

Reconstructing Generative Structure from Relational Trajectories

A Conformance Specification for the Study of Systems That Generate Their Own Units

Wanhong Huang

huangwanhong@serendip.ngo

Abstract

In a class of research objects the unit of analysis is produced by the process under study, and existing method treats the stability of that unit as a precondition to be checked rather than as an object to be modelled. Claims about change in such objects are therefore in part artefacts of the analyst's choice of unit, and no prevailing reporting norm requires anyone to say so. This paper offers one specification, not a method, for research on such objects. Three declarations are identified that no current standard requires: how the unit under comparison came to exist, how a variable at one scale is produced by and constrains another, and what an account of a process forbids. A formal presentation is then given in which a trajectory becomes analysable only under a declared partition of the state space; the grammar is stratified into epochs within which the production set is held fixed, with rewriting of the production set confined to epoch boundaries; non-terminals carry a scale index, so that the crossing of scales becomes a syntactic property open to audit; and the presentation is probabilistic, which is what makes model comparison well defined where only positive evidence exists. Three propositions follow. Membership is decidable within an epoch and undecidable if rewriting is permitted mid-derivation, which is why stratification is forced rather than stylistic. Acceptance under one partition entails nothing under a partition that does not refine it. A grammar containing no scale-crossing production derives no observation from a macro-level start symbol, so a claim to connect scales is checkable on the production set alone. The specification then states eight declarations that any implementation owes its reader, graded into requirements and recommendations, and it fixes which choices it leaves to the implementation. It is illustrated on a single trajectory. What the specification excludes, and the conditions under which each of its claims would fail, are stated at the end.

Keywords: methodology; unit of analysis; generative grammar; multi-scale analysis; grammar induction; falsifiability; reporting standards; relational systems.

A note on the standing of this paper. This paper proposes one specification among the several that would serve, and it neither claims that the approaches surveyed in §2 are inadequate nor that they were designed for the problem stated in §1. Its contribution is narrow by intention: a set of declarations that make a particular class of failure visible, together with a formal presentation precise enough that the declarations can be checked. The formal material in §5 is elementary, and it is meant to be; its propositions are stated so that they can be shown false. Section 8 gives the conditions of refutation. Objection, correction, and counter-evidence are welcome at huangwanhong@serendip.ngo.

Draft. This is a preprint draft. It is circulated for comment, and its arguments, formulations, and open questions are subject to revision. Correction and criticism are welcome at huangwanhong@serendip.ngo.

Licence. © 2026 Wanhong Huang. This work is licensed under a Creative Commons AttributionNonCommercial 4.0 International Licence (CC BY-NC 4.0).

You are free to share and adapt the material, under the following terms: you must give appropriate credit, provide a link to the licence, and indicate if changes were made; and you may not use the material for commercial purposes. The full text of the licence is at <https://creativecommons.org/licenses/by-nc/4.0/>.

Use of generative artificial intelligence. An AI assistant (Anthropic’s Claude) was used extensively in preparing this manuscript: for literature survey, for drafting prose from the author’s specifications, for adversarial review of the argument, for consistency auditing, and for typesetting. Formulations and objections arose in the course of that exchange as well as from the author. Every claim, argument, and citation was reviewed and decided by the author, who is responsible for the content and for any errors. Bibliographic details were checked against published sources for the journal and proceedings items; the monographs and a small number of remaining entries are checked for author, title, year, and publisher only.

Suggested citation: Wanhong Huang, *Reconstructing Generative Structure from Relational Trajectories: A Conformance Specification for the Study of Systems That Generate Their Own Units*, draft preprint, 2026.

1 THE UNIT IS GENERATED BY THE PROCESS UNDER STUDY

Comparison requires that the compared things be the same kind of thing, and a great deal of methodological care has been spent on securing that requirement. The care is warranted, and it has produced an apparatus: concept stretching and the ladder of abstraction (Sartori, 1970), and the systematic treatment of measurement validity that followed (Adcock & Collier, 2001). The apparatus asks whether a concept travels, whether the indicators chosen bear on the concept intended, and whether the cases assembled under one label are cases of one thing. It is a validity apparatus, and it does what it was built to do.

There is a class of objects for which it does something else. Consider a comparison of the arrangement between religion and public authority in one polity at the beginning of the twentieth century and in the same polity at the beginning of the twenty-first. The comparison presupposes that the term under which the two are gathered names the same relation at both dates. It does not. The relation at the later date was produced, in part, by the operation of the arrangement at the earlier one; the categories in which the later parties understand themselves were generated in the interval; and one of the generators was the legal and administrative practice that the comparison treats as its variable. The validity apparatus responds to this by stipulating equivalence, or by declining the comparison, or by descending the ladder of abstraction until a thinner concept travels. What it does not do, and was not built to do, is model the process by which the unit came to be the unit.

A second instance is vertical rather than horizontal. Work at the level of individual interaction treats collective identity as context, exogenous and given. Work at the level of institutions treats individual interaction as a sampling unit from which aggregates are formed. Multilevel modelling accommodates the nesting of the one in the other, and does so rigorously, but it presupposes that the levels are given and that they persist. The object of interest in the cases considered here is precisely the production of a level: the interval in which a pattern of interaction becomes something that constrains the interactions that produced it. Coleman's diagram of the macro-micro-macro transition names this problem exactly and has named it for four decades (Coleman, 1990). It remains, for the most part, a diagram.

The consequence is stated plainly. Where the unit of analysis is produced by the process under study, a claim about change in that object is in part an artefact of the analyst's choice of unit, and there is at present no reporting norm under which the analyst is required to say which part. This is not a charge of carelessness. It is an observation about visibility: the failure is of a kind that careful work and careless work display identically on the page, because the page contains no place where it would appear.

That last observation determines the form of what follows. A failure invisible under prevailing norms is not corrected by a better argument, since the better argument is also invisible; it is corrected by requiring a declaration. What this paper offers is therefore a specification and

not a method. It fixes what an implementation owes its reader and leaves to the implementation how the work is done, in the manner of an interface that constrains its implementations without containing one. The proposal is one specification among those that would serve, and the choice of the particular declarations below is defended in §3 rather than assumed.

The specification is framed around a formal presentation, developed in §5, in which a history is treated as a trajectory, a trajectory becomes a symbol sequence only under a declared partition, and generative structure is presented as a grammar over symbols that carry a scale index. The choice of a grammar is an inheritance from the generative tradition in linguistics (Chomsky, 1957, 1965), and the inheritance is partial in ways that §5 states before it states anything else. The formalism carries no claim that a society is a language.

One motivation may be declared and then set aside. The specification arose in the course of work on Generative Relational Being, a framework in which subjects, norms, and values are treated as generated in relational process rather than as antecedent to it. That framework is one adopter of the specification and is not its premise. Nothing below depends on it, and a reader who rejects it entirely may still hold that the declarations of §6 are owed.

The order of what follows is: what existing approaches settle (§2); the declarations none of them requires (§3); what the specification fixes and what it leaves open (§4); the formal presentation and its three propositions (§5); the declarations themselves (§6); a single trajectory carried through them (§7); and the conditions under which the argument fails (§8).

