

inform

Généralités et vocabulaire

antique

Repères chronologiques

On peut distinguer trois grandes phases dans l'histoire du calcul artificiel.

La première commence à la préhistoire et se termine dans les années 1930.

On y invente des outils pour faire de plus en plus facilement les opérations arithmétiques.

C'est pour éviter aux hommes des calculs longs et pénibles qu'on cherche à inventer des dispositifs matériels qui puissent le faire à leur place. C'est par exemple pour aider son père, collecteur des impôts, que Pascal inventa la pascaline.

Dans l'antiquité les sumériens utilisaient des boules et des jetons d'argile, des bouliers ont été utilisés dès 3000 ans avant notre ère.

1623 Machine à calculer de Wilhelm Schickard. Elle permet de faire des additions et des soustractions avec retenues. Cette machine, que Schickard appelait « horloge calculante » fonctionnait grâce à des roues dentées. C'est la première machine à calculer de l'histoire.

1641 A 19 ans **Blaise Pascal invente la pascaline.** Construite pour aider le père de Pascal, c'est la première machine à calculer à avoir réellement fonctionné. Elle ne permet d'effectuer que des additions.

1694 Machine à calculer de Gottfried Wilhelm von Leibniz. Elle effectue les 4 opérations.

1728 Invention des **cartes perforées par Jean-Baptiste Falcon** pour les métiers à tisser.

1806 Métiers à tisser à cartes perforées de Joseph-Marie Jacquard. D'abord détruits par les canuts, ils connaîtront ensuite un grand développement.

1820 Machine à différences de Charles Babbage. C'est une machine très complexe constituée de nombreux engrenages.

1822 Arithmomètre de Charles-Xavier Thomas de Colmar. Une machine modeste (3 chiffres en entrée et 6 en sortie) mais fabriquée en 1500 exemplaires de 1823 à 1878.

1833 Machine analytique de Babbage. Aidé de Ada Lovelace, Babbage conçoit une autre machine qui sera construite par son fils. Elle est considérée comme l'ancêtre des ordinateurs. Les deux machines de Babbage ne furent jamais terminées mais les plans étaient exacts comme une équipe le démontra en 1991 en réalisant un exemplaire.

1840 Ada LOVELACE décrit les composants d'une machine à calculer et invente le mot algorithme en l'honneur du mathématicien arabe AL KHOWARIZMI.

Elle explique qu'une machine à calculer doit comprendre un dispositif capable d'introduire des données numériques, une mémoire pour conserver les valeurs entrées, une unité de commande pour indiquer à la machine les tâches que l'on veut effectuer, un « moulin » pour réaliser les calculs et un dispositif pour connaître les résultats. On reconnaît là les principes qui seront à la base des premiers ordinateurs.

1854 Théorie de la **logique binaire** de **G. Boole**.

1867 Première **machine à écrire** inventée par deux américains. La disposition des lettres (QWERTY) est destinée à freiner l'ardeur des secrétaires.

1884 Machine à statistiques à cartes perforées de **H. Hollerith**.

Elle est conçue pour le recensement de 1890 aux Etats-Unis. Elle utilise des cartes perforées et est construite par la Tabulating Machine Company qui deviendra International Business Machine (IBM). Elle donnera naissance à toute une lignée de machines dites mécanographiques.

1928 Invention de la **carte perforée** d'IBM à 80 colonnes.

1930 Machine Enigma.

1936 Alan Turing définit la « machine de Turing », une machine imaginaire qui est un automate algorithmique universel.

1937 Réalisation d'un **additionneur binaire à relais** par **G. Stibitz** (le kitchen adder !).

1938 Z1, premier calculateur à relais par **Konrad Zuse**. Il est construit par Zuse dans la salle à manger de ses parents.

1939 ABC de **J. Atanasoff et Clifford Berry**. Première machine à utiliser le système binaire. Il faisait une multiplication par seconde. La mémoire est de 60 mots de 50 bits.

1940 Invention du **circuit imprimé**.

Les composants sont reliés entre eux par des pistes métalliques très fines.

La même année et séparément George Robert Stibitz aux USA et Konrad Zuse en Allemagne mettent au point ce que nous appelons le **calcul en virgule flottante**. (Calcul avec une mantisse et un exposant).

1941 Z3 de **K. Zuse**. Une machine composée de 2 600 relais, d'un lecteur de bandes et d'une console. Il travaille sur des nombres en virgule flottante. Il peut contenir 64 nombres de 22 chiffres.

1943 Harvard MARK 1 de **Howard Hathaway Aiken**.

Une machine de 5 tonnes composée de 765 000 éléments, refroidie par plusieurs tonnes de glace.

La deuxième phase de l'histoire du calcul artificiel commence dans les années 1930.

On y voit les grands calculateurs qui pèsent des tonnes succéder aux calculatrices de bureau.

Trois grands types de calculateurs coexistent pendant cette période : les calculateurs électromécaniques (par Stibitz, Aiken et Zuse), les calculateurs électroniques (par Atanasoff et l'équipe Eckert et Mauchly) et les calculateurs analogiques.

Les calculateurs analogiques disparaîtront rapidement pour céder la place au calcul digital et les calculateurs électromécaniques seront abandonnés à cause de leur lenteur.

1943 Colossus 1. Machine anglaise composée de 1 500 lampes et d'un lecteur de bandes capable de lire 5 000 caractères par seconde.

1944 Le lieutenant **Goldstine**, qui travaillait sur l'ENIAC, rencontre sur le quai de la gare d'Aberdeen **John von Neumann**. L'un travaille sur la technologie de l'avenir, l'autre sur l'architecture de l'avenir. Leur rencontre va accélérer la création de l'ordinateur.

1945 Publication par John von Neumann de « First Draft of a report on the EDVAC », un texte court mais d'une grande pertinence qui contient les principes sur lesquels fonctionnent encore les ordinateurs d'aujourd'hui. Von Neumann y décrit ce que nous appelons aujourd'hui une « architecture de Von Neumann » avec programme enregistré et unité de commande interne.

1946 ENIAC par J. Eckert et J. Mauchly (Electronic Numerical Integrator and Computer)

Commandé par l'armée des Etats-Unis en 43 pour effectuer les calculs de balistique, il remplaçait 200 personnes. Il occupait 23 m³, pesait 30 tonnes et consommait 200 kW. Il comporte 70 000 résistances, 10 000 condensateurs, 1 500 relais et 6 000 commutateurs. 17 000 lampes (19 000 changements en 9 ans.). A sa première démonstration publique, l'ENIAC additionna 5 000 nombres ensemble en une seconde.

