Neutral Citation Number: [2023] EWCA Civ 438

Case No: CA-2022-000108

IN THE COURT OF APPEAL (CIVIL DIVISION)
ON APPEAL FROM THE HIGH COURT OF JUSTICE, BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES, INTELLECTUAL PROPERTY LIST (ChD), PATENTS COURT
Mr Justice Meade
[2021] EWHC 3121 (Pat)

Royal Courts of Justice
Strand, London, WC2A 2LL

Date: 25 April 2023

Before :

LORD JUSTICE ARNOLD
LORD JUSTICE NUGEE
and
LORD JUSTICE BIRSS

- -

Between :

(1) OPTIS CELLULAR TECHNOLOGY LLC **(2) OPTIS WIRELESS TECHNOLOGY LLC** **(3) UNWIRED PLANET INTERNATIONAL LIMITED**	**Claimants/ Appellants**
- and -	
(1) APPLE RETAIL U.K. LIMITED **(2) APPLE DISTRIBUTION INTERNATIONAL LIMITED** **(3) APPLE INC.**	**Defendants/ Respondents**

- -
- -

James Abrahams KC, James Whyte and Michael Conway (instructed by **EIP Europe LLP**) for the **Appellants**
Lindsay Lane KC and Adam Gamsa (instructed by **Wilmer Cutler Pickering Hale and Dorr LLP**) for the **Respondents**

Hearing dates : 14-15 March 2023

- -

Approved Judgment

Lord Justice Arnold:

Introduction

1. This is an appeal by the Claimants ("Optis") from an order of Meade J dated 10 December 2021 revoking European Patents (UK) Nos. 2 093 953, 2 464 065 and 2 592 779 for the reasons given in the judge's judgment dated 25 November 2021 [2021] EWHC 3121 (Pat). The judge's decision was made following the third technical trial (Trial C) between Optis and the Defendants ("Apple") in their dispute over the terms of a FRAND licence of Optis' portfolio of allegedly standard-essential patents.

2. The patents in suit are all closely related and from the same family. It was and remains common ground that it is only necessary to consider claims 1 and 4 of European Patent (UK) No. 2 093 953 ("the Patent"). There is no challenge to the claimed priority date of 19 February 2008. Apple accepted that the Patent is essential to the LTE 4G mobile telecommunications standard, and therefore infringed if valid. The judge held that claims 1 and 4 of the Patent were both obvious over Slides R1-081101 entitled "PDCCH Blind Decoding - Outcome of offline discussions" presented at a meeting of RAN1 (Radio Access Network Working Group 1) on 11-15 February 2008 ("Ericsson"), although he rejected certain other attacks on the validity of the Patent. Optis appeal with permission granted by myself.

The skilled person

3. The judge found at [28]-[40] that the person skilled in the art to whom the Patent was addressed was a person engaged in work on RAN1.

The expert witnesses

4. Optis' expert witness was Johanna Dwyer. Apple's expert witness was Professor Angel Lozano. The judge found at [11]-[17] that Ms Dwyer was a good witness, but that the subject-matter of the Patent was "materially outside her area of expertise". As a result, her evidence was "of extremely limited help on the key issues in this case". By contrast, the judge found at [18]-[27] that Prof Lozano had relevant experience, had "a practical approach" and was "an excellent witness" who was "much more cogent … than Ms Dwyer". In general, therefore, the judge preferred Prof Lozano's evidence to that of Ms Dwyer, but warned himself that he should not accept anything Prof Lozano said uncritically.

5. It is worth considering why the judge was right to give himself that warning. It is because of the function of expert witnesses in patent cases. Experts usually have little or no expertise in the issues which confront the court, such as the obviousness of a claimed invention to a person skilled in the art. Thus the experts' primary function is to educate the court in the relevant technology. This was explained by Jacob LJ in a number of judgments. It is sufficient for present purposes to cite what he said in *SmithKline Beecham plc v Apotex Europe Ltd* [2004] EWCA Civ 1568, [2005] FSR 23:

> "52. … although it is inevitable that when an expert is asked what he would understand from a prior document's teaching he will

> give an answer as an individual, that answer is not as such all that helpful. What matters is what the notional skilled man would understand from the document. So it is not so much the expert's personal view but his reasons for that view—these the court can examine against the standard of the notional unimaginative skilled man. …
>
> 53. Thus in weighing the views of rival experts as to what is taught or what is obvious from what is taught, a judge should be careful to distinguish his views on the experts as to whether they are good witnesses or good teachers—good at answering the questions asked and not others, not argumentative and so on, from the more fundamental reasons for their opinions. Ultimately it is the latter which matter—are they reasons which would be perceived by the skilled man?"

Agreed common general knowledge

6. The judge set out the agreed common general knowledge at [47]-[104]. This included a topic which Optis did not agree was common general knowledge, but accepted would be apparent from Ericsson, namely collisions and blocking. I shall largely take that necessarily detailed exposition as read, and briefly summarise the points that matter for present purposes; but when it comes to collisions and blocking it is necessary to reproduce the judge's explanation in full.

PDCCH in LTE

7. In LTE what defines a channel depends on the level in the protocol stack. Between the Radio Link Control (RLC) and the Medium Access and Control (MAC) sub-layers, "logical channels" are used to carry information for certain purposes. At the MAC sub-layer, two or more logical channels may then be combined into a single "transport channel", for onward transmission to the physical layer. At the physical layer, transport channels are mapped to "physical channels", which are configured to have a particular structure and to use a particular set of resources.

8. LTE has some physical control channels which carry signaling necessary to configure transmissions on the physical layer, such as resource allocations. In the downlink, these include the Physical Downlink Control Channel (PDCCH). Depending on context, PDCCH may refer to a specific single control channel between the eNodeB (base station) and an individual UE (user equipment) or to all such channels.

9. In LTE radio resources are divided up according to a two-dimensional space in the time and frequency domains. In the time domain LTE uses units of 10 ms called a radio frame, each of which is further divided into ten 1 ms subframes.

10. The PDCCH carries Downlink Control Information (DCI), which contains critical information for the UE, because it informs the UE about its uplink resource allocation and where to find its information on the downlink. The DCI is sent in PDCCHs in the control region at the start of the downlink subframe. The design of the PDCCH includes the UE procedure for determining its PDCCH assignment. Aspects of the

PDCCH assignment procedure and how UEs would monitor the control region to find PDCCH for them to obtain DCI were still in development at the priority date.

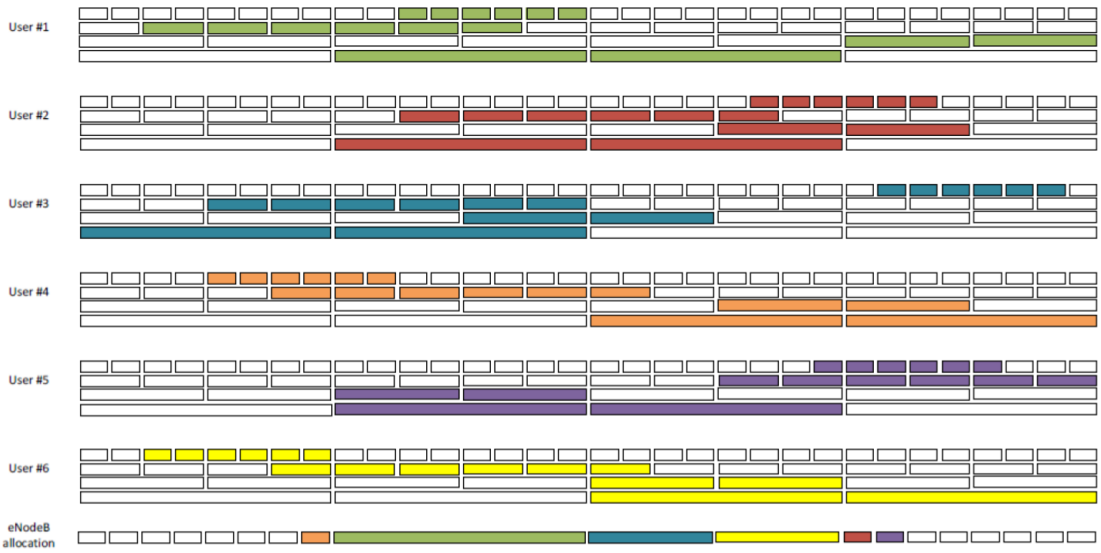
11. The PDCCH is organised into control channel elements (CCEs). A PDCCH message is transmitted on either 1, 2, 4 or 8 CCEs, called "CCE aggregations". A CCE aggregation that may carry a PDCCH message for a UE is referred to as a "PDCCH candidate".

12. Each subframe, a UE has to monitor the PDCCH to determine whether it contains DCI messages intended for it. To detect whether a PDCCH candidate contains control information for the UE, the UE searches for its unique identifier (referred to as the "UE ID"). The UE ID is allocated temporarily to the UE by the eNodeB when the UE enters that cell. The UE ID is encoded in the Cyclic Redundancy Check (CRC) of the PDCCH for that UE. The CRC is scrambled or "masked". The process of searching for the UE ID in the masked CRC of the PDCCH is referred to as "blind decoding".

13. Because it would be resource intensive to require each UE to monitor and attempt to blind decode the whole of the PDCCH each subframe, a UE is only required to monitor a subset of all possible PDCCH candidates, called a "search space". A search space can be defined by its starting point and size; and if the size is fixed, by just the starting point.

Collisions and blocking

14. In the specific context of search spaces on the PDCCH, a "collision" refers to the search spaces completely overlapping (i.e. starting at the same location). To illustrate this concept, the diagram below shows one way that the search space for six UEs might be arranged. The search space for each UE is shown in a different colour. In the diagram, the size of the search spaces at each aggregation level is 6 aggregations at each of aggregation levels 1 and 2 (the top two levels in each case, with 32 CCEs and 16 CCEs), and 2 aggregations at each of levels 4 and 8 (the bottom two levels in each case, with 8 and 4 CCEs).



Illustrative example of search spaces and actual allocation for 6 UEs

15. In each subframe, the eNodeB chooses where within each UE's search space to allocate DCI messages for that UE. One possible such choice is shown in the "eNodeB allocation" section of the diagram above. In this example, the eNodeB is sending one message for each UE. The eNodeB has chosen the aggregation level for each UE, and identified an arrangement for all the messages so that each UE's message is within that UE's search space at the appropriate aggregation level. This process is repeated in the next subframe.

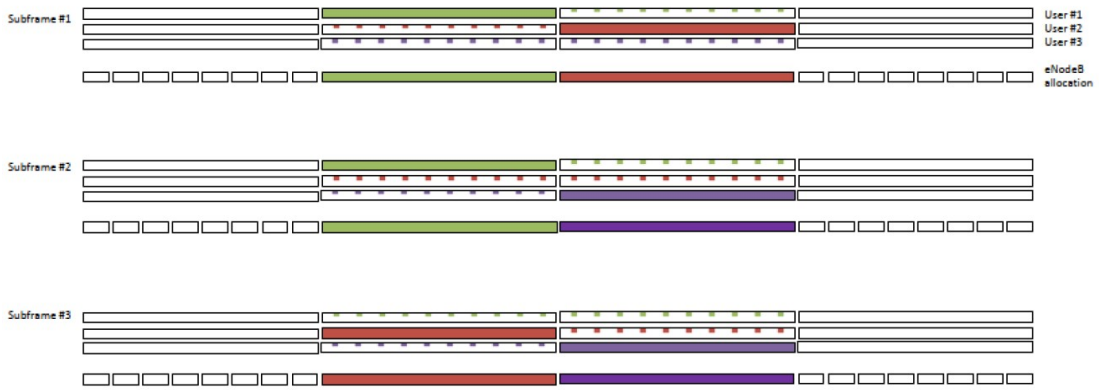
16. There may be circumstances in which the eNodeB is not able to find locations for all the messages it wishes to send. For example, in the diagram above, the eNodeB could not send a DCI message at aggregation level 8 to all of UEs 1, 2 and 5. The search space at aggregation level 8 for these three UEs completely overlaps, and comprises only two CCE aggregations. This is an example of "blocking". One of the three UEs is blocked; the eNodeB has to make a decision as to which UE's message it will not send in that subframe.

