Floating Point Representation

Question 1

- 1 In a particular computer system, real numbers are stored using floating-point representation with:
 - 10 bits for the mantissa
 - 6 bits for the exponent
 - two's complement form for both mantissa and exponent.
 - (a) Calculate the normalised floating-point representation of +192.5 in this system. Show your working.

				N	Mant	issa	l							Ехр	onen	ıt			
١	Norkin	g																	
																			[-]
(b)			he no	orma	alise	d floa	ating	-poin	it rep	rese	ntati	on o	f –19	92.5	in th	is sy	stem	. Sh	ow your
(b)	Calcu workir		he no	orma				-poin	it rep	rese	ntati	on o	f –19				stem	. Sh	ow your
(b)			he no	orma		d floa		-poin	t rep	rese	ntati	on o	f –19		one		rstem	. Sh	ow your
(b)			he no	orma				-poin	t rep	prese	ntati	on o	f –19				rstem	. Sh	ow your
(b)	workir	ng.			Mar	ntiss	a							Exp	one	nt			
(b)		ng.			Mar	ntiss	a							Exp	one	nt			
(b)	workir	ng.			Mar	ntiss	a							Ехр	one	nt			
(b)	workir	ng.			Mar	ntiss	a							Ехр	one	nt			

	(c)						prese as s			nas c	hang	ed. T	There	e are	now	12 k	oits fo	r the	mar	itissa a	and 4
								Man	tissa	a							Exp	onen	t		
										•			•		_					,	
		E	xplaiı	n wh	y +19	92.5	cann	ot be	acc	curate	ely re	pres	ente	d in t	his fo	orma	ıt.				
																		•••••			•••••
																					[3]
_			_																		
Qu	esti	on	2																		
1	In a							n, rea	al nu	ımbe	rs ar	e sto	red ι	using	floa	ting-	point	repr	esen	tation	with:
	•	4 bi	its fo	r the	e ma expo emer	onen	t	r botl	n ma	antiss	sa an	d ex	pone	ent.							
	(a)	The	e follo	owing	g floa	ating	-poin	t nur	nber	r stor	ed is	not	norn	nalise	ed.						
		Cal	culat	te the	e der	nary	value	e for	the f	floati	ng-po	oint r	numb	er. S	how	you	r wor	king.			
							N	/lanti	issa								Ехро	onen	t		
			0	0	0	0	1	1	0	0	0	0	0	0		0	1	0	1		
																				J	
		Wo	rking	l																	
		Der	nary	value	э																
																					[3]

	Wı	rite yo	ur ar	iswe	r in th	ne foll	lowing	g box	es.									
					M	antis	sa						ı	Expo	nent			
]						
												_						[2]
(i	i) De	scribe	e one	e pro	blem	that	can o	ccur	when	float	tina-r	oint i	numb	ers a	ire no	ot nor	malis	
,-																		
																		[2]
stic	n 3																	
n o r	orticu	lor oc	mnı	itor c	veto	m ro	al ni	ımbo	re or	o etc	rod i	ucina	floo	ting	ooint	ropr	ocon	tation with:
						iii, ie	aint	шье	is ai	e sic	neu i	using	IIUa	ung-	politic	repri	esen	iation, with.
						or bo	th ma	antis	sa an	nd ex	pone	ent.						
(a) (Calcul	ate th	ne de	enary	valu	ue for	the	follov	wing	floati	ing-p	oint ı	numl	oer. S	Show	your	wor	king.
						Man	tissa	ı							Expo	non	•	
															_,,,	men		
																men		
	0	1	0	1	0	0	0	0	0	0	0	0		0	1	1	0	
	0	1	0	1	0	0	0	0	0	0	0	0			1			
١	0 Workir				<u> </u>			<u> </u>			<u> </u>	<u> </u>		0				
١		ng												0		1	0	
		ng												0		1	0	
		ng												0		1	0	
		ng												0		1	0	
	Workir	ng												0		1	0	
	Workir	ng												0		1	0	
	s tic	n a particut 12 bits 4 bits two's c	n a particular co	n a particular compute 12 bits for the expense two's compleme	n a particular computer s 12 bits for the mantis 4 bits for the expone two's complement fo	(ii) Describe one problem stion 3 n a particular computer syste 12 bits for the mantissa 4 bits for the exponent two's complement form form	(ii) Describe one problem that a stion 3 In a particular computer system, research 12 bits for the mantissa 4 bits for the exponent two's complement form for bota) Calculate the denary value for	n a particular computer system, real number of the mantissa 4 bits for the exponent two's complement form for both mata). Calculate the denary value for the	(ii) Describe one problem that can occur stion 3 n a particular computer system, real number 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantiss	(ii) Describe one problem that can occur when	(ii) Describe one problem that can occur when float stion 3 n a particular computer system, real numbers are sto 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and ex a) Calculate the denary value for the following floating	(ii) Describe one problem that can occur when floating-passion 3 In a particular computer system, real numbers are stored at 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent a) Calculate the denary value for the following floating-passions.	(ii) Describe one problem that can occur when floating-point residual and particular computer system, real numbers are stored using 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. a) Calculate the denary value for the following floating-point residuals.	(ii) Describe one problem that can occur when floating-point numbers are stored using floating floating for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. a) Calculate the denary value for the following floating-point numbers are stored.	(ii) Describe one problem that can occur when floating-point numbers a strion 3 n a particular computer system, real numbers are stored using floating- 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. a) Calculate the denary value for the following floating-point number. See the strict of the st	(ii) Describe one problem that can occur when floating-point numbers are not stion 3 In a particular computer system, real numbers are stored using floating-point 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. (a) Calculate the denary value for the following floating-point number. Show	(ii) Describe one problem that can occur when floating-point numbers are not nor stion 3 In a particular computer system, real numbers are stored using floating-point representations for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. (a) Calculate the denary value for the following floating-point number. Show your	(ii) Describe one problem that can occur when floating-point numbers are not normalis stion 3 n a particular computer system, real numbers are stored using floating-point represent 12 bits for the mantissa 4 bits for the exponent two's complement form for both mantissa and exponent. a) Calculate the denary value for the following floating-point number. Show your work

