

## **3<sup>η</sup> Δραστηριότητα**

---

### **Μπορείς να το ξαναπείς! - Συμπίεση κειμένου**

#### **Περίληψη**

Δεδομένου ότι οι υπολογιστές έχουν μόνο ένα περιορισμένο χώρο όπου να αποθηκεύουν τις πληροφορίες, θα πρέπει να τις αναπαριστούν με τον πλέον αποτελεσματικό και εφικτό τρόπο. Χρησιμοποιούνται λοιπόν μέθοδοι συμπίεσης της πληροφορίας. Συμπιέζοντας τα δεδομένα προς απομνημόνευση και αποκωδικοποιώντας τα για την εκ νέου ανάγνωση, καθίσταται δυνατόν να διατηρηθούν περισσότερες πληροφορίες ή να σταλούν πιο γρήγορα μέσω του Διαδικτύου.

#### **Αντιστοιχία με το σχολικό πρόγραμμα \* :**

(\*Σημ. μτφ.: αναφέρεται στη Νέα Ζηλανδία)

- ✓ Αγγλικά: Αναγνωρίζοντας patterns με λέξεις και κείμενα.
- ✓ Τεχνολογία: Τεχνολογική γνώση και αντίληψη. Πως δουλεύουν οι υπολογιστές.

#### **Απαιτούμενες δεξιότητες**

- ✓ Αντιγραφή γραπτού κειμένου

#### **Ηλικία**

- ✓ Από 9 ετών και άνω

#### **Υλικό**

- ✓ Διαφάνειες για προβολή, φωτοτυπώντας το πρωτότυπο “Μπορείς να το ξαναπείς!” (σελ. 25)

Κάθε μαθητής θα πρέπει να έχει:

- ✓ Το φύλλο εργασίας: Μπορείς να το ξαναπείς! (σελ. 26)
- ✓ Το φύλλο εργασίας: extra για ειδικούς (σελ. 27)
- ✓ Το φύλλο εργασίας: Ο γιος του Ρουμπή, του Κουμπή (σελ. 28)
- ✓ Το φύλλο εργασίας: Extra για πραγματικούς ειδικούς (σελ. 29)

# Μπορείς να το ξαναπείς!

---

## Εισαγωγή

Οι υπολογιστές πρέπει να αποθηκεύουν και να αποστέλλουν πάμπολλα δεδομένα. Άρα λοιπόν, πρέπει να αποφεύγουν να χρησιμοποιούν πολύ χώρο ή να χρειάζονται πολύ ώρα για τη μετάδοση στα δίκτυα. Αυτή η δραστηριότητα δείχνει πως μπορούν να συμπιεσθούν τα δεδομένα.

## Επίδειξη και συζήτηση.

Προβάλλετε τη διαφάνεια στη σελ. 25. Προσέξτε ποιες σειρές γραμμάτων επαναλαμβάνονται μέσα στο κείμενο. Μπορεί να εντοπίσετε ομάδες των δύο ή περισσοτέρων γραμμάτων που επαναλαμβάνονται, ή και ολόκληρες λέξεις ή φράσεις; Αντικαταστήστε τις επαναλαμβανόμενες σειρές, σχεδιάζοντας ένα τετράγωνο τριγύρω τους, όπως στο διάγραμμα που ακολουθεί.

