CrossMark

ORIGINAL PAPER

On Quine on Arrow

Maurice Salles^{1,2,3}

Received: 23 January 2017 / Accepted: 17 March 2017 / Published online: 28 March 2017 © Springer-Verlag Berlin Heidelberg 2017

Abstract This paper describes an unknown episode in the development of the theory of social choice. In the Summer 1949, while at RAND, Quine worked on Arrow's (im)possibility theorem. This work was eventually published as a paper on (applied) set theory totally disconnected from social choice. The working paper directly linked to Arrow's work was never published.

I alluded to this (then unwritten) paper in a number of presentations I made on 'Logic and Social Choice' in Turku, Bucharest, Boston, Strasbourg and Munich, between October 2013 and January 2015. It was eventually first presented during a conference at Queen Mary, University of London, 19–20 June 2015, on 'Social Welfare, Justice and Distribution: Celebrating John Roemer's Contributions to Economics, Political Philosophy and Political Science', organized by Roberto Veneziani and Juan Moreno-Ternero. I am grateful to the participants for interesting reactions and comments, in particular Richard Arneson, Jon Elster, Marc Fleurbaey, Klaus Nehring and Gil Skillman. Jon Elster contacted Dagfinn Føllesdal, a well-known philosopher and a pre-eminent Quine scholar, who kindly responded to some queries. A more developed version was presented in Aix-en-Provence during the International Conference on Economic Philosophy and in Lund during the meeting of the Society for Social Choice and Welfare in June 2016. Comments of participants to these two events revealed to be very helpful, among which comments by Gilles Campagnolo, Christian List and John Weymark. While in Lund, I also greatly benefitted from conversations with Adrian Miroiu. Finally, I am very grateful to an Associate Editor of this journal for excellent suggestions and for detecting some very annoying slips.

Murat Sertel Center for Advanced Economic Studies, Bilgi University, Istanbul, Turkey



Maurice Salles
 maurice.salles@unicaen.fr

CREM (UMR-CNRS 6211), University of Caen-Normandy, 14032 Caen cedex, France

² CPNSS, London School of Economics, Houghton Street, London WC2A 2AB, UK

1 Introduction

In August 1985, while I was participating in the world meeting of the Econometric Society at MIT in Cambridge (Mass.), I had a browse in the MIT bookshop and came across the just published autobiography of Quine (1985). As many social choice theorists, I had (and still have) a side interest for logic and, by extension, for analytic philosophy in general. Then, I decided to get a copy and read the book. Reaching page 217, I discovered an astonishing fact: Quine already knew Arrow's work on social choice in 1949 and even made some research based on it (Quine 1949a, b, c). It is remarkable that a scientific link could be established between two of the major thinkers of the contemporary world of knowledge. However, in Quine (1985) one can read:

'Rand Corporation was an intelligence facility in the service of national defense. It was agreeably situated in Santa Monica, a coastal suburb of Los Angeles. My summer job as consultant there was unprecedentedly remunerative, but apart from that it was a mistake. Despite my top-secret clearance in the Navy, fresh clearance was required-reasonably enough, since a staunch anti-Nazi could be a communist. My new clearance did not arrive in time, so I was put on boondoggles. One of them concerned Kenneth Arrow's monograph on social reconciliation of individual preferences. My resulting memoranda included two theorems about Boolean functions, ultimately published in *Selected Logic Papers*. The other project was in game theory. My memorandum on this was subsequently incorporated by McKinsey and Krentel into an article under our three names.'²

It seems that the word 'Boondoggle' is rather disparaging. In another (shorter) autobiography (Quine 1986) Quine speaks of 'innocuous projects', which seems to me more appropriate.³ Anyway, boondoggle or not, Quine in 1966 judged that the paper was sufficiently interesting to be published in his *Selected Logic Papers*.

In 1986 I wrote to Quine and, to my great surprise, received a hand-written response in French. I asked Quine whether he had ever worked on Arrow's theorem after his 1949 stay at Rand corporation. He replied: 'Malheureusement je ne me suis plus occupé du résultat de Kenneth Arrow depuis cet époque tellement éloigné' (his French was excellent except that 'époque' is a feminine word and, accordingly, the correct spelling should have been 'depuis cette époque tellement éloignée'). With his letter, Quine

⁴ 'Unfortunately I never dealt with Kenneth Arrow's result since this so remote time.'



