Χρήση δέντρων στην ανάλυση αλγόριθμων και κάτω φράγματα στα προβλήματα αναζήτησης και διάταξης λίστας

Δέντρο απόφασης για τον αλγόριθμο A:

(0) ο A είναι σωστός

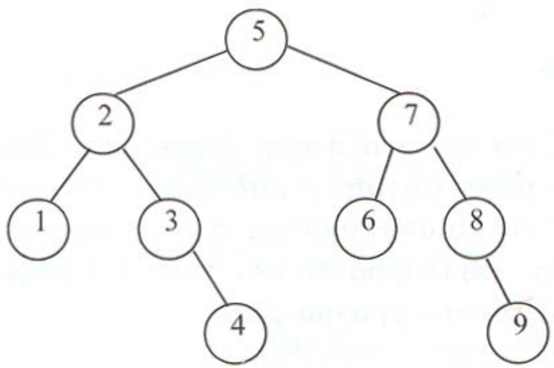
(1) ο A πρώτα διαλέγει έναν αρχικό δείκτη i_o, μεταξύ 1 και n, και κάνει την πρώτη σύγκριση (του x με το $L(i_o)$).

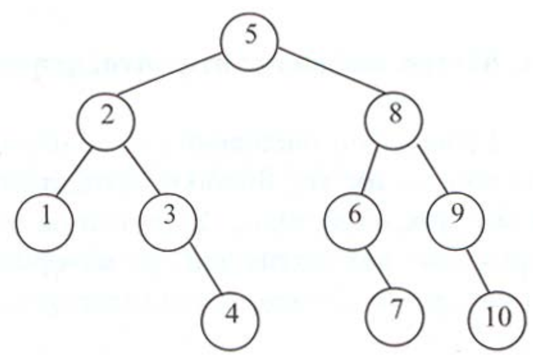
(2) μετά από κάθε σύγκριση του x με κάποιο $L(i)$ ο A κάνει **ένα** από τα παρακάτω:

[2φ] τερματίζει

[2α ή δ] διαλέγει **έναν από δυο το πολύ** δείκτες i_α ή i_δ (πάντα μεταξύ 1 και n) και κάνει την αντίστοιχη σύγκριση (του x με **ένα** από τα $L(i_\alpha)$ ή $L(i_\delta)$).

Για τον αλγόριθμο της δυαδικής αναζήτησης και για $n=9$, $n=10$ τα αντίστοιχα δέντρα απόφασης είναι αυτά





ΛΗΜΜΑ Λ1.11-1. Για κάθε $A \in Q$, το αντίστοιχο (δυαδικό) δέντρο απόφασης έχει τουλάχιστον n κόμβους.

Απόδειξη: Αν είχε λιγότερους, ας είναι i ένας από τους κόμβους που λείπουν. Ας θεωρήσουμε δύο διαφορετικά στιγμιότυπα:

(i) $x=L(i)$ και $x \neq L(j)$ για $j \neq i$,

(ii) $x \neq L(j)$ για $j \neq i$ και για $j = i$.

Αφού λείπει ο i, ο A δεν συγκρίνει ποτέ το x με το $L(i)$.

Άρα και στις δύο περιπτώσεις θα βγάλει το ίδιο αποτέλεσμα.

ο A δεν είναι σωστός, πράγμα που αντικρούει ότι $A \in Q$.

ΛΗΜΜΑ Λ1.11-2. Κάθε δυαδικό δέντρο βάθους d έχει το πολύ $2^{d+1} - 1$ κόμβους.
Προφανές.

ΛΗΜΜΑ Λ1.11-3. Για κάθε $A \in Q$, $ΠΧΠ_A(n) = d+1$, όπου $d =$ βάθος δέντρου του A.
Προφανές.

