Peter Abell

Narratives, Bayesian Narratives and Narrative Actions

(doi: 10.2383/25959)

Sociologica (ISSN 1971-8853)
Fascicolo 3, novembre-dicembre 2007

Ente di afferenza:
()

Narratives, Bayesian Narratives and Narrative Actions

by Peter Abell

doi: 10.2383/25959

Introduction

Human beings frequently claim to understand events when they manage to formulate a coherent story or narrative explaining how they believe an event was caused or, more often, how the world is causally transformed from one state to another by virtue of human agency/action. Social scientists, particularly those of a historical bent, also sometimes invoke action driven narratives as an explanatory device. They tend to do so in the absence of frequently recurring events which precludes systematic comparison and generalization and, thus, the use of statistical reasoning in order to validate any causal claims. Furthermore, a narrative understanding of the world often underpins the way in which people make decisions. They characteristically compare how their current situation is causally generated with the causal structure of past similar situations as a guide to "what to do next." Narratives consequently appear to lie at the very foundations of many of our cognitive procedures. Furthermore, whether used as an exploratory device or as a guide to action, in the absence of statistical reasoning, a Bayesian form of causal inference of necessity underpins narrative analysis.

The theory of comparative narratives [Abell 1987; 1993] was devised in order to describe and compare structures of action driven sequential events. Narratives, so conceived, provide, I claim, an appropriate analytical framework where events are of limited recurrence and case studies are the only source of information. Narratives depend upon a singular conception of causality [Abell 2001; 2004] which is to a degree controversial, but if sanctioned it does give narratives a distinctive role in both sociological explanation and decision making. The standard statistical approach to explanation requires that both generalization (even nomic generalization which is derived from a general counterfactual) and comparative method are prerequisites for any causal explanation. That is to say we cannot explain a phenomenon without either deductively or inductively subsuming its description under at least one generalization, the truth of which is determined by comparative method (i.e. observing cases of A that become B and not-A that fail to become B). Singular causality, on the other hand, inverts the explanatory components. Generalization and comparative method are, logically speaking, posterior to explanation. One asks whether a singular causal explanation can be generalized by comparing, a particular case, with others. I argue that action driven transformations in the world can be causally explained in singular terms. Causal inference becomes not a matter of establishing a co-variation in a population of events but of being assured "beyond all reasonable doubt" that a particular causal link exists, given the array of available evidence. This conception of causality leads to what I term Bayesian narratives.

Abbott [1995] has been in the forefront in developing what is variously called sequence analysis or optimal matching, which bears a close affinity with narrative analysis. Heise [1997] has promoted event structure analysis and the software ETHNO which, once again, provides a away of examining sequential structures of causally linked events. Franzosi [2003] has also developed a highly original way of analyzing what he terms narrative data. He converts qualitative (natural language) narrative data into a numerical scale by counting "subject-action-object" triples. Bates and his co-workers [1998] have sought a connection between narrative analysis and game theory. All of these authors seek, in one way or another, to capture the sequential nature of social phenomena – be it at the micro or macro level – but they do not promote their methods as peculiar to events of such limited occurrence that statistical reasoning is impossible.

In Abell [2003] I sketched a theory of cognitive narratives leading to what I termed narrative action theory (NAT). It is the foundations of this conception which I will develop later in this paper. In so doing I shall incidentally contrast the theory with the other two principle theories of action – rational action theory (RAT) and normative compliance theory (NCT). For a similar and much more formally developed theory, though not one based on narrative ideas see Gilboa and Schmeidler [2001].

It is not infrequently remarked that RAT is forward looking (i.e. prospective), actors are deemed to act "as if" they calculate the probabilities of the anticipated future valued outcomes (consequences) of their actions, whereas NCT is backward looking, actors follow pre-established guides to conduct [Elster 1989]. On the other hand, NAT, as we shall see, has it both ways. Actors can behave "as if" they pay attention to both an uncertain future and to similar precedents when deciding how to act. The precedents can, of course, provide clues about possible consequences.

"As if" assumptions have not proved popular in sociology. Rather the intellectual drive has almost invariably been towards an ever deeper phenomenological understanding of human actions. Whilst I believe this to be fundamentally mistaken – especially where individual actions are used to account for macro social outcomes – it is, nevertheless, important to exert some caution in embracing "as if" assumptions. The extreme version ignores all explanatory criteria other than predictability, invoking no reality constraints whatsoever. I am, however, advocating NAT as a theory which descriptively approximates actual cognitive processes – probably better than RAT – in some, but only some, circumstances. Actors are deemed to make more or less complex similarity comparisons between current and past situations (conceived as narratives) and then to adopt preference "as if" they made subsequent computations as described in the Appendix. These computations run from the rather simple to the more complex. The main conjecture is one whereby actors start with computationally simple rules and it is only if these do not furnish a clear preference that they will then move on to more complex rules. This should be contrasted with expected utility theory (RAT) where actors are deemed to make complex computations about weighted probabilities from the start. NAT, on the other hand, places far less computational demands upon the individual.

The paper proceeds as follows. First, I give a a semi-formal introduction to the concept of Narrative. Second, I develop the idea of a Bayesian Narrative which involves the insertion of non-frequentist causal links into a Chronology of events or actions; a causally related chronology then comprise a narrative. Armed with these ideas I, fourthly, outline the basics of Narrative Action Theory. The paper, thus, provides an introduction to both narrative reasoning as an explanatory device and to a theory of action which can complement RCT and NCT.