## Ο παπάς ο παχύς, έφαγε παχιά φακή

▽

Ο παπάς ο ...χύσ...έφαγε ....χά....κή

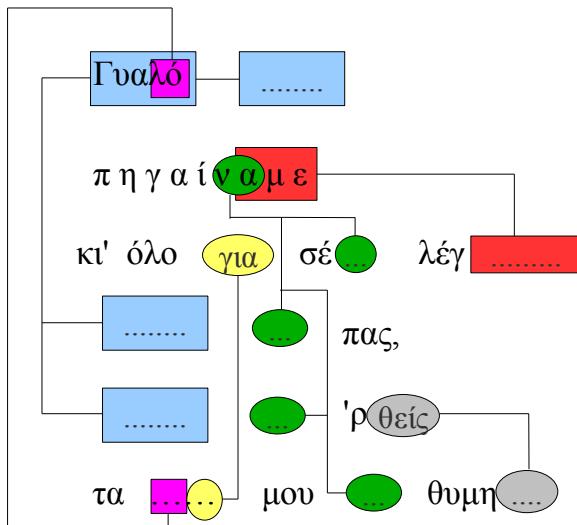


**Πρότυπο για φωτοτυπία: Μπορείς να το ξαναπείς!**

**Ο παπάς ο παχύς  
έφαγε παχιά φακή.  
γιατί παπά παχύ<sup>1</sup>  
έφαγες παχιά φακή ;**

## Φύλλο εργασίας: Μπορείς να το ξαναπείς!

Πολλά από τα γράμματα και τις λέξεις του ποιήματος που ακολουθεί, λείπουν. Μπορείτε να συμπληρώσετε τα κενά τόσο στα γράμματα όσο και στις λέξεις, ούτως ώστε να ολοκληρωθεί σωστά το κείμενο; Θα βρείτε τα γράμματα και τις λέξεις που λείπουν από τα πλαισια που δείχνουν τα βέλη.



Τώρα μπορείτε να διαλέξετε άλλα απλά ποιηματάκια ;ή νανουρίσματα και να δημιουργήσετε ένα παρόμοιο puzzle. Βεβαιωθείτε ότι τα βελάκια δείχνουν πάντα προς ένα προηγούμενο σημείο του κειμένου. Η ποίησή σας πρέπει πάντα να ανασυντίθεται ανατρέχοντας το κείμενο εξ αριστερών προς τα δεξιά, από πάνω προς τα κάτω, κατά τη συνηθισμένη φορά ανάγνωσης.

**Πρόκληση:** Να δείτε πόσες λίγες λέξεις και γράμματα της αρχικής ποίησης χρειάζεται να κρατήσουμε!

Άλλες ιδέες: χρησιμοποιήστε διάφορους γλωσσοδέτες, είτε ποιήματα ή λέξεις τραγουδιών με επαναλήψεις (“Φεγγαράκι μου λαμπρό, φέγγε μου να περπατώ...”, “Η γιαγιά μας η καλή, έχει κότες στην αυλή...”, κλπ).

**Υπόδειξη:** Προσπαθείστε να αποφύγετε έναν συνωστισμό βελών. Αφήστε πολύ χώρο ανάμεσα στα γράμματα και τις λέξεις, ούτως ώστε να περισσεύει και χώρος για τον σχεδιασμό των πλαισίων και των βελών που δείχνουν προς αυτά.

Είναι πιο εύκολο να σχεδιάσετε το puzzle, εάν πρώτα γράψετε το ποίημα και μετά αποφασίσετε που πρέπει να βρίσκονται τα πλαισια που θα σχεδιάσετε.

## Φύλλο εργασίας: Extra για ειδικούς

Πώς θα λύσετε αυτό το puzzle;

B a n - - -

Μερικές φορές, τα μέρη του κειμένου που λείπουν, δείχνουν, κατά κάποιον τρόπο, προς ένα τμήμα του. Σε αυτή την περίπτωση, μπορεί να αποκωδικοποιηθεί σωστά, εάν αντιγραφούν τα γράμματα εξ αριστερών προς τα δεξιά. Κατ' αυτό τον τρόπο, το κάθε γράμμα καθίσταται διαθέσιμο προς αντιγραφή, πριν καν χρειασθεί. Αυτή η μέθοδος είναι χρήσιμη όταν οι υπολογιστές συναντούν μεγάλες σειρές με ένα συγκεκριμένο γράμμα ή με μία ομάδα γραμμάτων.

Μπορείτε να δημιουργήσετε και άλλα παραδείγματα και από μόνοι σας.

Οι υπολογιστές αναπαριστούν τα πλαίσια και τα βελάκια με αριθμούς, π.χ.:

**Banana**

μπορεί να γραφεί σαν Ban(2,3). “2” σημαίνει ότι πρέπει να πάμε πίσω κατά 2 γράμματα, για να βρούμε το ακριβές σημείο απ’ όπου θα αρχίσει η αντιγραφή.

