Expressive Three-valued Truth Functions

Stephen Pollard*
Division of Social Science
Truman State University
spollard@truman.edu

Received by Hans van Ditmarsch
Published May 29, 2006
http://www.philosophy.unimelb.edu.au/ajl/2006

Abstract: The expressive truth functions of two-valued logic have all been characterized, as have the expressive unary truth functions of finitely-many-valued logic. This paper introduces some techniques for identifying expressive functions in three-valued logics.

1 BACKGROUND

We are going to explore a property of truth functions known as "expressiveness." Before subjecting readers to the long-winded explanation of what that property is, I would like to indicate why the property is worth investigating.

Pick any n-valued sentential logic. Suppose p is a sentential variable. Consider, for the moment, those sentences with occurrences of no sentential variables other than p. Call these the p-*sentences*. Let D_k be the set of p-sentences that receive a designated value when p receives the value k. D_k is the set of p-sentences that are true in some way when p is true or untrue in manner k. Such a set is sometimes called a THEORY. Suppose, for any values k and j, that $D_k \subseteq D_j$ only if $k = j$. This means the theories $D_1, \ldots, D_n$ are pairwise distinct: $D_k \neq D_j$ if $k \neq j$. It also means they are maximally satisfiable in the set of p-sentences: if ϕ is a p-sentence not in D_k, then, no matter what value p has, some member of $D_k \cup \{\phi\}$ will be untrue in some way. Since each p-sentence has one of only n^n possible truth tables, each theory D_k is finite modulo logical equivalence (i.e., modulo identity of truth tables). Take a representative from each logical equivalence class and form new sets $\Delta_1(p), \ldots, \Delta_n(p)$ by throwing out all the members of each D_k except any of those representatives that might be present. (So each member of D_k is represented by exactly one equivalent sentence in $\Delta_k(p)$.)

*Grateful thanks to the two referees who helped make this paper much more readable.

Now return to the full language of our logic. Given any sentence ϕ, form the set $\Delta_k(\phi)$ by replacing every occurrence of p in every member of $\Delta_k(p)$ with an occurrence of ϕ. We can show that an interpretation will assign a designated value to each member of $\Delta_k(\phi)$ if and only if it assigns k to ϕ. That is, the members of $\Delta_k(\phi)$ jointly affirm that ϕ has value k. The right-left direction is easy: an interpretation assigning k to ϕ will assign a designated value to each member of $\Delta_k(\phi)$. Here is a proof of the converse. $F(p)$ and $G(p)$ will be p-sentences. $F(\phi)$ will be the result of replacing each occurrence of p in $F(p)$ with an occurrence of ϕ. f will be the unary truth function whose graph is the same as the truth table for $F(p)$. $\mathrm{M}(\phi)$ will be the value an arbitrary interpretation M assigns to the sentence ϕ. Now suppose $G(p) \in D_k$. Let $F(p)$ be the member of $\Delta_k(p)$ equivalent to $G(p)$. Then $F(\phi) \in \Delta_k(\phi)$. Suppose M assigns a designated value to each member of $\Delta_k(\phi)$. Then $M(F(\phi))$ is designated, as is $f(M(\phi))$. So $F(p)$ receives a designated value when p receives the value $M(\phi)$. That is, $F(p) \in D_j$ where $j = M(\phi)$. So $G(p) \in D_j$. More generally, $D_k \subseteq D_j$. So $k = j = M(\phi)$.

A key assumption here was that $D_k \subseteq D_j$ only if $k = j$. We have seen that if $D_1, \ldots, D_n$ have this property, if they are pairwise distinct and maximally satisfiable in the set of p-sentences, then, for each value k and sentence ϕ, we can use finitely many sentences to assert that ϕ has value k. Here are two reasons to take an interest in this.

Forms of assertion and denial. Say that one denies a sentence when one attributes some form of untruth to it. In assessing the expressive capacities of some language, one might inquire about how many forms of denial it affords. If $D_1, \ldots, D_n$ are distinct and maximally satisfiable, then a partial answer is, "The language provides at least as many forms of denial as there are undesignated values." After all, to attribute an undesignated value j to a sentence ϕ, one need only assert each member of $\Delta_j(\phi)$. The same holds for forms of assertion. One can affirm that ϕ is true in some unspecified way by asserting ϕ. One can affirm that ϕ is true in some particular way by asserting each member of $\Delta_k(\phi)$ for some designated k. Students of many-valued resolution might find it helpful to recall the intended meaning of signed formulas or meta-language literals. (See, for example, [1].) A literal p^k is meant to assert that sentence p has value k. If $D_1, \ldots, D_n$ are distinct and maximally satisfiable, then the content of each such literal is captured by finitely many object-language sentences. We now consider one reason this might be convenient.

Formalizability. Suppose we can translate each meta-language literal using finitely many sentences of the object-language. Then a generalization of a well-known algorithm yields a deductive system that is sound and complete with respect to our logic. Beall and van Fraassen [2, pp. 182–185] provide a nice exposition. (See also [3] and [8].) They presuppose, for each value k, a predicate F_k such that $F_k(\phi)$ receives a designated value in an interpretation if and only if ϕ receives the value k in that interpretation. That is, they assume there is a single sentence announcing that ϕ has value k. It presents no problem, though, if several (but still only finitely many) sentences are needed to disclose the value of ϕ. A few easy modifications allow us to use the members of $\Delta_k(\phi)$ in place of the sentence $F_k(\phi)$.