Narratives

I have attempted a rather formal definition of Narratives elsewhere [Abell 1987].

Here I shall offer a semi-formal introduction which will, I trust, prove sufficient for the purposes of the rest of the argument.

The underlying intuition is rather simple: a narrative comprises a time ordered *chronology of situations (or states of the world) which are transformed by human agency (individual, collective, aggregate)*; they may be depicted as digraphs as in figure 1.

A narrative comprises:

(N1) A finite set of descriptive situations (states of the world) S.

Comment: the states of the world will characteristically contain descriptors both of the persistent conditions of action and conditions which are transformed by actions. In a formal sense we can envisage a world being propelled around various positions in a many dimensional state space. Since, however, the idea of narrative is probably most useful in an informal qualitative context where states are described in a natural language format, not too much reliance should be placed on such formal ideas. In what follows although I shall use letters to describe various features of narratives, it should be born in mind that they are place-holders often standing for what might be rather complex natural language descriptions. Though narrative analysis can take on a rather forbidding algebraic appearance it is essentially tied to a natural language depiction of the social world.

(N2) The elements contained in set S are weakly ordered in time providing a Chronology of events in the sense that earlier states are transformed to later states.

Comment: time might be continuous or discrete. States can be extended in time, so starts and finishes will then be ordered in time. A weak order allows for parallel paths in a narrative, as in figure 1.

(N3) A finite set of actors (individual or collective or an aggregate of actors), P.

Comment: narratives may provide a picture whereby individual actors become constituted as collective actors. Whether one starts with, on the one hand a collective or aggregate of actors or, on the other, with individuals depends upon the level of abstraction of the narrative (see below).

(N4) A finite set of actions A.

Comment: the same action may appear more than once but is distinct because of the time signature.

(N5) A mapping of P onto A giving pairs on PxA.

Comment: the mapping shows which actors perform which actions.

(N6) A mapping of elements of (PxA) onto (SxS).

Comment: The map shows how the states are transformed (or maintained) by the actions of named actors to later (the same) states.

(N7) The structure of a narrative can be depicted as a multi digraph N=(S; (PxA)).

Comment: An illustrative example of a simple digraph is given at figure 1. The states are depicted as points (vertices) and the actions carried out by named actors as the labeled directed lines (arcs) connecting pairs of points. A multi digraph allows for more than one arc between any pair of points indicating that more than one actor/action was responsible for the transition.

Structures of any given narrative posses certain properties:

1) They are a-cyclic; this seems reasonable since the chronology of states is ordered in time.

A state may, nevertheless, be revisited at a particular time but the time signature maintains an independent identity.

2) They are semi-connected; that is there exists at least one path (sequence of) actions (arcs), irrespective of direction, between all pairs of points (states) in the structure. This property guarantees the integral "wholeness" of the narrative structure.

3) They are "and" digraphs. That is, multiple arcs (i.e. actions) incident into a point (state) are read as being jointly sufficient for the state and multiple lines out of a state are also read as multiple actions prompted by the same state.

t_1-------------t_2----------------------t_3-----------t_4----------------------t_5

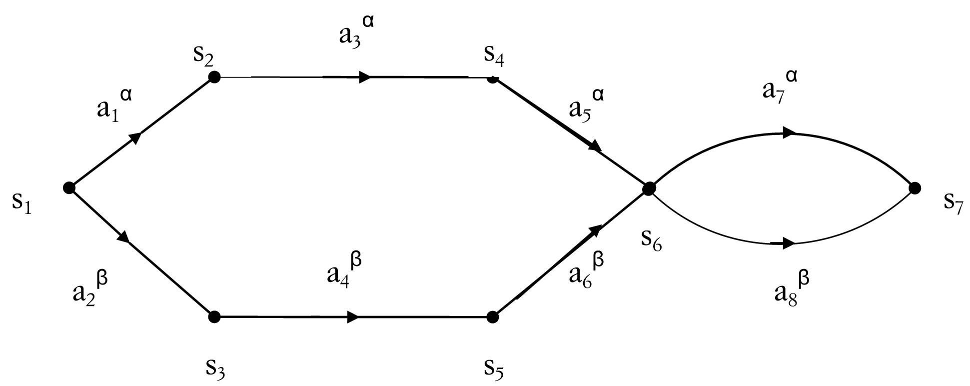

FIG. 1. A simple narrative structure.

So, in figure 1, S_1 at time t_1 prompts action a_1 by α and action a_2 by β. Action a_1 transforms state S_1 to state S_2 and action a_2 transforms S_1 to S_3 both at time t_2. Note, the dynamic system bifurcates at S_1. Actions a_5 by α and a_6 by β recombine the dynamic system at S_6 at time t_4.

Actions a_7 and a_8 are taken by α and β as a collective or aggregate actor. We may construe the narrative as furnishing a mechanism whereby an S_1 world at t_1 gets transformed to an S_7 world at t_5. This transformation or the state S_7 can be regarded as the explanandum. The paths of actions a_1-a_3-a_5-$a_7.a_8$ and a_2-a_4-a_6-$a_7.a_8$ then provide the explanans [Abell 2004]. It is often convenient to drop reference to the intervening states between S1 and S7 and to depict the narrative as an action skeleton as in figure 2. This depiction produces a narrative in terms of actions causing subsequent actions. I shall adopt the action skeleton depiction in the following section.

