

Characters in Elsevier Bibliographic databases output

Version: 1.5

Created: 16/08/2007

Changed: 14/09/2007

Introduction

The Elsevier XML files use the Unicode character set, version 4.1.0. Only a subset of the full Unicode collection is used: the characters to be expected are listed in the table below. All non-ASCII characters are encoded as explicit numeric character references, for example, as `©` or `©`.

Internally, the bibliographic database uses a limited character set which corresponds to the characters in the table below. One may expect to find many Unicode characters, the ones that do not have a representative in this limited character set, to have been input as an ASCII character with some descriptive text enclosed in braces. This is not done consistently throughout all the data, however, and never in author's names. For instance, what was originally a dotless i in an abstract may be represented in XML output as `i{dotless}` in some items, and as `i` in some other items.

Table columns

- Unicode – the Unicode using the hexadecimal notation
- Decimal - the corresponding character using the decimal position.
- OPSBK – internal Elsevier id
- Character - picture of the character, included where available (represented by a if the character is not supported by this version of MSWord).
- W3c/MathML entity – The corresponding W3c/MathML entity
- Description - verbal description of the character

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
20	 	001			space
21	!	002	!		exclamation mark
22	"	003	"	"	quotation mark
23	#	004	#		number sign
24	$	005	\$	$	dollar sign (EW=Bhc)
25	%	006	%		percent sign
26	&	007	&	&	ampersand (EW=Bha)
27	'	008	'		apostrophe
28	(009	(left parenthesis
29)	010)		right parenthesis
2a	*	011	*	&z.ast;	asterisk
2b	+	012	+		plus sign
2c	,	013	,		comma
2d	-	014	-		hyphen-minus
2e	.	015	.		full stop
2f	/	016	/		solidus (EW=Pf1)
30	0	017	0		digit zero

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
31	1	018	1		digit one
32	2	019	2		digit two
33	3	020	3		digit three
34	4	021	4		digit four
35	5	022	5		digit five
36	6	023	6		digit six
37	7	024	7		digit seven
38	8	025	8		digit eight
39	9	026	9		digit nine
3a	:	027	:		colon
3b	;	028	;		semicolon
3c	<	029	<	<	less-than sign
3d	=	030	=		equals sign
3e	>	031	>	>	greater-than sign
3f	?	032	?		question mark
40	@	033	@		commercial at
41	A	034	A		latin A
42	B	035	B		latin B
43	C	036	C		latin C
44	D	037	D		latin D
45	E	038	E		latin E
46	F	039	F		latin F
47	G	040	G		latin G
48	H	041	H		latin H
49	I	042	I		latin I
4a	J	043	J		latin J
4b	K	044	K		latin K
4c	L	045	L		latin L (EW=Pfl)
4d	M	046	M		latin M (EW=Pem)
4e	N	047	N		latin N
4f	O	048	O		latin O
50	P	049	P		latin P
51	Q	050	Q		latin Q
52	R	051	R		latin R
53	S	052	S		latin S
54	T	053	T		latin T
55	U	054	U		latin U
56	V	055	V		latin V
57	W	056	W		latin W
58	X	057	X		latin X
59	Y	058	Y		latin Y
5a	Z	059	Z		latin Z
5b	[060	[left square bracket
5c	\	061	\	\	reversed solidus (EW=Pll)
5d]	062]		right square bracket
5e	^	063	^		circumflex accent (spacing character)
5f	_	064	_		low line (spacing character)
60	`	065	`		grave accent (spacing character)
61	a	066	a		latin a (EW=Paa)
62	b	067	b		latin b (EW=Pab)
63	c	068	c		latin c (EW=Pac)
64	d	069	d		latin d (EW=Pad)
65	e	070	e		latin e (EW=Pae)
66	f	071	f		latin f (EW=Paf)
67	g	072	g		latin g
68	h	073	h		latin h (EW=Pah)
69	i	074	i		latin i (EW=Pai)
6a	j	075	j		latin j (EW=Paaj)
6b	k	076	k		latin k (EW=Pak)
6c	l	077	l		latin l (EW=Pal)
6d	m	078	m		latin m (EW=Pam)
6e	n	079	n		latin n (EW=Pan)
6f	o	080	o		latin o (EW=Pao)
70	p	081	p		latin p (EW=Pap)
71	q	082	q		latin q (EW=Paq)
72	r	083	r		latin r (EW=Par)
73	s	084	s		latin s (EW=Pas)
74	t	085	t		latin t (EW=Pat)
75	u	086	u		latin u (EW=Pau)
76	v	087	v		latin v (EW=Pav)