(b) (i) Normalise the floating-point number given in part (a).

	(b)			pera are															floa	ting-po	oint
		(i)		culate oatter												t nur	nber	that	has	the sa	me
				Ехро	onen	t						ا	Man	tissa	l						
			0	1	0	1		0	0	0	0	0	0	0	0	0	1	1	0		
			Wor	king																	
			Den	ary v	alue	•••••															[3]
	(ii)			two point			that	can	occ	ur d	ue t	o the	e ch	ange	in i	the i	epre	sen	tatio	n of th	те
		Prol	olem	1																	
		Prol	olem	2																	
																					 [2]
Qu	est [.]	ion	4																		
1				comp	uter	syste	m, re	eal nu	ımbe	rs ar	e sto	red ι	using	float	ting-p	ooint	repre	esen	tatio	n with:	
	•	4 bits	s for t	the race	cpone	ent	or bo	th ma	antiss	sa an	ıd ex	pone	nt.								
	(a)			ving fl										ed.							
		Calc	ulate	the d	enar	y valu	ue foi	the	floati	ng-p	oint r	numb	er. S	how	your	wor	king.				
							Man	tissa	l						ı	Expo	nent	t			
		0	0	0	0	1	1	0	0	0	0	0	0		0	1	0	1			
		Work	ing .																		
		Dena	ary va	alue .																[3]]

			Write	you	r ans	wer ii	n the	follo	wing	boxe	es.											
							Mar	ntiss	a							E	Ξхр	one	nt			
			_																			[2]
		(ii)	Desc	cribe	one	proble	em tr	nat ca	an oc	cur v	vhen	float	ing-p	oint	nu	mb	ers	are	not r	norm	alise	d.
																						[2]
Qu	est	ion	5																			
	ln s		nuto	r 0.40	tom	rool	nun	abore	oro	oto	od 1	ıcina	nor	mali	000	1 fl	ooti	na r	oint	ropr	000	atation
1	with		ipute	rsys	item,	rear	nun	ibers	are	Stor	ea t	ising	HOI	Шан	sec	. III	oau	ng-p	OITIL	repr	ese	ntation
	•					manti ooner																
	The	man	tissa	and	expo	nent	are l	ooth	in tw	o's c	omp	leme	nt fo	rm.								
	(a)	Calc	ulate	the	dena	ry va	lue f	or th	e foll	owin	g bin	ary f	loati	ng-p	oin	t n	um	oer.				
		Sho	w you	ır wo	rking	J.																
						Mant	issa	ı											Ехр	oner	nt	
	1	0	0	1	0	1	1	1	0	0	1	1						0	1	1	1	
		Wor	king .																			
		Ansı	wer																			[3]

(b) (i) Normalise the floating-point number given in part (a).

(l	o)	Cal	culat	e the	norm	alise	d flo	ating	-poir	nt rep	rese	ntatio	on of +1.5625 in	this s	ystem.		
		Sho	ow yo	ur wo	orking	J.											
		Wo	rking														
						Man	tissa	1							Expon	ent	
Γ																	7
L																	」 [3]
(c)	(i)	Write	e the	large	est n	ositi	ve n	umbe	er th	at c	an be	e stored as a i	norma	lised fl	oating-	noint
(0)	,	(-)		ber u												oamig	Politic
					ľ	Mant	issa							E	Expone	ent	
																	[2]
	(ii)	Write	e the	sma	allest	non	ı-zer	о ро	sitive	e nu	mber	that can be	stored	as a	norma	alised
			float	ing-po	oint n	umb	er us	ing t	his fo	orma	t.						
					ľ	Mant	issa							E	Expone	ent	
																	[2]
(d)	The	e dev	/elope	er of	a ne	w pr	ogra	mmir	ng la	ngua	iae d	ecides that all r	eal nu	ımbers	will no	w be
•	•	sto	red u	sing	20-bit	t nor	nalis	ed fl	oatin	g-po	int re	prese	entation. She maxponent.	ust de	cide ho	w man	y bits
													umber of bits fo	r tha r	nantice	a ora	large
				of bit					oning e	71U1G1	a ia	ige ii	uniber of bits to	1 1116 1	Haritiss	a, or a	large
																	ا

	· •	eight bits for the mantissa four bits for the exponent.	
	The		
		mantissa and exponent are both stored in	two's complement format.
	(i)	Calculate the denary value of the following	floating-point number.
		Show your working.	
		Mantissa	Exponent
		0 0 1 1 0 1 1 1	0 1 0 1
		Working	
		Answer	LO.
			[3]
(ii)	Stat	e why the floating-point number in part	(a)(i) is not normalised.
iii)			in normalised two's complement format.
,		Mantissa	Exponent