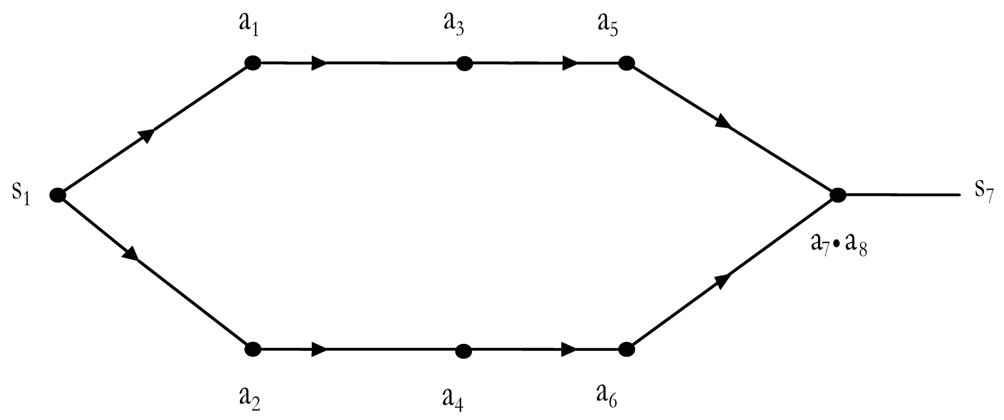

FIG. 2. An action projection.

Narratives can be mapped (translated) to each other by:

1) Identifying points (states or actions) in sub-digraphs (the semantic aspect of translation); whilst

2) preserving connectivity of action paths (the syntactical aspect of translation).

Translation can involve both similarity and simplification (abstraction). A simple illustrative example is provided at figure 3. There, two narrative structures are each abstracted (the upward dashed lines). We can then ask whether, at different degrees of abstraction, the narratives can be mapped (that is they are similar) to each other – the horizontal dashed lines. An obvious intuition is that the more abstract the narrative is the easier it will be to locate similarity. The abstraction mappings not only identify points but maintain connectivity. I am not here concerned with technical details but the diagram captures the idea whereby narratives can be abstracted and the compared by surrendering local detail [see Abell 1987]. I shall later make use of this aspect of narratives in the formulation of NAT

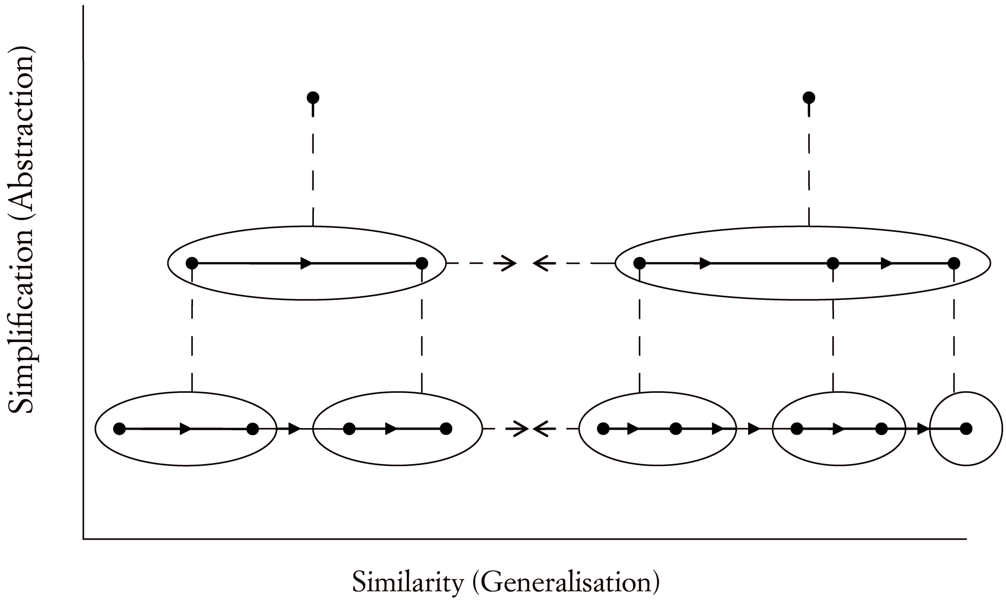

FIG. 3. Simplification (abstraction) and similarity (generalization).

Bayesian Causal inference and Narratives

It is standard practice in large-N studies to adopt a Humean, constant conjunction, interpretation of any causal link C (cause) → E (effect). When suitable controls are in place repeated cases of C leading to E, along with cases which are non-Cs leading to not-E (i.e. comparative method) need to be observed. Technical details apart, the underlying logic is transparent; both comparative method (i.e. comparing units of analysis or cases) and generalisation (preferably perhaps nomothetic) are prerequisites for the discovery of any causal explanation (and prediction).

Thus, in this framework, in the absence of a "known" causal generalisation (some might say law) there can be no causal explanation without inter-case comparison. Single cases can, as a consequence, only play a role in the discovery of a cause if causal relations can be conceived as singular [Cartwright 1989]. Much depends upon finding a satisfactory conception of singular causality and specifying the conditions under which it can be observed.

