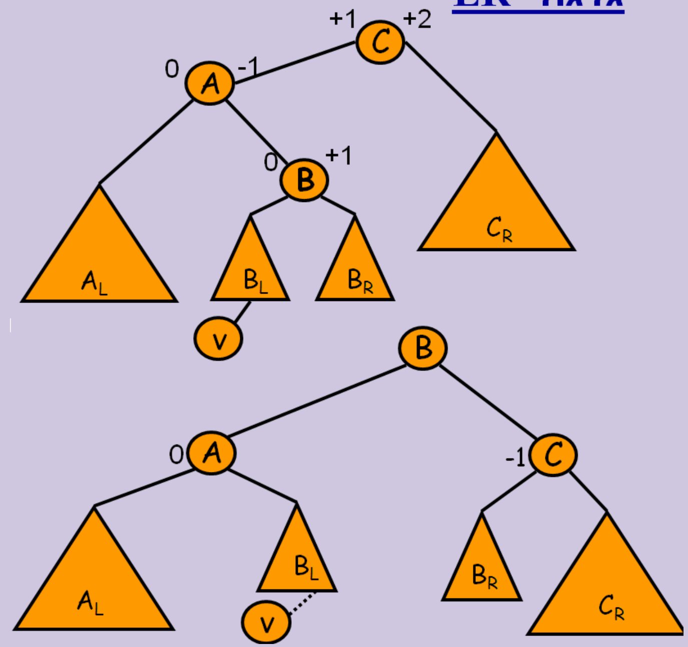
דוגמא להכנסת ערך x לעץ AVL



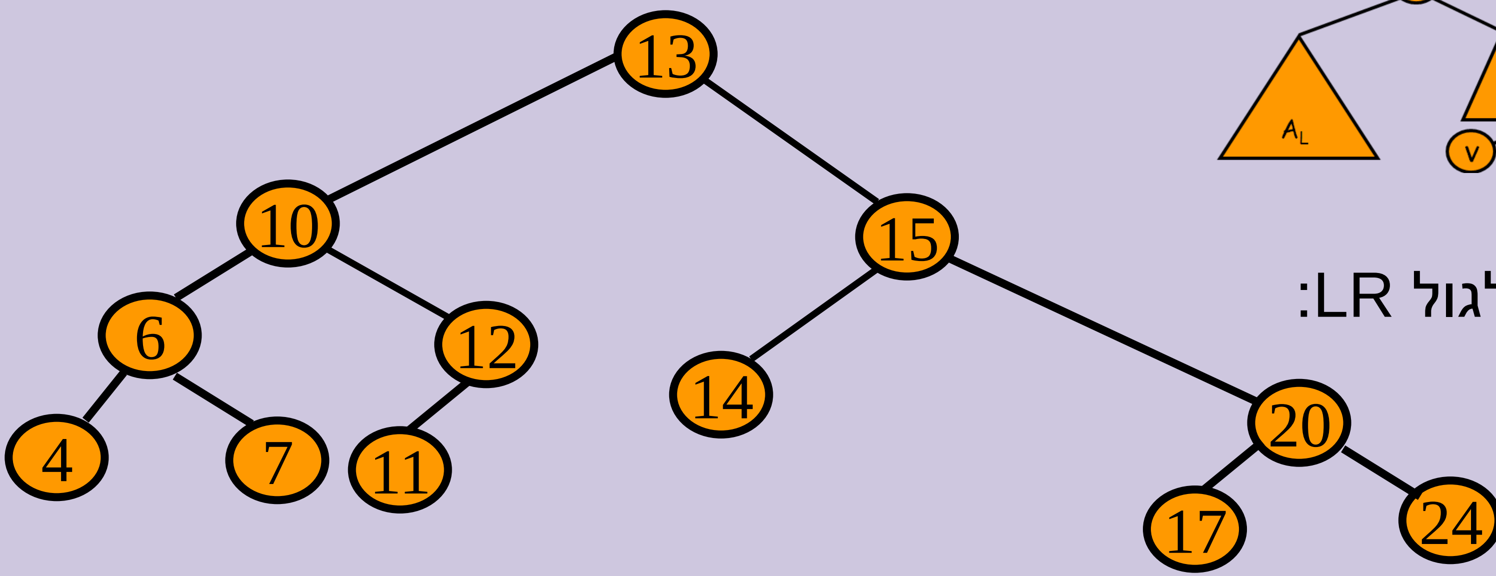
הוסף 11 :



