# **BÀI 3: MỘT SỐ KIỂU DỮ LIỆU VÀ DỮ LIỆU VĂN BẢN**

Môn học: Tin Học; Lớp: 10

Thời gian thực hiện: 2 tiết

**I**. **MỤC TIÊU**

***1****.* ***Kiến thức:***

* Nêu được các loại thông tin và các kiểu dữ liệu gặp trong chương trình tin học phổ thông
* Biết được các bảng mã thông dụng ASCII và Unicode.
* Giải thích sơ lược về việc số hóa văn bản.

***2****.* ***Kỹ năng:***

- Năng lực tự chủ và tự học

- Năng lực giao tiếp và hợp tác

- Năng lực sáng tạo và giải quyết vấn đề

***3****.* ***Phẩm chất:*** Nghiêm túc, tập trung, tích cực chủ động.

**II**. **THIẾT BỊ DẠY HỌC VÀ HỌC LIỆU**

Sgk, Sbt, giáo án.

**III**. **TIẾN TRÌNH DẠY HỌC**

**1. HOẠT ĐỘNG KHỞI ĐỘNG (MỞ ĐẦU)**

- **Mục tiêu**: Tạo hứng thú học tập cho học sinh

- **Nội dung**: Hs dựa vào hiểu biết để trả lời câu hỏi

- **Sản phẩm**: Từ yêu cầu Hs vận dụng kiến thức để trả lời câu hỏi GV đưa ra

- **Tổ chức thực hiện**: GV giới thiệu và dẫn dắt vào bài

Thông tin đưa và bộ nhớ máy tính dưới dạng các dãy bit. Như vậy khi đưa vào máy tính, phải mã hóa thông tin dưới dạng nhị phân. Tùy theo bản chất của thông tin được mã hóa mà dữ liệu tương ứng có cách biểu diễn riêng, từ đó hình thành nên các kiểu dữ liệu khác nhau. Vậy trong máy tính có các kiểu dữ liệu nào?

