A Recipe for Self-Consistent Projection

1.

Most scientists, if asked to define induction, will say something like, 'When we have observed that a certain set of circumstances has always been accompanied by a certain observation, induction allows us to infer that those circumstances will always be accompanied by that observation.' Philosophers have many times shown that this casual definition of induction is inadequate. In particular, Goodman[1] demonstrated that it is not self-consistent. In his presentation of the *New Riddle of Induction*, he defined *grue* to be green if already observed, but blue if not yet observed. Given that all emeralds already observed are green, we see that they are all grue as well. Therefore the casual definition of induction allows us to infer not only that all emeralds yet to be observed will be green, but also that they will be grue, which is to say that they will not be green.

Most of us would rather not project grue. Many efforts have therefore been made to determine a criterion by which well-behaved predicates such as green can be distinguished from ill-behaved predicates such as grue. Armed with such a criterion, we could compose a definition of induction that would prohibit the projection ill-behaved predicates, and thus solve the New Riddle of Induction. The chief obstacle to obtaining such a criterion is the analytical duality between grue and green. To illustrate this duality, Goodman defined *bleen* as blue if already observed, but green if not yet observed. Thus green can be described as grue if already observed, but bleen if not yet observed, thereby expressing green in terms of grue and bleen in exactly the same way that grue was originally expressed in terms of green and blue. As a consequence of this duality, all efforts to identify ill-behaved predicates by purely analytical means have failed.

Efforts have also been made to identify ill-behaved predicates by a combination of analytical, empirical, and even metaphysical means, but none of these has gained wide acceptance. Goodman himself suggested that the 'relative entrenchment' of a predicate might be used to determine whether it is worthy of projection, but this amounts to nothing more than trusting our instincts.

I propose a simple and purely empirical criterion for identifying well-behaved predicates. Loosely speaking, I argue that only well-behaved predicates can be represented directly in a set of observations that contains no duplication or combination of observations. By means of this criterion, I extend the casual definition of induction so that it forbids the projection of ill-behaved predicates. In the course of my discussion, I show how this extended definition solves Goodman's problem with emeralds, and also solves other, more sophisticated manifestations of the New Riddle of Induction.

2.

In the definition of grue, it has become popular to say, 'before the year 2000,' instead of, 'already observed.' The observations defined in the modern version of Goodman's riddle are,

green(n) = true if and only if (iff) emerald n is green,
blue(n) = true iff emerald n is blue,
before(n) = true iff it is before the year 2000 when emerald n is examined,
grue(n) = [green(n) and before(n)] or [blue(n) and not before(n)].
bleen(n) = [blue(n) and before(n)] or [green(n) and not before(n)].

From these we obtain the analytical dualities,

green(n) = [grue(n) and before(n)] or [bleen(n) and not before(n)],
blue(n) = [bleen(n) and before(n)] or [grue(n) and not before(n)].

Every inductive exercise begins with a set of observations upon which inductive inferences are to be based. Let us refer to this set as the *data*. For a study of emeralds, the data might be (where T=true and F=false):

n	**green(n)**	**grue(n)**	**before(n)**	**blue(n)**	**bleen(n)**
1	T	T	T	F	F
2	T	T	T	F	F
3	T	T	T	F	F

Table 1: Data for a Study of Emeralds

We now investigate the *empirical*, rather than the analytical, dependency between the observations in Table 1. Suppose it is close to midnight on the last day of 1999, but we have lost track of the time, and do not know whether it is the year 2000 yet or not. Looking at a green emerald, we can tell immediately that it is green. *But looking at a grue emerald, we cannot tell that it is grue*. A grue emerald is either blue or green, depending upon the date. If our emerald is green, we can be sure that it is grue only if we are sure that the time is before midnight. If our emerald is blue, we can be sure that it is grue only if we are sure that the time is after midnight. But by assumption we do not know the time, so we cannot tell that a grue emerald is grue.

Thus, despite the analytical duality between grue and green, the two can be distinguished by considering their empirical dependence upon the date. Given a grue emerald, we must observe before(n) in order to observe grue(n). Even if we build a machine that determines grue(n) for us, the machine will have to contain some sort of clock. But a machine that determines green(n) requires no clock.