17. In other circumstances, search spaces can completely overlap - i.e. there is a collision - but there is no blocking. For instance, in the example in the previous paragraph, if the eNodeB only wanted to send messages at aggregation level 8 to two of the three UEs, it could do so.

18. Blocking may also arise because of the interaction between aggregation levels. In the diagram above, the eNodeB could send an aggregation level 8 message to each of UEs 1-4, but doing so blocks any messages to UEs 5 and 6.

19. If only two UEs are being considered and there are enough CCEs available, there will never be a situation where a message to one UE blocks a message being sent to the other. In practice, blocking arises because there are more than two UEs and/or there are too few CCEs. (This assumes that the search space at each aggregation level contains more than one CCE aggregation at that aggregation level.)

20. From a system viewpoint, it is the level of blocking, not collisions, that is important. Blocking does not mean that all communication is prevented, but that some is. In the example above, each of the three UEs can still receive a DCI message in two out of every three subframes. This is illustrated further in the diagram below.



Persistent collisions, blocking and the effect on the connection between the UE and the eNodeB.

21. The diagram shows three subframes with 32 CCEs for PDCCH in each, and illustrates the search spaces for three particular UEs at aggregation level 8. In this illustration, all other aggregation levels are omitted, and the CCEs forming part of the search space but not used for a DCI message are shown shaded. In this case, there is a persistent collision between the three UEs, meaning that each UE has the same search space in each of the three subframes (CCE#8 to 23, i.e. two aggregations at aggregation level 8). So the eNodeB cannot send a DCI message at aggregation level 8 to each of these UEs in each of these subframes. There is blocking: some communication is prevented. Nevertheless, the eNodeB could still send DCI messages to each UE in two out of the three subframes. The overall effect is therefore to degrade the connection between the eNodeB and each UE, without necessarily breaking it.

Hashing functions and random number generation

22. A hashing function is any function that can be used to map data of potentially arbitrary size to fixed sized values. In other words, it is a function that allocates a large number of inputs to a small known number of outputs.

23. A random number is a number selected from a range of numbers where each number in the range has a certain chance of being selected, but each selection of a number is completely unpredictable. A pseudo-random number is a number generated in software using an algorithm (a Random Number Generator or RNG) which is deterministic. Random numbers and pseudo-random numbers have uses in many fields, including communication technologies. In the remainder of this judgment I shall follow the judge's example of referring to pseudo-random numbers as random numbers.

Disputed common general knowledge

24. The judge made findings as to a number of disputed areas of common general knowledge at [105]-[141]. It is important to note that, for convenience, the judge included in this section of his judgment findings as to information which was not common general knowledge, but which the judge held the skilled person would ascertain as a matter of routine when considering Ericsson.

Modular arithmetic

25. Modular arithmetic is arithmetic involving the modulo or "mod" operation. This is simply the exercise of finding a remainder, so 17 mod 3 = 2, or 1000 mod 111 = 1. It was agreed that this was common general knowledge. The judge found that the "distributive property" of modular arithmetic would also be common general knowledge. This property was explained by Prof Lozano as follows:

    (X+Y) mod C = (X mod C + Y mod C) mod C.

26. In the left hand side of the equation you add X and Y and take the remainder after dividing by C. In the right hand side of the equation you divide by C and take the remainder for X and Y separately and then add the remainders. But that might be greater than C, so you perform the mod C operation on the total. In each case you are just getting rid of all the multiples of C.

LCGs

27. What follows was not common general knowledge, but information the skilled person would find out if they consulted Knuth (as to which, see below).

28. A Linear Congruential Generator (LCG) is a random number generator with the following form:

$X_{n+1} = (AX_n + B) \text{ mod } D.$

29. The equation set out above shows that an LCG is recursive. The LCG will generate a sequence of numbers in the following way:

    i) You take the previous value in the sequence, which was X_n.

    ii) You multiply it by A. A is thus the multiplier.

    iii) You add B. B is thus the increment.

    iv) You take mod D of the result. D is the modulus.

    v) The process is repeated.

30. The sequence has to start somewhere, and that is called the seed (or start value, or initial value, or similar). Eventually the sequence will repeat. The number of iterations before repeating is called the period. The period cannot exceed the modulus, but it may be less. Maximising the period depends on the parameters A, B and D chosen.

Experience of RNGs

31. There was an issue between the parties as to how much experience the skilled person would have of practically using RNGs. The judge found that, although the skilled person would not necessarily have any previous experience of using RNGs, they would nevertheless be able to do so. The Patent assumed that the skilled person would be able to determine alternative values for A, B and D to those specified in claim 4 with only modest assistance from the specification. Thus the skilled person would think about and understand the choices they were presented with in NRC (as to which, see below) and Knuth, and they would not lack confidence so that they only used off-the-shelf solutions.

Hashing functions v random number generators

32. Optis argued that hashing functions and RNGs were regarded as separate and distinct concepts. The judge rejected this for reasons he expressed as follows:

> "120. I think this was artificial; an attempt to create a conceptual difference that would not be seen by the skilled person to matter. My reasons include:
>
> i) Counsel for Optis had opened the case by saying that 'Hashing functions do involve randomisation as a

means of achieving an even or uniform distribution, as your Lordship has obviously got'.

ii) The goal of a hashing function is to spread a large number of inputs evenly over a smaller number of outputs. The even spread may be achieved by mimicking a random spread.

iii) Ms Dwyer accepted that hashing functions that are uniform and random are good.

iv) The terms 'hashing function' and 'randomisation function' are used interchangeably in the art, including in the secondary evidence relied on by Optis, as Ms Dwyer also accepted. Ms Dwyer's own written evidence in relation to Ericsson referred to the 'randomization' and 'randomness' of the hashing function.

v) Knuth's chapter on hashing cross-refers to the chapter on random numbers. Optis sought to downplay this, but Ms Dwyer accepted that the direction was to look to the chapter on random numbers in connection with getting a hashing function whose overall output was random.

121. I therefore reject the argument that it was CGK (or would emerge from routine research) that hashing functions and random number generators were separate and distinct from one another."

Recursive v self-contained

33. Optis argued that there was a fundamental difference between a recursive function (which an LCG is, because the production of each number uses the previous answer as an input as explained above) and a free-standing or "self-contained" function such as that in Ericsson, where the n^{th} number in the sequence produced can be calculated directly without deriving the previous ones first.

34. The judge found that, although there was a difference between self-contained and recursive functions, there was no reason for the skilled person to care, or to be worried by using a recursive function, or to regard the difference as a radical or practically significant one. Working out the one millionth term of a recursive function might be computationally too intensive, but in the circumstances of the PDCCH all that would be needed would be to work out and store ten values for each aggregation level. That would not present any difficulty. Thus there was no practical problem which would put the skilled person off from using an LCG merely because it was recursive.

Simulations

35. The judge found that it was common general knowledge to use simulations to test proposals that were made during RAN1 work using appropriate metrics.

Max hits

36. A metric that is used in the simulations presented in the Patent is maximum number of hits or "max hits". The judge found that this was not common general knowledge.

Ericsson

37. Ericsson is a seven-page presentation. The judge reproduced pages 1, 2 and 5, which summarise where the development of the PDDCH had got to at the time of the RAN1 meeting, at [161]-[163]. The key page is page 6:

38. The subframe number is simply a number from 0 to 9, since there are 10 subframes in a frame. The idea is that the output of the hashing function which defines the start of the search space for each UE should be different in each subframe. The judge referred to the floor(#CCEs/aggregation_level) part of the hashing function as C for brevity and ease of comparison with the function in the Patent, and I will follow his example. C is essentially the number of possible CCE aggregations at a particular aggregation level. (If the number of CCEs is not precisely divisible by the aggregation level, the "floor" operation takes only the integer part.) Taking mod C gives the remainder after dividing by C, and so the result is a number between 0 and C-1.

The Patent

39. The invention described and claimed in the Patent is intended to address the problems of collisions and blocking on the PDCCH described above. The judge pithily summarised the teaching of the specification at [142]-[148] as follows. It begins with some general teaching about LTE and the PDCCH at [0004]-[0006]. It identifies Ericsson at [0007]. The specification then goes at length into the meaning and use of the function of claim 1, which involves an LCG and a mod C operation.

40. From [0096] onwards, the specification starts to describe the choice of parameters for the LCG. It explains at [0099] that it will use the concept of number of "hits" (which essentially means collisions) as a criterion. At [0103] it explains that it will be looking at average number of hits and maximum number of hits, as well as whether the range 0 to C-1 is uniformly covered by the start positions generated, and the variance of probabilities that values between 0 and C-1 will be generated (another measure of uniformity).

41. Tables 2, 3 and 4 are then presented, giving values for those metrics for various combinations of the parameters A, B, C and D.

42. At [0107]-[0108] the specification gives some guidance as to the choice of parameter D, which is the modulus. It refers to the size of D and to whether D is prime. At [0109] it recommends choosing D = 65537 when the UE ID is a 16-bit number.

43. At [0112] the specification recommends that B is set to zero.

44. Prof Lozano's evidence was that [0107]-[0108] were unreliable, or at least very badly written, because (paraphrasing for simplicity) they misunderstood or mis-explained the importance of the size of D on the one hand, and whether it is prime on the other. The judge accepted that evidence.

45. Prof Lozano also gave evidence that Tables 3 and 4 (in particular the latter) contain errors. A particular problem is that they do not specify C, with the result that they cannot be replicated. Prof Lozano attempted to work out what had been done, and did manage to verify that the max hits column in Table 4 would make sense if C were 16; but that would mean the other columns were wrong. Again, I understand the judge to have accepted this evidence.

The claims

46. Broken down into integers, claim 1 is as follows:

> "[a] A method for a User Equipment, UE, to receive control information through a Physical Downlink Control Channel, PDCCH, the method comprising:
>
> [b] receiving control information from a base station through the PDCCH in units of Control Channel Element, CCE, aggregations, each of the CCE aggregations including at least one CCE in a control region of subframe 'i'; and
>
> [c] decoding the received control information in units of search space at subframe 'i',
>
> **characterized in that**
>
> [d] the search space at subframe 'i' starts from a position given based on a variable x_i and a modulo 'C' operation, wherein 'C' is a variable given by: $C = \text{floor}(N_{CCE} / L_{CCE})$,
>
> and wherein 'x_i' is given by: $x_i = (A * x_{i-1} + B) \bmod D$,

> wherein A, B and D are predetermined constants, and x_{-1} is initialized as an identifier of the UE, and N_{CCE} represents the total number of CCEs at subframe 'i', and L_{CCE} is the number of CCEs included in the CCE aggregation, and floor(x) is a largest integer that is equal to or less than x."

47. Claim 1 uses the notation x_i, while the specification refers to Y_k. This does not make any substantive difference, but needs to be borne in mind when reading the Patent. The "i" or "k" denotes the subframe number.

48. As the judge explained at [155]-[157], essentially what claim 1 is saying is that, for each subframe, the start position of a PDCCH search space is found using an LCG, with the output of the LCG being subjected to a mod C operation. For the initial "seed" value for the LCG, the UE ID is used. For subsequent subframes the LCG works recursively. C is the number of possible start positions, and is found by taking the number of CCEs in the subframe and dividing by the aggregation level L. So if there are e.g. 64 CCEs and the aggregation level is 8, then there are 8 possible starting positions. Taking mod C will give a number from 0 to C-1. The reasons why this is an improvement over the Ericsson function are explained below.

49. Claim 4 is as follows:

> "The method according to claim 1 or 2, wherein D, A, and B are 65537, 39827, and 0, respectively."

This claims a "bespoke" LCG, in the sense that the values have been selected specifically for this purpose. It is not an "off-the-shelf" LCG which can be found in the reference works referred to below.