1947 Invention du transistor par Bardeen, Brattain et Shockley (qui recevront pour cette découverte le prix Nobel en 1956). Progressivement le transistor va remplacer les lampes, c'est un gain de taille et de fiabilité, c'est aussi une consommation d'énergie beaucoup plus faible et donc un échauffement réduit.

1947 EDVAC Premier calculateur à programme enregistré de John von Neumann. C'est le descendant de l'ENIAC. 4 000 tubes. Mémoire de 1 024 mots de 44 bits.

1947 Premier langage : l'assembleur.

1948 Shannon publie sa « Théorie de l'information ». Il y introduit le bit (Binary digIT) comme unité de mesure de la quantité d'information.

1948 IBM SSEC de Wallace Eckert. Une vitrine pour IBM (20 000 relais et 12 500 tubes)

1948 Manchester Mark 1 (The baby). Conçu par Neumann en Angleterre. C'est la première machine à suivre un programme enregistré.

Après la guerre commence la troisième phase de notre histoire, celle des ordinateurs.

1949 EDSAC par Maurice WILKES

1949 Invention de la **bande magnétique**.

1951 UNIVAC par **J. Presper Eckert et John Mauchly**.

Il utilise des bandes magnétiques. Il comporte 5 000 tubes sa mémoire est de 1 000 mots de 12 bits. Il peut réaliser 8 000 additions ou 500 multiplications par secondes. Il occupe 25 m².

1951 Premier langage compilé, le A0.

1954 Invention du **transistor au silicium** plus simple à fabriquer et moins cher que les composants au germanium.

1954 Invention des mémoires vives à **tores de ferrites**.

1955 Premier **calculateur transistorisé**.

1955 Invention du **tambour magnétique**.

1956 Invention du **disque magnétique**.

1956 Naissance du **Fortran** (FORmula TRANslation)

1959 Invention du **circuit intégré** par Jack Kilby. Une date importante car le circuit intégré va permettre la miniaturisation progressive des ordinateurs et la baisse de leur coût.

1959 IBM 1401. Transistors et tores de ferrites. Il est associé à une imprimante rapide (600 lignes minutes). 12 000 exemplaires furent vendus (prévisions 1000)

1959 Création du Cobol par G. Hopper. (COMmon Business-Oriented Language)

1959 LISP, un langage créé par John Mc Carthy qui contribuera au développement de la programmation objet.

1960 Invention de la **diode électroluminescente**.

1960 PDP-1.

1960 Algol par Edsger Wybe Dijkstra (ALGORithmic Language)

1963 Invention de la **souris** par Douglas Engelbart.

1964 IBM 360. 6 ordinateurs et 44 périphériques compatibles (la technologie est très classique : transistors et mémoires à tores)

1964 PL/1.

1965 PDP-8 de DEC. C'est un mini-ordinateur (18 000\$) 50 000 exemplaires vendus.

1965 Basic.

1968 Invention des **mémoires à semi-conducteurs.**

1968 Pascal de Niklaus Wirth.

1969 Naissance d'**Internet.**

1969 UNIX.

1970 Invention de la **technologie M.O.S.** La course à la miniaturisation commence.

1971 Premier microprocesseur le 404 d'Intel. (2 300 transistors, 60 000 opérations par seconde à une fréquence de 108 kHz. Sa puissance est égale à celle de l'ENIAC.

1971 Première **imprimante matricielle.**

1972 Création du langage **C.**

1972 Prolog (PROgrammation en LOGique) par Alain Colmerauer (le langage descriptif de l'Intelligence Artificielle).

1973 Micral-N de R2E

Premier micro-ordinateur du monde inventé par **A. Truong**, fondateur français de la société R2E.

1973 Présentation de l'**Alto de Xerox.** Il tient compte des idées proposées au Palo-Alto Research Center (PARC). Il introduit les notions de fenêtres, d'icône et de souris. En raison de son coût, il ne sera pas un succès commercial.

1974 Invention de la **carte à puce** par Rolland Moreno.

1974 CP/M-DOS de G. Kildall. (Control Program for Microcomputers - Disk Operating System). Premier système d'exploitation pour micro-ordinateur. L'ancêtre de MS-DOS.

1975 Imprimante laser.

1976 CRAY 1. Créé par Seymour Clay. Premier ordinateur à architecture vectorielle.

1977 Apple II par S. Jobs et S. Wozniak. Premier ordinateur à recevoir un succès grand public.

1978 DEC VAX 11/780. Premier modèle de supermini. Ordinateur 32 bits. Son système d'exploitation est VMS.

1978 Wordstar par John Barnaby. Un des premiers traitement de texte.

1979 Visicalc (VISIble CALCulator) par Dan Bricklin et Bob Franston. Le premier tableur.

1979 ADA par Jean Ichbiah.

1981 Sinclair ZX81 Micro-ordinateur à moins de 1 000 F. Il utilise l'écran de télévision.

1981 IBM-PC. La réaction d'IBM à la micro-informatique.

1981 Lotus 1.2.3. Premier logiciel de gestion intégré.

1981 MS-DOS. Système d'exploitation mis au point par Microsoft.

1982 Invention du compact disc.

1982 PostScript d'Adobe.

1983 Lisa d'Apple. Echec commercial mais interface graphique et souris.

1984 Macintosh d'Apple. Début de la P.A.O.

1984 Invention du **CD-ROM.**

1985 Windows de Microsoft.

1986 Premier ordinateur parallèle composé de 16 000 processeurs.

1987 RNIS (Réseau numérique à intégration de services).

Voix, images et données circulent désormais en numérique.

1989 Naissance du **web** inventé par T. Berners-Lee du CERN (HTML).

1991 Linux. par Linus Torvalds.

1994 PowerMac d'Apple.

1995 Invention du **DVD-ROM**. Environ 8,5 Go de données.

1995 Java (Sun microsystems).

1996 XML (eXtensible Markup Language) Sous-ensemble simplifié de SGML

1998 Naissance du logiciel libre.

1998 iMac d'Apple.

2000 C# par Microsoft.

2001 Clé USB. Puce mémoire dotée d'un connecteur USB, dépasse les capacité d'un CD-ROM.

2003 L'École Estienne crée la Licence « Flux numérique ».

Les ordinateurs les plus puissants d'aujourd'hui

Pour avoir une idée des ordinateurs les plus puissants actuellement en fonctionnement, nous avons utilisé le site « TOP500 SUPERCOMPUTERS SITES » qui est administré par les universités de Mannheim et du Tennessee et par un organisme américain, le NERSC/LBNL. Ce classement est évidemment discutable (la méthodologie de classement est exposée sur le site) mais elle permet de se faire une idée d'ensemble de la situation.

