

**Lausanne Conference on Sequence Analysis (LaCOSA), June 6<sup>th</sup>-8<sup>th</sup> 2012**

**Introduction to the conference:  
Sequence Analysis in 2012**

Philippe Blanchard, Felix Buhlmann and Jacques-Antoine Gauthier (alphabetical)

University of Lausanne

Updated June 3<sup>rd</sup>, 2012

The aim of this document is to nurture and to guide the debates during the Lausanne Conference on Sequence Analysis (LaCOSA). From this perspective, we will first try to summarise the main epistemological principles of sequence analysis (SA) as developed by Andrew Abbott and others. This includes the ontology of sequences of states considered *narratives*, the complexity of sequential models, and the main properties of narratives. Second, we will trace the main developments in SA over the last twenty years. We argue that due to the specific structure of the scientific field of SA we can distinguish a relatively consensual core program and a series of innovative but fragile alternatives (in terms of data, methods for dissimilarity calculation, and associated methods). Third, using, among others, the papers and abstracts of the participants, we point to a series of issues that are particularly debated at the moment. These include issues of sequence generation as well as questions of context and visualisation. Finally, we would like to formulate some provocative questions raised by the gap between the promises and the current achievement of SA.

## **1. Theoretical and Epistemological Issues**

### **1.1. “Social reality happens in stories”**

The centrality of order and processes in social science theory makes Abbott (1983) claim that “social reality happens in stories”, which he calls *narratives*. Becker (1963) also pleads for sequential models of social reality (for example, deviance). He insists that explanatory factors

(causes) are not simultaneously influential. A social process is divided into phases that have their specific explanations. The final behavior (or state of affair) can only be explained by the chronological succession of these phases. From this perspective, the various elements of such narratives are bound not only to immediate past, but also to remote past.

But considering social processes as “stories” does not mean that they are free of constraints and that, for instance, individuals orient their lives wherever and whenever they want, as some post-modernists’ stances may support. This also does not mean that social processes are somehow predetermined. They unfold in a web of constraints and opportunities.

Following the work of Levy and collaborators aiming at offering an interdisciplinary perspective on the life course, most narratives may be formally described using four elements: **trajectory**, **stage**, **event**, and **transition** (Levy et al., 2005). Let us give a short definition of these terms. A **trajectory** may be seen as a “model of stability and long-term change” or as a “sequence of profiles of insertion” that represents the variation over time of one or more roles an individual holds in various life spheres. It is used to describe the movements or developments occurring during the whole life span. **Stage** refers to a life period of relative structural or functional, for example, stability (family life stages, child development stages, periods of economic growth). **Transition** refers the idea of change at the end of a stage (for example, role and structural change during the transition to parenthood or during a political revolution). An **event** is what occurs at a given time in a given place and that is provided with meaning by human beings. They can be normative (birth, marriage, graduation) or non-normative (disease, war, death). Important events are often associated with a transition.

Sequences outside the life course paradigm, such as rituals, interpersonal interactions, oral and written texts, and historical evolutions of regimes or groups, may involve alternative units of analysis or time scales. They may also develop within a symbolic time differing from standard clock time, yet they are social processes as well. They form *trajectories* and under certain conditions go through *stages*, *events*, and *transitions*.

## 1.2. Ontology of narratives

The term *ontology* refers to the most general feature of being, what makes it be what it is. In this sense, narratives are structured patterns that have a beginning, a middle, and an end. Given the huge number of possible expressions of a specific social phenomenon, looking for

ideal types is a convenient solution to apprehend it as a narrative. Indeed, social facts (think, for instance, of transition to adulthood or political commitment) contain typical events that unfold in a typical order. This order is influenced by factors such as personality traits, family configuration, type of occupation, health status, or institutional power relationships. The ontology of a narrative is therefore, in most cases, both endogenous and exogenous. The relative weight of these two structuring factors may vary according to which social phenomenon is considered (for example, biological, psychological, sociological, historical or political). To be scientifically relevant, however, narratives must have a unity and a coherence that is analogous to that of variables (cf. standardisation, professionalization). This unity of narratives forms the basis for generalization. This is precisely why, once generalized, a narrative may become a variable again (for an empirical example, see Levy, Gauthier, & Widmer, 2006). Narratives form coherent units because the events that build them concern a **central subject** (an individual, a couple, an occupation, an organization, or a corpus of texts). Another way to describe the exogenous and endogenous structure of narratives is given by David Hull (1989). For Hull, a central subject must be seen on the one hand as a linkage between the events entering the narrative (internal structuring dynamic). On the other hand, one has to consider the part/whole (that is, systemic) relationship that links the central subject to its surrounding (for example, historical, political, organisational) narratives (external structuring dynamic). For structuralists, the elements of a narrative are defined by their function in the narrative line, considered as a whole. These elements are therefore not context independent. Systemically taking the macro-micro link into account is therefore central to an understanding of social narratives.

