

Attività 3

Puoi dirlo nuovamente! — *La compressione del testo*

Sommario

Dato che i computer hanno uno spazio limitato per poter immagazzinare le informazioni, devono rappresentarle nel modo più efficiente possibile. Si usano quindi metodi di compressione dell'informazione. Comprimendo i dati per memorizzarli e decomprimendoli per rileggerli, si possono conservare più informazioni o spedirle più velocemente attraverso Internet.

Competenze richieste

- ✓ Copiare testo scritto

Età

- ✓ A partire da 9 anni

Materiale

- ✓ Immagini da proiettore ottenute fotocopiando il prototipo "puoi dirlo nuovamente!" (pagina 37)

Ogni studente deve avere:

- ✓ Il foglio di lavoro: puoi dirlo nuovamente! (pagina 38)
- ✓ Il foglio di lavoro: extra per esperti (pagina 40)
- ✓ Il foglio di lavoro: Apelle, figlio di Apollo (pagina 41)
- ✓ Il foglio di lavoro: extra per veri esperti (pagina 42)

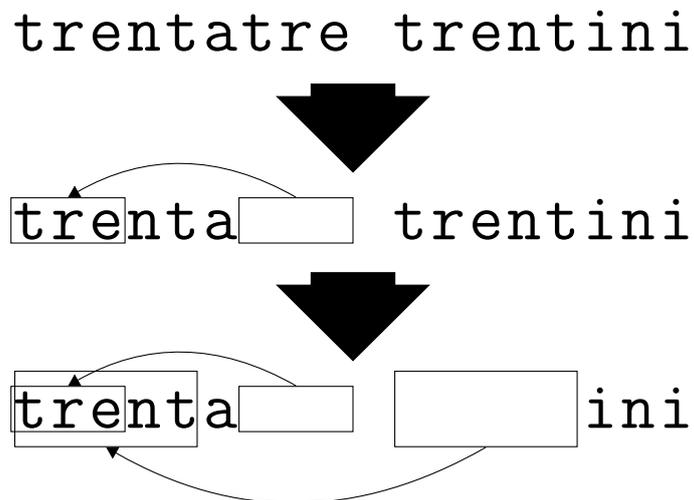
Puoi dirlo nuovamente!

Introduzione

I computer devono memorizzare e trasmettere tantissimi dati. Quindi per ottimizzare il lavoro, devono evitare di usare troppo spazio o impiegare troppo tempo per trasmetterli lungo le reti. Questa attività mostra come si possono comprimere i testi.

Dimostrazione e discussione.

Mostrate lo scioglilingua di pagina 37. Osservate quali sequenze di lettere sono ripetute nel testo. Potete trovare i gruppi di due o più lettere sono ripetute o anche intere parole o gruppi di parole. Sostituite le sequenze ripetute tracciando loro un rettangolo attorno come nel diagramma che segue.



**Immagine da proiettare (o fotocopiare):
puoi dirlo nuovamente!**

trentatre trentini

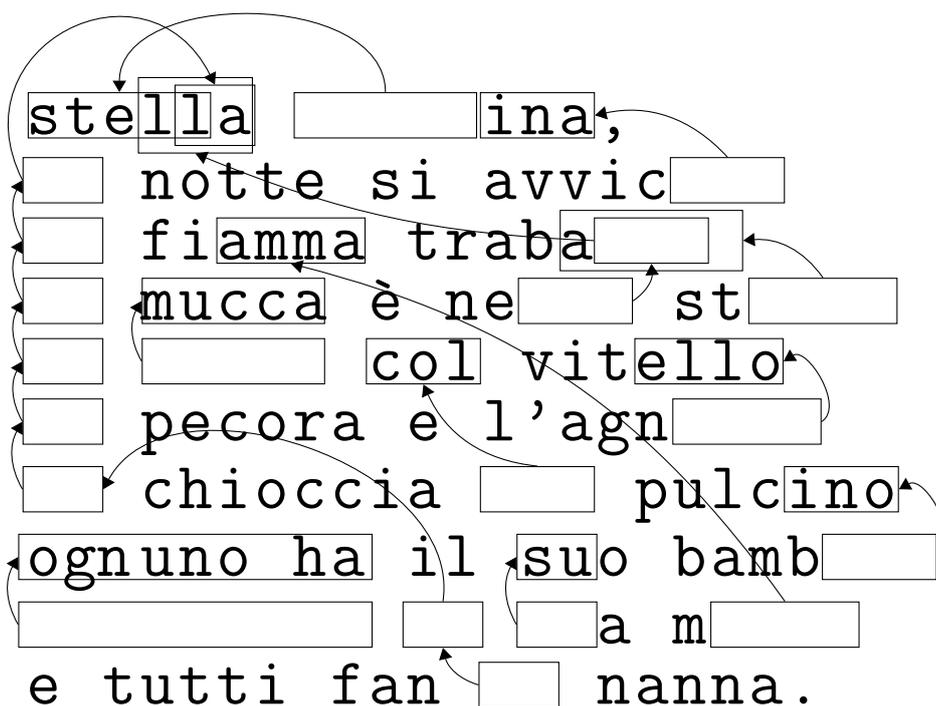
entrarono a Trento

tutti e trentatre

trotterellando

Foglio di Lavoro: puoi dirlo nuovamen- te!

