Chapter 3

Algebras for Logic

3.1 Boolean and Heyting Algebras

3.1.1 Boolean Operations

A ***Boolean operation*** is a finitary operation on the set $2 = \{0, 1\}$.

In particular, for each natural number n, an n-ary Boolean operation is a function $f : 2^n \to 2$, of which there are 2^{2^n} such. The two zeroary operations or constants are the truth values 0 and 1. The four unary operations are identity x, negation $\neg x$, and the two constant operations $\lambda a.0$ and $\lambda a.1$. There are 16 binary Boolean operations, 256 ternary operations, and so on.

Each element $(a_0, \ldots, a_{n-1})$ of the domain 2^n of f may be understood as an n-tuple of bits, or a subset $G \subseteq n$ where n is taken to denote the set $\{0, 1, \ldots, n-1\}$, or a function $g : n \to 2$. This last can be understood as the characteristic function of G, but in logic it is more often regarded as an *assignment* of truth values to the elements of n construed as *variables* $x_0, \ldots, x_{n-1}$. When $f(a_0, \ldots, a_{n-1}) = 1$ we call $(a_0, \ldots, a_{n-1})$ a ***satisfying*** assignment.

However there is nothing in the definition of Boolean operation to restrict its applicability exclusively to logic. The notion of basis will help us sharpen this idea.

A ***basis*** for a set of operations is a subset of those operations generating the whole set by composition.

We start with an example that is clearly motivated by logic.

Proposition 1 *The set of Boolean operations consisting of the constant 0 and the binary operation $a \to b$ defined as 0 when $a = 1$ and $b = 0$ and 1 otherwise forms a basis for the set of all Boolean operations.*

Proof: Negation $\neg a$ is obtainable as $a \to 0$. We may now obtain binary disjunction $a \vee b$ as $\neg a \to b$, and binary conjunction $a \wedge b$ as $(a \to \neg b)\neg$. Binary conjunction can then be iterated to n-ary conjunction, $n > 2$, via $a \wedge (b \wedge c)$ and so on, and similarly for n-ary disjunction. Unary conjunction and disjunction are the same operation, namely the identity, obtainable as $a \wedge a$. We already have zeroary disjunction, namely 0, and zeroary conjunction is obtainable as its negation $\neg 0$. We now have n-ary conjunction and disjunction for all $n \geq 0$, and also negation. We may now obtain every Boolean operation as a disjunction of conjunctions of ***literals*** (possibly negated variables), as follows.

Each n-ary assignment determines a unique n-ary operation which is satisfied by that assignment and no other. This operation may be obtained as the n-ary conjunction of the n unary operations x_i or $\neg x_i$, each chosen according to whether a_i is 1 or 0 respectively in this assignment. For example the satisfying assignment $(1, 0, 1)$ determines the operation which is representable as as $x_0 \wedge \neg x_1 \wedge x_2$.

We may now obtain any n-ary Boolean operation $f : 2^n \to 2$ as the disjunction, over all satisfying assignments of f, of the n-ary operations associated as above to those assignments. ∎

The standard Boolean basis is $(\vee, \wedge, \neg, 0, 1)$.[1] That this is a basis follows from the preceding proposition and either of the identities $a \to b = \neg a \vee b = \neg(a \wedge \neg b)$. This shows further that it is not a ***minimal basis***, in that either $\vee, \neg$ or $\wedge, \neg$ suffice. One reason for keeping both $\vee$ and $\wedge$ in the basis is that the omission of $\neg$ yields the operations of a lattice. Now the lattice $(2, \vee, \wedge, 0, 1)$ is distributive. Thus the Boolean algebra $(2, \vee, \wedge, \neg, 0, 1)$ may be understood as the distributive lattice with units $(2, \vee, \wedge, 0, 1)$ with signature expanded to include Boolean complement. This forms the basis for one definition of Boolean algebra in the next section.

Our next example of a basis is motivated not by logic but by number theory. Consider the ring $(2, +, 0, \cdot, 1)$ of integers mod 2 with additive unit 0 and multiplicative unit 1. In logical terms addition is exclusive-or—$a \oplus b = 1$ just when $a \neg b$—while multiplication is simply conjunction, and their respective units are 0 and 1.

We can use the preceding proposition to show that these ring operations forms a basis for the Boolean operations by obtaining $a \to b$ as the polynomial $ab + a + 1$ (where ab abbreviates $a \cdot b$ as usual). This equivalence can be seen either by exhaustively considering all four values for a and b, or by the reduction $ab + a + 1 = a(b + 1) + 1 = \neg(a \wedge \neg b) = a \to b$.