### **1.3. Complexity of sequential models**

Schematically there are various levels of complexity for sequential models. We can differentiate with respect to the recurring or non-recurring nature of events and according to the number of dimensions taken into account. Abbott (1996) distinguished three levels of complexity:

- **Stage theories or natural histories** describe biographical or historical stages through which most of the units of observation seem to pass. They obey an internal-process logic, relatively independent of contextual influences (for example, child

developmental stages, compulsory education in highly standardized countries, legislative or public agenda processes).

- **Career theories** show more variability than natural histories. They are dependent on the context, more contingent, and also include recurrent events and states. They typically include trajectories of education, occupational careers, residential trajectories, or activist careers.
- **Interactional field theories:** At this most complex level, the conceptual model integrates a whole network of interdependent sequences belonging to various relevant dimensions that form a system (family, occupation, residence, health, welfare regime, political stability).

#### 1.4. Main properties of narratives

We may now summarize the main properties of narratives (Abbott, 2001): **Enchainment**, defined as the nature of the narrative link from one step to the next; this is the narrative analogue to causality. **Order** (the dimension on which Abbott mainly focuses) is the degree to which a social story presupposes a particular, exact order of events. The notion of **convergence** corresponds to the degree to which a social sequence approaches a steady state, for which non-narrative methods would then be appropriate. Institutionalization is a good example, as it tends to produce a foreseeable series of enchainment (for example, tripartition of the life course). It is interesting that the degree of convergence may also vary historically (see, for instance, the hypotheses of standardization vs. destandardization applied to the institutionalisation of the life course or to the transition to parenthood). In short, social reality may be seen as a narrative order characterized by sequential dependency and coupled interdependencies.

## 2. The Main Developments of Sequence Analysis: The Core Program and Its Alternatives

### 2.1. The core program: How did we end up here?

The research trend in sequence analysis was structured in the 1990s–2000s around a core program composed of a reduced number of methodological options. These studies concern contemporaneous human units with yearly measurements and even lengths. The samples comprise hundreds to thousands individuals. The treatment of sequences runs from explorations to coding, optimal matching analysis (OMA), clustering, and description of clusters. A pairwise comparison of sequences is operated, based on the principles of aligning common patterns, comparing strings state by state, using even and fixed insertion-deletion costs (*icosts*) and “objective” or transition substitution costs (*scosts*). Finally, the Needleman and Wunsch (1970) recursive algorithm is used for the selection of the optimal sequence of elementary operations. Although not fully consensual, this core program provided the backbone of 80% to 90% of the approximately 120 papers making substantial use of SA published since the 1980s in the social sciences.

One may wonder how these standard options emerged. They result from a process of convergence and selection between research programs led by scholars from different disciplines and different countries. This process was not globally managed or coordinated by some central individual or team. The one who could have played the role of a conductor or a referee, Andrew Abbott, left the field at the end of the 1990s without formally proposing a best line to follow. His *Time Matters* (Abbott, 2001), in the shape of a testimony of his work on sequences, deals with the limits of standard methods used to treat longitudinal data in the social sciences and what objectives social sciences should seek to attain with sequences. Abbott neither proposed nor intended to propose a precise methodological guide. His expectation was that his own empirical experiments on sequences were limited, probably not optimal; and that diverse techniques and tools would burgeon, compete with each other, and eventually contribute to improving our understanding of this new object (Abbott, 2011).

Actually, the collective aggregation and convergence of efforts around a standard line resulted from a series of factors (Blanchard, forthcoming 2013):

1. the dominance of Abbott’s legacy from his methodological and empirical papers, still the most cited papers in the field;

2. a theoretical need for sequences in social statistics that had to be fulfilled empirically;
3. the strength of applications in sociology of work and their specific expectations in terms of sequences (work was Abbott's original affiliation, although not an exclusive one at all—see Abbott & Forrest, 1986; Abbott & Deviney 1992; and Abbott & Barman, 1997);
4. the availability of good-quality sequential datasets based on surveys, whether retrospective questionnaires or panels;
5. the modest number of software implementations, which gave a decisive advantage to the first to be developed for social sciences; and
6. the limited size of the community of sequence analysts facing a much larger population of sociologists using traditional methods, which compelled the former to stick to basics and promote a minimum core program so as to pass the journals' reviewers—chosen mainly from among the latter.

All these factors reinforce each other and explain the present state of the sequential art: a good cumulativeness, yet limited to some kind of data, research questions, research designs, and treatment strategies. Is this good or bad? Is it a matter of timing in a collective scientific strategy: Keeping a slow pace in order to look progressively more “normal” (Kuhn, 1962) and taking more risks at a later stage of development? Or does it come from a lack of imagination?