As we saw above the basic building blocks of narratives (expressed as action skeletons) are causal links of the form "α doing X_1" causes "β to do X_2." Single cases can be constructed in order to tell a causal story as long as they furnish *a connected action narrative path* as in figure 2.

The action path concerns the endogenous transformation of one state of the world to another. The essential points are that: *a*) the transformation is by human agency (individual or collective) and *b*) that prior actions *or* their consequences (intended *or* unintended) *or* both provide partial grounds for subsequent actions. Partial in the sense we allow for conjoint additional exogenous conditions. So, the basic building blocks of a narrative are paths which take the form:

in situation C action a_1 by α causes action a_2 by β,

where α and β may be individual or collective and may or may not be distinct. We shall assume the chronology of actions is known ("observed") and the question arises as to the nature and probative force of evidence in favour of a causal link running between them. The evidence is often of the form γ said "β did a_2 because α did a_1" which might derive from participants or close observers of the interaction. In the context of historical inquiry historians may well provide the evidence.

In general the question we wish to pose is, how do the odds that the causal link is present alter given various items of evidence which have been assembled? The items of evidence favouring and opposing the presence of a causal link may be either independent or dependent conditional upon the link (e.g. testimony of two or more independent observers as opposed to colluding observers). The objective of causal analysis is to find a way of combining the separate items of evidence in a consistent manner to demonstrate that the odds of the presence of the link is "beyond all reasonable doubt."

Let us call the hypothesis that a causal link does in this case pertain A and its denial ¬A. So, we now ask what is the evidence, in this particular case, for the truth of A? Assume, initially, one item of evidence only, *b* (e.g. a report by β that she did a_2 because α did a_1), in substantiating P(A|b), the probability that a causal link is present given that β reports that "she did a_2 because α did a_1." So, using Bayes law we have:

$$P(b)\cdot P(A\,|\,b) = P(A)\cdot P(b\,|\,A)$$

and

$$P(b)\cdot P(\neg A\,|\,b) = P(\neg A)\cdot P(b\,|\neg A) \tag{1}$$

Where P(A) is the probability of A, P(b) is the probability of b, P(b|A) is the probability of b given A and P(b|¬A) is the probability of b given ¬A.

Thus,

$$\frac{P(A\,|\,b)}{P(\neg A\,|\,b)} = \frac{P(A)\cdot P(b\,|\,A)}{P(\neg A)\cdot P(b\,|\neg A)} \tag{2}$$

So,

(3) $$Odds((A:\neg A)\,|\,b) = Odds(A:\neg A)\cdot L_b$$

where,

$$L_b = P(b\,|\,A)/P(b\,|\neg A)$$

is the likelihood ratio of the evidence *b* on A and ¬A.

So,

(4) $$LogL_b = LogOdds((A:\neg A)\,|\,b) - LogOdds(A:\neg A)$$

$LogL_b$ [Good 1983] gives a measure of how the odds of A as against ¬A alters as a consequence of the evidence b. It provides a measure of the evidential support that b gives for the existence of the causal link A by changing the prior odds to the posterior odds. Expressing equation (3) in logarithmic terms (equation (4)) is clearly optional it merely allows an interpretation of the change in odds in terms of a difference and locates no evidential impact at zero.

$LogL_b=0$ (i.e. $L_b=1$) if b has no impact on the odds; $(P(b|A)=P(b|\neg A)$,

$LogL_b>0$ (i.e. $L_b>1$) if b has a positive impact on the odds; $(P(b|A)>P(b|\neg A)$,

$LogL_b<0$ (i.e. $L_b<1$) if b has a negative impact on the odds; $(P(b|A)<P(b|A)$.

In practice there may be two or more items of evidence which bear upon the likelihood of the truth or falsity of hypothesis A. Start by assuming there are two items, b_1 and b_2 mutually independent conditional upon A and ¬A. (e.g. two entirely independent reports).

Then we have,

$$P(A\,|\,b_1,b_2)\cdot \mathrm{P}(b_1)\cdot \mathrm{P}(b_2) = P(A)\cdot P(b_1\,|\,A)\cdot P(b_2\,|\,A)$$

and

(5) $$P(\neg A\,|\,b_1,b_2)\cdot P(b_1)\cdot P(b_2) = P(\neg A)\cdot P(b_1\,|\neg A)\cdot P(b_2\,|\neg A)$$

Dividing the first equation by the second:

$$\frac{P(A\,|\,b_1,b_2)\cdot P(b_1)\cdot P(b_2)}{P(\neg A\,|\,b_1,b_2)\cdot P(b_1)\cdot P(b_2)} = \frac{P(A)\cdot P(b_1\,|\,A)\cdot P(b_2\,|\,A)}{P(\neg A)\cdot P(b_1\,|\neg A)\cdot P(b_2\,|\neg A)}$$

Cancelling P(b1) and P(b2) gives:

(6) $$\frac{Odds\big((A:\neg A)\,|\,b_1,b_2\big)}{Odds(A:\neg A)} = \frac{P(b_1\,|\,A)\cdot P(b_2\,|\,A)}{P(b_1\,|\neg A)\cdot P(b_2\,|\neg A)} = L_{b1}\cdot L_{b2}$$

So,

(7) $$LogOdds((A:\neg A)\,|\,b_1,b_2) - LogOdds(A:\neg A) = LogL_{b1} + LogL_{b2}$$

Let,

(8) $$L_B = L_{b1} \cdot L_{b2}$$

Where, L_B is the likelihood ratio of b_1 and b_2 conditional upon A and ¬A.