Zhegalkin polynomials. Unlike other moduli, the integers mod 2 satisfy $a^2 = a$ and $a + a = 0$. Hence all polynomials reduce to polynomials whose coefficients are 0 and 1 and whose terms are square-free, e.g. a^2b^3 reduces to ab. We call such polynomials ***Zhegalkin polynomials***.

Proposition 2 *(Zhegalkin, 1927) Each Boolean operation is uniquely representable as a Zhegalkin polynomial.*

We now raise the question of how to tell when a set of Boolean operations is a basis for the set of all Boolean operations. We shall answer this in terms of five structural properties each of which some operation in the basis must possess.

Affineness. The affine Boolean operations are those representable as affine Zhegalkin polynomials, those whose terms contain only one variable. These have the form of a sum of some subset of the variables and possibly 1, namely the parity operation or its complement, with the zeroary parity operation being the constant 0. All Boolean operations of arity at most one are automatically affine. The affine binary Boolean operations are just those that are 1 on an even number of inputs: thus the constant functions 0 and 1 are 1 on respectively zero and four out of the four possible inputs while the two identity operations a and y, the two complement operations $\neg a$ and $\neg y$, and the exclusive-or operation $a + y$ and its complement logical equivalence $a \equiv y$ are 1 on two out of the four inputs.

Strictness and Costrictness. We call an n-ary Boolean operation ***strict*** when it maps 0^n to 0, and ***costrict*** when it maps 1^n to 1. Thus $\vee$ and $\wedge$ are both strict and costrict, the constant 0 is strict but not costrict, and the constant 1 and the operation $\to$ are costrict but not strict. We will use strictness and costrictness in the following treatment of Boolean bases.

Duality. The ***dual*** of a Boolean operation $f(x_1, \ldots, x_n)$ is the operation $\neg f(\neg x_1, \ldots, \neg x_n)$.

Example 1 *The dual of $\wedge$ is $\vee$ and vice versa. The dual of $a + y$ is its complement $a \equiv y$. The dual of the two-out-of-three threshold function $ab \vee bc \vee ca$ is itself.*

[1] The constants are frequently omitted, but this has the unfortunate side effect that the formula 2^{2^n} for number of n-ary Boolean operations fails at $n = 0$: there being no zeroary terms. $x \vee \neg x$ is certainly constant but is a unary term.

An operation is ***self-dual*** when it is its own dual.

Example 2 *We saw above that $ab \vee bc \vee ca$ is self-dual. The only operations of arity less than 2 that are self-dual are the nonconstant unary operations, namely the identity and complement. The four binary self-dual operations are the two identities a and b and the two complements $\neg a$ and $\neg b$.*

Monotonicity. A Boolean operation $f : 2^n \to 2$ is ***monotone*** when it is monotone with respect to the inclusion order on 2^n and the order $0 < 1$ on 2. $\neg a$ is the only Boolean operation of arity at most one that is not monotone. Only 6 of the 16 binary Boolean operations are monotone, namely 0, 1, a, b, $a \wedge b$, and $a \vee b$.

Proposition 3 *(Post) A necessary and sufficient condition that a set of operations form a basis for the nonzeroary Boolean operations is that it contain operations that are respectively nonaffine, nonmonotone, nonstrict, noncostrict, and nonselfdual.*

For example the single operation $\neg(a \wedge b)$ passes this test because it meets all five conditions on its own. However $a \to b$ fails the test because it is costrict, though it meets the other four conditions. The constant 0 is not costrict, so $\to$ together with 0 forms a basis. The unary operation $\neg a$ and the ternary operation $ab + bz + za$ are both selfdual

Proof: The constant is needed because there is no way to construct a zeroary operation from only nonzeroary operations, and sufficient in the presence of $\neg$. Hence it suffices to verify that the remaining operations generate all nonzeroary operations.

Necessity. From Exercise 3 we infer that if any of the five properties are missing from all operations of the basis then it is missing from the operations obtainable from the basis by composition. But for each of these properties there exists a Boolean operation with that property – in fact the NAND operation has all of them. So a set of operations all of which lack one of the properties could not form a basis.

Sufficiency. If we identify all the arguments of a nonstrict function we get a unary function mapping 0 to 1. This will be either the constant unary operation 1 or complement. Similarly we may obtain from a noncostrict function either the constant unary operation 0 or complement.

Suppose we have not obtained complement in this way. Then we have both 0 and 1 (as constant unary operations). Now a monotonic operation must be monotonic in each variable separately. Hence a nonmonotonic n-ary operation must be nonmonotonic in some variable for some assignment of values to the $n - 1$ other variables. Assign the constants from the previous paragraph to those $n - 1$ variables so as to produce a nonmonotonic unary operation in the remaining variable. Complement is the only nonomonotonic unary operation.