Now what does this have to do with expressiveness? Suppose f is an n-valued truth function with our (as yet undefined) property: suppose f is expressive. Actually, "expressive" is shorthand for "expressive with respect to a set of designated values." So suppose f is expressive with respect to the designated values of our logic. Suppose, further, that the graph of f is the same as the truth table of some sentence in our logic. If f is unary, then $D_1, \ldots, D_n$ will be maximally satisfiable in the set of p-sentences and, as long as $D_1, \ldots, D_n$ are pairwise distinct, our logic will have the nice properties just discussed. This is the case no matter what n is. (See Theorem 4.3 of [8].) But suppose now that $n = 3$, our values being 1, 2, and 3 with just 1 designated. Suppose f is expressive with respect to $\{1\}$. If $f(1 \ldots 1) \neq 1$, then, again, $D_1, \ldots, D_n$ will be maximally satisfiable in the set of p-sentences and, as long as $D_1, \ldots, D_n$ are all distinct, our logic will have the nice properties just discussed. (See Theorem 6 below.) There are various results of this sort. The general idea is that we can confirm that a logic has certain desirable properties (say, an elegant formalization with a straightforward Henkin-style completeness proof) if we can show that an expressive function of one kind or another is definable in the logic. This makes it desirable to have techniques for identifying expressive functions.

2 CLOSURE SPACES

Since expressiveness began life as a property of closure spaces, it will be helpful to review a bit of closure space theory. Pick some universe of discourse S. A CLOSURE SPACE C on this universe is a set of subsets of S closed under intersection. That is, $\bigcap W \in C$ whenever $W \subseteq C$ (letting $\bigcap \emptyset = S$). The members of C are known as the CLOSED SETS. A subset W of C REDUCES a closed set B if and only if $B = \bigcap W$. B is REDUCIBLE if and only if it is reduced by some subset of $C \setminus \{B\}$. If $A \subseteq S$, then $Cl(A)$, the CLOSURE of A, is the intersection of all the members of C that contain A. In a FINITARY closure space, if $x \in S$ and $A \subseteq S$, then x will belong to $Cl(A)$ only if it belongs to the closure of some finite subset of A. If $B \in C$, then B is MAXIMALLY CONSISTENT if and only if B sits just below S in the lattice of closed sets: that is, S is the only member of C that properly contains B.

Now, at last, we come to our first definition of expressiveness. A finitary closure space is EXPRESSIVE if and only if all its irreducible sets are maximally consistent. (There is a more general definition that applies to non-finitary closure spaces, but we can make do with the simpler, more restricted version since we consider only finitary logics. For the more general version, see [5, p. 121].)

Return, now, to n-valued logic. Say that a sentence ϕ is a CONSEQUENCE of a set of sentences A if and only if every interpretation assigning a designated value to each member of A also assigns a designated value to ϕ. Let C consist of the sets of sentences closed under consequence. (Those sets whose consequences are already members.) Then C is a finitary closure space. (See [9, pp. 142–144]; [10, 11].) For each interpretation M, let the THEORY D_M be the set of sentences assigned a designated value by M. Each irreducible set is a theory. (S, the set of all sentences, is reducible because $S = \bigcap \emptyset$. Every other closed set is the intersection of the theories that contain it.) So C is expressive if and only if each of its irreducible theories is maximally consistent. We are interested in the circumstances under which this occurs. In particular, we are interested in whether the definability of certain truth functions might guarantee expressiveness.

3 EXPRESSIVE FUNCTIONS

Suppose our logic is two-valued. Suppose the function g returns a T when we feed it nothing but F's and returns F for at least one sequence of inputs. Say, for example, that $g(FFFF) = T$ and $g(TFFT) = F$. Suppose there is a sentence θ in our language whose truth table is exactly the graph of g. So $\theta(\phi\psi\psi\phi)$ is T when ϕ and ψ are both F, while $\theta(\phi\psi\psi\phi)$ is F when ϕ is T and ψ is F. Then we can show that our logic has a property a bit stronger than expressiveness: each of its consistent theories (each of its theories other than S) is maximally consistent. To confirm this, suppose $D_M \neq S$. If M assigns F to only one sentence, then D_M is maximally consistent and we are done. Suppose M assigns F to at least two sentences. Suppose D_N properly contains D_M. Let ϕ be a sentence that M considers false but N considers true. Let ψ be any other sentence that M considers false. Then M thinks $\theta(\phi\psi\psi\phi)$ is true. N thinks a sentence is true whenever M thinks it is true. So N thinks $\theta(\phi\psi\psi\phi)$ is true. But this means N cannot think ψ is false. ψ was an arbitrary falsehood of M. So N thinks all of M's falsehoods are truths. So N thinks everything is true. That is, $D_N = S$. So S is the only theory that properly contains D_M. So S is the only closed set that properly contains D_M. So D_M is maximally consistent.

Every two-valued logic with a sentence expressing g is expressive. This leads us to say that g itself is expressive in the sense of "expressiveness guaranteeing" or "expressiveness producing." To be pedantic, g is expressive with respect to the choice of $\{T\}$ as the set of designated values (a qualification that can be omitted without fear of confusion in the two-valued case). We are interested in characterizing expressive n-valued truth functions in general. The two-valued ones have all been characterized [7], as have all the unary ones for each choice of n [8]. We are going to investigate some of the remaining functions, with special emphasis on the three-valued ones. First, however, a brief detour that will further motivate this project.

4 ASSIMILATORS

Intuitionist and classical connectives do not always mix well. If you have a million connectives obeying the rules for intuitionist negation and even one of them satisfies the classical principle of double negation elimination, then they all do. If you have a million connectives obeying the rules for the intuitionist conditional and even one of them satisfies Peirce's law, then they all do. Our investigations into expressive functions help us to identify deductive properties that give connectives this power of assimilation.

For example, there is an algorithm that takes us from the truth table for an expressive two-valued function like g to a pair of sequents that will "de-intuitionize" each intuitionist conditional and negation. To see an example of how this works, consider again our two-valued logic and our function g. We first introduce a new connective $\odot$ with the stipulation that $M(\phi \odot \psi) = g(M(\phi)M(\psi)M(\psi)M(\phi))$. If A and B are sets of sentences, we say that $A \models B$ if and only if B intersects each theory that contains A. (That is, an interpretation assigning T to each member of A will assign T to at least one member of B.) Then we can verify the following.