Let us say that a set of observations is *distinct* if and only if each member of the set can be obtained without using or duplicating any other member. To prove that a set of observations is distinct, we must provide, for each observation in the set, a procedure by which it can be obtained without using or duplicating any of the other observations. Table 1 is not distinct because it is impossible to tell that a grue emerald is grue without knowing the date. It is also impossible to tell that a grue emerald is not bleen without knowing the date. On the other hand, it *is* possible to observe that a grue emerald is green without knowing the date.

It is also possible to observe that an emerald is not blue without observing that it is green in the process. For example, we might look at the emerald through a blue filter. When the emerald appears black through the filter, we would know that it was not blue. But we would not know if it were green, red, or even opaque. It is not possible, however, to determine that an emerald is green without observing that it is not blue in the process.

Applying our definition, we see Table 2 does not form a distinct set of observations.

n	**grue(n)**	**before(n)**
1	T	T
2	T	T
3	T	T

Table 2: Indistinct Observations

Table 2 is not distinct because grue(n) cannot be obtained without using or duplicating before(n).

n	grue(n)	green(n)
1	T	T
2	T	T
3	T	T

Table 3: More Indistinct Observations

Table 3 is not distinct either. There is no way to verify that an emerald is grue without using or duplicating an observation of green(n). We note that the converse is not the case. By depriving ourselves of knowledge of before(n), we guarantee that no observation of grue(n) is made when we verify that an emerald is green.

To solve the New Riddle of Induction, we generalize the empirical distinction between green and grue, and so produce a criterion for distinguishing between ill-behaved and well-behaved predicates. As a first step towards composing this criterion, let us propose that the projection of a predicate is justified only if observations of the predicate are to be found in a distinct subset of our data. This restriction prevents us from projecting grue(n) by means of Tables 1 or 2, since neither table is distinct.

Our restriction does not, however, prevent us from projecting grue(n) using the following table.

n	grue(n)
1	T
2	T
3	T

Table 3: Distinct but Unrepresentative Observations

Table 3 contains no observations of before(n), so each observation of grue(n) can be obtained without using or duplicating any other observation in the table. Thus Table 3 represents a distinct subset of our data, and contains observations of grue(n).

Table 3 does not, however, contain enough information for us to deduce all of the observations in Table 1. Knowing only grue(n), we cannot deduce green(n) or before(n). We shall say that a set of observations is *representative* of our data if we can deduce all our data from the observations it contains. Table 3, therefore, is not representative of Table 1. Table 4, however, is both distinct and representative of Table 1.

n	green(n)	before(n)
1	T	T
2	T	T
3	T	T

Table 4: Distinct and Representative Observations

A sufficient criterion for preventing the projection of ill-behaved predicates can be obtained by combining the requirements of distinct subsets with those of representative subsets. Consequently, the *rule of distinct observation* states that the projection of a predicate is justified by our data only if observations of that predicate can be found in a subset of the data that is both distinct and representative of the data.

When applied to our example, the rule of distinct observation allows us to project green(n), using Table 4, but not grue(n), since observations of grue(n) are not to be found in any distinct and representative subset of Table 1. Alternative definitions of grue are also dealt with effectively by the same rule. As mentioned above, in Goodman's original definition of grue, he used 'already observed' instead of 'before 2000.' If by 'already observed' he meant 'first observed before today,' then we see that we must know when an emerald was *first* observed in order to determine grue(n). If he meant 'is observed before today,' then today's date plays the same role as the year 2000 did above. On the other hand, if by 'already observed' he meant 'has been observed', then grue and green become, for all practical purposes, identical: the moment we look at an emerald, it has been observed, so every green emerald we observe is also guaranteed to be grue.

So far, I have allowed the notion of a 'well-behaved predicate' as opposed to an 'ill-behaved predicate' to be defined by our intuition. Let us adopt stricter definitions for both terms. We shall say that a predicate is *well-behaved* with respect to the data if and only if its observations are present in a distinct and representative subset of the data. We shall say that a predicate is *ill-behaved* with respect to the data if and only if its observations are present in the data, but it is not well-behaved with respect to the data.