¹ A non-Anglo-Saxon mainstream economist may have never heard of Quine. Although I will not endorse this type of ranking, Quine was ranked fifth in a ranking of the most important philosophers of the past 200 years-after Wittgenstein, Frege, Russell and Mill, but before Kripke, Nietzsche and Marx. This ranking is based on a survey by Brian Leiter, an American philosopher and Professor of Jurisprudence at the University of Chicago. See: http://leiterreports.typepad.com/blog/2009/03/so-who-is-the-most-important-philosopher-of-the-past-200-years.html.

² See Quine (1966, 1995) for the *Selected Logic Papers* and Krentel et al. (1951) for the game theory paper. McKinsey was a mathematical logician and game theorist located at RAND in 1949 and later at Stanford. He authored one of the first textbooks on game theory (McKinsey 1952).

³ Although the shorter autobiography was published later than *The Time of my Life*, it was, according to Quine, written several years before.

was so kind as to send me his page-proofs (with his hand-written corrections) of the relevant chapter in his *Selected Logic Papers*.⁵

I subsequently left this question aside until I recently decided to work on a partly historical monograph on logic and social choice.

I will first present in this paper the chapter *On Boolean functions* as it is in the *Selected Logic Papers*. Then I will describe the Rand memoranda related to Arrow's impossibility theorem.

2 On Boolean functions

It is very difficult to perceive how Chapter XVI of *Selected Logic Papers* is related to Arrow's impossibility theorem. Neither Arrow nor the aggregation of individual preferences are mentioned and there are no references. We are left somewhat in the dark with the vague phrase 'utility concept' in the first sentence:

'As auxiliaries to a concurrent study of the utility concept, two theorems of set theory are proving useful. The envisaged application of these theorems concerns only relations, or sets of ordered pairs, whereas the theorems hold for sets generally. Consequently the two theorems are most conveniently set forth separately from the eventual memorandum in which they are to be applied.'

The two theorems are, as Quine said, theorems of set theory, but I do not know whether they have been used or applied somewhere.

A *Boolean function* is defined as a set which is specifiable in terms of given sets by means exclusively of intersection, union and by taking complements.⁶

Quine then proposes to link general functions taking sets as arguments (variables) and as values to Boolean functions.

Precisely, let ϕ be any *n*-ary function.

Definition 1 The function ϕ is said to be a parametric Boolean function if there are sets H_1, \ldots, H_r for some r and an (n+r)-ary Boolean function ψ such that for all sets F_1, \ldots, F_n ,

$$\phi(F_1, \ldots, F_n) = \psi(H_1, \ldots, H_r, F_1, \ldots, F_n)$$

Two sets are said to agree in an object x when x belongs to both or to neither.

Definition 2 An *n*-ary function ϕ is said to preserve agreement when the following law holds:

If $F_1, \ldots, F_n, G_1, \ldots, G_n$ are any sets and x is an object in which F_i and G_i agree pairwise for each i, then $\phi(F_1, \ldots, F_n)$ and $\phi(G_1, \ldots, G_n)$ agree in x.

⁶ Quine writes: 'Any set which is specifiable in terms of the given sets F_1, \ldots, F_n by means exclusively of intersect, union, and complement is called a *Boolean function* of F_1, \ldots, F_n .' I agree with the Associate Editor's comment that it would be better to write that a Boolean function assigns sets given a well specified domain. Furthermore, the now standard way to define Boolean functions is the following: a Boolean function is a function $f: \{0, 1\}^n \to \{0, 1\}$. The choice of 0 and 1 is purely symbolic; it could be $\{T, F\}$ or $\{-1, 1\}$ etc. as long as we have two-alternative sets.



⁵ I purchased a copy of the enlarged edition when it appeared in 1995.

Then we can state the first theorem.

Theorem 1 A function is a parametric Boolean if and only if it preserves agreement.

It seems rather obvious that agreement preservation is related to independence of irrelevant alternatives and that Theorem 1 provides an equivalence to independence.

The second theorem is even more easily interpretable in the context of social choice.

