

## فعالیت ۳

### دوباره بگو! فشرده‌سازی متن

#### خلاصه

کامپیوترها فضای محدودی برای نگهداری اطلاعات دارند، بنابراین باید اطلاعات را به گونه‌ای ارائه کنند که کمترین فضا را اشغال کند. به این کار فشرده‌سازی گفته می‌شود. در فشرده‌سازی داده‌ها قبل از ذخیره رمزگذاری می‌شوند یعنی داده‌های کوتاه‌تری جایگزین آن‌ها می‌شوند، سپس هنگام بازیابی داده‌های اصلی رمزگشایی می‌شوند. کامپیوتر با کمک این فرایند می‌تواند داده‌های بیشتری ذخیره کند یا ارسال آن‌ها از طریق اینترنت را سریع‌تر انجام دهد.

#### پیوندها با برنامه درسی

- ✓ زبان انگلیسی: تشخیص الگوهای موجود در کلمه‌ها و متن‌ها
- ✓ تکنولوژی: کاهش فضای اشغال شده توسط داده‌های تکراری

#### مهارت‌ها

- ✓ کپی برداری از متن‌های نوشته شده

#### رده سنی

- ✓ ۹ سال به بالا

#### مواد لازم

- ✓ اسلاید برای نمایش: دوباره بگو! (صفحه ۲۸)

#### برای هر دانش‌آموز:

- ✓ کاربرگ ۱: دوباره بگو! (صفحه ۲۹)
- ✓ کاربرگ ۲: تمرین بیشتر برای حرفه‌ای‌ها (صفحه ۳۰)
- ✓ کاربرگ ۳: کوتاه و شیرین (صفحه ۳۱)
- ✓ کاربرگ ۴: تمرین بیشتر برای فوق حرفه‌ای‌ها (صفحه ۳۳)

کامپیوترها داده‌های بسیار زیادی را ذخیره و تبادل می‌کنند، پس مجبورند از کمترین فضای ذخیره‌سازی استفاده کنند یا برای انتقال اطلاعات در یک اتصال شبکه‌ای کمترین زمان را صرف کنند. بنابراین متن‌ها را به شیوه‌ای تقریباً مشابه شیوه ریز فشرده می‌کنند.

نمونه زیر را اجرا و درباره آن بحث کنید

اسلاید شعر "The Rain" (صفحه ۲۸) را نمایش دهید. به دنبال الگوهای حروف در این شعر بگردید. می‌توانید گروه‌های دو یا چند حرفی، یا یک کلمه یا متن را پیدا کنید که تکرار شده‌اند؟ (کادر سفیدی را مانند تصویر زیر بر روی آن‌ها قرار دهید.)

Pitter patter

↓

Pitter pa

# The Rain

---

**Pitter patter**

**Pitter patter**

**Listen to the rain**

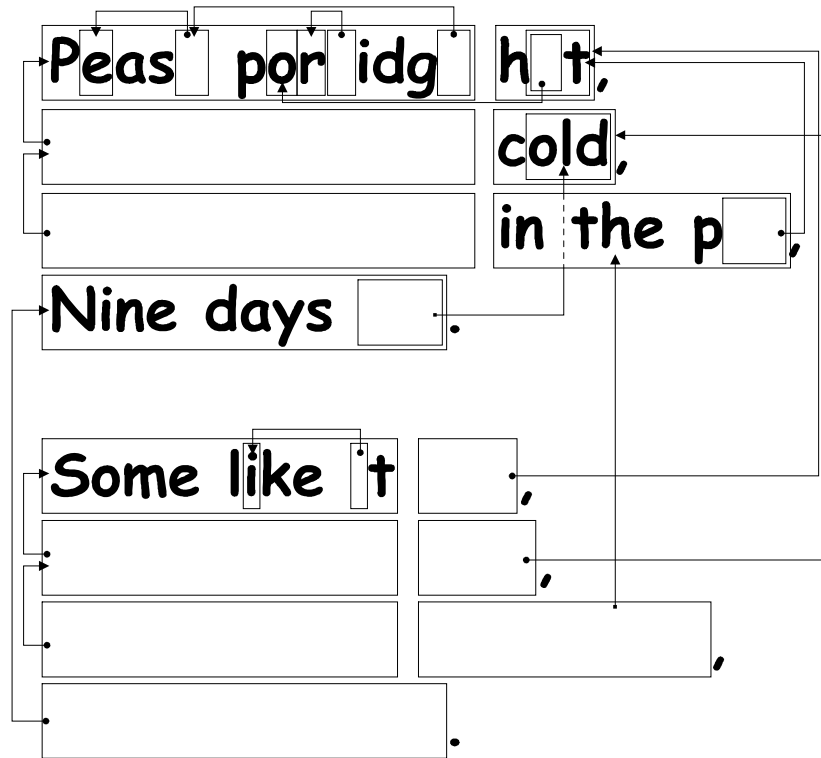
**Pitter patter**

**Pitter patter**

**On the window pane**

## کاربرگ ۱: دوباره بگو!

در شعر زیر بسیاری از کلمه ها و حروف حذف شده اند. آیا می‌توانید جای کلمه‌ها و حرف‌های گم شده را پر کنید و شعر را کامل کنید؟ هر قسمت حذف شده را در کادری که با فلش نشان داده شده پیدا کنید.



