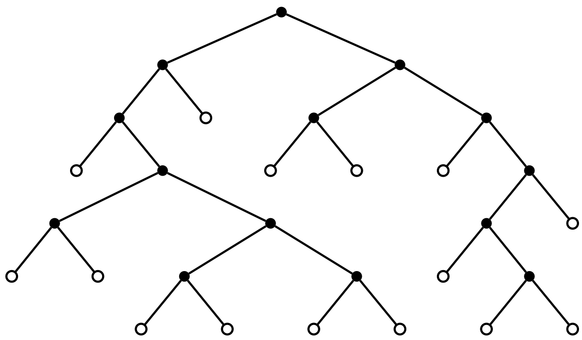
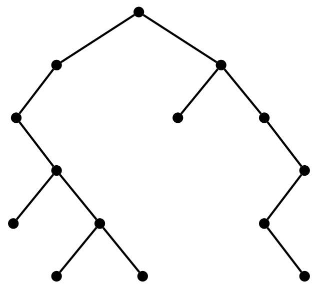
οπότε $T'$ είναι το δένδρο:



Το ζητούμενο δένδρο $T$ προκύπτει σβήνοντας όλα τα φύλλα του δένδρου $T'$ οπότε έχουμε το δένδρο $T$:



Από τις προηγούμενες κατασκευές μπορεί να αποδειχθεί η επόμενη πρόταση.

**Πρόταση 2.4.1.** *Ο αριθμός των λέξεων Dyck μήκους $2n$ ισούται με τον $n$-στό αριθμό Catalan, δηλαδή $|\mathcal{D}_n| = C_n$.*

Ο τύπος του Segner $C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}$, $C_0 = 1$ έχει το μειονέκτημα ότι στον υπολογισμό κάποιου όρου της ακολουθίας των αριθμών Catalan $(C_n)$ απαιτείται η γνώση όλων των προηγούμενων όρων.

Με τη βοήθεια των λέξεων Dyck μπορούμε να βρούμε (και να αποδείξουμε) ένα απλούστερο τύπο για τους αριθμούς Catalan.

**Πρόταση 2.4.2.** *Ο $n$-οστός αριθμός Catalan $C_n$ δίδεται από τον τύπο*

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n-1}.$$

*Απόδειξη.* Έστω $\mathcal{A}_n$ το σύνολο των δυαδικών λέξεων που αποτελούνται από $n$ μηδενικά και $n$ άσσους, χωρίς κανένα άλλο περιορισμό.

Το σύνολο $\mathcal{A}_n$ διαμερίζεται στα σύνολα:

- των λέξεων μήκους $2n$ στις οποίες σε κάθε αρχικό τμήμα τους τα 1 είναι τουλάχιστον όσα τα 0, δηλαδή λέξεις Dyck μήκους $2n$. (Συμβολισμός: $\mathcal{D}_n$.)

- των λέξεων μήκους $2n$ στις οποίες κάποιο αρχικό τμήμα τους έχει περισσότερα 0 από ότι 1. (Συμβολισμός: $\overline{\mathcal{D}}_n$.)

Επομένως, ισχύει ότι

$$|\mathcal{A}_n| = |\mathcal{D}_n| + |\overline{\mathcal{D}}_n| = C_n + |\overline{\mathcal{D}}_n|$$

Για τον υπολογισμό του πληθάριθμου του $\mathcal{A}_n$ παρατηρούμε τα εξής: Κάθε δυαδική λέξη με $n$ μηδενικά και $n$ άσσους καθορίζεται μονοσήμαντα αν επιλέξουμε σε ποιες από τις $2n$ θέσεις της βρίσκονται τα $n$ μηδενικά. Άρα, ο πληθάριθμος του $\mathcal{A}_n$ ισούται με τον αριθμό των συνδυασμών $2n$ στοιχείων ανά $n$.

$$|\mathcal{A}_n| = \binom{2n}{n}.$$

Για τον υπολογισμό του πληθάριθμου του $\overline{\mathcal{D}}_n$ θα ακολουθήσουμε μια προσέγγιση, η οποία ονομάζεται **αρχή της ανάκλασης**. Θα αποδείξουμε ότι κάθε δυαδική λέξη του $\overline{\mathcal{D}}_n$ απεικονίζεται μονοσήμαντα σε μια δυαδική λέξη με $n-1$ άσσους και $n+1$ μηδενικά χωρίς άλλους περιορισμούς.

