Indicative:

If this sample *is* burning green, then if it *is* a sodium salt

(a) it *is* a sodium salt burning green

(b) it *is* burning yellow

Subjunctive:

If this sample *were* burning green, then if it *were* a sodium salt

(a) it *would* be a sodium salt burning green

(b) it *would* be burning yellow

I think that the natural interpretation in each case could choose the first continuation as trivially true and reject the second as trivially false. Nevertheless, in the subjunctive case (though not in the indicative) I can imagine appropriate promptings and side remarks that would lead me to take some variant of the counterfactual in the Stalnaker way:

If this sample were burning green (say it was barium) then it would still be true that had it been sodium it would have burned yellow.

The question of what cues in English lead you to take a counterfactual one way rather than the other is, I think, a very complicated business.

Appendix 4
Nonstandard Analysis and Infinitesimal Probabilities

Compactness

From Henkin's completeness proof we know that first-order logic is *compact*. If a set of sentences is such that every finite subset of it has a model, then the set in question has a model. The Henkin proof in no way depends on the assumption that the set of sentences or the set of constants of the language in question is denumerable. (That every set with the property that every finite subset of it has a model can be extended to a maximal set with that property, holds for sets of arbitrary cardinality by transfinite induction. The union of a chain of sets with that property must have that property.) It does depend on the sentences of the language being of finite length, and the logical constants being limited to the truth functions, identity, and first-order quantifiers. (For the proof that the model associated with the maximal consistent set is a model of each sentence in it is by induction on the length of the sentences.) Thus by Henkin's proof (or by a variation on it on the level of the models, the ultraproduct construction) we see that we have compactness for rich, nondenumerable first-order languages (e.g., for a first-order language with a name for every real number and operation, and relation symbols for every operation and relation on the reals). Compactness fails for higher-order quantification if second-order quantifiers are given the "natural" interpretation of having as their domain the power set of the domain of the first-order quantifiers, and so forth. Given the natural interpretation of second-order quantifiers, we can in second-order logic write a sentence, A, which has all the truths of arithmetic as logical consequences. Then $\{A, (\exists x)(\sim Fx), F1, F2, F3, \ldots\}$ is an infinite set which has no model but such that each finite subset of it has a model. But if we relax the interpretation of higher-order quantifiers, so that a permissible model (Henkin calls these *general* models) results whenever the higher-order quantifiers are taken as ranging over a *subset* of their natural domain, then the Henkin strategy for constructing the domain of a model from the constants occurring in a maximally consistent set of sentences succeeds for higher-order domains as well. In this sense, higher-order logic with the usual quantifier rules is complete and compact. But, once general models are admitted, higher-order logic cannot categorically characterize arithmetic anymore than first-order logic can. Nonstandard general models of arithmetic (and analysis) are now possible. It is a mark of Abraham Robinson's genius that he turned this logical defect into a powerful tool of discovery.

Nonstandard Arithmetic

For the time being we will restrict our attention to first-order logic. Let us choose as our language of arithmetic a first-order language containing a name for every number operation; symbols for successor, plus, and times; and a relational symbol for less than. Let *Arithmetic* be the set of all true sentences of this language. Take a variable, *y*. Consider the theory Arithmetic $\cup \{y \neq 0, y \neq 1, y \neq 2, \ldots\}$. By compactness, this theory has a model (and by the Lowenheim-Skolem theorem a denumerable model). Designate the elements of this model *numbers** and in general designate the denotation assigned by this model to any constant by the constant followed by an asterisk. Thus *less than** is the extension assigned by the nonstandard model to the less-than relation.

Since the axioms for a linear ordering are first-order, and since the nonstandard model *is* a model of arithmetic, *less than** linear orders the *numbers**. Since the claim that zero is the least number is a first-order truth, and since the nonstandard model *is* a model, 0* is the *least* number**. Likewise, 1* is the *next-to-least* number**, and so forth. The nonstandard model, considered under the order relation *less than**, begins with an ω-series: 1*, 2*, 3*. . . . The elements of this series together with the asterisked operations and relations on them are isomorphic

to the numbers, because the nonstandard model is a model (e.g., since "2 + 2 = 4" is a sentence of *Arithmetic*, 2* + *2* = 4*.) We may then, for all intents andpurposes, call the elements assigned as denotations to the numerals, 1*, 2*, 3*, . . . , the *numbers* (or, for emphasis, the *standard numbers*) and the structure consisting of them together with the restriction to them of the asterisked relations and operations, the *standard model of arithmetic*. The nonstandard model is, then, an end-extension of the standard model.