**Ban---**

και “3” σημαίνει ότι πρέπει να αντιγραφούν τρία γράμματα στη σειρά.

**Bana--**



**Banan-**

**Banana**

Καθώς χρησιμοποιούνται 2 αριθμοί για την κωδικοποίηση αυτών των λέξεων, συνήθως, αξίζει να συμπιεσθούν μόνον ομάδες 2 ή παραπάνω γραμμάτων, ειδάλλως δεν υπάρχει καμία εξοικονόμηση χώρου. Πράγματι, το μήκος ενός αρχείου θα μάκραινε εάν χρησιμοποιούσαμε 2 αριθμούς για την κωδικοποίηση ενός γράμματος.

Γράψτε μερικές δικές σας λέξεις, με τον τρόπο που θα τις έγραφε ο υπολογιστής, εάν ήταν συμπιεσμένες. Θα καταφέρουν οι φίλοι σας να τις αποκωδικοπούσουν;

## **Φύλλο εργασίας: Ο γιος του Ρουμπή, του Κουμπή...**

**Πόσες λέξεις είναι πραγματικά απαραίτητες εδώ;**

Φανταστείτε ότι είστε ένας υπολογιστής που θέλει να αποθηκεύσει στο δίσκο όσες περισσότερες πληροφορίες μπορεί. Βρείτε όλες τις σειρές 2 ή περισσότερων γραμμάτων που εμφανίσθηκαν ήδη μέσα στο κείμενο. Αυτές οι σειρές δεν είναι πλέον χρήσιμες και μπορούν να αντικατασταθούν από ένα καταδεικτικό (ένα κενό πλαίσιο, με ένα βέλος που να καταδεικνύει προς ένα άλλο πλαίσιο, με τα γράμματα προς αντιγραφή). Ο στόχος σας είναι να εντοπίσετε όσο περισσότερες σειρές μπορείτε και να τις αντικαταστήσετε.