Moreover, in spite of this—implicit—consensus around a set of core options, no recipe book is available. Presently the beginner may refer to some parts of Abbott and his colleagues' introductory articles (Abbott 1995; Abbott & Tsay, 2000; Macindoe & Abbott, 2004); two rather short textbooks, one in German (Aisenbrey, 2000) and one in French (Robette, 2011), and a few papers partly following some pedagogic purpose (Lesnard, 2006; Blanchard, 2011). Courses within method schools in social and political science are proposed in German (2007), British (Essex, 2007; Manchester, 2009), French (Lille, 2009, 2010, 2011), Swiss (Geneva, 2009, 2012), Austrian (ECPR-Vienna, 2012), US (Columbia, 2011) and Canadian (Ontario 2012) universities. Yet no comprehensive textbook has been published, a criterion for setting up a paradigm, in Kuhn's (1962) view. Do specialists have the feeling that the method was not fully ready and that a textbook would be too quickly outdated? Exploring new lines in SA would look more fruitful than recording the standard state of the art for a shared community. Or does the absence of a transcription of the core program into a proper doctrine mean that this program is still weak and that its proponent dare not confront the critics?

## 2.2 Variants to the core program: *What have we done for him lately?*

Many variants and alternatives to the standard method may be imagined. Among these several have been proposed, many parts of which have also been tested empirically. The following is a short review of sequence data, methods for the calculation of dissimilarity between sequences and methods associated with the former.

First, data may diverge from the standard frame. They may relate to non contemporaneous individuals (for example, Dribe et al. LaCOSA 2012; Mercklé & Zalc LaCOSA 2012). They may not deal with human individual data (for example, Abbott & Barman, 1997; Stovel, 1997). They may use time units other than years or months (for example, Lesnard, 2008). They may not relate directly to standard clock time (for example, Abbott & Forrest, 1986; Buton et al., LaCOSA 2012; Buton et al., 2012). Data may also encompass small-N populations (for example, Abbott & Deviney, 1992), and sequence lengths may be very uneven (Fillieule & Blanchard, forthcoming 2012). Some of these alternative data bring new challenges to the standard approach. For example, sequences related to non-human objects (such as bills in Borghetto, LaCOSA 2012) radically change the perspective: instead of a person making decisions in successive contexts, they represent complex interactions between diverse actors through the actions they take over the objects, which modifies the underlying genesis of sequences. Other challenges are proposed: Non-standard time measurement creates specific event sequences requiring specific tools (Ritschard, LaCOSA 2012); small-N populations change the way the results are generalized to a larger population (if the sample is not the full population, as in Abbott & Deviney, 1992); and sequences of uneven lengths affect the interpretation of elementary operations applied by OM, especially in terms of balance between icosts and scosts.

A second field of variants comes with the way sequences are processed. It may take directions quite different from the one preferred today in the majority of studies. Much has been written about the possible variations regarding costs, including empirically machine-trained scosts (Gauthier et al. 2009), indel costs varying with states, scosts varying with data-based time (Lesnard 2010), or indel costs varying with the state's neighbourhood in the sequence (Hollister 2009), although no consensual doctrine has yet emerged on costs. Less attention has been paid to other variations. For example, instead of incremental, state-by-state comparison, as in optimal matching, sequence comparison may use holistic criteria based on the search for chunks of common subsequences (for example, longest common subsequence, longest

common prefix—see Elzinga, 2006, implemented in CHESA [Elzinga, 2007] and TraMineR [Gabadinho et al., 2011]). One may also compare without aligning, either by introducing alternative sets of elementary operations (swaps, discarding of isolated states, reduction of stable subsequences), by comparing pairs of ordered elements, or by applying correspondence analysis to the matrix of time spent by period in each state (for example, Deville & Saporta, 1980; Dijkstra & Taris, 1995; Elzinga, 2003; Lesnard, 2010). These different approaches rely on diverging conceptions of sequences and give privilege to different mathematical tools. As a result, they put the accent on distinct aspects of social time, as Robette and Bry (LaCOSA 2012) conclude in their synthetic comparison of sequence metrics over simulated and real datasets. The strategy of comparison itself may not rely on complete pairwise OM but on comparison with *ex ante* ideal-typical sequences (Wiggins et al., 2007) or on a time-varying OM-based assignment of individuals to clusters by means of year-by-year recalculation of dissimilarities (Wilson, LaCOSA 2012). The lengths of sequences may be standardized (Abbott & Hrycak, 1997; Halpin & Chan, 1998), and the dissimilarity matrix normalized (Han & Moen, 1999), in both cases with noticeable consequences on how to interpret the result. One may even imagine an optimal path inside the OM pairwise comparison array that would be different than simply the cheapest one resulting from the Needleman and Wunsch algorithm, which would lead again to a different pairwise dissimilarity structure.