Definition 3 An *n*-ary Boolean function ϕ is thoroughly commutative if for all sets F_1, \ldots, F_n and every permutation G_1, \ldots, G_n of them $\phi(F_1, \ldots, F_n) = \phi(G_1, \ldots, G_n)$.

The *incidence* of x in the sets F_1, \ldots, F_n is the number of sets from F_1, \ldots, F_n to which x belongs. It will be denoted by $I_x(F_1, \ldots, F_n)$.

The second relates the thorough commutativity to some integer.

Theorem 2 An n-ary function ϕ , whose arguments and values are sets, is a thoroughly commutative Boolean function if and only if there is a class N of natural numbers $\leq n$ such that for all sets F_1, \ldots, F_n , and all objects x,

x is an element of $\phi(F_1, \ldots, F_n)$ if and only if $I_x(F_1, \ldots, F_n)$ belongs to N.

3 The Rand memoranda

In 1949 Quine authored three Rand memoranda:

- A theorem on parametric Boolean functions, RM-196, dated 27 July 1949,
- Commutative Boolean functions, RM-199, dated 10 August 1949 and
- On functions of relations, with especial reference to social welfare, RM-218, dated 19 August 1949.

The concatenation of the first two memoranda essentially constitutes the paper published in *Selected Logic Papers*. So the remaining part of my paper will be devoted to RM-218. It is rather curious that, according to their dates, the more abstract papers precede the memorandum which is, in some sense, an application, and, as a consequence, less abstract or more precisely more interpretable from an intuitive standpoint. But one can suppose that RM-218 was conceived before the other two. This research memorandum includes a reference to 'Kenneth J. Arrow *The possibility of a universal welfare function* RAD(L) 289. This is probably the same document as *The possibility of a universal social welfare function* P-41 dated 26 September 1948 (Arrow 1948). I have been unable to trace RAD(L) 289. In Amadeo (2003, p. 85), one can read: This puzzle led to Arrow's initial formulation of his impossibility theorem, titled "Social Choice and Individual Values," RAND report RM-291, 28 July 1949. I have also been unable to find this document whose title is strangely identical to

Arrow's monograph (Arrow 1951, 1963) and whose date and number are not consistent with Quine's RM-218. As a consequence I even doubted that Arrow and Quine met at RAND in 1949 until, in a private communication, Arrow wrote me 'I did meet him (Quine) briefly at RAND. I did not meet him when I was at Harvard...'.



3.1 Postulate P1: preservation of agreement

In the first section of RM-218, Quine introduces the notion of aggregation of individual orderings function, denoted by F (then we have a social ordering given by $F(R_1, \ldots, R_n)$ where the R_i s are the individual orderings) mentioning that such a function must satisfy suitable conditions and justifies his work by stating "The purpose of the present paper is to transform certain such conditions into logical equivalents with a view to illuminating those conditions and making them easier to work with."

Quine formally considers binary relations (which Quine calls 'relations'), that is subsets of the Cartesian product of a basic set X by itself or sets of ordered pairs. Then two binary relations R and S are said to agree in a pair < x, y > when < x, y > is an element of both R and S or of neither. Although Quine does not mention the phrase *Independence of irrelevant alternatives* (IIA) but only mentions Arrow's Condition 2 (in RAD(L)-289) his P1 is obviously IIA.

P1. For all relations $R_1, \ldots, R_n, S_1, \ldots, S_n$ and all objects x and y, if R_1 agrees with S_1 in $\{x, y\}$, and R_2 with S_2, \ldots , and R_n with S_n , then $F(R_1, \ldots, R_n)$ agrees with $F(S_1, \ldots, S_n)$ in $\{x, y\}^7$

One can remark that this is basically similar to Definition 2.8

Then Quine defines *Boolean functions* and *parametric Boolean functions* essentially as in his previous RAND memorandum RM-196 and as in Quine (1995).

'A function of R_1, \ldots, R_n is called *Boolean* if constructed of R_1, \ldots, R_n by means exclusively of intersect, union, and complement.'

'An n-ary function ψ is called a *parametric Boolean function* if there are relations H_1, \ldots, H_r and an (n + r)-ary Boolean function ϕ such that, for all choices of R_1, \ldots, R_n ,

$$\psi(R_1,\ldots,R_n) = \phi(H_1,\ldots,H_r,R_1,\ldots,R_n).$$

Quine then states:

'Now P1 is equivalent to this:

F is a parametric Boolean function.'