Note that we are not assured of having 0 and 1 at this point since both the nonstrict and noncostrict operations may have been complement. To take care of this case, let f be an n-ary nonself-dual operation in the basis. A witness w to failure of self-duality is an n-tuple $(w_1, \ldots, w_n)$ for which $f(w_1, \ldots, w_n) = f(\neg w_1, \ldots, \neg w_n)$. Two cases arise.

Case (i). $f(0, \ldots, 0) = \neg f(1, \ldots, 1)$. In this case the witness w is not constant. Substitute variables a and b for the variables of f, putting a where w is 0 and b where w is 1. This yields a binary operation mapping 00 and 11 to different values (the premise of this case) and 01 and 10 to the same value (since w witnesses nonselfduality). The only such operations are $\wedge$, $\vee$, and their complements. Any one of these together with complement (which we already have) forms a basis and we are done.

Case (ii). $f(0^n) = f(1^n)$. In this case we may identify all the variables to yield a constant unary operation. Since we have complement we may then obtain the other constant unary operation.

Now take a nonaffine operation and consider its Zhegalkin polynomial. Since it is nonaffine it must contain a monomial in which two or more variables a and b occur. Factor the polynomial as $\alpha ab + \beta$ where α is a nonzero polynomial independent of a and b and no monomial in the Zhegalkin polynomial of β contains both a and b. Since α is not identically zero we may set the variables it depends on so as to make $\alpha = 1$. Set the remaining variables except a and b arbitrarily, making β one of the eight affine functions of a and b. It may now be verified by inspection that the constants, $\neg$, and any of the eight operations $ab + \beta$ form a basis. ∎

3.1.2 Boolean Algebras

Given an element a in a lattice with 0 and 1, an element b is a ***complement*** of a when $a \vee b = 1$ and $a \wedge b = 0$.

Proposition 4 *In a distributive lattice complements are unique. That is, if b and b' are complements of the same element a then $b = b'$.*

Proof: $b' = b' \wedge (b \vee a) = (b' \wedge b) \vee (b' \wedge a) = b' \wedge b$. By symmetry $b = b' \wedge b$ also, whence $b = b'$. ∎

A ***complemented*** lattice is one in which every element has a complement.

The previous proposition then tells us that for any complemented distributive lattice L there exists a uniquely determined complement for every element a. We denote this element $\neg a$, thereby defining a unary operation $\neg$ on the lattice.

A ***Boolean algebra*** is a complemented distributive lattice with units.

Equivalently a Boolean algebra $(A, \vee, \wedge, \neg, 0, 1)$ is an algebra satisfying the following equations. (i) The equations making each of the operations $\vee$ and $\wedge$ the binary operation of a semilattice. (ii) The two absorption laws. (iii) One distributivity equation (either one will do). (iv) The complementation equations $x \vee \neg x = 1$ and $x \wedge \neg x = 0$.

Proposition 5 *Every identity of the two-element Boolean algebra is derivable from the axioms of a complemented distributive lattice.*

Proof: We use the fact that equivalent terms have the same maximal disjunctive normal form (MDNF) up to permutations of conjunctions and disjunctions. We show that these axioms suffice to put both sides into MDNF, and these two normalized terms can then be made using commutativity of $\wedge$ and $\vee$ if and only if they denoted the same term to begin with.

A literal is a variable x or its negation $\neg x$. A term is in disjunctive normal form (DNF) when it is a disjunction of conjunctions of literals. A V-term is in maximal disjunctive normal form when every disjunct (argument of the disjunction) contains all variables of V, each occurring either positively or negatively. When V is empty we write 1 for the unique conjunction of literals, namely the empty conjunction.

We first derive the following useful identities from the equations defining a complemented distributive lattice.

(i) $x = \neg\neg x$.

Proof. $1 = \neg x \vee \neg\neg x$, so $x = x \wedge (\neg x \vee \neg\neg x) = (x \wedge \neg x) \vee (x \wedge \neg\neg x) = x \wedge \neg\neg x$, that is, $x \leq \neg\neg x$. The dual argument shows $x = x \vee \neg\neg x$, that is, $\neg\neg x \leq x$.

(ii) $\neg(x \vee y) = \neg x \wedge \neg y$ and dually $\neg(x \wedge y) = \neg x \vee \neg y$.

Proof. $(\neg x \wedge \neg y) \wedge (x \vee y) = (\neg x \wedge \neg y \wedge x) \vee (\neg x \wedge \neg y \wedge y) = 0$. $(\neg x \wedge \neg y) \vee (x \vee y) = (\neg x \vee x \vee y) \wedge (\neg y \vee x \vee y)$ $= 1$. This makes $\neg x \wedge \neg y$ the complement of $x \vee y$. The dual is obtained dually.

(iii) $\neg 0 = 1$ and $\neg 1 = 0$.

Proof. Both laws follow immediately from $0 \vee 1 = 1$, $0 \wedge 1 = 0$.