2 EXISTING METHOD CHECKS UNIT STABILITY AND DOES NOT MODEL IT

This section states what the adjacent literatures settle, so that the residue claimed in §3 is narrow. They are grouped by their treatment of the unit, which is the criterion at issue, rather than by discipline or by date. The grouping is a device of exposition and several bodies of work belong to more than one group.

2.1 *Units fixed, then compared*

Comparative historical analysis, set-theoretic and configurational methods, and the regression tradition share a sequence in which units are constituted before the analysis begins. Comparative historical work has developed a sophisticated treatment of sequence, timing, and path dependence within that sequence (Mahoney & Rueschemeyer, 2003; Pierson, 2004), and qualitative comparative analysis provides an explicit apparatus for conjunctural causation, in which conditions produce outcomes in combination rather than singly and several distinct combinations may produce the same outcome (Ragin, 1987, 2008). The last point matters below: equifinality is not a discovery of this paper but a working assumption of that method, and QCA supplies the most developed existing treatment of the combinatorial space that §5 approaches from another direction.

This group also holds the most developed treatment of temporal structure available in the social sciences. Increasing returns, self-reinforcing sequences, and the distinction between a critical juncture and the period of reproduction that follows it are its results, and they establish that the order of events carries explanatory weight independent of their content (Pierson, 2004; Mahoney & Rueschemeyer, 2003). Nothing in what follows displaces this. A specification that required the declaration of epoch boundaries is, in one reading, a request that the boundary between juncture and reproduction be stated in a form an auditor could check, and the concept it asks to be made checkable belongs to this literature.

The group has moreover faced the problem of the travelling concept directly and produced more than one answer to it. Beyond the ladder of abstraction there are the strategies of family resemblance and of radial categories, in which a concept is held together by overlapping rather than shared attributes and the secondary category is defined by reference to a primary one (Collier & Mahon, 1993). These are genuine responses to conceptual change, and they are responses at the level of the concept.

The treatment of the unit is nonetheless a check rather than a model. The analyst asks whether the cases are cases of one thing and answers yes, no, or not at this level of abstraction; where the answer is no, the remedy is to adjust the concept. The process that produced the divergence is not thereby represented, and in the objects considered here that process is frequently the explanandum.

2.2 *The unit as process, at one scale*

A second group takes process as its object. Process tracing reconstructs the sequence of events and mechanisms within a case and has developed explicit tests for the evidentiary weight of within-case observations (Beach & Pedersen, 2013). Sequence analysis treats whole trajectories, rather than variables measured on them, as the units to be compared and clustered (Abbott, 1995a), and it grew from a critique of variable-based analysis that anticipates much of §1 (Abbott, 1988). Genealogy in Foucault's sense takes the formation of a category as its object directly (Foucault, 1977). Microhistory reconstructs the generation of social positions at a resolution the aggregate methods cannot reach.

One member of this group requires something close to the declaration this paper asks for, and the concession should be made before the claim is stated rather than after. Process tracing in its formalised version assesses within-case evidence against rival explanations using tests graded by the evidence's necessity and sufficiency for a hypothesis, so that an item of evidence may be doubly decisive, may eliminate a rival without confirming the favoured account, or may do neither (Collier, 2011; Bennett & Checkel, 2015). To apply such a test the analyst must say in advance what would be found if the hypothesis were false. That is a falsification surface, it is required rather than commended within that tradition, and Claim 3.3 below is narrowed accordingly.

What the narrowing leaves standing is the difference between a hypothesis and a generative structure. The tests operate on the evidentiary standing of particular propositions within a case; they do not ask what the reconstructed process as a whole excludes, and a reconstruction can pass every test on every proposition while remaining compatible with any trajectory whatever at the level of the structure. Sequence analysis reaches the level of the whole trajectory and pays for it in a different currency, since its comparisons are conducted at one scale and its distance measures are defined over a fixed alphabet of states whose formation is not itself the object.

This group models unit formation, and in the genealogical case models little else. What it does not carry is an apparatus for the passage between scales, nor, at the level of the reconstructed structure rather than the individual hypothesis, a statement of what has been excluded.

2.3 Relation prior to entity

The claim that entities are constituted by the relations they inhabit, rather than entering relations already formed, is held and has been held for some time. Emirbayer's statement of the relational position against substantialism is explicit and programmatic (Emirbayer, 1997); actor-network theory holds that an actor is an effect of a network rather than a node within one (Latour, 2005); assemblage theory develops a related ontology with attention to scale (DeLanda, 2006). The ontological premise on which the present paper's motivating framework rests is therefore not new, and is not claimed here.

The strongest statement of the specific problem of this paper also belongs here rather than to any programme built upon it. Abbott's treatment of boundaries argues that social entities are not given and then bounded, but are formed when boundaries of difference are yoked together into a unit, so that the making of the thing and the making of its edge are one process (Abbott, 1995b). That is the problem of §1 stated precisely, and it was stated three decades ago.

What this group characteristically declines is a criterion that would discriminate among formations, and in the case of actor-network theory the refusal is deliberate and methodological. The consequence for present purposes is that the ontology is available, the problem has been named, and the audit is not. Where a criterion is offered it is offered as an interpretive judgement, and interpretive judgements are not the kind of thing a reader can check against the analyst's own machinery.

2.4 Explanation by generation

The fourth group explains a macro pattern by generating it. Agent-based modelling in the generative programme takes the sufficiency of a specified micro process to produce an observed

macro regularity as the form of explanation, in Epstein's formulation that a pattern not grown has not been explained (Epstein, 2006). This is the nearest neighbour of the present proposal, it shares its vocabulary, and it is a mature programme with an established literature.

Its known difficulty is the one that bears directly here. Generation establishes sufficiency and does not establish identification: distinct micro specifications routinely produce indistinguishable macro output, so that a successful simulation constrains the space of candidate mechanisms much less than the demonstration suggests.

That literature has answered the difficulty, and the answer must be conceded because it is close to the proposal made here. Pattern-oriented modelling requires a model to reproduce several patterns simultaneously, observed at different hierarchical levels of the system, on the reasoning that each additional pattern is a further filter on the space of admissible specifications (Grimm et al., 2005). Multiple patterns at multiple levels used as a filter is a cross-scale adequacy requirement in substance, and it is in working use. The second answer is protocol: a standardised description under which a model can be read, replicated, and compared, revised in the light of a decade of use (Grimm et al., 2010). Both answers are the right kind of answer, and the present paper is in the same genre as the second.

Two things are nonetheless absent from both. The filter operates on patterns the modeller has selected as targets, and the selection is not itself declared as part of the hypothesis space; and neither answer asks the modeller to state what the model forbids, as distinct from which patterns it reproduces. A model reproducing five patterns has been shown adequate to five patterns. The set of trajectories it rules out is a different object and is not reported.

2.5 *The residue*

Taken together the four groups already hold: that units are relationally constituted and that entity and boundary are formed together; that trajectories rather than variables may be the unit of comparison; that conjunctural and equifinal causation requires its own apparatus; that formation can be reconstructed genealogically; that within-case evidence can be graded by what it would show were the hypothesis false; and that generative sufficiency is a form of explanation, filtered by patterns at several levels and reported under a protocol.