حالا یک شعر ساده یا یک لالایی کودکانه را انتخاب کنید و خودتان یک پازل طراحی کنید. مطمئن شوید که فلش‌هایتان همواره به قسمت قبلی متن اشاره کنند. شعر شما باید بتواند به همان طریقی که خوانده می‌شود از راست به چپ و از بالا به پایین رمزگشایی شود.

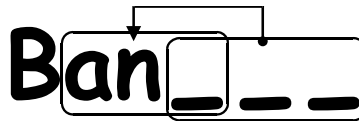
**چالش:** توجه کنید که چه تعداد کمی از کلمات شعر اصلی را لازم است نگه دارید!

نکته: سعی کنید پازلتان زیادی با فلش‌ها شلوغ نشود. هنگام نوشتن در اطراف حروف و کلمه‌ها جای زیادی بگذارید تا برای کادرهایی که داخل کادره‌های دیگر قرار می‌گیرند و فلش‌هایی که به آنها اشاره می‌کنند فضای کافی وجود داشته باشد.

اگر ابتدا شعر را بنویسید و سپس برای جای کادرها تصمیم بگیرید طراحی پازل برایتان راحت‌تر خواهد شد.

## کاربرگ ۲: تمرین بیشتر برای حرفه‌ای‌ها

شما این پازل را چگونه حل می‌کنید؟



گاهی متن حذف شده به قسمتی از خودش اشاره می‌کند. در این صورت اگر حروف از چپ به راست کپی شوند، متن به درستی رمزگشایی می‌شود. بنابراین هر حرف قبل از آن که نیاز باشد، برای کپی کردن در دسترس است. این شیوه در کامپیوترها هنگامی که یک توالی طولانی از یک کاراکتر یا الگوی خاص وجود داشته باشد، استفاده می‌شود.

سعی کنید تعدادی از متن‌های خودتان را طراحی کنید.

در کامپیوترها، کادرها و فلش‌ها با اعداد نشان داده می‌شوند. برای مثال Banana را می‌توان به صورت Ban(2,3) نوشت. "2" یعنی دو کاراکتر به عقب بروید تا به نقطه شروع برای کپی کردن برسید:

**Ban...**

و "3" یعنی سه کاراکتر متوالی را کپی کنید:



**Bana..**

**Banan.**

**Banana**

از آنجا که دو عدد برای رمزگذاری این کلمه‌ها استفاده می‌شود، معمولاً تنها گروه‌های دو تایی یا بیشتر از حروف ارزش فشرده‌سازی دارند، در غیر این صورت صرفه‌جویی در فضای ذخیره‌سازی اتفاق نمی‌افتد. در واقع اگر از دو عدد برای رمزگذاری تنها یک حرف استفاده کنیم، اندازه فایل بیشتر هم می‌شود.

چند کلمه انگلیسی را انتخاب کنید و به روش کامپیوتری فشرده‌سازی کنید. ببینید آیا دوستان تان می‌توانند آن‌ها را رمزگشایی کنند؟

## کاربرگ ۳: کوتاه و شیرین

وانمود کنید کامپیوتر هستید و سعی دارید تا جای ممکن در دیسک خود اطلاعات جای دهید. در متن صفحه بعد گروه‌های دوتایی و بیشتر از حروف که تکرار شده‌اند را پیدا و حذف کنید. چون دیگر نیازی به آنها نیست و می‌توان با یک فلش آن‌ها را جایگذاری کرد. هدف شما این است که تا جای ممکن حروف حذف شده داشته باشید.

در نهایت به چند کلمه نیاز دارید؟

I know an old lady who swallowed a bird  
How absurd! She swallowed a bird!  
She swallowed the bird to catch the spider  
That wriggled and jiggled  
and tickled inside her  
She swallowed the spider to catch the fly  
I don't know why she swallowed a fly  
Perhaps she'll die...

## کاربرگ ۴: تمرین بیشتر برای فوق حرفه‌ای‌ها

برای یک فشرده‌سازی درست و حسابی آماده‌اید؟

داستان زیر توسط یک برنامه کامپیوتری اجرا شد، و مشخص شد که دست کم ۱۶۳۳ حرف برای حذف کردن وجود دارد. شما چند حرف را برای حذف کردن می‌توانید پیدا کنید؟ به یاد داشته باشید که تنها حروف تکراری در گروه‌های دوتایی و بیشتر را می‌توان حذف کرد. موفق باشید!

Once upon a time, long, long ago, three little pigs set out to make their fortunes. The first little pig wasn't very clever, and decided to build his house out of straw, because it was cheap. The second little pig wasn't very clever either, and decided to build his house out of sticks, for the "natural" look that was so very much in fashion, even in those days. The third little pig was much smarter than his two brothers, and bought a load of bricks in a nearby town, with which to construct a sturdy but comfortable country home.