Wikipedia, NRC and Knuth

50. Apple's obviousness case based on Ericsson also involved three standard reference works:

i) Wikipedia;

ii) Press *et al*, *Numerical Recipes in C: The Art of Scientific Computing* (2nd ed, 1992) ("NRC2") and *Numerical Recipes: The Art of Scientific Computing* (3rd ed, 2007) ("NRC3") (collectively "NRC");

iii) Knuth, *The Art of Computer Programming*, vol. 2 *Seminumerical Algorithms* (2nd ed, 1981), specifically Chapter 3 "Random Numbers" ("Knuth").

51. The reason for this is that, as explained in more detail below, it was Prof Lozano's evidence that, having found that the function described in Ericsson did not have the "desired properties", the skilled person would look to replace the K*x + L part of the Ericsson function with an RNG. In order to find a suitable RNG, the skilled person would turn to either Wikipedia or NRC, both of which, he said, refer to Knuth. In that way the skilled person would find the description of LCGs in Knuth. As the judge recorded at [236], Prof Lozano said that he thought that NRC would be the most natural thing to turn to first, and that is what he himself had done.

52. It is important to note certain complications concerning each of these works. So far as Wikipedia is concerned, the pages to which Prof Lozano referred in his first report were the pages on RNGs and LCGs. The former links to the latter, and the latter cites Knuth as a reference. The page which this Court was shown, however, was the page on "hash function" (as at 15 October 2007). I will explain the significance of this below.

53. So far as NRC is concerned, it turned out that, when preparing his first report, Prof Lozano had consulted NRC2, because that was what he had had on his bookshelves. Optis only found NRC3 shortly before trial, after all the written evidence had been exchanged. The judge found at [238], however, that the skilled person would use NRC3, which was published about six months before the priority date. As explained below, NRC3 is materially different to NRC2.

54. As for Knuth, the chapter relied on is in volume 2 of a three-volume treatise. The second edition of that volume was published in 1981. Even though a third edition of that volume was published in 1997, Apple relied upon (the first 40 pages of Chapter 3 in) the second edition not only as part of its obviousness case starting from Ericsson, but also as an independent starting point. (The judge rejected Apple's case starting from Knuth, and there is no challenge by Apple to that rejection.) It is common ground that the trial proceeded on the assumption that there was no material difference between the second and third editions so far as Chapter 3 is concerned. A separate point is that volume 3 of the treatise (subtitled *Sorting and Searching*, the second edition of which was published in 1998) includes Chapter 6, "Searching", which in turns includes section 6.4, "Hashing".

Wikipedia

55. The judge did not set out the disclosure of Wikipedia, no doubt because he found at [239] that it was significantly less likely that the skilled person would turn to Wikipedia than NRC. It is not clear from the judgment which page of Wikipedia the judge had in mind. I do not think that he can have envisaged that the skilled person would have gone directly to the page on LCGs, because the skilled person could only consult that page if they knew about LCGs; but on the judge's findings they would not have done so at that stage of the analysis. It is possible that the judge envisaged that the skilled person would first go to the page on RNGs, but then the question would arise as to why they would then go to the page on LCGs when the RNG page also describes at least four other classes of RNG. The judge provides no answer to this question, nor were we shown any evidence which does.

56. I presume that this is why we were shown the page on "hash function". As its title suggests, this describes hashing functions. It says, in the context of describing the application of hash functions in hash tables (tables which enable the fast lookup of a data record given its key), that "[h]ash functions that are truly random with uniform output … are good". Although there are many cross-references to other Wikipedia entries, however, there is no reference to RNGs. The only reference to Knuth is towards the end of the page, under the heading "Origins of the term":

> "Knuth notes that Hans Peter Luhn of IBM appears to have been the first to use the concept, in a memo dated January 1953, and that Robert Morris used the term in a survey paper in

> CACM which elevated the term from technical jargon to formal terminology.[4]"

57. Reference 4 is "Knuth, Donald (1973), *The Art of Computer Programming*, volume 3: *Sorting and Searching*". As can be seen, this is the first edition of volume 3 of the treatise (and this topic is indeed discussed in section 6.4).

58. It follows that, if the skilled person consulted the "hash function" page, they would not be led to volume 2 of Knuth, let alone to Chapter 3. It further follows that Wikipedia can be disregarded as a route to Knuth.

NRC

59. There is no dispute that NRC was a standard reference work. As I have noted, the judge found that the skilled person would use NRC3 rather than NRC2. There is an important difference between them: unlike NRC2, NRC3 warns the reader against using LCGs. Although the judge set out most of the material at [241]-[246], in view of its importance I must do so again and a little more fully.

60. Although there is a chapter in NRC3 about hashing functions, for the reasons touched on above and explained more fully below, the relevant chapter of NRC3 for the purposes of Apple's case is Chapter 7, "Random Numbers". In section 7.0, "Introduction" (pages 340-341), the authors say:

> "It may seem perverse to use a computer, that most precise and deterministic of all machines conceived by the human mind, to produce 'random' numbers. More than perverse, it may seem to be a conceptual impossibility. After all, any program produces output that is entirely predictable, hence not truly 'random'.
>
> Nevertheless, practical computer 'random number generators' are in common use. We will leave it to philosophers of the computer age to resolve the paradox in a deep way (see e.g. Knuth [1] §3.5 for discussion and references). …
>
> The pragmatic point of view is thus that randomness is in the eye of the beholder (or programmer). What is random enough for one application may not be random enough for another. Still, one is not entirely adrift in a sea of incommensurable applications programs: There is an accepted list of statistical tests, some sensible and some merely enshrined by history, that on the whole do a very good job of ferreting out any nonrandomness that is likely to be detected by an applications program (in this case, yours). Good random number generators ought to pass all of these tests or at least the user had better be aware of any that they fail, so that he or she will be able to judge whether they are relevant to the case at hand.
>
> For references on this subject, the one to turn to first is Knuth [1]. Be cautious about any source earlier than about 1995, since the field progressed enormously in the following decade."

Reference 1 is (the third edition) of Knuth "especially §3.5".

61. In section 7.1, "Uniform Deviates" (pages 341-342), the authors say:

> "Uniform deviates are just random numbers that lie within a specified range …
>
> The state of the art for generating uniform deviates has advanced considerably in the last decade and now begins to resemble a mature field. …
>
> The greatest lurking danger for a user today is that many out-of-date and inferior methods remain in general use. Here are some traps to watch for:
>
> - Never use a generator principally based on a *linear congruential generator* (LCG) or a *multiplicative linear congruential generator* (MLCG). We say more about this below.
>
> - Never use a generator with a period less than $\sim 2^{64} \approx 2 \times 10^{19}$, or any generator whose period is undisclosed.
>
> - Never use a generator that warns against using its low-order bits as being completely random. That was good advice once, but it now indicates an obsolete algorithm (usually a LCG).
>
> …
>
> If all scientific papers whose results are in doubt because of one or more of the above traps were to disappear from library shelves, there would be a gap on each shelf about as big as your fist.
>
> You may also want to watch for indications that a generator is overengineered, and therefore wasteful of resources:
>
> - Avoid generators that take more than (say) two dozen arithmetic or logical operations to generate a 64-bit integer or double precision floating result.
>
> - Avoid using generators (over-)designed for serious cryptographic use.
>
> - Avoid using generators with period $> 10^{100}$. You *really* will never need it, and, above some minimum bound, the period of a generator has little to do with its quality.
>
> Since we have told you what to avoid from the past, we should immediately follow with the received wisdom of the present:

> An acceptable random generator must combine at least two (ideally, unrelated) methods. The methods combined should evolve independently and share no state. The combination should be by simple operations that do not produce results less random than their operands.

> If you don't want to read the rest of this section, then use the following code to generate all the uniform deviates you'll ever need. This is our suspenders-and-belt, full-body-armor, never-any-doubt generator; and it also meets the above guidelines for avoiding wasteful, overengineered methods. …"

62. In section 7.1.1, "Some History" (pages 343-344), the authors say:

> "With hindsight, it seems clear that the whole field of random number generation was mesmerized, for far too long, by the simple recurrence equation [for LCGs as set out in paragraph 28 above]. …
>
> The idea of LCGs goes back to the dawn of computing, and they were widely used in the 1950s and thereafter. The trouble in paradise first began to be noticed in the mid-1960s …
>
> [Various problems with LCGs are described]
>
> Looking back, it seems clear that the field's long preoccupation with LCGs was somewhat misguided. There is no technological reason that the better, non-LCG, generators of the last decade could not have been discovered decades earlier, nor any reason that the impossible dream of an elegant 'single algorithm' generator could not also have been abandoned much earlier (in favor of the more pragmatic patchwork in combined generators). As we will explain below, LCGs and MLCGs can still be useful, but only in carefully controlled situations, and with due attention to their manifest weaknesses."

63. Reference was made in some of the oral evidence at trial to section 7.1.7, but none of the content of that section was included in the judge's exposition. Nor was that section included in the five-page extract from Chapter 7 which was put before this Court. It follows that it must be disregarded for the purposes of the appeal.

Knuth

64. The judge helpfully summarised the teaching of Knuth at [165]-[176]. I shall reproduce his account with a few small clarifications.

65. After section 3.1, "Introduction", there follows section 3.2, "Generating Uniform Random Numbers". The first method introduced in 3.2.1 is the Linear Congruential Method. LCGs are described (at page 9) as "[b]y far the most popular random number generators in use today". After setting out the form of the LCG, Knuth observes that

choosing the "magic numbers", meaning A, B, D and the starting value, appropriately will be covered later in the chapter. He explains that sequences from LCGs have a period and that "[a] useful sequence will have a relatively long period."

66. At the top of page 10, Knuth says that "[t]he special case c = 0 deserves explicit mention". He uses c to denote the increment i.e. B in the form of equation set out above. He explains that this case is referred to as "multiplicative", and says that it is quicker but tends to reduce the period.

67. Thereafter, Knuth gives advice on the choice of modulus in 3.2.1.1 (page 11 onwards). He says this should be "rather large" and notes that its choice affects speed of generation. He deals with choosing a modulus when the word size of the computer in question is *w* (the word size is 2^e for an e-bit binary processor) and discusses the case where the modulus is set to *w*+1 or *w*-1. He also provides Table 1, which gives the prime factorisations for various values of e.

68. Choice of multiplier is discussed in 3.2.1.2 (page 15 onwards). Knuth explains that the intention is to choose the multiplier so as to give the period of maximum length, and that "we would hope that the period contains considerably more numbers than will ever be used in a single application".

69. Further specific advice is given for particular cases over the following pages. For example, getting a long period with an increment of zero (the multiplicative case) is covered at page 19, and Knuth says that with an increment of zero an effectively maximum period can be achieved if the modulus is prime.

70. Section 3.2.2 (page 25 onwards) introduces "Other Methods", i.e. other than LCGs. This contains the caveat that a common fallacy is to think that a small modification to a "good" generator can make it even more random, when in fact it makes it much worse.

71. Section 3.3 (page 38 onwards) then introduces statistical tests to test if sequences produced are in fact random.

72. Section 3.6, "Summary" (page 170 onwards), describes "a simple virtuous generator" and gives advice for the choice of the seed, modulus, multiplier and increment. It recommends that the modulus "should be large, say at least 2^{30}".

73. It is worth noting that, although NRC3 twice cites section 3.5 of (the third edition of) Knuth, the judge did not include any material from (the second edition of) that section in his summary. Nor did either side rely upon that section before this Court. As the passage I have quoted from NRC3 in paragraph 60 above suggests, however, section 3.5 addresses the question "What is a random sequence?". It is a detailed consideration of what is meant by randomness in mathematical terms.

