

Groupe de travail Réseau
Request for Comments : 20
STD 80

Vint Cerf, UCLA
16 octobre 1969
Traduction Claude Brière de L'Isle

Format ASCII pour les échanges réseau

Pour être concret, on suggère l'utilisation de l'ASCII standard à 7 bits incorporé dans un octet de 8 bits dont le bit de poids fort est toujours 0. Cela conduit au code standard donné ci-dessous, copié de la norme USAS X3, 4-1968. Ce code sera utilisé sur les connexions primaires d'hôte à hôte. Les caractères de coupure seront définis par l'hôte distant receveur, par exemple, SRI utilise "." (ASCII X'2E' ou 2/14) comme caractère de fin de ligne, tandis que UCLA utilise X'OD' ou 0/13 (retour chariot).

ASCII : Code standard des USA pour les échanges d'information

1. Domaine d'application

Le présent jeu de caractères codé est à utiliser pour les échanges généraux d'informations entre les systèmes de traitement de l'information, les systèmes de communication, et les équipements qui leur sont associés.

2. Code standard

B \ b7 ----->					0	0	0	0	1	1	1	1
I \ b6 ----->					0	0	1	1	0	0	1	1
T \ b5 ----->					0	1	0	1	0	1	0	1
S					----->							
Colonne-->					0	1	2	3	4	5	6	7
b4	b3	b2	b1	Rangée								
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(8	H	X	h	x
1	0	0	1	9	HT	EM)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	
1	1	0	1	13	CR	GS	-	=	M]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL

3. Représentation de caractère et identification de code

La représentation de caractère standard à 7 bits, avec b7 comme bit de poids fort et b1 comme bit de moindre poids, est montrée ci dessous :

Exemple : La représentation binaire du caractère "K," positionné dans la colonne 4, rangée 11, est

b7	b6	b5	b4	b3	b2	b1
1	0	0	1	0	1	1

La position du tableau de code pour le caractère "K" peut aussi être représentée par la notation "colonne 4, rangée 11" ou aussi par "4/11". L'équivalent décimal du nombre binaire formé par les bits b7, b6, et b5, forme collectivement le numéro de colonne, et l'équivalent décimal du nombre binaire formé par les bits b4, b3, b2, et b1, forme collectivement le numéro de rangée.

Le code standard peut être identifié par l'utilisation de la notation ASCII ou USASCII.

La notation ASCII (prononcée aski) ou USASCII (prononcé u.s.aski) devrait ordinairement être comprise comme signifiant le code prescrit par la plus récente version de la norme. Pour désigner explicitement une version particulière (peut-être plus ancienne) les deux derniers chiffres de l'année de production peuvent être ajoutés, comme, "ASCII 63" ou "USASCII 63".

4. Légende

4.1 Caractères de contrôle

NUL Nul	DLE (<i>Data Link Escape</i>) échappement de ligne de données (CC)
SOH (<i>Start of Heading</i>) début d'en-tête (CC)	DC1 (<i>Device Control 1</i>) commande d'appareil 1
STX (<i>Start of Text</i>) début de texte (CC)	DC2 (<i>Device Control 2</i>) commande d'appareil 2
ETX (<i>End of Text</i>) fin de texte (CC)	DC3 (<i>Device Control 3</i>) commande d'appareil 3
EOT (<i>End of Transmission</i>) fin de transmission (CC)	DC4 (<i>Device Control 4</i>) commande d'appareil 4 (Stop) Arrêt
ENQ (<i>Enquiry</i>) demande (CC)	NAK (<i>Negative Acknowledge</i>) accusé de réception négatif (CC)
ACK (<i>Acknowledge</i>) accusé de réception (CC)	SYN (<i>Synchronous Idle</i>) repos synchrone (CC)
BEL (<i>Bell</i>) (signal audible ou d'attention)	ETB (<i>End of Transmission Block</i>) fin de bloc de transmission (CC)
BS (<i>Backspace</i>) espace arrière (FE)	CAN (<i>Cancel</i>) annule
HT (<i>Horizontal Tabulation</i>) tabulation horizontale (FE)	EM (<i>End of Medium</i>) fin de support
LF (<i>Line Feed</i>) saut à la ligne (FE)	SUB (<i>Substitute</i>) substitut
VT (<i>Vertical Tabulation</i>) tabulation verticale (FE)	ESC (<i>Escape</i>) échappement
FF (<i>Form Feed</i>) page suivante (FE)	FS (<i>File Separator</i>) séparateur de fichier (IS)
CR (<i>Carriage Return</i>) retour chariot (FE)	GS (<i>Group Separator</i>) séparateur de groupe (IS)
SO (<i>Shift Out</i>) glissement externe	RS (<i>Record Separator</i>) séparateur d'enregistrement (IS)
SI (<i>Shift In</i>) glissement interne	US (<i>Unit Separator</i>) séparateur d'unité (IS)
	DEL (<i>Delete</i>) supprime [1]