These four equations permit any term containing $\neg$ applied to other than a variable to be rewritten so that the $\neg$'s are moved closer to the variables. Every equation should be understand as a rewrite rule changing the left hand side to the corresponding right hand side.

The distributivity law then permits all instances of $\wedge$ that have a disjunction as one of their arguments to be rewritten so as to put the disjunction above the conjunction. This replaces such a conjunction with two or more conjunctions each having few disjunctions below it than the replaced conjunction had. Hence this rewriting process terminates with no conjunction above a disjunction.

Associativity allows $x \vee (y \vee z)$ to be understand as the ternary disjunction $x \vee y \vee z$, and likewise for conjunctions. We may then view the result as a disjunction of conjunctions of literals.

We now derive a further law:

(iv) $x = (x \wedge y) \vee (x \wedge y')$.

Proof. $x = x \wedge (y \vee y') = (x \wedge y) \vee (x \wedge y')$.

With this law, any conjunction φ lacking a variable x can be rewritten as two conjunctions $\varphi \wedge x$ and $\varphi \wedge \neg x$. In this way a conjunction lacking variables can be expanded to a disjunction of two or more conjunctions containing the missing variables.

Repetitions of variables (whether or not of the same sign) may then be eliminated using $x \vee x = x$, $x \wedge x = x$, $x \wedge \neg x = 0$, and $x \vee \neg x = 1$. All 0's and 1's may then be removed with $x \wedge 0 = 0$, $x \wedge 1 = 1$, $x \vee 0 = x$, and $x \vee 1 = 1$, unless we are left with a single 0, which we leave stand. (We cannot be left with 1 since this would not be maximal unless V is empty.)

We have now shown that every term can be rewritten using the equations defining "complemented distributive lattice" to maximal disjunctive normal form (MDNF).

But the disjuncts of an MDNF term must correspond to those values of the variables at which the term evaluates to 1. Hence two terms denoting the same operation must have the same disjuncts, whence they can be made equal by suitable permutations within and between disjuncts. We then have the promised derivation of their identity. ∎

3.1.3 Heyting Algebras

In 1908 Brouwer proposed that mathematics be conducted more constructively. This proposal involved constructive aspects of both truth values and data. For truth values Brouwer took objection to $p \vee \neg p = 1$, the law of the excluded middle, as not being constructively sound. For data he developed notions of lawlike and lawless sequences that we would today relate best to in terms of recursive enumerability. Here we shall consider just truth values.

Although Brouwer did not have in mind a mathematically precise alternative to Boolean logic, the concept of Heyting algebra provides a very nice model that has come to be accepted as the basis for intuitionistic propositional logic.

A Heyting algebra is a distributive lattice expanded with an operation $p \to q$. Just as we can define $\wedge$ uniquely in terms of $\vee$ and $\neg$ with $p \wedge q = \neg(\neg p \vee \neg q)$, we would like to define $\to$ uniquely in terms of $\wedge$ and $\vee$. A basic property we are after is $(p \wedge q) \to r = p \to (q \to r)$. This is a Boolean identity under the interpretation of $p \to q$ as $\neg p \vee q$. (If $p \to q$ is read as the set of functions from the set p to the set q then the identity can also be read as the so-called Curry or Schönfinkel-Curry equivalence between two sets of functions, where a function of two arguments can be made a function of one argument returning a function of the other.)

Unfortunately the equation mentions $\to$ too often to determine it uniquely. We could satisfy $(p \wedge q) \to r = p \to (q \to r)$ by taking $p \to q$ to be the operation that always returns q, or even a constant operation. What we really want is a definition of → that determines it uniquely for any given distributive lattice

We strengthen the desired equation by rewriting it as $(p \wedge q) \leq r \equiv p \leq (q \to r)$. This is no longer an ordinary equation, but rather says that $p \wedge q$ is below r in the lattice ordering if and only if p is below $q \to r$.

A Heyting algebra is a lattice L with a least element 0 such that for any two elements p, q of L, $\max\{r | p \wedge r \leq q\}$ exists (i.e. this set contains an element greater than or equal to every one of its elements). We denote that element by $p \to q$; that is, $\to: L \to L$ is a binary operation on S. We may think of $p \to q$ as the weakest precondition for the inference rule of modus ponens, from p and $p \to q$ infer q.

Equivalently we may define → via the following equations.

$$\begin{aligned} p \to p &= 1 \\ p \wedge (p \to q) &= p \wedge q \\ p \to (q \wedge r) &= (p \to q) \wedge (p \to r) \\ q &\leq p \to q \end{aligned}$$

Exercise 9 asks to prove the equivalence of these two definitions of →.

Since the second definition is purely equational we have the following.