**2. HÌNH THÀNH KIẾN THỨC MỚI**

***Hoạt động 1:*** Tìm hiểu cách phân loại và biểu diễn thông tin trong máy tính

**- Mục Tiêu:** + Biết thông tin được phân thành mấy loại và cách biểu diễn thông tin trong máy tính

**- Nội dung**: HS quan sát SGK để tìm hiểu nội dung kiến thức theo yêu cầu của GV

-  **Sản phẩm:** Hs hoàn thành tìm hiều kiến thức

**- Tổ chức thực hiện:**

| **Sản phẩm dự kiến** | **Hoạt động của giáo viên và học sinh** |
| --- | --- |
| **1. PHÂN LOẠI VÀ BIỂU DIỄN THÔNG TIN TRONG MÁY TÍNH**  - Khi đưa vào máy tính thông tin được chuyển thành dữ liệu. Dữ liệu trên máy cũng cần được phân loại cho phù hợp với các phép xử lí trong máy tính.  Ví dụ, đối với các dữ liệu là số có thể tính toán và so sánh. Còn đối với các dữ liệu dạng văn bản thì có thể tách, ghép, so sánh.  - Việc mã hóa thông tin thành dữ liệu nhị phân được gọi là biểu diễn thông tin. Biểu diễn thông tin là bước đầu để có thể đưa thông tin vào máy tính.     * Biểu diễn thông tin là cách mã hóa thông tin. * Các kiểu dữ liệu thường gặp là văn bản, số, hình ảnh, âm thanh và logic. * Việc phân loại xử lí dữ liệu để có cách biểu diễn phù hợp nhằm tạo thuận lợi cho việc xử lí thông tin trong máy tính. | **\*‌ ‌Bước‌ ‌1:‌ ‌Chuyển‌ ‌giao‌ ‌nhiệm‌ ‌vụ:‌ ‌**  **GV:** Nêu đặt câu hỏi  Hình 3.1 minh họa thẻ căn cước công dân. Trên đó có những thông tin gì?  Hãy chia những thông tin đó thành các nhóm, ví dụ nhóm các thông tin có thể tách ghép được hay so sánh được để tìm kiếm và nhóm các thông tin có thể thực hiện được với các phép tính số học.    **?1.** Theo em số căn cước công dân có kiểu số hay kiểu văn bản?  **?2.** Kiểu số thực thường dùng để biểu diễn các số có phần thập phân (phần lẻ). Em hãy cho ví dụ một loại hồ sơ có dữ liệu kiểu số thực.  **HS:** Thảo luận, trả lời  **\*‌ ‌Bước‌ ‌2:‌ ‌Thực‌ ‌hiện‌ ‌nhiệm‌ ‌vụ:‌ ‌ ‌**   ‌+‌ ‌HS:‌ ‌Suy‌ ‌nghĩ,‌ ‌tham‌ ‌khảo‌ ‌sgk‌ ‌trả‌ ‌lời‌ ‌câu‌ ‌hỏi‌ ‌  +‌ ‌GV:‌ ‌quan‌ ‌sát‌ ‌và‌ ‌trợ‌ ‌giúp‌ ‌các‌ ‌cặp.‌ ‌ ‌  **\*‌ ‌Bước‌ ‌3:‌ ‌Báo‌ ‌cáo,‌ ‌thảo‌ ‌luận:‌ ‌ ‌**  +‌ ‌HS:‌ ‌Lắng‌ ‌nghe,‌ ‌ghi‌ ‌chú,‌ ‌một‌ ‌HS‌ ‌phát‌ ‌  biểu‌ ‌lại‌ ‌các‌ ‌tính‌ ‌chất.‌ ‌ ‌  +‌ ‌Các‌ ‌nhóm‌ ‌nhận‌ ‌xét,‌ ‌bổ‌ ‌sung‌ ‌cho‌ ‌nhau.‌ ‌ ‌  **\*‌ ‌Bước‌ ‌4:‌ ‌Kết‌ ‌luận,‌ ‌nhận‌ ‌định:‌ ‌‌**GV‌ ‌chính‌ ‌xác‌ ‌hóa‌ ‌và‌ ‌gọi‌ ‌1‌ ‌học‌ ‌sinh‌ ‌nhắc‌ ‌lại‌ ‌kiến‌ ‌thức‌**‌** |