[2]

(b)	(i)	Convert the dena	ary number +11.6	25 into a normalised fl	loating-point numb	er.
		Show your worki	ng.			
		Working				
		Ma	antissa		Exponent	
						[3]
	(i	i) Convert the de	enary number -11.	.625 into a normalised	floating-point numb	oer.
	•	Show your wo			3 .	
		ı	Mantissa		Exponent	
						[3]

	(c)	A stud	dent e	enter	s the	follo	owing	g into	an	inter	prete	er:									
		C	UTPU	JT (C).2	* 0.	.4)														
		The s	tuder	nt is	surpr	ised	to s	ee th	at th	e in	erpr	eter o	utpu	ts the	follo	owing	g:				
		C	.080	0000	0000	0000	0000	2													
		Expla	in wh	y the	e inte	rpre	ter o	utpu	ts thi	is va	lue.										
																					[3
Que	In a with	twelv four mant Calc	puter ve bits bits fo	s for or the and the	the ree expo	manti ooner nent ry va	issa nt. are	both	in tw	/o's (comp	oleme	nt fo	rm.				oint	repre	esent	ation
					١	Mani	tissa	1									ı	Expo	onen	t	
	1	0	0	1	0	1	1	1	0	0	1	1					0	1	1	1	
		Work	king																		
		Ansv	ver																		 [3

(b)	C	alcı	ulate	the r	norm	alised	d floa	ating	poin	t rep	res	sei	ntati	or	of +1.5625 in t	his s	ystem.		
	S	hov	v you	r woı	rking														
	٧	Vork	ing																
															•••••				
						Mant	issa										Expone	ent	
														_					[3]
(c)		(i)	Write	e the	e lar	gest	posit	ive r	numl	oer t	ha	t c	an I	be	e stored as a r	orm	alised fl	oating-	point
			num	ber ι	using	this 1	forma	at.											
						Man	tissa	ı									Expone	ent	
														_					[2]
	(ii)											umbe	er	that can be s	store	d as a	norma	alised
			floati	ing-p	oint	numb	er u	sing	this	form	at.								
						Man	tissa	1									Expone	ent	
														_					[2]
(d)	Т	-hΔ	dovo	lonei	r of	a new	v pro	orar	nmin	n la	naı	II 2	ne d	אםו	cides that all re	al n	umhars	will no	w ha
(u)	S	tore	d usi	ng 2	0-bit	norm	alise	ed flo	atin	g-po	int	re	pres	er	ntation. She mu				
	Ţ	o us	e tor	tne r	manı	issa a	and r	10W I	man	y Dits	S TC	or t	ne e	exp	ponent.				
						off bet the ex			ng e	ithe	a	lar	rge r	าน	mber of bits for	the	mantissa	a, or a	large
														•••					
																			[3]

1	Real r	numbers are stored using floating-point representation in a computer system.	
	This r	epresentation uses:	
		bits for the mantissa, followed by bits for the exponent.	
	Two's	complement form is used for both the mantissa and the exponent.	
	(a) (i) A real number is stored as a 12-bit normalised binary number as follows:	
		Mantissa Exponent	
		0 1 0 1 0 0 1 0	
		Calculate the denary value for this binary number. Show your working.	
		Working	
		Denary value	
			[3]
	(ii)	Calculate the normalised binary number for -3.75. Show your working.	
	(ii)	Calculate the normalised binary number for -3.75. Show your working. Mantissa Exponent	
	(ii)		
	(ii)		
	(ii)		
	(ii)	Mantissa Exponent	
((Mantissa Exponent	
(I	o) The	Mantissa Exponent Working	
((o) The	Mantissa Exponent Working	
(I	o) The	Mantissa Exponent Working	[3]

	(C)	State	e why so	ome b	inary	repres	sentati	ons c	an lea	d to ro	undin	g erro	rs.				
																	. [1]
	(d)	Com	plete th	e follo	owing	descri	ptions	by in	sertinç	g the t	wo m	issing	terms				
							. can	occur	in the	expon	ent of	a floa	ıting-p	oint n	umbei	r, whe	n the
		expo	nent ha	ıs bec	ome t	oo lar	ge to l	oe rep	resen	ted us	ing th	e num	ber of	bits a	availat	ole.	
		A cal	lculatior	n resu	lts in a	a num	ber so	smal	that i	t cann	ot be	repres	sented	by th	e nun	nber o	f bits
			able. Th											,			
		avaii	abie. Ti	113 13 (Janeu		•••••										[2]
Ωu	est	ion	9														
Qu		1011															
8	(a)	com	followinglemen	t forn	n. The	twelv	e mo	st sigr	nifican	t bits							
		Mos	st														
		sigr	nificant	bit											sig	nifica	_east
		sigr	nificant	bit											sig	nifica	
			nificant	bit 1	1	0	0	0	0	0	0	0	0	1	sig	nifica 0	
		\downarrow		1						0	0	0	0	1			nt bit
		0	1	1 the b		value	of the	expo	nent.						1	0	t bit
		0	1	1 / the k	inary	value	of the	expo	nent.						1	0	t bit
		(i)	1 Identify	1 / the k	inary	value	of the	expo	nent.						1	0	1 [1]
		(i)	1 Identify	the k	oinary oinary oinary	value	of the	expo	nent.						1	0	1 [1]
		(i) (ii)	ldentify	the to the to the total th	pinary Dinary Dinary Er the	value	of the	expo	nent.	ve or n	negativ	/e. Ju	stify yo	our ch	oice.	0	nt bit 1 [1]
		(i) (ii)	Identify State v	the by th	pinary pinary pinary er the	value value numb	of the	expo	nent.	ve or n	negativ	/e. Ju	stify yo	our ch	oice.	0	[1]
		(i) (ii)	Identify Identify State v Positiv	the to the to the total th	pinary pinary pinary er the	value value numb	of the	expo	nent.	ve or n	egativ	/e. Ju	stify yo	our ch	oice.	0	[1]
		(i) (ii)	Identify Identify State v Positiv	the to the to the total th	pinary pinary er the egativ	value value numb	of the	expo	nent.	ve or n	egativ	/e. Ju	stify yo	our ch	oice.	0	[1]