Molte delle lettere e delle parole della seguente poesia sono mancanti. Potete inserire le lettere e le parole mancanti per completare correttamente il testo? Troverete le lettere e le parole mancanti nel riquadro puntato dalla freccia:



Ora potete scegliere altre semplici poesie o ninna nanne e creare un puzzle simile. Controllate che le frecce puntino sempre ad una parte precedente del testo. La vostra poesia deve sempre essere ricomposta scorrendo il testo da sinistra a destra, dall'alto in basso, nell'ordine solito della lettura.

Sfida: Provate a vedere quante poche parole e lettere della poesia originale occorre tenere!

Altri suggerimenti: usate degli scioglilingua (sopra la panca..., se il coniglio gli agli piglia..., ...) oppure poesie o parole di canzoni con ripetizioni.

Suggerimento: Tentate di evitare un sovraffollamento di frecce. Lasciate molto spazio fra le lettere e fra le parole così che ci sia lo spazio sufficiente per disegnare i riquadri e le frecce che puntano ad esse.

È più facile disegnare il puzzle se prima scrivete la poesia e poi decidete quali debbano essere i riquadri da tracciare.

Foglio di Lavoro: extra per esperti

Come risolvereste questo puzzle?

Talvolta le parti mancanti del testo puntano a una porzione di se stesse. In questo caso occorre copiare lettera per lettera da sinistra a destra per poter ritrovare il testo originario. In questo modo ogni lettera è già disponibile quando è necessaria per proseguire. Questo metodo è utile quando i computer si trovano lunghe sequenze di un particolare carattere o di un gruppo di caratteri.

Potete provare a creare altri esempi per conto vostro.

I computer rappresentano i riquadri e le frecce con numeri, per esempio:

Banana

può essere scritto come **Ban(2,3)**. "2" significa che occorre tornare indietro di due caratteri per iniziare la copia.

Ban___

e "3" significa che occorre copiare tre caratteri consecutivi.

Bana__

Banan_

Banana



(cosa stai facendo?)

Sto comprimendo le banane)

Siccome vengono utilizzati due numeri per indicare che ci sono caratteri ripetuti, solo gruppi di due o più lettere vengono compressi, altrimenti non c'è alcun risparmio di spazio. Infatti la lunghezza di un testo si allungherebbe se utilizzassimo due numeri al posto di una lettera.

Scrivete alcune parole nel modo usato da un computer per comprimerle, cercate le parole più interessanti per la compressione. Riescono i vostri amici a ricostruire le parole che avete compresso?

Foglio di lavoro: Apelle, figlio di Apollo

Quante parole sono veramente necessarie qui?

Pensate di essere un computer che voglia mettere sul proprio disco più informazioni possibili. Scorrete il testo e trovate tutte le sequenze di due o più lettere che sono già comparse nel testo. Queste sequenze non sono più utili e possono essere sostituite con un puntatore (un riquadro vuoto e una freccia che indica un altro riquadro con le lettere da copiare). Il vostro scopo è di trovare più sequenze che potete.

Apelle figlio di Apollo

Fece una palla di pelle di pollo

Tutti i pesci vennero a galla

Per vedere la palla di pelle di pollo

Fatta da Apelle figlio di Apollo

Foglio di lavoro: extra per veri esperti

Siete pronti per una compressione veramente dura?

La favola che segue è stata analizzata da un programma per computer (uno di quelli veri con la spina) che ha trovato che ci sono almeno 499 sequenze e 2281 lettere che possono essere eliminate perché appartengono a sequenze ripetute (la divisione fra righe e paragrafi viene considerata come uno spazio). Quante sequenze riuscite a trovare? Quanti caratteri in totale potete eliminare? Ricordatevi che solo le sequenze di due o più caratteri ripetuti possono essere eliminate. Buona Fortuna!

C'erano una volta, tanto tempo fa, tre porcellini che volevano andare a vivere da soli. Il primo porcellino non era molto furbo e decise di costruire la sua casa con la paglia perché costava poco. Anche il secondo porcellino non era furbo e costruì la sua casa di legno perché aveva un look ecologico ed era di moda. Il terzo porcellino era più furbo dei suoi due fratelli e comperò un carico di mattoni in una città vicina coi quali costruì una casa modesta ma confortevole.