ΘΕΩΡΗΜΑ Θ1.11-4. Για κάθε $A \in Q$ και για κάθε n

$$\lfloor log n \rfloor + 1 \leq ΠΧΠ_A(n).$$

Απόδειξη:

$$\begin{aligned} n &\leq \text{πλήθος κόμβων δέντρου} && \text{(από } \textit{Λήμμα Λ1.11-1}\text{)} \\ &\leq 2^{d+1} - 1 && \text{(από } \textit{Λήμμα Λ1.11-2}\text{)} \\ &= 2^{ΠΧΠ_A(n)} - 1 && \text{(από } \textit{Λήμμα Λ1.11-3}\text{)} \end{aligned}$$

$$ΠΧΠ_A(n) \geq log(n+1)$$

Κάτω φράγμα για το πρόβλημα της διάταξης

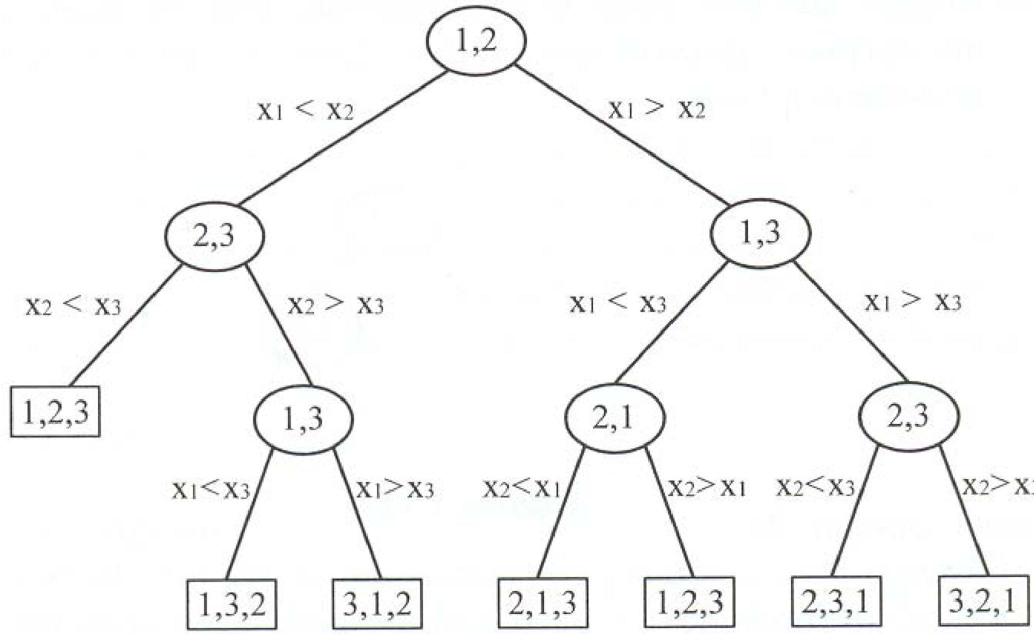
Θεωρούμε την κατηγορία Q αλγορίθμων που είναι σωστοί και αποτελούνται από εντολές του εξής είδους:

i) σύγκρινε x_i με x_j,

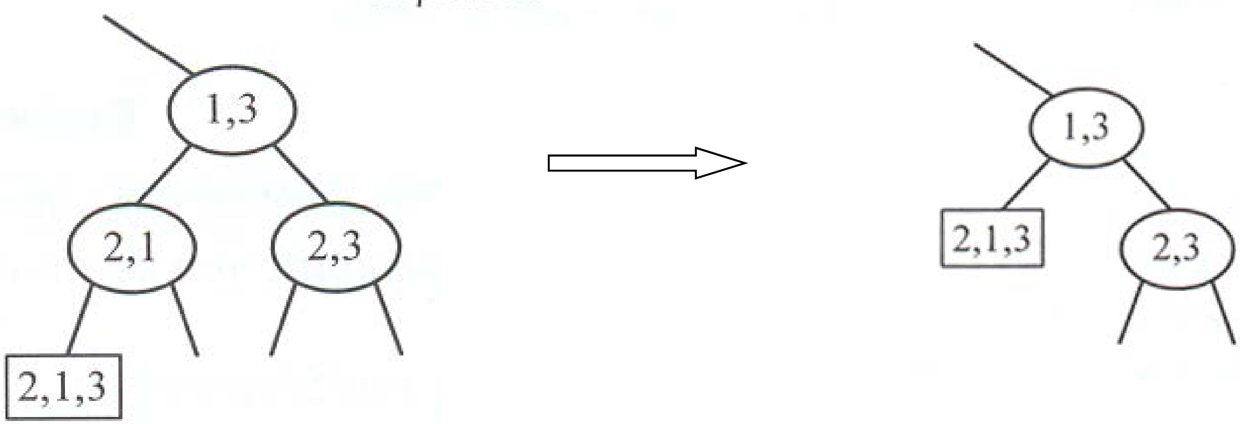
ii) δώσε αποτέλεσμα {δηλαδή την διαταγμένη λίστα. Αν η αρχική λίστα είναι $x_1,\ldots,x_n$, η διαταγμένη λίστα θα είναι μια αναδιάταξη αυτών των στοιχείων δηλαδή $x_{i_1},\ldots,x_{i_n}$ },