	(iv)	Convert the binary floating-point number in part (a) into denary. Show your working.
		Working
		Denary value
		[3]
	(b)	The number of bits used for the exponent is increased to eight, and the number of bits used for the mantissa is decreased to eight.
		State the effects of this change.
		[2]
Qu	esti	on 10
1	In a with	computer system, real numbers are stored using normalised floating-point representation :
	•	12 bits for the mantissa
	•	4 bits for the exponent Two's complement form for both mantissa and exponent.
	(a)	Find the denary value for the following binary floating-point number.
		Mantissa Exponent
	1	0 1 1 1 0 0 1 1 0 1
		Show your working.
		Working
		Answer
		[3]

	(b)		culate king.	the	nori	malis	ed fl	oatir	ng-po	int r	eprese	entati	on of	5.25	in th	nis s	ysten	n. Sh	ow y	our/
		Wo	king																	
			<u> </u>			Ivian	tissa	I 							ſ	-	=xpo	nent		
																				[3]
	(c)	The	size	of th	e ma	ntiss	a is c	decr	ease	d an	d the s	size	of the	expor	nent	is ind	creas	ed.		
		State	e hov	v this									e num						yste	m car
		repr	esent	i.																
																				[2]
Qu	est	ion	11																	
3	In a with		pute	rsys	stem,	real	num	ber	s are	stoi	ed us	ing	norma	lised-	floati	ng p	oint	repre	sent	ation
	•	4 bit	s for to s for to s com	the e	xpor	ent	for b	oth i	mant	issa	and ex	kpon	ent.							
	(a)	Calc		the	norm	nalise	ed flo	atin	g-poi	nt re	preser	ntatio	n of +	- 21.7	'5 in	this	syste	em. Sl	now	your
		Worl	king .																	
											•••••									
						Man	tissa								Expo	nen	t			

	(b)	Find	the o	dena	ıry va	alue f	or the	e fol	lowin	ng bin	ary fl	oatin	g-poi	int r	numb	er.			
						Man	itissa	a							ı	Ехро	onen	t	
			1	0	1	1	0	0	0	0					1	1	1	0	
		Sho	w you	ur wo	orkin	g.					_			·					
		Wor	king																
		Ans	wer .																
																			[3]
Qu	esti	on	12																
2	(a)		ompu nbers			m sto	res r	eal r	numb	ers u	sing	floatii	ng-po	oint	repr	eser	ntatio	n. Th	e floating-point
		:				the m													
		The	man	tissa	and	l exp	onen	t are	e both	n in tv	vo's c	comp	leme	nt fo	orm.				
		(i)	Calc	culate	e the	den	ary va	alue	of th	e foll	owin	g floa	ting-	poir	nt nu	mbe	r.		
						Mar	ntiss	а								Ехр	oner	nt	
			0	0	1	1	1	0	0	0					0	1	1	1	
			Sho	w yo	ur w	orkin	g.												
			Wor	king															
			Ans	wer															
	(ii) :	State	how	you	ı kno	w the	e floa	ating-	-point	num	ıber i	n pa ı	rt (a	ı)(i)	is no	t nor	mali	[3] sed.
																			[1]

((iii)	Normalise the floating-point number in part (a)(i).	
		Mantissa	Exponent
			[2]
(b)	(i)	Write the largest positive number that this system ca floating-point number in this format.	n represent as a normalised
		Mantissa	Exponent
			[2]
	(ii)	Write the smallest positive number that can be stored number in this format.	as a normalised floating-point
		Mantissa	Exponent
			[2]
(0)	The	ne number of bits available to represent a real number is inc	
(c)		•	
		ate the effect this has on the numbers that can be represented in the:	ed, if the additional four bits are
	(i)	mantissa	
			[1]
	(ii)	exponent	
	()	·	
			[1]
(d)	A st	student enters the following code into an interpreter.	
		X = 0.1 Y = 0.2	
		Z = 0.3	
		OUTPUT (X + Y + Z)	
	The	e student is surprised to see the output:	
		0.600000000000001	
	Exp	plain why this is output.	