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
77	w	088	w		latin w (EW=Paw)
78	x	089	x		latin x (EW=Pax)
79	y	090	Y		latin y (EW=Pay)
7a	z	091	Z		latin z (EW=Paz)
7b	{	092	{		left curly bracket
7c	|	093			vertical bar
7d	}	094	}		right curly bracket
7e	~	095	~		tilde (spacing character)
a0	 	096		 	non-breaking space
a1	¡	097	¡	¡	inverted exclamation mark
a2	¢	098	¢	¢	cent sign
a3	£	099	£	£	pound sign
a5	¥	100	¥	¥	yen sign
a7	§	101	§	§	section sign
a9	©	102	©	©	copyright sign
ab	«	103	«	«	left double angle quote
ac	¬	296	¬	¬	not sign
ae	®	104	®	®	registered sign
b0	°	105	°	°	degree sign consistent with º
b1	±	106	±	±	plus-minus sign
b6	¶	107	¶	¶	pilcrow (paragraph) sign
b7	·	108	·	·	middle dot
Bb	»	109	»	»	right double angle quotation mark
Bf	¿	110	¿	?	inverted question mark
c6	Æ	111	Æ	Æ	latin AE
d7	×	112	×	×	Multiplication sign
d8	Ø	113	Ø	Ø	latin O with stroke
df	ß	114	ß	ß	latin sharp S
e6	æ	115	Æ	æ	latin ae
f7	÷	116	÷	÷	division sign
f8	ø	117	Ø	ø	latin o with stroke
l41	Ł	276	Ł	Ł	latin L with stroke
l42	ł	277	ł	ł	latin l with stroke
l52	Œ	278	Œ	Œ	latin ligature OE
l53	œ	279	œ	œ	latin ligature oe
300	̀	286	˘	`	combining grave accent
301	́	287	˙	´	combining acute accent
302	̂	288	ˆ	ˆ	combining circumflex accent
303	̃	289	˜	&tild;	combining tilde
304	̄	290	ˆ	¯	combining macron
307	̇	291	˙	˙	combining dot above
308	̈	292	¨	¨	combining diaeresis
309	̉	□	◌		combining hook above
30a	̊	293	◌	˚	combining ring above
30c	̌	294	◌	ˇ	combining caron
327	̧	295	◌	¸	combining cedilla
387	·	108	·		Greek Ano Teleia (00B7 preferred character)
391	Α	143	Α	Α	greek Alpha
392	Β	144	Β	Β	greek Beta
393	Γ	145	Γ	Γ	greek Gamma
394	Δ	146	Δ	Δ	greek Delta
395	Ε	147	Ε	&Epsi;	greek Epsilon
396	Ζ	148	Ζ	Ζ	greek Zeta
397	Η	149	Η	Η	greek Eta
398	Θ	150	Θ	Θ	greek Theta
399	Ι	151	Ι	Ι	greek Iota
39a	Κ	152	Κ	Κ	greek Kappa
39b	Λ	153	Λ	Λ	greek Lambda
39c	Μ	154	Μ	Μ	greek Mu
39d	Ν	155	Ν	Ν	greek Nu
39e	Ξ	156	Ξ	Ξ	greek Xi
39f	Ο	157	Ο	Ο	greek Omicron
3a0	Π	158	Π	Π	greek Pi
3a1	Ρ	159	Ρ	Ρ	greek Rho
3a3	Σ	160	Σ	Σ	greek Sigma
3a4	Τ	161	Τ	Τ	greek Tau
3a5	Υ	162	Υ	Υ	greek Upsilon
3a6	Φ	163	Φ	Φ	greek Phi
3a7	Χ	164	Χ	Χ	greek Chi

