

Atividade 3

Você pode dizer isso de novo! –*Compressão de Texto*

Sumário

Desde que os computadores possuíam apenas uma quantidade limitada de espaço para armazenar informações, estes precisavam armazená-las da forma mais eficiente possível. Isso é chamado compressão. Codificando os dados antes de armazená-los e os decodificando assim que estiverem sendo recuperados, permite ao computador armazenar mais dados ou enviá-los mais rapidamente através da Internet.

Correlações curriculares

- ✓ Português: Reconhecimento de padrões em palavras e texto.
- ✓ Tecnologia: Conhecimento sobre o funcionamento dos computadores.

Habilidades

- ✓ Cópia de textos escritos

Idades

- ✓ A partir de 9 anos

Material

- ✓ Transparência feita a partir da Transparência-mestre: Você pode dizer isso de novo! (página 25)

Cada criança precisará de:

- ✓ Planilha de Atividade: Você pode dizer isso de novo! (página 26)
- ✓ Planilha de Atividade: Extra para especialistas (página 27)
- ✓ Planilha de Atividade: Curto e Grosso (página 28)
- ✓ Planilha de Atividade: Extra para verdadeiros especialistas (página 29)

Você pode dizer isso de novo!

Introdução

Os computadores devem conservar e transmitir diversos dados. Para não usarem muito espaço de armazenamento, ou levar muito tempo para enviar informações através de uma conexão via modem, eles comprimem o texto mais ou menos dessa forma.

Demonstração e discussão

Mostre o poema "A Aranha e a Jarra" (Transparência-mestre da página 25). Identifique padrões de letras neste poema. Você consegue encontrar grupos de 2 ou mais letras repetidas, ou mesmo palavras inteiras ou frases? (Substitua esses padrões por caixas em branco conforme mostrado no diagrama abaixo.)

Aranha arranha



Aranha ar

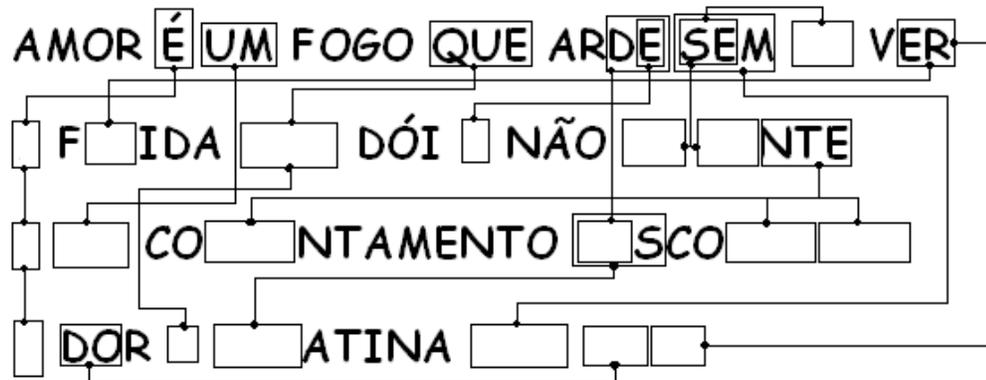
Transparência-Mestre: Você pode dizer isso de novo!

A Aranha e a Jarra

Debaixo da cama tem uma jarra,
Dentro da jarra tem uma aranha.
Tanto a aranha arranha a jarra,
Como a jarra arranha a aranha.

Planilha de Atividade: Você pode dizer isso de novo!

Muitas palavras e letras estão faltando neste lindo poema do ilustre poeta português Luís Vaz de Camões. Você pode preencher as letras e palavras que faltam para completá-lo corretamente? Você as encontrará nos retângulos apontados pela seta.



Agora escolha um poema ou versinho infantil e construa seu próprio enigma. Certifique-se de que as setas apontem sempre para uma parte anterior do texto. O seu poema deve ser capaz de ser decodificado da esquerda para a direita e de cima para baixo, da mesma forma que lemos um texto.

Desafio: Veja quão poucas palavras do texto original você precisa conservar!

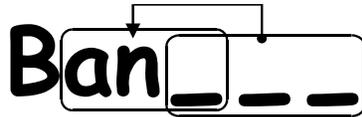
Aqui vão algumas sugestões: "A Casa" e "O Pato" de Vinícius de Moraes.

Dica: Tente evitar o uso demasiado de setas. Deixe bastante espaço livre em volta das letras e palavras ao escrever de modo que você tenha espaço para os retângulos e as setas que apontam para estes.

É mais fácil construir o enigma se, primeiro, você escrever o poema por extenso e depois decidir onde ficarão os retângulos.

Planilha de Atividade: Extra para Especialistas

Como resolver este enigma ?



Às vezes, textos incompletos apontam para parte dele mesmo. Neste caso, o texto pode ser decodificado corretamente se as letras forem copiadas da esquerda para a direita. Assim, cada letra está disponível para ser copiada antes de ser necessária. Isso é útil em computadores nos quais existe uma cadeia grande de caracteres ou padrões específicos.

Tente desenhar alguns de seus próprios padrões.