- $\{\phi, (\phi \odot \psi)\} \models \{\psi\}$
- $\emptyset \models \{\phi, \psi, (\phi \odot \psi)\}$

Now go to any deductive system featuring the corresponding sequents.

- $\phi, (\phi \odot \psi) \vdash \psi$
- $\vdash \phi, \psi, (\phi \odot \psi)$

Suppose this system also features a connective $\neg$ obeying the intuitionistically valid principle that $\psi_1, \ldots, \psi_k \vdash \neg\phi$ whenever $\phi, \psi_1, \ldots, \psi_k \vdash \neg\phi$. Then we can argue as follows.

1. $\phi, (\phi \odot \neg\phi) \vdash \neg\phi$
2. $(\phi \odot \neg\phi) \vdash \neg\phi$
3. $\vdash \phi, \neg\phi, (\phi \odot \neg\phi)$
4. $\vdash \phi, \neg\phi$

Since this last line is a version of excluded middle, we see that $\odot$ has assimilated or "de-intuitionized" $\neg$. $\odot$ has a similar effect on any connective satisfying modus ponens and the deduction theorem. This means that anyone who wants to speak an intuitionist language and a language expressing g will have to prevent the latter from infiltrating the former. This holds for any expressive two-valued truth function.

Things are just a bit more complicated when we add more truth-values. Here is an example. Go to a three-valued logic where 2 and 3 are designated, but 1 is not. Suppose the function h has a matrix of the following form. (The idea is that an empty cell can be filled by any value drawn from $\{1, 2, 3\}$.)

$$\begin{array}{c|c|c} 3 & 2 & - \\ \hline - & 3 & - \\ \hline 1 & - & - \end{array}$$

This means that $h(11) = h(22) = 3$, $h(12) = 2$, and $h(31) = 1$. Introduce the connective $\star$ with the stipulation that $M(\phi \star \psi) = h(M(\phi)M(\psi))$. It turns out that h is expressive. One proof of this draws attention to the following facts.

- $\{\phi, (\phi \star \psi), (((\psi \star \phi) \star \phi) \star \psi)\} \models \{\psi\}$
- $\emptyset \models \{\phi, \psi, (\phi \star \psi)\}$
- $\emptyset \models \{\phi, \psi, (((\psi \star \phi) \star \phi) \star \psi)\}$

As before, go to any deductive system featuring the corresponding sequents.

- $\phi, (\phi \star \psi), (((\psi \star \phi) \star \phi) \star \psi) \vdash \psi$
- $\vdash \phi, \psi, (\phi \star \psi)$
- $\vdash \phi, \psi, (((\psi \star \phi) \star \phi) \star \psi)$

Suppose we encounter a connective $\neg$ behaving as above. ($\psi_1, \ldots, \psi_k \vdash \neg\phi$ whenever $\phi, \psi_1, \ldots, \psi_k \vdash \neg\phi$.) Then we can argue as follows.

1. $\phi, (\phi \star \neg\phi), (((\neg\phi \star \phi) \star \phi) \star \neg\phi) \vdash \neg\phi$
2. $(\phi \star \neg\phi), (((\neg\phi \star \phi) \star \phi) \star \neg\phi) \vdash \neg\phi$
3. $\vdash \phi, \neg\phi, (\phi \star \neg\phi)$
4. $(((\neg\phi \star \phi) \star \phi) \star \neg\phi) \vdash \phi, \neg\phi$
5. $\vdash \phi, \neg\phi, (((\neg\phi \star \phi) \star \phi) \star \neg\phi)$
6. $\vdash \phi, \neg\phi$

So h is an assimilator and anyone wishing to speak an intuitionist language and a language expressing h will have to prevent the latter from infiltrating the former. All expressive functions are assimilators. (See [5, pp. 161–162 and 203], and [6, pp. 122–124].) When we identify expressive functions, we identify logical notions that would-be intuitionists need to keep carefully at bay. This is yet another reason to take an interest in these functions.

5 SOME LEMMAS

We now define our terms with greater care and proceed to some results. A THREE-VALUED LOGIC is a quadruple $\langle D, L, CON, \natural\rangle$ where D (the set of designated values) is a non-empty proper subset of $\{1, 2, 3\}$, L is a non-empty set of variables, CON is a non-empty set of connectives, and $\natural$ is an operator that assigns a member of $\{1, 2, 3\}$ to each 0-ary member of CON and assigns a function $f : \{1, 2, 3\}^k \to \{1, 2, 3\}$ to each k-ary member of CON when $k > 0$. The sentences of such a logic are the members of L, the 0-ary members of CON, and any expressions $F(\phi_1, \ldots, \phi_k)$ where $\phi_1, \ldots, \phi_k$ are sentences and F is a k-ary member of CON. An INTERPRETATION is any homomorphism that assigns members of {1,2,3} to sentences. That is, if M is an interpretation in a logic $\langle D, L, CON, \natural\rangle$, then $M(F(\phi_1, \ldots, \phi_k)) = F\natural(M(\phi_1) \ldots M(\phi_k))$ whenever $\phi_1, \ldots, \phi_k$ are sentences and F is a k-ary member of CON. If $p_1, \ldots, p_k$ are variables and $\psi(p_1, \ldots, p_k)$ is a sentence, then $\psi(p_1, \ldots, p_k)$ EXPRESSES the function $f : \{1, 2, 3\}^k \to \{1, 2, 3\}$ in a logic if and only if $M(\psi(p_1, \ldots, p_k)) = f(M(p_1) \ldots M(p_k))$ for every interpretation M in that logic. A logic EXPRESSES a truth function if and only if one of its sentences does. We also say that a connective F expresses the function $F\natural$.