The inclusion of one nonstandard element in the model forces the inclusion of many others, since *Arithmetic* requires that every number have a successor greater than itself. Let us partition the *numbers** into equivalence classes by considering the equivalence relation *differs* by a standard number* (i.e., there is a standard, z, such that x +* z = y or y +* z = x). Call these equivalence classes *Blocks*. The standard numbers form one *Block*. There must be at least one *greater* Block* since there are nonstandard numbers. But for every nonstandard *Block* there must be a greater *Block* to which we may pass by *multiplying** by *2**. For if *2* times* y* were in the same *Block* as *y*, then by the definition of *same Block* they would *differ** by a standard number, but they *differ** by *y* ("Twice *y* less *y* is *y*" is a sentence of *Arithmetic*), contradicting the assumption that *y* is nonstandard. A like argument will show that there is no least *Block* and that the *Blocks* are densely ordered by *less than**. Consider the operation, *halb*, of approximate division by two. *Halb d* = *c* iff 2*d* = c or 2*d* + 1 = *c*. Now if *d* is in a nonstandard *Block*, *halb* d* is in a *lesser** nonstandard *Block*. For if they were in the same *Block*, *d* would have to be standard, contrary to hypothesis. And if *halb* d* were standard, *d* would be also, contrary to hypothesis. Likewise, *between** any two *Blocks* there must be another, for if *c* and *d* are in different *Blocks* *halb* c* +* *d* must be in another *Block* (all of whose members are) *between* c* and *d*. The *Blocks* must then be ordered with dense order: no first, no last element. Inside the *Blocks*, we may rely on the sentences of *Arithmetic* that say that every number has a successor and that nothing comes after the number but before its successor, and that every number other than zero has a predecessor, and that no number comes after its predecessor but before it. This, together with the constraints of the equivalence relation defining the *Blocks*, tells us that internally each *Block* is ordered as the negative integers, zero, and positive integers. If we confine ourselves to *countable* nonstandard models, this then fixes the order type of the nonstandard model.

Nonstandard Rationals

Let the language for the rationals contain a name for each rational; operations of addition, multiplication, and division; a less-than relation; and the predicate "is a natural number." Let *RAT* consist of all the true sentences of this language. Consider the theory $RAT \cup (y > r_1, y > r_2, y > r_i, \ldots)$ for some enumeration (I) of the rationals. Every finite subset of this set has a model in the rationals, so by compactness it does as well. (And, again, by the Lowenheim-Skolem theorem, it has a countable model.) Once more we will identify the denotata of the r_is as the standard rationals. So the nonstandard model contains an infinite element *greater than** all the standard rationals. Multiply it by its divisor to get an infinite natural *number** and the argument establishing the structure of the nonstandard natural *numbers** can proceed as with nonstandard arithmetic. The countable nonstandard model of the rationals is an extension of the countable nonstandard model of arithmetic. Since a proposition of *RAT* asserts that every rational has a reciprocal and that taking the reciprocals of two numbers inverts the order, the infinite elements must have reciprocals which are *less than** any positive standard rational and *greater than* 0**. That is, we have infinitesimal elements. Notice that we have quite a rich structure of infinitesimals due to the interaction of closure under addition, multiplication, and division. Thus, if we have an infinite element, *N*, we have the corresponding infinitesimal $\epsilon = 1^*/N$, as well as ϵ^{2*}, ϵ^{3*}, $1^*/N +^* \epsilon$, etc. Let us say that a nonstandard rational is *finite* if it is bounded above and below by standard rationals other than *zero**, *infinite* if it is *greater than** any finite number,

and *infinitesimal* if it is *less than** any finite number. Let us say that two nonstandard rationals are of the same *Order* if their quotient is finite. *Orders* provide a coarser partitioning of the infinite elements than the *Blocks* considered in the discussion of nonstandard arithmetic. In fact, . . . *N*/4, *N*/2, N, 2*N*, 4*N* . . . are all of the same *Order*. Nevertheless, for every infinite *Order*, there is a *greater** one. N^{2*} is greater than *N* and of a different *Order*.[1] Furthermore, for every infinite *Order* there is a *lesser** one. Let us say that *y* is an *approximate square root of* x if $x \leqslant^* y^{2*} \leqslant^* x +^* 1^*$. A sentence of *RAT* says that everything has at least one approximate square root.[2] An approximate square root of an infinite number cannot be finite, since the finite numbers are closed under addition and multiplication, and it cannot be of the same *Order* as x since its square is. Along the same lines, we can show that *between** each two infinite *Orders* there is another one using as the leading idea "an approximate geometric mean." Since two infinite elements are of the same *Order* just in case their reciprocals are, and since the *Order* of the reciprocals is the inverse of the *Order* of the infinite elements, this shows that the *Orders* of infinitesimals (excluding zero) are densely ordered with no first or last element.

