inevitably reflect our own experience. For example Curry's work will be given more space than Church's simply because we know it better (and the reader probably does not), not because we think it more important.

A disclaimer on priorities: the question of who invented an idea first is very natural to ask, but often very hard to answer: important ideas may be invented independently almost simultaneously, and in this field several key papers were circulated informally for some years before being published. So when we give a publication date below, we should not be interpreted as asserting a priority, merely stating a historical fact.

1.1 Acknowledgements

For useful information the authors are very grateful to Peter Andrews, Corrado Böhm, Steven Givant, Peter Hancock, Per Martin-Löf, Bill Lawvere, John Shepherdson, Jonathan Seldin, Christian Thiel, Philip Wadler, and [??] and many other helpful colleagues. For some useful general advice we are very grateful to the mathematical historian Ivor Grattan-Guinness.

This article has also been much helped by material from other published historical works, especially Barendregt's [29], Crossley's [113], Curry and Feys' [133, §§0D, 1S, 2S, 3S, etc.], Gandy's [179], van Heijenoort's source-book [211], Hodges' [220], Kalman's [241], Kleene's [253], Laan's [267], Manzano's [318], Nederpelt's and Geuvers' [356], Rosser's [415], Seldin's [448, 449], and Sieg's [452].

However, any errors in what follows are our own responsibility.

Last but not least, we express our gratitude for financial support and facilities from the Dipartimento di Informatica of the University of Turin, and the British Council, which made a crucial consultation visit possible in May 2000.

2 Pre-history

Notations for function-abstraction and substitution go back at least as far as 1889. In Giuseppe Peano's book on axioms for arithmetic [370, §VI], for any term α containing a variable x, the function of x determined by α was called $\alpha[x]$, and for $\phi = \alpha[x]$ the equation $\phi\, x' = \alpha[x]x'$ was given and its right-hand side was explicitly stated to mean the result of substituting x' for x in α. Later, Peano used other function-abstraction notations instead of $\alpha[x]$; these included $\alpha\,\overline{x}$ and $\alpha|x$, in [372, p.277] and [371, Vol.3 §11] respectively.

In 1891 Gottlob Frege discussed the general concept of function and introduced the notion of a function as a graph (*Wertverlauf* in [169]). Two years later, notations for abstraction and application appeared in his [170, Vol.1 §9], where the graph of a function $\Phi(\;)$ was denoted by $\grave{\varepsilon}\;\Phi(\varepsilon)$, and the function's application to an argument Δ was denoted by $\Delta \cap \grave{\varepsilon}\;\Phi(\varepsilon)$. Frege also proved the equation $a \cap \grave{\varepsilon}\; f(\varepsilon) = f(a)$ [170, Vol.1 p.73, §54].

Notations for class-abstraction too appeared at around the same time. Peano denoted the set of all x satisfying condition α by $[x\,\varepsilon]\,\alpha$ in his 1889 work [370, §V], and by $\overline{x\,\varepsilon}\,\alpha$ from 1890 onward, see [372, Vol.2 p.110]. The latter notation was also used by his colleague Burali-Forti, see [70, p.71 §3]. Another class-abstraction notation, $\hat{z}(\phi\, z)$, appeared in 1913 in Whitehead and Russell's [418, Ch.1]. The same authors also used $\varphi\,\hat{x}$ for the unary propositional function determined by $\varphi\, x$.

However, none of these authors gave formal definitions of substitution and conversion, which are characteristic of CL and λ proper. These did not appear until the 1920s, and, although λ might seem to express the concepts involved more directly than CL, it was actually CL that was invented first.

of avoiding bound variables, and a finite axiomatization of set theory.

4 1930s: Birth of λ and Youth of CL

4.1 Early λ-calculus

The λ-calculus was invented in about 1928 by Alonzo Church, and was first published in 1932 in [73]. Church was born in 1903 in Washington D.C., and studied at Princeton University, though in 1928–29 he visited Göttingen and Amsterdam.[12]

Around 1928 he began to build a formal system with the aim of providing a foundation for logic which was more natural than Russell's type theory or Zermelo's set theory, and which would not contain free variables (for reasons he explained in [73, pp. 346–347]). He chose to base it on the concept of function rather than set, and his primitives included abstraction $\lambda x[M]$ and application $\{F\}(X)$ (which we shall call here "$\lambda x.M$" and "(FX)").

(By the way, why did Church choose the notation "λ"? It was said by Rosser in [415, p.338] to have come from the notation "$\hat{z}(\phi z)$" used for class-abstraction by Whitehead and Russell, by first modifying "$\hat{z}$" to "$\wedge z$" to distinguish function-abstraction from class-abstraction, and then changing "$\wedge$" to "λ" for ease of printing. On the other hand, two enquirers (one was Seldin in 1982) were told by Church at different times that the choice was more accidental: a letter was needed and "λ" just happened to be chosen.)