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
3a8	Ψ	165	Ψ	Ψ	greek Psi
3a9	Ω	166	Ω	Ω	greek Omega
3b1	α	118	α	α	greek alpha
3b2	β	119	β	β	greek beta
3b3	γ	120	γ	γ	greek gamma
3b4	δ	121	δ	δ	greek delta
3b5	ε	122	ε	ε	greek epsilon
3b6	ζ	123	ζ	ζ	greek zeta
3b7	η	124	η	η	greek eta
3b8	θ	125	θ	θ	greek theta
3b9	ι	126	ι	ι	greek iota
3ba	κ	127	κ	κ	greek kappa
3bb	λ	128	λ	λ	greek lambda
3bc	μ	129	μ	μ	greek mu
3bd	ν	130	ν	ν	greek nu
3be	ξ	131	ξ	ξ	greek xi
3bf	ο	132	ο	ο	greek omicron
3c0	π	133	π	π	greek pi
3c1	ρ	134	ρ	ρ	greek rho
3c2	ς	135	ς	ς	greek final sigma
3c3	σ	136	σ	σ	greek sigma
3c4	τ	137	τ	τ	greek tau
3c5	υ	138	υ	υ	greek upsilon
3c6	φ	139	φ	φ	greek phi
3c7	χ	140	χ	χ	greek chi
3c8	ψ	141	ψ	ψ	greek psi
3c9	ω	142	ω	ω	greek omega
5d0	א	275	א	ℵ	hebrew aleph
2020	†	167	†	†	dagger
2021	‡	168	‡	‡	double dagger
2022	•	169	•	•	bullet
2030	‰	170	‰	‰	per mille sign
2032	′	171	'	′	prime
2033	″	172	''	″	double prime
2034	‴	173	'''	‴	triple prime
20ac	€	280	€	€	euro sign
2102	ℂ	322	Ⓒ	&Cdbl;	double-struck capital C
210b	ℋ	332	Ⓗ	ℋ	script capital H
210d	ℍ	323	Ⓓ	&Hdbl;	double-struck capital H
210f	ℏ	282	Ⓗ	ℏ	planck constant over two pi
2110	ℐ	333	Ⓘ	ℐ	script capital I
2111	ℑ	283	Ⓙ	𝔍	black-letter capital J
2112	ℒ	334	Ⓛ	ℒ	script capital L
2113	ℓ	284	ℓ	ℓ	script small l
2115	ℕ	324	Ⓝ	&Ndbl;	double-struck capital N
2118	℘	336	Ⓟ	𝒫	script capital P
2119	ℙ	325	Ⓟ	&Pdbl;	double-struck capital P
211a	ℚ	326	Ⓠ	&Qdbl;	double-struck capital Q
211b	ℛ	337	Ⓡ	ℛ	script capital R
211c	ℜ	285	Ⓡ	ℜ	black-letter capital R
211d	ℝ	327	Ⓡ	&Rdbl;	double-struck capital R
2122	™	281	™	™	trade mark sign
2124	ℤ	328	Ⓣ	&Zdbl;	double-struck capital Z
2127	℧	321	Ω	℧	inverted ohm sign
212c	ℬ	329	Ⓑ	ℬ	script capital B
2130	ℰ	330	Ⓔ	ℰ	script capital E
2131	ℱ	331	Ⓕ	ℱ	script capital F
2133	ℳ	335	Ⓜ	ℳ	script capital M
2190	←	176	←	←	leftwards arrow
2191	↑	177	↑	↑	upwards arrow
2192	→	178	→	→	rightwards arrow
2193	↓	179	↓	↓	downwards arrow
2194	↔	180	↔	↔	left right arrow
2195	↕	303	↕	↕	up down arrow
2196	↖	181	↖	↖	north west arrow
2197	↗	182	↗	↗	north east arrow
2198	↘	183	↘	↘	south east arrow
2199	↙	184	↙	↙	south west arrow
219e	↞	302	↔	↞	leftwards two headed arrow
21a0	↠	301	↔	↠	rightwards two headed arrow

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
21bc	↼	299	☐	↼	leftwards harpoon with barb upwards
21bd	↽	300	☐	↽	leftwards harpoon with barb downwards
21c0	⇀	297	☐	⇀	rightwards harpoon with barb upwars
21c1	⇁	298	☐	⇁	rightwards harpoon with barb downwards
21c4	⇄	185	☐	&rlarr2;	right over left arrow N.B. rlarr in ES grid
21c6	⇆	186	☐	&lrarr2;	left over right arrow N.B. lrarr in ES grid
21cb	⇋	187	☐	&lrhar2;	left over right harpoon N.B. lrhar in ES grid
21cc	⇌	188	☐	&rlhar2;	right over left harpoon N.B. rlhar in ES grid
21d0	⇐	189	☐	⇐	leftwards double arrow
21d2	⇒	190	☐	⇒	rightwards double arrow
21d4	⇔	191	☐	⇔	left right double arrow
2200	∀	226	☐	∀	for all
2202	∂	227	∂	∂	partial differential
2203	∃	228	☐	∃	there exists
2204	∄	229	☐	∄	there does not exist
2205	∅	220	☐	∅	empty set
2207	∇	230	☐	∇	nabla
2208	∈	231	☐	∈	element of
2209	∉	232	☐	∉	not an element of
220b	∋	233	☐	∋	contains as member
220c	∌	234	☐	∌	does not contain as member
220f	∏	235	∏	∏	n-ary product
2210	∐	236	☐	∐	n-ary coproduct
2211	∑	237	∑	∑	n-ary summation
2213	∓	192	±	∓	minus-or-plus sign
221a	√	221	√	√	square root
221d	∝	238	☐	∝	proportional to
221e	∞	193	∞	∞	infinity
2220	∠	239	☐	∠	angle
2225	∥	240	☐	∥	parallel to
2226	∦	241	☐	∦	not parallel to
2227	∧	242	☐	∧	logical and
2228	∨	243	☐	∨	logical or
2229	∩	244	∩	∩	intersection
222a	∪	245	☐	∪	union
222b	∫	218	∫	∫	integral
222e	∮	219	☐	∮	contour integral
2234	∴	254	☐	∴	therefore
223c	∼	194	☐	∼	tilde operator
2241	≁	195	☐	≁	not tilde
2243	≃	196	☐	≃	asymptotically equal to
2244	≄	197	☐	≄	not asymptotically equal to
2245	≅	198	☐	≅	approximately equal to
2247	≇	199	☐	≇	neither aproximately nor actually equal to
2248	≈	200	≈	≈	almost equal to
2249	≉	201	☐	≉	not almost equal to
224a	≊	202	☐	≊	almost equal or equal to
2252	≒	203	☐	≒	approximately equal to or the image of
2253	≓	204	☐	≓	image of or approximately equal to
2259	≙	312	☐	&wedgedq;	estimates
2260	≠	205	≠	≠	not equal to
2261	≡	206	≡	≡	identical to
2262	≢	207	☐	≢	not identical to
2264	≤	208	≤	≤	less-than or equal to
2265	≥	209	≥	≥	greater-than or equal to
2266	≦	304	☐	≦	less-than over equal to
2267	≧	305	☐	≧	greater-than over equal to
226a	≪	210	☐	≪	much less-than
226b	≫	211	☐	≫	much greater-than
226e	≮	212	☐	≮	not less-than
226f	≯	213	☐	≯	not greater-than
2270	≰	214	☐	≰	neither less-than nor equal to
2271	≱	215	☐	≱	neither greater-than nor equal to
2272	≲	216	☐	≲	les-than or equivalent to
2273	≳	217	☐	≳	greater-than or equivalent to