iii) τερμάτισε.

ΠΑΡΑΔΕΙΓΜΑ Π1.11-2. Να δοθεί το δέντρο που αντιστοιχεί στον *bubblesort* για $L=\{x_1, x_2, x_3\}$.

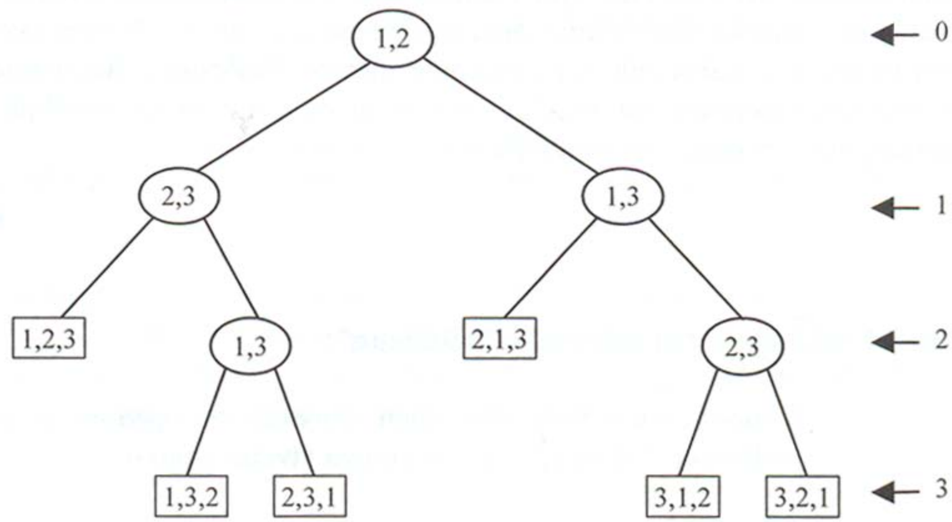


"κλάδεμα" Κλαδιά στα οποία δεν είναι δυνατόν να φθάσει ο αλγόριθμος κόβονται.



δέντρα b ή **2-δέντρα** κάθε εσωτερικός κόμβος έχει ακριβώς δυο παιδιά

Το κλαδεμένο δέντρο



ΠΧΤΠ(n) = *d* (το βάθος του δέντρου) και

ΠΜΟ(n)= μέσος όρος μηκών των διαδρομών (αν όλες οι περιπτώσεις είναι ισοπίθανες)

ΠΧΤΠ(n)=3 *ΠΜΟ(n)=1/6 (2 + 3 + 3 + 2 + 3 + 3) = 16/6 = 8/3*

$\varphi(d)$ το πλήθος των φύλλων σε ένα δυαδικό δέντρο βάθους d.

ΛΗΜΜΑ Λ1.11-5. Ισχύει ότι $\varphi(d) \leq 2^d$.

Απόδειξη: Επαγωγή.

ΛΗΜΜΑ Λ1.11-6. Ισχύει ότι $d \geq \lceil \log \varphi(d) \rceil$.

Απόδειξη: Από Λήμμα *Λ1.11-4*.

ΛΗΜΜΑ Λ1.11-7: (Κάτω φράγμα). Για κάθε $A \in Q$ $\lceil \log(n!) \rceil \leq ΠΧΠ(n)$.