**Ο γιος του Ρουμπή, του Κουμπή**

**του ρουμπο - κομπολογή,**

**βγήκε να ρουμπέψει, να κουμπέψει,**

**να ρουμποκομπολογέψει,**

**και τον πιάσαν οι ρουμπήδες, οι κουμπήδες,**

**οι ρουμπο-κομπολογήδες**

## Φύλλο εργασίας: Extra για ειδικούς

**Είστε έτοιμοι για μία αληθινά σκληρή συμπίεση;**

Το παραμύθι που ακολουθεί αναλύθηκε από ένα πρόγραμμα υπολογιστή και βρήκε πως υπάρχουν τουλάχιστον 499 αλληλουχίες και 2281 γράμματα που μπορούν να απαλειφθούν, καθόσον ανήκουν σε επαναλαμβανόμενες σειρές (ο διαχωρισμός ανάμεσα σε σειρές και παραγράφους του κειμένου, εκλαμβάνεται ως ένα διάστημα). Πόσες θα καταφέρετε να διαγράψετε; Να θυμάστε πως μόνο οι σειρές με 2 ή παραπάνω επαναλαμβανόμενους χαρακτήρες μπορούν να διαγραφούν. Καλή τύχη!

~~~~~ ταν μια φορά κι' έναν καιρό, πριν πολλά πολλά χρόνια, τρία γουρουνάκια που θέλανε να πάνε να ζήσουν μόνα τους. Το πρώτο γουρουνάκι δεν ήταν και πολύ πονηρό και αποφάσισε να χτίσει το σπίτι του από άχυρο, γιατί ήταν φθηνό. Και το δεύτερο γουρουνάκι δεν ήταν πονηρό και έχτισε το σπίτι του από ξύλο, διότι έτσι είχε ένα οικολογικό look και ήταν και της μόδας. Το τρίτο γουρουνάκι ήταν πιο πονηρό από τους άλλα δύο αδέλφια του και αγόρασε ένα φορτίο τούβλα από μία κοντινή πόλη και έχτισε ένα σπίτι ταπεινό αλλά άνετο.

Λίγο καιρό μετά τη γιορτή εγκαινίων του σπιτιού, το πρώτο γουρουνάκι καθόταν σε μία καρέκλα και ήταν απορροφημένο διαβάζοντας ένα βιβλίο, όταν άκουσε να χτυπάνε στη πόρτα. Ήταν ο κα-κός λύκος, φυσικά.

"Γουρουνάκι, γουρουνάκι, άσε με να μπω!" φώναξε ο λύκος.

"Μα ούτε για μία τρίχα από το δέρμα μου!" απήντησε το πρώτο γουρουνάκι.

"Τότε θα φυσήξω, θα ξεφυσήξω και θα γκρεμίσω το σπίτι σου" μούγκρισε ο λύκος και πήρε μια βαθιά ανάσα, φύσηξε και το σπίτι γρήγορα γκρεμίστηκε. Το πρώτο γουρουνάκι έτρεξε όσο πιο γρήγορα μπορούσε μέχρι το ξύλινο σπίτι και γρήγορα βρήκε εκεί καταφύγιο με τον αδελφό του. Άλλα, πριν περάσει πολλή ώρα, ο λύκος έφθασε και εκεί.

"Γουρουνάκι, γουρουνάκι, άσε με να μπω!" φώναξε ο λύκος.

"Μα ούτε για μία τρίχα από το δέρμα μου!" απήντησε το δεύτερο γουρουνάκι.

"Τότε θα φυσήξω, θα ξεφυσήξω και θα γκρεμίσω το σπίτι σου" μούγκρισε ο λύκος και πήρε μια βαθιά ανάσα, φύσηξε και, σύντομα, δεν έμεινε από το σπίτι παρά μόνο ένα μάτσο ξύλα για το τζά-κι. Τα δύο τρομοκρατημένα γουρουνάκια τρέζανε όσο πιο γρήγορα μπορούσανε μέχρι το τούβλινο σπίτι. Άλλα ο λύκος ήταν συνεχώς ξοπίσω τους και μόλις τα δύο γουρουνάκια μπήκανε στο σπίτι με τον αδελφό τους, αμέσως κατέφθασε κι' ο λύκος στο κατώφλι.

"Γουρουνάκι, γουρουνάκι, άσε με να μπω!" φώναξε ο λύκος.

"Μα ούτε για μία τρίχα από το δέρμα μου!" απήντησε το τρίτο γουρουνάκι.

"Τότε θα φυσήξω, θα ξεφυσήξω και θα γκρεμίσω το σπίτι σου" μούγκρισε ο λύκος και πήρε μια βαθιά ανάσα, φύσηξε και το σπίτι, φυσικά, επειδή ήταν χτισμένο με τούβλα, άφησε τον λύκο χωρίς μπαφιασμένο. Τότε, του ήρθε μία ιδέα: το τζάκι! Σκαρφάλωσε σε μία οξιά και, από κει, ανέβηκε στη σκεπή. Άλλα το τρίτο γουρουνάκι, που είχε μεγάλες οικολογικές εναισθησίες προς το περιβάλλον, αντί για τζάκι είχε εγκαταστήσει φωτοβολταϊκά πάνελ! Τσατισμένος ο λύκος, γλίστρησε κι' έπεσε από τη σκεπή κι' έσπασε το αριστερό του πόδι, καταρρακώνοντας έτσι την αυτοεκτίμησή του. Όταν ο λύκος έφυγε, τα τρία γουρουνάκια ύρχισαν να γελάνε και σκέφτηκαν πόσο πιο ασφαλές είναι να ζεις στην πόλη, όπου μπορείς να βρεις λύκο μονάχα στο ζωολογικό κήπο. Κι' έτσι έπραξαν και, φυσικά, ζήσανε αυτοί καλά κι' εμείς καλύτερα.