.....[3]

1	(a)	A computer system uses floating-point representation to store real numbers. The floating-ponumbers have:	int
		8 bits for the mantissa8 bits for the exponent	
		The mantissa and exponent are both in two's complement form.	
		(i) Calculate the denary value of the following floating-point number. It is not in normalis form.	ed
		Mantissa Exponent	
		0 0 1 0 1 0 1 0 0 0 0 1 0 1	
		Show your working.	
		Working	
		Answer	 [3]
	(ii)	Convert the denary number +7.5 into a normalised floating-point number.	
	()	Show your working.	
		Mantissa Exponent	
	Γ		
	L		
		Working	

[3]

		Show	you	r wor	king.													
				Mar	ntiss	a						E	Expo	nent				
		Work	ing .															
A r	nor	malise	ed flo	ating	g-poi	nt nu	mbe	r is s	hown									
				Man	tissa	a						ı	Ехро	nen	t			
	0	1	1	1	1	1	1	1		0	1	1	1	1	1	1	1	
			1						[1	1	1	1	1	1	1	
(i)		1 State t	1						[/ num		1	1	1	1	1	1	1	
(i)			1						/ num		1	1	1	1	1	1	1]
(i)			1							ber.								
			1 he s	ignifi	canc	e of	this t	pinary		ber.								
		State t	1 he s	ignifi	canc	e of	this t	pinary		ber.								
		State t	1 he s	ignifi	canc	e of	this t	pinary	num	ber.	s add	ded t	to thi	s nur	mber			
		State t	1 he s	ignifi	canc	e of	this t	pinary	num	ber.	s add	ded t	to thi	s nur	mber			
(i) (ii)		State t	1 he s	ignifi	canc	e of	this t	pinary	num	ber.	s add	ded t	to thi	s nur	mber			

				9 P	oint re Mant i		critati	011 01	72.0		o oyo	tom.	Onov		onent	
	•															
				ı												<u> </u>
••••									•••••							
(b) (Calcul	ate th	ne flo	ating-	point r			tion of	-2.5	in th	is sys	stem.	Shov			
					Man	tissa	1					1		Expo	onent	
	•															

(c) Find the denary value for the following binary floating-point number. Show your working. Mantissa **Exponent** 0 0 0 1 1 0 1 1 0 0 0 0 0 0 0 0[3] (d) (i) State whether the floating-point number given in part (c) is normalised or not normalised.[1] (ii) Justify your answer given in part (d)(i).[1] (e) The system changes so that it now allocates 8 bits to both the mantissa and the exponent. State two effects this has on the numbers that can be represented. 1 2

	the expor		or both i	mantiss	sa and	d exp	onen	t				
Calculate	e the floati		nt repre	sentati	on of	+3.5	in thi					work
	Mant	issa			[Expo	nent		
T												
	•••••											
		•••••										
Calculat	te the float	ting-poi	int repre									
Calculat		ting-poi	int repre					is sys	tem.		your	
Calculat			int repre					is sys	tem.	Show	your	
Calculat			int repre					is sys	tem.	Show	your	
•		tissa		esenta	tion of	f –3.5	in th	is sys	tem.	Show	your	worki
•	Man	tissa		esenta	tion of	f –3.5	in th	is sys	tem.	Show	your	worki
•	Man	tissa		esenta	tion of	f –3.5	in th	is sys	tem.	Show	your	worki
•	Man	tissa		esenta	tion of	f –3.5	in th	is sys	Expo	Show	your	worki

				Man	tissa								Expo	nent			
	0 •	1	1	1	0	0	0	0		0	0	0	0	0	1	0	C
						•	•					•					
															•••••		
q)		Sta	ate wh	nethe	r the f	loatin											
d)	(i)	Sta			r the f		•	int nu	mber	given	in p a	art (c)) is no	ormali	sed o	or not	nor
								int nu	mber	given	in p a	art (c)) is no	ormali	sed o	or not	nor
	(i) (ii)							int nu	mber	given	in p a	art (c)) is no	ormali	sed o	or not	noi
			stify y	our a	nswe	r give	n in p	oart (c	nber	given	in p a	art (c)) is no	ormali	sed o	or not	noi
	(ii)	Jus	stify y	our a	nswe	r give	n in p	oart (mber (given	in p a	nrt (c)) is no	ormali	sed o	or not	nor
	(ii)	Jus	stify y	our a	nswe	r give	n in p	oart (mber (given	in p a	nrt (c)) is no	ormali	sed o	or not	nor
	(ii)	Jus	stify y	our a	nswe	r give	n in p	oart (mber (given	in p a	ve nu) is no	ormali	sed o	or not	noi
	(ii)	Jus	stify y	our a	nswe	r give	n in p	oart (mber (given	in p a	ve nu) is no	ormali	sed o	or not	no

- 1 In a particular computer system, real numbers are stored using floating-point representation with:
 - 8 bits for the mantissa, followed by
 - 8 bits for the exponent

Two's complement form is used for both mantissa and exponent.