Given any logic and any of its interpretations M, we let the THEORY D_M be the set of sentences assigned a designated value by M. D_M is CONSISTENT if and only if there are sentences that do not belong to D_M. A theory is MAXIMALLY CONSISTENT if and only if it is consistent but is not a proper subset of any consistent theory. A set of sentences is SATISFIABLE if and only if it is a subset of some theory. A theory is MAXIMALLY SATISFIABLE if and only if it is not a proper subset of any theory. (If every theory is consistent, then maximal consistency is equivalent to maximal satisfiability.) A theory is REDUCIBLE if and only if it is the intersection of theories distinct from itself. If S is the set of all our sentences, we let $\bigcap \emptyset$=S. So any inconsistent theory would be reducible (since it would be the result of applying $\bigcap$ to the empty set of theories). A logic is expressive if and only if all its irreducible theories are maximally consistent.

Say that a D-LOGIC is one whose set of designated values is D. A truth function is D-EXPRESSIVE if and only if every D-logic that expresses the function is expressive. Unlike properties such as functional completeness, the expressiveness of a function can depend on our choice of designated values.

If a logic has just one variable, we adopt a special notation for its interpretations and theories. We let M_i be the interpretation that assigns truth-value i to the one variable. We let D_i be the set of sentences assigned a designated value by M_i.

We now establish four lemmas that apply to any finitely-many-valued logics (not just the three-valued ones).

LEMMA 1 *If* $k \notin D$ *and* $f(k \ldots k) = k$*, then* f *is not* D*-expressive.*

Proof: To say that f is not D-expressive is to say that some non-expressive

D-logic expresses f. We need just one example, the simpler the better. So consider a logic with just one variable p and three connectives F, G, and H. Let F express f. We assume that $k \notin D$ and $f(k \ldots k) = k$. Let G express a function that returns k for every input. Let H express a function that returns k when given a k, but otherwise returns a designated value. Then M_k assigns k to every sentence. So D_k is empty. If i is any value other than k, then $H(p) \in D_i$. Every theory is consistent, since $G(p)$ does not belong to any theory. So D_k is irreducible but not maximally consistent.

LEMMA 2 *If* $f(i \ldots j) \in D$ *only if at least one of* $i, \ldots, j$ *belongs to* D*, then* f *is not* D*-expressive.*

Proof: Again, we only need one example of a non-expressive D-logic that expresses f. Consider a logic with just one variable p and two connectives F and G. Let F express a function f that returns a designated value only when at least one of its inputs is designated. Let G express a function g such that $g(j)$ is undesignated for every input j. Pick any undesignated value k. Then M_k assigns an undesignated value to every sentence. (Proof: an easy induction on the complexity of sentences. Our assumption about f justifies the step from "$\phi_1, \ldots, \phi_m$ are all undesignated" to "$F(\phi_1, \ldots, \phi_m)$ is undesignated.") So D_k is empty whenever k is undesignated. On the other hand, p belongs to D_j whenever j is designated. Every theory is consistent, since $G(p)$ does not belong to any theory. So if k is undesignated, D_k is irreducible but not maximally consistent.

LEMMA 3 *If* $f(i \ldots j) \in D$ *whenever at least one of* $i \ldots j$ *belongs to* D*, then* f *is not* D*-expressive.*

Proof: Once again, we describe a simple non-expressive logic with a connective expressing an f of the indicated sort. Consider a logic with just one variable p and two connectives F and G. Let F express a function f that returns a designated value whenever it receives at least one designated input. G is to be a 0-ary connective (that is, an individual constant). Pick any undesignated value k and let $M_j(G) = k$ for every value j. If i is designated, then every sentence in which p occurs belongs to D_i. A sentence in which p does not occur will belong to every theory or to none. So $D_i = D_j$ whenever i and j are both designated. Furthermore, $D_i \subset D_j$ whenever j is designated but i is not. (We get proper containment, $\subset$ rather than $\subseteq$, because $p \in D_j$ if and only if j is designated.) Every theory is consistent, since G does not belong to any theory. So if i is undesignated and D_i is not a proper subset of any D_j where j is undesignated, then D_i is irreducible but not maximally consistent.

From now on, the sets D_i are understood to be the theories of a logic with just one variable p and one connective expressing the function f. The idea will be that inspection of these simple logics allows us to tell whether certain functions are expressive.

LEMMA 4 *If* $D_j \cup \{p\}$ *is unsatisfiable whenever* $j \notin D$*, then every theory in every* D*-logic that expresses* f *will be maximally consistent.*

Proof: We consider a logic whose vocabulary includes p and F, the latter expressing f. We want to show that an arbitrary theory D_M is maximally consistent. Suppose $\phi \notin D_M$. Let $M(\phi) = j$ where j is undesignated. For any sentence θ, let $\theta(p/\phi)$ be the result of replacing every occurrence of p in θ with an occurrence of ϕ. Then $\theta(p/\phi) \in D_M$ whenever $\theta \in D_j$. Suppose $D_M \cup \{\phi\} \subseteq D_N$. Let $N(\phi) = k$ where k is designated. Pick $\theta \in D_j$. Then $\theta(p/\phi) \in D_N$. So $\theta \in D_k$. More generally, $D_j \cup \{p\} \subseteq D_k$. Suppose $D_j \cup \{p\}$ is unsatisfiable whenever $j \notin D$. Then $D_M \cup \{\phi\}$ is unsatisfiable whenever $\phi \notin D_M$. So D_M is not a proper subset of any theory. Suppose D_M is inconsistent. Then D_k is inconsistent if $M(p) = k$. But this would mean that $D_j \cup \{p\}$ is always satisfiable, contrary to our hypothesis. So D_M must be consistent.