Third, the choice and combination of methods associated with the dissimilarity calculation also vary notably. These complementary methods may take diverse roles relative to the SA core program, as described previously: They may be more or less autonomous from SA or reversely specific to SA; they may be used at different stages of the treatment process, either before or after the calculation of pairwise dissimilarities (in case any dissimilarity is calculated). For example, measures of the complexity of individual sequences or groups of sequences are often imported from other fields but adapted to social sequences (Elzinga, 2010). They can be used at the exploration stage, or after some OM-based clusters have been distinguished. Multidimensional scaling, not specific to sequences either, is useful to sort out both sequence populations and clusters (Piccarreta & Lior 2010). Validity and sensitivity analyses (for example, Hollister, 2009; Aisenbrey & Fasang, 2010) as well as simulation experiments (Bison, 2009; Halpin, LaCOSA 2012; Robette & Bry, LaCOSA 2012), at the end of the treatment process, can improve tremendously the robustness of OMA and SA.

Looking back, it appears that many of these alternative data and methods had been explored or at least suggested by Abbott and his colleagues' empirical (Abbott & Barman, 1997;

Abbott & Deviney, 1992; Abbott & Forrest, 1986; Abbott & Hrycak, 1990) and methodological (Abbott 1995, Abbott & Tsay, 2000) works. Although retired, the pioneer is still influential. The sequence community (if we may call it so) has been right to give more concrete and systematic insights into what was often left at the stage of intuitions by the pioneer. Yet does any global picture emerge from these variants? Do any federating concepts bring unity back to the picture?

### **3. Some Currently Debated Issues**

#### **3.1 Genesis and explanation of sequences**

According to Abbott, the core of SA is the identification of sequential patterns and not the genesis of these sequences. However, several participants of the conference seem to be interested in the pattern-generating questions and even think, as does Katherine Stovel (Stovel, LaCOSA 2012), that the two questions should not be separated. We may consider two types of sequence genesis:

- Sequences could be generated endogenously, that is, certain sub-sequences, stages, events, or specific episodes could have an influence on the further enchainment of the sequential elements. In this vein, Ritschard (LaCOSA 2012), for example, examines particularly discriminating events, Gabadinho (LaCOSA 2012) proposes an extension of Markov modeling and Han (LaCOSA 2012) studies the potential for motif as both a thematically unifying and formally salient part of sequences.
- When it comes to exogenous mechanisms for generating sequence patterns, we might think of structural/objective and individual/subjective effects. Structural constraints can be the organisations in which occupational career occurs, the welfare regime in which one goes through the transition from school-to-work (Salmelo-Aro et al., 2011) or from work to retirement (Barban et al. LaCOSA 2012) or legal framework of a legislative processes (Borgettho LaCOSA 2012). These constraints then interact with features of the individuals under study, such as socio-demographic variables (gender, education, and ethnicity) in life-course studies—variables that seem often to generate specific forms of sequences.

A method that does not primarily use a typology approach but is able to relate such variables with certain sequential patterns is the recently developed discrepancy analysis (Studer et al. 2011). Are such new approaches important steps for a closer articulation of the “pattern question” and the “generating question”? Or are they a step back towards “variable sociology” as has been opposed by Abbott? Still other scholars, especially in a life-course or political sociology framework, think that it is necessary to include subjective interpretations and individual decisions as important factors that generate typical sequential patterns. Even though Abbott never lacked an opportunity to affirm the continuity of his thinking with the Chicago School of Sociology, it is conspicuous that the “subject aspect”—central, for example, in the works of Hughes or Becker—seldom plays a role in the “narrative positivism” approach, as its founder called it. How could these aspects of agency be integrated into SA? Is it, for example, possible to depict sequences of subjective/psychological states and to link them to other trajectories? What about the combination of SA with qualitative interviews in the framework of mixed-mode designs?

### **3.2 Context of sequences**

One of the reasons that brought Abbott to think about the importation of SA to social sciences was the lack of contextuality of variable-based sociology. Causal linear reality, such as the one underlying standard regression models because of its reliance upon variables, has decontextualized sociological analysis spatially, temporally, and even sociologically. As an answer, Abbott (1996) distinguishes three analytical levels—stage theories, career theories, and theories of interactional fields—with different degrees of spatial and temporal embeddedness. Whereas stage theories are temporally contextualised, they are not very open to structural and spatial influences. Career theories are more contingent and theories of interactional fields are both spatially and temporally embedded. How does SA fit into this picture? Whereas it is certainly very well embedded in temporal terms, it struggles to link individual biographies to their structural and spatial context, for instance, by relating it to other trajectories, networks, space and social structure. However, several attempts have been made to develop the method in this direction: multi-channel analysis (Bühlmann, 2008; Gauthier, 2010; Pollock, 2007) in certain cases connects a trajectory to trajectories of other people and can—in the sense of Elder’s “linked lives” (1985)—help to embed it in a wider array of trajectories. It is also possible to combine, for instance, trajectories of occupational

positions, trajectories along types of firms, and trajectories of functions within the firm. Another, more direct strategy to contextualise trajectories spatially is to study residential trajectories (Stovel & Bolan, 2004). Still other scholars try to combine SA with network analysis (Bison, LaCOSA 2012) or have invented measures of turbulence/entropy that allow for an aggregate characterisation of group trajectories (Elzinga, 2010; Gabadinho et al, 2012, implemented in TraMineR), and so forth. Others propose to introduce the context to sequence individual plots as events that “align” sequences (Blanchard, 2010; Colombi & Paye, LaCOSA 2012), to study paired or grouped sequences (couples in Lesnard, 2008; households in Buton et al., 2011), or to show how distinct the biographies of successive cohorts of activists are (Fillieule & Blanchard, forthcoming 2012). What could be further developments along these paths? Is it, for example, possible to compare the trajectories of different cohorts and to relate historical changes to changes in trajectories?