Απόδειξη: Από Λήμμα *Λ1.11-6* και από $\varphi = n!$

$$log(n!) = log(1 \cdot 2 \cdot \ldots \cdot n) = \sum^{n} log\, j \geq \int_1^n log(x)dx =$$

$$= \int_1^n log(e) \cdot ln(x)dx = log\, e \cdot (n\, ln(n) - n + 1) = n\, log\, n - n\, log\, e + log\, e$$

και με $log\, e \approx 1.44$,

$$log(n!) \approx n\, log\, n - 1.44n + 1.44 \approx n\, log\, n - 1.5n$$

ΘΕΩΡΗΜΑ Θ1.11-8. Κάτω φράγμα για την διάταξη λίστας και την Π×Π.

Για κάθε $A \in Q$

$$ΠΧΠ(n) \geq \lceil log(n!) \rceil \approx \lceil n\, log\, n - 1.5n \rceil.$$

Μέση Περίπτωση

Υπάρχουν $n!$ διαδρομές.

a το άθροισμα των μηκών όλων των διαδρομων απο την ρίζα σε φύλλο.

$$ΠΜΟ(n) = a/n!$$

Για κάθε αλγόριθμο ταξινόμησης βασιζόμενο μόνο σε συγκρίσεις ο μέσος αριθμός συγκρίσεων είναι αυστηρά μεγαλύτερος του

$$\lfloor \lceil \log n! \rceil \rfloor \approx \lfloor n \log n - 1.5n \rfloor$$

- Σε κάθε επίπεδο ενός δέντρου b, εκτός από το τελευταίο, υπάρχει τουλάχιστον ένας κόμβος με δυο παιδιά.
- Αν σ' ένα δέντρο b, που έχει όλα τα φύλλα του στα δύο τελευταία επίπεδα, ισχύει $\varphi = 2^p$ τότε $p = d$ (p ακέραιος)
- Σ' ένα δέντρο b που έχει όλα τα φύλλα του στα δύο τελευταία επίπεδα, ισχύει $d - 1 = \lfloor log\varphi \rfloor$ και $d = \lceil log\varphi \rceil$

ΛΗΜΜΑ Λ1.11-9. Απ' όλα τα δέντρα b με φ φύλλα, αυτά για τα οποία τα φύλλα είναι μαζεμένα σε δύο συνεχόμενα επίπεδα (άρα στα δύο τελευταία επίπεδα d-1 και d) έχουν το μικρότερο a.

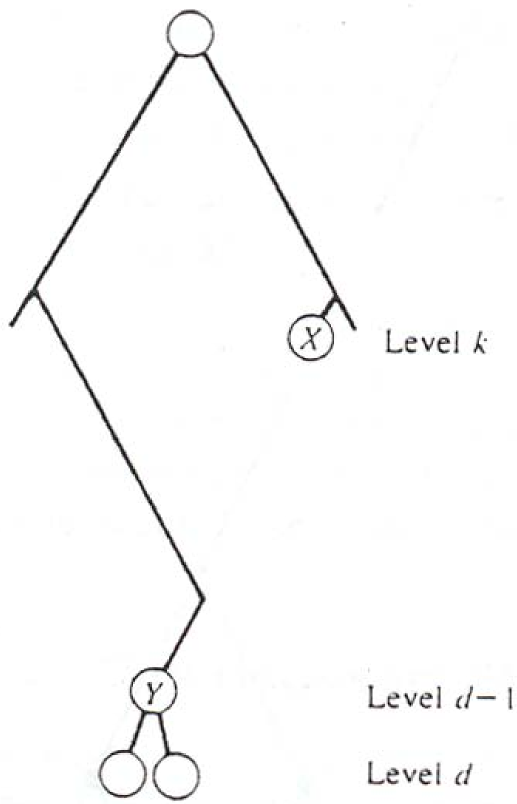
Απόδειξη: Ας θεωρήσουμε ένα δέντρο b που έχει φύλλο X σε κάποιο επίπεδο k, άλλο από τα δύο τελευταία δηλαδή