(a) (i) A real number is stored as the following two bytes:

			Man	ıtissa								Expo	onent			
)	0	1	0	1	0	0	0		0	0	0	0	0	0	1	1
		Calcu	late t	he der	nary v	alue d	of this n	umb	er. Sh	ow yo	our wo	rking.				
	(ii)	Exp	lain w	vhy the	floati	ing-po	oint nun									[3
	(iii)	Nor	malis	e the fl	oatin	g-poir	nt numb	er ir								
	(iii)	Nor		e the fl antissa		g-poir	nt numb	er ir					ponen			
	(iii)	Nor				g-poir	nt numb	er ir								
(b)	(iii)) (i)	Wri	Ma	antissa	est po	ositive	nt numb		part	(a)(i).		Ex	ponen	t		
(b)		Wri	Ma te the	antissa	est po	ositive			part	(a)(i).		Ex en as	ponen	rmalis		
(b)		Wri	Ma te the	e large	est po	ositive			part	(a)(i).		Ex en as	ponen	rmalis		
(b)		Wrinum	te the	e large n this	est poforma	positive at.		er tl	n part	an be	writte	Expen as	a no	rmalis	sed flo	pating-p
(b)) (i)	Wrinum	Mate the	e large n this t	est poforma	positive at.	numb	er tl	n part	an be	writte	Expen as	a no	rmalis	sed flo	pating-p

(iii)	If a p	ositive	e numb	er is a	added	to the	numb	er in I	part (b)(i) ex	(plain	what v	vill ha	ppen.	
															[2]
(c)	A s	studer	nt write	es a pr	ogram	to ou	itput ni	umber	s usin	g the f	ollowir	ng cod	le:			
		FOR	X ←	0 - 0 T X +		00										
	Th	e stud	lent is	surpri	sed to	see t	hat the	e progi	ram oı	utputs	the fol	lowing	sequ	ence:		
	0.0	0.1	0.2	0.29	99999	0.3	99999	9								
	Ex	plain v	why th	nis outp	ut has	s occu	ırred.									
•																,o]
Quest																
1 In a	par	ticular	comp	outer s	ystem	, real	numb	ers ar	e stor	ed usii	ng floa	ating-p	oint r	epres	entation v	vith:
•				antiss xponer		owed	by									
						forb	oth ma	ntiood	s and s	0.420.00	ont					
				form is												
(a)	(i)	A rea	al nun	nber is	store	d as t	he foll	owing	12-bi	t binar	y patt	ern:				
		0	1	1	0	1	0	0	0		0	0	1	1		
		Calc	ulate	the de	nary v	/alue (of this	numb	er. Sh	ow yo	ur woi	rking.				
					-							_				
																[31

(ii)	G	live the normalised binary pattern for +3.5. Show your working.
		[3]
(iii)	G	Sive the normalised binary pattern for –3.5. Show your working.
		[3]
The	nu	mber of bits available to represent a real number is increased to 16.
(b)	(i)	If the system were to use the extra 4 bits for the mantissa, state what the effect would be on the numbers that can be represented.
		[1]
	(ii)	If the system were to use the extra 4 bits for the exponent instead, state what the effect would be on the numbers that can be represented.
		[1]

(c)	A student enters the following expression into an interpreter:	
	OUTPUT (0.1 + 0.2)	
	The student is surprised to see the following output:	
	0.30000000000001	
	Explain why this output has occurred.	
		•••
		•••
		•••
		•••
		3

Answer 1

1(a)	= (0)11000000.1 (conversion to binary) = 0.110000001×2^8 (evidence of shifting binary point appropriately) = $0110000001 001000$ (stored as mantissa and exponent)	[1] [1] [1]	3
1(b)	1001111110 (one's complement of 10 bit mantissa) 1001111111 (two's complement of 10 bit mantissa) 1001111111 001000 (stored as mantissa and exponent)	[1] [1] [1]	3
1(c)	Any three from: Exponent too large to fit in 4 bits as a two's complement number Exponent will turn negative/–8 therefore, point moves the wrong way Value will be approx. +0.0029(296875)		3

Answer 2

1(a)	Exponent = 5 (conversion of exponent to denary) 0.00011 or 0.09375 or 3/32 (value of mantissa) //moving of binary point 3 (answer)	3
1(b)(i)	Mantissa = 011000000000 Exponent = 0010	2
1(b)(ii)	Any two from Precision lost Redundant leading zeros in the mantissa Bits lost off right hand end / least significant end Multiple representations of a single number	2

1(a)	Exponent = 6 (conversion of exponent to denary) 0.101 or 0.625 or 5/8 (value of mantissa) // moving of binary point 40 (answer)	3
1(b)(i)	Exponent = 5 (conversion of exponent to denary) 0.0000000110 or 3/1024 (value of mantissa) // moving of binary point 0.09375 or 3/32 (answer)	3
1(b)(ii)	Any two from The number calculated will change The same bit pattern is for a different number Software may crash (if not updated)	2

1(a)	Exponent = 5 (conversion of exponent to denary) 0.00011 or 0.09375 or 3/32 (value of mantissa) //moving of binary point 3 (answer)	3
1(b)(i)	Mantissa = 011000000000 Exponent = 0010	2
1(b)(ii)	Any two from Precision lost Redundant leading zeros in the mantissa Bits lost off right hand end / least significant end Multiple representations of a single number	2

1(a)	2 marks for working shown 1 mark for the correct answer	3
	Working: ∞ Correct calculation of <u>negative</u> value (any method) (= −0.11010001101) ∞ Correctly moving the binary point 7 places (= −01101000.1101) // Exponent 7	
	Answer: ∞ -104.8125 // -104 13/16	
1(b)	2 marks for working shown 1 mark for the correct answer	3
	Working: ∞ Correct conversion to binary (01.1001) ∞ Correct calculation of exponent (1)	
	Answer: ∞ (Mantissa) 0110 0100 0000 (Exponent) 0001	
1(c)(i)	1 mark per bullet point	2
	 Mantissa = 0111 1111 1111 Exponent = 0111 	
1(c)(ii)	1 mark per bullet point	2
	∞ Mantissa = 0100 0000 0000 ∞ Exponent = 1000	
1(d)	1 mark per bullet point to max 3	3
	 The trade-off is between range and precision Any increase in the number of bits for the mantissa, means fewer bits available for the exponent // Any decrease in the number of bits for the mantissa, means more bits available for the exponent More bits used for the mantissa will result in better precision More bits used for the exponent will result in a larger range of numbers Fewer bits used for the mantissa will result in worse precision Fewer bits used for the exponent will result in a smaller range of numbers 	