### **3.3 Visualisation of sequences**

Visualisation is another topic that is both old and still debated. Synoptic capacity of visualization has an enormous impact on the diffusion of research findings, especially if these findings are based on innovative analytical approaches (Tufte, 1997). Second, as stated above, we do not have a consensus on how to best represent sequences. Third, the fact that an increasing number of large longitudinal data sets are now available to researchers in social sciences may boost the use of optimal-matching analysis and therefore the need for adequate and effective ways to think graphically about sequences and communicate the results. Thus, there is a need to evaluate the possible mapping strategies that may apply to results stemming from OMA, their respective strengths and weaknesses, and their specific application range.

When considering the mapping of a given type of trajectory on OMA (or other approaches), at least three semantic sources appear simultaneously. The first is reading across the sequence, which informs on individual variations over time (for example, which social status follows which one, and at what pace). The second is transversal reading (reading down, or up, see Piccarreta & Lior 2010]); that is, the inter-individual variation at fixed points in time. The third is the image produced by the alignment. It represents a global meaning for the entire population or a significant subgroup.

In comparison with the methodological and theoretical literature available on SA in the social sciences, few attempts have been made to address the way trajectories considered in light of empirical categorization should be best represented (although, see Gauthier, 2007 ; Piccareta and Lior 2010; Billari 2001). One of the essential features of the way pairwise as well as multiple SA is performed is the fact that two or more sequences are aligned in a way that minimizes the transformation cost needed to superimpose one entire sequence on the other. At a higher level of generalization, the same happens, but less formally, when comparing different types of occupational trajectories, and even more when considering typical graphs stemming from multichannel SA.

## **4. Some Provocative Theses on the Promises of Sequence Analysis**

### **4.1 Sequence analysis does not account for the richness of life-course sociology**

Users of SA claim that the method is particularly suited to understanding biographical phenomena. Usually SA is then compared to event history analyses and praised as the method that is able to overcome a restriction to transitions and investigate “whole trajectories” (Abbott, 1995; Aisenbrey & Fasang, 2010). However, life-course sociology cannot be reduced to the opposition between transition and trajectory. Life-course sociology is considerably richer than that, and it is in many cases doubtful if SA in its current state is well adapted for most of the theoretical principles of life-course sociology. For instance, one of the important developments of recent life-course research is theories of cumulative advantage (Dannefer, 2003; DiPrete & Eirich, 2006). These theories state that inequalities are created over the life course according to different mechanisms of simple accumulation, path dependency, or cumulative exposure. The same may fit sequential phenomena beyond life-course studies, such as accumulation of organizational resources for leading activists (Blanchard, 2011) or reputational resources for successful bills (Borghetto, LaCOSA 2012). SA struggles to engage with these kinds of theories because it is based on notions of capital (resources, assets) and not on stages. How could SA engage with this theory, which also has important bridges to other important theoretical frameworks of the social sciences (such as Bourdieu’s theory, see Pallas refs)? Is it possible, for example, to integrate the notion of asset/capital and its different modes of accumulation into the framework of SA?

## **4.2 Sequence analysis has not managed to lodge itself in substantial debates**

Methodologically, as Aisenbrey and Fasang (2010) have shown, the second wave of SA has made noteworthy advances. However, it sometimes confines itself to methodological debate and has not really managed to thoroughly reorient substantial theoretical debates. Ironically, the substantial area where SA has been most successful is transitions; namely, the school-to-work transition (Brzsinsky-Fay, 2007; Billari, 2001; Salmelo-Aro et al., 2011; Roberts & Pollock, 2009; Elzinga & Liefbroer, 2007; Robette, 2010; Lesnard et al., 2010) and the transition from work to retirement (Fasang, 2009; Barban et al., LaCOSA 2012) have been extensively covered. However, many other substantial debates (on social mobility, family trajectories, political careers, etc.) that imply longer trajectories would benefit from a concerted intrusion of SA. A multiplication of substantial contributions making use of SA would contribute to legitimatising the method and building its reputation. In other words, the reason SA has partially failed to redeem its promises (for example, such as formulated by Abbott in 1995) is probably not merely “technical”. It is perhaps also due to a lack of coordination and common strategy among SA scholars. They are united through a common interest in the method but have failed to build a joint substantial-methodological framework, such as life-course and event history analysis.