So, in general, for n conditionally independent items of evidence we will then have:

(9) $$L_B = L_{b1} \cdot L_{b2} \cdot \ldots\ldots \cdot L_{bn}$$

and

(10) $$LogL_B = \sum_n log\,b_{n}.$$

Equation (10) details how all the items of conditionally independent evidence combine in support of or against the hypothesis A (alternatively, support of or against hypothesis ¬A). As we shall see below L_B gives an estimate of how the combined (conditionally independent on A and ¬A) items of evidence alter the log-odds of A as against ¬A; that is to say, the log-odds for the existence of a causal link between the actions/events specified in the chronology.

Now assume that b_1 and b_2 are not independent conditional upon A and ¬A (e.g. two partially collaborative reports)

So,

$$P(A\,|\,b_1,b_2)\cdot P(b_1,b_2) = P(A)\cdot P(b_1\,|\,A)\cdot P(b_2\,|\,A,b_1)$$

(11) $$P(\neg A\,|\,b_1,b_2)\cdot P(b_1,b_2) = P(\neg A)\cdot P(b_1\,|\,\neg A)\cdot P(b_2\,|\,\neg A,b_1)$$

Dividing and taking logs,

(12) $$LogOdds\big((A:\neg A)\,|\,b_1,b_2\big) - LogOdds(A:\neg A) = LogL_{b1} + LogL_{b2/b1}$$

where $L_{b2/b1}$ is the likelihood ratio of b_2 given b_1 conditional on A and ¬A. Thus, with multiple conditionally dependent items of evidence we have equations similar to (6) and (9) but reflecting the pattern of conditional dependence amongst the items of evidence.

Comparing equations (12) with (7) demonstrates that the conditional dependence of items of evidence does not materially alter the picture.

Inferential Procedures

Given a conjectured causal link in a chronology (A above) which is linked to item(s) of evidence (b_1, b_2,… above), we wish to attach a value to the posterior odds of A or ¬A given the evidence. The value, in the absence of repeated observation (frequencies), will of necessity be derived from the "degrees of belief" of the analyst [Schum 1994]. The analyst, in turn, may rely upon the estimates of informed respondents, for example, knowledgeable historians. There is of course a lengthy tradition in sociology of placing careful reliance upon the testimony of experts or "key informants." March *et al.* [1991] in a provocatively entitled paper *Learning from Samples of One or Fewer*, write as follows:

> Theories of historical inference tend to emphasise pooling of observations. Pooling over observers appears to have advantages in some common situations but in the absence of a clearer formulation of the gains and losses involved, it is hard to specify the precise conditions favouring one strategy or the other.

Bayesian narratives, I conjecture, might offer some help in this respect.

Since an estimate of the posterior odds, conditional upon the evidence, is our ultimate objective the analyst could of course ask any key respondent to make a stab at its value. We could perhaps, following Simon, even average the values given by a number of such respondents. Giving such estimates, however, would prove both difficult for the respondents to achieve with any reliability and would not allow for comparative estimates of the impact of individual items of evidence upon the posterior odds. Rather, respondents should be offered a framework in which they can assemble their estimates of the odds from the evidence whilst checking upon the consistency of their reasoning.

The general inferential structure from the forgoing analysis takes the form:

$$\text{(13)} \qquad \frac{Odds((A:\neg A)\,|\,b_1,b_2,\ldots b_n)}{Odds(A:\neg A)} = L_B = L_{b1}\cdot L_{b2/b1}\cdots L_{bk/b1\ldots bk-1}\cdots L_{bn/b1\ldots bn-1}$$

In order to estimate the posterior

$$Odds((A:\neg A)\,|\,b_1,b_2,\ldots b_n)$$

from L_B we still require an estimate of the prior odds. I shall return to that issue presently.

Expert informants may make estimates of:

1) The constituent likelihoods in the appropriate equation for L_B.

2) The global likelihood L_B.

The analyst can than search for consistency between the these estimates before accepting the estimate of the global L_B. Alternatively, the likelihoods could be estimated as though the items of evidence are severally conditionally independent on A and ¬A and any diversity between their product and the global estimate of L_B could then be ascribed to conditional dependence amongst the items of evidence. In general, the analyst will discount those estimates by key informants whom are neither consistent nor embrace a wide range of evidential items.

Since the objective is to estimate $Odds((A:\neg A)|b_1,b_2\ldots,b_n)$ we still need an estimate if the prior odds of A to ¬A. There are probably two reasonable approaches to this estimation:

- assume the odds are 1, (i.e. in the absence of any evidence A and ¬A are equally probable);
- ask the key informant for an estimate [see Schum 1994 for a legal analysis].

Then in either case the posterior odds can be computed from the prior odds and the appropriate likelihood ratios.

Finally, do the odds so computed permit us to infer "beyond all reasonable doubt" that either A or ¬A is true? As with any conception about causality based upon frequencies one can impose more or less demanding criteria in respect of significance.

The values of P(A) and P(¬A) are given by :

$$P(A) = X/(1+X) \tag{14}$$

$$P(\neg A) = 1/(1+X) \tag{15}$$

where X:1 is the value of the odds that A is true.