For the proof, Quine refers to his previous memorandum. He also mentions that, as a by-product of this proof, 'P1 guarantees the existence of relations K_1, \ldots, K_{2^n} such that, for all choices of R_1, \ldots, R_n

$$F(R_1, \ldots, R_n) = \sum_{i=1}^{2^n} K_i R^i$$
.

One needs some explanations regarding the R^i and about the meaning of Σ and of the contiguity of relations.



⁷ In the RAND document by Arrow I was able to read, IIA is Condition 3 and is presented in terms of choice sets. According to Quine the binary-pairwise-version has been shown to be equivalent to the choice sets version by Norman Dalkey.

 $^{^8}$ The notations P1, P2... are for Postulate 1, Postulate 2... as can be checked before the statement of P2.

⁹ Strangely, the roles of ψ and ϕ have been reversed.

Furthermore, Σ is the symbol used for union, contiguity means intersection and an overline is for complements. Accordingly, with standard notations, we should have:

$$F(R_1, \dots, R_n) = \bigcup_{i=1}^{2^n} K_i \cap R^i \text{ and }$$

$$R^1 \text{ is equal to } R_1 \cap R_2 \cap \dots \cap R_n$$

$$R^2 = \overline{R_1} \cap R_2 \cap \dots \cap R_n$$

$$\vdots$$

$$R^{n+1} = R_1 \cap R_2 \cap \dots \cap \overline{R_n}$$

$$R^{n+2} = \overline{R_1} \cap \overline{R_2} \cap R_3 \cap \dots \cap R_n$$

$$\vdots$$

$$R^{2^n} = \overline{R_1} \cap \overline{R_2} \cap \dots \cap \overline{R_n}.$$

Quine does not offer any comments. Of course, in 1949, the Arrovian condition of independence of irrelevant alternatives and its consequences had been hardly discussed and even completely understood. ¹⁰

 $^{^{10}}$ Although in Arrow's memorandum one can understand the meaning of 'Condition 3 implies that in considering C(S), we can disregard all preferences among alternatives not in S', in the memorandum as in Arrow (1951) it is shown that Borda's rule does not satisfy IIA, which, of course, is true, through an example where the IIA which is considered is related to the Nash version of IIA which states that if the chosen elements in a set belong to some subset of this set, the chosen elements in the subset are identical to the chosen elements in the set. Some scholars believe that Nash (1950) borrowed his notion of IIA from Arrow's notion, but this is dubious since Nash never used, to the best of my knowledge, the phrase 'independence of irrelevant alternatives.' These two notions are different mathematical objects, one being linked to choice functions (Nash and many later developments on set-theoretic revealed preference theory) and the other being linked to aggregation of individual data, for instance, preferences. Kuhn (see Kuhn and Nasar 2002) in his Editor's introduction to Nash (1950) explicitly mentions the phrase independence of irrelevant alternatives. At this time, I am still unable to determine the date when the phrase 'independence of irrelevant alternatives' has been introduced in the context of a non-Arrovian framework.



3.2 Postulate P2: symmetry over objects

According to Quine, the construction from a (binary) relation of a new relation where the objects x and y are interchanged but 'leaving all other objects unchanged' is due to Dalkey. Quine denotes by R^{xy} the relation obtained from R by interchanging x and y. In fact this is obviously a transposition in combinatorial theory and we know that a permutation is the product of transpositions. Quine mentions that 'an interesting postulate on F that Dalkey has propounded is this:

P2 For all
$$R_1, ..., R_n, x, y,$$

 $[F(R_1, ..., R_n)]^{xy} = F(R_1^{xy}, ..., R_n^{xy}).'$

Quine justly observes that 'The substance of this postulate is that $F(R_1, \ldots, R_n)$ is to make use of no special features of objects over and above their manner of occurence in R_1, \ldots, R_n .' This notion is now known as *neutrality*. However, one must distinguish a form of neutrality which can be called *combinatorial neutrality* from the more standard form of neutrality which one deals with in social choice, neutrality which can be called *preference-based neutrality* or, by analogy with *welfarism*, *preferencism*.¹¹ To clarify this distinction consider Borda's rule with three individuals 1, 2 and 3, four objects x, y, z, w and the following relations:

$$xR_1yR_1wR_1z$$

 $xR_2yR_2wR_2z$
 $yR_3xR_3wR_3z$.