6 SOME WELL-BEHAVED FUNCTIONS

The rest of this paper focuses on three-valued logics. Let $\alpha(k) = f(k \ldots k)$ for each value k. Let $\alpha^{m+1} = \alpha \circ \alpha^m$ where α^0 is the identity function. The function α has FIXED POINTS if and only if $\alpha(k) = k$ for some value k. Lemma 1 says that k is not a fixed point if f is D-expressive and $k \notin D$. If $\alpha(1) = i$, $\alpha(2) = j$, and $\alpha(3) = k$, then we say $\alpha = ijk$ (that is, we identify α by listing its outputs). There are eight unary functions $\alpha : \{1, 2, 3\} \to \{1, 2, 3\}$ with no fixed points. They are: 211, 212, 231, 232, 311, 312, 331, 332. If α is D-expressive, then so is f (since every logic that expresses f expresses α). So it is worthwhile to identify the expressive α's. This work has already been done [8, p. 100]. Here are the results. The two derangements 231 and 312 guarantee that every theory is maximally consistent no matter what D is. 211 and 311 guarantee that every theory is maximally consistent if D is either $\{1\}$ or $\{2, 3\}$. If D is either $\{1\}$ or $\{2, 3\}$, then the four remaining functions are not D-expressive.

We turn now to techniques that allow us to identify expressive f's even when the corresponding α's are not expressive.

LEMMA 5 *If* α *has no fixed points and* $f(j \ldots k) = 1$ *for some* m*-tuple* $(j \ldots k) \in \{2, 3\}^m$*, then neither* $D_2 \subset D_3$ *nor* $D_3 \subset D_2$ *when D is either* $\{1\}$ *or* $\{2, 3\}$.

Proof: Suppose $\alpha(1) = 3$. (The argument is the same if $\alpha(1) = 2$.) If $\alpha(2) = 1$, then α is either 311 or 312 and we are done (since, by the result mentioned just above, D_2 and D_3 are maximally consistent). Suppose $\alpha(2) = 3$. $\alpha(3)$ is either 1 or 2. If the former, then we are done, since $\alpha^2(2) = \alpha(3) = 1$ and $\alpha^2(3) = 3 = \alpha(2)$. (Note that if $D = \{1\}$, then D_2 is not a subset of D_3 since $\alpha^2(2) = 1$ and $\alpha^2(3) = 3$, while D_3 is not a subset of D_2 since $\alpha(3) = 1$ and $\alpha(2) = 3$. Similar reasoning applies if $D = \{2, 3\}$.) Suppose $\alpha(3) = 2$. Then we can define a binary function g such that $g(23) = 1$ while $g(33) = 2$ and $g(22) = 3$. (For example, if $f(22323) = 1$, we let $g(jk) = f(jjkjk)$.) Let $h(k) = g(k\alpha(k))$. If $h(3) = 3$, then we are done, since $h(2) = 1$ while $h(\alpha(3)) = 1$ and $h(\alpha(2)) = 3$. Suppose $h(3)$ is 1 or 2. Then $g(2\alpha(h(2))) = g(23) = 1$ and $g(3\alpha(h(3))) = g(33) = 2$. So D_2 is not a subset of D_3 if $D = \{1\}$, while D_3 is not a subset of D_2 if $D = \{2, 3\}$. Furthermore, $g(\alpha^2(h(3))3) = g(23) = 1$ and $g(\alpha^2(h(2))2) = g(22) = 3$. So D_3 is not a subset of D_2 if $D = \{1\}$, while D_2 is not a subset of D_3 if $D = \{2, 3\}$.

THEOREM 6 *If* 1 *is not a fixed point and* f *is* $\{1\}$*-expressive, then every theory in every* $\{1\}$*-logic that expresses* f *is maximally consistent.*

Proof: Lemmas 1 and 2 imply that f satisfies the conditions of Lemma 5. So D_2 and D_3 are irreducible and, hence, are maximally consistent. D_1 is consistent since $f(1 \ldots 1) \neq 1$. So $D_2 \cup \{p\}$ and $D_3 \cup \{p\}$ are unsatisfiable. Now apply Lemma 4.

In an expressive logic, every irreducible theory is maximally consistent. If a $\{1\}$-logic expresses a function satisfying the conditions of Theorem 6, all its theories are both irreducible and maximally consistent.

Here is another way to think about Theorem 6. Let the expresssion $A \models B$ mean that B intersects every theory that contains A. That is, every interpretation that assigns a designated value to every member of A will assign a designated value to at least one member of B. Call this situation an ENTAILMENT. Go to any single-variable $\{1\}$-logic in which all three of the D_i's are maximally consistent. We can pick a sentence $\theta_2(p)$ that belongs to D_2 but not to D_1. We can also pick a sentence $\theta_3(p)$ that belongs to D_3 but not to D_1. This yields the following entailments.

- $\emptyset \models \{p, \theta_2(p), \theta_3(p)\}$
- $\{p, \theta_2(p)\} \models \emptyset$
- $\{p, \theta_3(p)\} \models \emptyset$

We can generalize this beyond the single-variable case. Suppose that f is $\{1\}$-expressive and that $f(1 \ldots 1) \neq 1$. This guarantees the maximal consistency of the three D_i's in the single-variable logic for f. So we can find $\theta_2(p)$ and $\theta_3(p)$ as above. Now let ϕ be any sentence in any $\{1\}$-logic expressing f. Let $\theta_2(\phi)$ and $\theta_3(\phi)$ be the result of replacing every occurrence of p in $\theta_2(p)$ and $\theta_3(p)$ with an occurrence of ϕ. Then we obtain the following entailments.

- $\emptyset \models \{\phi, \theta_2(\phi), \theta_3(\phi)\}$
- $\{\phi, \theta_2(\phi)\} \models \emptyset$
- $\{\phi, \theta_3(\phi)\} \models \emptyset$

Pick any theory D_M. If ϕ is not a member of D_M, either $\theta_2(\phi)$ or $\theta_3(\phi)$ will be a member. In either case, $D_M \cup \{\phi\}$ will be unsatisfiable. So D_M is maximally consistent if it is consistent. Since $f(1 \ldots 1) \neq 1$, M cannot assign 1 to every sentence. So D_M is consistent.