Έστω μια δυαδική λέξη $\alpha$ που ανήκει στο $\overline{\mathcal{D}}_n$. Επειδή τα μηδενικά είναι περισσότερα από τους άσσους, υπάρχει κάποιο αρχικό τμήμα $\beta$ της $\alpha$ στο οποίο για πρώτη φορά τα 0 είναι περισσότερα από τα 1. Ας συμβολίσουμε με $\gamma$ το υπόλοιπο τμήμα της $\alpha$ που ακολουθεί το $\beta$.

$$\alpha = \underbrace{\beta}_{\text{Το ελάχιστο πρόθεμα με περισσότερα 0 από 1}} \gamma$$

Αν το $\beta$ περιέχει $k$ άσσους και $k+1$ μηδενικά, όπου $k+1 \leqslant n$, τότε το $\gamma$ περιέχει $n-k$ άσσους και $n-k-1$ μηδενικά.

Στη λέξη $\alpha$ εφαρμόζουμε τον εξής μετασχηματισμό: Αφήνουμε το τμήμα $\beta$ όπως είναι και αλλάζουμε κάθε άσσο του τμήματος $\gamma$ σε μηδέν και κάθε μηδενικό του τμήματος $\gamma$ σε άσσο. Η λέξη $\alpha'$ που προκύπτει έχει $k+(n-k-1) = n-1$ άσσους, $k+1+(n-k) = n+1$ μηδενικά και δεν έχει άλλους περιορισμούς.

$$\alpha' = \underbrace{\beta}_{\text{Το ελάχιστο πρόθεμα με περισσότερα 0 από 1}} \underbrace{\gamma'}_{\text{Εναλλάσουμε τα 0 και τα 1 της } \gamma}$$

Για παράδειγμα, στη λέξη $\alpha = 1100010011 \in \overline{\mathcal{D}}_5$ έχουμε $\beta = 11000$ και $\gamma = 10011$, οπότε $\alpha' = 1100001100$. Επίσης, για τη λέξη $\alpha = 0011011010 \in \overline{\mathcal{D}}_5$ έχουμε $\beta = 0$ και $\gamma = 011011010$, οπότε $\alpha' = 0100100101$.

Αντίστροφα, έστω $\alpha'$ μια δυαδική λέξη με $n-1$ άσσους και $n+1$ μηδενικά.

Τότε επειδή τα μηδενικά είναι περισσότερα από τους άσσους, θα υπάρχει κάποιο αρχικό τμήμα $\beta$ της $\alpha'$ στο οποίο για πρώτη φορά τα 0 είναι περισσότερα από τους 1, Ας συμβολίσουμε με $\gamma'$ το υπόλοιπο τμήμα της $\alpha'$ που ακολουθεί το $\beta$.

$$\alpha' = \underbrace{\beta}_{\text{Το ελάχιστο πρόθεμα με περισσότερα 0 από 1}} \gamma'$$

Αν το $\beta$ αποτελείται από $k$ άσσους και $k+1$ μηδενικά, όπου $k+1 \leqslant n+1$, τότε το $\gamma$ περιέχει $n-1-k$ άσσους και $n+1-(k+1) = n-k$ μηδενικά.