According to these definitions, green(n) and before(n) are well-behaved with respect to Table 1, while grue(n), bleen(n), and blue(n) are ill-behaved, and red(n) is neither ill-behaved nor well-behaved, since no observations of it are present in Table 1. Furthermore, the rule of distinct observation may be restated as requiring that any projected predicate be well-behaved with respect to the data.

There is a further complication to be addressed in the application of the rule of distinct observation. Suppose that our data consists *only* of observations of grue(n).

That is, suppose our data consisted solely of the observations in Table 3. Such a set of observations is distinct and representative of itself, so grue(n) is well-behaved with respect to the data, and we can project it, saying that all emeralds are grue. Provided we know the relationship between grue, green, blue, and the date, we will be able to infer that emeralds will turn blue in the year 2000, and be green before that. But if we know this relationship, we will also know that grue(n) is ill-behaved in the presence of before(n). That is to say, we will know that we *cannot* project grue(n) if we add observations of before(n) to the data.

We shall say that an observation is *relevant* to our data if its addition would prevent a predicate from being projected. By excluding relevant observations from our data, we can project predicates that are ill-behaved with respect to the full extent of our knowledge. We shall make it a principle, therefore, to include all relevant observations in our data. Applying this principle, we would not be justified in using Table 3 as the data in a study of emeralds, since observations of before(n) are relevant to Table 3, and must therefore be included in the data.

The rule of distinct observation does not prevent us from projecting before(n) on the basis of Table 1. This is a problem of accidental generalization, and therefore not the subject of my discussion. The rule of distinct observation is phrased as a necessary, but not a sufficient, condition for projection. It can be combined with other rules that prevent accidental generalization.

3.

When a person who prefers the predicate grue meets a person who prefers the predicate green, they might end up arguing about which predicate is the best one to use when describing emeralds. Their debate will take on practical importance when the two speakers claim that the best predicate for projection is the one that they themselves prefer. The debate can then proceed in several directions.

In the first case, suppose the speakers are unaware of the true relationship between grue and green. The grue speaker knows when to call an emerald grue, but he does not know when to call an emerald green. Likewise, the green speaker knows when to call an emerald green, but she does not know when to call an emerald grue. They must therefore determine a practical relationship between grue and green by pointing at objects. It is before 2000, so every grue object is also green. By projecting their observations of the meaning of green and grue (subject to the rule of distinct observation if necessary), they conclude that grue will always have the same meaning as green. This conclusion is erroneous, but it nevertheless ends any contention between the two speakers.

In the second case, suppose both speakers are aware of the relationship between grue and green, and that the grue-speaker admits that he cannot observe grue without knowing the date. In this case, the grue-speaker must concede that the rule of distinct observation forbids the projection of grue. If he adopts the rule of distinct observation, he cannot project grue. If he does not adopt the rule, he will have to find some other way to make his use of induction self-consistent. Either way, the green-speaker is comfortable in her position. Her projection of green, and her refusal to project grue, are both justified by her use of the rule of distinct observation. Even if all emeralds turn blue in 2000 AD, she will still have been justified in projecting green. The rule of distinct observation by no means guarantees correct projections, nor is its adoption or continued use contingent upon its always producing correct projections.

In the third case, suppose both speakers are aware of the relationship between grue and green, but the grue-speaker believes with absolute certainty that he can observe grue without knowing the date, and the green speaker believes with absolute certainty that she can observe green without knowing the date. The grue-speaker can only assume that the green-speaker is mistaken. Likewise, the green-speaker can only assume that the grue-speaker is mistaken. The two speakers reach an impasse. Personally, I am absolutely certain that I can tell an emerald is green without knowing the date, and I do not hesitate to say that all grue-speakers are mistaken in believing otherwise. It may be that they have an instinctive knowledge of the date, of which they are unaware, or it may be that they are lying, but it is impossible for them to be correct.

4.

The rule of distinct observation allows us to project only well-behaved predicates. To project a relationship between observations, we must represent the relationship with a predicate, add observations of this predicate to our data, and demonstrate that the new predicate is well-behaved with respect to the data. If we can represent a relationship with a well-behaved predicate, then we shall say the relationship itself is *well-behaved*. Let us investigate the circumstances under which the equality of two observations is a well-behaved relationship.