Nonstandard Analysis

In the preceding sections we took pains to keep to a denumerable model, but here we will not, so we may allow the luxury of starting with a truly opulent first-order language. Let our language of analysis include a name, c_r, for every real, *r*; a relational symbol for every relation on the reals; and an operational symbol for every operation on the reals. Let *ANA* be the set of all true sentences of this language, and consider the theory which is the union of *ANA* with the set of all sentences of the form $c_r < y$ for

1. Notice that this shows that we did not look at *all* the nonstandard natural numbers in the argument which established the order type of the denumerable nonstandard model of arithmetic.

2. Note that we could approximate *infinitely* close to $\sqrt{2}$, for instance, since a translation of $(z)(x)(x \neq \supset (\exists y)(|y^2 - x| \leqslant z))$ is in RAT.

each real r. By compactness, this theory has a model, a nonstandard model of the reals. Again, the function which maps each real, r, on to $c_r{}^*$, the denotation in the nonstandard model of its name, is an isomorphism. Each nonstandard model contains an isomorphic copy of the reals. Working within the model we will simply call these the *standard reals*. The denotation of the less-than relation, $<^*$, totally orders the nonstandard reals, R^*, since the axioms of total order are first-order. The nonstandard reals form a field (that is, $\langle R^*, 0^*, 1^*, +^*, \cdot^* \rangle$ is a field), since the properties of a field are expressible by first-order axioms. It need not have properties which require second-order axioms for their expression. It is not Archimedean. It does not have the least-upper-bound property. The standard reals provide an example of a set in the model which has an upper bound but no least upper bound.

If the set of standard reals which satisfies F is unbounded in the standard reals, then $^*F$ has an infinite element in the nonstandard model. For then the first-order sentence $(x)[Fx \supset (\exists y)\ (Fy \text{ and } y > x)]$ is in *ANA*, and taking x as the infinite element that we constructed with the model forces F to contain an infinite element. Thus the set of nonstandard natural numbers, N^*, and the set of nonstandard rationals, Ra^*, contain infinite elements. The statement that every real is bounded by natural numbers is in *ANA*. So we can repeat the arguments of the previous sections to show that *Blocks* and *Orders* of infinite elements are densely ordered with no first or last elements, and likewise for the infinitesimals (excluding 0^*). This, however, no longer settles the question of order type since the model is nondenumerable.

Let us say that x if *infinitely close to y* if $|x - {}^*y|^*$ is infinitesimal. Infinitely close to (symbolically $\simeq$) is an equivalence relation on R^*. For standard reals $r \simeq s$ implies $r = s$, since zero is the only standard infinitesimal. If $x \not\simeq y$ and at least one of them is finite, then there is a standard real, q, between x and y. For suppose that $0^* \leqslant^* x \leqslant^* y$. By the definition of $\simeq$ there is a standard, b, such that $0^* \leqslant^* b \leqslant^* y -^* x$. Choose the least standard integer,

m, such that $mb >^* x$. Then $x <^* mb <^* y$. *Every finite nonstandard real is infinitely close to a unique standard real.* For consider the set of all *standard* reals less than x. It has a standard upper bound (since x is finite), so it has a least upper bound, r, in the standard reals. Then x is infinitely close to r, since there is no standard q between x and r. Furthermore, r is the *unique* standard real that is infinitely close to x; since if y is infinitely close to r, then by transitivity y is infinitely close to x, and if y as well as x is standard, then they can be infinitely close only if identical.