Στη λέξη $\alpha'$ εφαρμόζουμε τον εξής μετασχηματισμό: Αφήνουμε το τμήμα $\beta$ όπως είναι και αλλάζουμε κάθε άσσο του τμήματος $\gamma'$ σε μηδέν και κάθε μηδενικό του τμήματος $\gamma'$ σε άσσο. Η λέξη $\alpha$ που προκύπτει έχει $k+(n-k) = n$ άσσους, $k+1+(n-1-k) = n$ μηδενικά και ανήκει στο σύνολο $\overline{\mathcal{D}}_n$.

$$\alpha = \underbrace{\beta}_{\text{το ελάχιστο πρόθεμα με περισσότερα 0 από 1}} \underbrace{\gamma}_{\text{εναλλάσουμε τα 0 και τα 1 της } \gamma'}$$

Για παράδειγμα, στην λέξη $\alpha' = 10100\ 0011010$ που περιέχει $n-1 = 5$ άσσους και $n+1 = 7$ μηδενικά έχουμε $\beta = 10100$ και $\gamma' = 0011010$, οπότε $\alpha = 10100\ 1100101$.

Επομένως, επειδή η διαδικασία αυτή είναι μονοσήμαντη προκύπτει ότι ο αριθμός των λέξεων του $\overline{\mathcal{D}}_n$ ισούται με τον αριθμό των δυαδικών λέξεων με $n-1$ άσσους και $n+1$ μηδενικά.

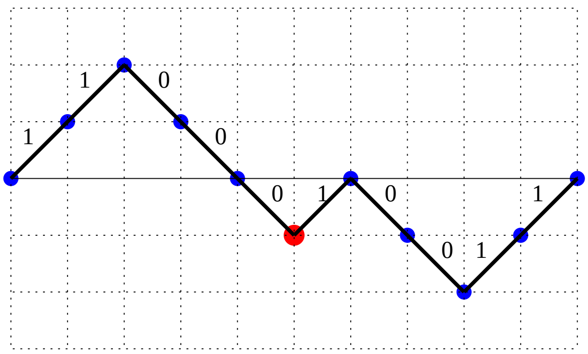
Ο αριθμός των δυαδικών λέξεων με $n-1$ άσσους και $n+1$ μηδενικά, υπολογίζεται εύκολα, όπως ο πληθάριθμος του $\mathcal{A}_n$. Κάθε τέτοια λέξη καθορίζεται μονοσήμαντα αν επιλέξουμε σε ποιες από τις $2n$ θέσεις της βρίσκονται οι $n-1$ άσσοι, οπότε υπάρχουν $\binom{2n}{n-1}$ τέτοιες λέξεις, άρα

$$|\overline{\mathcal{D}}_n| = \binom{2n}{n-1}.$$

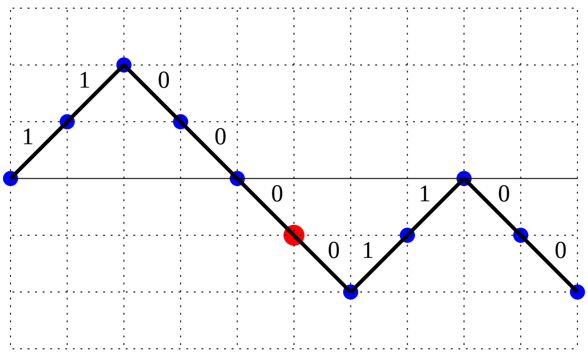
Επομένως,

$$\begin{aligned} C_n &= |\mathcal{D}_n| = |\mathcal{A}_n| - |\overline{\mathcal{D}}_n| = \binom{2n}{n} - \binom{2n}{n-1} \\ &= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!} = \left(1 - \frac{n}{n+1}\right)\frac{(2n)!}{n!n!} = \frac{1}{n+1}\binom{2n}{n}. \end{aligned}$$ □

**Παρατήρηση:** Η ονομασία αρχή της ανάκλασης ερμηνεύεται αν αναπαραστήσουμε τις δυαδικές λέξεις από μονοπάτια με βήματα $(1,1)$ για κάθε άσσο και $(1,-1)$ για κάθε μηδενικό. Τα στοιχεία του $\mathcal{A}_n$ αντιστοιχούν σε μονοπάτια που αρχίζουν από το $(0,0)$ και τελειώνουν στο $(2n,0)$ χωρίς άλλους περιορισμούς. Τα στοιχεία του $\mathcal{D}_n$ αντιστοιχούν σε μονοπάτια του $\mathcal{A}_n$ που δεν διέρχονται κάτω από τον οριζόντιο άξονα $x$, ενώ τα στοιχεία του $\overline{\mathcal{D}}_n$ αντιστοιχούν σε μονοπάτια του $\mathcal{A}_n$ που διέρχονται τουλάχιστον μια φορά κάτω από τον οριζόντιο άξονα $x$.