Theorem 1 *The class of Heyting algebras is a variety.*

Note that in our definition of Heyting algebra we omitted the requirement that the lattice be distributive. Nevertheless Exercise 6 shows that every Heyting algebra is a distributive lattice.

Examples

1. Any chain having a greatest element 1 forms a Heyting algebra; when $p \leq q$, $p \to q$ must be 1, and otherwise it must be q.

2. Since the property of being a Heyting algebra is equational, direct products of Heyting algebras are Heyting algebras. In particular $(2^X, \wedge, \vee, 1)$, that is, the power set of X under union and intersection, with 1 being X itself and therefore the maximum element of 2^X, forms a Heyting algebra. Similarly so does the real line $(R, \vee, \wedge, 1)$, the real plane $(R^2, \vee, \wedge, 1)$, and higher powers.

3. The simplest nontrivial Heyting algebra is the two-element lattice $\mathbf{2} = (\{0,1\}, \vee, \wedge, 1)$. By the above rule for calculating → in a chain we find that $1 \to 0$ is 0 and the other three cases are all 1.

Now when we considered **2** as the lattice $(2, \vee, \wedge)$ we found that its theory was not that of lattices but rather of distributive lattices. For $(2, \vee, \wedge, \to)$ we may ask an analogous question: is its theory just that of Heyting algebras, or something larger? Just as with the lattice $(2, \vee, \wedge)$, the answer is that there another equation we need. Consider $(p \to q) \to p$. For each of the four assignments of 0 and 1 to p and q we find that it evaluates to p. Hence $(p \to q) \to p = p$ holds of **2**.

Now consider the next biggest Heyting algebra we know of, namely the lattice $\mathbf{3} = (\{0,1,2\}, \wedge, \vee)$. This differs from **2** in having a "middle" in addition to its two extremal elements. When $p = 1$ and $q = 0$, $(p \to q) \to p = 2$, so the equation we found for **2** fails for **3**. The general form of this counterexample for an arbitrary chain is any assignment satisfying $q < p < 1$, for which $(p \to q) \to p$ is then the maximal element. Turning this around we see that the only chain satisfying $(p \to q) \to p = p$ is **2**. So among chains at least, $(p \to q) \to p = p$ amounts to a law of the excluded middle, in the sense that there are no middle values between 0 and 1.

For a Heyting algebra with least element 0 we may define negation $\neg p$ as $p \to 0$. This notion of negation has certain familiar properties, such as $\neg(p \vee q) = \neg p \wedge \neg q$, but lacks others such as $\neg(p \wedge q) = \neg p \vee \neg q$ and $\neg\neg p = p$.

3.1.4 Boolean Heyting Algebras

Theorem 2 *The following three conditions on a lattice L are equivalent, where Heyt(L) abbreviates "L is a Heyting algebra with pseudocomplement $\neg$."*

(i) L is a Boolean algebra, with complement $\neg$.

(ii) Heyt(L) and $\neg\neg p = p$ for all p in L.

(iii) Heyt(L) and $p \vee \neg p = 1$.

Proof: (i)→(ii) Define $p \to q$ to be $\neg p \vee q$. If $r \leq p \to q$ then

$$\begin{aligned} p \wedge r &\leq p \wedge (\neg p \vee q) \\ &= (p \wedge \neg p) \vee (p \wedge q) \\ &\leq q. \end{aligned}$$

Conversely if $p \wedge r \leq q$ then

$$\begin{aligned} r &\leq (\neg p \vee p) \wedge (\neg p \vee r) \\ &= \neg p \vee (p \wedge r) \\ &\leq \neg p \vee q. \end{aligned}$$

Hence defining → in this way shows that L is a Heyting algebra; furthermore $\neg p = \neg p \vee 0 = p \to 0$, so $\neg$ is pseudocomplement. Hence Heyt(L).

Now in a Heyting algebra $p \leq \neg\neg p$, by (6) above. Moreover

$$\begin{aligned} \neg\neg p &= (p \vee \neg p) \wedge \neg\neg p \\ &= (p \wedge \neg\neg p) \vee (\neg p \wedge \neg\neg p) \\ &= (p \wedge \neg\neg p) \end{aligned}$$

so $\neg\neg p \leq p$. Hence $\neg\neg p = p$.

(ii)→(iii) $p \vee \neg p = \neg\neg(p \vee \neg p) = 1$ by (11).

(iii)→(i) We have $p \wedge \neg p = 0$ by (5), which together with $p \vee \neg p = 1$ makes $\neg$ complement. ∎

A regular element of a Heyting algebra is one satisfying $\neg\neg p = p$.

We remark that by (7) the regular elements of a Heyting algebra are just those of the form $\neg p$ for some p in H, that is, the image $\neg H$ of H under $\neg$ is the set of regular elements of H.