Consider an experiment in which we examine the color of a collimated beam of light both at its source and where it hits a paper target. We obtain the following observations, where S(n) is the color at the source, and T(n) is the color at the target, at trial n.

n	S(n)	T(n)
1	green	green
2	red	red
3	blue	blue

Table 5: Observations of a Colored Beam of Light

We are inclined to infer that the beam will always have the same color at both ends. But we cannot project S=T from the table, because the table does not contain observations of S=T. Let us add observations of S=T, and see what happens.

n	S(n)	T(n)	S=T(n)
1	green	green t	rue
2	red	red	true
3	blue	blue	true

Table 6: Extended Observations of a Colored Beam of Light

The predicate S=T(n) is analytically dependent upon S(n) and T(n). Nevertheless, Table 6 is distinct and representative of itself. The rule of distinct observation does not refer to analytical dependency, but to empirical dependency. It is possible to observe the equality of S and T without using or duplicating observations of either S or T, as the following procedure demonstrates.

Suppose I have a color-blind friend and a box of unmarked color filters. Each filter allows only one color of the rainbow to pass through. My color-blind friend, whom I invite into the laboratory to assist me, does not know which filters correspond to which colors, nor is he going to bother to find out. We turn on the beam, I set the source color to green, observe that the target shows green as well, and instruct my friend to start trying filters. He is color blind, so he does not know what color the beam is. He tries one filter, but he cannot see the source through it. He tries another, and another, until he finds one through which he can see the source. He then looks at the target through the same filter, and sees its light as well.

'They are the same color,' he says.

'What color is the light?'

'I have no idea,' he says.

'What about the filter, does it say on that?'

‘Not a letter on the thing.’

I am convinced, by this point, that no duplication of my observations of S or T has been made, and therefore that the equality of S and T can be observed without duplicating observations of S and T.

Let us discuss the implications of the above procedure in more detail. We begin with an analogous situation in thermodynamics. The Zeroth Law states, ‘Given two objects in thermal equilibrium with a third object, the two objects will be in thermal equilibrium with one another as well.’ That is, whenever we form two sides of a triangle of thermal equilibrium, the third side is guaranteed. In order to confirm this law, we can observe the thermal equilibrium between each pair of objects independently. The Zeroth Law makes it meaningful to define a scale of temperature, since all objects in thermal equilibrium with a certain reference object will be in thermal equilibrium with one another, so we can say they are all ‘at the same temperature.’ Now consider the Zeroth Law as stated in terms of temperature, ‘Given two objects whose temperatures are equal to the temperature of a third object, the two objects will have temperatures equal to one another as well.’ In this second form, it far less obvious that we can compare the three temperatures independently. The very mention of a scale of temperature tacitly assumes the truth of the Zeroth Law, and in doing so obscures the possibility that the three temperatures might be independently comparable. But the fact is that the three temperatures are independently comparable. We simply connect them with a conducting medium and see if any heat passes between them. If no heat is exchanged, they are at the same temperature.

We can propose a Zeroth Law of Colors as well. It states that if two colors pass through the same filter as a third color, then the two colors pass through the same filter as one another as well. The Zeroth Law of Color makes it meaningful to define a list of reference colors by which to name colors, since all colors that pass through the same filter as a reference color will pass through the same filter as one another. Thus we find ourselves saying that two colors are the same because they both pass through the same filter as a certain reference color. But we can always compare two colors directly by making sure that they pass through the same filter as one another, and this comparison will be independent of our determining the names of the two colors.

Given that Table 6 is distinct and representative of itself, we can project S=T, concluding that the beam of light will always be the same color at both ends. Whenever two observations can form a distinct set with observations of their equality, we say that the two observations are *of the same type*. Observations of colors are of the same type, as are observations of temperatures, durations, distances, forces, masses, and intensities. By definition, a predicate representing the equality of two observations of the same type is a well-behaved relationship.

5.