The residue is correspondingly narrow, and narrower than it appeared before the concessions were made. What none of the four requires is that the analyst state how the compared unit came to exist, that the passages between scales asserted by the account be identifiable in the account's own machinery rather than in its prose, and that the account state what it excludes at the level of the structure rather than at the level of a hypothesis within it. The next section states these three and defends the claim that each is currently unrequired.

3 DECLARATIONS THAT NO CURRENT STANDARD REQUIRES

Claim 3.1. (Unit generation). *Where the unit of analysis is produced by the process under study, the analyst is at present required to establish that the unit travels and is not required to state how it was produced. The two are different requirements, and satisfying the first is consistent with any answer to the second.*

The distinction can be tested on the apparatus itself. A demonstration of measurement validity establishes that the indicators bear on the concept and that the cases assembled are cases of one kind. It is silent on whether the kind was generated in the interval under study, and a comparison can satisfy every validity criterion while comparing a relation with its own descendant. Where the generating process is among the variables, the silence is not innocuous: the analyst's unit and the analyst's explanandum have a common cause that the design does not represent.

Claim 3.2. (Scale production). *Multi-level designs require that the levels be specified and do not require that the production of a level be represented. A study may therefore claim a cross-level result while representing only the nesting of one level in another.*

Coleman's transition names the missing representation and does not supply it (Coleman, 1990); the analytical-sociology programme that took up the problem has supplied mechanism vocabulary for the downward and upward arms (Hedström & Ylikoski, 2010) without making their representation a condition of reporting. The declaration at issue is modest: which passages between scales the account asserts, and by what mechanism, stated in a form an auditor could check against the account's own machinery. Section 5 makes it checkable.

Claim 3.3. (Falsification surface). *Where a falsification surface is required at all, it is required of hypotheses within a case and not of the reconstructed structure as a whole. No prevailing requirement obliges an account of a generative process to state which trajectories the account excludes, and an account that excludes nothing is consistent with every history and discriminates among none.*

The first clause records the concession made in §2: the graded evidentiary tests of process tracing do require the analyst to say what would be found were the hypothesis false (Collier, 2011). The claim is that this requirement has not been carried up a level. A reconstruction may satisfy it proposition by proposition and still, taken as a structure, be compatible with every course the history could have taken, since the tests range over evidence for parts and not over the trajectories the whole admits.

This is the point at which the present proposal parts from the generative programme rather than extending it. Where explanation takes the form of demonstrated sufficiency, the demonstration establishes that the specified process can produce the observed pattern, and the space of other processes that would produce it equally well is left unmeasured (Grimm et al., 2005). The remedy proposed here does not require the analyst to obtain negative evidence, which

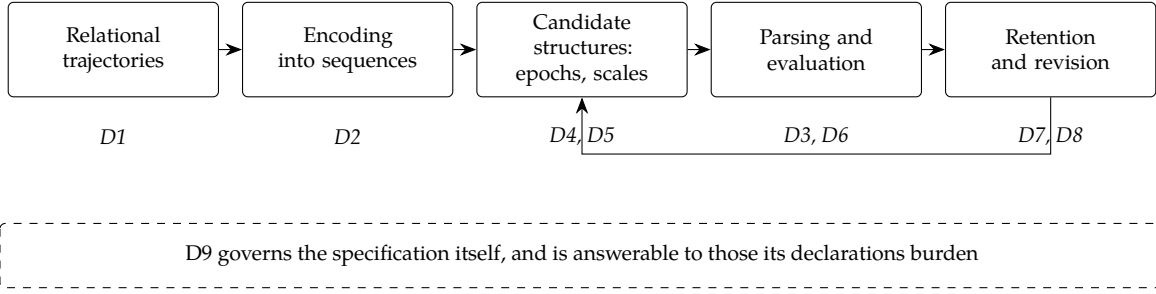


Figure 1: Where each declaration attaches. The chain is the ordinary course of an analysis and is not prescribed by the specification; what the specification fixes is the tag beneath each step. The return arrow marks revision in the light of what was not accounted for, and nothing in the specification requires that the loop terminate.

history does not furnish. It requires the analyst’s model to generate negative claims, which any model with genuine structure does, and to state them.

The three claims are the specification’s warrant. Each is a claim about prevailing requirements rather than about prevailing practice: careful analysts do some of this some of the time, and that is precisely the situation in which a standard, rather than an argument, is the appropriate instrument.

4 THE SPECIFICATION FIXES DECLARATIONS, NOT PROCEDURES

The specification is an interface. It states what an implementation must make available and leaves the implementation free in everything else, on the understanding that an interface with no contract constrains nothing: a list of operation names admits any behaviour whatever, and the content lies in the conditions attached.

Five operations are named, and naming them is not the content of the specification. An implementation *generates* candidate structures, *evaluates* them against trajectories, *retains* them in some form, *transforms* them, and *re-evaluates* them under evidence acquired later. The cycle is not novel and is not claimed as such; the archive-and-revisit posture in particular restates positions long held in the philosophy of science, where the retention of currently inferior rivals is argued for directly (Lakatos, 1970; Feyerabend, 1975; Laudan, 1977). What the specification adds is the contract. Figure 1 shows where each declaration attaches to the ordinary course of an analysis, and Table 1 states the division of labour.

Two features of the table deserve comment. The left column contains no procedure, which is what permits an interpretive reconstruction, a statistical estimation, and a machine induction to be held to one contract and their results compared. The right column contains the whole of what is ordinarily called method, which is what permits the specification to be adopted without adopting any theory, including the one from which it arose.

Table 1: What the specification fixes and what it leaves to the implementation.

Fixed by the specification	Left to the implementation
That a partition of the state space is declared	Which partition is chosen
That epochs and their boundaries are marked	Where the boundaries are placed
That scale-crossing productions are identified	How many scales are distinguished
That the excluded set is stated	Whether exclusion is structural or by a declared likelihood threshold
That complexity is reported against coverage on withheld trajectories	Which complexity measure is used
That suspension carries a stated condition of re-entry	When a structure is suspended
That the grounds of sampling and of pruning are recorded	How sampling and pruning are performed
That the specification's own revision procedure is stated	—

5 GENERATIVE STRUCTURE IS PRESENTABLE AS A STRATIFIED GRAMMAR

5.1 *A grammar states what can be derived, and thereby what cannot*

Readers already at home with generative grammar may pass to §5.2. The apparatus used below is elementary and is set out here so that nothing in the propositions depends on prior acquaintance with it.

A grammar has four parts. *Terminals* are the observable symbols, the things that appear in the record. *Non-terminals* are auxiliary symbols that appear in no record and serve to organise the derivation. *Productions* are rules of the form $A \rightarrow \beta$, read as permission to replace the symbol A by the string β . One distinguished non-terminal is the *start symbol*. A *derivation* begins at the start symbol and applies productions until only terminals remain; the sequence so produced is *generated* by the grammar, and the set of all sequences it can generate is the grammar's *language*.