Let F be a function on the reals. Then the standard definition of "F converges to a at b" is:

$$(\epsilon)[\epsilon < 0 \supset (\exists\delta)(x)(|x - a| \neq 0 \;\&\; |x - a| < \delta \supset |b - F_*| < \epsilon].$$

The nonstandard definition is: F converges to a at b iff whenever x is infinitely close to (but different from) a, $F^*(x)$ is infinitely close to b. The two definitions are equivalent. *The standard definition implies the nonstandard one:* Suppose that the standard definition holds. Then its true instances are in *ANA*. That is, for a fixed real number, E, greater than zero, we will have a sentence:

$$(x)[|x - a| < D \supset |b - F(x)| < E]$$

in *ANA* where D is a fixed real number. These sentences must be satisfied in the nonstandard model. Then if x is infinitely close to a, for each such sentence we have $|x -^* a^*|^* <^* D^*$ and thus $|b^* -^* F^*(x)|^* <^* E^*$. But since we have such a sentence for every standard real with E naming that real, $F^*(x)$ is infinitely close to b. *The nonstandard definition implies the standard one.* Suppose the nonstandard definition holds. Then consider any instantiation of the standard definition to a fixed *standard* real, E:

$$(\exists\delta)(x)[|x - a| < \delta \supset |b - F(x)| < E]$$

This sentence is true in the nonstandard model, as can be seen by choosing δ infinitesimal. It is a first-order sentence, so it holds in the standard model. Thus, reasoning about the behavior of infinitesimals in the nonstandard model can yield truths about limits in the standard model. Many of the arguments of Newton and Leibniz can, from our vantage point, now be seen to have just this character.

Robinson's interest in the nonstandard model was centered on its use as a tool to prove theorems about the standard model.

Nonstandard Measure Theory

Here we assume we have a nonstandard general model of analysis, where the first-order language of analysis of the previous section is extended to type theory. (Types are taken as follows: The set of types is the smallest set such that (1) 0 is a type and (2) if $t_1, t_2, \ldots t_n$ are types, then $(t_1, t_2, \ldots t_n)$ is a type.) The general model allows the higher-order quantifiers to have as their domain some subset of their "natural domain."[3] For instance, quantifiers of type (0) may not range over all subsets of R^*. The elements of the model that are within the domains of the quantifiers are called *internal*.

A relation $R(xy)$ is said to be *finitely satisfiable* if for every set of elements in its domain, $a_1, a_2, \ldots a_n$, there is a y such that $R(a_i, y)$ holds for each a_i. If R is a finitely satisfiable relation on the standard reals (or the higher-order structure built up from them), then, by compactness, a nonstandard model can be found such that it contains an element, y, such that given any standard x in the domain of R, $R^*(x^*, y)$. The arguments of the previous sections that established the existence of nonstandard models with infinite elements are a special case, with the finitely satisfiable relation taken us $<$. The nonstandard model can be arranged to contain elements such as y which simultaneously satisfy the relation R for all standard elements in their domain, for as many finitely satisfiable relations as you please, since the set of sentences $\{R_j(a_{ij}, y_j)$ for every a_{ij} in the standard domain of $R_j\}$ still has the property that every finite subset of it has a standard model.

The leading idea of nonstandard measure theory is the use of *finite samples. The second-order predicate "is finite" has as its extension in the nonstandard model a set

3. The "natural domain" of the quantifiers of higher type induced by a domain for type zero is defined as follows: the domain of type 0 is the natural domain of type zero ($= ND(t_0)$). If $t = (t_1, t_2, \ldots t_n)$, then $ND(t) =$ the set of all subsets of $ND(t_1) \times ND(t_2) \times \ldots \times ND(t_n)$.

of sets of nonstandard reals, χ. We call the members of χ the **finite* sets of nonstandard reals. Some *finite sets contain an infinite number of elements. (For instance, the set of nonstandard natural numbers less than some infinite nonstandard natural number is infinite, yet *finite.) The relation "is the cardinality of" assigns each *finite set a nonstandard natural number as its *nonstandard cardinality*. This allows us to use a *counting measure* for infinite sets which are *finite. A *finite set, F, is called a sample. Relative to such a fixed sample, we can take the *measure of any internal set*, A, as the nonstandard cardinality of its intersection with the sample, F, divided by the nonstandard cardinality of the sample. If A and B are internal sets, we can take the *conditional probability* of B on A as the nonstandard cardinality of the intersection of A, B, and F over the nonstandard cardinality of the intersection of A and F.