$\mapsto$




$\alpha = 11000\ 10011$
(5 άσσοι, 5 μηδενικά)

$\alpha' = 11000\ 01100$
(4 άσσοι, 6 μηδενικά)

Η διάσπαση που χρησιμοποιήθηκε στην προηγούμενη απεικόνιση, εντοπίζει το αριστερότερο σημείο ενός μονοπατιού του $\mathcal{D}_n$ το οποίο βρίσκεται κάτω από τον οριζόντιο άξονα $x$, και αντανακλά τα βήματα που ακολουθούν μετά από αυτό το σημείο αλλάζοντας κάθε $(1,1)$ σε $(1,-1)$ και αντιστρόφως, οπότε το μονοπάτι που ξεκινά από αυτό το σημείο αντανακλάται πάνω - κάτω.

Επιπλέον, από τον προηγούμενο τύπο προκύπτει μια νέα αναδρομική σχέση για τους αριθμούς Catalan, η οποία είναι χρήσιμη για υπολογισμούς

$$\frac{C_{n+1}}{C_n} = \frac{\frac{1}{n+2}\binom{2(n+1)}{n+1}}{\frac{1}{n+1}\binom{2n}{n}} = \frac{\frac{(2(n+1))!}{(n+1)!(n+2)!}}{\frac{(2n)!}{n!(n+1)!}} = \frac{(2n+1)(2n+2)}{(n+1)(n+2)} = \frac{2(2n+1)}{n+2}.$$

$$\boxed{C_{n+1} = \frac{2(2n+1)}{n+2}C_n, \text{ για κάθε } n \geqslant 0.}$$

**Κατασκευή των λέξεων Dyck μήκους $2n$**

Κάθε λέξη Dyck μήκους $2n$ καθορίζεται μοναδικά από τις θέσεις στις οποίες εμφανίζονται οι $n$ άσσοι της, επομένως μια εναλλακτική κωδικοποίηση των λέξεων Dyck είναι η ακολουθία $n$ ακεραίων $a_1, a_2, \ldots, a_n$ όπου $a_i$ είναι η θέση εμφάνισης του $i$-οστού άσσου.
**Παράδειγμα:** Οι 14 λέξεις Dyck μήκους 8

| | | | | |
|---|---|---|---|---|
| 11110000 | 11101000 | 11100100 | 11100010 | 11011000 |
| 11010100 | 11010010 | 11001100 | 11001010 | 10111000 |
| 10110100 | 10110010 | 10101100 | 10101010 | |

μπορούν αντίστοιχα να κωδικοποιηθούν από τις ακολουθίες

| | | | | |
|---|---|---|---|---|
| 1234 | 1235 | 1236 | 1237 | 1245 |
| 1246 | 1247 | 1256 | 1257 | 1345 |
| 1346 | 1347 | 1356 | 1357 | |

Η κωδικοποίηση των λέξεων Dyck μέσω της ακολουθίας των άσσων είναι πολύ βολική για την λεξικογραφική κατασκευή. Συγκεκριμένα, ισχύει η επόμενη πρόταση.