<u>The judge's assessment of obviousness over Ericsson</u>

The law

74. The judge set out the applicable legal principles at [178]-[198]. Optis do not criticise his exposition of the law. It included a distillation of the principle expressed by Lord

Diplock in *Technograph Printed Circuits Ltd v Mills & Rockley (Electronics) Ltd* [1972] RPC 346 at 362:

> "The cross-examination of the respondents' expert followed with customary skill the familiar 'step by step' course. I do not find it persuasive. Once an invention has been made it is generally possible to postulate a combination of steps by which the inventor might have arrived at the invention that he claims in his specification if he started from something that was already known. But it is only because the invention has been made and has proved successful that it is possible to postulate from what starting point and by what particular combination of steps the inventor could have arrived at his invention. It may be that taken in isolation none of the steps which it is now possible to postulate, if taken in isolation, appears to call for any inventive ingenuity. It is improbable that this reconstruction a posteriori represents the mental process by which the inventor in fact arrived at his invention, but, even if it were, inventive ingenuity lay in perceiving that the final result which it was the object of the inventor to achieve was attainable from the particular starting point and in his selection of the particular combination of steps which would lead to that result."

The Pozzoli *analysis*

75. The judge adopted the structured *Pozzoli* approach to the assessment of obviousness. The judge considered question 3 at [201]-[205]. He held that the difference between Ericsson and claim 1 was the use of the function of that claim instead of the Ericsson function. He held that the difference between Ericsson and claim 4 consisted in addition of the combination of values A=39827, B=0, and D=65537.

Claim 1

76. As the judge explained at [206], Apple's case was that the skilled person would:

> "i) Verify whether the Ericsson function would provide 'the desired properties' and conclude that it would not.
>
> ii) Observe that the problem lay with the randomisation part of the Ericsson function (and not the mod C part).
>
> iii) Do a literature search to find another appropriate hashing/randomisation function.
>
> iv) Identify, ultimately from Knuth, LCGs as a good choice."

77. As the judge explained at [193]-[194], there was a dispute before him as to the number of steps which Apple's case involved. Optis contended that it required as many as 19 steps, while Apple said that it only involved three. The judge held at [195]-[196] that what mattered was not so much the number of steps, which depended on how they were formulated and counted, but whether "the gap between the prior art

and patent is being deconstructed in such a way as to build in hindsight". I agree with this.

78. As part of their argument before this Court, Optis characterised the judge's assessment of the obviousness of claim 1 as involving seven steps. For the reason I have just given, the number of steps does not matter, particularly since it is debatable whether some of the steps are truly distinct from each other. It is nevertheless convenient to adopt Optis' numbering when summarising the judge's reasoning, since it helps to identify the targets of Optis' grounds of appeal.

79. *Step 1*. Ericsson itself invited the reader to assess the function proposed, and that is what the skilled person would do (judgment at [208]-[210]). The skilled person would realise, either using analytical common sense and/or by performing simulations, that the function would not work for two reasons ([211]-[220]). First, if two UEs collided in one subframe, they would collide in every following subframe. This is because, for each UE, x would be incremented by the same amount in each subframe. This problem (referred to as "persistent collision" or "lockstep") arises because of the distributive property of modular arithmetic. Secondly, the function would not work for C=16.

80. It is worth noting that the judge did not spell out precisely why the function would not work for C=16. This is because x is obtained by multiplying the UE ID by 16 and then adding the subframe. Taking mod C of UE ID*16 will give no remainder where C=16 and so the value of x in each subframe will be the same. Due to the distributive property of modular arithmetic, taking mod C of K*x + L has the same result. Thus taking mod C of K*x + L would not give a random distribution where C=16, but rather would again lead to persistent collision or lockstep. Nugee LJ posed the question during the course of argument: is the same not true where C=2, 4 and 8? The parties agreed that the answer to that question is yes. This does not seem to have been spotted by the parties, the experts or the judge in the court below. This is of some relevance to the issue of hindsight considered below.

81. *Step 2*. The skilled person would realise that the mod C part of the Ericsson function was there to "squeeze" the random output of the K*x + L part down from a large number to a range from 0 to C-1, that the mod C part was therefore necessary and was working all right, and that any change should therefore be to the random K*x + L part. Accordingly, an obvious route was to retain mod C, on the basis that it was adequately performing a well-understood and necessary task, and look to remedy the problem with the K*x + L randomisation part ([221]-[227]).

82. *Step 3*. The skilled person would decide to replace K*x + L with an RNG ([228]-[233]).

83. *Step 4*. The skilled person would look in the literature for an appropriate RNG ([228]-[233]).

84. *Step 5*. The skilled person would find their way to Knuth, and in particular the discussion of LCGs, most probably via NRC3 ([234]-[250]).

85. *Step 6*. Having looked at Knuth, the skilled person would conclude that LCGs had much to commend them. They would see Knuth as a reliable source of teaching about

how to implement LCGs. The negative comments about LCGs in NRC3 were of low relevance to obviousness. The skilled person would consider that LCGs were well known, had been widely used for a long time, easy to implement and suitable for low power devices. Although LCGs were not of good enough randomness for demanding applications such as cryptography, the skilled person would appreciate that the problem presented by Ericsson was not a demanding application. Thus there was no relevant prejudice against LCGs. ([251]-[257]).

86. Various points raised by Optis did not undermine the judge's view that using an LCG would be an obvious thing for the skilled person to do ([258]-[267]).

87. *Step 7*. The skilled person would therefore decide to replace the K*x + L part of the Ericsson function with an LCG, and retain the mod C part to convert the result of the LCG to a value between 0 and C-1 ([268]-[270]).

88. Two other points raised by Optis did not detract from this conclusion ([271]-[277]).

89. The secondary evidence did not militate against this conclusion; if anything, it supported Apple's case ([278]-[291]).

90. Although there were alternative routes which the skilled person could follow, they would not all be attractive as a way forward. In any event, the fact that LCGs were so well known and understood, and still regarded as useful where the randomness needed was not too great, meant that they would be at the very least a leading option ([292]-[293]).

91. *Summary*. The judge summarised his conclusion as follows:

> "294. I accept Apple's case. The problem with the Ericsson function would be identified readily and by routine analysis. It would be clear without invention that the mod C part of the function was all right and the Kx + L needed remedy. A literature search would readily throw up LCGs as the best known option and although they would be known to have limitations they would be regarded as adequate for the PDCCH task in hand.
>
> 295. The secondary evidence does not displace this conclusion and nor do the existence of other alternatives; in any event an LCG would be at the forefront of any list of options.
>
> 296. In reaching this conclusion I have borne very much in mind that Apple's case involves a number of sequential steps. But I find that they represent systematic uninventive work and not the use of inappropriate hindsight."

Claim 4

92. Optis characterise the judge's assessment of the obviousness of claim 4 as involving four further steps. Again, I shall adopt this numbering for the purposes of summarising the judge's reasoning.

93. *Step 8*. Although one option would be to use what Knuth suggested in his summary, another would be for the skilled person to think about values specifically suitable for the PDCCH situation ([300]).

94. *Step 9*. B=0 was a known useful category of LCGs. It would be one obvious option, and probably the most obvious option ([301]).

95. *Step 10*. D=65537 was an obvious choice, and probably the most obvious choice since it stood out clearly from Table 1 on page 13 of Knuth once one understood the appropriate thinking ([302]-[314]).

96. *Step 11*. A=39827 was one of a number of values that worked which could be identified by routine simulations. Moreover, it would be obvious to choose a multiplier which maximised period length, which 39827 did for a modulus of 65537 ([315]-[320]).

97. The judge concluded at [321]:

> "Thus each value in claim 4 is at least *an* obvious value. Although I have had to explain my reasoning in a sequence of steps I am satisfied there is no improper salami-slicing. Each step flows into the next for logical and routine reasons. The choices are related in ways which support their combination being obvious. This is basically careful implementation work of the kind for which the notional skilled person is inherently well-equipped, and this was part of the case where I thought Prof Lozano's advantage over Ms Dwyer was particularly striking. Claim 4 is obvious."

Appeals on obviousness

98. Obviousness involves a multi-factorial evaluation, and therefore this Court is not justified in intervening in the absence of an error of law or principle on the part of the judge: see *Actavis Group PTC EHF v ICOS Corp* [2019] UKSC 15, [2019] Bus LR 1318 at [78]-[81] (Lord Hodge). This accords with the general approach of this Court to appeals against evaluative decisions: see *Re Sprintroom Ltd* [2019] EWCA Civ 932, [2019] BCC 1031 at [72]-[78] (McCombe, Leggatt and Rose LJJ).

99. In the present case Optis not only face this obstacle, which confronts all appeals on obviousness, but two more specific difficulties. The first is the judge's assessment of the expert witnesses, which Optis cannot and do not challenge. The second is that, on its face, the judgment contains a very careful, detailed and nuanced consideration of the evidence and the issues. I will explain below why Optis say that these difficulties are not insuperable.

Grounds of appeal

100. Optis appeal on eight grounds. Grounds 1-6 concern claim 1, while grounds 7a and 7b concern claim 4. Ground 2 can be ignored because it is the consequence of grounds 1 and 3-6.

101. Ground 1 is that the judge was wrong not to find that the skilled person would have perceived hashing functions and RNGs as separate and distinct from one another, and ought to have found that the skilled person would have seen a relevant conceptual difference between them. This concerns step 2.

102. Ground 3 is that the judge erred in concluding that the skilled person starting from Ericsson would retain the mod C part of the function and focus on changing the K*x + L part. This also concerns step 2.

103. Ground 4 is that the judge erred in finding that the skilled person would look in the literature for an appropriate RNG with which to replace K*x + L. This concerns steps 3 and 4.

104. Ground 5 is that the judge erred in finding that the statements in NRC3 deprecating LCGs were of low relevance to obviousness. This concerns steps 5 and 6.

105. Ground 6 is that the judge failed to stand back and consider whether, taken as a whole, the steps from Ericsson to claim 1 were obvious.

106. Ground 7a is that the judge erred in finding that the skilled person would ignore the teaching in Knuth that, in selecting D, the period should be much larger than the numbers that would actually be used. This concerns step 10.

107. Ground 7b is that there was no evidential foundation for the judge's finding that A=39827 was an obvious choice, particularly given his finding that the "max hits" metric for simulations was not common general knowledge. This concerns step 11. In oral submissions counsel for Optis did not pursue this ground, but did rely upon the point about "max hits" as part of his argument in support of ground 6.

Ground 1

108. On its face this ground appears to be a challenge to a primary finding of fact by the judge at [121]. That could only succeed if there was no evidence to support the finding, which is not the case. In their skeleton argument in support of the appeal Optis tried to get around that difficulty by arguing that the judge had misunderstood the distinction they were making. In the course of argument Birss LJ reformulated the submission in a better way. This is that, even accepting everything the judge says in [120], those points do not justify the breadth of the conclusion expressed by the judge at [121]. What each of the points made in [120] shows is that the skilled person would understand hashing functions to involve randomisation, as Optis accept. It does not follow that the skilled person would not think that hashing functions and RNGs were separate and distinct from one another. In response, counsel for Apple accepted that hashing functions and RNGs are different things, but argued that the skilled person would not regard them as conceptually distinct in a way which mattered in this context.

109. In considering these submissions I think it is important not to get bogged down by semantic points. The judge did not say that hashing functions were the same, or would have been perceived by the skilled person as the same, as RNGs, nor can he have meant that. He explained the basic difference between hashing functions and RNGs at [88]-[89], which I have reproduced in paras 22-23 above. As explained there, a

hashing function is a tool that operates on externally-generated data, whereas an RNG is an algorithm which generates a long series of random numbers from a single seed. Moreover, as I read [118]-[121] and [122]-[130], the judge accepted that the skilled person would understand that RNGs are usually recursive whereas hashing functions are not. What I understand the judge to have rejected is the proposition that the two would be perceived by the skilled person who was trying to improve Ericsson as being conceptually unrelated to one another. He found that the skilled person would understand that they are conceptually related in that a good hashing function involves randomisation of the data. Understood in that way, I see no error in the judge's conclusion.

110. I would nevertheless add that I do not understand the relevance of Chapter 6 of volume 3 of Knuth, which the judge cited at [120(v)], to this issue given that its contents were not found to be common general knowledge and that it would not have been found by the skilled person when consulting NRC3 for a suitable RNG as posited by Apple. This does not matter, however.