גלגול LR

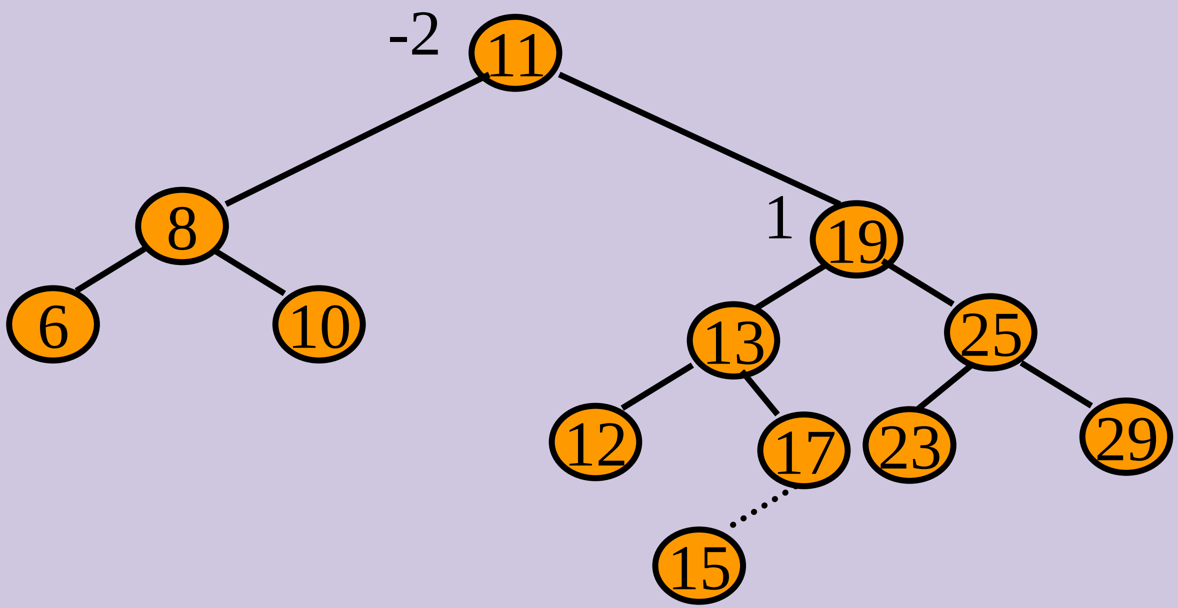


לאחר גלגול LR:

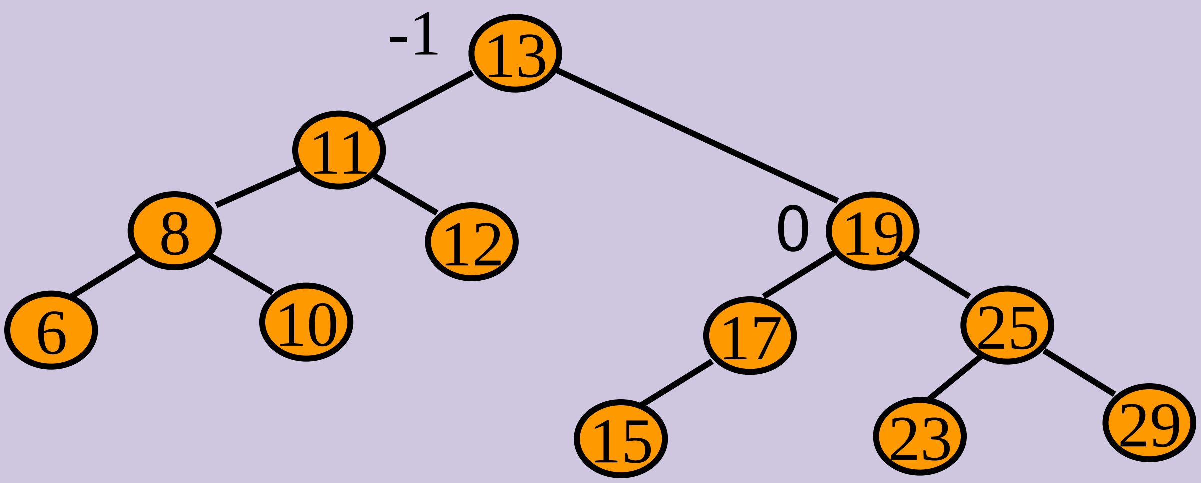


דוגמא להכנסת ערך x לעץ AVL

הוסף 15 (נחוץ גלגול RL)



לאחר גלגול RL:



אלגוריתם להכנסת ערך x לעץ AVL

1. הכנס את x כמו לעץ חיפוש בינרי. יהי v העלה שהוסף.
2. $h(v) = 0$
3. כל עוד $v \neq root$ בצע:
    4. $p = parent(v)$
    5. אם $h(p) \geq h(v) + 1$ סיים.
    6. $h(p) = h(v) + 1$
    7. אם ב- p הופר האיזון, בצע גלגול וסיים.
    8. אחרת $v = p$

איך נחשב את parent(v) ?