By setting the prior odds to unity we assume, in the absence of any evidence, that P(A) and P(¬A) are identical in value and therefore both equal to 0.5. In the presence of evidence we might reasonably require that when the posterior odds are 100:1 then A is true "beyond all reasonable doubt" (a causal connection exists between a_1 and a_3). Furthermore, when the odds are 1:100 then ¬A is true "beyond all reasonable doubt" (there is no causal connection between a_1 and a_3). Thus, the range of change in odds is four log units.

The reader might feel rather disconcerted about the subjective nature of these various estimates but my claim is that, by being explicit about the inferences and incorporating the internal consistency checks, a systematic inferential procedure is at hand which permits causal inference when inter-case comparison is not feasible.

A single non recurrent case depicted as a chronology of events or actions is now open to an analysis of each of the conjectured causal links. Such a case will offer a putative, always revisable in the face of additional evidence, picture of how human agency has transformed states of the world.

Although there is an extensive literature devoted to the virtues of detailed case studies the treatment there of causal inference is either missing or underspecified. Bayesian narratives provide one method for bringing case studies within the ambit of acceptable causal analysis and thus scientific procedure. The awkward disjuncture between historical and social science analyses is bridged.

Occasionally situations arise where a small number of cases do invite comparison but are not found in sufficient numbers to enable the detection of statistical co-variation. Here Bayesian analysis enjoins a singular causal analysis of each case prior to the formulation of the question as to whether the narratives lend themselves to generalisation.

If single cases can provide at least a provisional estimate of causal efficacy then they may also be used by decision makers as guides to action where, once again, statistical reasoning is not feasible. So we now turn to Narrative Action Theory (NAT).

Narrative Action Theory (NAT)

It does appear that decision makers often tend both to locate what they are currently doing within a narrative framework and to compare their present situation with previously encountered (memorized) narratives. Since these narratives are often not recurrent to a sufficient degree to permit statistical treatment, decision makers inevitably furnish the chronologies they compare with Bayesian causal inferences. NAT is an initial attempt to capture this aspect of decision making. If it is to play any role as a theory of decision making (action) then one needs to find its empirical implications which separate it from RCT and NCT. Unfortunately, this falls beyond the compass of this paper. Here I merely outline the central ideas.

(NA1) An actor faces a (current) situation S(c) conceived as the latest state in the current evolving narrative N(c). Label this as S(c)/N(c) The actor needs to choose the next action (what should she now do?) at this juncture in the current narrative.

Comment: N(c) may embody (in the mind / cognition) of the actor more or less detail (abstraction, figure 3), more or less causal history and more or less anticipated future beyond the current state S(c), consequential upon what she now does. Selections of actions at S(c) are, thus, characteristically embedded in an unfolding narrative. First and last moves are special cases.

(NA2) The actor has a finite reference set of narratives Π each element of which comprises a chronology of states and state transforming actions. (e.g. figure 1).

Comment: The reference set Π may be regarded as the total recollected cognitive resource available to the actor. It will of course alter over time as new narratives are incorporated and some are lost. The narratives will characteristically be fragments at differing levels of abstraction (figure 3).

(NA3) The actor imposes a similarity relation L (like) on (S(c)/N(C)) and elements of Π. Call the set which is similar to S(c)/N(c) the similarity set.

Comment: the idea here is that the actor has some way of comparing (mapping) the current narrative state onto some (at least one) of the narratives in the reference set. It is important to emphasise that actors do not just compare states for similarity, but the structure of narratives which generated them. People seem more certain in their ascription of similarity the more they know about the history of a situation and the detail of the generative processes. This probably makes particular sense where the narrative describes human interactions leading to the current situation.

(NA4) The actor chooses at S(c)/N(c), according to some decision rule, an action from the set A of actions taken in those narratives in Π which are similar to S(c)/N(c).

I will distinguish between pure frequency based rules and value based rules; each of which may be used in differing circumstances.

Pre Frequency Based Rules

(F1) Choose the action with the maximum frequency of use in the similarity set.

Comment: this rule, in effect, merely use a Boolean L (see Appendix) relation as a threshold for inclusion in the reference set. The actor at S(c)/N(c) surveys all the similar narratively generated situations in this set and merely chooses the action that has been most frequently used at a similar juncture in the past. More complex rules (below) may be resorted to in the case of a tie. This simple frequency based rule has the virtue of simplicity but, note, it may select an action which has come to be used most frequently in the past on the basis of a more complex frequency or value based criteria. Indeed, frequency based myopic switching rules are often found in evolutionary game theoretic dynamics which converge to a Nash equilibrium and where the frequency of a strategy (i.e. action) in the current population is a proxy for past success (value/utility). Norms may, of course, often be transmitted in a population by a process of frequency emulation.

(F2) Choose the action with the maximum weighted (by L) frequency of use in the similarity set.

Comment: This rule is slightly more complex than (F1) requiring some measure of L in order to weight the action (see Appendix). So actions used in the reference set of narratives which are of greater similarity to S(c)/N(c) will contribute more to the choice than those used in less similar narratives. It is possible for an action of seldom use, but with high similarity to the current narrative to outperform a frequently used action with lower similarity.

(F3) Choose the action with the maximum weighted (by L) mean frequency of use in the similarity set.

Comment: This rule is more complex than (F2) requiring, in addition to a measure on L, an averaging over the weightings.