Légende : (CC) Contrôle de communication
(FE) Effecteur de format
(IS) Séparateur d'informations

Note [1] Au sens strict, DEL n'est pas un caractère de contrôle. (Voir au paragraphe 5.2)

4.2 Caractères graphiques

Colonne/Rangée	Symbole	Nom
2/0	SP	Espace (Normalement non imprimable)
2/1	!	Point d'exclamation
2/2	"	Guillemets (tréma [2])
2/3	#	Numéro [3], [4]
2/4	\$	Dollar
2/5	%	Pour cent
2/6	&	Éperluette
2/7	'	Apostrophe (Accent aigu, guillemet simple [2])
2/8	(Parenthèse ouverte
2/9)	Parenthèse fermée
2/10	*	Astérisque

2/11	+	Plus
2/12	,	Virgule (Cédille [2])
2/13	-	Trait d'union (Moins)
2/14	.	Point
2/15	/	Barre oblique
3/10	:	Deux points
3/11	;	Point virgule
3/12	<	Moins que (Inférieur à)
3/13	=	Égal
3/14	>	Supérieur à (Plus que)
3/15	?	Point d'interrogation
4/0	@	Arobase (à commercial) [3]
5/11	[Crochet ouvrant [3]
5/12	\	Barre oblique inverse [3]
5/13]	Crochet fermant [3]
5/14	^	Accent circonflexe [2], [3]
5/15	~	Souligné
6/0	`	Accent grave [2], [3] (Guillemet simple ouvrant)
7/11	{	Accolade d'ouverture [3]
7/12		Ligne verticale [3]
7/13	}	Accolade de fermeture [3]
7/14	~	Tilde [2], [3]

Notes :

- [2] L'utilisation des symboles 2/2, 2/7, 2/12, 5/14, /6/0, et 7/14 comme marques diacritiques est décrite dans l'Appendice A5.2 de X3.4-1968
- [3] Ces caractères ne devraient pas être utilisés dans les échanges internationaux sans avoir déterminé si il y a accord entre l'expéditeur et le receveur. (Voir l'Appendice B4 de X3.4-1968.)
- [4] Dans les applications où le symbole # n'est pas exigé, le symbole £ (Livre Sterling) peut être utilisé dans la position 2/3.

5. Définitions

5.1 Généralités

- (CC) Contrôle de communication : caractère fonctionnel destiné à contrôler ou faciliter la transmission des informations sur les réseaux de communication.
- (FE) Effecteur de format : caractère fonctionnel qui contrôle les informations de disposition ou de positionnement des informations dans les appareils d'impression ou d'affichage.
- (IS) Séparateur d'informations : caractère qui est utilisé pour séparer et qualifier des informations dans un sens logique. Il y a un groupe de quatre de ces caractères, qui sont à utiliser dans un ordre hiérarchique.

5.2 Caractères de contrôle

NUL (Nul) : c'est le caractère tout à zéro qui peut servir à réaliser un bourrage temporel ou du support.

SOH (*Start of Heading*) début d'en-tête : caractère de contrôle de communication utilisé au début d'une séquence de caractères qui constitue une adresse significative pour la machine ou une information d'acheminement. Une telle séquence est appelée "l'en-tête". Un caractère STX a pour effet de terminer un en-tête.