The construction is due to Chomsky (Chomsky, 1957, 1965), where the terminals are words and the language is the set of admissible sentences. Nothing in the formal apparatus requires that reading. The terminals may be any finite set of observable symbols, and in what follows they are the cells of a partition of a space of relational states.

The property that matters here is the last one. A grammar does not merely describe the sequences it was built from; it determines a language, and everything outside that language is thereby excluded. Figure 2 gives the smallest example that shows this. The terminals are three kinds of recorded event: a dispute at the level of a single school, an episode of consultation with an organised body, and a change in the standing rules. The grammar generates any number of disputes followed by consultation and then a rule change. It generates the sequence in which two disputes precede consultation and a rule change. It does not generate

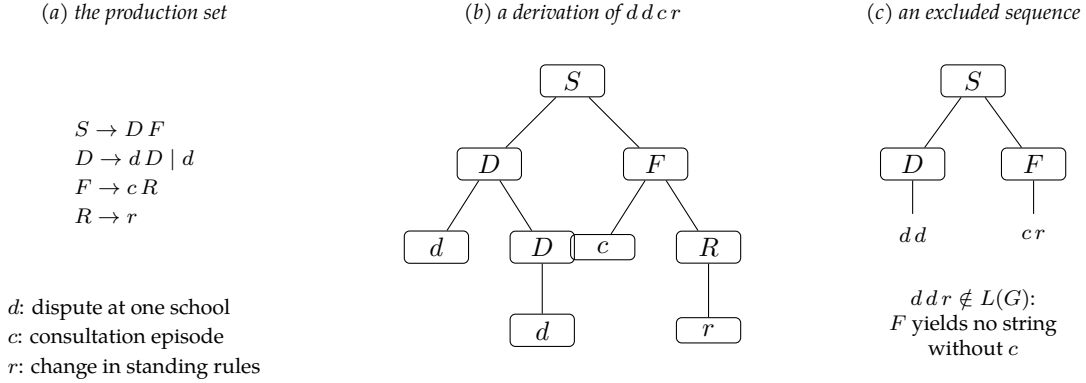


Figure 2: A grammar over three kinds of recorded event. The production set (a) generates sequences of one or more disputes followed by consultation and a rule change, and (b) gives the derivation of one such sequence. The sequence in (c), in which a rule change follows disputes with no consultation, has no derivation and lies outside the language. The exclusion is a property of the production set, not a remark added by the analyst.

the sequence in which a rule change follows disputes with no consultation between, because no derivation reaches that sequence: the only production introducing a rule change requires that consultation be introduced with it.

That last sentence is the whole point of using the apparatus. The claim “rule changes here do not occur without prior consultation” is not an interpretation added to the analysis afterwards; it is a property of the production set, present whether or not the analyst notices it, and checkable by anyone who has the production set. This is what §6 asks analysts to state, and §5.6 makes precise for the case in which the grammar is probabilistic rather than categorical.

5.2 What the presentation does not deliver

The limits are stated before the machinery, because they are conditions of its design rather than qualifications upon it.

The formalism does not deliver identification. Distinct grammars parse the same trajectory set, and nothing below removes that; the equifinality already noted in configurational and generative work (Ragin, 2008; Grimm et al., 2005) is a property of the object.

It does not deliver learnability in the classical sense. Historical evidence is positive evidence: what occurred is recorded and what could not have occurred is not. Gold’s result establishes that no class of languages containing all finite languages and at least one infinite language is identifiable in the limit from positive presentation alone (Gold, 1967), and the regular and context-free classes are of this kind. This is why the presentation in §5.6 is probabilistic: for stochastic grammars, learnability from positive data is recoverable in a Bayesian setting (Horning, 1969), and the problem is thereby moved from identification to model comparison, where it can be honestly conducted.



Figure 3: One history, two partitions, two sequences. The cell boundaries of π and π' interleave, so neither partition refines the other and the two sequences are strings over disjoint alphabets. The apparent stability of the history under π' , which records a single excursion, and its volatility under π , which records two, are properties of the encoding.

It does not deliver decidability if the production set may be rewritten during a derivation. Proposition 5.4 states the consequence, which is that stratification is forced.

It does not deliver an encoding. A history is not a symbol sequence until a partition is imposed, and the partition is a choice. This is familiar from symbolic dynamics, where a Markov partition yields a subshift and the correspondence between sofic subshifts and regular languages is exact, while the partition itself is not canonical (Lind & Marcus, 1995). Proposition 5.3 states what follows for the analyst.

5.3 Trajectories become analysable only under a declared partition

Definition 5.1. (Encoding). Let a relational history be a map $H : T \rightarrow \Sigma$ from an ordered index set to a state space. A finite partition π of Σ with cells labelled from an alphabet A_π induces a symbol sequence $w_\pi(H) \in A_\pi^*$. All subsequent analysis is conducted on $w_\pi(H)$ and is relative to π .

Proposition 5.2. (Encoding relativity). Let π and π' be partitions of Σ such that neither refines the other. Then for a grammar G over A_π , the acceptance or rejection of $w_\pi(H)$ entails nothing concerning $w_{\pi'}(H)$, since the latter is not a string over A_π . A statement of the form “ G accounts for H ” is accordingly well formed only as “ G accounts for H under π ”.

Figure 3 shows what is at stake. The proposition is immediate and its interest is not mathematical. It fixes what the analyst may assert. A finding stable across a family of partitions is a finding about the history; a finding that moves with the partition is a finding about the encoding; and an analysis that never varies its partition cannot tell which it has. The declaration required in §6 is correspondingly of two parts: the partition, and the range of partitions over which the reported findings were checked.

Defeat sketch. The proposition is defeated for an important family of cases, and the defeat should be stated because it bears on how cheap the corresponding declaration is. Suppose every partition under consideration arises as a coarsening of one finest partition π^* . Each

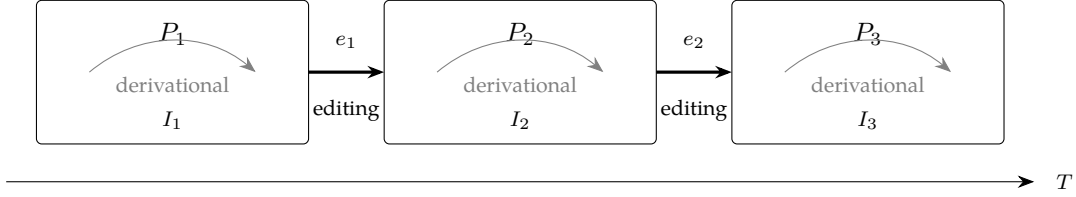


Figure 4: Two sites, kept disjoint. Derivational operations act on strings within an epoch, where the production set is held fixed; editing operations act on the production set and apply only at a boundary. Permitting an editing arrow to enter the interior of a box makes the system a general rewriting system, and by Proposition 5.4 failure to account for a trajectory then ceases to be possible.

coarsening then induces a letter-to-letter homomorphism $h : A_{\pi^*} \rightarrow A_{\pi}$, the image of a context-free language under such a homomorphism is context-free, and acceptance transfers along h by construction. Within any family of coarsenings of a declared π^* , Proposition 5.2 has no force. What survives is the case in which the partitions in play have no common refinement that the analyst is willing to declare, which is the ordinary case when a periodisation is introduced for one purpose and a second is introduced later for another. The constructive consequence is a strengthening rather than a weakening: an analyst who can exhibit π^* and derive the working partitions from it has discharged D2 in the strongest available form, and an analyst who cannot has thereby located an unexamined commitment.