$$k < d-1$$

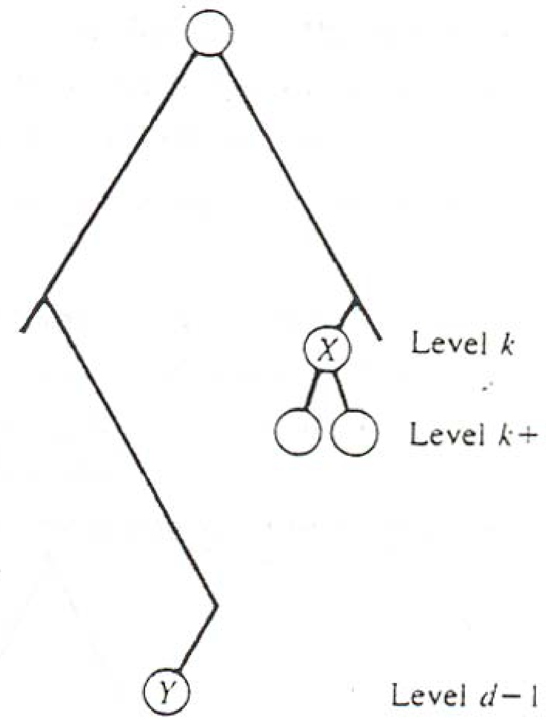
Αρκεί να κατασκευάσουμε ένα νέο δέντρο b, που έχει τον ίδιο αριθμό φύλλων αλλά μικρότερο a:

Διαλέγουμε έναν κόμβο Y, βαθμού 2, στο επίπεδο d-1,
παίρνουμε τα δύο παιδιά του (φύλλα) και τα κάνουμε παιδιά (φύλλα) του X.
Y είναι φύλλο και ο X έχει βαθμό 2 (δεν είναι πλέον φύλλο).

Πώς άλλαξε το a;
συνολικά μειώθηκε κατά $(2d+k)-[d-1+2(k+1)] = d-k-1$



The given 2-tree with l leaves.



Modified 2-tree with l leaves and external path length decreased by $d-k-1$.

ΛΗΜΜΑ Λ1.11-10. Για δέντρα b με φ φύλλα το ελάχιστο a είναι

$$a_{\varepsilon\lambda\alpha\chi.} = \varphi\lfloor log\varphi \rfloor + 2(\varphi - 2^{\lfloor log\varphi \rfloor}).$$

Απόδειξη: Αφού ψάχνουμε για το $a_{\varepsilon\lambda\alpha\chi.}$ περιοριζόμαστε σε δέντρα b που έχουν όλα τα φύλλα τους στα επίπεδα d-1 και d (Λήμμα *Λ1.11-9*).

$\varphi=2^p$, p ακέραιος. $\Longrightarrow$ $p=d$ όλα τα φύλλα είναι στο επίπεδο d