With Borda's scores given by (3, 2, 1, 0) (3 for top, 2 for second, 1 for third and 0 for last), x has a score of 8 y a score of 7.

Now with the following relations:

$$yR_1xR_1wR_1z$$

 $yR_2xR_2wR_2z$
 $xR_3yR_3wR_3z$.

The objects x and y have been interchanged in the sense of Dalkey/Quine and now the scores are 8 for y and 7 for x. Postulate P2 has been satisfied.

Now consider the following relations:

$$yR_1xR_1wR_1z$$
$$yR_2xR_2wR_2z$$
$$xR_3wR_3zR_3y.$$

¹¹ It is this notion of *preferencism* and its implications which is at the origin of the developments on the non-welfaristic approaches in social choice following Sen (1970a, b). For a recent discussion of neutrality by Sen, see Sen (2014).



One can observe that, preference-wise, *x* and *y* were again interchanged (compared with the first set of three relations). But the score of *x* is 7 and the score of *y* is 6, so that, in spite of the interchange, the result (collective preference) regarding *x* and *y* is the same, showing that Borda's rule does not satisfy *preferencism* although it satisfies *combinatorial neutrality*.

The relation of *identity* is then defined as the set of pairs of the form $\langle x, x \rangle$ and is denoted by I. Quine shows that adding P2 to P1 amounts to adding the following to the former result: F has no parameters but I. I quote from RM-218: 'In other words, P1-2 are equivalent to saying that there is an (n+1)-ary Boolean function ϕ such that, for all R_1, \ldots, R_n ,

$$F(R_1, ..., R_n) = \phi(I, R_1, ..., R_n).'$$

In Quine's words, this is 'a considerable simplification.' However, since scoring rules, such as Borda's rule or plurality rule, do not satisfy IIA, and since combinatorial neutrality is not imposed in the Arrovian framework, the functions F which can be considered will be rather limited (pairwise voting rules where the names of objects/candidates do not matter, for instance majority rule or pairwise voting rules associated with a simple game structure). 12

3.3 Postulate P3: thorough commutativity

Again Quine refers to Dalkey who proposed a postulate which will be Quine's third postulate, denoted P3.

P3 For all relations R_1, \ldots, R_n , and every permutation S_1, \ldots, S_n thereof,

$$F(R_1, ..., R_n) = F(S_1, ..., S_n).$$

In Quine's words, 'the order of arguments of F is immaterial; F is, in a word, thoroughly commutative.' Obviously this postulate of thorough commutativity is what social choice theorists call anonymity. Then Quine adds P3 to his two other postulates. He considers the complement \overline{I} of his identity relation, that is the set of pairs $\langle x, y \rangle$ such that $x \neq y$ and if a pair is an element of exactly i of the relations R_1, \ldots, R_n, i is said to be the *incidence* of the pair in R_1, \ldots, R_n .

Considering the three postulates together, Quine demonstrates the following result: 'There are classes A and B of natural numbers $\leq n$ such that, for all choices of R_1, \ldots, R_n , $F(R_1, \ldots, R_n) =$ the set of all identity pairs whose incidence in R_1, \ldots, R_n belongs to A, and all diversity pairs whose incidence in R_1, \ldots, R_n belongs to B.'

Of course, assuming anonymity in addition to P1 and P2 still reduces the set of possible F. For instance, if we consider voting games previously mentioned, we will have to restrict ourselves to the so-called *quota games* (voting games where coalitions are winning according to the number of voters belonging to the coalitions).

¹² See, for instance, Martin and Salles (2013).



In his last, Quine considers relations of the form 'better than' (strict preference) and relations of the form 'no worse than' (weak preference). He obviously assumes that the relation 'no worse than' is complete. He then reformulates his result: if we choose the 'better than' interpretation we exclude all identity pairs and if we choose the 'no worse than' interpretation we include them all. In both cases, the so-called class A disappears from the statement of the result.