Em computadores as caixas e flechas são representadas por números. Por exemplo,

Banana

pode ser escrita como **Ban (2,3)**. "2" significa contar dois caracteres para trás para encontrar o ponto de partida para a cópia,

Ban...

e "3" significa copiar três caracteres consecutivos:

Bana..

Banan..

Banana



Elabore algumas palavras e as escreva na forma que um computador faria se estas fossem comprimidas. Seus amigos conseguem decodificá-las ?

Planilha de Atividade: Curto e Grosso

Quantas palavras você realmente precisa ?

Finja que você é um computador tentando colocar o máximo de informações em seu disco. Descarte todos os grupos de duas ou mais letras que já tenham ocorrido. Estes deixaram de ser necessários, uma vez que poderiam ser substituídos por um ponteiro. Seu objetivo é conseguir marcar a maior quantidade de letras possíveis.

Banana (Nelma Sampaio)

Banana, bananeira, bananada

A macacada se alvoroça quando vê

E sai correndo dando pulos, dando saltos

Pra ver quem chega lá no topo pra comer

Na bananeira tem um cacho bem maduro

E Vovó colhe as bananas pra fazer

A bananada que é um doce bem escuro

Em quantidade que é pra dar e pra vender.

Planilha de Atividade: Extra para Verdadeiros Especialistas

Prontos para uma compressão realmente difícil ?

Quantas letras você consegue marcar para compressão na história dos Três Porquinhos ? Lembrem-se: apenas grupos de dois ou mais caracteres repetidos podem ser eliminados. Boa sorte!

Era uma vez, três porquinhos que viviam com a sua mãe. Um dia ela disse-lhes: - Queridos filhos, penso que já são grandinhos para viverem sozinhos e poderem trabalhar. Têm braços fortes e não lhes falta inteligência para pensar o que é melhor para cada um. Primeiro têm que construir suas próprias casas perto uns dos outros e se prepararem para todos os perigos que possam surgir.

Os porquinhos puseram-se a caminho. De todos os três o pequeno era o mais trabalhador, o do meio era trapalhão e o maior era preguiçoso. Como o maior era preguiçoso, fez a casa de palha para ser mais rápido. O do meio fez a casa de madeira, pois também não gostava muito de trabalhar. Mas, o mais novo, o mais trabalhador, fez a sua com cimento e tijolos. Um dia, apareceu o lobo que com um sopro derrubou a casa do mais velho e com outro sopro deitou abaixo a casa do porquinho do meio. Os porquinhos muito amedrontados correram para casa do irmão mais novo com o lobo correndo atrás deles. O porquinho abriu-lhes a porta rapidamente e, os irmãos, entraram muito admirados por verem uma casa tão forte e tão bonita. O lobo pensava que a derrubava, soprava e dizia:

- Soprarei, soprarei e a sua casa derrubarei!

Mas a casa era forte e ele não a conseguiu derrubar. Por isso, muito envergonhado fugiu e não voltou mais. Os três porquinhos continuaram morando na casa de tijolos e viveram felizes para sempre.

O que é tudo isso afinal ?

A capacidade de armazenamento dos computadores cresce num ritmo incrível— nos últimos 25 anos, esta quantidade de armazenamento em um computador típico cresceu cerca de um milhão de vezes — mas ainda encontramos mais coisas para colocarmos em nossos computadores. Os computadores podem armazenar livros ou até mesmo bibliotecas e, agora, música e filmes também, se houver espaço. Arquivos grandes também são um problema na Internet porque demoram muito tempo para serem baixados. Além disso, tentamos reduzir os computadores cada vez mais — até mesmo um relógio de pulso ou celular pode armazenar grandes quantidades de informação !

Entretanto, existe uma solução para este problema. Ao invés de comprar mais espaço de armazenamento ou um modem mais rápido, podemos *comprimir* os dados para que estes ocupem menos espaço. Este processo de compressão e descompressão de dados é normalmente feito automaticamente pelo computador. O que percebemos é que o disco tem mais espaço disponível, ou que as páginas da web são exibidas mais rapidamente, porém o computador está, na verdade, realizando mais processamento.

Vários métodos de compressão foram inventados. O método utilizado nesta atividade, baseado no princípio de apontar para ocorrências anteriores de blocos de texto, é muitas vezes referido como “codificação Ziv-Lempel” ou “codificação LZ”, e foi inventado por dois professores israelenses na década de 1970. Este método pode ser usado para qualquer idioma e pode facilmente reduzir à metade o tamanho dos dados a serem compactados. É, por vezes, referido como “zip” em computadores pessoais, é também utilizado em imagens do tipo “GIF”, bem como em modems de alta velocidade. No caso dos modems, este método reduz a quantidade de dados que precisam ser transmitidos através da linha telefônica, sendo assim bem mais rápido.

Alguns outros métodos são baseados na idéia de que as letras usadas mais freqüentemente deveriam ter códigos menores do que as outras. O Código Morse utilizou esta idéia.

Soluções e dicas

Você pode dizer isso de novo! (página 26)

**Amor é fogo que arde sem doer,
É ferida que dói e não se sente,
É contentamento descontente,
É dor que desatina sem doer.**

Luís Vaz de Camões