



Des bugs¹ riches d'enseignement



Sommaire

1 Donkey Kong.....	2
2 Ariane 5.....	3
3 Affichage bizarre de page html (« arrête ton charset »).....	4

1 Un bug ou bogue est un défaut de conception d'un programme informatique à l'origine d'un dysfonctionnement (source Wikipédia).

1 Donkey Kong

Donkey Kong, jeu vidéo d'Arcade créé par Nintendo en 1981 est l'un des premiers jeux de plates-formes. Donkey Kong introduit le personnage de Mario qui deviendra une véritable icône du genre. Ce jeu calcule pour chaque niveau la durée dont le joueur dispose pour finir le niveau en question.

Le calcul est simple et repose sur la formule suivante :

$$\text{duree_niveau [s]} = \text{niveau} \times 10 + 40$$

La durée du niveau est stockée dans un octet et est codée en binaire naturel.

Lors de sa sortie aucun joueur n'a pu dépasser le niveau 22 car l'écran de fin apparaissait trop tôt.



Figure 1 : Donkey Kong

Q1. Quelle est la durée du niveau 22 du « point de vue du jeu » ?



Un nombre entier codé sur n bits en binaire naturel ne peut représenter que les valeurs entières allant de 0 à $2^n - 1$.

Q2. En déduire pourquoi aucun joueur n'a pu dépasser le niveau 22 lors de la sortie de Donkey Kong.

Q3. Proposer une manière de résoudre ce bug.

Pour les curieux voici un lien vers un dépôt github qui propose un correctif :

<https://gist.github.com/guebe/e074b165ab6fd291b1fa3f424c18a893>

2 Ariane 5

Le vol inaugural 501 a eu lieu le 4 juin 1996 à Kourou (Guyanne).

Le lanceur fut détruit au bout de 37 secondes de vol. L'échec était dû à une erreur informatique. Le coût estimé des pertes s'élève à 500 millions de dollars.

L'enquête a révélé que lors du vol une exception s'est produite. Cette exception est due à la conversion d'un nombre décimal codé sur 64 bits et représentant l'accélération horizontale du lanceur en un nombre entier signé codé sur 16 bits.

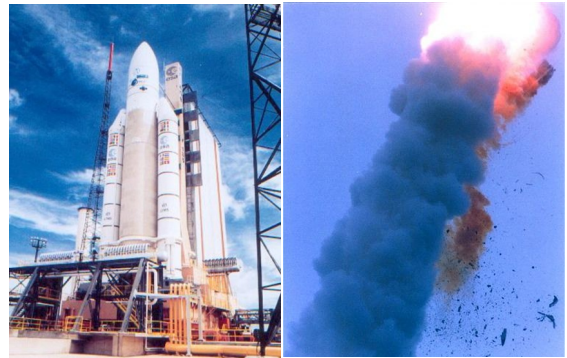


Figure 2 : Ariane 5

Cette exception a provoqué la panne du système de navigation puis celui du système de secours. Le pilote automatique a alors ordonné une violente correction de trajectoire. La fusée a alors dévié brutalement de sa trajectoire ce qui a déclenché une autodestruction préventive de la fusée.



Les nombres entiers signés sont codés :

- en binaire naturel pour les nombres positifs ;
- en complément à 2 pour les nombres négatifs.

Cela signifie par exemple qu'un octet représente en binaire signé les nombres compris entre -128 et +127.

Q1. Que vaut le MSB en binaire signé d'un entier positif codé sur 8 bits. Même question pour un entier négatif codé sur 8 bits.

Entiers naturels	Binaire				Entiers relatifs
16	1	0	0	0	
15		1	1	1	-1
14		1	1	1	-2
13		1	1	0	-3
12		1	1	0	-4
11		1	0	1	-5
10		1	0	1	-6
9		1	0	0	-7
8		1	0	0	-8
7		0	1	1	7
6		0	1	1	6
5		0	1	0	5
4		0	1	0	4
3		0	0	1	3
2		0	0	1	2
1		0	0	0	1
0		0	0	0	0

Figure 3 : binaire signé sur 4 bits

Q2. Vérifier en binaire signé que l'addition de 5 et de -5 donne bien zéro sur 4 bits.