Ground 3

111. Optis' starting point here is a finding by the judge at [233] that the Ericsson function, although described as a hashing function, was "not of an existing type that was well-understood". Why then, Optis ask rhetorically, would the skilled person, having found that the Ericsson function did not have the desired properties, retain one part of the function and change another part when there was nothing in their common general knowledge to suggest that? In particular, why would the skilled person retain the mod C part when the problems with the Ericsson function were due to the distributive properties of the mod C operation, as the skilled person would appreciate? Optis accept that the skilled person would want a result between 0 and C-1, but say that there were other ways of getting that and Prof Lozano agreed with this. Optis contend that Prof Lozano's evidence, which the judge accepted, was based on hindsight, knowing that the claimed function included a mod C operation as well as an LCG (which itself includes a mod operation).

112. At first blush, this appears to be quite a powerful argument. It is therefore necessary to consider with some care Prof Lozano's evidence and the judge's acceptance of it.

113. In his first report Prof Lozano simply said at paragraph 206:

> "The skilled person would recognise that Ericsson was using mod *C* to get a number between 0 and *C*-1 and would consider that the obvious and natural way to achieve that end. They would also see that Ericsson is using K*x + L to generate a random number, seeking to select 'big enough' values for K and L to make the result random."

He did not explain why the skilled person would see this, and in particular why they would see that Ericsson was using K*x + L to generate a random number. On its face, it does not generate a random number.

114. The judge quoted at [223]-[224] two passages from Prof Lozano's evidence in cross-examination. The first is as follows:

> "So, my Lord, a hashing function has two purposes. One is to map a big number of inputs down to a smaller number of outputs, so there is a squeezing process, and the other one is to randomise these mappings so that two very similar inputs do not get mapped to two very similar outputs to minimise confusions. So the mod C, the outer mod C, is doing the squeezing down part of the hashing, and it is a standard way of doing it. It has been used in 3GPP before. It had been proposed already many months before the priority date to do this squeezing down by Motorola, and it is actually in Ms. Dwyer's report. So that is well understood, the squeezing down. The discussion here, and the work that was taking place, was around the randomisation part, the randomisation part. So the randomisation part here is K x+L, right, because the rest is the mod C which is doing the squeezing down, concentrating the many input into the few outputs. So randomisation is done by K x+L and is not doing a good job at that."

115. The second is as follows:

> "Well, like I said, the Skilled Person here is looking for something that randomises properly; that is it. That is all that is missing here. The rest is fine. The outputs are 0 to C-1 as they should be, so that part is functioning well. What is not functioning well is the randomisation part, so one would look to randomise things so you look at the book and see what it says about the number generators and pick an off-the-shelf solution."

116. The judge then said:

> "225. I accept this and think that while the skilled person would certainly have to think about the overall effect of the whole function, it would stand out clearly that the mod C part had the object and effect for which Apple contends. It would not require insight to retain mod C if possible (there is a specific point about using mod C which interfaces with the issues on the specified claims and with which I deal below).
>
> 226. Ms Dwyer came close to accepting much of this. She accepted that mod C would be seen as having the 'squeezing down' effect to which I have referred. She said '*I think people would look to change part of the equation, agreed, and the mod C does map the output to the range that is desired. So I think that if they were trying to keep it similar to the original format that is true.*'
>
> 227. I therefore accept Apple's contention that an obvious route was to retain mod C, on the basis that it was adequately performing a well-understood and necessary task, and look to remedy the

> problem, apparent at this stage, with the Kx + L randomisation part."

117. It may be observed that nowhere in the passages quoted by the judge does Prof Lozano explain how the K*x + L part of the Ericsson function produced a random number. Nor does the judge explain this. Nor was counsel for Apple able to explain it to this Court. It seems to me that it is the interaction between the K*x + L part and the mod C part which must have been intended by Ericsson to produce the randomisation desired of a hashing function, as I think counsel for Apple accepted.

118. Furthermore, the answer given by Ms Dwyer quoted by the judge does not assist Apple because it is based on the premise "if they were trying to keep it similar to the original format". Counsel for Apple also relied upon another answer given by Ms Dwyer, but that does not assist Apple either because it was merely an acceptance of what the skilled person could do in this respect. In any event, Apple's reliance upon Ms Dwyer's evidence is inconsistent with their reliance upon the judge's finding that she was technically out of her depth.

119. In my judgment, however, these points do not invalidate the judge's conclusion. This is for two reasons. First, the fact that the mod C part had a role in randomisation does not mean that it did not also have the function of "squeezing down" the inputs to 0 to C-1. On the contrary, it seems clear that it did have that effect. Secondly, I think there is a clue in Prof Lozano's statement that K*x + L was "not doing a good job" at randomisation. It seems to me that what he may well have meant was that K*x + L must have been intended by Ericsson, in combination with mod C, to produce randomisation, but failed properly to do so.

120. On that basis, I consider that the judge was entitled to conclude that an obvious way forward for the skilled person was to retain mod C and look for a better way to achieve randomisation than K*x + L, in combination with mod C, did.

Ground 4

121. Optis argue that, even if the skilled person decided to look for a better way to achieve randomisation than K*x + L, it does not follow that they would think of replacing K*x + L with an RNG rather than looking for a better hashing function (e.g. by consulting Wikipedia and in that way finding section 6.4 in Chapter 6 of Knuth). At best (from Apple's perspective) the skilled person would look for a better way to achieve the randomisation desired of a hashing function. Optis say that there was no precedent in the common general knowledge for using an RNG as the source of randomness in a hashing function. Optis point out that the judge's conclusion at [227] (quoted in paragraph 116 above) was that the skilled person would "look to remedy the problem … with the Kx + L randomisation part", yet the judge proceeded to record in [228] that it was Apple's case that the skilled person "would look in the literature for an appropriate RNG" and to accept that case. That, Optis argue, involved a leap and was tainted by hindsight.

122. It is convenient before addressing this argument to get out of the way a separate point made by Optis, which is to question why the skilled person would search the literature at this stage, rather than at an earlier stage. It seems to me that the judge's reasoning on this point is perfectly clear and justified, namely that the skilled person would turn

to the literature once they had identified what they were looking for. That would certainly be an obvious approach, even if other approaches were possible.

123. Again, at first blush the argument I have summarised in paragraph 121 above appears to be quite a powerful argument. Again, therefore, it is necessary to consider with some care Prof Lozano's evidence and the judge's acceptance of it.

124. In his first report Prof Lozano simply said:

> "210. Once the skilled person had verified that not all the desired properties are met [by the Ericsson function], they would look to replace *K*x+L* with a random number generator having the desired properties.
>
> 211. The skilled person would look for the random number generator in the literature, for instance, by turning to a standard reference text. …"

He did not explain why the skilled person would look to replace K*x + L with an RNG.

125. Having recorded Ms Dwyer's evidence that this is not what the skilled person would do since it was a leap, and her acceptance that it was nevertheless something the skilled person could do, the judge reasoned as follows:

> "230. Unsurprisingly, Counsel for Optis submitted that 'could' was not good enough. I agree that in itself it is not, but my task is to weigh the evidence of Prof Lozano who clearly said that is what the skilled person would do (there being other possibilities, of course, and I have to weigh that up as well), against the evidence of Ms Dwyer who would not go that far, although my sense at the time was that she was as close as may be to accepting 'would'.
>
> 231. I prefer Prof Lozano's evidence; I found him the more persuasive expert for reasons given in my overall assessment of the witnesses above. One sensible thing to do would be to look in the literature for an established and understood way to generate randomness. I think it would be the most natural way forward, and certainly one of the leading ones. It is the reliable, routine, systematic approach of the uninventive skilled person.
>
> 232. Prof Lozano was fair in putting this forward. He did not reject other options as being possible. Ms Dwyer's idea of modifying x was not really explored with him, but he was asked about the possibility of varying K and/or L by subframe, the idea put to him being that it would create more decorrelation between subframes. Prof Lozano agreed that changing K and/or L this way was something that the skilled person might do, and indeed it was discussed in RAN1 ...
>
> 233. I do not think the existence of such other options makes it any less natural or obvious that the uninventive skilled person would look for

> an established RNG. …"

126. I have to say that I do not find this reasoning convincing. It is one thing to say that an obvious choice for the skilled person would be "to look in the literature for an established and understood way to generate randomness" in place of K*x + L. It is another thing to say that "the uninventive skilled person would look for an established RNG". Even accepting, as I do, that the skilled person would perceive hashing functions and RNGs to be conceptually related in that they both involve randomisation, nowhere does the judge explain why it would be obvious to the skilled person to use an RNG for this purpose. The fact remains that hashing functions and RNGs are different, not least because RNGs are recursive whereas hashing functions are not.

127. As for Optis' point that there does not appear to have been any precedent in the common general knowledge for using an RNG to produce randomness in a hashing function, the judge addressed this at [262]:

> "The second [argument] was that it was unknown to use an RNG within a hashing function. … I reject the point because the Ericsson function itself, which is clearly a hashing function, while not explicitly expressing its teaching in terms of an RNG, used what the skilled person would recognise as a source of randomness (Kx + L), coupled with the mod C operation. It was also known, and taught in Knuth, to take only part of a large random number generated by an LCG by various means which included taking the least significant bits using the mod function, and this is essentially hashing. … "

128. Again, I have to say that I find this unconvincing. I agree with the judge that the Ericsson function is described as a hashing function, but as discussed above the suggestion that K*x + L would be recognised as a source of randomness, and therefore the skilled person would replace it with an RNG, is problematic. As for Knuth, the judge's reasoning depends on the skilled person having gone to Knuth to find out about LCGs and then appreciating, based on what is said in 3.2.1.1, that they can use an LCG to do something that is "essentially hashing" even though that is not what the text says in terms. But this is backwards: the question is why the skilled person would decide they should replace K*x + L in the Ericsson hashing function with an RNG in the first place.

129. As Prof Lozano accepted, Knuth does not suggest using an RNG as the source of the randomness desired of a hashing function (whether in Chapter 3 or Chapter 6). The nearest Knuth gets to this is that section 6.4 says (at page 515) that:

> " … it is not difficult to produce a pretty good imitation of random data, by using simple arithmetic as we have discussed in Chapter 3. And in fact we can often do even better, by exploiting the nonrandom properties of actual data to construct a hash function that leads to fewer collisions than truly random keys would produce.

> Consider, for example, [a hash function is described]. Experiments with actual data show, in fact, that this 'middle square' method isn't bad, provided that the keys do not have a lot of leading or trailing zeros; but it turns out that there are safer and saner ways to proceed, just as we found in Chapter 3 that the middle square method is not an especially good random number generator.
>
> Extensive tests on typical files have shown that two major types of hash functions work quite well. One is based on division, and the other is based on multiplication. …"

Leaving aside the question of why the skilled person would be reading section 6.4 in the first place if they are looking for an RNG, this is not a teaching to use an RNG in a hashing function. If anything, it is the opposite.

130. Nevertheless, I do not feel able to say that the conclusion the judge reached was not open to him. There is no error of principle on the judge's part, and there was some evidence to support his conclusion.

Ground 5

131. It is common ground that, if the skilled person got this far, an obvious thing to do would be to consult NRC3 to find a suitable RNG. Optis argue that the skilled person who did this would be firmly put off using an LCG, and therefore an LCG cannot have been an obvious choice. As Optis point out, NRC3 is emphatic: "Never use a generator principally based on … LCG". "An acceptable random generator must combine at least two (ideally, unrelated) methods." "… the field's long preoccupation with LCGs was somewhat misguided". Optis acknowledge that it also says that "LCGs … can still be useful, but only in carefully controlled situations, and with due attention to their manifest weakness". Read in context, Optis say, it is plain that what this is saying is that an LCG can only be used as part of a combined generator, and even then only with considerable care.