Les unités utilisées pour mesurer les performances des super calculateurs sont le Mflop : million d'opérations en virgule flottante par seconde (en 64 bits), le Gflop vaut 1 milliard d'opérations par seconde et le Tflop vaut mille milliards d'opérations par secondes.

La machine classée première est actuellement le BlueGene/L DD2 beta-System (0,7 GHz PowerPC 440) située à Rochester aux Etats-Unis. Fabriquée en 2004 par IBM, cette machine utilise 32 768 processeurs.

Sa puissance théorique est de 91 750 Gflop, sa puissance mesurée de 70 720 Gflop.

Cet ordinateur sera utilisé pour des calculs de simulation dans différents domaines : biologie, hydrodynamique, modélisation de phénomènes en 3D,...

Il devrait encore grandir pour atteindre 360 Tflop en 2005.

Les caractéristiques finales de la machine sont les suivantes :

16 To de mémoire vive.

400 To de mémoire sur disque.

131 072 processeurs.

2 MW de consommation électrique.

La **deuxième machine** est le SGI Altix 1,5 GHz, Voltaire Infiniband de Mountain View aux U.S.A.

Cette machine a une puissance de 51 870 Gflop. Elle est utilisée par la NASA.

Elle comporte 20 SGI Altix de 512 processeurs interconnectés soit 10 240 processeurs au total. La mémoire total est de 20 To.

La **troisième machine** est le « Earth Simulator Center ». Installée en 2002 à Yokohama au Japon.

Elle a une puissance de 35 860 Gflop.

Elle comprend 5 120 processeurs et une mémoire totale de 10 To.

La **quatrième machine** est « MareNostrum » du Barcelona Supercomputer Center de Barcelone en Espagne et la cinquième « Thunder » du « Lawrence Livermore National lab » à Livermore aux USA.

Si l'on s'intéresse à l'ensemble du classement on constate qu'environ 400 machines dans le monde dépassent le Tflop.

IBM est le principal constructeur avec 43 % des systèmes et la moitié de la puissance installée.

267 des 500 machines les plus puissantes sont aux Etats-Unis.

L'ordinateur

Mémoire virtuelle

Zone d'un support de mémoire de masse considérée par un logiciel ou le système d'exploitation comme de la mémoire vive.

Mémoire morte

Circuit mémoire dans lequel les informations peuvent être lues mais non modifiées. Les mémoires mortes conservent leurs informations même en l'absence d'alimentation électrique. (on emploie souvent le sigle anglais ROM)

Microprocesseur

Ce terme désignait à l'origine les processeurs miniaturisés, il est aujourd'hui un quasi synonyme du mot processeur.

Un micro-ordinateur est généralement constitué d'une unité centrale et de périphériques.

L'unité centrale est un boîtier qui regroupe :

- Une **alimentation** (alim), qui alimente tous les autres éléments en courant électrique. Elle utilise le courant du secteur (alternatif 220 V) et le transforme pour délivrer du courant continu (5 V pour les composants électroniques, 12 V pour certains périphériques).
- Une **carte mère**. C'est un circuit imprimé qui regroupe le processeur, de la mémoire, et des dispositifs d'entrée/sortie.
- Un ou des **ventilateurs** permettent d'évacuer la chaleur produite par le fonctionnement des composants. Des radiateurs facilitent la dissipation de la chaleur.
- Au moins un **disque dur** qui contient l'essentiel du système d'exploitation de la machine.

Les périphériques dépendent de l'utilisation de l'ordinateur. Les plus fréquents sont :

- les souris
- les trackball
- les écrans
- les tablettes graphiques
- les disques externes
- les lecteurs/graveurs de CD ou de DVD
- Les imprimantes
- les convertisseurs analogique/numérique
- les clés USB
- les crayons optiques

- les écrans tactiles

Des appareils de plus en plus nombreux peuvent maintenant être connectés à un ordinateur (appareils photo numériques, caméscopes numériques,...) et deviennent, de fait, des périphériques.

La présentation qui vient d'être faite de l'ordinateur était parfaitement exacte il y a quelques années. Avec la généralisation de l'outil informatique, l'ordinateur prend aujourd'hui des formes très variées et disparaît même, de plus en plus souvent, en tant qu'appareil distinct et visible. Sa miniaturisation permet son intégration à des dispositifs industriels ou à des produits grand public dans lesquels il devient peu ou pas visible (jouets, automobile, machines-outils, dispositifs de signalisation, avions, missiles,...)

Le succès de l'informatique se soldera probablement par la disparition quasi complète des ordinateurs en tant qu'objets clairement identifiables.

Le traitement de l'information

UNIX

Système d'exploitation multitâche créé par AT&T.

Systemes d'exploitation

Sans logiciel un ordinateur ne sait rien faire. Pour pouvoir exécuter des programmes, il a besoin d'un logiciel de base qu'on appelle **Systeme d'exploitation**.

Pour mieux faire comprendre son rôle, nous allons décrire ses principales fonctions.

- C'est lui qui lance les programmes et règle leur accès au processeur.
- **C'est lui qui gère la mémoire vive.** Il la répartit entre les applications qui sont lancées et s'il en manque il utilise une zone du disque qui est vue par le processeur comme de la mémoire vive. C'est la **mémoire virtuelle**.
- **Il gère les entrées/sorties.** Grâce à des pilotes (gestionnaires de périphériques) il permet aux applications d'utiliser les périphériques.
- **Il peut détecter le mauvais fonctionnement d'une application** et l'interrompre sans qu'il soit nécessaire d'arrêter la machine.
- **Il gère les droits d'accès** des utilisateurs, garantissant la sécurité.
- **Il gère le système de fichiers.** Quand on fait « enregistrer » dans un logiciel, Adobe InDesign par exemple, ce n'est pas ce logiciel qui enregistre. InDesign fait un appel au système d'exploitation qui réalise l'opération demandée.

On parle de système d'exploitation **multitâche** quand un système peut accomplir plusieurs tâches en même temps et de système préemptif quand il sait répartir le temps machine entre les processus qui en font la demande avec des règles précises.

Les systèmes d'exploitation modernes (Windows NT, Windows XP, MacOS X, Linux, UNIX, VMS) sont tous multitâche préemptifs et multi-utilisateurs.

Langages

Généralités

Nous allons nous intéresser à la manière dont les applications sont fabriquées et à la manière dont elles fonctionnent.

Les processeurs ne peuvent exécuter que des programmes en langages machine. Ce sont des suites de 0 et de 1 qui sont incompréhensibles pour les humains. Il n'est pas possible de rédiger des programmes complexes en langage machine.