	1	
1(a)(i)	2 marks for working 1 mark for correct answer	3
	Working:	
	 = 0. 0110111 x 2⁵ places // exponent = 5 = 1101.11 (moving bp 5) 	
	Answer:	
1(a)(ii)	The first two bits of the mantissa are 0 / the same / not different / are not 01	1
1(a)(iii)	1 mark per bullet point ∞ Mantissa = 01101110 ∞ Exponent = 0100	2
1(b)(i)	2 marks for working 1 mark for correct answer	3
	Working:	
	 01011.101 0.1011101 × 2⁴ // showing calculation of exponent = 4 	
	Answer:	
1(b)(ii)	2 marks for working 1 mark for correct answer	3
	Working: □ 10100.011 // 10100011 correct use of two's complement or other method □ Exponent = 4	
	Answer: ∞ 10100011 0100	
1(c)	1 mark per bullet point (max 3)	3
	0.2 has been represented by a value just greater than 0.2 // 0.4 has been represented by a value just greater than 0.4	
	π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying these two representations together increases the π Therefore multiplying the second multiply multiplying the second multiply multiplying the second multiply multiplying the second multiply multi	
	difference difference after the calculation is significant enough to be seen (given the	
	number of positions after the decimal place)	

1(a)	2 marks for working shown 1 mark for the correct answer Working:	3
	 Correct calculation of <u>negative</u> value (any method) (= −0.11010001101) Correctly moving the binary point 7 places (= −01101000.1101) // Exponent 7 	
	Answer:	
	∞ −104.8125 // −104 13/16	
1(b)	2 marks for working shown 1 mark for the correct answer	3
	Working:	
	© Correct conversion to binary (01.1001)	
	∞ Correct calculation of exponent (1)	
	Answer:	
	∞ (Mantissa) 0110 0100 0000 (Exponent) 0001	
1(c)(i)	1 mark per bullet point	2
	∞ Mantissa = 0111 1111 1111	
	∞ Exponent = 0111	
1(c)(ii)	1 mark per bullet point	2
	∞ Mantissa = 0100 0000 0000	
	∞ Exponent = 1000	
1(d)	1 mark per bullet point to max 3	3
	The trade-off is between range and precision	
	Any increase in the number of bits for the mantissa, means fewer bits available for the exponent // Any decrease in the number of bits for the mantissa, means more bits available for the exponent	
	More bits used for the mantissa will result in better precision	
	More bits used for the exponent will result in a larger range of numbers	
	 Fewer bits used for the mantissa will result in worse precision Fewer bits used for the exponent will result in a smaller range of 	
	numbers	

1(a)(i)	1 mark per bullet point	3
	• Exponent 0010 = 2 • Mantissa 0.1010010 becomes 010.10010 // $\frac{41}{64}$ // $2 + \frac{1}{2} + \frac{1}{16}$ • Answer $2\frac{9}{16}$ // 2.5625	
1(a)(ii)	 1 mark per bullet point -3.75 = 100.01000 // -4 + 1/4 / 0.25 100.01000 becomes 1.0001000 Exponent = +2 Answer: Mantissa = 10001000 Exponent = 0010 	3
1(b)	Only the range is increased (no effect on precision)	1
1(c)	 1 mark per bullet point to max 1 There is no exact binary conversion for some numbers More bits are needed to store the number than are available 	1
1(d)	First term: Overflow Second term: Underflow	2

8(a)(i)	1101	1
8(a)(ii)	011100000000	1
8(a)(iii)	mark for positive, 1 for justification Positive the most significant / first bit in the mantissa is 0	2
8(a)(iv)	 1 mark per bullet point Exponent = 1011 = -3 // binary point moved 3 places left Mantissa 0.111 becomes 0.000111 // ⁷/₈ // ¹/₂ + ¹/₄ + ¹/₈ // 2⁻¹ + 2⁻² + 2⁻³ Answer: 7 / 64 // 0.109375 	3
8(b)	mark per bullet point Increases the range Decreases the precision	2