We now consider whether the derivatives of numerical predicates are well-behaved relationships. Take as an example the study of the time, t, it takes for a golf ball to fall from a height, h. Suppose we have already established that the time taken for the golf ball to fall is a function only of the height from which it is dropped. We might do this be dropping the ball several times from the same height and projecting the equality of these times, as we projected the equality of the colors at both ends of a beam of light. Continuing our observations, we obtain the following results, where h is in meters, and t is in seconds.

h (m)	**t (s)**	$\mathbf{t^2=h/5}$	$\mathbf{t=h^4-5.962h^3+10.750h^2-5.338h}$
0.00	0.00	true	true
1.00	0.45	true	true
2.00	0.63	true	true
3.00	0.77	true	true

Table 7: Some Observations of Gravity

Looking at Table 7, we might conclude that $t^2=h/5$. But $t^2=h/5$ is ill-behaved with respect to the table: there is no way to observe that t^2 is equal to h/5 without using or duplicating at least one observation of h or t. Similarly, $t=h^4-5.962h^3+10.750h^2-5.338h$ is ill-behaved. Therefore the rule of distinct observation forbids the projection of either relationship.

The following table, however, which contains observations of the derivative of t with respect to h, is distinct and representative of itself.

h (m)	**t (s)**	**dt/dh** ($\mathbf{sm^{-1}}$)
0.00	0.00	0.45
1.00	0.45	0.18
2.00	0.63	0.14
3.00	0.77	0.14

Table 8: Expanded Observations of Gravity

To see why dt/dh is well-behaved with respect to Table 8, consider the following hypothetical experiment conducted by two people, each of whom has an accurate timer, and using only one golf ball. The first experimenter climbs to a window and drops a golf ball from the level of the sill. The moment she lets go of the ball, she starts her timer. The second experimenter watches the golf ball descend, and starts his timer the moment the golf ball strikes the ground. He then retrieves the ball and throws it up to his partner. She catches the ball, waits however long she likes, and drops the ball again. This time, however, she drops it from a height one meter above the sill. The moment she drops the ball, she stops her timer. The man on the ground again watches the golf ball descend, and stops his timer the moment it strikes the ground. The reading on the man's timer, minus the reading on the woman's timer, divided by one meter, is an observation of dt/dh (albeit not an exact one). This observation is coincident with the height, h, of the window sill, and the time, t, it takes the ball to fall to the ground, but it was obtained without using or duplicating observations of either h or t.

Determining dt/dh without observing h or t is even easier if we have two golf balls. The woman at the window simply drops both golf balls at the same time, the first from the height of the sill, and the second from one meter above the sill. The man on the ground starts his timer when the first ball hits the ground, and stops it when the second ball hits the ground. The reading on his timer, divided by one meter, is an observation of dt/dh.

These hypothetical experiments demonstrate that dt/dh can be obtained without observations of h or t, but they by no means constrain us to estimate dt/dh by means of such experiments. We can use any data to estimate dt/dh, including the data contained in Table 8 itself. This is exactly what I have done to provide values for dt/dh, and explains why the entries for h=2 m and h=3 m are the same.

Despite the fact that Table 8 is distinct and representative of itself, we cannot project dt/dh because it is not constant. If, however, we were to represent the time a golf ball takes to fall with t^2 instead of t, then the derivative of the new representation would be constant. But the derivative of t^2 with respect to h *cannot* be observed without duplicating an observation of t. Our hypothetical experiments were successful in measuring dt/dh without duplicating observations of t or h only because t is a *linear* predicate. That is, when we represent the passage of time with t, we can divide any interval of time into an arbitrary collection of sub-intervals, and the sum of the lengths these sub-intervals will always be equal to the length of the entire interval. The predicate t^2 is not linear. An interval of 4 s, when squared, is 16 s^2, while two sub-intervals of 2 s, when squared, are each 4 s^2, and add up to only 8 s^2. The only other linear representations of time, besides t itself, are obtained by multiplying t by a constant, k. But using kt instead of t will not make the derivative constant in Table 8.