Theorem 3 *The set $\neg H$ of regular elements of a Heyting algebra H form a Boolean algebra, with complement the $\neg$ of H and with $p \to q$ in H being $\neg p \vee q$ in $\neg H$.*

Proof: We show that $\neg H$ under the ordering on H is a lattice. Since $\neg p \wedge \neg q = \neg(p \vee q)$ in H, the glb in H of $\neg p$ and $\neg q$ is in $\neg H$, and so *a fortiori* $\neg(p \vee q)$ is their glb in $\neg H$, so finite glb's in $\neg H$ are as in H. For lubs, note by (6) that $\neg\neg(p \vee q)$ is an upper bound on p and q.

Furthermore, for any upper bound r on p and q that is in $\neg H$ we have $p \vee q \leq r$, whence $\neg\neg(p \vee q) \leq \neg\neg r = r$ ($\neg\neg$ is monotone, by (4) twice), so $\neg\neg(p \vee q)$ must be the *least* upper bound on p and q in $\neg H$. Hence $\neg H$ forms a lattice.

Distributivity follows from $\neg\neg(p \vee (q \wedge r)) = \neg\neg((p \vee q) \wedge (p \vee r)) = \neg\neg(p \vee q) \wedge \neg\neg(p \vee r)$.

Since $p \wedge \neg p = 0$ by (5) and $\neg\neg(p \vee \neg p) = 1$ by (11) $\neg$ is complement, establishing that $\neg H$ is a Boolean algebra with complement $\neg$. Now if $p \wedge r \leq q$ then $\neg\neg(p \wedge r) \leq \neg\neg q$, so $\neg\neg p \wedge \neg\neg r \leq \neg\neg q$. Hence if p and q are in $\neg H$, $p \wedge \neg\neg(p \to q) \leq q$. So $\neg\neg(p \to q) \leq p \to q$ since $p \to q$ is the greatest such. Combining this with (6) yields $\neg\neg(p \to q) = p \to q$. But in a Boolean algebra $p \to q$ is $\neg p \vee q$. ■

3.1.5 General Properties of Heyting Algebras

Theorem 4 *The following are all properties of Heyting algebras.*

$$r \leq p \to q \quad \text{iff} \quad p \wedge r \leq q \tag{1}$$

$$p \to q \leq p \to (q \vee r) \tag{2}$$

$$(p \vee r) \to q \leq p \to q \tag{3}$$

$$\neg(p \vee q) \leq \neg p \tag{4}$$

$$p \wedge \neg p = 0 \tag{5}$$

$$p \leq \neg\neg p \tag{6}$$

$$\neg\neg\neg p = \neg p \tag{7}$$

$$\neg 0 = 1 \tag{8}$$

$$\neg 1 = 0 \tag{9}$$

$$\neg(p \vee q) = \neg p \wedge \neg q \tag{10}$$

$$\neg\neg(p \vee \neg p) = 1 \tag{11}$$

Proof: Since $p \wedge (p \to 0) \leq 0$ we have $p \wedge \neg p \leq 0$, whence (5). So $x \leq (\neg x \to 0)$, whence (6). Substituting $\neg x$ for x in (6) yields $\neg x \leq \neg\neg\neg x$, while antimonotonicity of $\neg$ (expressed by (4)) converts (6) to $\neg x \geq \neg\neg\neg x$, whence (7). (8) and (9) are easy. For (10)

$$\begin{aligned}
\neg p \wedge \neg q \wedge (p \vee q) &= (\neg p \wedge \neg q \wedge p) \vee (\neg p \wedge \neg q \wedge q) \\
&= 0, \qquad \text{by(5)}
\end{aligned}$$

so $\neg p \wedge \neg q \leq (p \vee q) \to 0 = \neg(p \vee q)$. Also $\neg(p \vee q) \leq \neg p \wedge \neg q$ by (4), giving us (10). We now use (10) to get $\neg(p \vee \neg p) = \neg p \wedge \neg\neg p = 0$, whence $\neg\neg(p \vee \neg p) = \neg 0 = 1$, establishing (11). ■

The following is Heyting's axiomatization (1930) of intuitionistic logic.

I. $p \to (p \wedge p)$.
II. $(p \wedge q) \to (q \wedge p)$.
III. $(p \to q) \to ((p \wedge r) \to (q \wedge r))$.
IV. $((p \to q) \wedge (q \to r)) \to (p \to r)$.
V. $q \to (p \to q)$.
VI. $(p \wedge (p \to q)) \to q$.
VII. $p \to (p \vee q)$.
VIII. $(p \vee q) \to (q \vee p)$.
IX. $((p \to r) \wedge (q \to r)) \to ((p \vee q) \to r)$.
X. $\neg p \to (p \to q)$.
XI. $((p \to q) \wedge (p \to \neg q)) \to \neg p$.