LaCOSA intends to contribute to filling in this gap by bringing together scholars from different places, different disciplines, and different statistical cultures. Some of them are part the life-course research trend and, as such, they have already been encouraged to use SA. Some came across one of Abbott’s pioneer papers or some stimulating application of SA to their own domains. Others were attracted by the specifically methodological challenge of a new and promising method. All of them should benefit from a common effort to set up a more unified intellectual framework.

## References

- Abbott, A. (1983). Sequences of social events: Concepts and methods for the analysis of order in social processes. *Historical Methods*, 4, 129–147.
- Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, 21, 93–113.
- Abbott, A. (2001). *Time Matters: On Theory and Method*. Chicago, IL: University of Chicago Press.
- Abbott, A., & DeViney, S. (1992). The Welfare State as Transnational Event: Evidence from Sequences of Policy Adoption." *Social Science History*, 16(2), 245–74.
- Abbott, A., & Hrycak, A. (1990). Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers. *American Journal of Sociology*, 96, 144–85.
- Abbott, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology. Review and Prospect. *Sociological Methods and Research*, 29 (1), 3–33.
- Aisenbrey, S. (2000). *Optimal Matching Analyse*. Opladen, Germany: Leske-Budrich.
- Aisenbrey, S., & Fasang, A. (2010). New Life for Old Ideas: The 'Second Wave' of Sequence Analysis Bringing the 'Course' Back Into the Life Course. *Sociological Methods and Research*, 38(3), 420–462.
- Barban N., Svensson I., Billari F., de Luna X. & Lundholm E. Matching lifecourses for causal inference: The effect of early retirement on health using (Swedish) register data. Lausanne: LaCOSA.
- Becker, H. S. (1966). *Outsiders: Studies in the Sociology of Deviance*. New York: Simon and Schuster.
- Billari, F. C. (2001). The Analysis of Early Life Courses: Complex Description of the Transition to Adulthood. *Journal of Population Research*, 18(2), 119–42.
- Billari, F. C., & Piccarreta, R. (2005). Analyzing Demographic Life Courses through Sequence Analysis. *Mathematical Population Studies*, 12, 81–106.
- Bison, I. (2009). OM matters: the interaction effects between indel and substitution costs. *Methodological Innovations Online*, 4(2), 53–67.
- Bison, I. (2012). Sequence analysis and Network analysis: a tentative to represent and study sequences by using NetDraw. Lausanne: LaCOSA.
- Borghetto, E. Legislative processes as sequences: exploring the temporal dimension of law-making by means of sequence analysis. Lausanne: LaCOSA.

- Brzinsky-Fay, C. (2007). Lost in Transition? Labour Market Entry Sequences of School Leavers in Europe.” *European Sociological Review*, 23 (4), 409–22.
- Blanchard, P. (2010). *Analyse séquentielle des carrières militantes à Aides* [Sequence Analysis of Militant Careers in Aides]. HAL-Open Archives. Retrieved from <http://hal.archives-ouvertes.fr/hal-00476193/fr>
- Blanchard, P. (2011). Sequence Analysis for Political Science”. *Working Papers of the Committee on Concepts and Methods, International Political Science Association*. Retrieved from <http://www.concepts-methods.org/WorkingPapers/PDF/1082>
- Blanchard, P. (Forthcoming, 2013). “Validity, falsifiability, parsimony, consistency, precision, and so on”: Les vicissitudes de l’innovation méthodologique [“Validity, falsifiability, parsimony, consistency, precision, and so on”: The Difficulties of Methodological Innovation]. In M. Jouvenet & D. Demazière (Eds.). *La sociologie d’Andrew Abbott*. Paris: EHESS.
- Bühlmann, F. (2008). The Corrosion of Career?—Occupational Trajectories of Business Economists and Engineers in Switzerland. *European sociological Review*, 24(5): 601-616.
- Bühlmann, Felix. 2010. Routes into the British Service Class. Feeder Logics According to Gender and Occupational Groups. *Sociology*, 44(2): 195-212.
- Borghetto, E. (2012). “Legislative processes as sequences: Exploring the temporal dimension of law-making by means of sequence analysis”, Lausanne: LaCOSA.
- Buton, F., Lemerrier, C., & Mariot, N. (Forthcoming, 2012). The Household Effect on Electoral Participation. A Contextual Analysis of Voter Signatures from a French Polling Station (1982–2007). *Electoral Studies*. Preprint doi:10.1016/j.electstud.2011.11.010
- Casper, G., & Wilson, M. (2011). Bargaining within and across Crises. Seattle, WA: American Political Science Association Conference.
- Colombi D., & Paye S. (2012). Synchronising sequence data to idiosyncratic events: a method to explore relationships between events and structures. Lausanne: LaCOSA.
- Dannefer, D. (2003). Cumulative Advantage/Disadvantage and the Life Course: Cross-Fertilizing Age and Social Science Theory, *Journals of Gerontology*, 58(6), 327–37.
- Deville, J. C., & Saporta, G. (1980). Analyse harmonique qualitative, [Qualitative Harmonic Analysis], in E. Diday (Ed.), *Data Analysis and Informatics* (pp. 375–89). Amsterdam: North Holland Publishing.