Poco tempo dopo la festa di inaugurazione della casa, il primo porcellino era assorto su una sedia leggendo un libro quando udì bussare alla porta. Era il lupo cattivo, naturalmente.

"Porcellino, porcellino, lasciami entrare!" gridò il lupo.

"Neanche per un pelo della mia cotenna!" rispose il primo porcellino.

"Allora sbufferò, soffierò e abatterò la tua casa" ruggì il lupo e sbuffò, soffiò e la casa in breve crollò. Il primo porcellino corse più forte che poté verso la casa di legno e presto fu al sicuro lì col fratello. Ma non passò molto che il lupo arrivò anche lì.

"Porcellino, porcellino, lasciami entrare!" gridò il lupo.

"Neanche per un pelo della mia cotenna!" rispose il secondo porcellino. "Allora sbufferò, soffierò e abatterò la tua casa" ruggì il lupo e sbuffò, soffiò e presto della casa non rimase che un mucchio di legna per il camino. Il due porcellini terrorizzati corsero più forte che poterono verso la casa di mattoni, ma il lupo era sempre alle loro calcagna e appena i due porcellini arrivarono alla casa del loro fratello, anche il lupo arrivò ancora una volta sulla soglia.

"Porcellino, porcellino, lasciami entrare!" gridò il lupo.

"Neanche per un pelo della mia cotenna!" rispose il terzo porcellino.

"Allora sbufferò, soffierò e abatterò la tua casa" ruggì il lupo e sbuffò, soffiò e, naturalmente, la casa era fatta di mattoni e il lupo rimase senza fiato. Allora ebbe un'idea: il camino! Si arrampicò su una quercia e salì sul tetto. Ma il terzo porcellino che era molto sensibile ai temi ambientali invece del camino aveva installato pannelli solari. Frustrato il lupo scivolò e cadde dal tetto rompendosi la zampa sinistra, danneggiando irreparabilmente la propria autostima. Quando il lupo se ne andò i porcellini risero e pensarono a quanto sicuro sia vivere in città dove i lupi si possono trovare solo allo zoo. E così fecero e, naturalmente, vissero felici e contenti.

Cosa c'entra tutto questo?

La capacità di memorizzazione dei computer sta crescendo ad un ritmo incredibile negli ultimi 25 anni la quantità di dati immagazzinabili è cresciuta di circa un milione di volte, ma continuiamo a trovare informazioni da mettere nei computer. I computer possono contenere interi libri o intere biblioteche e ora musica e film a patto che abbiano spazio a sufficienza. I grandi archivi (file) rappresentano un problema sulla rete Internet perché hanno necessità di lunghi tempi per poter essere trasferiti. I computer inoltre stanno diventando sempre più piccoli: anche un telefono cellulare o un orologio da polso contengono moltissime informazioni.

C'è comunque una soluzione per risolvere questo problema. Invece di acquistare sempre più spazio per immagazzinare dati, per esempio dischi fissi o removibili, oppure invece di acquistare modem più veloci e comperare linee a velocità più elevata, possiamo comprimere i dati. Il processo di compressione e di decompressione dei dati è effettuato automaticamente dal computer. Tutto quello che possiamo notare è che il disco contiene più dati e che le pagine web vengono trasferite più velocemente. Quello che cambia è che il computer deve fare più calcoli.

Sono stati inventati molti metodi di compressione per i testi. Il metodo usato in questa attività, sfruttando il principio di riferirsi alla precedente occorrenza di sequenze ripetute, è noto con il nome di "codifica di Ziv-Lempel" o più semplicemente come codifica "ZL" perché è stata inventata da due professori israeliani (il prof. Ziv e il prof. Lempel), negli anni '70. È qualche volta indicata anche come "zip" nei personal computer e viene anche usata nelle immagini GIF e nei modem ad alta velocità. Nel caso dei modem, riduce la quantità di dati che devono essere scambiati sulla linea telefonica e quindi il collegamento risulta più veloce.

Alcuni altri metodi sono basati sull'idea che alcune lettere sono più frequenti di altre e quindi vengono associati codici più corti alle lettere più frequenti. Il codice Morse usa questo principio.

Soluzioni e suggerimenti

Puoi dirlo nuovamente! (pagina 38)

stella stellina,
la notte si avvicina
la fiamma traballa
la mucca è nella stalla
la mucca col vitello
la pecora e l'agnello
la chioccia col pulcino
ognuno ha il suo bambino
ognuno ha la sua mamma
e tutti fan la nanna.