4 Conclusion

In my view, the most important part of Quine's work is his alternative definition of IIA. The proof of the equivalence between P1, i.e. IIA, and the fact that F is a parametric Boolean function is far from trivial. Since Quine's declared purpose was to 'illuminate' IIA—and also combinatorial neutrality and anonymity—and to 'make them easier to work with', the next step should have been to offer another proof of Arrow's theorem. My conjecture is that, while at RAND, Quine tried to prove Arrow's theorem, but did not succeed, and considering, back at Harvard, that he had more important, or at least more interesting from his point of view, things to do, he abandoned this research. So the question remains: Is it possible to derive a Quinean proof of Arrow's theorem? Furthermore, are there Quinean proofs of the impossibilities of various preference aggregation functions? Related to these questions, we may wonder whether we can derive a Quinean version of preferencism.

In the more abstract setting where sets are considered rather than relations, another issue arises: is it possible to have a set-theoretic construction which would convert the Arrovian framework in this abstract setting? In case of a positive answer, would it be possible to prove an inconsistency? Finally, would it be possible to obtain something which is interpretable from an intuitive point of view?¹³

There is a small literature on the use of Boolean functions in social choice, for instance Kalai (2002) and O'Donnell (2014). My comments here apply to O'Donnell. O'Donnell considers three alternatives and a Boolean function (of the variety described in footnote 6) for every two-alternative subsets (the Boolean function is not necessarily identical for each two-alternative subset). For each two-alternative subset, the Boolean function selects one alternative. The problem then is to obtain a Condorcet winner (an alternative selected against the two others). This is quite remote from Quine's analysis, but I think that some further explorations could be undertaken. However, the restriction to two values, $\{0, 1\}$, or $\{-1, 1\}$ in O'Donnell, amounts to restrict the analysis to asymmetric binary relations (no indifference). Also, we cannot contemplate an analysis concerning transitivity or transitivity-type properties of the preferences. Incidentally, I believe that Quine came up against this difficulty: how can transitivity be rendered uniquely in terms of intersection, union and complementation of sets?



¹³ These last remarks were prompted by Christian List's comments in Lund.

References

Amadeo SM (2003) Rationalizing capitalist democracy. The cold war origins of rational choice liberalism. The University of Chicago Press, Chicago

Arrow KJ (1948) The possibility of a universal social welfare function. RAND P-41

Arrow KJ (1951) Social choice and individual values. Wiley, New York

Arrow KJ (1963) Social choice and individual values, 2nd edn. Wiley, New York

Kalai G (2002) A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. Adv Appl Math 29:412–426

Krentel WD, McKinsey JCC, Quine WV (1951) A simplification of games in extensive form. Duke Math J 18:885–900

Kuhn HW, Nasar S (eds) (2002) The essential John Nash. Princeton University Press, Princeton

Martin M, Salles M (2013) Social choice and cooperative game theory: Voting games as social aggregation functions. Int Game Theory Rev 15:1–17

McKinsey JCC (1952) Introduction to the theory of games. McGraw-Hill, New York

Nash JF (1950) The bargaining problem. Econometrica 18:155–162

O'Donnell R (2014) Analysis of Boolean functions. Cambridge University Press, Cambridge

Quine WV (1949a) A theorem on parametric Boolean functions, RM-196, dated 27 July 1949

Quine WV (1949b) Commutative Boolean functions, RM-199, dated 10 August 1949

Quine WV (1949c) On functions of relations, with especial reference to social welfare, RM-218, dated 19 August 1949

Quine WV (1966, 1995) On Boolean functions. In: Quine WV (ed) Selected logic papers. Harvard University Press, Cambridge

Quine WV (1985) The time of my life. MIT Press, Cambridge

Quine WV (1986) Autobiography of W. V. Quine. In: Hahn LE, Schilpp PA (eds) The philosophy of W. V. Quine. Open Court, La Salle

Sen AK (1970a) The impossibility of a Paretian liberal. J Polit Econ 78:152-157

Sen AK (1970b) Collective choice and social welfare. Holden-Day, San Francisco

Sen AK (2014) Arrow and the impossibility theorem. In: Maskin E, Sen A (eds) The Arrow impossibility theorem. Columbia University Press, New York

