

Name(s) _____ Period _____ Date _____

Activity Guide - Text Compression



Objectives

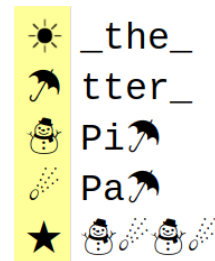
- Compress a piece of text using the Text Compression Widget
- Explain the factors that make compression challenging.
- Explain why the “best” compression is impossible or “hard” to identify.
- Create your own heuristic for compressing data.

Text Compression Tool

If you have not already done so, open up the Text Compression Tool in Code Studio and watch the video explaining how to use it. Then choose one poem and build a simple dictionary using the table below.

Tips:

- Look for patterns that repeat and enter each in the “dictionary”
- Look for patterns of patterns - the dictionary can refer to itself (see right) making for a powerful amount of compression!
- Compare with others compressing the same text
- Try to develop a general strategy that will lead to “good” compression,



What’s your best?

Copy and paste the best compression you made. (If you have a digital copy of this activity guide you could copy/paste a screenshot here, or just copy paste the text from the tool itself.)

Poem Name: _____

Compressed Text:

Dictionary:

Compression Stats:

Reflection

Respond to these prompts.

1. What made compressing text hard to do?
2. Describe the thinking process you used in solving this challenge - what was your strategy for compressing the text. Could you explain it to someone who had never done this before?

Develop a Heuristic

Continue working on compressing your poem. As you do so, develop a set of rules, or a “heuristic” that generally seems to provide good results. Record the steps of your heuristic in some way that will allow you to exchange with another group. Make sure your rules are as clear as possible, so the other group will always know what to do.

Exchange Heuristics

Trade your heuristics with another group. Are they clear and specific enough that you always know what to do? If not, provide feedback to one another and improve your heuristics to provide clearer instructions.

Test Heuristics

You should now have another group’s heuristic. Using the heuristic, attempt to compress the poems below. Record the compression rates you achieve.

Poem	Compression Rate
I Need a Dollar	
Pitter Patter	
A Tutor	
She Sells Sea Shells	
I Know an Old Lady	
Pease Porridge	

Record Your Conclusions

Use the data you collected to respond to the questions below.

1. Do you think it’s possible to describe (or write) a specific set of instructions that a person could follow that would always result in better text compression than your heuristic? Why or why not?
2. Is there a way to know that a compressed piece of text is compressed the most possible? If yes, describe how you could determine it. If no, why not?
3. If you send the compressed poem would your friend be able to read it? Why is the dictionary important?

Vocabulary

Compress: to decrease the number of bits used to represent a piece of information

Algorithm: a precise sequence of instructions designed to complete a task

Heuristic: a specific type of algorithm, usually used when exact solutions are difficult or impossible. Heuristics are generally simple to use and are designed to provide reasonably good results without guaranteeing a perfect solution.