למשל, נוציא אותו ממחסנית בה נמצאים כל הצמתים על המסלול מהשורש ועד v.

זמן ההכנסה לעץ AVL

כיוון שהצומת בו עושים גלגול לא משנה את גובהו,
מבצעים רק גלגול אחד.

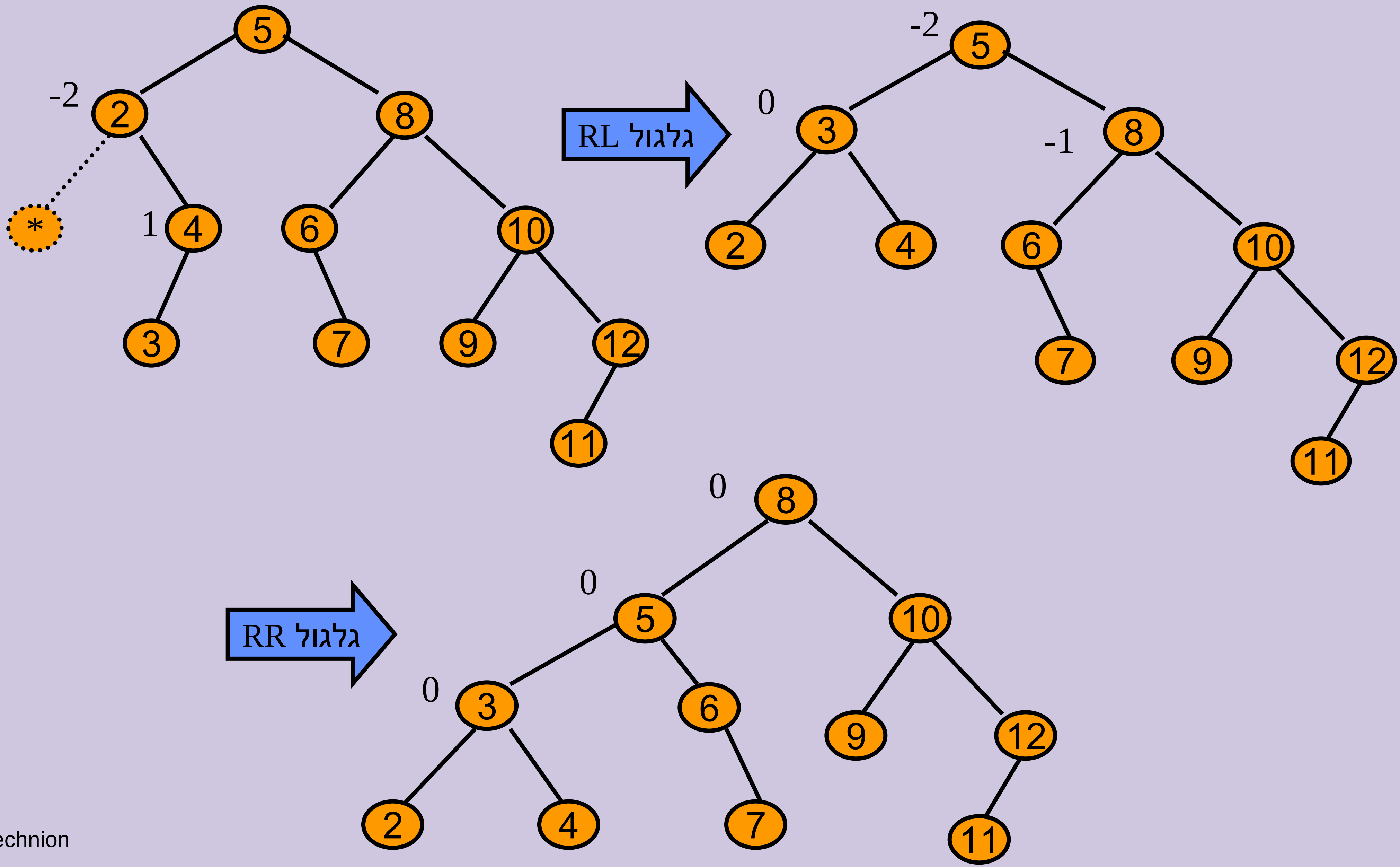
מציאת המקום הדרוש להכנסה	$O(h)$
הוספת הצומת	$O(1)$
מציאת המקום בו מופר האיזון	
(אם מופר)	$O(h)$
תיקון האיזון	$O(1)$
סה"כ	$O(h) = O(\log n)$