5.4 Rewriting is confined to epoch boundaries

The object of interest includes the alteration of the generative structure itself, and a formalism that fixed the production set for all time could not represent it. A formalism that allowed the production set to change at an arbitrary derivation step, however, would represent it at the cost of representing everything else as well.

Definition 5.3. (Stratified grammar). A stratified grammar over an index set $1, \dots, n$ is a sequence $\mathcal{G} = \langle (I_k, P_k) \rangle_{k=1}^n$ of epochs, where I_k is an interval of the index set and P_k a production set held fixed throughout I_k , together with a set of editing operations e_k , each mapping P_k to P_{k+1} and applying only at the boundary between I_k and I_{k+1} . Derivational operations act on strings with P_k fixed; editing operations act on production sets and are disjoint from the derivational operations.

The two-site construction is adopted from the treatment of grammar rewriting in the author’s Metamodel series, where its function is to keep the measure of collapsed distinctions non-circular (Huang, 2026). Its function here is different and simpler.

Proposition 5.4. (Stratification). Let each P_k be context-free. Then membership of w in the language generated within epoch k is decidable. If instead editing operations may be interleaved with derivational operations at arbitrary steps, the combined system is a general rewriting system, membership is undecidable in the general case, and there exist such systems whose generated language is A^* .

The first part is the standard decidability of the context-free membership problem; the sec-

and is the standard undecidability of the word problem for unrestricted rewriting (Hopcroft et al., 2007). The methodological corollary is the point. Any measure of how well a structure accounts for a trajectory presupposes that failing to account for a trajectory is possible, and under unrestricted mid-derivation rewriting it is not. Stratification is therefore a condition of the evaluation being a measure at all, and the placement of epoch boundaries becomes a substantive analytic claim rather than a convenience: to say where the generative structure was rewritten is to say something that can be wrong.

The proposition establishes that stratification suffices and does not establish that it is necessary. Restricted rewriting systems with tractable decision problems are known, and a formulation permitting interleaved editing under suitable restrictions may well be available. Should one be exhibited, what survives is the requirement rather than the device: some restriction must be declared under which failure to account for a trajectory remains possible, and stratification is the simplest such restriction rather than the only one.

Defeat sketch. A concrete candidate can be named, and it is worth naming because its cost is instructive. Let the rule inventory be a fixed finite set R , let every production set be a subset of R , and let editing operations be *monotone*, that is, permitted to add a production and not to remove one. Then the sequence of production sets reachable during a derivation is a chain in the finite lattice of subsets of R , the number of distinct sets encountered is bounded by $|R|$, and membership decomposes into finitely many context-free membership problems and remains decidable. Editing may then be interleaved freely with derivation, and the necessity half of Proposition 5.4 fails. The cost is that monotone editing cannot represent the withdrawal of a distinction: a production once available remains available, so a system that loses the capacity to make a distinction it formerly made has no representation. That loss is precisely the phenomenon the two-site construction was built to capture elsewhere (Huang, 2026). The choice between stratification and monotone interleaving is therefore not a choice between a correct and an incorrect formalism, but a choice about which class of historical change the analysis intends to be able to state.

5.5 Scale is carried by the symbols

Definition 5.5. (Scale typing). Let each non-terminal A carry an index $\sigma(A) \in \{1, \dots, L\}$ and each terminal carry index 0. A production $A \rightarrow \beta$ is *scale-preserving* if every symbol of β has index $\sigma(A)$, and *scale-crossing* if some symbol of β has index strictly less than $\sigma(A)$.

Proposition 5.6. (Scale monotonicity). Let G contain only scale-preserving productions and let S be a start symbol with $\sigma(S) = \ell > 0$. Then no derivation from S yields a string of terminals, and $L(G) = \emptyset$. Consequently a grammar that derives observations from a macro-level start symbol contains at least one scale-crossing production, and the set of such productions is determined by inspection of G alone.

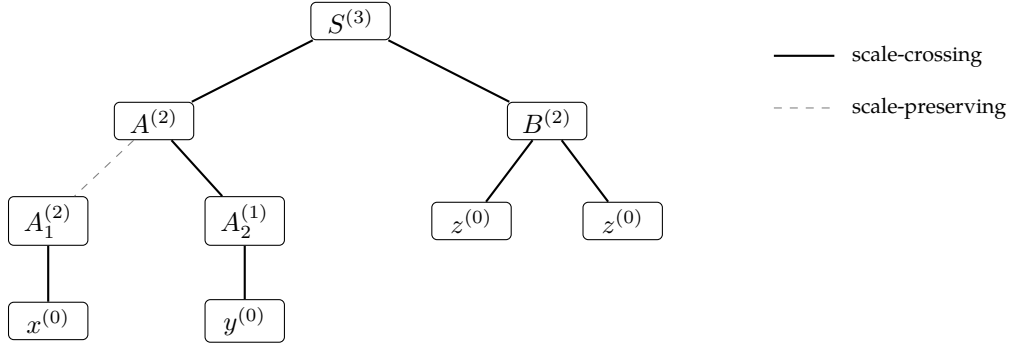


Figure 5: A derivation with scale indices carried by the symbols. Superscripts give σ ; terminals carry 0. The production $A^{(2)} \rightarrow A_1^{(2)} A_2^{(1)}$ is scale-crossing on its right conjunct only. A grammar in which every edge were of the dashed kind would reach no terminal, so the presence and the placement of the solid edges is what a cross-scale claim amounts to, and it is read off the production set rather than from the surrounding prose.

The proof is immediate: every symbol reachable from S carries index ℓ by induction on derivation length, and terminals carry index $0 \neq \ell$. Figure 5 shows the two kinds of edge. The value of the proposition lies entirely in its corollary. A claim to have connected micro and macro is ordinarily assessed by reading the author’s narrative; under Definition 5.5 it is assessed by inspecting the production set, and a structure whose scale-crossing productions are absent, or are confined to a single link inserted to make the derivation go through, displays that fact syntactically. This is the sense in which a criterion that is usually rhetorical becomes auditable.

Defeat sketch. The audit is defeated by a single assignment, and the defeat is easy enough that it must be closed rather than noted. Let $\sigma(A)$ be the depth at which A occurs in the derivation, with L the height of the tree. Every production is then scale-crossing by construction, since children lie deeper than their parent, Proposition 5.6 is satisfied for any grammar whatever, and the inspection reports nothing. The same happens under any assignment made after the structure is fixed and in the light of it. The repair is to require that the index be anchored outside the grammar: the index of a non-terminal must be the scale at which the pattern it stands for is observed in the material declared under D1, so that the number of indices is fixed by the observation design and the assignment of a symbol to an index is answerable to it. D5 is stated to carry this requirement. Where the anchoring is absent, Proposition 5.6 remains true and the audit built on it is empty, which is the more dangerous of the two failures because it leaves the appearance of a check.