Some desirable properties follow from the *finiteness condition on the sample: the measure of the null set is zero; measure is monotonic, if $A \subseteq B$, the measure $A \leq$ the measure B; measure is additive. Other desirable properties can be secured by judicious selection of the sample. Bernstein and Wattenberg have shown that there is a sample with the following characteristics: the associated measure is defined for all subsets of the unit interval which is infinitely close to Lebesgue measure for all Lebesgue measurable sets; it assigns each nonempty set a positive (possibly infinitesimal) measure; and it is translation-invariant up to an infinitesimal. (The strategy of the existence proof is to show that the appropriate relation, specifying the desirable properties, is finitely satisfiable.) Parikh and Parnes have carried through the analogous investigation, putting the constraints directly on the conditional probability function and showing the existence of samples which yield associated conditional probabilities with desirable properties. Loeb has studied nonstandard measures on abstract spaces.[4]

4. For details, see these three papers and the references to related work cited therein: A. Bernstein and F. Wattenberg, "Non-Standard Measure Theory," in W. Luxemburg, ed., *Applications of Model Theory to Algebra, Analysis, and Probability* (New York: Holt, Reinhart and

Infinitesimal Probabilities

The Vitali-Hausdorff example of a nonmeasurable set shows that no sigma-additive, translation-invariant, real-valued measure can be defined on all subsets of the interval [0, 1). A wheel of fortune is spun and comes to rest with some point or other at the lowest point. We can give the wheel unit circumference and label its points with the numbers in [0, 1). Assume the equiprobable distribution. More specifically, assume that, for any point set, the probability that the wheel stops with a point in that set as the bottommost one is equal to the probability for any point set gotten from the first by displacing each member of the first through a fixed angle, 0 (translation-invariance under addition modulo 1.) Consider the equivalence relation: $x - y$ is rational. This partitions [0, 1), and thus the points in the circumference of our wheel, into equivalence classes. Consider a choice set, C, containing one member of each of these classes. For each rational in [0, 1) let C_r be the set gotten by adding (modulo 1) r to each member of C (i.e., by translating the point set a rational distance around the circumference). There are a denumerable number of C_rs; they are mutually exclusive; and their union is [0, 1). They are equiprobable by translation invariance modulo 1. If propensities are real-valued, then either $\Sigma P_r(C_r) = 0$ or $\Sigma P_r(C_r) = \infty$. If propensities are, furthermore, sigma-additive, then $0 = 1$ or $\infty = 1$.

These difficulties can be overcome if the requirement of sigma additivity is relaxed to that of finite additivity. Banach has shown (as a corollory to the Hahn-Banach theorem) that Lebesgue measure can be extended to a finitely additive, translation-invariant measure defined in all subsets of [0, 1) (indeed, defined on all subsets of the reals). The measure so defined is not, however, regular (strictly coherent). There are sets other than the empty set

Winston, 1969), pp. 171–85; R. Parikh and R. Parnes, "Conditional Probabilities and Uniform Sets," in P. Loeb, ed., *Victoria Symposium on Non-Standard Analysis, 1972* (Heidelberg: Springer, 1974), pp. 180–94; P. Loeb, "A Non-Standard Representation of Borel Measures and σ-Finite Measures," in *Victoria Symposium on Non-Standard Analysis, 1972*, pp. 144–52.

which have Lebesgue measure zero. Thus, on this approach, the probability in our example of the pointer stopping on a rational would be zero, although this might happen. Then the conditional probability (conceived of in the Kolmogoroff way) of stopping on a rational in [0, 1/2) given that it stops on a rational would be undefined. But, as De Finetti and Savage have emphasized, intuitively it should be defined.

The problem is more general than has so far been indicated, however. Lebesgue measure aside, there is no finitely additive, translation-invariant, real-valued measure defined on all subsets of [0, 1) that is *regular*. For, considering the Vitali-Hausdorff example again, the C_rs are equiprobable by translation invariance.[5] If they have a non-zero, real-valued probability, then by the Archimedean property of the reals there is an integer, n, such that n times their probability is greater than 1. Thus finite additivity leads to a contradiction. So the C_rs have zero probability, and the measure is not regular.

If the measure has values in a non-Archimedean ordered field, as in nonstandard analysis, then the paradox is avoided. And, as Bernstein and Wattenberg have shown, there is a finitely additive, almost translation invariant, regular measure defined on all subsets of [0, 1). Nonempty sets of Lebesgue measure zero, as well as the C_rs, then receive infinitesimal measure. And we can say that the probability of our pointer hitting a rational in [0, 1/2), given that it hits a rational, is the ratio of the two appropriate infinitesimals and equals 1/2.

5. Notice that translation invariance implies translation invariance modulo 1.