Let us consider a particularly simple case where we can let $\theta_2(\phi)$=$\theta_3(\phi)$. Suppose f is a binary function of the following form.

2	–	1
–	1	–
–	–	2

Note that $f(\alpha^2(2)2) = f(\alpha^2(3)3) = 1$ while $f(\alpha^2(1)1) = 2$. So, if F expresses f and A expresses α in some $\{1\}$-logic, we have the following entailments.

- $\emptyset \models \{\phi, F(A(A(\phi))\phi)\}$
- $\{\phi, F(A(A(\phi))\phi)\} \models \emptyset$

$F(A(A(\phi))\phi)$ behaves like the classical negation of ϕ. This allows us to show that every theory in a $\{1\}$-logic that expresses f will be maximally consistent.

If we let 1 be our only undesignated value (if we let $D = \{2, 3\}$), each entailment is replaced by its dual. So in place of the above entailments we would have the following.

- $\{\phi, F(A(A(\phi))\phi)\} \models \emptyset$
- $\emptyset \models \{\phi, F(A(A(\phi))\phi)\}$

Since these are the very entailments we had before, it is especially obvious in this case that the dual entailments force every theory to be maximally consistent. More generally, if $f : \{1, 2, 3\}^k \to \{1, 2, 3\}$ is $\{1\}$-expressive, $f(1 \ldots 1) \neq 1$, and ϕ is any sentence in a $\{2, 3\}$-logic that expresses f, then there are sentences $\theta_2(\phi)$ and $\theta_3(\phi)$ with the following properties in that $\{2, 3\}$-logic.

- $\{\phi, \theta_2(\phi), \theta_3(\phi)\} \models \emptyset$
- $\emptyset \models \{\phi, \theta_2(\phi)\}$
- $\emptyset \models \{\phi, \theta_3(\phi)\}$

If ϕ is not a member of the theory D_M, both $\theta_2(\phi)$ and $\theta_3(\phi)$ are members and $D_M \cup \{\phi\}$ is unsatisfiable. So D_M is maximally consistent if it is consistent. Since $\{\phi, \theta_2(\phi), \theta_3(\phi)\}$ is unsatisfiable, M cannot assign a designated value to every sentence. So D_M is consistent. This confirms the following result.

THEOREM 7 *If* 1 *is not a fixed point, then* f *is* $\{1\}$*-expressive only if every theory in every* $\{2, 3\}$*-logic that expresses* f *is maximally consistent.*

Now suppose $D = \{2,3\}$. Then $p \in (D_2 \cap D_3)$, but $p \notin D_1$. So D_1 is irreducible, as are D_2 and D_3 if they are consistent. This means that if D_2 and D_3 are consistent, the following are equivalent.

- f is $\{2,3\}$-expressive.
- D_1, D_2, and D_3 are maximally consistent.
- Every theory in every $\{2,3\}$-logic that expresses f is maximally consistent.

The following theorem holds because its hypothesis identifies a sufficient (and, in fact, necessary) condition for the consistency of D_2 and D_3.

THEOREM 8 *If neither* 2 *nor* 3 *is a fixed point and* $f(j \ldots k) = 1$ *for some* m*-tuple* $(j \ldots k) \in \{2,3\}^m$*, then* f *is* $\{2,3\}$*-expressive only if every theory in every* $2,3$*-logic that expresses* f *is maximally consistent.*

We can apply reasoning about entailments to obtain the following.

THEOREM 9 *If* α *has no fixed points and* $f(j \ldots k) = 1$ *for some* m*-tuple* $(j \ldots k) \in \{2,3\}^m$*, then* f *is* $\{1\}$*-expressive if and only if* f *is* $\{2,3\}$*-expressive.*

7 AN ALGORITHM

Given any function $f : \{1,2,3\}^k \to \{1,2,3\}$ such that $f(1 \ldots 1) \neq 1$, there is a mechanical (though not particularly elegant) way to test whether f is $\{1\}$-expressive. Lemma 1 says that f is not $\{1\}$-expressive if either 2 or 3 is a fixed point of α. So we first confirm that $f(2 \ldots 2) \neq 2$ and $f(3 \ldots 3) \neq 3$. If f passes this test, we check whether f returns 1 in response to at least one k-tuple of 2's and 3's. If f does not do so, if it returns 1 only when we feed it at least one 1, then, by Lemma 2, f is not $\{1\}$-expressive. If, on the other hand, f does pass this test, then we have more work to do. If neither D_2 nor D_3 is a subset of D_1, then Lemma 4 implies that f is $\{1\}$-expressive. So we should check whether D_2 and D_3 each have members absent from D_1. To do so, we first list all the unary functions definable from f by composition. Here is one way to accomplish that. We go to a logic with just one variable and one connective, the latter expressing f. We start listing sentences in order of complexity, keeping track of the function expressed by each sentence. Since there are only 27 unary functions in a three-valued logic, there must be a j such that every unary function is defined by a sentence of complexity less than j. How do we know when we have reached this upper bound? Well, since an upper bound exists, we will eventually encounter a run of complexities from j to $j \times k$ where no new functions are expressed. That is, we will reach a complexity j such that each function expressed by a sentence of complexity less than $j \times k$ is already expressed by a sentence of complexity less than j. We can then show that no new function will be expressed by a sentence of higher complexity. (Induction step: if $\phi_1, \ldots, \phi_k$ are each equivalent to a sentence of complexity less than j, then $F(\phi_1, \ldots, \phi_k)$ is equivalent to a sentence of complexity less than $j \times k$.) Now that we have listed all the unary functions definable from f by composition, we need to conduct two searches to determine whether the theories D_2 and D_3 are subsets of D_1.