- Dijkstra, W., & Taris, T. (1995). Measuring the Agreement between Sequences. *Sociological Methods and Research*, 24, 214–31.
- DiPrete, T., & Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. *Annual Review of Sociology*, 32, 271–97.
- Dribe, Martin, Manfredini Matteo, Oris Michel & Ritschard Gilbert. (2012). Access to Reproduction in Pre-transitional Europe. A Sequential Approach. Lausanne: LaCOSA.
- Durbin, R., Eddy, S., Krogh, A., & Mitchison, G. (1998). *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge: Cambridge University Press.
- Elder, G. H. (1985). Perspectives on the Life Course. In Glen H. Elder Jr. (Ed.), *Life course dynamics, trajectories and transitions* (p. 23–49). Ithaca NY: Cornell University Press.
- Elder, G. H. (1999). *Children of the Great Depression: Social Change in Life Experience*. Boulder, CO: Westview Press.
- Elzinga, C. H. (2003). Sequence Similarity: A Nonaligning Technique. *Sociological Methods and Research*, 32, 3–29.
- Elzinga, C. H. (2006). *Sequence Analysis: Metric Representations of Categorical Time Series*. Amsterdam, NL: Department of Social Science Research Methods, Vrije Universiteit.
- Elzinga, C. (2007). *CHESA 2.1 User Manual*. Amsterdam: Frije Universiteit.
- Elzinga, C. H. (2010). Complexity of Categorical Time Series, *Sociological Methods & Research*, 38, 463–81.
- Elzinga, C. H., & Liefbroer, A.C. (2007). De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis'', *European Journal of Population*, 23, 225–50.
- Fasang, A. (2012). Retirement Pattern and Income Inequality, *Social Forces*. Preprint doi: 10.1093/sf/sor015
- Fillieule, O., & Blanchard, P. (Forthcoming, 2012). Fighting Together. Assessing Continuity and Change in Social Movement Organizations Through the Study of Constituencies' Heterogeneity. In Niilo Kauppi (Ed.), *The New European Political Sociology*. London: Routledge.
- Gabadinho, A. (2012). Sequence analysis with Variable Length Markov Chains. Lausanne: LaCOSA.

- Gabadinho, A., Ritschard, G., Studer, M., & Müller, N. (2011). *Mining sequence Data in R with the TraMineR Package: A User's Guide*. Geneva, CH: Department of Econometrics and Laboratory of Demography, University of Geneva.
- Gauthier, J. (2007). *Empirical categorizations of social trajectories: A sequential view on the life course*. Lausanne, CH: University of Lausanne.
- Gauthier, J., Widmer, E., Bucher, P., and Notredame, C. (2009). How Much Does It Cost? Optimization of Costs in Sequence Analysis of Social Science Data.” *Sociological Methods and Research*, 38(1), 197–231.
- Gauthier, J., Widmer, E., Bucher, P., & Notredame, C. (2010). Multichannel Sequence Analysis Applied to Social Science Data”. *Sociological Methodology*, 40, 1–38.
- Halpin, B. (2012). Multiple Imputation for Life-Course Sequence Data. Lausanne: LaCOSA.
- Halpin, B., & Chan, T. W. (1998). Class Careers as Sequences: An Optimal Matching Analysis of Work-Life. *European Sociological Review*, 14(2), 111–30.
- Han S. K. (2012). *Motif of Sequence, Motif in Sequence*. Lausanne: LaCOSA.
- Hollister, M. (2009). Is Optimal Matching Suboptimal? *Sociological Methods and Research*, 38, 235–64.
- Hull, D. L. (1989). *The Metaphysics of Evolution*. NY: SUNY Press.
- Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lemercier, C. (2005). Les carrières des membres des institutions consulaires parisiennes au XIXe siècle [The Careers of Economic Elites in Paris in the 19th Century]. *Histoire et Mesure*, XX(1/2), 59–95.
- Lesnard, L. (2006). Optimal Matching and Social Sciences. *CREST Working Papers 2006-01, INSEE: Paris*.
- Lesnard, L. (2008). Off-Scheduling within Dual-Earner Couples: An Unequal and Negative Externality for Family Time.” *The American Journal of Sociology*, 114(2), 447.
- Lesnard L. (2010). Setting Cost in Optimal Matching to Uncover Contemporaneous Socio-Temporal Patterns”, *Sociological Methods and Research*, 38, 389–419.
- Lesnard, L., Cousteaux, Chanvril, F., & Le Hay, V. (2010). Do Transitions to Adulthood Converge in Europe? An Optimal Matching Analysis of Work-Family Trajectories of Young Adults from 20 European Countries. *Notes & Documents, 2010-04/ Paris: OSC, Sciences Po/CNRS*.
- Levy, R., Gauthier, J. A., & Widmer, E. (2006). Entre contraintes institutionnelle et domestique: les parcours de vie masculins et féminins en Suisse [Institutional and