L'**assembleur** est une première étape pour rendre la programmation abordable. On écrit avec des codes qui sont ensuite traduits en langage machine par un programme qu'on appelle un assembleur. Mais la programmation reste ainsi très aride.

L'assembleur est encore très proche du langage machine qu'il contribue seulement à rendre lisible. Il dépend du processeur utilisé et un programme écrit en assembleur pour une machine ne pourra donc pas fonctionner sur une autre.

Les **langages évolués** sont très nombreux, ils permettent d'écrire des programmes beaucoup plus facilement. Ils sont relativement proches des langages naturels mais possèdent évidemment une syntaxe très précise qu'il faut apprendre.

Il est très difficile de classer les langages informatiques car on peut se placer de différents points de vue qui ont chacun leur pertinence. Quel que soit le principe utilisé, un langage appartient souvent à 2 catégories.

Pour simplifier, nous distinguerons grossièrement 3 grandes catégories : **les langages de programmation** (qui permettent d'effectuer un traitement de certaines données en fournissant une série d'instructions), **les langages de définition de données** (qui permettent d'enregistrer des données structurées) et **les langages de définition de page** (qui permettent de définir l'apparence de documents imprimés ou diffusés sur le web).

Algorithme :

(d'al-Kharezmi, mathématicien arabe) Méthode de calcul dont les étapes sont déterminées avec précision, permettant de parvenir à un résultat final. On parle par exemple d'un algorithme de compression de données, d'un algorithme permettant de générer une texture.

Assembleur

1. Programme informatique traduisant en langage machine les programmes écrits en langage d'assemblage.
2. langage d'assemblage dépendant du processeur utilisé.

BASIC

(Beginner's All purpose Symbolic Instruction Code)

Langage de programmation créé en 1964 par Johnny Kemeny et Thomas Kurz du Dartmouth College. Longtemps utilisé pour initier à la programmation, il est aujourd'hui abandonné pour des langages plus modernes ne présentant pas ses inconvénients (absence de récursivité, instruction GOTO menant à des programmes impossibles à maintenir, pauvreté des types,...)

FORTRAN

(Abréviation de FORmula TRANslation) Langage informatique, l'un des premiers conçus pour les applications scientifiques.

ADA

(du prénom de la Comtesse Lovelace, fille de Lord Byron, considérée comme la première programmeuse de l'histoire) Langage de programmation créé par J. Ichbiah en 1979 pour l'armée américaine.

Javascript

Langage de script créé par Netscape. Javascript n'est pas un sous-ensemble de java même s'il lui ressemble. Il ne peut être utilisé qu'à l'intérieur d'une page web (i.e. à l'intérieur d'un document HTML).

PERL

Langage de programmation développé par Larry Wall. Il est utilisé, en particulier, pour écrire des scripts CGI.

Programmation orientée objet

Type de programmation moderne dans laquelle les définitions de données et les procédures qui les concernent sont regroupées dans des ensembles appelés objets.

Bytecode

(on écrit aussi Byte-Code) Code intermédiaire produit par un compilateur Java pour être utilisé sur une plate-forme quelconque par la machine virtuelle Java appropriée qui le transformera en code machine exécutable.

Machine virtuelle java

Logiciel traduisant le bytecode en code exécutable sur une machine donnée. Elle peut être insérée dans un navigateur web mais ce n'est pas une nécessité.

SGML

(Standard Generalised Markup Language) Langage permettant de structurer des documents hypertexte. HTML est défini en SGML.

XML :

(Abréviation anglaise pour eXtensible Markup Language) Langage de balisage permettant d'enregistrer et de transmettre de l'information de façon structurée. XML est dérivé de SGML. On peut en XML définir ses propres balises mais selon une syntaxe très précise.

HTML

(HyperText Markup Language) Langage de balisage qui permet de créer des sites web.

UML :

(Abréviation anglaise pour Unified Modeling Language) Langage de modélisation de troisième génération servant à représenter un système objet. Il est utilisé lors de la phase de développement.

IDE :

(Integrated Development Environment) Environnement de développement intégré. Ensemble de logiciels aidant le programmeur à développer.

Les langages de programmation

On peut les classer en 3 catégories :

Les langages interprétés

Ils sont écrits dans un langage évolué et sont traduits au fur et à mesure de leur exécution en langage machine. Ils nécessitent donc en permanence un programme de traduction qu'on appelle un **interpréteur**. Les applications écrites dans un langage interprété sont donc toujours plus lentes que celles écrites dans un langage compilé. Ils sont par contre plus facilement modifiables.

BASIC, Prolog ou Perl sont des langages interprétés.

Les langages compilés.

Dans ce type de langage, le programme écrit par le programmeur est traduit une fois pour toute en langage machine par un programme appelé **compilateur**. Le compilateur produit un programme qui peut fonctionner sans aide extérieure. On l'appelle un **exécutable**.