- We search our list for a function g such that $g(2) = 1$ but $g(1) \neq 1$.
- We search our list for a function h such that $h(3) = 1$ but $h(1) \neq 1$.

If each of our searches is successful (if we manage to show that D_2 and D_3 are not subsets of D_1), then, by Lemma 4, f is {1}-expressive. Otherwise, by Theorem 6, it is not.

Here is an example. Suppose f is the following binary function.

2	1	2
1	1	3
2	3	2

We determine by inspection that the diagonal function α has no fixed points and that there is an ordered pair of 2's and 3's that causes f to return a 1. (In particular, $f(22) = 1$.) We then list the unary functions definable from f.

	α^0	α^1	α^2	$f(\alpha^0\alpha^1)$	$f(\alpha^1\alpha^2)$	$\alpha^1 f(\alpha^0\alpha^1)$
1	1	2	1	1	1	2
2	2	1	2	1	1	2
3	3	2	1	3	1	2

There is no function h on our list such that $h(3) = 1$ but $h(1) \neq 1$. So in a logic with just one variable and one connective expressing f, D_3 would not be maximally consistent. So f is not {1}-expressive.

There are 729 binary functions whose diagonal values $\alpha(j)$ are 212. Using our algorithm (and a variety of short-cuts), we can establish that exactly 713 of them are {1}-expressive. Theorem 9 implies that these functions are also the {2, 3}-expressive ones. By permuting the values 2 and 3, we obtain the {1}-expressive and {2, 3}-expressive binary functions whose diagonal values are 331.

There are 729 binary functions whose diagonal values $\alpha(j)$ are 332. Using our algorithm (and a variety of short-cuts), we can establish that exactly 398 of them are {1}-expressive. Theorem 7 implies that all 398 are {2, 3}-expressive. But there are more: an additional 255 are {2, 3}-expressive. We now consider a technique for identifying these functions.

8 ANOTHER TECHNIQUE

Suppose we find sentences $\theta(\phi\psi)$ and $\theta^*(\phi\psi)$ with truth tables of the following form.

ϕ	ψ	$\theta(\phi\psi)$	$\theta^*(\phi\psi)$
1	1	2 or 3	2 or 3
2	1	1	–
3	1	–	1

Then the following entailments will hold when $D = \{2, 3\}$.

- $\emptyset \models \{\phi, \psi, \theta(\phi\psi)\}$
- $\emptyset \models \{\phi, \psi, \theta^*(\phi\psi)\}$
- $\{\phi, \theta(\phi\psi), \theta^*(\phi\psi)\} \models \{\psi\}$

These entailments guarantee that every consistent theory is maximally consistent. For suppose D_M is consistent but not maximally consistent. There must be at least two sentences ϕ and ψ that do not belong to D_M. $\theta(\phi\psi)$ and $\theta^*(\phi\psi)$ will belong to D_M. So ψ will belong to every theory that contains $D_M \cup \{\phi\}$. ψ could have been any sentence not in D_M. So every non-member of D_M will belong to every theory that contains $D_M \cup \{\phi\}$. So no consistent theory will contain $D_M \cup \{\phi\}$ and D_M must be maximally consistent after all.

To see how this works in particular cases, let f be a binary function of the following form.

3	–	–
1	3	–
2	–	2

If F expresses f and A expresses α, we have the following assignments.

ϕ	ψ	$F(\phi\psi)$	$F(A(\phi)\psi)$
1	1	3	2
2	1	1	2
3	1	2	1

So the following entailments will hold when $D = \{2, 3\}$.

- $\emptyset \models \{\phi, \psi, F(\phi\psi)\}$
- $\emptyset \models \{\phi, \psi, F(A(\phi)\psi)\}$
- $\{\phi, F(\phi\psi), F(A(\phi)\psi)\} \models \{\psi\}$

Since these entailments guarantee that every consistent theory is maximally consistent, f is $\{2,3\}$-expressive.

Using this technique, we can show that 653 binary functions with diagonal values 332 are $\{2,3\}$-expressive. Lemma 3 implies that 64 of the remaining 76 are not $\{2,3\}$-expressive. Theorem 9 guarantees that an additional seven are not $\{2,3\}$-expressive. Here are the five that are left.

$$\begin{array}{c|c|c} 3 & 3 & 1 \\ \hline 3 & 3 & 2 \\ \hline 1 & 2 & 2 \end{array} \qquad \begin{array}{c|c|c} 3 & 3 & 1 \\ \hline 3 & 3 & 2 \\ \hline 2 & 2 & 2 \end{array} \qquad \begin{array}{c|c|c} 3 & 3 & 1 \\ \hline 3 & 3 & 2 \\ \hline 3 & 3 & 2 \end{array} \qquad \begin{array}{c|c|c} 3 & 3 & 2 \\ \hline 3 & 3 & 2 \\ \hline 1 & 2 & 2 \end{array} \qquad \begin{array}{c|c|c} 3 & 3 & 3 \\ \hline 3 & 3 & 3 \\ \hline 1 & 2 & 2 \end{array}$$

How do we show that these functions are not $\{2,3\}$-expressive?

Consider a logic with just two variables p and q and two connectives F and G. Let M_{jk} be the interpretation that assigns j to p and k to q. Let D_{jk} be the set of sentences assigned a designated value by M_{jk}. Suppose F expresses one of our five functions. Let G express the following function.

$$\begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 2 & 2 & 3 \\ \hline 3 & 3 & 3 \end{array}$$

Then we can show: $M_{11}(\theta) = 1$ only if $M_{21}(\theta) \neq 3$; $M_{11}(\theta) = 2$ only if $M_{21}(\theta) = 2$; $M_{11}(\theta) = 3$ only if $M_{21}(\theta) = 3$. So $D_{11} \subset D_{21}$. Furthermore, $D_{11} \subset D_{12}$. But $D_{11} \neq (D_{12} \cap D_{21})$, since $G(pq) \in (D_{12} \cap D_{21})$ while $G(pq) \notin D_{11}$. $F(pq)$ and $F(qp)$ prevent D_{11} from being a subset of either D_{13} or D_{31}. D_{22}, D_{23}, D_{32}, and D_{33} are all inconsistent. So D_{11} is irreducible and our function is not $\{2,3\}$-expressive.

