**What is computer science?**

Computer science is concerned with information. Much as physics is concerned with matter and energy or biology is concerned with living organisms, the practice of computer science is grounded by its inherent and intimate involvement with information. Other disciplines certainly use information as well. Biologists capture information about the DNA structure of genomes which can be described in a simple four letter alphabet. Engineers often use information-based modeling tools to help understand and design mechanical or electrical artifacts. Even artists use information in creating novel works of digital art. However, computer science is distinctive in that information is the object of its study while information in other disciplines is important, but secondary. For example, while a biologist uses the information about a genome's DNA, the biologist's focus is on the genome and its properties, not on the information and its properties.

Representing and manipulating information are two key activities in computer science. Representation is important because information cannot be computed about until it has been made manifest in a concrete form. Three important aspects of representing information are modeling, data structuring, and displaying. Modeling concerns what information is being represented, how the pieces of information are related to each other, and how the information may change under different circumstances. A great deal of creativity and, often, experience is needed to determine the "right" model of information in a given situation. Creativity is required to guarantee that the information model has fidelity (corresponds to what is being modeled), is minimal (does not contain extraneous elements), is complete (does not ignore an essential aspect). Data structuring concerns how the modeled information can be represented in a computerized form. Computing devices have simple, primitive memories. The rich and complex model of information must be mapped onto these simple, primitive memories. An extensive body of knowledge exists about data structures. However, ingenuity is often needed to determine the "best" data structure to use in a given situation because there are often several, or many, ways to record a given information structure in the memory of a computing device. Each way of structuring the data implies what computing resources will be consumed to store and compute about the information. A careful understanding of the tradeoffs among these different data structures is critical. Displaying information for the human user is increasingly an important activity in computer science. Any one who has struggled with a mystifying user interface understands the importance of how information is displayed. Correspondingly, anyone who has used a natural, intuitive interface appreciates the significance of information display to an effective computer system.

The ability to automatically manipulate information is a defining characteristic of computer science. Mathematics, like computer science, has a rich sense of manipulating symbols. These symbols have meaning and the symbol manipulation yields new meaning. However, computer science focuses on manipulations that can be performed automatically by computing machines. The theoretical limit of computer science is defined by what is possible to compute in an automated way. Three major kinds of manipulations are transformation, transmission, and storage. Transformation involves manipulating the information to produce new information. An example of a simple

transformation is when a spreadsheet application determines the sum of a specified set of numbers. A critical aspect of computer science involves developing algorithms, the clearly defined, step-by-step procedures that dictate how a transformation is carried out. Simple transformations, like summation, have simple algorithms. Very complex transformations (e.g., matching the genes across two genomes) require very complex algorithms. An extensive body of knowledge exists about algorithms for different purposes and new algorithms are being developed as new transformations are needed. The transmission of information is a second critical way in which information is manipulated. In a networked world, where information is located and where it is needed are often on different computing devices. The two devices may be nearby (two devices in the same room) or as far apart as the opposite sides of the planet (or off planet if we take into account space stations and satellites). Moving people or goods between distant places involves complex transportation systems, vehicles, laws, infrastructure, and organizations. Similarly, the movement of information requires complex network systems, packets, protocols, naming systems, and standards bodies. Storing information for later use is a third way of manipulating information. Example of stored information include the contents web pages, music files, on-line books, patient medical information, and electronic records or all kinds. Storage as an act of manipulating information has a great deal of overlap with data structuring, but introduces additional issues due to is permanence including how it is named, how it may be searched, what limits are imposed on its sharing, how it is protected against device failure, and others.

Computer science can be involved wherever there is information inherent to artifacts or systems of interest. Importantly, the breadth and diversity of such artifacts and systems is enormous. The scope of these artifacts and systems means that computer science has natural collaborations with many other disciplines of study. The scope includes such general categories as natural, engineered, social, and artistic systems. The example given above of representing DNA information is an example of how computer science relates to entities in the natural world. Engineered systems are increasingly being designed and tested using computational processes. The recent inaugural flight of the newest Boeing airliner is but one example. Computer science has a multifaceted interface with social systems. In some cases computer systems create and sustain social systems through social networking tools, computer-supported collaborative work systems, and a variety of multimedia communication tools. In other cases algorithms are used to detect and study possible relationships among people (political parties, terrorist networks, peer-to-peer systems, medical provider networks) forming social systems. Finally, materials and media have, throughout history, been used by artists to achieve esthetic expression. Turning a mobile phone into a musical instrument and creating a work of art in a 3D virtual environment are only two examples of how computing technology can be appropriated for artistic purposes.

These observations lead to a working definition of computer science:

**Definition**: Computer science is the study of how to represent and manipulate the information inherent to natural, engineered, social, and artistic systems.