As mentioned earlier, Church was not the first to introduce an explicit notation for function-abstraction. But he was the first to state explicit formal conversion rules for the notation and to analyse their consequences in depth.[13]

Church's system as it appeared in 1932 in [73] was a type-free logic with unrestricted quantification but without the law of excluded middle. Explicit formal rules of λ-conversion were included. Positive integers were represented by pure λ-terms and the Peano axiom for induction was proved, though no further arithmetic was developed. However, almost immediately after publication a contradiction was found in it. The system was revised in 1933, [74], and at the end of the revised paper Church stated a hope that Gödel's recent incompleteness theorems for the system of Whitehead and Russell [418] did not extend to his revised system, and that a finitary proof of the consistency of this system might be found [74, pp. 842–843]. (This hope seems naïve nowadays, but at that time the full power of Gödel's theorems was only slowly becoming understood.)

From 1931 to 1934 in Princeton University, Church received the help of two outstanding graduate students, Stephen Kleene and Barkley Rosser, and in a remarkable four years of collaboration the three made a series of major discoveries about both Church's 1933 logic and the underlying pure λ-calculus. Unfortunately for the 1933 logic, one of these discoveries was its inconsistency; a variant of the Richard paradox was proved by Kleene and Rosser in [256].

In contrast, the pure λ-calculus, which had at first seemed shallow, turned out to be surprisingly rich. (Well, Kleene was surprised, anyway, [253, p.54].) The study of its reduction-theory was begun using the method of residuals, which was applied to prove the Church-Rosser confluence theorem [84, p.479], and this ensured the consistency of the pure system. The system's relation to combinatory logic was also clarified, see Rosser's [414], and a simple proof of the confluence of combinatory

[12] For more on Church's life, see M. Manzano's [318, §2] and H. Enderton's obituary [154]. For more on his work, see [318], [13], [155], [452], and (on the impact of λ-calculus) [29]. He died in 1995.

[13] By the way, Church did not read Frege's work until 1935 (letter, 1979, cited by D. Scott in [438, p.260]). We assume he did not see Peano's 1889 notation either.

Church's own interests underwent some evolution after his 1941 book [80] appeared. Much of his time and energy was spent on the business of the Association for Symbolic Logic, of which he had been a co-founder and in which he was still a driving force, and on editing the Journal of Symbolic Logic, especially its very comprehensive Reviews and Bibliography sections. On the research side he became interested in intensional logic and the logic of relevant implication. For the former he made axiomatizations in [81] in 1951 and [83] in 1973, both based on modifications of his type theory. For relevant implication he gave a formulation, in 1951 in [82], which was essentially equivalent to the simple type theory of pure λI-terms.

5.2 Abstract reduction theory

Probably the first publication on abstract reduction-theory was a 1910 paper by the mathematician Axel Thue on rewriting finite trees, [479]. Thue's paper even included a lemma that local confluence implies confluence if every contraction strictly decreases a finite measure, [479, §4 Satz 3], anticipating what is often nowadays called Newman's Lemma [362, Thm. 3].[22]

But the main inspiration for later abstract work in this field was Newman and his 1942 paper [362]. Some years previously, Newman had proved a crucial confluence conjecture in combinatorial topology, and his interest in logic (and probably his link with Turing) had led him to study the Church-Rosser confluence theorem in [84] and the method of residuals used in its proof.[23] In his [362] he analysed confluence proofs in a very general setting, and deduced confluence from a set of abstract axioms about the concept of residual. Unfortunately, due to an error involving variables, his axioms failed to cover λ as a special case. This was corrected about 20 years later by David Schroer in his unpublished thesis [424] (which also included the finiteness-of-developments theorem for $\lambda\beta$.) Newman's analysis led in the 1970s to general methods for proving confluence and other key theorems such as standardization, for systems of λ augmented by extra operators, for example a primitive recursion operator. (A summary of these results is in [217, pp.269–270].)

Other abstract studies of confluence were Barry Rosen's [411] in 1973 and Gérard Huet's [223] in 1980. But in the 1980s the abstract theory of reductions was reorganised completely and developed by Jan-Willem Klop in his thesis [258], and from this work the abstract theories of *combinatory reduction systems* and *term rewriting systems* evolved. (For surveys see [259] and [260].)

5.3 Reductions in CL and λ

Besides abstract reduction theory, there was also progress in analysing the two most important concrete reductions, weak reduction in CL and β-reduction in λ. As mentioned in 4.1 above, Church and Rosser had begun the study of these in the 1930s, see [84] for example.