We can now tally up the D-expressive binary functions f where α has no fixed points and D is either $\{1\}$ or $\{2,3\}$. We include a column giving the number of Sheffer (that is, functionally complete) functions [4].

α	$\{1\}$-expressive	$\{2,3\}$-expressive	Sheffer
211	729	729	389
212	713	713	389
231	729	729	720
232	398	653	389
311	729	729	389
312	729	729	720
331	713	713	389
332	398	653	389
Total	5138	5648	3774

If we consider functions with fixed points, we find an additional 2222 $\{1\}$-expressive binary functions (for a total of 7360) and an additional 6356 $\{2,3\}$-expressive binary functions (for a total of 12004). Readers will here be spared the many pages of bookkeeping from which these numbers are derived. A full tally appears below.

α	$\{1\}$-expressive	$\{2,3\}$-expressive	Sheffer
111	540	0	0
112	643	0	0
113	0	0	0
121	0	0	0
122	0	0	0
123	0	0	0
131	643	0	0
132	396	0	0
133	0	0	0
211	729	729	389
212	713	713	389
213	0	704	0
221	0	648	0
222	0	584	0
223	0	581	0
231	729	729	720
232	398	653	389
233	0	661	0
311	729	729	389
312	729	729	720
313	0	648	0
321	0	704	0
322	0	661	0
323	0	581	0
331	713	713	389
332	398	653	389
333	0	584	0
Total	7360	12004	3774

9 A QUESTION

We have been investigating those truth functions expressible only in logics whose irreducible theories are all maximally consistent. These, of course, are the expressive functions. A good many of these functions have been identified: all the classical ones [7], all the unary finitely-many-valued ones [8], all the binary three-valued ones. Now here is a curious fact: at most one reducible theory appears in any logic expressing any of the functions just listed. When such a theory appears it is always the set of all sentences. So, in any logic expressing one of these functions, every consistent theory is maximally consistent.

Here is the question. Is this always the case? Are expressive functions expressible only in logics whose consistent theories are all maximally consistent? Do expressive functions make it impossible for a reducible theory to be consistent? I would guess this is so. I would guess there is a simple reason why this is so. I would dearly love to know what that reason is.

References

[1] Baaz M. and C. G. Fermüller, "Resolution for Many-Valued Logics," in *Logic Programming and Automated Reasoning*, A. Voronkov, ed., Springer-Verlag, Berlin, 1992, pp. 107–118.

[2] Beall, J. C., and B. C. van Fraassen, *Possibilities and Paradox*, Oxford University Press, Oxford, 2003.

[3] Carnielli, W. A., "Systematization of Finite Many-Valued Logics Through the Method of Tableaux," *The Journal of Symbolic Logic*, vol. 52 (1987), pp. 473–493.

[4] Martin, N. M., "The Sheffer Functions of 3-Valued Logic," *The Journal of Symbolic Logic*, vol. 19 (1954), pp. 45–51.

[5] Martin, N. M., and S. Pollard, *Closure Spaces and Logic*, Kluwer Academic Publishers, Dordrecht, 1996.

[6] Pollard, S., and N. M. Martin, "Closed Bases and Closure Logic," *The Monist*, vol. 79 (1996), pp. 117–127.

[7] Pollard, S., "The Expressive Truth Conditions of Two-Valued Logic," *Notre Dame Journal of Formal Logic*, vol. 43 (2002), pp. 221–230.

[8] Pollard, S., "The Expressive Unary Truth Functions of n-Valued Logic," *Notre Dame Journal of Formal Logic*, vol. 46 (2005), pp. 93–105.

[9] van Fraassen, B. C., *Formal Semantics and Logic*, Macmillan, New York, 1971.

[10] Weaver, G., "Compactness Theorems for Finitely-Many-Valued Sentential Logics," *Studia Logica*, vol. 37 (1978), pp. 413–416.

[11] Woodruff, P. W., "On Compactness in Many-Valued Logic: Part I," *Notre Dame Journal of Formal Logic*, vol. 14 (1973), pp. 405–407.

The *Australasian Journal of Logic* (ISSN 1448-5052) disseminates articles that significantly advance the study of logic, in its mathematical, philosophical or computational guises. The scope of the journal includes all areas of logic, both pure and applied to topics in philosophy, mathematics, computation, linguistics and the other sciences.

Articles appearing in the journal have been carefully and critically refereed under the responsibility of members of the Editorial Board. Only papers judged to be both significant and excellent are accepted for publication.

The journal is freely available at the journal website at

`http://www.philosophy.unimelb.edu.au/ajl/`.

All issues of the journal are archived electronically at the journal website.

SUBSCRIPTIONS Individuals may subscribe to the journal by sending an email, including a full name, an institutional affiliation and an email address to the managing editor at `ajl-editors@unimelb.edu.au`. Subscribers will receive email abstracts of accepted papers to an address of their choice. For institutional subscription, please email the managing editor at `ajl-editors@unimelb.edu.au`.

Complete published papers may be downloaded at the journal's website at `http://www.philosophy.unimelb.edu.au/ajl/`. The journal currently publishes in pdf format.

SUBMISSION The journal accepts submissions of papers electronically. To submit an article for publication, send the LaTeX source of a submission to a member of the editorial board. For a current list of the editorial board, consult the website.