Inspiring Greatness

ISD

INDEPENDENCE SCHOOL DISTRICT

Computer Science Virtual Learning

PLTW Computer Science Principles

May 13, 2020

Lesson: May 13, 2020
Visualizing Data

Learning Target:

The goal of this lesson is for students to be able to create visualizations to analyze sets of large data and to meaningfully interpret the patterns they uncover. They draw conclusions about themselves from relevant data, including local weather, the economics of their community, and naming trends.

Introduction

In the last activity, you considered some of the social issues with collecting and analyzing the explosive amounts of data now being produced. But the era of Big Data also has technical challenges.

To understand some of these technical issues, think about your school's library. How does the library organize the information stored in books? How do you find information in this library? How does the library prevent books from being lost or stolen? Would these methods work if the library had a million times as many books? Would these methods work if that many new books arrived every day? What problems would arise if the library tried to scale up a millionfold?

Practice: Introduction

These problems describe digital data, too. In the current decade, each year brings an enormous increase in the digital data being collected. Methods that were able to send, store, retrieve, and analyze data ten years ago no longer work. What are the new methods?

Write down your thoughts in your Computer Science journal. Discuss your thoughts with your family and friends.

Practice: Visualizing data

So much data is being collected that new technical issues arise. Big Data refers to many situations in which data is difficult to manage. Let's break it down. Watch this video and find out!

Practice: Visualizing Data

Open this presentation about Big Data and take notes in our Computer Science notebook.

The definition of Big Data is poorly agreed upon and is also changing as computers get more powerful. Four reasons describe why data might be considered Big Data.

- **Volume: Big Data has too many bytes to be stored by one computer.**
- **Velocity: Big Data is produced faster than one computer can store it.**
- **Variety: Big Data combines several different and conflicting sources.**
- **Volcanism: Big Data requires explosive amounts of human or computer processing to be useful.**

Practice: Data Visualization

Using the information given to you in the Big Data presentation, make a copy of the Big Data conversion activity and practice converting data. Check your work here.

Practice: Data Visualization

In the last few years, most new computers have been manufactured with dual or quad processors, and they mostly use a task parallel approach by working on different threads. The graphics card in a computer uses hundreds of parallel processors collectively called the graphics processing unit (GPU). The GPU processors share an instruction bus but operate on different streams of data. The Cuda platform allows you to write programs for the Central Processing Unit (CPU) that will assign data-parallel processing tasks to the GPU. Read https://developer.nvidia.com/about-cuda. That web site's list of application domains of parallel processing is repeated here. Pick one and skim a Wikipedia page about the application. Summarize what is being accomplished with parallel processing in that application.

- **Bioinformatics**
- **Computational Structural Mechanics**
- **Electronic Design Automation**
- **Medical Imaging**
- **Computational Finance**
- **Defense**
- **Numerical Analytics**
- **Computational Fluid Dynamics**
- **Data Science**
- **Computational Chemistry**
- **Electronic Design Automation**
- **Weather and Climate Simulation**

Practice: Data Visualization

Open this presentation over The Brain: Biological Parallel Processing and take notes in our Computer Science notebook.

The retina is light-sensitive brain tissue at the back of each eyeball. Light is directed by the lens to create an image on the retina. The neurons in the retina include light sensitive rods and red-, green-, or blue-sensitive cones. These neurons turn the image into a set of electrical signals. The signals are processed by another layer of neurons and are then sent to the back of the brain. The signal is sent through the optic nerve to the occipital lobe, the back of the brain. Before reaching the occipital lobe, the image is processed to identify dots, lines, and faces in all areas of the field of vision.

Practice: Data Visualization

Based on the presentation over The Brain: Biological Parallel Processing and any other research that you may need to do over the subject, answer the following questions in your Computer Science notebook. Discuss your answers with your family and friends. :

a. **The parallel processing from retina to occipital lobe identifies dots, lines, and faces in all areas of the field of vision. In what way is that description describing a task parallel approach? In what way is the approach data parallel?**

Practice: Data Visualization

Continued

b. The retina, optic nerve, and occipital lobe are made of neurons, the type of cell that sends electrical output to muscles or other neurons based on sensory input or electrical input from other neurons. Neurons transmit about 10 bits per second.
Each human retina has about 126 million light sensitive neurons. The signal from the retina is processed and carried by only 1 million neurons in the optic nerve. Estimate the image compression ratio as the percentage decrease in the size of the image in bytes.

Conclusion: Reflection

Open a page in the Computer science Notebook, and consider this question for your activity reflection.

Estimate how many bytes of data are collected about you and your immediate surroundings per day. Consider cameras, microphones, keyboard and mouse input, car sensors, etc.