Not long after his housewarming party, the first little pig was curled up in a chair reading a book, when there came a knock at the door. It was the big bad wolf, naturally.

"Little pig, little pig, let me come in!" cried the wolf.

"Not by the hair on my chinny-chin-chin!" squealed the first little pig.

"Then I'll huff, and I'll puff, and I'll blow your house down!" roared the wolf, and he *did* huff, and he *did* puff, and the house soon collapsed. The first little pig ran as fast as he could to the house of sticks, and was soon safe inside. But it wasn't long before the wolf came calling again.

"Little pig, little pig, let me come in!" cried the wolf.

"Not by the hair on my chinny-chin-chin!" squealed the second little pig.

"Then I'll huff, and I'll puff, and I'll blow your house down!" roared the wolf, and he *did* huff, and he *did* puff, and the house was soon so much firewood. The two terrified little pigs ran all the way to their brother's brick house, but the wolf was hot on their heels, and soon he was on the doorstep.

"Little pig, little pig, let me come in!" cried the wolf.

"Not by the hair on my chinny-chin-chin!" squealed the third little pig.

"Then I'll huff, and I'll puff, and I'll blow your house down!" roared the wolf, and he huffed, and he puffed, and he huffed some more, but of course, the house was built of brick, and the wolf was soon out of breath. Then he had an idea. The chimney! He clambered up a handy oak tree onto the roof, only to find that there *was* no chimney, because the third little pig, being conscious of the environment, had installed electric heating. In his frustration, the wolf slipped and fell off the roof, breaking his left leg, and severely injuring his pride. As he limped away, the pigs laughed, and remarked how much more sensible it was to live in the city, where the only wolves were in the zoo. And so that is what they did, and of course they all lived happily ever after.



ظرفیت ذخیره سازی کامپیوترها با سرعتی باورنکردنی رو به افزایش است- در ۲۵ سال گذشته، میزان ذخیره سازی یک کامپیوتر معمولی یک میلیون برابر افزایش پیدا کرده است، اما ما هنوز هم داده های بیشتری برای جا دادن در کامپیوترهایمان پیدا کنیم. کامپیوترها در صورتی که فضا داشته باشند می توانند همه کتابها یا حتی کتابخانه ها و موسیقی ها و فیلمها را نیز در خود ذخیره کنند. فایل های بزرگ در اینترنت هم مشکل محسوب می شوند چرا که دانلود آنها زمان زیادی می برد. از طرفی، ما هم همیشه تلاش کرده ایم کامپیوترها را کوچکتر کنیم- حالا دیگر حتی از یک تلفن همراه یا ساعت مچی هم انتظار می رود که اطلاعات زیادی در خود ذخیره کنند!

البته برای این مشکل یک راه حل وجود دارد. به جای خریدن فضای ذخیره سازی بیشتر یا افزایش دادن سرعت اینترنت، می توانیم داده ها را فشرده کنیم تا فضای کمتری اشغال کنند. فرآیند فشرده سازی و به حالت عادی برگرداندن داده ها معمولاً به صورت خودکار به وسیله کامپیوترها انجام می شود. تمام چیزی که ما ممکن است از این فرایند متوجه شویم این است که دیسکها اطلاعات بیشتری ذخیره می کنند و صفحه های وب سریع تر به نمایش در می آیند، اما کامپیوتر در واقع پردازش بیشتری انجام می دهد.

روش های فشرده سازی متعددی تا کنون ابداع شده اند. روش مورد استفاده در این فعالیت، یعنی اشاره به قسمتی از متن که قبل تر ذکر شده است، اغلب «رمزگذاری زیو-لمپل» (Ziv-Lempel coding) یا «رمزگذاری LZ» نامیده می شود که به وسیله دو پروفیسور اسرائیلی در دهه ۱۹۷۰ ابداع شده است. این روش برای هر زبانی قابل استفاده است و می تواند به راحتی اندازه داده ها را با فشرده سازی به نصف برساند. در کامپیوترهای شخصی، این برنامه "zip" هم خوانده می شود و برای تصویرهای "GIF" و "PNG" هم کاربرد دارند و در مودم های سرعت بالا نیز استفاده می شوند. این روش در مودمها اندازه داده هایی که باید از خط تلفن منتقل شوند را کاهش می دهد، در نتیجه داده ها سریع تر مبادله می شوند.

بعضی از روش های دیگر فشرده سازی بر پایه این ایده توسعه یافته اند که حروفی که بیشتر استفاده مورد استفاده قرار می گیرند باید کدهای کوتاهتری نسبت به بقیه حروف داشته باشند. کد مورس همین ایده را به کار گرفته است.

## پاسخ ها و نکته ها

دوباره بگو! (صفحه ۲۹)

Pease porridge hot,

Pease porridge cold,

Pease porridge in the pot,

Nine days old.

Some like it hot,

Some like it cold,

Some like it in the pot,

Nine days old.