5.6 The probabilistic form, and what the structure forbids

Definition 5.7. (Probabilistic form). Each production in P_k carries a weight, with weights normalised over the productions sharing a left-hand side, inducing a distribution over derivations and

hence over strings. Comparison between candidate structures is conducted by penalised description length, or equivalently by posterior probability under a prior over grammars (Rissanen, 1978; Grünwald, 2007).

Two consequences follow. The first is the one already stated: model comparison under positive evidence is well posed where identification is not (Horning, 1969), and the induction methods that would implement it are established, whether by Bayesian merging of non-terminals (Stolcke & Omohundro, 1994) or by compression-driven discovery of hierarchical structure (Nevill-Manning & Witten, 1997), the latter yielding a hierarchy for a single sequence rather than a grammar that generalises. The specification does not select among them.

The second concerns exclusion, and it requires care that the deterministic case does not.

Definition 5.8. (Excluded set). *The excluded set of a structure is either the complement of its generated language, where the analyst’s claim is structural, or the set $\{w : P_G(w) < \theta\}$ for a threshold θ declared in advance, where the claim is probabilistic. In the probabilistic case every string in the language carries positive probability, so exclusion without a declared θ has no content.*

Proposition 5.9. (Vacuity). *A structure whose excluded set is empty is consistent with every trajectory over its alphabet and discriminates among none. Its agreement with the observed record is therefore not evidence in its favour.*

The excluded set does not survive coarsening, and the direction of the failure should be recorded because it constrains what an exclusion claim can mean. Let h be the projection from a finer partition to a coarser one, as in the defeat sketch of Proposition 5.2. If some w' with $h(w') = h(w)$ lies in the generated language, then $h(w)$ lies in the coarser language even where w is excluded at the finer level. Coarsening can therefore only shrink the excluded set, never enlarge it. An exclusion is accordingly a claim made at the finest partition the analyst has declared, and it weakens monotonically as the reader coarsens. Whether a notion of exclusion invariant across a family of partitions can be defined is left open in §8.

This is the formal statement of Claim 3.3, and it is the reason the corresponding declaration in §6 is a requirement rather than a recommendation. It is also the point at which the specification does the work that a standard is for: an analyst who cannot state an excluded set has not been shown to be wrong, but has been shown to have made no claim of the kind the analysis purports to make.

5.7 Retention is indexed by coverage, not by score

A single scalar quality measure over candidate structures suppresses precisely the information the specification is designed to preserve, since two structures with equal scores may account for disjoint portions of the record. A retained population should therefore be indexed by the profile of what each structure covers: which trajectories, at which scales, under which

partitions. The practice is established in evolutionary computation, where archives indexed by behavioural descriptors retain the best structure found in each region rather than the best overall (Lehman & Stanley, 2011; Mouret & Clune, 2015), and Pareto methods are available where several evaluation dimensions must be kept separate (Deb et al., 2002). A structure discarded for low aggregate score cannot be recovered when later evidence favours it; a structure retained as the best in some region can be.

Diversity is a property of a population and not of an individual structure, and cannot be a term in the evaluation of one. This is stated because the contrary formulation is a natural error.

6 EVERY IMPLEMENTATION OWES THE SAME DECLARATIONS

The declarations follow. They are graded in the manner of technical specifications: a requirement is a condition of conformance, and a recommendation may be departed from where the departure is stated (Bradner, 1997). The grading exists because a specification more expensive than the research it governs is not adopted, and the observation that reporting standards are followed to the letter and evaded in substance is well documented in the fields that have them.

D1. Trajectory (requirement). *Declare the trajectories analysed, the index set that orders them, the resolution of that ordering, and the scales at which they are observed. Evidence that supplies no resolved ordering across at least two scales does not support the claims this specification governs. Cross-sectional data, single-scale network snapshots, and equilibrium statistics fail this declaration.*

Conforming. Three scales named, annual resolution, the state at each scale tabulated for every year of the period. *Non-conforming.* Two cross-sections a decade apart, with the interval supplied narratively and no ordering resolved within it.

D2. Encoding (requirement). *Declare the partition under which histories were rendered as sequences, and the range of alternative partitions over which the reported findings were checked. Report which findings moved. Where the partitions in use are coarsenings of a common finest partition, declare that partition, since the declaration then discharges this requirement in its strongest form. A finding reported without this declaration is not distinguishable from a property of the encoding (Proposition 5.2).*

Conforming. The cells of the partition listed, and a table reporting which findings survived two coarser partitions and which did not. *Non-conforming.* A periodisation into phases whose boundaries enter through the narrative, are never varied, and carry the result.

D3. Excluded set (requirement). *State what the structure forbids: either the structural exclusion, or the likelihood threshold and the set it induces. Exhibit at least one trajectory that is excluded and that a reader could plausibly have expected to observe. A structure with an empty excluded set makes no claim (Proposition 5.9).*

Conforming. A named class of trajectories the structure cannot derive, with one instance a reader might have expected to observe. *Non-conforming.* A statement that the account is consistent with the record and can accommodate variation, with nothing named that it cannot accommodate.

D4. Epochs (requirement where generative change is asserted). *Where the analysis asserts that the generative structure itself changed, mark the epochs, the boundaries, and the editing operations applied at each boundary. Rewriting within an epoch is not permitted, and an analysis that requires it has asserted an epoch boundary it has not declared (Proposition 5.4).*

Conforming. Two boundaries marked, with the productions added and removed at each stated explicitly. *Non-conforming.* An account of continuous transformation that relies, at one point, on a rule unavailable earlier in the same derivation.

D5. Scale crossing (requirement where a cross-scale claim is made). *Identify the scale-crossing productions on which the cross-scale claim rests, and state the mechanism each represents. Declare the assignment of scale indices and its anchor in the observation design of D1, so that the index of a symbol is the scale at which the pattern it stands for is observed; an assignment made in the light of the finished structure makes every production scale-crossing and the identification vacuous. The identification is by inspection of the structure and is not satisfied by narrative assertion (Proposition 5.6).*

Conforming. Two scale-crossing productions named, each with the mechanism it stands for. *Non-conforming.* The sentence that repeated interactions aggregate into a collective norm, with no element of the structure corresponding to the aggregation.

D6. Complexity against coverage (requirement where compression is claimed). *Report the size of the structure against its coverage, and report coverage on trajectories withheld from the construction of the structure. Compression measured on the trajectories used to build the structure is not evidence of compression.*

Conforming. Rule count against coverage, reported on two trajectories withheld from construction, including the loss of coverage. *Non-conforming.* The report that the structure accounts for all twelve cases from which it was assembled.

D7. Suspension (recommendation). *Where a candidate structure is set aside, record the condition of its re-entry: the evidence that would return it to consideration. Setting a structure aside without such a condition and setting it aside permanently are indistinguishable in the record, and a programme that can neither abandon nor recall its rejected candidates has no mechanism of self-correction. Where several candidate structures are retained, index the retained set by what each covers rather than by an aggregate score, so that a structure best in some region of the record is recoverable from the archive.*

Conforming. A candidate set aside, together with the class of trajectory whose appearance would return it to consideration. *Non-conforming.* An early candidate that ceases to be mentioned.