STX (*Start of Text*) début de texte : caractère de contrôle de communication qui précède une séquence de caractères qui est à traiter comme une entité et est entièrement transmise à la destination ultime. Une telle séquence est appelée un "texte". STX peut être utilisé pour terminer une séquence de caractères qui commence par SOH.

ETX (*End of Text*) fin de texte : caractère de contrôle de communication utilisé pour terminer une séquence de caractères commencée par STX et transmise comme une entité.

EOT (*End of Transmission*) fin de transmission : caractère de contrôle de communication utilisé pour indiquer la conclusion d'une transmission, qui peut avoir contenu un ou plusieurs textes et des en-têtes associés.

ENQ (*Enquiry*) demande : caractère de contrôle de communication utilisé dans les systèmes de communication de données comme une demande de réponse de la part d'une station distante. Il peut être utilisé comme "qui êtes vous ?" (WRU,

Who Are You) pour obtenir une identification, ou peut être utilisé pour obtenir l'état de la station, ou les deux.

- ACK (*Acknowledge*) accusé de réception : caractère de contrôle de communication transmis par un receveur comme réponse affirmative à un envoyeur.
- BEL (*Bell*) sonnerie : caractère à utiliser lorsque il y a besoin d'attirer l'attention d'une personne. Il peut contrôler des appareils d'alarme ou qui attirent l'attention.
- BS (*Backspace*) espace arrière : effecteur de format qui commande le mouvement de la position d'impression d'un espace d'impression vers l'arrière sur la même ligne d'impression. (Applicable aussi aux appareils d'affichage.)
- HT (*Horizontal Tabulation*) tabulation horizontale : effecteur de format qui commande le mouvement de la position d'impression sur la suivante dans une série de positions prédéterminées sur la ligne d'impression. (Aussi applicable aux appareils d'affichage et à la fonction de saut sur une carte perforée (*une séquence binaire*)).
- LF (*Line Feed*) saut à la ligne : effecteur de format qui commande le mouvement de la position d'impression à la ligne d'impression suivante. (Aussi applicable aux appareils d'affichage.) Lorsque approprié, ce caractère peut avoir la signification "Nouvelle ligne" (NL), un effecteur de format qui commande le mouvement du point d'impression sur la prochaine ligne d'impression à la ligne d'impression suivante. L'utilisation de cette convention exige un accord entre l'envoyeur et le receveur des données.
- VT (*Vertical Tabulation*) tabulation verticale : effecteur de format qui commande le mouvement de la position d'impression à la suivante d'une série de lignes d'impression prédéterminée. (Aussi applicable aux appareils d'affichage.)
- FF (*Form Feed*) saut de page : effecteur de format qui commande le mouvement de la position d'impression à la première ligne d'impression prédéterminée de la page suivante. (Aussi applicable aux appareils d'affichage.)
- CR (*Carriage Return*) retour chariot : effecteur de format qui commande le mouvement de la position d'impression à la première position d'impression sur la même ligne d'impression. (Aussi applicable aux appareils d'affichage.)
- SO (*Shift Out*) glissement externe : caractère de contrôle qui indique que les combinaisons de code qui suivent doivent être interprétées comme étant en dehors du jeu de caractères du tableau de code standard jusqu'à ce que survienne un caractère Shift In.
- SI (*Shift In*) glissement interne : caractère de contrôle qui indique que les combinaisons de code qui suivent doivent être interprétées conformément au tableau de code standard.
- DLE (*Data Link Escape*) échappement de ligne de données : caractère de contrôle de communication qui va changer la signification d'un nombre limité de caractères contigus suivants. Il est exclusivement utilisé pour fournir des commandes supplémentaires dans les réseaux de communication de données.
- DC1, DC2, DC3, DC4 (*Device Controls*) commandes d'appareil : caractères pour le contrôle des appareils auxiliaires associés aux systèmes de traitement de données ou de télécommunication, et plus précisément les appareils de commutation "marche" ou "arrêt". (Si une seule commande "stop" est requise pour interrompre ou arrêter les appareils auxiliaires, DC4 est l'allocation préférée.)
- NAK (*Negative Acknowledge*) accusé de réception négatif : caractère de commande de communication transmis par un receveur comme réponse négative à l'envoyeur.
- SYN (*Synchronous Idle*) repos synchrone : caractère de contrôle de communication utilisé par un système de transmission synchrone en l'absence de tout autre caractère pour fournir un signal à partir duquel la synchronisation peut être réalisée ou conservée.
- ETB (*End of Transmission Block*) fin de bloc de transmission : caractère de contrôle de communication utilisé pour indiquer la fin d'un bloc de données pour les besoins de la communication. ETB est utilisé pour mettre en bloc des données lorsque la structure de bloc n'est pas nécessairement en relation avec le format de traitement.
- CAN (*Cancel*) annule : caractère de contrôle utilisé pour indiquer que les données avec lesquelles il est envoyé sont erronées ou sont à ignorer.
- EM (*End of Medium*) fin de support : caractère de contrôle associé aux données envoyées qui peut être utilisé pour identifier la fin physique du support, ou la fin de la portion utilisée, ou voulue, des informations enregistrées sur un support.