Q3. Quelle valeur positive a dépassé le nombre décimal codé sur 64 bits pour que cela déclenche une exception lors de sa conversion ?

3 Affichage bizarre de page html (« arrête ton charset »)

Avez-vous jamais remarqué des pages html qui affichent les accents de manière bizarre ?



Figure 4 : mauvais rendu d'une page html

Cela est d      un mauvais d  codage du texte par le navigateur.

Dans l'exemple pr  c  dent la page est encod  e en UTF-8 or le navigateur pense traiter de l'ISO 8859-1 parce que le jeu de caract  res (charset en anglais) est mal sp  cifi   dans la page html.

Voici un extrait de chaque table de caract  re (source Wikip  dia):

ISO/CEI 8859-1																
	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9	xA	xB	xC	xD	xE	xF
0x	positions inutilis��es															
1x																
2x	SP	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3x	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4x	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5x	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6x	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7x	p	q	r	s	t	u	v	w	x	y	z	{		}	~	
8x	positions inutilis��es															
9x																
Ax	NBSP	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��
Bx	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��
Cx	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��
Dx	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��
Ex	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��
Fx	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��	��

Figure 5 : table ISO 8859-1

Des bugs riches d'enseignement

Type	Caractère	Point de code (hexadécimal)	Valeur scalaire		Codage UTF-8		
			décimal	binaire	binaire		hexadécimal
Contrôle (C0)	[NUL]	U+0000	0	0	00000000		00
	[US]	U+001F	31	1·1111	00011111		1F
Texte (US-ASCII)	[SP]	U+0020	32	10·0000	00100000		20
	0	U+0030	48	11·0000	00110000		30
	9	U+0039	57	11·1001	00111001		39
	?	U+003F	63	11·1111	00111111		3F
	@	U+0040	64	100·0000	01000000		40
	A	U+0041	65	100·0001	01000001		41
	Z	U+005A	90	101·1010	01011010		5A
	a	U+0061	97	110·0001	01100001		61
	z	U+007A	122	111·1010	01111010		7A
	~	U+007E	126	111·1110	01111110		7E
Contrôle (C0 et C1)	[DEL]	U+007F	127	111·1111	01111111		7F
	[PAD]	U+0080	128	1000·0000	11000010 10000000		C2 80
	[APC]	U+009F	159	1001·1111	11000010 10011111		C2 9F
Texte (PMB)	[NBSP]	U+00A0	160	1010·0000	11000010 10100000		C2 A0
	ı	U+00BF	191	1011·1111	11000010 10111111		C2 BF
	À	U+00C0	192	1100·0000	11000011 10000000		C3 80
	é	U+00E9	233	1110·1001	11000011 10101001		C3 A9
	ı	U+07FF	2047	111 1111·1111	11011111 10111111		DF BF
	ı	U+0800	2048	1000 0000·0000	11100000 10100000 10000000		E0 A0 80
	€	U+20AC	8 364	10·0000 1010·1100	11100010 10000010 10101100		E2 82 AC
	ı	U+D7FF	55 295	1101·0111 1111·1111	11101101 10011111 10111111		ED 9F BF

Figure 6 : table UTF-8

Q1. À partir des tables fournies expliquer l'affichage erroné du mot « Actualités » (indiquez il faut partir de la table UTF-8 dans un 1^{er} temps) :

Q2. Donnez la valeur hexadécimale des caractères 'è' (e accent grave) et 'ï' (i tréma) codés en UTF-8 (indiquez il faut partir de la table ISO 8859-1 cette fois) :

De nos jours 95% des sites web sont encodés en UTF-8 et les erreurs d'affichage des pages html sont de plus en plus rares.

Pour convertir des caractères en UTF-8 :

<https://www.cogsci.ed.ac.uk/~richard/utf-8.cgi?input=C3+A8&mode=bytes>