Nevertheless, we can define our best estimate of dt/dh at a height, h, as begin equal to $(h_2-h_1)/(t_2-t_1)$, where h_1 and h_2 are two heights for which we have observations t_1 and t_2 of t respectively, and are such that there is no other height between h_1 and h_2 at which t has been observed. We now have the following observations.

h (m)	**t (s)**	**dt/dh (sm^{-1})**
0.00	0.00	0.45
0<h<1	-	0.45
1.00	0.45	0.18
1<h<2	-	0.18
2.00	0.63	0.14
2<h<3	-	0.14
3.00	0.77	-

Table 9: Completed Observations of Gravity

Table 8 contains observations of dt/dh for *all* values of h up to three meters, as well as isolated observations of t at four different heights. Since we have already established that t is a function only of h, we infer that dt/dh is also a function only of h, and thus our observations of dt/dh give us enough information to project a functional relationship between t and h for values of h up to three meters. This relationship is, of course, the linear interpolation between observed values of t.

If we acquired another observation of t for some height between zero and three meters, and this observation contradicted the observations of dt/dh contained in Table 9, we would have to include the new observation in Table 9 and adjust our observations of dt/dh accordingly. Following this procedure for every new observation of t, our linear interpolation might eventually imply that t^2 is nearly equal to h/5, allowing us to use $t^2=h/5$ as a working relationship between t and h. Even with the linear interpolation obtained from the observations of Table 7, we can say that $t^2=h/5$ calculates t to within a tenth of a second of the linear interpolation.

Suppose we try to thwart our inductive system by representing t with a non-invertible function, f(t), as in Table 10.

h (m)	t (s)	f(t)
0.00	0.00	1.00
1.00	0.45	1.00
2.00	0.63	1.00
3.00	0.77	1.00

Table 10: Distorted Observations of Gravity

Although f(t) is constant in the table, let us assume that it varies between our observed values of t.

Because f(t) is not linear, we cannot project its derivative. But we might be tempted to project f(t) itself. The function is, however, ill-behaved with respect to Table 10. We cannot determine f(t) without observing t, so a distinct subset of Table 10 must exclude observations of t if it is to include observations of f(t). But such a subset would not be representative of Table 10, because t cannot be deduced from f(t). We might leave observations of t out of the data, choosing to represent time only with f(t). But our observations of t would still be relevant to the remaining data, because their inclusion prevents the projection of f(t). Consequently, our commitment to include relevant observations forbids us to exclude observations of t.

In summary, we have found that the only relationship between t and h we can project without violating the rule of distinct observation (combined a commitment to include relevant observations) is the linear interpolation of time with respect to height. Should we attempt to use a non-linear representation of either time or height, we find that we cannot project the derivative of time with respect to height. Should we attempt to use non-invertible representations of either time or height, we find that we cannot project either of these representations or their derivatives. Therefore we are justified in preferring the traditional, linear representations of time and height in our experiment, because they are the only ones that allow us to project a functional relationship between the time it takes a golf ball to fall and the height from which it was dropped.

6.

No Ill-behaved predicate can be observed without first observing a well-behaved predicate. The converse, however, is not true. Thus there exists an empirical difference between ill-behaved and well-behaved predicates. The rule of distinct observation exploits this difference so as to forbid the projection of ill-behaved predicates, while permitting the projection of well-behaved predicates. When induction is subject to this rule, grue cannot be projected, but green can. Thus we can justify our belief that all emeralds are green. Similarly, we can justify our belief that the sun will come up tomorrow, because the predicate ‘rises every day until tomorrow’ is not well behaved: it cannot be observed without knowing whether or not tomorrow has arrived.

If a relationship between observations can be represented by a well-behaved predicate, then the rule of distinct observation allows us to project it, and we say that it is a well-behaved relationship. The derivatives of linear predicates are well-behaved, as are relationships of equality between observations of the same type. By projecting the derivatives of predicates, we can arrive at linear interpolation between data points. By projecting the equality of observations, we can conclude that these observations will always be the same.

By no means have I proved that science can proceed along *all* of its accustomed paths without violating the rule of distinct observation. Nor have I offered an a priori reason for adopting this rule in preference to any others that might solve the New Riddle of Induction with different results. Nevertheless, I have demonstrated that the rule assures self-consistent and scientifically agreeable inferences over a wide range of inductive exercises.

Footnotes

[1] Goodman, N. (1965), *Fact, Fiction and Forecast*, Second Edition. New York: Bobbs-Merrill.