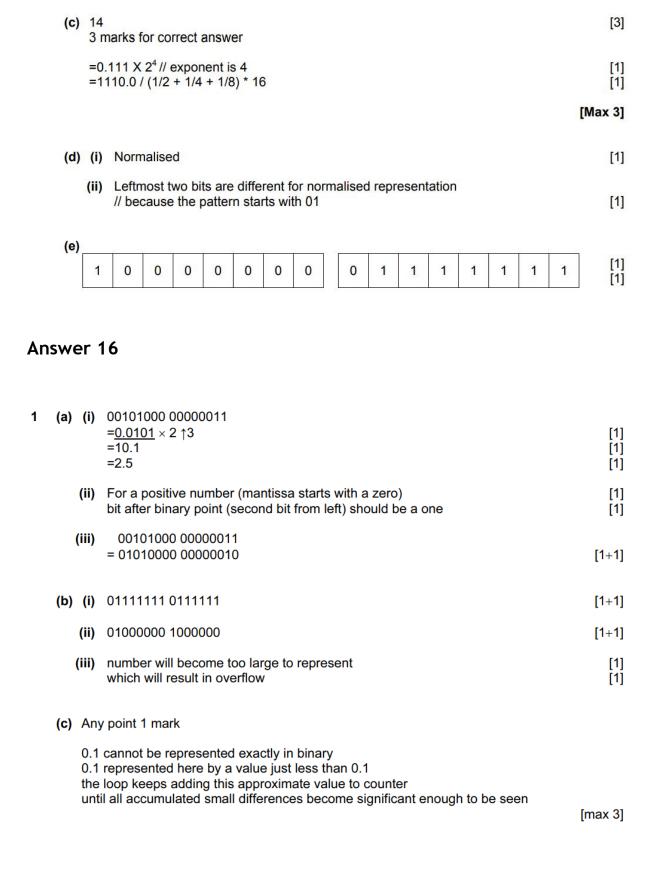
1(a)	1 mark per bullet max 2	3
	 ∑ Use exponent to denormalise mantissa 1 mark for correct answer ∑ = -17 19/32 // -17.59375 	
1(b)	1 mark per bullet	3
1(c)	1 mark per bullet ∞ (Size of mantissa decreased means that) precision is reduced ∞ (Size of exponent is increased means that) range is increased	2

3(a)	1 mark per bullet	3
	 21.75 = 010101.11 (conversion to correct binary) 0.1010111 × 2⁵ (evidence of shifting binary point appropriately) 01010111 0101 (stored as mantissa and exponent) 	
3(b)	1 mark per bullet, max 2	3
	 1110 = -2 (conversion of exponent to denary) 1.011000 = -0.101 (conversion of mantissa to negative binary number)// - 0.625 (denary value of mantissa)// -5/8 -0.00101 (binary value) // 	
	Or Use exponent to denormalise mantissa	
	1 mark for correct answer ■ -5/32 // -0.15625	

2(a)(i)	1 mark per bullet point:	3				
2(a)(ii)	2(a)(ii) The two most significant bits are 0 in the mantissa // In mantissa, 2nd bit is not the inverse of 1st bit					
2(a)(iii)	1 mark per bullet point:	2				
2(b)(i)	1 mark per bullet point:	2				
2(b)(ii)	1 mark per bullet point:	2				
2(c)(i)	Precision of numbers represented will increase					
2(c)(ii)	Range of numbers represented will increase					
2(d)	1 mark per bullet point to max 3: □ 0.1/0.2/0.3 cannot be represented exactly in binary / rounding errors adding two or more inaccurate representations together increases the probability of inaccuracy □ giving an answer where the difference is significant enough to be seen	3				

1(a)(i)	1 mark per bullet point: Correct value for exponent identified e.g. (0.010101 × 2^)5 Used to give correct value e.g. 1010.1 or 21/64 x 32 Correct answer i.e. 10.5 // 10½	3
1(a)(ii)	mark per bullet point: Correct binary value i.e. 111.1 Value for exponent identified e.g. (0.1111 × 2^)3 Correct answer i.e. 01111000 00000011	3
1(a)(iii)	1 mark per bullet point:	3
1(b)(i)	Largest (positive) number (in this format)	1
1(b)(ii)	Overflow // too large to represent // would become negative	1

1	(a)	+2.5 = 010100000000 0010		
	Give full marks for correct answer (normalised or not normalised)			
		= 10.1 = 0.101×2^2 // evidence of shifting binary point appropriately	[1] [1]	
			[Max 3]	
	(b)	-2.5 10110000000 0010 Give full marks for correct answer		
		One's complement of 12-bit mantissa of +2.5	[1] [1]	
			[Max 3]	
	(c)	3 Give full marks for correct answer	[3]	
		= 0.011 X 2 ³ // exponent is 3 = 11.0 // (1/4+1/8) * 8	[1] [1]	
			[Max 3]	
(d) (i) Not normalised	[1]	
	[1]			
(e		educed accuracy acreased range	[1] [1]	
An	SW	er 15		
1	(a)	+3.5 01110000 00000010 Give full marks for correct answer (normalised or unnormalised)	[3]	
	[1] [1]			
			[Max 3]	
	(b)	-3.5 10010000 00000010 3 marks for correct answer	[3]	
		One's complement of 8-bit mantissa for +3.5 10001111 - allow f.t. +1 to get two's complement 10010000	[1] [1]	
			[Max 3]	



1	(a)	(i)	01101000 0011 = 0.1101 (or $1/2 + 1/4 + 1/16$) × 213 = 110.1 = 6.5		[1+1] [1]	
		(ii)	+3.5 = 11.1 = 0.111 × 2†2 (or indication of moving binary point correctly) = 01110000 0010			
		(iii)	01110000 10001111 10001111 +1 = 10010000 001	Allow f.t. from (ii) One's complement on mantissa Two's complement	[1] [1]	
	(b)	(i) (ii)		cy of numbers represented will increase	[1] [1]	
	(c) Any point, 1 mark (max. 3) 0.1/0.2 cannot be represented exactly in binary // rounding error 0.1 represented by a value just greater than 0.1 // 0.2 represented by a value just greater than 0.2 adding two representations together adds the two differences summed difference significant enough to be seen					
					[Total: 14]	