132. This is a powerful argument. How then did the judge reach the opposite conclusion? He reasoned as follows:

> "252. Although initially very striking, I think the statements in NRC are of low relevance, at most, to the issue of obviousness in this case. I accept Prof Lozano's evidence that what is under consideration in NRC is demanding situations where very long sequences of very random numbers are needed ('very random' in the sense that they pass extremely stringent tests intended to identify even the smallest signs of a pattern; an example was called 'Diehard'). He was clear that sometimes sequences as long as 10^{30} or 10^{40} were needed, and that for cryptography sequences were needed and were produced that were '*longer than the [age] of the universe measured in seconds*', but that that was '*way beyond what anyone at RAN1 would even think about. At RAN1 we have never seen sequences of more than a few thousand or maybe tens of thousands of repetition of period*.' He said that the

more sophisticated RNGs in NRC were '*way beyond anything that is required in a mobile device*'.

253. Accordingly I think the skilled person's attitude to LCGs would be that they were well known, widely used for a long time, easy to implement and suitable for low power devices, but not of good enough randomness for demanding applications. ….

254. Bringing this understanding to bear on the problem presented by Ericsson, the skilled person would know that there were about 65,000 UEIDs that needed to be distributed randomly, and that that was vastly lower than the scales relevant to cryptography etc. The skilled person would readily understand that LCGs were unsuitable if very large sequences of very random numbers were needed, but that that was not the relevant requirement for the PDCCH.

255. Prof Lozano explained this cogently in his written and oral evidence. …"

133. In my judgment this reasoning does involve an error of principle. As counsel for Optis submitted, the first question with respect to NRC3 is one of interpretation. The interpretation of NRC3, as with any other technical document, is a question for the court once educated as to the identity and attributes of the skilled person through whose eyes the document is to be read and as to their common general knowledge (including the meaning of any technical terms). Expert evidence is admissible, and usually essential, to assist the court to understand those matters; but it is not admissible, let alone determinative, as to the meaning of the document: see *Terrell on the Law of Patents* (19th ed) at 9-181 to 9-195 and the authorities cited. It is, of course, true that expert evidence is also admissible, and often vital, on the question of what the skilled person would think and do after reading the document, but that is a separate question: see *Terrell* at 9-179. With respect to the judge, I consider that he has elided the two questions here.

134. The judge did not identify anything in the attributes or common general knowledge of the skilled person which would cause them to read NRC3 in any special way. On the contrary, he found that LCGs were not common general knowledge, and so the skilled person reading NRC3 would be learning about them for the first time. The relevant passages are expressed in ordinary English, and are perfectly clear. The primary message is that one should never use an LCG on its own. The secondary message is that an LCG may nevertheless be used as part of a combined generator if sufficient care is taken. NRC3 does not say that these warnings only apply to demanding applications such as cryptography or that LCGs are useful for less demanding applications. It does say that one should avoid generators that are overengineered, and therefore wasteful of resources, such as generators designed for serious cryptographic use; but that is a distinct recommendation from the recommendation not to use an LCG. Nor does it say that LCGs are acceptable for periods shorter than a particular length. It does say that one should never use a generator with a period of less than 2^{64} (which is a point relied upon by Optis in relation to claim 4) and avoid using a generator with a period greater than 10^{100}; but again those are distinct recommendations from the recommendation not to use an LCG. (The fact that NRC3 also says "[b]e cautious about any source earlier than about 1995", yet (the second

edition) of Knuth is dated 1981 must be ignored for the reason explained in paragraph 54 above.)

135. The judge stated in [252] that he accepted Prof Lozano's evidence that "what is under consideration in NRC[3] is demanding situations where very long sequences of very random numbers are needed". No doubt Prof Lozano was both sincere and persuasive in his evidence, but it does not follow that his interpretation of NRC3 was correct, or even admissible. There is nothing in the text of NRC3 to support that interpretation.

136. Once NRC3 has been correctly interpreted, the next question is what the skilled person would think and do in the light of NRC3. The fact that NRC3 warns against using LCGs is not determinative of that question. In principle, it could still be the case that an obvious course for the skilled person would be to ignore those warnings and to decide that LCGs were nevertheless worth pursuing. But the skilled person would have to have reasons for ignoring the warnings based on their common general knowledge.

137. As I read the judgment, the judge's statement in [253] that "the skilled person's attitude to LCGs would be that they were well known, widely used for a long time, easy to implement and suitable for low power devices, but not of good enough randomness for demanding applications" is largely based upon his earlier finding at [248]:

> "… it was clear from Prof Lozano's evidence, accepted by Ms Dwyer and also supported by NRC and Knuth (and Wikipedia for what it is worth) that LCGs were very well known, had a long history, and were fast and easy to understand and implement."

138. It is indeed clear from NRC3 that LCGs were very well-known, had a long history, were fast and were easy to understand and implement, but NRC3 is equally clear that they nevertheless should never be used because they produce poor results. Thus what matters is whether the skilled person would think that LCGs were of good enough randomness for undemanding applications despite that warning. In so far as the judge's finding to that effect was expressed to be based upon NRC3, NRC3 simply does not support it for the reasons explained above. As for Knuth, the skilled person's postulated route to Knuth is via NRC3, and so the skilled person reads NRC3 first. Even supposing that the skilled person turns to Knuth after reading NRC3, and that the skilled person reads the whole of Chapter 3 rather than just section 3.5, Knuth gives the skilled person no reason to disregard the warnings about LCGs in NRC3. Moreover, Knuth teaches the skilled person about other forms of RNG. Wikipedia does not assist Apple for the reasons explained above. Ms Dwyer's evidence does not assist Apple for two reasons. First, as the judge recorded at [249], she disagreed that LCGs were common general knowledge; and as I have noted, the judge did not find that they were. Secondly, and in any event, as I have also noted, the judge found that Ms Dwyer was technically out of her depth.

139. As for Prof Lozano's evidence, it does not follow that the skilled person, who has never encountered an LCG before, would react to NRC3 in the way that Prof Lozano did. The only reason given by Prof Lozano for ignoring the warnings, and then only in cross-examination, was that the skilled person would think that the present application

was an undemanding one, and therefore would conclude that an LCG could be used despite what NRC3 says. But as discussed above that evidence was based on a misreading of NRC3. Why else would the skilled person think that this application was sufficiently undemanding that an LCG could be used without the hindsight knowledge that it does in fact work? In my view no sufficient reason is given by the judge for concluding that the skilled person would ignore the warnings in NRC3. Still less is any sufficient reason given by the judge for concluding that the skilled person would not merely ignore the warnings in NRC3, but follow up the cross-reference to Knuth to find out more about LCGs and decide to use one. The obvious thing for the skilled person to do would be to heed those warnings and to use the RNG that NRC3 recommends as being a "never-any-doubt" generator.

140. I would therefore allow the appeal on this ground. Even if the skilled person got this far, in my judgment it would not be obvious to them to use an LCG in the light of the warnings in NRC3 against doing so. Simply put, NRC3 teaches away from the invention. Thus claim 1 was not obvious over Ericsson.

141. For completeness, I should mention that Optis also have a point about step 7. This is that, even if the skilled person got this far, what they would do would be to go to Knuth and follow its teaching, particularly in section 3.6, with the result that they would simply use an LCG and not take mod C of that. This point was primarily deployed to support ground 3, which I have already rejected. If the judge was right about steps 2 to 6, then I cannot see that he made any error at step 7. The point was also deployed in support of ground 7a, which I will address below.

Ground 6

142. Having regard to the conclusion I have reached on ground 5, I can deal relatively briefly with ground 6. It is clear from the judgment that the judge was acutely conscious of the need to avoid a step-by-step analysis which involved hindsight, and that he was satisfied that he had avoided that trap. As is well known, however, hindsight is very difficult entirely to avoid even if one is conscious of the problem and tries hard to avoid it. This is true even for experienced patent judges. In my view Apple's case smacks strongly of hindsight. The claimed invention lies essentially in the use of a particular mathematical method. Once the utility of a mathematical method is known, it may be easy to suggest how it can be derived from first principles in a logical series of steps; but it does not necessarily follow that it would have been obvious to the skilled person before that. I am struck by the fact that Apple's case was first articulated by Prof Lozano in his first report in a series of steps without any real explanation as to why it would have been obvious to the skilled person to take some of those steps, still less as to why it would have been obvious to take all of them. Nor did Prof Lozano consider any other options, although he accepted that there were alternatives when this was put to him in cross-examination, as the judge noted at [232]. Furthermore, Prof Lozano's evidence was clearly coloured by the fact that he had relied upon NRC2, which did not include the warnings in NRC3, when forming his opinion. Counsel for Optis also submitted that Prof Lozano's reliance upon "max hits" in his first report in relation to the choice of A in claim 4 showed that his approach was tainted by hindsight because it could only have come from the Patent, albeit that Prof Lozano justified the choice of A in another way in a later report, but I am not persuaded that this adds anything so far as claim 1 is concerned. In conclusion, I doubt that this ground of appeal on its own would justify this Court in reversing the

judge's decision, but it lends support to the conclusion I have reached in paragraph 140.

Ground 7a

143. It follows that it is not necessary to consider this ground in any detail. It suffices to say that, if the judge was correct that claim 1 was obvious over Ericsson, I am not persuaded that he made any error of principle in concluding that the value of D in claim 4 was an obvious choice in the sense it would only involve routine work to determine.

Conclusion

144. For the reasons given above, I would allow the appeal and set aside the revocation of the patents in suit.

Postscript

145. Following circulation of the Court's judgments in draft Apple complained that the Court had not addressed two of their arguments on the appeal. The first was that the judge had accepted Apple's submission that, if there was a "prejudice" against LCGs, the Patent did not overcome it. I did not mention this argument because the judge did not find that the skilled person had a prejudice against LCGs (see [257]), and Optis did not argue on the appeal that he was wrong about that. On the contrary, Optis' case was that, given the judge's unchallenged finding that LCGs were not common general knowledge, the skilled person reading NRC3 would be learning about them for the first time. Furthermore, the skilled person reading the Patent has not (or at least not necessarily) read NRC3 first. The second argument is that the judge was entitled to rely in [252] upon Prof Lozano's evidence that "[a]t RAN1 we have never seen sequences of more than a few thousand or maybe tens of thousands of repetition of period". I would point out that I did consider the judge's reasoning in [252], as did Nugee LJ and Birss LJ. In case it is not obvious from what I have said above, however, I will add that this specific part of Prof Lozano's evidence does not assist Apple because it does not affect the skilled person's reading of NRC3, which is what matters.

Lord Justice Nugee:

146. I have had the great advantage of reading in draft both the judgment of Arnold LJ above and that of Birss LJ below.

147. I agree with them both that ground 1 should be dismissed.

148. I have been more troubled by ground 3. This concerns Optis' step 2: see paragraph 81 above where Arnold LJ summarises step 2, paragraph 102 where he summarises ground 3 and paragraphs 111 to 120 where he considers it. My concern was caused by the Judge having accepted Apple's case which included describing the K*x + L part as having "a random output", or as being the "random part" or the "randomisation part" of the Ericsson function. Thus in addition to referring to it in his judgment at [227] as "the Kx + L randomisation part" (set out at paragraph 116 above), he said at [222]:

> "222. …Apple's position was that the skilled person would realise that the mod C part of the Ericsson function was there to "squeeze" the random output of the (K*x + L) part down from a large number to a range from 0 to C-1, that the mod C part was therefore necessary and was working all right, and that any change should therefore be to the random, (K*x + L) part."

149. It is true that Professor Lozano's evidence described the K*x + L part in similar terms: see the evidence cited by the Judge at [223]-[224] (set out at paragraphs 114 and 115 above). He had said the same in his first report where he said (at [206]):

> "They [the skilled person] would also see that Ericsson is using K*x+L to generate a random number, seeking to select "big enough" values for K and L to make the result random."

150. Like Arnold LJ, however, the difficulty I have with these statements is that the output or result of K*x + L is anything but random if by that is meant, as I would understand it to mean, a scattering of the outputs showing no discernible pattern. K*x + L by itself does not do this, and it is, as Arnold LJ has said, only the interaction of the results of this part of the Ericsson function with the mod C part that produces, or was at any rate intended to produce, random outputs (see paragraph 117 above).