After 1950 their programme was continued by Curry. His research had been suspended completely during the Second World War, like that of many other academics, and did not recover momentum until quite late in the 1940s. Then for about 20 years he put his main effort into compiling two large monographs containing much new work, [133] in 1958 and [134] in 1972. Part of this effort consisted of a detailed study of $\lambda\beta$-reduction, including the first proof of its standardization

[22] See M. Steinby's and W. Thomas' [459], which contains an English summary of [479].

[23] If a term X contains redex-occurrences $R_1, ..., R_n$, and $X \triangleright_1 Y$ by contracting R_1, certain parts of Y can be considered to be descendents of $R_2, ..., R_n$ in a sense; they are called *residuals* of $R_2, ..., R_n$; for details in λ, see [133, §§4B1–2] or [219, Def. A1.3], or [84, pp.473–475]. If we reduce Y by contracting one of these residuals, and then a residual of these residuals, etc., the reduction $X \triangleright_1 Y \triangleright ...$ is called a *development* of $R_1, ..., R_n$.

theorem, [133, §4E1]. He also made the first study of $\lambda\beta\eta$-reduction, including proofs of its confluence and η-postponement theorems, [133, §4D]. (But his proof of the latter contained a gap, which was noticed and remedied in 1973 by Robert Nederpelt, [355, Thm. 7.28].)

Weak reduction in CL was given a systematic analysis in [134, §11B].

Curry also defined a *strong reduction* for CL analogous to $\lambda\beta\eta$-reduction, [133, §6F].[24] His purpose was to make later discussions of logic applicable equally to languages based on CL and on λ. Strong reduction indeed allowed this to be done, in [134], but its metatheory turned out to be complicated, and although simplifications were made by Bruce Lercher and Roger Hindley in the 1960s which produced some improvement, it eventually seemed that wherever strong reduction could be used, it would probably be easier to use λ instead of CL. Strong reduction was therefore more or less abandoned after 1972. However, in the 1980s a β-strong reduction was defined and studied by Mohamed Mezghiche, see [341], and the idea of finding a combinatory reduction which is preserved by abstraction but avoids the complexity of bound variables is still tantalisingly attractive.

5.4 Illative systems

Outside Church's group the study of applied CL and λ was pursued in the 1950s mainly by Curry and Fitch. Fitch continued the system-building programme described earlier, but his systems seemed to have relatively little effect on the ideas of other logicians.

Curry's work on applied systems (which he called *illative*, from the Latin "illatum" for "inferred") was published in the books [133] in 1958 and [134] in 1972.

Perhaps his most important contribution in the illative sections of the 1958 book [133] was to introduce Gentzen's techniques into type theory. As we mentioned in §4.2, this approach was not as obvious in 1958 as it is now. Curry had emphasized the importance of Gentzen's methods in a series of lectures on first-order logic at the University of Notre Dame in 1948, and in [133, §9F] he defined a sequent-based system of type-assignment. In [133, §9F4] he proved a cut-elimination theorem for this system, from which he deduced the first published proof of the WN theorem for simple type theory, [133, §9F6, Corollary 9F9.2].[25] Other type-theoretic properties described in this book included the propositions-as-types correspondence in [133, §9E], see §4.2 above and §6.1.4 below, and the link with grammatical categories in [133, §8S2], see §6.1.1 below.

In the 1972 book [134], Curry's main aim was to make a thorough analysis of basic logical concepts in the settings of λ and CL, with as few restrictions as possible and an emphasis on constructive methods.[26] The book contained a detailed comparison of logical systems based on three different concepts: (i) F, the functionality constant (§4.2 above), (ii) Ξ, where ΞAB was equivalent to $(\forall x)(Ax \supset Bx)$, and (iii) $\forall$ and $\supset$. The formulation of these systems benefitted greatly from results and simplifications introduced by Jonathan Seldin in his thesis [446] in 1968, but many variants of each system were discussed and the resulting treatment was rather diffuse.

In 1969 in his thesis [69], Martin Bunder developed an illative system in which the axioms of Zermelo-Fraenkel set theory could be deduced.

[24] In CL, $\beta\eta$-strong reduction $\succ\!\!-_{\beta\eta}$ is defined by adding to the rules defining weak reduction the extra rule (ξ): *if* $X \succ\!\!-_{\beta\eta} Y$ *and* x *is any variable, then* $[x]_{abcf}.X \succ\!\!-_{\beta\eta} [x]_{abcf}.Y$, where $[x]_{abcf}$ is evaluated by abstraction-algorithm (abcf).

[25] Curry's WN theorem was for CL with $\beta\eta$-strong reduction. It was deduced via a λ-CL translation from his cut-elimination theorem, which was for $\lambda\beta$.

[26] For further comments see [448, pp.22–27].