***Hoạt động 2:*** Tìm hiểu biếu diễn dữ liệu văn bản

**a) Mục tiêu:** Nắm được một số bảng mã trong biểu diễn dữ liệu văn bản

**b) Nội dung:** HS quan sát SGK để tìm hiểu nội dung kiến thức theo yêu cầu của GV.

**c) Sản phẩm:** HS hoàn thành tìm hiểu kiến thức

**d) Tổ chức thực hiện:**

| **Sản phẩm dự kiến** | **Hoạt động của giáo viên và học sinh** |
| --- | --- |
| **2. BIỂU DIỄN DỮ LIỆU VĂN BẢN**  Việc đưa văn bản vào máy tính như thế nào không chỉ phụ thuộc vào kiểu dữ liệu là kí tự, xâu kí tự hay tệp văn bản mà còn phụ thuộc vào các kí tự ấy được mã hóa như thế nào? Cách mã hóa được quy định trong bảng kí tự.  **a) Bảng mã ASCII**  - là bảng mã được dùng phổ biến nhất trong tin học - “bảng mã chuẩn của Mĩ để trao đổi thông tin”  - Ban đầu bảng mã này dùng các mã 7 bit, với 128 (27) mã khác nhau nên chỉ thể hiện được đúng 128 kí tự.  - Bảng mã ASCII mở rộng (8 bit), cho phép mã hoá 256 kí tự, trong đó giữ nguyên 128 kí tự cũ. 128 vị trí được thêm vào trong bảng mã 8 bit so với bảng mã 7 bit được gọi là phần mở rộng của bảng mã ASCII. Trong bảng này, muốn lấy mã nhị phân của một kí tự thì chỉ cần ghép 4 bit ở chỉ số hàng với 4 bit ở chỉ số cột tương ứng với kí tự. Ví dụ mã nhị phân của "A" (có số thứ tự là 65) là 01000001.  **b) Bảng mã Unicode và tiếng Việt trong Unicode**  - Unicode thực tế là một bộ tiêu chuẩn biểu diễn kí tự văn bản trong máy tính, cho phép biểu diễn kí tự thuộc nhiều ngôn ngữ khác nhau trên thế giới.  - Các kí tự Unicode có thể mã hóa nhờ một số hệ thống định dạng chuyển đổi (tiếng anh là UTF), trong đó phổ biến nhất là UTF-8 (UTF 8 bit). UTF-8 là hệ thống mã hóa kí tự với độ dài khác nhau (từ 1 tới 5 byte) dành cho Unicode.  Từ năm 2017, VIệt Nam đã ban hành quy định bắt buộc sử dụng UTF-8 để biểu diễn các kí tự Việt trong máy tính, trong đó sử dụng bảng mã ASCII để mã hóa các kí tự La tinh không dấu, sử dụng 2 byte để mã hóa nguyên âm có dấu, các kí tự đ Đ và chỉ dùng 3 byte để mã hóa một số rất ít kí tự đặc biệt  Như vậy, hiểu một các ngắn gọn, các bảng mã ASCII và Unicode quy định cách biểu diễn kí tự  **c) Số hóa văn bản**  Tệp văn bản là định dạng lưu trữ ở bộ nhớ ngoài. Việc số hóa văn bản được thực hiện bằng các phần mềm soạn thảo văn bản như Word (của Microsoft) hay writer (của Open Office). Gần đây người ta có thể nhập văn bản bằng nhận dạng tiếng nói.  => Ghi nhớ  - Bảng mã ASCII mở rộng sử dụng 8 bit để biểu diễn một kí tự  - Unicode là bảng mã hợp nhất quốc tế, cho phép tạo ra các ứng dụng đa ngôn ngữ. Mỗi kí tự unicode có thể được mã hóa bởi nhiều byte. | **\*‌ ‌Bước‌ ‌1:‌ ‌Chuyển‌ ‌giao‌ ‌nhiệm‌ ‌vụ:‌ ‌**  **GV:**  **? 1**. Bảng kí tự La tinh có những kí tự nào?  **?2**. Trong tin học, mỗi nguyên âm có dấu thanh của tiếng Việt là một kí tự. Hãy kể tên các kí tự có trong tiếng Việt không có trong bảng kí tự La tinh. Có bao nhiêu kí tự như vậy?  **HS:** Thảo luận, trả lời  **HS:** Lấy các ví dụ trong thực tế.  **\*‌ ‌Bước‌ ‌2:‌ ‌Thực‌ ‌hiện‌ ‌nhiệm‌ ‌vụ:‌ ‌ ‌**  ‌+‌ ‌HS:‌ ‌Suy‌ ‌nghĩ,‌ ‌tham‌ ‌khảo‌ ‌sgk‌ ‌trả‌ ‌lời‌ ‌câu‌ ‌hỏi‌ ‌  +‌ ‌GV:‌ ‌quan‌ ‌sát‌ ‌và‌ ‌trợ‌ ‌giúp‌ ‌các‌ ‌cặp.‌ ‌ ‌  **\*‌ ‌Bước‌ ‌3:‌ ‌Báo‌ ‌cáo,‌ ‌thảo‌ ‌luận:‌ ‌ ‌**  +‌ ‌HS:‌ ‌Lắng‌ ‌nghe,‌ ‌ghi‌ ‌chú,‌ ‌một‌ ‌HS‌ ‌phát‌ ‌  biểu‌ ‌lại‌ ‌các‌ ‌tính‌ ‌chất.‌ ‌ ‌  +‌ ‌Các‌ ‌nhóm‌ ‌nhận‌ ‌xét,‌ ‌bổ‌ ‌sung‌ ‌cho‌ ‌nhau.‌ ‌ ‌  **\*‌ ‌Bước‌ ‌4:‌ ‌Kết‌ ‌luận,‌ ‌nhận‌ ‌định:‌ ‌‌**GV‌ ‌chính‌ ‌xác‌ ‌hóa‌ ‌và‌ ‌gọi‌ ‌1‌ ‌học‌ ‌sinh‌ ‌nhắc‌ ‌lại‌ ‌kiến‌ ‌thức‌  ?1. Mã nhị phân và mã thập phân của các kí tự S, G, K trong bảng mã ASCII là gì?  ?2. Trong bảng mã Unicode tiếng Việt, mỗi kí tự được biểu diễn bởi bao nhiêu byte?  a. 1 byte  b. 2 byte  c. 3 byte  d. từ 1 đến 3 byte |