$a_{\varepsilon\lambda\alpha\chi.}=d\varphi.$

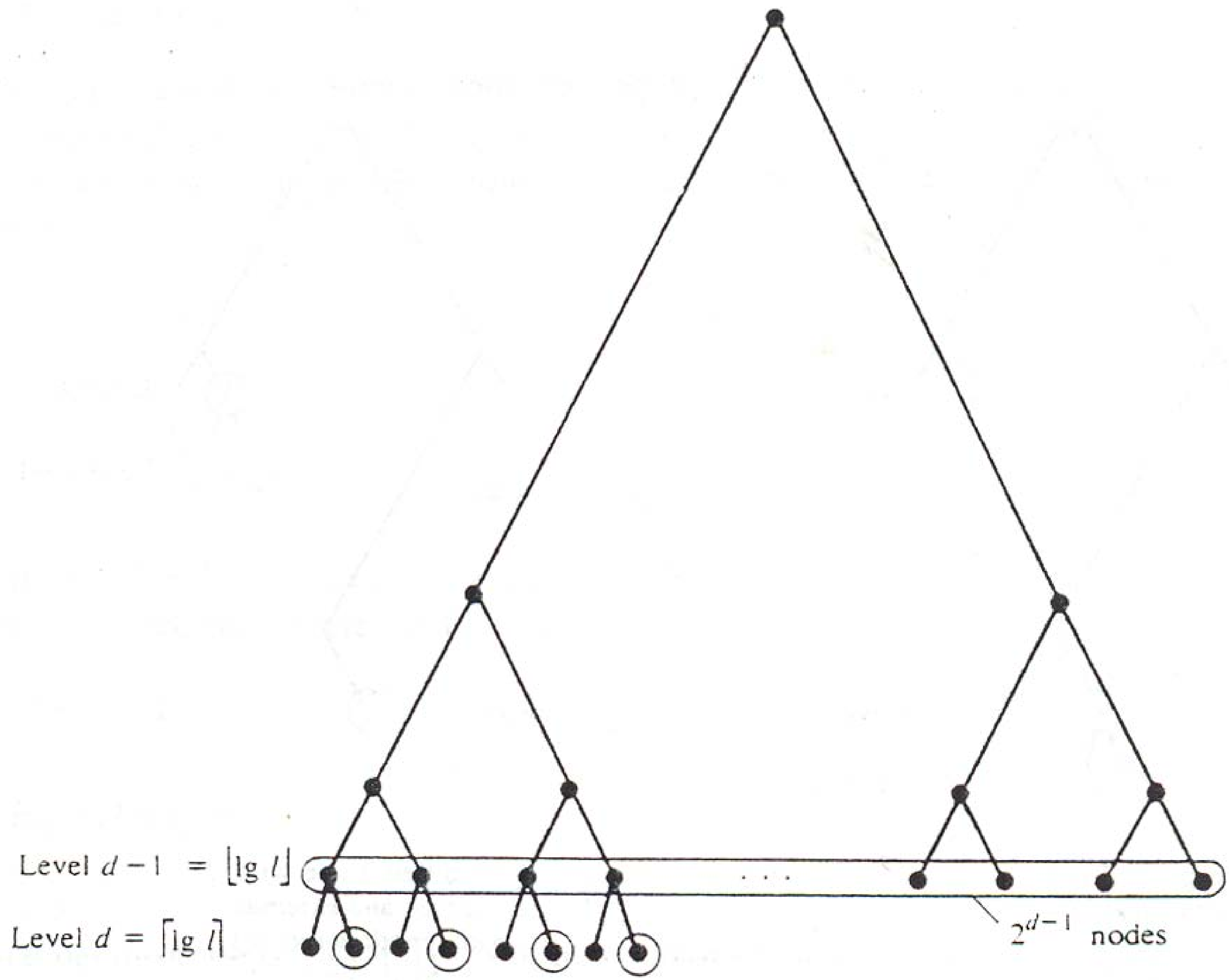
$\lfloor log\varphi \rfloor = \lfloor log 2^d \rfloor = \lfloor d \rfloor = d$, οπότε $a_{\varepsilon\lambda\alpha\chi.} = d\varphi + 2(\varphi - 1) = d\varphi.$

Αν ο φ δεν είναι δύναμη το 2 τότε έχουμε φ_1 φύλλα στο επίπεδο d-1 και φ-φ_1 φύλλα στο d.

$$a_{\varepsilon\lambda\alpha\chi.} = \varphi_1(d-1)+(\varphi-\varphi_1)d.$$

φ_1 φύλλα στο d-1 $\Longrightarrow$ 2^{d-1} - φ_1 εσωτερικοί κόμβοι στο d-1

υπάρχουν $2(2^{d-1} - \varphi_1)$ φύλλα στο d. Άρα, φ-φ_1 $= 2^d$ - $2\varphi_1$


Level $d-1 = \lfloor \lg l \rfloor$
Level $d = \lceil \lg l \rceil$
. . .
2^{d-1} nodes

Κάτω όριο στη ΠΜΟ των αλγόριθμων ταξινόμησης

- Επομένως:
  - $\varphi_1 = 2^d - \varphi$
  - Άρα:
  - $\alpha_{\varepsilon\lambda\alpha\chi} = 2^d + \varphi d + \varphi$
  - Όμως, d = $\lfloor \log \varphi \rfloor + 1$, Άρα

$$a_{\varepsilon\lambda\alpha\chi} = \varphi \lfloor \log \varphi \rfloor + 2(\varphi - 2^{\lfloor \log \varphi \rfloor})$$

  - Έχουμε φ=n!, και ΠΜΟ(n) ≥ $\alpha_{\varepsilon\lambda\alpha\chi}/\varphi \geq \lfloor \log \varphi \rfloor$
  - Άρα ΠΜΟ(n) $= \lfloor \log n! \rfloor \approx \lfloor n \log n - 1.5n \rfloor$