Uni code	decimal	OPSBK	char	w3/MathML entity	Description
227a	≺	310	□	≺	precedes
227b	≻	311	□	≻	succeeds
2282	⊂	246	□	⊂	subset of
2283	⊃	247	□	⊃	superset of
2284	⊄	248	□	⊅	not a subset of
2285	⊅	249	□	⊅	not a superset of
2286	⊆	250	□	⊆	subset of or equal to
2287	⊇	251	□	⊇	superset of or equal to
2288	⊈	252	□	⊈	neither a subset of nor equal to
2289	⊉	253	□	⊉	neither a superset of nor equal to
2295	⊕	222	□	⊕	circled plus
2296	⊖	223	□	⊖	circled minus
2297	⊗	224	□	⊗	circled times
2299	⊙	225	□	⊙	circled dot operator
22a2	⊢	256	□	⊢	right tack
22a3	⊣	257	□	⊣	left tack
22a4	⊤	258	□	⊤	down tack
22a5	⊥	259	□	⊥	up tack
22b4	⊴	313	□	⊴	normal subgroup of or equal to
22b5	⊵	314	□	⊵	contains as normal subgroup or equal to
22c9	⋉	320	□	⋉	left normal factor semidirect product
22ca	⋊	319	□	⋊	right normal factor semidirect product
22d8	⋘	308	□	⋘	very much less-than
22d9	⋙	309	□	⋙	very much greater-than
22ee	⋮	306	□	⋮	vertical ellipsis
22ef	⋯	307	□	&mellip;	midline horizontal ellipsis
2308	⌈	317	□	⌈	left ceiling
2309	⌉	318	□	⌉	right ceiling
230a	⌊	315	□	⌊	left floor
230b	⌋	316	□	⌋	right floor
25a0	■	260	■	&sqf;	black square
25a1	□	261	□	□	white square
25b2	▲	264	▲	▴	black up-pointing triangle
25b3	△	265	□	▵	white up-pointing triangle
25b6	▶	268	□	▸	black right-pointing triangle
25b7	▷	269	□	▹	white right-pointing triangle
25bc	▼	266	▼	▾	black down-pointing triangle
25bd	▽	267	□	▿	white down-pointing triangle
25c0	◀	270	□	◂	black left-pointing triangle
25c1	◁	271	□	◃	white left-pointing triangle
25c6	◆	262	□	♦	black diamond
25c7	◇	263	□	⋄	white diamond
25cb	○	272	○	○	white circle
2640	♀	274	♀	♀	female sign
2642	♂	273	♂	♂	male sign
2aa1	⪡	210	□	⪡	much less than (double) (EW=Bkl)
2aa2	⪢	211	□	⪢	much greater than (double) (EW=Bml)
3008	〈	174	<	⟨	left angle bracket
3009	〉	175	>	⟩	right angle bracket
30fb	・	108	·		Katakana middle dot
1d4c1	𝓁	333	□		script capital I