3.1.6 Exercises

1. Given a Boolean term t, let t' be the result of replacing $\wedge$ by $\vee$ and vice versa throughout t. Show that t and t' denote dual Boolean operations.

2. Show that a necessary condition that an n-ary Boolean operation be affine is that the number of inputs on which it is 1 is an integer multiple of 2^{n-1}. Show that this condition is sufficient if and only if $n \leq 2$. (Here $2^{n-1} = \frac{1}{2}$ as usual.)

3. Show that composition preserves each of strictness, costrictness, monotonicity, affineness, and selfduality. (That is, an expression built up from variables using strict operations defines a strict operation, etc.)

4. Give a minimal Boolean basis having four operations. Show that there is no minimal basis having five operations.

5. Show that if the requirement "nonzeroary" is omitted from the definition of functional completeness then we may replace "nonself-dual" by "zeroary" in the statement of Post's theorem.

6. Every Heyting algebra is a distributive lattice.

7. Prove properties (1)-(4) of the properties of Heyting algebras listed above.

8. Show that Heyting algebras contain a maximal element.

9. Prove the equivalence of the two definitions of $\rightarrow$ in a Heyting algebra.

10. Show that every complete completely distributive lattice is a Heyting algebra but not conversely.

11. Show that the lattice of positive divisors of $p^m q^n$, where p, q are prime and m, n are positive integers, is a Heyting algebra with $(i \rightarrow j) \vee (j \rightarrow i) = 1$ for all elements i, j.

12. Show that all of Heyting's axioms are equal to 1 in a Heyting algebra.

3.2 Modal Logic and Kripke Structures

3.2.1 Modal Logic

Modal or intensional logic originated with Aristotle in around 330 BC. It expands logic with a unary operation $\Diamond p$, called a *modal operator* or a *modality.* Propositional modal logic so expands propositional logic, first order modal logic so expands first order logic, and so on. We confine our attention here to propositional modal logic.

In propositional modal logic, a formula A is either (i) a propositional variable, (ii) a Boolean combination of formulas such as $B \vee C$, $B \wedge C$, or $\neg B$, or (iii) $\Diamond B$. Dual to $\Diamond$ is the operator $\Box$, which can either be taken as an abbreviation of $\neg\Diamond\neg$ or considered as an unary operation in its own right sibling $\Diamond$.

Normal modal logic, the only kind we shall consider, can be axiomatized as for propositional logic, along with an additional axiom K and a second rule to accompany propositional logic's rule of Modus Ponens. These may be presented in Hilbert style as follows.

K. $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$
R2. From A infer $\Box A$

Equivalently they may be presented as equations.

K. $\Diamond(x \vee y) = \Diamond x \vee \Diamond y$
R2. $\Diamond 0 = 0$

Treating $\Box$ as the abbreviation of $\neg\Diamond\neg$, these two equations could just as well be written,

K. $\Box(x \wedge y) = \Box x \wedge \Box y$
R2. $\Box 1 = 1$

Theorem 5 *These three forms of axiom K and rule R2 are equipollent.*

Proof: The second and third are trivially equivalent. The equivalence of all forms of R2 is justified by taking the theorems of propositional modal logic to be those terms x satisfying $x = 1$. The Hilbert style rule then says that $A = 1$ implies $\Box A = 1$, i.e. that $\Box 1 = 1$.

To deduce the first version of axiom K from the third, rephrase the first version as $(\Box(A \to B) \wedge \Box A) \to \Box B$ and then by the third version we have $(\Box((A \to B) \wedge A) \to \Box B$ which simplifies to $(\Box(B \wedge A) \to \Box B$ and thence (using the third version again) to $(\Box B \wedge \Box A) \to \Box B$, a propositional tautology. The other direction is left as an exercise. ∎

3.2.2 Kripke Structures

A ***modal algebra*** is a Boolean algebra expanded with a unary operation $\Diamond$ satisfying K and $R2$ presented as equations as in the second version above of the axiomatization of normal modal logic. As for the equational theory of Boolean algebras, modal algebras have a finitely based equational theory, in the sense that finitely many equations suffice to axiomatize that theory.

Whereas the unary operation of negation is easily understood in terms of its action on the two-element lattice, $\Diamond p$ is not since every possible unary operation on $\{0, 1\}$ is already accounted for as a Boolean operation. Thus any lattice of truth values in which modal logic is discernibly different from Boolean logic must have at least three values.

A basic way in which such a lattice can be constructed is in terms of the following notion. A ***Kripke structure*** (W, R, P) consists of a set W of ***possible worlds***, a binary relation $R \subseteq W^2$ called the ***accessibility relation***, and a concrete Boolean algebra $(P, \cup, \cap, W-)$ consisting of subsets of W and closed under finite union, finite intersection, and complement relative to W.)