Value Based Rules

Now assume that the actor places value/utility on the ultimate states of each reference narrative (e.g. S_7 in figure 1) or indeed on all states subsequent to the state similar to S(c)/N(c).

(V1) Choose the action which leads to the maximum value/utility of outcomes in the similarity set.

Comment: Here the actor treats the reference set as Boolean, as with (F1), and merely selects the action which was most successful in the past (the single best action). Again ties will call for some resolution. The rule is completely devoid of any frequency considerations – a frequently used action with slightly inferior outcome value will not be chosen. The rule although looking backwards does take account of future states consequent upon an action at a particular juncture in similar narratives.

(V2) Choose the action which leads to the maximum weighted (by L) value/ utility in the similarity set.

Comment: Here the actor can be modeled as if she computes for every action the weighted sum of the outcomes each time the action was chosen in the reference set and then chooses the best sum.

(V3) Choose the action which lead to the maximum weighted (by L) mean value in the similarity set.

Comment: Here the actor behaves as under (V2) but computes an average value for the outcomes of each action. Note that both (V2) and (V3) do have frequency implications in the sense that more frequently chosen actions in the reference set are likely to contribute more to the value of the action.

NAT clearly allows for a number of different decision rules, but all have in common cognitive processes whereby actors refer to past experience.

Narrative, Rational, and Normative Actions

When faced with a decision about how to act the first conceptual resort, according to NAT is remote from both the conceptual components and computations of RCT (i.e. expected utility maximisation). It would be difficult to conceive of either rule F1 or V1 as rational or, indeed, even as satisfying. Rule F1 requires individuals, firstly, to compare the narratives in which they find themselves involved to a reference set of (Boolean) similar narratives, then to look at the actions taken at a particular juncture in these narratives and, finally, to select the one most frequently used. As I noted this rule is innocent of consequences let alone probabilities. It can, nevertheless, be deemed rational if the decision maker has reason to believe that those narratives in his reference set were the outcome of rational deliberations by either himself or others on previous occasions, as long as the similarity to the current narrative is close enough to warrant such a conclusion.

Whether F1 should be regarded as a norm is a moot point. Although it is devoid of prospective thought [Elster 1989] it does not provide a recipe of unreflective response. Pure density dependent myopic dynamics are, however, sometimes found in evolutionary theories of norms so some level of critical scrutiny is, on this count, consistent with normative constraint. Indeed emulating the majority in a reference group could be conceived as a meta-norm in the sense that adopting the rule does not enjoin a particular course of action but a decision procedure.

Rule V1 is again backward looking but consequential in orientation as it picks out the action, if it exists, which procured the "best" outcome in the reference set. From a rational point of view it suffers, however, from the obvious problem that the action delivering the best outcome in the past may have been atypical.

Rules F2 and V2 both require the decision maker to weight and aggregate – certainly something they don't literally do. I think we must regard these rules very much in the spirit of an "as if" theory. The similarity weighting outlined in the Appendix is as close as I can get to actual cognitive procedures. The basic idea might be that people move from F1 and V1 if these rules don't provide a clear guide to action. Rules F3 and V3 are even more complex requiring in addition to similarity weighting and aggregation an averaging computation.

The conjecture of this exploratory exercise is that in relatively complex decision environments actors will sometimes explore a range of backward looking rules of increasingly complexity in choosing how to act. One may further conjecture that it is only if these do not provide a clear guide to action that RCT becomes the cognitive resort. This description of decision making is also not consistent with NCT as usually conceived. NCT selects a course of action (the norm). NAT selects a rule which, given varied historical experience, can suggest a range of alternative actions.

Narratives

I have outlined the basic ideas underpinning the concept of Bayesian Narratives. They offer the prospect of providing explanations of changes in the world when statistical reasoning is precluded because of limited recurrence of events. Furthermore, decisions are often taken where forward looking rational calculation is prohibitively difficult. It is sometimes averred that in such situations rational or satisfying norms tend to evolve. However, it proves difficult to reconcile the diversity in behavior with norms that enjoin particular courses of action. NAT provide an alternative model. Actors compare their current with past narratives and use the latter as guides to action. Their comparative narratives will reflect the idiosyncracies of their past experiences and only in special circumstances will enjoin uniform actions. Inevitably the causal links in such narratives are of a Bayesian nature. Bayesian narratives, consequently, should underpin our analytical approaches to explanation and to an understanding of decision making in a world where near unique events flourish. Historians find it relatively easy to agree upon chronologies of events but much more difficult to insert causal links amongst them. It is intellectually hopeless to advocate a Humean constant conjunction (statistical) interpretation of these links. All we can ask for is a method of connecting items of evidence to a conjectured causal link in a systematic and internally consistent manner. Bayesian Narratives provide one route to this objective.

Appendix: Narrative Similarity and Decision Rules

The basic idea behind NAT is one whereby actors pose the question as to how similar the current situation, S(c), in narrative, N(c), is to a finite set of previously encountered narratives where differing actions were taken. The actor then chooses the action which satisfies one of the decision rules outlined in the text.

In various ways these rules depend upon a measure of the similarity between N(c) and the comparators.

As we have seen, however, narratives, including N(c), can themselves be abstracted (simplified) (figure 3). Thus, narratives expressed at varying levels of abstraction can also be compared for similarity.