(La position de ce caractère ne correspond pas nécessairement à la fin physique du support.)

SUB (*Substitute*) substitut : caractère qui peut être substitué à un caractère qui est déterminé comme invalide ou erroné.

ESC (*Escape*) échappement : caractère de contrôle destiné à fournir une extension de code (caractères supplémentaires) dans un échange d'informations générales. Le caractère Échappement lui-même est un préfixe qui affecte l'interprétation d'un nombre limité de caractères contigus suivants.

FS (*File Separator*) séparateur de fichier, GS (*Group Separator*) séparateur de groupe, RS (*Record Separator*) séparateur d'enregistrement, et US (*Unit Separator*) séparateur d'unité : ces séparateurs d'informations peuvent être utilisés dans des données de façon facultative, mais leur relation hiérarchique doit être : FS est le plus inclusif, puis GS, puis RS, et US est le moins inclusif. (Le contenu et la longueur d'un fichier, groupe, enregistrement, ou unité ne sont pas spécifiés.)

DEL (*Delete*) supprime : ce caractère est principalement utilisé pour "écraser" ou "oblitérer" des caractères erronés ou non voulus dans une bande perforée (*une séquence binaire*). (Au sens strict, DEL n'est pas un caractère de contrôle.)

5.3 Caractères graphiques

SP (*Space*) espace : caractère graphique normalement non imprimable utilisé pour séparer les mots. C'est aussi un effecteur de format qui commande le mouvement de la position d'impression, une position d'impression plus loin. (Aussi applicable aux appareils d'affichage.)

6. Considérations générales

- 6.1 La présente norme ne définit pas les moyens par lesquels l'ensemble codé est enregistré sur un support physique ; il ne comporte aucune redondance et ne définit pas de techniques de contrôle d'erreur. De plus, la présente norme ne définit pas de structure de caractères de communication de données, de formats de communication de données, de techniques d'extension de code, ou de représentation graphique des caractères de contrôle.
- 6.2 Toute déviations de la norme peut créer de sérieuses difficultés dans les échanges d'informations et ne devrait être utilisée qu'en pleine connaissance de cause par les parties impliquées.
- 6.3 La séquence relative de deux caractères quelconques, lorsque utilisée comme base de collationnement, est définie par leurs valeurs binaires.
- 6.4 Aucune signification spécifique n'est prescrite pour le graphisme du tableau de code à part qu'elle soit comprise par les utilisateurs. De plus, la présente norme ne spécifie pas un type de style pour l'impression ou l'affichage des divers caractères graphiques. Dans des applications spécifiques, il peut être désirable d'employer un style distinctif des graphismes individuels pour faciliter leur usage pour des fins spécifiques comme, par exemple, pour styliser le graphisme dans les positions de code 2/1 et 5/15 selon celui associé fréquemment avec le OU logique (|) et le NON logique (252), respectivement.
- 6.5 Les appendices (*non inclus dans la RFC*) de la présente norme contiennent des information supplémentaires sur la conception et l'usage de ce code.

[Cette RFC a été mise en forme machine pour être placée dans l'archive des RFC en ligne par Robbie Bennet 9/99]