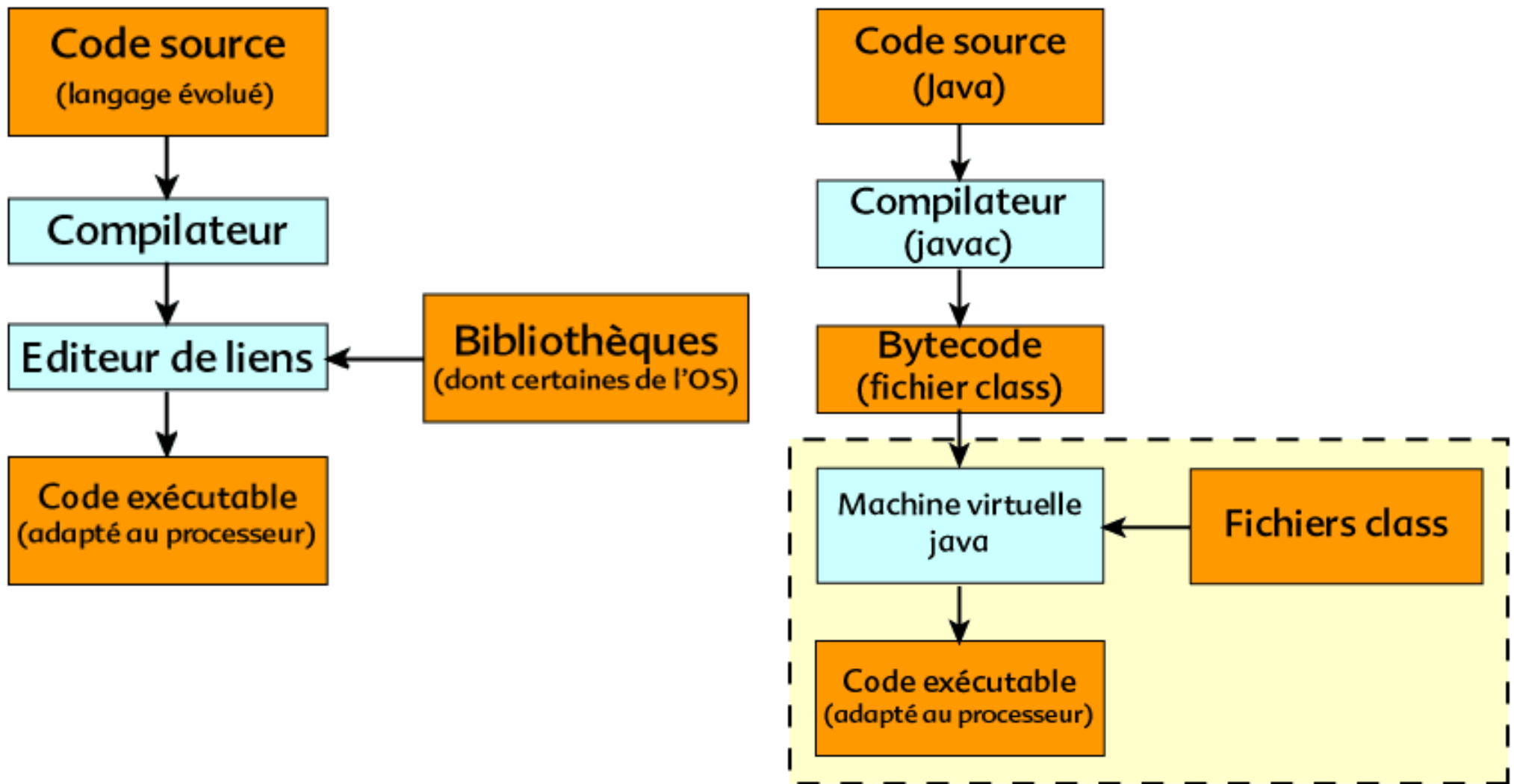
Ce type de programme est rapide (on n'a pas besoin de faire la traduction en langage machine) mais la mise au point est plus difficile car à chaque modification, il faut refaire la traduction en langage machine, la **compilation**. Il présente aussi l'avantage de la confidentialité, il est très difficile de comprendre le fonctionnement d'un programme compilé.

Le schéma ci-dessous présente en le précisant le processus de compilation.

Le code source écrit dans un langage évolué est compilé c'est à dire traduit en code exécutable par le processeur concerné sur la plate-forme concernée. C'est un compilateur qui réalise cette traduction.

L'éditeur de lien incorpore des portions de code déjà écrites, il fait appel aux bibliothèques du système d'exploitation concerné (Microsoft Foundation classes sous Windows, MacOS Toolbox sous MacOS). Il en résulte un code exécutable sur une seule plate-forme. (XPress existe sur Mac et sur PC, mais ce n'est pas le même code qui s'exécute sur les deux machines).

C, C++, COBOL, FORTRAN, ADA ou Pascal sont des langages compilés.



Bézier

(Courbes de) Ces courbes qui doivent leur nom au mathématicien et ingénieur français Pierre Bézier, sont à la base du langage de description de page Postscript (et accessoirement de la société américaine Adobe) Une courbe de Bézier peut être enregistrée de manière très concise car chaque portion de courbe est déterminée par seulement 4 points (2 points d'ancrage et 2 points directeurs)

Bézier (Pierre)

Ingénieur en mécanique et en électricité, Pierre Bézier est né le 1^{er} septembre 1910 à Paris. Il est mort le 25 novembre 1999.

Diplômé de l'École Nationale Supérieure d'Arts et Métiers en 1930 et de l'École supérieure d'électricité (Supélec) en 1931, il recevra le titre de docteur en mathématique de l'université de Paris en 1977.

Entré chez Renault en 1933, il y fera toute sa carrière jusqu'en 1975. Il y conçoit en 1945 des machines transferts pour la ligne de fabrication de 4CV et en 1958 l'une des premières machines à commande numérique d'Europe. Il enseigne la productique au Conservatoire national des arts et métiers de 1968 à 1979.

Il est surtout connu pour son invention des courbes de Bézier qui sont couramment utilisées en informatique.

Bitmap

1. Se dit en général d'une image pixellisée par opposition aux documents « vectoriels » (i.e. définis à l'aide d'un langage de description de page comme le Postscript).
2. Dans Photoshop uniquement, se dit d'une image à deux niveaux (blanc et noir).

Vectoriel

Se dit, par opposition à bitmap, d'une page ou d'une illustration qui est décrite avec un langage de description de page comme le Postscript et non divisée en pixels dont on enregistre les couleurs.

Les langages intermédiaires (type Java)

Intéressons-nous au fonctionnement de Java, l'un des langages les plus importants aujourd'hui.

Le développeur écrit son application en Java et la compile à l'aide du compilateur javac mais le code résultant n'est pas exécutable directement sur une machine, il s'agit d'un code intermédiaire (appelé **byte code**) qui aura besoin pour s'exécuter d'une sorte de traducteur, d'adaptateur qui va le rendre exécutable sur une plate-forme donnée.

C'est ce logiciel qui s'appelle une **machine virtuelle Java**.

L'intérêt de cette méthode c'est que le byte code constitue une application inter-plates-formes. Le développeur ne travaille plus pour une plate-forme particulière, il n'a pas à se préoccuper des appels aux bibliothèques des systèmes d'exploitation. Java dispose de ses propres bibliothèques mais elles sont indépendantes des plates-formes.

Si Java a été très apprécié par ceux qui développent des sites Internet c'est qu'ils s'adressent à des internautes dotés de machines différentes utilisant des systèmes d'exploitation différents. Java est donc une excellente solution pour faire tourner des programmes sur la machine de tous les internautes.

C'est possible depuis que tous les navigateurs récents intègrent une machine virtuelle Java et que HTML permet d'inclure un appel à un programme Java à l'intérieur d'une page. Ces programmes sont appelés des applets (ce mot est féminin, il devrait normalement être traduit en français par « appliquette » mais...)

Les langages de définition de données.

Le plus important est sans doute SGML (Standard Generalised Markup Language - Langage normalisé de balisage généraliste). Il a été créé en 1986. Il s'agit d'un métalangage qui permet de créer des langages de description de pages. SGML est utilisé dans le monde de l'édition professionnelle, en particulier dans la documentation technique.