**3. HOẠT ĐỘNG LUYỆN TẬP**

**a. Mục tiêu:** Củng cố, luyện tập kiến thức vừa học.

**b. Nội dung:** HS đọc SGK làm các bài tập.

**c. Sản phẩm:** Bài làm của học sinh, kĩ năng giải quyết nhiệm vụ học tập.

**d. Tổ chức thực hiện:**

Gv Cho HS nhắc lại KT:

Hs: Nhắc lại các vấn đề đã học

Câu 1. Giấy chứng nhận sở hữu xe máy có các thông tin nêu ở cột bên trái của bảng sau. Hãy ghép mỗi thông tin ở cột bên trái với kiểu dữ liệu thích hợp ở cột bên phải

|  |  |
| --- | --- |
| **Thông tin** | **Kiểu dữ liệu** |
| Họ và tên | Văn bản |
| Số căn cước công dân | Số nguyên |
| Biến số xe | Hình ảnh |
| Dung tích xy-lanh (cm3) | Âm thanh |
| Số khung, số máy |  |

Câu 2. Câu trả lời nào đúng cho câu hỏi “Tại sao cần có Unicode?”

1. Để đảm bảo bình đẳng cho mọi quốc gia trong ứng dụng tin học
2. Bảng mã ASCII mã hóa mỗi kí tự bởi 1 byte. Giá thành thiết bị lưu trữ ngày cành rẻ nên không cần phải sử dụng các bộ kí tự mã hóa bởi 1 byte.
3. Dùng 1 bảng mã chung cho mọi quốc gia, giải quyết vấn đề thiếu vị trí cho bộ kí tự của 1 số quốc gia, đáp ứng nhu cầu dùng nhiều ngôn ngữ đồng thời trong cùng 1 ứng dụng
4. Dùng cho các quốc gia sử dụng chữ tượng hình.

**4. HOẠT ĐỘNG VẬN DỤNG**

**a. Mục tiêu:** Vận dụng các kiến thức vừa học quyết các vấn đề học tập và thực tiễn.

**b. Nội dung:**.

**c. Sản phẩm:** HS vận dụng các kiến thức vào giải quyết các nhiệm vụ đặt ra.

**d. Tổ chức thực hiện:**

Gv đưa câu hỏi về nhà:

Câu 1. Dựa trên bảng mã ASCII, Việt Nam xây dựng bảng mã VSCII (Vietnamese standard Code for infomation interchange), còn gọi là TCVN 5712: 1993. Hãy tìm hiểu bảng mã này trên Internet theo những gợi ý sau:

* Bảng mã có đủ cho tất cả kí tự tiếng Việt không?
* Bảng mã có bảo toàn bảng mã ASCII 7 bit không?

Câu 2. Có 2 bộ gõ tiếng Việt thông dụng là Unikey và Vietkey. Nếu mở bảng điều khiển của 2 phần mềm này ta sẽ thấy rất nhiều bảng mã tiếng Việt trong đó có TCVN3. Em hãy tìm hiểu trên Internet để biết bảng mã TCVN3 là gì.

**5. Hướng dẫn học sinh tự học:**

**- Hướng dẫn học bài cũ:**

***-* Hướng dẫn chuẩn bị bài mới:**

**..........................................................................................................................................**