## Τι σχέση έχουν όλα αυτά;

---

Η ικανότητα αποθήκευσης/ απομνημόνευσης του υπολογιστή αυξάνεται με απίστευτους ρυθμούς, και στα τελευταία 25 χρόνια, η ποσότητα των δεδομένων που μπορούν να αποθηκευθούν, πολλαπλασιάσθηκε κατά ένα εκατομμύριο φορές!

Παρ' όλα ταύτα, όλο και συνεχίζουμε να βρίσκουμε νέες πληροφορίες προς περαιτέρω αποθήκευση στον υπολογιστή. Οι υπολογιστές μπορούν να περιέχουν ολόκληρα βιβλία ή ολόκληρες βιβλιοθήκες, και πλέον μουσική και ταινίες, υπό την προϋπόθεση να διαθέτουν επαρκή χώρο. Τα μεγάλα αρχεία (files) αποτελούν ένα πρόβλημα στα πλαίσια του διαδικτύου, διότι απαιτούνται μεγάλοι χρόνοι για την μεταφορά τους. Οι υπολογιστές γίνονται επίσης με τον καιρό όλο και πιο μικροί, σε σημείο που ακόμη και ένα κινητό τηλέφωνο ή ένα ρολόϊ χειρός περιέχουν πάμπολες πληροφορίες.

Υπάρχει πάντως μία λύση γι' αυτό το πρόβλημα. Αντί να αγοράζουμε όλο και παραπάνω χωρητικότητα για την αποθήκευση δεδομένων, υπό μορφή π.χ. σκληρών δίσκων ή εξωτερικών σκληρών, ή πάλι, αντί να αγοράζουμε πιο γρήγορα modem και γραμμές με μεγαλύτερη ταχύτητα, μπορούμε να συμπλέσουμε τα δεδομένα. Η διαδικασία συμπίεσης και αποσυμπίεσης των δεδομένων, επιτελείται αυτόματα από τον υπολογιστή. Αυτό που εμείς μπορούμε να διαπιστώσουμε είναι ότι ο δίσκος περιέχει περισσότερα δεδομένα και ότι οι ιστοσελίδες εμφανίζονται πιο γρήγορα, αλλά ο υπολογιστής πρέπει να κάνει πιο πολλούς υπολογισμούς και επεξεργασία.

Έχουν επινοηθεί πολλοί τρόποι συμπίεσης για τα κείμενα. Η μέθοδος που θα χρησιμοποιηθεί σε αυτή τη δραστηριότητα, με την αρχή της παραπομπής σε προηγούμενη σύμπτωση επαναλαμβανόμενων σειρών μέσα στο κείμενο, είναι γνωστή με το όνομα “κωδικοποίηση κατά Ziv-Lempel” ή, επί το απλούστερο, σαν κωδικοποίηση “LZ”, διότι επινόθηκε από δύο Ισραηλινούς καθηγητές (τον καθηγ. Ziv και τον καθηγ. Lempel) κατά τη δεκαετία του '70. Μπορεί να χρησιμοποιηθεί σε οποιαδήποτε γλώσσα και μπορεί εύκολα να περικόψει κατά το ήμισυ το μέγεθος των δεδομένων προς συμπίεση. Μερικές φορές αναφέρεται και ως “zip” στους προσωπικούς υπολογιστές και χρησιμοποιείται επίσης στις εικόνες GIF και στα modem υψηλής ταχύτητος. Στη περίπτωση των modem, ελαττώνει την ποσότητα δεδομένων που πρέπει να ταξιδεύσουν στην τηλεφωνική γραμμή και, άρα, η σύνδεση γίνεται πιο γρήγορη.

Άλλες μέθοδοι βασίζονται στην ιδέα ότι κάποια γράμματα είναι πιο συχνά από άλλα και συνεπώς, θα πρέπει να έχουν βραχύτερους κώδικες σε σχέση με τα άλλα. Τα σήματα Morse χρησιμοποιούν αυτή την ιδέα.

## Λύσεις και υποδείξεις

---

**Μπορείς να το ξαναπείς! (σελ. 26)**

Γυαλό – γυαλό, πηγαίναμε,

κι' όλο για σένα λέγαμε,

Γυαλό να πας, γυαλό να 'ρθείς,

τα λόγια μου να θυμηθείς.