151. My concern was that characterising the K*x + L part as "the randomisation part" makes it easier to suggest that the mod C part, being there to squeeze down the outputs, "was working all right" (Judgment at [222]), or in Professor Lozano's words was "functioning well" (Judgment at [224]). That in turn makes it easier to accept that since what was not working well was the randomisation part, it was obvious that it was that part (and that part alone) that needed replacing, and that what was needed was therefore something that did the randomisation better. If instead one starts from the position that it is the interaction between the K*x + L part and the mod C part which was intended to produce random outputs, it does not seem anything like as apparent that what is wrong with the function and needs replacing is simply the K*x + L part. One might think that although the mod C part was no doubt doing a good job of squeezing the outputs into the desired range (0 to C-1), it was not working in terms of interacting with the K*x + L part.

152. Nevertheless the Judge undoubtedly had evidence to support his conclusion, and I do not think it can be said to be one that was not open to him. When Professor Lozano referred to the K*x + L part as the randomisation part, I think what he must have meant is that this part was intended to produce outputs to which the mod C function could then be applied so as to generate the desired random results. It was that part which was "not doing a good job" in the sense explained by Arnold LJ at paragraph 119 above. I am in the end therefore persuaded that ground 3 should be dismissed for the reasons given by him. And leaving aside ground 5 on which Arnold and Birss LJJ disagree, I do not wish to add anything on the other grounds. The issue, as Birss LJ says, is ground 5.

153. Faced with two well-reasoned judgments from very experienced patents judges which nevertheless lead to opposite conclusions, I have naturally not found it easy to decide this issue. But I will say straightaway that I have come to the conclusion, in

agreement with Arnold LJ, that the appeal should be allowed on this ground. I will try and explain why I have reached this view.

154. The critical finding of the Judge on this issue was that in [253]:

> "Accordingly I think the skilled person's attitude to LCGs would be that they were well known, widely used for a long time, easy to implement and suitable for low power devices, but not of good enough randomness for demanding applications."

This was of course in the context of a finding that LCGs were not CGK. That means that whatever the skilled person would have discovered about LCGs would have come from reading about them in their literature search. For the reasons explained by Arnold LJ at paragraph 58 above (with which Birss LJ agrees), Wikipedia can be left out of account for this purpose, leaving NRC3 and Knuth. So the question is whether the skilled person would have been able to derive from NRC3 or Knuth the knowledge or belief that LCGs were "suitable" although "not of good enough randomness for demanding applications". Or to put it another way, the belief that as long as your application was not demanding, LCGs would do a good enough job.

155. I agree with Arnold LJ that one cannot find this, or anything approaching it, in NRC3 (paragraph 134 above). There is certainly support there for LCGs being "well known" and "widely used for a long time": see the extracts cited by Arnold LJ at paragraph 62 above which refer to the idea of LCGs going back to the dawn of computing, to them being widely used in the 1950s and thereafter, and to the field's long pre-occupation with them. But none of this begins to suggest that they can still be suitable so long as the applications are undemanding. On the contrary, as shown by the extracts he cites, NRC3 is quite uncompromising about not using LCGs. The only exception (at the end of the passage cited in paragraph 62 above) is that the authors of NRC3 say that LCGs can be used in "carefully controlled conditions" but we were not shown anything to suggest that what was meant by that was anything other than that they could be used as part of a combined generator.

156. I also accept that Ms Dwyer accepted in cross-examination that the skilled person would discover that LCGs were "easy to implement because they have few operations that you have to carry out". And I agree that the fact that she was found by the Judge to be technically out of her depth does not prevent that part of her evidence nevertheless being evidence that the Judge could accept. Indeed it does not by itself seem surprising: it is easy to see that an LCG, as opposed, say, to a combined generator using two unrelated functions, has few operations to carry out and is therefore likely to be easy (in terms of computing power required) and hence fast to implement. So the conclusion of the Judge at [248] that:

> "LCGs were very well known, had a long history, and were fast and easy to understand and implement"

is one that was supported by the evidence and was open to him.

157. But fast and easy to understand and implement are not the same as suitable for undemanding applications, and I do not think we were shown anything to suggest that Ms Dwyer gave evidence that they were. The only evidence to that effect was from

Professor Lozano whose evidence was (or was thought by the Judge to be) that NRC3 was dealing with very demanding applications. The question is whether there was (or, as Birss LJ puts it, there could have been) any basis for this view.

158. This point was not dealt with in the reports at all because NRC3 was only discovered shortly before trial. So the only potential evidence in support of this conclusion was that given by Professor Lozano in cross-examination. There are three relevant exchanges as follows.

(1) First, from Day 1:

> "Mr Justice Meade: … You pointed us to 7.1.7 "When You Only Have 32-Bit Arithmetic".
>
> A. Yes, I am with you.
>
> Q. And it says, "get a better compiler! But if you ... must live [with it] here are some options. None of these individually pass Diehard." What does that last sentence mean?
>
> A. It is a test for randomness, but we are talking about extreme randomness here, my Lord. Sometimes people need to generate sequences that string along, like 10 to the 30, 10 to the 40. Cryptography, for instance, requires sequences that are longer than the edge of the universe measured in seconds, and one can do that actually. That is way beyond what anyone at RAN1 would even think about. At RAN1 we have never seen sequences of more than a few thousand or maybe tens of thousands of repetition of period."

(2) Also from Day 1, shortly afterwards:

> "Mr Abrahams: So you are referring to, on page 186, it has "LCG 6 Modulo 232"; yes?
>
> A. Yes.
>
> Q. You see, in the description of this particular method, it says "can use as random: not recommended; can use in bit mix: not recommended." Yes?
>
> A. Yes. I mean, like I say my Lord, the requirements that are being put here are extreme. I mean, they are way beyond anything that is required in a mobile device.
>
> Mr Justice Meade: Sorry, professor, what did you say, way beyond anything required for a what device?
>
> A. A mobile device. So there is this principle that people used to have to justify decisions, which is that perfect is the enemy

> of good, so I was being sought for solutions that got the job done and were as simple as possible."

(3) From Day 2:

> "Mr Abrahams: So I am right, am I not, that this book makes clear that you should never use an LCG outside of a combined generator; correct?
>
> A. No, later on that, as will be explained, LCGs and MLCGs can still be useful only in carefully controlled situations.
>
> Q. It makes clear that those carefully controlled situations are within a combined generator?
>
> A. Well, the thrust of these sections is on generating extremely long sequences as I explained yesterday, with periods there are many, many orders of magnitude above what anyone would need for RAN1. So anybody looking into solving the problem that we are discussing here would not be interested in this complicated generators, they would be looking for solutions, such as the ones given in Knuth."

159. These are evidently the passages that the Judge had in mind at [252] and we were not referred to anything else relevant in the evidence. I quite understand Birss LJ's "carbonara" point that it is possible for a text to say something in apparently absolute terms ("never use LCGs" or "never use LCGs except as part of a combined generator"), and yet for the skilled person reading the text to understand that "never" did not actually mean "never" but only "hardly ever", and that the text was really dealing with a more high-end situation than he himself was faced with. But I do not think that Professor Lozano ever quite said that the skilled person would read the strictures in NRC3 about not using LCGs as not really intended for less demanding situations.

160. In the first passage he is answering a question from the Judge about what the reference to Diehard is. There is no reason to doubt his answer, but he does not here suggest that the skilled person would have this knowledge. Far less does he suggest that the whole of the relevant discussion of RNGs in NRC3 is concerned with sequences of the complexity demanded in cryptography (that is, of the order of 10^{30} or 10^{40}), or that the skilled person would so understand it. Indeed if the skilled person reads the advice in section 7.1 of NRC3 (cited by Arnold LJ at paragraph 60 above) which contains the introductory advice as to what is and what is not acceptable in an RNG, he will there find advice to *avoid* using an RNG designed for serious cryptographic use, but also of course advice never to use an LCG, and never to use a generator whose period is less than 2^{64} (about 10^{19}). That combination of statements could not plausibly be understood by the skilled person as conveying the message that "never use LCGs" only applies to very long sequences of 10^{30} or more such as are used in cryptography, nor does Professor Lozano say that it could, or would, be so understood.

161. In the second passage Professor Lozano says, again no doubt correctly, that the requirements in the particular passage he is looking at (which it appears comes from section 7.1.7 of NRC3) are beyond anything required in a mobile device, and that it is a good principle not to let the perfect be the enemy of the good, so what is required is something that got the job done and was as simple as possible. But this does not explain how the skilled person would be able to conclude that an LCG would be satisfactory for his purposes, or even assert that he would. One might readily expect the skilled person to understand that an LCG would be simple – indeed NRC3 itself refers to the whole field of RNGs being "mesmerized, for far too long, by the simple recurrence equation"; and to "the impossible dream of an elegant 'single algorithm' generator": see the passages cited by Arnold LJ in paragraph 62 above. But that is not enough to explain how or why the skilled person could or would conclude that the LCG was not only as simple as possible but was also good enough and something that got the job done. What NRC3 teaches (this is also from section 7.1) is that:

> "It is now reasonable to expect to get "perfect" deviates in no more than a dozen or so arithmetic or logical operations per deviate, and fast, "good enough" deviates in many fewer operations than that."

So NRC 3 itself distinguishes between the perfect and the good enough, but gives no encouragement to the skilled person to think that an LCG, even if not perfect, is "good enough".

162. In the third passage Professor Lozano repeats his evidence from the first passage that "these sections" are dealing with complicated combined generators with periods of very high orders of magnitude. I do not think it materially adds to the first passage. There is evidence here that a skilled person, used to the periods in RAN1, might conclude that he did not need such complicated generators. But I do not think there is evidence in this passage that the skilled person would understand that when NRC3 said "never use LCGs" it did not mean that to apply to less demanding tasks.

163. I have looked at the evidence in detail because I am naturally conscious that the Judge had advantages that we can never have, expressed a clear preference for Professor Lozano's evidence, and reached clear conclusions having, as Birss LJ says, confronted the issues head on. There was as I have said evidence supporting his finding at [248] that LCGs were "very well known, had a long history, and were fast and easy to understand and implement." But I have not been able to find support in the evidence for his conclusion at [253] that the skilled person's attitude to LCGs would be that they were "well known, widely used for a long time, easy to implement and suitable for low power devices, but not of good enough randomness for demanding applications." That is dependent on his summary of Professor Lozano's evidence at [252] that "that what is under consideration in NRC is demanding situations where very long sequences of very random numbers are needed" and for the reasons that I have given I do not think the Professor ever said, or meant, that this applied to the entirety of NRC3.

164. For those reasons I agree with Arnold LJ that NRC3 teaches away from using an LCG. I accept that the skilled person would get from NRC3 to Knuth, but nothing in Knuth is I think sufficient to overcome this, nor did the Judge find that it was. I agree with the way it is put by Arnold LJ in paragraph 138 above that Knuth gives the

skilled person no reason to disregard the warnings about LCGs in NRC3. I therefore agree with him that Optis' ground 5 is well-founded and that the appeal should be allowed on this ground.

Lord Justice Birss:

165. I am grateful to Lord Justice Arnold for the thorough review of the judgment in this case. I agree with almost all of it, but I am unable to agree with the conclusion that the appeal should be allowed. I would dismiss this appeal. Given Arnold LJ's judgment, my reasons can focus on the point of disagreement.

166. I agree with all of my Lord's judgment explaining the background, the issues and the judgment below, and I agree with the conclusions that grounds 1, 3 and 4 of Optis' grounds of appeal should be dismissed. I also agree that ground 6 would not succeed on its own without ground 5, and that if ground 5 succeeds then so too would ground 7a against the specified claims. On the other hand, if ground 5 fails, then so does ground 7a.

167. The issue is ground 5. To recap, the starting point is that as a result of the failure of grounds 1, 3 and 4, the judge below was entitled to conclude that given the Ericsson prior art, the skilled person acting without invention, would retain the mod C part of the approach proposed by Ericsson and would look in the literature for an appropriate RNG with which to replace K*x + L. Moreover an obvious thing to do was consult NRC3 to find a suitable RNG, as explained from paragraph 130 onwards above. Thus far the argument has reached step 4 (paragraph 83 above). The question is – what happens next?