SGML ne définit pas de balises, le créateur d'un document SGML peut inventer ses propres balises ou utiliser une bibliothèque de balises pré-définie. On l'appelle une DTD (Document Type Definition). Quand nous écrivons une page pour le web en HTML, nous utilisons un langage qui est défini par une DTD en SGML.

XML (eXensible Markup Language) est aussi un métalangage, il s'agit d'un SGML allégé. Il s'est imposé comme l'un des langages de balisage les plus utilisés.

XML a été mis au point par le WC (World Wide Web Consortium) dès 1996.

Un document XML décrit des données mais non leur formatage.

Un document suivant les règles XML est un **document bien formé**.

Un document XML possédant une DTD et étant conforme à celle-ci est un **document valide**.

Pour décoder un document XML on utilise un analyseur (Parser)

De nombreux langages spécialisés sont définis en XML : MathML (Mathematical Markup Language) est consacré à la représentation des formules mathématiques, CML (Chemical Markup Language) est utilisé pour décrire des formules chimiques.

Les langages de description de page.

Il permettent de décrire l'apparence, le formatage d'une page.

PostScript a pris une place considérable dans les industries graphiques, d'abord seul, ensuite intégré dans les fichiers PDF.

HTML est à la base du développement du web.

Unités.

Dans un ordinateurs, les informations sont traitées sous forme binaire. La plus petite quantité d'information (un 0 ou un 1) est appelé bit (Binary digIT). Elle est représentée physiquement de différentes manières (présence ou absence d'une trou dans une carte, aimantation ou non aimantation d'une zone, valeur d'une caractéristique électrique dépassant ou non une valeur donnée.

L'octet (en anglais Byte, abréviation B capitale) est une suite de 8 bits.

Quand on a une suite de n bits, on peut représenter 2^n valeurs différentes.

Exemples :

Avec 2 bits on peut représenter 2^2 valeurs soit 4 (00, 01, 10 et 11)

Avec 8 bits (1 octet) on peut représenter 2^8 valeurs soit 256.

Dans toutes les sciences, on utilise les préfixes suivant pour construire des unités (préfixes SI) :

kilo (k) 10^3

méga (M) 10^6

giga (G) 10^9

téra (T) 10^{12}

péta (P) 10^{15}

exa (E) 10^{18}

zetta (Z) 10^{21}

yotta (Y) 10^{24}

On devrait donc avoir 1 ko = 1 000 octets comme on a 1kg = 1 000 g. Malheureusement les informaticiens ont pris la mauvaise habitude de travailler avec des puissances de 2 et donc pour eux :

1 kilo-octet (ko ou kB) = 2^{10} octets soit 1024 octets.

1 Méga-octet (Mo ou MB) = 2^{20} octets soit 1 048 576 octets (1 024 ko).

1 Giga-octet (Go ou GB) = 2^{30} octets soit 1 073 741 824 octets (1 024 Mo).

1 Téra-octet (To) = 2^{40} octets soit 1 099 511 627 776 octets (1 024 Go).

Et ainsi de suite.

Normalement, ce système ne devrait plus être utilisé depuis 98 mais dans la réalité il n'en est rien.

Signalons tout de même qu'on devrait normalement utiliser les préfixes SI dans leur sens habituel (kilo pour 10³, méga pour 10⁶)... et utiliser les nom suivants pour les préfixes binaires :

kibi (KI ou K) pour 2^{10} (1 024)

mébi (Mi) pour 2^{20}

gibi (Gi) pour 2^{30}

tébi (Ti) pour 2^{40}

pébi (Pi) pour 2^{50}

exbi (Ei) pour 2^{60}

Les nombres binaires étant peu maniables, les informaticiens utilisent souvent l'hexadécimal (notation en base 16).

Codage

Quand on veut représenter des caractères on doit utiliser un code. On utilise de manière à peu près universelle le code ASCII (American Standard Code for Information Interchange pour Code Américain Standard pour l'Échange d'Informations). Ce code représente les caractères avec 7 bits (voir tableau ci-dessous)

On remarquera que les premiers codes sont des codes de contrôle et non des caractères à proprement parler.

ASCII

(American Standard Code for Information Interchange) On prononce « à ski ». Code permettant de transmettre les caractères. Ce code est très répandu, c'est lui qui permet l'échange des données entre des plates-formes différentes. Seuls les 128 premiers codes (sur 7 bits) sont universels.

UNICODE

Code destiné à remplacer l'ASCII. Chaque caractère est codé sur 2 octets ce qui autorise 65 536 signes et donc l'utilisation d'un grand nombre d'alphabets. Java utilise Unicode en standard.

On notera aussi que ce code a été établi pour la langue anglaise, il ne comporte donc pas de caractères accentués. Pour coder les textes français il a fallu l'étendre (et utiliser un octet par caractère) malheureusement l'ASCII étendu est moins standard que l'ASCII proprement dit (il dépend de la plate-forme utilisée).

Comme on a souvent besoin sur Internet et dans les industries graphiques d'alphabets non latins, on a mis au point en 1991 le code UNICODE qui va s'imposer peu à peu. Il utilise 2 octets pour coder chaque caractère ce qui donne 65 536 caractères différents. Ainsi ont pu être intégrés les alphabets arabe, arménien, cyrillique, grec, hébreu...