- Domestic Constraints: Male and Female Life Course in Switzerland]. *The Canadian Journal of Sociology*, 31(4), 461–89.
- Levy, R., Ghisletta, P., Le Goff, J. M., Spini, D., & Widmer, E. (2005). Towards an interdisciplinary perspective on the life course. *Advances in life course research*. Amsterdam: Elsevier JAI.
- Macindoe, H., & Abbott, A. (2004). Sequence Analysis and Optimal Matching Techniques for Social Science Data 387–406). In M. Hardy and A. Bryman (Eds.). *Handbook of Data Analysis*. Thousand Oaks, CA: Sage.
- Mercklé Pierre & Zalc Claire. (2012). Trajectories of persecuted during World War 2 : contributions of the microhistory of the holocaust. Lausanne: LaCOSA.
- Needleman, S., & Wunsch, C. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48, 443–53.
- Notredame, C., Higgins, D., & Heringa, J. (2000). T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment. *Journal of Molecular Biology* 8, 302(1), 205-17.
- Piccarreta, R., & Lior, O. (2010). Exploring sequences: a graphical tool based on multi-dimensional scaling. *Journal of the Royal Statistical Society Series A* 173(part 1), 165–84.
- Pollien, A., & Joye, D. (2011). A la poursuite du répondant? Essai de typologie des séquences de contact dans les enquêtes [Searching For the Respondent ? Tentative Typology of Contact Sequences in Surveys] (pp. 189-212). In D. Joye, C. Pirinoli, D. Spini, & E. Widmer (Eds.), *Parcours de vie et insertions sociales*. Zurich: Seismo.
- Pollien, A., & Joye, D. (2012). *Pattern of contacts attempts in surveys*, Lausanne: LaCOSA.
- Pollock, G. (2007). Holistic trajectories: A study of combined employment, housing and family careers by using multiple-sequence analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170, 167–83.
- Pollock, G. (2007). Employment trajectories in the South Caucasus during the transition from communism to post-communism: a comparison of OMA and non-sequence based methods of producing typologies, Lausanne: LaCOSA.
- Ritschard, G. (2012). Exploring the sequencing and timing of life events, Lausanne: LaCOSA.
- Roberts, K., Pollock, G., Manasyan, H., & Tholen, J. (2008). School-to-Work Transitions after Two Decades of Post-Communist Transition: What’s New? *Eurasian Journal of Business and Economics*, 1 (2), 103–29.

- Robette, N. (2011). *Explorer et décrire les parcours de vie : les typologies de trajectoires*, [Exploring and Describing the Life Course : Typologies of Trajectories]. Paris: CEPED.
- Robette, N., Thibault, N., & Dutreuilh, C. (2008). Comparing qualitative harmonic analysis and optimal matching. An exploratory study of occupational trajectories. *Population-E* 63(4), 533–56.
- Robette N. (2010). The diversity of pathways to adulthood in France: Evidence from a holistic approach. *Advances in Life Course Research*, 15, 89–96.
- Robette, N., & Bry Xavier. (2012). Harpoon or maggot ? A comparison of various metrics to fish for sequence patterns. Lausanne: LaCOSA.
- Salmela-Aro, K., Kiuru, N., Nurmi, J., & Eerola, M. (2011). Mapping pathways to adulthood among Finnish university students: Sequences, patterns, variations in family- and work-related roles. *Advances in Life Course Research*, 16, 25–41.
- Stovel K. (2001). Local Sequential Patterns: The Structure of Lynching in the Deep South, 1882-1930. *Social Forces*, 79, 843-880.
- Stovel K. (2012). The Nature of Sequences. Lausanne: LaCOSA.
- Stovel, K., & Bolan, M. (2004). Residential Trajectories: The Use of Sequence Analysis in the Study of Residential Mobility. *Sociological Methods and Research*, 32, 559–98.
- Studer, M., Ritschard, G., Gabadinho, A., & Müller, N. (2011). Discrepancy Analysis of State Sequences. *Sociological Methods and Research*, 40, 471–510.
- Wiggins, R., Erzberger, C., Hyde, M., Higgs, P., & Blane, D. (2007). Optimal Matching Analysis Using Ideal Types to Describe the Lifecourse: An Illustration of How Histories of Work, Partnerships and Housing Relate to Quality of Life in Early Old Age. *International Journal of Social Research Methodology*, 10(4), 259–78.
- Wilson, M. (2012). Temporal Determinants of Democracy: Beyond Duration. Lausanne: LaCOSA.
- Wilson, W. C. (1998). Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A*, 30, 1017–38.