אלגוריתם הוצאה

- הוצא צומת v כפי שהפעולה מתבצעת בעץ חיפוש בינרי.
- תקן את גורמי האיזון בצורה הבאה. לכל צומת v לאורך המסלול החל מלמטה ועד לשורש בצע:
  - עדכן את BF(v)
  - אם $|BF(v)| = 2$, בצע גלגול והמשך כלפי מעלה.
  - אם גובה תת העץ ששורשו v לא השתנה, סיים.
  - אם גובה תת העץ השתנה ו- BF(v) תקין, המשך כלפי מעלה.

בהוצאה יתכן יותר מגלגול אחד.

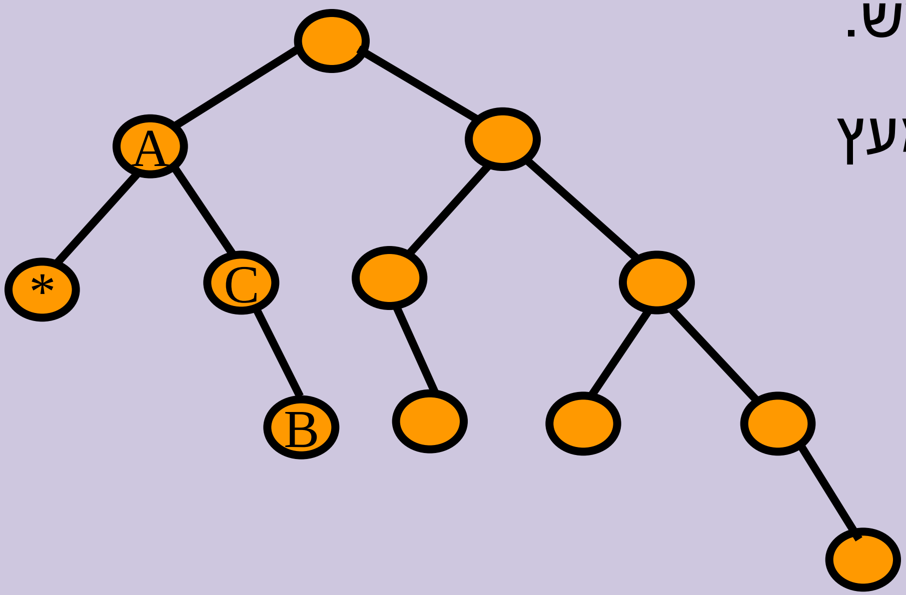
דוגמא להוצאה מעץ AVL



דוגמא: הוצאה מעץ פיבונצ'י

בהוצאה יתכן גלגול בכל צומת על המסלול בין הצומת המוצא ובין השורש.

לדוגמא כאשר מוציאים עלה ראשון מעץ פיבונצ'י - F_i.



<u>תרגיל לבית</u>: מהי סדרת הגלגולים המתבצעת כאשר מוציאים עלה ראשון מעץ פיבונצ'י - F_i.

זמן הוצאה מעץ AVL

מציאת המקום הדרוש להוצאה $O(h)$

מציאת המקום בו מופר האיזון

(אם מופר) $O(h)$

תיקון האיזון $O(h)$

(לכל היותר פעם בכל רמה)

סה"כ $O(h) = O(\log n)$

עצי AVL מאפשרים חיפוש, הכנסה, הוצאה בזמן $O(\log n)$

סיבוכיות הפעולות:

פעולה	function	סיבוכיות
אתחול	create(D)	$O(1)$
חיפוש	find(D,x)	$O(log\ n)$
הוספה	insert(D,x,info)	$O(log\ n)$
הוצאה	delete(D,x)	$O(log\ n)$
מינימום	min(D)	$O(log\ n) \rightarrow O(1)$
עוקב	next(D,x)	$O(log\ n) \rightarrow O(1)$