Code ASCII

NUL (Null)	000	!	33	21	B	66	42	c	99	63
SOH (Start of heading)	101	"	34	22	C	67	43	d	10	64
STX (Start of text)	202	#	35	23	D	68	44	e	101	65
ETX (End of text)	303	\$	36	24	E	69	45	f	102	66
EOT (End of transmission)	404	%	37	25	F	70	46	g	103	67
ENQ (Enquiry)	505	&	38	26	G	71	47	h	104	68
ACK (Acknowledge)	606	'	39	27	H	72	48	i	105	69
BEL (Bell)	707	(40	28	I	73	49	j	106	6A
BS (Backspace)	808)	41	29	J	74	4A	k	107	6B
TAB (tabulation horizontale)	909	*	42	2A	K	75	4B	l	108	6C
LF (Line Feed, saut de ligne)	100A	+	43	2B	L	76	4C	m	109	6D
VT (Vertical tabulation, tabulation verticale)	110B	,	44	2C	M	77	4D	n	110	6E
FF (Form feed)	120C	-	45	2D	N	78	4E	o	111	6F
CR (Carriage return, retour à la ligne)	130D	.	46	2E	O	79	4F	p	112	70
SO (Shift out)	140E	/	47	2F	P	80	50	q	113	71
SI (Shift in)	150F	0	48	30	Q	81	51	r	114	72
DLE (Data link escape)	1610	1	49	31	R	82	52	s	115	73
DC1 (Device control 1)	1711	2	50	32	S	83	53	t	116	74
DC2 (Device control 2)	1812	3	51	33	T	84	54	u	117	75
DC3 (Device control 3)	1913	4	52	34	U	85	55	v	118	76
DC4 (Device control 4)	2014	5	53	35	V	86	56	w	119	77
NAK (Negative acknowledgement)	2115	6	54	36	W	87	57	x	120	78
SYN (Synchronous idle)	2216	7	55	37	X	88	58	y	121	79
ETB (End of transmission block, fin de bloc de transmission)	2317	8	56	38	Y	89	59	z	122	7A
CAN (Cancel, annulation)	2418	9	57	39	Z	90	5A	{	123	7B
EM (End of medium, fin du médium)	2519	:	58	3A	[91	5B		124	7C
SUB (Substitute, substitut)	261A	;	59	3B	\	92	5C	}	125	7D
ESC (Escape, caractère d'échappement)	271B	<	60	3C]	93	5D	~	126	7E
FS (File separator, séparateur de fichier).....	281C	=	61	3D	^	94	5E	(Touche de		
GS (Group separator, séparateur de groupe)	291D	>	62	3E	_	95	5F	suppression)		
RS (Record separator, séparateur d'enregistrement)	301E	?	63	3F	`	96	60			
US (Unit separator, séparateur d'enregistrement)	311F	@	64	40	a	97	61			
SP (Space, espace)	3220	A	65	41	b	98	62			

L'internet

Minitel

Jadis et seulement en France, appareil permettant de recevoir très lentement des informations et servant à faire la fortune d'une société de téléphone publique et de quelques sociétés privées.

HTTP

(Hyper Transfer Protocol) Protocole de transmission pour les documents hypermedia utilisé sur internet.

Hypertexte

Texte comportant des liens et permettant donc une circulation non linéaire dans l'information.

Hypermedia

Se dit des documents comportant des liens à partir de portions de texte mais aussi d'autres éléments en particulier des images.

URL

(Uniform Resource Locator) Adresse électronique d'une ressource sur internet.

Adresse IP

(Internet protocol) Adresse identifiant une machine sur Internet. Elle est constitué de 4 nombres compris entre 0 et 255, séparés par des points. Dans la pratique, on utilise des noms symboliques qui sont traduits en adresse IP par les DNS.

ADSL

(Asymmetric Digital Subscriber Line) Technologie proposée par le CNET pour transmettre plus d'informations que les lignes classiques (elle utilise une bande passante de 100 kHz à 1,1 MHz).

Bande passante :

(on dit aussi largeur de bande)

1. Intervalle des fréquences transmises sans distortions sur un canal de transmission ou un support.
2. Par extension, s'agissant d'un réseau, et d'internet en particulier, débit de transmission maximum.

Comment ça marche ?

L'internet est un réseau qui réunit un très grand nombre de machines dans le monde entier. Comme il réunit aussi de nombreux réseaux qui ont leur existence propre, on l'appelle parfois le réseau des réseaux.

L'information circule sur ce réseaux grâce à des **protocoles** qui permettent d'avoir accès à des services, à des ressources différents. C'est parce que le web est le plus connu et le plus apprécié des services proposé qu'il est parfois employé, à tort, comme un synonyme de l'internet.

Le protocole qui est la base de tous les échanges sur Internet d'appelle TCP/IP. C'est lui qui permet des relier des ordinateurs qui ne possèdent pas le même système d'exploitation et n'ont en fait pas grand chose en commun d'un point de vue logiciel.

D'autres protocoles de plus haut niveau permettent d'utiliser internet pour obtenir des ressources, des services différents :

l'IRC permet le dialogue en direct

FTP permet le transfert de fichiers.

HTTP permet de consulter des sites web.

Comment faire un site ?

Généralités.

On distingue les **sites web statiques** composées de pages web en HTML stockées sur le serveur et **les sites web dynamiques** dont les pages sont créées « à la volée » en exécutant un programme sur le serveur qui interroge souvent une base de données. Des solutions intermédiaires sont possibles.

Le langage de base reste l'HTML même si une page statique peut intégrer une applet en java ou une séquence Flash.

Navigateur

(on dit aussi butineur, brouteur, explorateur, fouineur, fureteur ou (moins bien) browser). Logiciel permettant de consulter des serveurs web sur Internet, de télécharger des fichiers et, pour les dernières versions, de gérer son courrier et de participer aux groupes de discussion. Netscape Communicator, Microsoft Internet Explorer Firefox ou Safari sont des navigateurs.

Cookie :

Suite de caractères envoyée par un serveur Internet au navigateur qui l'interroge. Stockée sur l'ordinateur du « client » (dans le fichier MagicCookie sur un Mac), ces informations permettront de tenir compte de la consultation actuelle lors d'interrogations ultérieures.

CSS

(Cascading Style Sheet) Sur le web, feuille de style simplifiant la mise en forme typographique des pages.

DNS

(DOMAINE NAME SYSTEM) Système indispensable au routage des messages sur Internet. Il convertit les adresses « en clair » du type « ruses.com » en adresse IP du type « 163.5.67.8 ». Les DNS se mettent à jour en échangeant leurs données.

Pare-feu

(On dit aussi, moins bien, firewall) Système de sécurité placé entre Internet et un réseau privé qui sert à protéger le serveur d'une entreprise contre les intrusions malveillantes.

FTP

(File Transfer Protocol) Protocole utilisé sur Internet pour télécharger les fichiers. C'est lui qui nous permet de mettre à jour le site de la tour des ruses.

Image map

1. Définition de différentes zones dans une image avec les liens hypermedias qui leur correspondent.
2. Image possédant une image map au sens de la définition précédente.

POP

(Post Office Protocol) Protocole utilisé sur Internet qui régit la manière dont les messages sont envoyés à leur destinataire.

Pour la génération des pages des sites dynamiques, de nombreux langages peuvent être utilisés. A côté de solutions propriétaires (.NET, Coldfusion, Java,...) le logiciel libre prend une place croissante. C'est en fait une véritable plate-forme constitué de logiciels libre qui est maintenant proposée.

On l'appelle LAMP (Linux pour le système d'exploitation, Apache pour le serveur web, MySQL pour la base de données et PHP pour le langage de programmation. On trouvera quelques informations ci dessous sur HTML, CSS et Javascript.

HTML

HTML est le langage qui a permis le développement du web, il reste à la base de toutes les pages publiées même si de nombreuses techniques sont venues palier ses insuffisances.