D8. Sampling and pruning (recommendation). *Record which trajectories were selected and on what grounds, and which couplings were excluded from consideration before evaluation and on what grounds. The discoverable structure is bounded by the range of scales and contexts present in the material examined, so these grounds are part of the hypothesis space and not part of the validation.*

Conforming. The trajectories considered and set aside, with grounds, and the couplings excluded before evaluation, with grounds. *Non-conforming.* The trajectories used, only.

D9. Revision of the specification (requirement upon the specification). *The specification states its own procedure of amendment and the standing to invoke it, which belongs to those on whom its declarations impose a burden. A specification not revisable by those it burdens has the form it exists to prevent, namely a structure that fixes the terms under which others must present themselves and is not answerable to them.*

The last declaration is directed at a foreseeable pathology rather than at a hypothetical one. A reporting requirement becomes, over time, a criterion of competence; conformity to it is then read as quality, and the substantive matter it was introduced to make visible resumes its invisibility behind a satisfied checklist. That standards are not neutral instruments but infrastructures that render some things visible and others not is the settled finding of the literature on classification and standardisation (Bowker & Star, 1999; Timmermans & Berg, 2003; Lampland & Star, 2009), and a specification that ignored it would be refuted by its own subject matter.

7 ONE TRAJECTORY, FULLY DECLARED

The illustration is deliberately thin. Its purpose is to show the shape of a conforming report, not to conduct an analysis, and the substantive study to which it gestures is a separate undertaking.

Take as the trajectory the accommodation of a minority religious practice within a state system of public education, in one polity, across roughly four decades. No state is named, since the point of the illustration is the form of the declarations and a named case would invite assessment of the substance instead.

Under D1 the analyst declares three scales: the interaction between a family and a school; the negotiation between an organised religious body and an administrative authority; and the constitutional or statutory settlement. The index set is annual, the ordering is resolved to the year, and the trajectory is the sequence of states at each scale.

Under D2 the analyst declares a partition of the state space at the middle scale into cells such as absence of any standing forum, an ad hoc forum convened per dispute, a standing consultative body without decision rights, and a body with delegated decision rights. The declaration records that the reported findings were checked against a coarser partition that merges the last two, and that one finding, concerning the timing of the transition, did not survive the merger; that finding is accordingly reported as encoding-relative.

Under D4 the analyst marks two epoch boundaries where the production set was edited, rather than treating the whole four decades as one generative regime, and states which productions were added or removed at each.

Under D5 the analyst identifies the scale-crossing productions: the production by which a repeated pattern of school-level dispute yields a standing forum at the middle scale, and the production by which a settlement at the upper scale constrains which forums are available below. Two arrows, both inspectable.

Under D3 the analyst states an exclusion, and this is where the illustration earns its place. Suppose the structure contains no production deriving a delegated-rights body directly from the upper scale without a prior middle-scale standing body. Then the structure forbids trajectories in which delegated decision rights appear by legislative or judicial imposition where no consultative body preceded them. Such trajectories are not hard to imagine, and if one is exhibited the structure is wrong. That sentence is what the specification exists to obtain.

Under D6 the analyst reports the size of the structure against its coverage on the trajectories of two further polities withheld from its construction, and reports honestly if coverage falls.

A second trajectory is set beside the first because the two exercise different branches of Definition 5.8. Take the emergence of a distinct occupational category within an industry across two decades: a set of tasks initially dispersed among existing roles, then named, then held by identifiable post-holders, then certified, and finally required. Here the analyst has many instances of the same passage across many firms, and the structure is estimated in probabilistic form rather than asserted structurally.

The exclusion under D3 accordingly takes the second branch. The analyst declares a threshold in advance of examining the withheld firms, and the excluded set is the set of trajectories whose likelihood under the structure falls below it. The substantive content might be that trajectories in which certification appears without any prior inter-firm body that could confer it fall below the threshold. This is a weaker statement than the structural exclusion of the first illustration, and its weakness should be visible in the report rather than concealed by it: a threshold chosen after inspection of the withheld material excludes nothing, and the declaration is therefore of the threshold and of the point at which it was fixed.

The contrast between the two illustrations is the reason both are given. A structural exclusion is sharp and is available only where the analyst is prepared to assert that a passage cannot occur. A probabilistic exclusion is available far more widely and carries a discipline of its own, which is that the threshold must be fixed before the evidence that will test it is seen.

8 CONDITIONS OF REFUTATION AND OPEN QUESTIONS

Claim 3.1 fails upon the exhibition of a prevailing methodological standard under which an account of the generation of the compared unit is required rather than commended. Claim 3.2

fails likewise for the representation of scale production. Claim 3.3 fails upon the exhibition of a standard requiring accounts of process to state their exclusions. In each case the claim is about requirements, and a single reporting guideline that imposes the requirement refutes it.

The three formal propositions carry their own defeat sketches in §5, and the sketches are the more informative part of them. Two of the three are already defeated in a restricted setting by constructions given there, namely the projection homomorphism over coarsenings of a common partition and monotone editing over a finite rule inventory, and the third is defeated by any assignment of scale indices made in the light of the finished structure. What survives in each case is a requirement rather than a theorem, and the requirements are what §6 states. A reader who wishes to attack the formal material should attack the requirements: exhibit a family of partitions with a declared common refinement under which D2 is nonetheless idle, an analysis of generative change that needs neither stratification nor monotone editing nor any other declared restriction, or an anchoring of scale indices in an observation design that still permits the audit of D5 to be passed vacuously.

The whole proposal fails, in a way no single proposition captures, if the declarations turn out to be satisfiable at negligible cost by work that has not done what they are meant to secure. That is the characteristic failure of reporting standards, and no formal result prevents it.

The questions left open are the following, and the first two are narrower than they were before the defeat sketches were written.

The individuation of scales. D5 closes the vacuity of the cross-scale audit by anchoring the indices in the observation design, and in doing so it transfers the problem rather than solving it. What the anchoring does not supply is a criterion by which two patterns observed under different arrangements count as the same pattern at the same scale, nor any principle fixing how many scales an observation design ought to distinguish. An analyst who multiplies scales will find scale-crossing productions everywhere, and the anchoring constrains that only to the extent that the observation design was itself fixed for independent reasons.

The adjudication of epoch boundaries. Definition 5.3 makes the placement of a boundary a substantive claim, and Proposition 5.4 makes something depend on it, but nothing here indicates how a disputed placement is settled. The natural proposals are unattractive in different ways: penalised model comparison makes the boundary an artefact of the penalty, and interpretive judgement returns the question to the condition the specification was written to escape.

The invariance of exclusion. Exclusion weakens monotonically under coarsening, by the remark following Proposition 5.9, so an exclusion claim is always relative to the finest declared partition. Whether a notion of exclusion invariant across a family of partitions can be defined, and whether it would be strong enough to be worth stating, is unsettled. This is the most technically tractable of the open questions and the one on which progress would most improve D3.