168. The relevant part of the judge's judgment is the section called "Getting to Knuth" at paragraphs 234 to 257. The judge's reasoning was in three parts. First: the skilled person would find their way to Knuth, most probably via NRC3. Second: the skilled person's attitude to LCGs would be assessed from NRC3, Knuth and "to a lesser extent" Wikipedia. Third: having done that, the skilled person would think that although LCGs did not give good enough randomness for demanding applications, this was not such a demanding application, and that LCGs had much to commend them and had been widely used for a long time. Therefore for those reasons, despite the clear negative statements about using an LCG in NRC3, nevertheless for the skilled person, using an LCG as the RNG was obvious. These three parts are aspects of steps 5 and 6, summarised at paragraphs 84-85 above. I agree that a route based on Wikipedia cannot assist Apple on this appeal. In other words if the appeal should be allowed in relation to the approach based on NRC3, then Wikipedia is irrelevant.

169. The critical point made by Optis is that NRC3 itself is unequivocal. It tells the skilled person that while LCGs had a long history and were widely used, they should no longer be used, at least on their own. The text of NRC3 does not contain any qualification of the kind found in the judgment. It does not state that LCGs can be used in less demanding situations. It essentially states they should never be used. In fact it does contemplate using an LCG in a two-part RNG but that is not good enough for Apple.

170. How then did the judge reach the conclusion he did? The first thing to note from an appellate point of view is that the judge confronted this problem head on. It was not

missed.

The first part – find their way to Knuth

171. The skilled person reads NRC3. It is a common general knowledge source. The judge held that the skilled person would then read Knuth. The judge did not find that the skilled person would go to Knuth looking for LCGs or because of what NRC3 may or may not have said about LCGs. On the contrary, the reason the skilled person goes from NRC3 to Knuth is because, as the judge recognised in paragraph 241, NRC3 expressly refers a reader interested in RNGs to Knuth in an entirely general manner, stating that "For references on this subject, the one to turn to first is Knuth". The "this subject" NRC3 is talking about is not LCGs, it is the whole topic in general. One attribute of the skilled person which the judge had also identified (at paragraph 231) was that looking in the literature was the skilled person's "reliable, routine and systematic" approach.

172. Therefore the clear negative words in NRC3 about LCGs do not undermine the judge's conclusion that the skilled person starting from NRC3 would read Knuth. And I believe the judge was entitled to conclude that the skilled person would get that far.

173. There is a curiosity about dates in that NRC3 says do not trust anything before 1995 and the version of Knuth in the court papers was from 1981, but that point goes nowhere because it was common ground at trial that there was a post-1995 edition of Knuth and the case proceeded on the basis that nothing turned on using the earlier version.

174. Again at the risk of labouring the point, so far LCGs have played no part in the reasoning. I also therefore disagree with Arnold LJ that the skilled person's attitude to LCGs will be based solely on NRC3. They have got to Knuth without focussing on LCGs. Nor does it require any reading of NRC3 in a special way. No hindsight is involved.

The second part – what is the basis for whatever it is the skilled person now thinks about LCGs?

175. The judge held that the skilled person would find out that LCGs were a kind of RNG. In paragraph 239, and now ignoring Wikipedia, the judge held that the skilled person's attitude to LCGs would be assessed from NRC3 and Knuth (see also paragraph 248). I do not agree that the judge has made the error of forgetting that LCGs as such were not common general knowledge. The judge's conclusion is that what the skilled person learns about LCGs comes from these sources – NRC3 (which is common general knowledge) and Knuth (which they would read looking for RNGs).

176. I therefore disagree with the start of paragraph 138 above in which the finding of the judge at paragraph 248, that LCGs were fast and easy to use and understand etc., is characterised as being formed at a stage of the analysis before the skilled person read Knuth. That is not what the judge did and in my judgment the reasoning that got this far was open to him without hindsight.

The third part: what does the skilled person think about LCGs?

177. The judge identified Prof Lozano's evidence as being that the skilled person would think that what was under consideration in NRC3 was very demanding situations, and (in paragraph 253) the judge concluded that what the skilled person would think about LCGs was that they were well known, widely used for a long time, easy to implement and suitable for low power devices, but not of good enough randomness for demanding applications.

178. Just focussing on "well known" briefly, of course LCGs were not well known to the skilled person before they read NRC3 (and the judge did not find that they were). But having read these two sources the skilled person can see that LCGs are in fact well known to those involved with RNGs even though the skilled person had never previously heard of them.

179. The judge's approach was to accept the evidence of Prof Lozano about what the skilled person would think about LCGs. Part of this involves the skilled person thinking that what was under consideration in NRC3 was very demanding situations unlike the one they were faced with.

180. As my Lord explains, it is well established that construction of a document is for the court, but expert evidence about what the skilled person would think in the light of a document, is admissible expert evidence. I accept that in one place the judge's language could be read as making that error but as I shall explain below I do not believe that the judge, with long experience of patent law, fell into the trap of eliding these two things.

181. One question is about what reason is there or could there be for the Professor's opinion. I say "could there be" because I have not had my attention drawn to any cross-examination of the Professor in which he was asked the rhetorical question posed in paragraph 139 above. In other words: what reason is there, absent hindsight and on the premise that LCGs are not part of their common general knowledge before reading NRC3, for the opinion that the skilled person would think LCGs could be used in undemanding applications whereas NRC3's prohibition was about more demanding cases? There was a generalised point put in cross-examination that the Professor's evidence was tainted with hindsight but he did not accept that. As a result, on appeal, there is some difficulty. If one could be satisfied that the only possible reason for this view was hindsight then an appeal court would be right and entitled to overturn the finding that that evidence should be accepted. However if there is another reason, untainted with hindsight, then the position is different.

182. The professor's evidence might have been based on hindsight but in my judgment that is not necessarily so. Imagine a modern recipe book by a famous celebrity chef. It contains a recipe for spaghetti carbonara and states that the only ingredients should be spaghetti, bacon, egg and parmesan. The recipe book might observe that in the 1970s many British home cooks made carbonara sauce for spaghetti by mixing bacon and supermarket cream cheese, but this was very poor, and today it should never be used in any circumstances.

183. An expert might well give evidence that the thinking of a skilled person reading the book, for whom the cream cheese method was not part of their common general

knowledge, was that although the text is clear that cream cheese should not be used, they would think that the author was really focussed on high end restaurant dishes and cooking to impress. Therefore, they might think, for someone looking to make themselves a simple meal at the end of a busy day using cream cheese would suffice. I am not trying to make a perfect analogy. My point is simply that it does not require hindsight to explain that a reader of a book might, in context, take a softer view of what is an unequivocal prohibition in the text.

184. Moreover the skilled person reads Knuth as well as NRC3. And that provides clear support for the Professor's position. For example we were shown a passage of cross-examination in which he was pressed on the point that even for what NRC3 called "quick and dirty" applications LCGs did not meet what NRC3 regarded as the minimum level of randomness, which was to pass a test for randomness called Diehard. The skilled reader would therefore understand why NRC3 only contemplated using an LCG in combination. The Professor's response was that the thrust of those sections of NRC3 was on generating extremely long sequences with periods much longer than what anyone would need for RAN1, that the skilled person would look to Knuth, and that Knuth explains that LCGs are a very good solution for low complexity applications (Day 2 p106-109). For example in this passage the Professor said:

> "So anybody looking into solving the problem that we are discussing here would not be interested in [*these*] complicated generators, they would be looking for solutions such as the ones given in Knuth."

185. The "complicated generators" referred to are the ones in NRC3. Now this view may be right or wrong, but it is in my judgment admissible evidence of the skilled person's thinking given NRC3 and Knuth, and it does not depend on prior knowledge of LCGs.

186. Turning to the thinking about the task in hand, the idea is that each phone in a cell has to have a different pattern. In a cell the phones each have a numerical user ID ("UEID"). It is a sixteen bit number and so there are 2^{16} of them (i.e. about 65,000). As judgment paragraph 254 explains that is vastly lower than the scales relevant to "cryptography etc." (by which the judge was referring back to paragraph 252 and sequences lengths of 10^{30}-10^{40} or more). The other aspect is that C varies between 0 and 95 (paragraph 255), so the largest mod value is 95. The judge had evidence from the Professor that the skilled person would consider it unhelpful to work with larger numbers than are necessary for reasons of computational simplicity.

187. Overall, I believe the judge was entitled to reach the conclusion he did. He gave ample reasons why the skilled person's thinking, having been informed by NRC3 and Knuth, was not to accept the clear NRC3 prohibition on using LCGs alone, but in fact to see them as useful for the application the skilled person had in mind. That is enough to reject ground 5.

188. However there was also further relevant evidence, of Optis' own expert Ms Dwyer. As the judge noted in paragraph 249, Ms Dwyer had actually accepted that if the skilled person did want to use an RNG they would look it up and would come across LCGs as one of the categories, and would find out that LCGs were fast, easy to implement and easy to understand. Ms Dwyer also did not agree that NRC was a

common general knowledge source but that is now irrelevant because the judge found that it was.

189. Ms Dwyer's evidence does also assist Apple and I do not agree with the contrary view expressed at paragraph 138. The fact Ms Dwyer's view was that LCGs were not themselves common general knowledge is not what paragraph 249 is concerned with. That view does not undermine the utility for Apple of her acceptance that if a skilled person did look up RNGs (as the judge was entitled to find that they would), then they would find out about LCGs, and would find they were easy, fast etc.. The premise which Ms Dwyer did not accept was that any of NRC or Knuth or Wikipedia were themselves common general knowledge, but as I say the judge had found that NRC3 was a common general knowledge source.

190. The other reason given at paragraph 138 why this part of Ms Dwyer's evidence was not of help to the judge was because she was out of her depth. I cannot reach that conclusion on appeal. The judge did not find that nothing whatever of Ms Dwyer's evidence was of any utility. He did find her evidence was of extremely limited help on the key issues but that does not preclude the judge putting some weight on a specific part of her evidence despite her limitations as a witness, which the judge clearly had in mind. That is the kind of nuanced approach to the evidence that is a matter for a trial judge. An appellate court cannot gainsay it simply because the judge had also held that the witness's evidence had extreme limitations.

191. I have also had the advantage of reading Nugee LJ's judgment in draft. There is only one point I would add to what I have said above. The skilled person is not a blank slate who only acquires knowledge from the cited text. What the skilled person thinks is the result of the combination of their pre-existing common general knowledge, their thinking about the problem to be solved, and what they acquire from the text (which in this case is two texts NRC3 and Knuth together). The finding that LCGs were not common general knowledge was not a conclusion that the skilled person lacked any mathematical skill or insight. On the contrary the judge rejected Optis's attempt to characterise the skilled person as uncomfortable with modular arithmetic, hashing functions or random numbers ([37]). Rolled up in the crucial finding at [253] is a view that the skilled person would be able to apply their skill and form an opinion about how numerically and computationally demanding the task they had was relative to other situations, such as those referred to in these texts. Paraphrasing, the conclusion that LCGs would be suitable included forming opinions: (i) that the task was not so numerically demanding in terms of sequence lengths etc., (ii) that whatever was used needed to be computationally feasible on a mobile telephone, and (iii) that the strictures in NRC3 (which would be understood as unqualified) nevertheless would not put the skilled person off using an LCG. As I have already said, I can see no error in the judge's approach to this.

Reflections

192. The judge's approach is a step by step one and when I first read the judgment I must say I was struck by that and wondered if I would have reached the same one if I had been the trial judge. However the Supreme Court in ***Actavis v ICOS*** (cited above) makes clear that inventions can be obvious based on a step by step approach provided of course those steps are not driven by hindsight. The judgment below is careful and closely reasoned. The judge was well aware of what he was doing and considered the

matter with care both individually and then stepping back and looking at the exercise as a whole. There are reasons in support of each step absent hindsight and I believe there is no ground for this court to interfere.