On consultera le W3C (World Wide Web Consortium) pour connaître son histoire.

HTML est un langage de balisage qui rappelle les langages utilisés naguère en photocomposition.

Les balises (des codes placés entre crochets <comme ça>) n'indiquaient au début que la structure du texte, elles décrivaient son contenu.

<H1>Le site des renards bleus</H1>

indiquait que le texte inséré entre les deux balises était un titre de premier niveau.

<cite>« Il faut aimer les profils ICC » </cite>

indiquait que le texte placé entre les balises était une citation.

 Trait d'Union Graphique

indiquait que le texte placé entre les balises était particulièrement important et méritait une « accentuation ».

Ces balises permettent ce que l'on appelle un balisage logique. C'est le navigateur qui les interprète et qui décide en dernier ressort de l'apparence du texte (et donc de la page sur l'écran). On ne dispose pas, avec ce type de balisage d'un réel contrôle sur l'apparence de la page.

Au fil des normes successives, HTML s'est enrichi à la demande des concepteurs de sites de nombreuses

SMTP

(Simple Mail Transport Protocol) Protocole utilisé sur Internet dans le domaine des messageries pour assurer la transmission des messages entre les serveurs.

TCP/IP

(Transmission Control Protocol / Internet Protocol) Protocole de communication utilisé sur Internet.

Web

L'un des services disponibles sur Internet que l'on confond trop souvent avec le réseaux des réseaux. C'est l'ensemble des pages construites en HTML consultables grâce aux navigateurs.

balises qui indiquent directement quelle doit être l'apparence du texte concerné à l'écran, on parle alors de balisage physique.

`Ingrid` est arrivée.

permet de mettre Ingrid en gras.

`<I>Cécile</I>` est au jardin

permet de mettre Cécile en italique.

Cette évolution a fini par rendre HTML ambigu, il est maintenant condamné.

Le W3C a adopté XHTML Basic. Cette nouvelle norme tient compte des problèmes rencontrés avec HTML 4. XHTML permet de travailler pour les ordinateurs mais aussi pour les assistants personnels et les téléphones portables. Il assurera une transition vers XML.

L'apparition d'éditeurs tel écran tel écrit comme DreamWeaver ou Golive a permis à des concepteurs qui ne connaissent pas HTML de bâtir des sites. Si la connaissance au moins partielle de ce langage permet parfois de se sortir d'une difficulté ponctuelle, elle ne constitue plus un préalable à la création de sites professionnels. Il est par contre nécessaire de se tenir au courant de l'évolution des normes et des techniques.

CSS

Ce sigle est l'acronyme de Cascading Style Sheet (feuille de style en cascade). Le principe est le même que celui des feuilles de style dans les logiciels de PAO ou de bureautique. Il s'agit d'appliquer une mise en forme pré définie à une portion de texte. Comme une mise en forme exige en général plusieurs paramètres, l'utilisation des CSS permet de gagner du temps.

Il y a 3 manières d'appliquer des feuilles de style à un texte. IL s'agit de savoir quels seront les sélecteurs c'est à dire de quelle manière seront délimitées les zones de texte auxquelles les styles doivent s'appliquer.

- La méthode la plus courante et la plus sûre consiste à utiliser les balises HTML (généralement celles qui assurent un balisage logique) comme sélecteurs. L'intérêt de cette méthode est de conserver une

compatibilité avec les navigateurs qui ne savent pas interpréter les CSS. On va par exemple structurer les différents niveaux de titres en utilisant les balise H1, H2, H3... La hiérarchie des titres sera respectée par tous les navigateurs et les navigateurs compatibles CSS les afficheront avec les attributs typographiques spécifiques que nous aurons choisis avec les styles.

- Il est également possible d'appliquer un style à une zone de texte qui ne correspond pas à un balisage HTML en créant une classe. Cette technique est couramment utilisée aujourd'hui.
- Enfin, il est aussi possible de créer des indentifiants qui permettront d'appliquer un style à une seule zone de texte. La prudence est recommandée.

Il est possible de placer une feuille de style dans un fichier spécifique (on parle alors de feuille de style externe) auquel des pages ou même toutes les pages d'un site feront appel.

On insère dans l'en-tête de la page la ligne suivante :

```
<link rel: 'STYLESHEET' href='style.css' type='text/css'>
```

qui permet d'appeler le fichier externe qui définit les styles.

Javascript

Javascript est un langage qui permet d'incorporer des scripts dans une page web. HTML a de grandes qualités mais il a aussi ses limites, en particulier il ne sait réaliser seul que des pages statiques. La seule interactivité possible est celle de l'utilisation des liens.

Javascript va permettre l'exécution de commandes soit au moment du chargement de la page, soit quand un événement créé par l'utilisateur se produit (survol d'une zone avec le curseur, clic sur une zone, frappe au clavier...)

Javascript est un langage interprété, ce qui signifie que le navigateur traduit une à une les instructions d'un script et les exécute.

L'intérêt de Javascript est de réaliser sur le poste de l'internaute des traitements qui sans lui nécessiteraient l'intervention du serveur avec les inconvénients qui en résultent en terme d'utilisation de la bande passante, de temps d'attente et de charge pour le serveur.

Zoologie :

Discipline scientifique consacrée à l'étude des animaux en particulier ceux que l'on rencontre dans le monde des arts graphiques : lapins, puces, hirondelles, ours, chiens, souris, mulots, chameau, hamsters, bourdons... et renards.

Un des exemples les plus courants que l'on puisse donner est celui du traitement d'un formulaire. Il est souvent utile de vérifier que l'internaute a correctement rempli certains champs.

Javascript va nous permettre de vérifier que certains champs ne sont pas vides, que les adresses e-mail comportent bien le signe @, que le numéro d'une carte bleue est « vraisemblable » (algorithme de Luhn)... Mais il peut également permettre d'effectuer le calcul d'un devis, de programmer de petits jeux, d'ouvrir des fenêtres avec des attributs prédéterminés. Bref, il apporte de sérieuses possibilités de calcul et d'interactivité.

Son inconvénient majeur est l'absence de norme. Un script ne se comportera pas forcément de la même façon dans Netscape et dans Microsoft Internet Explorer car les interpréteurs sont différents. En fait, ce sont deux versions différentes et concurrentes du même langage qui ont été développées et soutenues par les deux éditeurs.



SYLVAIN
RENARD

Sylvain Renard

Formation et réalisations multimédias

58, Bd Henri Dunant. Bâtiment A

91100 Corbeil-Essonnes

Tél. : 01 64 96 02 31

srenard@ruses.com

www.ruses.com

v 1.0.