In a Kripke structure, the interpretation of a proposition is that it holds (is true) in just those worlds belonging to it. Thus union, intersection, and complement provide interpretations for conjunction, disjunction, and negation respectively.

So far we have not used the accessibility relation. The interpretation of $\Diamond x$ is the set $\{u \in W | \exists v \in x.uRv\}$. That is, $\Diamond x$ holds in all worlds u such that x holds in some world v "accessible from" u via R. We think of R as relating the world u in which we utter $\Diamond x$ to all conceivable alternative worlds to u. Thus $\Diamond x$ means that x is possible in the sense that it is true in some world that we can conceive of as alternative to the present one.

The usual notion of propositional Boolean logic corresponds to taking W to be a singleton: with only two subsets of W available as propositions, the only propositions are true and false, interpreted as respectively W and the empty set. The only relations R on a singleton are the identity relation, which makes $\Diamond x = x$, the identity operation, and the empty relation, which makes $\Diamond x = 0$, the constantly zero operation. Neither adds anything new to Boolean logic.

With two possible worlds we start to see more interesting diversity. Here we have four possible propositions, which we could identify with the propositions *true*, *false*, A, and $\neg A$, i.e. the free Boolean algebra on one generator. It remains the case that taking R to be the identity relation or the empty relation makes $\Diamond$ respectively the identity operation or the constantly false operation.

Taking it to be the clique however does not make $\Diamond$ the constantly true operation: for one thing this would contradict R2. However this is all it contradicts, and as the reader may calculate, $\Diamond x = 1$ provided $x \neq 0$.

It is natural to think of the world we are in as one of its own alternatives. This corresponds to the requirement that R be reflexive. In this case we have

T. $\Box A \rightarrow A$

In equational terms T becomes $x \leq \Diamond x$ or dually $\Box x \leq x$. This axiom is sound in the sense that it must be true in every possible world when R is reflexive.

It is reasonable to ask what other axioms should be introduced at this point. If none then we would have a complete axiomatization of reflexive Kripke structures. In retrospect we see that we neglected to ask whether the axiomatization of normal modal logic was complete for arbitrary Kripke structures. Let us now address both of these questions.

Theorem 6 *Axiom system K is complete for arbitrary Kripke structures.*

We take this to mean that every identity (universally true equation) holding in every Kripke structure is an equational consequence of the equations of system K. It is straightforward to obtain the corresponding result treating K as a Hilbert style system and taking the criterion for theoremhood to be truth in every world of every Kripke structure.

Proof: Every equation $s = t$ is interderivable with $s \oplus t = 0$ where $\oplus$ is exclusive or, by purely Boolean reasoning. Thus it suffices to show that any modal term t for which $t = 0$ is not derivable has a Kripke structure in which t does not evaluate to 0, i.e. is nonempty, i.e. holds in some world of that structure, i.e. is *satisfiable.*

Normalize term t by pushing negations down through the logical connectives and modalities to the propositional variables. We now have a formula built from literals (possibly negated variables) using $\vee$, $\wedge$, $\Diamond$, and $\Box$. We proceed to construct a Kripke structure satisfying t.

Take W to consist of the term t itself along with all subterms of t of the form $\Diamond u$ for some term u (itself therefore a subterm of t). W is partially ordered by the subterm relation: $u \leq v$ just when u occurs as a subterm of the term v. Take R to be the transitive hull of $\geq$. That is, uRv holds just when v is an immediate subterm of u, in the sense that there is no other subterm of u having v as a subterm. The idea is that we wish to move through W via R as we pass from the outside to the inside of any given subterm $\Diamond s$.

∎

3.2.3 Exercises

1. Definitions. A ***modal bounded lattice*** is as for a modal algebra, but with "bounded lattice" (a lattice with constants 0 and 1) in place of "Boolean algebra". (Thus the signature is $(\wedge, \vee, 0, 1, \Diamond)$, satisfying the equations for a bounded lattice plus K and R2.) ***Modalic*** is as for "modal" but without R2. An ***atom*** of a poset with 0 is an element x such that $[0, x] = \{0, x\}$, where the interval $[x, y]$ is defined as $\{z | x \leq z \leq y\}$.

(i) Describe the initial modal bounded lattice, and show that it is (a) distributive, (b) complete, and (c) contains no atoms.

(ii) Describe the initial modalic bounded lattice and show that it is (a) distributive, (b) not complete, and (c) contains an atom.

2. A modal algebra is *representable* when it is isomorphic to the modal algebra defined by some Kripke structure. (That is, as a Boolean algebra it is isomorphic to some power set 2^W under union, and $\Diamond$ is the modality associated to some binary relation R on W.) Show that the direct product of two representable modal algebras is representable.