The residual underdetermination. The specification improves the discipline with which candidate structures are reported and does not address identification. Distinct structures conforming fully to every declaration may account for the record equally well, and nothing here says what would further constrain the space; the pattern-oriented answer of filtering on several patterns at several levels (Grimm et al., 2005) is available and is not obviously sufficient.

A direction is worth naming separately from the questions, because it bears on all of them and is available now. The specification cannot at present be tested, since history furnishes no case in which the generative structure is known independently of the analysis that reconstructs it. A simulated society of interacting artificial agents furnishes exactly that. Multi-agent systems built from language models now produce extended records of interaction, role formation, and institutional change at more than one scale (Park et al., 2023), and the generative programme in social science has long used simulation as its instrument (Epstein, 2006). What a simulation adds here is not realism but *ground truth*: the experimenter fixes the rules by which the artificial society operates, and can then run the pipeline of Figure 1 on the record it leaves and ask whether the declared structure recovers what was in fact installed. Sensitivity to the encoding can be measured by varying the partition; the cost of each declaration can be measured by the difference between conforming and non-conforming reports of the same run; and the excluded set can be tested directly, since the experimenter can attempt to produce a trajectory the reconstructed structure forbids.

Two limits on that direction should be stated with it. Recovering a structure from a simulation tests the induction machinery and not the applicability of the apparatus to history, and a positive result under artificial conditions establishes nothing about a record generated by no rule-writer. Further, a simulator's own specification is already a rule system of the kind the analysis seeks, so a favourable result may reflect the affinity between the two rather than the adequacy of either; the informative design is accordingly one in which the artificial society's regularities emerge from agent interaction rather than being written into the simulator, and in which the experimenter's expectations about what will be recovered are recorded before the reconstruction is attempted.

Adoption. A specification issued by a single author has no adopting body and is therefore a proposal. What would convert declarations of this kind into requirements is a question about the institutions of a field rather than about method, and it is not one this paper can answer.

A final remark on standing. The approaches surveyed in §2 are achievements, and the class of failures identified here is not a deficiency in any of them: the failure is invisible precisely because it falls between the requirements each was built to enforce. Nothing here supports the inference that an account is improved by being expressed in the vocabulary of generation, and a structure declared under this specification is not thereby a good one. It is a claim that can be found wrong, which is the whole of what is asked.

REFERENCES

- Abbott, Andrew (1988). Transcending General Linear Reality. *Sociological Theory*, 6(2), 169–186.
- Abbott, Andrew (1995a). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, 21, 93–113.
- Abbott, Andrew (1995b). Things of Boundaries. *Social Research*, 62(4), 857–882.
- Bennett, Andrew, and Jeffrey T. Checkel (eds.) (2015). *Process Tracing: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press.
- Adcock, Robert, and David Collier (2001). Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review*, 95(3), 529–546.
- Beach, Derek, and Rasmus Brun Pedersen (2013). *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.
- Bowker, Geoffrey C., and Susan Leigh Star (1999). *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bradner, Scott (1997). *Key Words for Use in RFCs to Indicate Requirement Levels*. RFC 2119, Internet Engineering Task Force.
- Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Collier, David (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44(4), 823–830.
- Collier, David, and James E. Mahon Jr. (1993). Conceptual “Stretching” Revisited: Adapting Categories in Comparative Analysis. *American Political Science Review*, 87(4), 845–855.
- Coleman, James S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- DeLanda, Manuel (2006). *A New Philosophy of Society: Assemblage Theory and Social Complexity*. London: Continuum.
- Deb, Kalyanmoy, Amrit Pratap, Sameer Agarwal, and T. Meyarivan (2002). A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197.
- Emirbayer, Mustafa (1997). Manifesto for a Relational Sociology. *American Journal of Sociology*, 103(2), 281–317.
- Epstein, Joshua M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press.
- Feyerabend, Paul (1975). *Against Method*. London: New Left Books.
- Foucault, Michel (1977). *Discipline and Punish: The Birth of the Prison*. Trans. Alan Sheridan. New York: Pantheon.

- Gold, E. Mark (1967). Language Identification in the Limit. *Information and Control*, 10(5), 447–474.
- Grimm, Volker, Eloy Revilla, Uta Berger, et al. (2005). Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, 310, 987–991.
- Grimm, Volker, Uta Berger, Donald L. DeAngelis, et al. (2010). The ODD Protocol: A Review and First Update. *Ecological Modelling*, 221(23), 2760–2768.
- Grünwald, Peter D. (2007). *The Minimum Description Length Principle*. Cambridge, MA: MIT Press.
- Hedström, Peter, and Petri Ylikoski (2010). Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, 36, 49–67.
- Hopcroft, John E., Rajeev Motwani, and Jeffrey D. Ullman (2007). *Introduction to Automata Theory, Languages, and Computation*. 3rd edn. Boston: Pearson Addison-Wesley.
- Horning, James Jay (1969). *A Study of Grammatical Inference*. PhD thesis, Stanford University.
- Huang, Wanhong (2026). *A Metamodel for Ontologically Heterogeneous Social Dynamics — Part II: Grammar Rewriting*. Draft preprint.
- Lakatos, Imre (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 91–196.
- Lampland, Martha, and Susan Leigh Star (eds.) (2009). *Standards and Their Stories*. Ithaca: Cornell University Press.
- Latour, Bruno (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Laudan, Larry (1977). *Progress and Its Problems: Towards a Theory of Scientific Growth*. Berkeley: University of California Press.
- Lehman, Joel, and Kenneth O. Stanley (2011). Abandoning Objectives: Evolution through the Search for Novelty Alone. *Evolutionary Computation*, 19(2), 189–223.
- Lind, Douglas, and Brian Marcus (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge: Cambridge University Press.
- Mahoney, James, and Dietrich Rueschemeyer (eds.) (2003). *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.
- Mouret, Jean-Baptiste, and Jeff Clune (2015). Illuminating Search Spaces by Mapping Elites. arXiv:1504.04909.
- Nevill-Manning, Craig G., and Ian H. Witten (1997). Identifying Hierarchical Structure in Sequences: A Linear-Time Algorithm. *Journal of Artificial Intelligence Research*, 7, 67–82.

- Park, Joon Sung, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein (2023). Generative Agents: Interactive Simulacra of Human Behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. New York: ACM.
- Pierson, Paul (2004). *Politics in Time: History, Institutions, and Social Analysis*. Princeton: Princeton University Press.
- Ragin, Charles C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ragin, Charles C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.
- Rissanen, Jorma (1978). Modeling by Shortest Data Description. *Automatica*, 14(5), 465–471.
- Sartori, Giovanni (1970). Concept Misformation in Comparative Politics. *American Political Science Review*, 64(4), 1033–1053.
- Stolcke, Andreas, and Stephen Omohundro (1994). Inducing Probabilistic Grammars by Bayesian Model Merging. In R. C. Carrasco and J. Oncina (eds.), *Grammatical Inference and Applications (ICGI-94)*. Lecture Notes in Artificial Intelligence 862. Berlin: Springer, 106–118.
- Timmermans, Stefan, and Marc Berg (2003). *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care*. Philadelphia: Temple University Press.