**Πρόταση 2.4.3.** *Αν $(a_i)_{i\in[n]}$ είναι η ακολουθία των θέσεων εμφάνισης του $i$-στου άσσου σε μια λέξη Dyck τότε*

i) *$a_i < a_{i+1}$ για κάθε $i \in [n-1]$*

ii) *$i \leqslant a_i \leqslant 2i-1$ για κάθε $i \in [n]$ ισχύει ότι*

*Σε κάθε ακολουθία $(a_i)_{i\in[n]}$ με τις παραπάνω ιδιότητες αντιστοιχεί μια λέξη Dyck.*

Με την αναπαράσταση αυτή

- η μικρότερη λεξικογραφικά λέξη Dyck μήκους $2n$ είναι η λέξη $12\cdots i\cdots n$.
- η μεγαλύτερη λεξικογραφικά λέξη Dyck μήκους $2n$ είναι η λέξη $13\cdots(2i-1)\cdots(2n-1)$.
- ο $i$-οστός άσσος βρίσκεται μεταξύ των θέσεων $i$ και $2i-1$.
- κάθε λέξη Dyck (εκτός από την μεγαλύτερη) γράφεται κατά μοναδικό τρόπο στην μορφή

$$a_1a_2\cdots a_i\underbrace{(2i+1)(2i+3)\cdots(2n-1)}_{n-i \text{ όροι}} \text{ όπου } i \leqslant n \text{ και } a_i < 2i-1.$$

  Με άλλα λόγια σε κάθε λέξη Dyck (εκτός από την μεγαλύτερη) υπάρχει ο δεξιότερος άσσος ο οποίος δεν βρίσκεται στην δεξιότερη δυνατή θέση που μπορεί να εμφανισθεί. Μετά από αυτόν, αν υπάρχουν άλλοι άσσοι αυτοί βρίσκονται όλοι στις δεξιότερες δυνατές θέσεις που μπορεί να εμφανίζεται ο καθένας.

- για κάθε λέξη Dyck (εκτός από την μεγαλύτερη) με κωδικοποίηση

$$a_1a_2\cdots a_i\underbrace{(2i+1)(2i+3)\cdots(2n-1)}_{n-i \text{ όροι}} \text{ όπου } i \leqslant n \text{ και } a_i < 2i-1$$

  η (αμέσως) επόμενη είναι η λέξη

$$a_1a_2\cdots(a_i+1)\underbrace{(a_i+2)(a_i+3)\cdots(a_i+(n-i)+1)}_{n-i \text{ όροι}}.$$

Με άλλα λόγια η επόμενη λέξη Dyck προκύπτει μεταφέροντας τον δεξιότερο άσσο της διάσπασης μια θέση δεξιά και όλους τους υπόλοιπους άσσους που ακολουθούν αμέσως μετά τον άσσο αυτό.

Με βάση τις προηγούμενες παρατηρήσεις ένας γρήγορος τρόπος κατασκευής όλων των λέξεων Dyck μήκους $2n$ δίδεται στο επόμενο πρόγραμμα.

```
    #generate the first dyck word
    #note that dyck[pos] = a_(pos+1)
    dyck = [i for i in range(1,n+1)]
    print(dyck)
    #count words
    counter = 1
    #read the word from right to left
    #pos is the position of the rightmost 1 that
    #can be moved one position to the right
    pos = n
    while(pos > 1):
        pos -=1
        #if a_i < 2i - 1 aka dyck[pos] < 2(pos+1) -1
        #increase a_i by 1
        #set all remaining 1 (if any) to their min values
        if(dyck[pos] < 2*(pos+1)-1):
            dyck[pos] += 1
            for j in range(pos+1,n):
                dyck[j] = dyck[j-1]+1
            print(dyck)
            counter += 1
            pos = n
    return counter

n = 4
print("In total, there are", genDyck(n), "Dyck words of length",2*n)
```

Output:

```
[1, 2, 3, 4]
[1, 2, 3, 5]
[1, 2, 3, 6]
[1, 2, 3, 7]
[1, 2, 4, 5]
[1, 2, 4, 6]
[1, 2, 4, 7]
[1, 2, 5, 6]
[1, 2, 5, 7]
[1, 3, 4, 5]
[1, 3, 4, 6]
[1, 3, 4, 7]
[1, 3, 5, 6]
[1, 3, 5, 7]
In total, there are 14 Dyck words of length 8
```