Assume a given narrative N can be increasingly but finitely abstracted producing a complete strictly ordered set of abstraction as follows:

$$N_i < N_j < N_k \ldots\ldots\ldots < N_n$$

The degree of similarity between any pair of abstractions can be conceived as departure from isomorphism. Two narratives are isomorphic if and only if they posses identical node sets and are identically connected.

A similarity relation, L, on the abstraction order will have the following properties:

Symmetric - if N_i L N_j then N_j L N_i for all i and j;

Reflexive - N_i L N_i for all i;

Insertion - If N_i L N_k then N_i L N_j and N_j L N_k where $i < j < k$.

L will not, however, generally be transitive. That is, if N_i L N_j and N_j L N_k then N_i will not always be similar N_k. Similarity is, thus, a tolerance rather than an equivalence relation.

Similarity across distinct narratives, N(1) and N(2), each of which may abstracted in an order will also define a tolerance.

These assumptions require that if two distinct narratives, $N(1)_i$ and $N(2)_j$, respectively at abstraction levels i and j, are similar then all those abstractions in N(1) similar to $N(1)_i$ and in N(2) similar to $N(2)_j$ are mutually similar.

The rules in the text which only require a Boolean measure follow the above assumptions at some threshold value of similarity. Decision makers may, however, conceive an ordinal measure of similarity. They then think in terms of "very similar," "similar" and "dissimilar" narratives derivative of the abstraction order. They also may adopt something along the lines of what I shall term a "weakest link calculus" with respect to the degree of transitivity of similarity between different levels of abstraction.

So let,
"very similar" = H(igh)
"similar" = W(eak)
"dissimilar" = O

Then,
if N_i H N_j and N_j H N_k then N_i H N_k,
if N_i W N_j and N_j W N_k then N_i W N_k,
If N_i H N_j and N_j W N_k then N_i W N_k,
if N_i W N_j and N_j H N_k then N_i W N_k

Since some of the decision rules in the text depend upon addition and averaging it is necessary to ascertain how in an essentially ordinal world actors achieve these operations. The objective here is to find the simplest possible "as if theory." The obvious metric scoring H =1, W = 0.5 may not be too far off the mark. These values can be inserted in the appropriate rule given below.

I should like to raise an issue not covered in the text. The more abstract the narratives are when they are compared then the easier it is to see them as similar. We may envisage all narratives as similar when enough detail is lost! Thus, similarity should count less the more abstract the level at which it is established when contributing to a decision rule. It might be useful to think in terms of metric measures of similarity (L) and abstraction (abs) each of which varies in the unit interval. Then similarity could enter the appropriate decision rule in the form:

$$L(1-abs)$$

The value base rules in the text also require a definition of:

O(k) be the outcome of narratives N(k) contained.
V{O(k)} is the utility/value to the actor of O(k).

References

Abbott, A.

1995 "Sequence Analysis: New Methods for Old Ideas." *Annual Review of Sociology* 21: 93-113.

Abell, P.

1987 *The Syntax of Social Life*. Oxford: Oxford University Press.

1993 "Some Aspects of Narrative Method." *Journal of Mathematical Sociology*. 18: 93-134.

2001 "Causality and Low Frequency Complex Events." *Sociological Methods and Research* 30: 57-80.

2003 "The Role of Rational Choice and Narrative Action Theories in Sociological Theory. The Legacy of Coleman's Foundations." *Revue Française de Sociologie* 44: 255-274.

2004 "Narrative Explanation: An Alternative to Variable Centered Explanation?" *Annual Review of Sociology* 30: 287-310.

Bates, R.H., Greif, E., Levi, M., Rosenthal, J.L., and Weingast B.R.

1998 *Analytical Narratives*. Princeton: Princeton University Press.

Cartwright, N.

1989 *Natures Capacities and Their Measurement.* Oxford: Oxford University Press.

Elster, J.

1989 *The Cement of Society*. Cambridge: Cambridge University Press.

Franzosi, R.

2003 *From Words to Numbers*. Cambridge: Cambridge University Press.

Gilboa, I., and Schmeidler, D.

2001 *The Theory of Case Based Decisions*. Cambridge: Cambridge University Press.

Goode, I.J.

1983 *Good Thinking: The Foundations of Probability and is Applications*. Minneapolis: University of Minnesota Press.

Heise, D.R., and Durig, A.

1997 "A frame for Organisational Actions and Macro Actions." *Journal of Mathematical Sociology* 22: 95-123.

March, J.G., Sproull, L.S., and Tamuz, M.

1991 "Learning from Samples of One and Fewer." *Organisation Science* 2: 1-13.

Schum, D.A.

1994 *The Evidential Foundations of Probabilistic Reasoning.* Evanston: North Western University Press.

Narratives, Bayesian Narratives and Narrative Actions

Abstract: A narrative comprises of singular causal links connecting a sequence of actions/events (i.e. a chronology). A technique which inserts these links using Bayesian inferences from items of evidence is outlined. The so constructed Bayesian narratives underpin both the analysis of historically based case studies and a possible theory of action which complements the standard theories of normative and rational actions.

Keywords: Bayesian narratives, causal inference, action theory, small-n, learning.

Peter Abell is professor in the managerial economics and strategy group, London School of Economics and professor of social research methods, Copenhagen Business School. Interests: mathematical sociology, social interacton and network models, rational choice, formalising qualitative methods.