

Conflict and Computation on Wikipedia: A Finite-State Machine Analysis of Editor Interactions

Simon DeDeo

SFI WORKING PAPER: 2016-06-012

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu



SANTA FE INSTITUTE

Article

Conflict and Computation on Wikipedia: a Finite-State Machine Analysis of Editor Interactions

Simon DeDeo^{1,2,3,4*}

¹ Center for Complex Networks and Systems Research, Department of Informatics, Indiana University, 919 E 10th St, Bloomington, IN 47408, USA

² Program in Cognitive Science, Indiana University, 1900 E 10th St, Bloomington, IN 47406, USA

³ Ostrom Workshop in Political Theory and Policy Analysis, 513 N Park Avenue, Bloomington, IN 47408, USA

⁴ Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

* Author to whom correspondence should be addressed; simon@santafe.edu, telephone (812) 856-2855, fax (812) 856-3825

Version June 27, 2016. Typeset by \LaTeX using class file *mdpi.cls*

Abstract: What is the boundary between a vigorous argument and a breakdown of relations? What drives a group of individuals across it? Taking Wikipedia as a test case, we use a hidden Markov model to approximate the computational structure and social grammar of more than a decade of cooperation and conflict among its editors. Across a wide range of pages, we discover a bursty war/peace structure where the systems can become trapped—sometimes for months—in a computational subspace associated with significantly higher levels of conflict-tracking “revert” actions. Distinct patterns of behavior characterize the lower-conflict subspace, including tit-for-tat reversion. While a fraction of the transitions between these subspaces are associated with top-down actions taken by administrators, the effects are weak. Surprisingly, we find no statistical signal that transitions are associated with the appearance of particularly anti-social users, and only weak association with significant news events outside the system. These findings are consistent with transitions being driven by decentralized processes with no clear locus of control. Models of belief revision in the presence of a common resource for information-sharing predict the existence of two distinct phases: a disordered high-conflict phase, and a frozen phase with spontaneously-broken symmetry. The bistability we observe empirically may be a consequence of editor turn-over that drives the system to a critical point between them.

Keywords: conflict; cooperation; finite-state machine; social grammar; tit-for-tat; critical transition; hidden Markov model; relaxation time; memory; social norms; knowledge commons; Wikipedia

Many societies are characterized by periods of cooperation separated by significant periods of intra-group conflict. Both cooperation and conflict are themselves complex phenomena: cooperation is not the absence of conflict, nor is the presence of conflict incompatible with cooperation. One can talk of a lull in fighting, for example, or, conversely, of antagonistic interactions that are part of the internal logic of a resilient and cooperative society. Two questions, then suggest themselves: can we define the boundary between cooperation and conflict, and can transitions across this boundary be explained, controlled, or predicted before they happen?

We approach this pair of fundamental questions by examining a social system in terms of how it processes information. As we shall show, a system's computational process—how it moves through an abstract state space that both responds to past antagonism and predicts its future development—connects directly to large-scale features that are immediately perceptible to human participants.

Our system of study is social behavior observed in the debates on the website Wikipedia. Wikipedia is an ideal case study for the computational properties of conflict. By design, actions on the system are logged at extremely high resolution, allowing us to track the actions and interactions of tens of thousands of users second-by-second for nearly fifteen years. A natural binary classification of actions (“revert” and “non-revert”), that tracks conflict between users, allows us to use sequence analysis and the language of symbolic timeseries [1,2] to operationalize the boundary between conflict and cooperation with Hidden Markov Models (HMMs).

Using HMMs to detect hidden structure in social behavior is a natural extension to how they have been used in the engineering, linguistic, and biological sciences. HMMs are in wide use when we believe that a noisy observable signal tracks some more complex underlying process, such as when the garbled sound of a noisy room includes someone speaking a sentence (speech recognition [3,4]), the individual words in a sentence give clues to that sentence's hidden syntactic structure (part of speech tagging [5,6]), or the A, C, T, and G symbols of DNA code for genes [7]. The volume of data on Wikipedia makes it possible to model the information processing of the system itself, an approach becoming increasingly common in large-scale studies of social media [8]. At the same time, our understanding of the products of this analysis are enhanced by ethnographic and quantitative studies of the system [9] that allow us to interpret individual actions, users, and their social contexts, and allow us to test narrative accounts of the causes and patterns of conflict.

Our analysis of Wikipedian conflict through the lens of the revert/non-revert distinction reveals two long-lived behavioral patterns associated with higher and lower levels of these conflict-tracking revert action. While the modes themselves are defined by reference to short sequences, or motifs, of interactions that can take only a few minutes to happen, the system as a whole will stay in one mode or another for weeks, and sometimes even months or years. This separation of timescales provides a bottom-up definition of the boundary between cooperation and conflict that depends on the existence of persistent, hidden system memories. In the first part of this paper, we demonstrate the existence of these unexpectedly long-term memories, and the distinct local patterns of interaction that they give rise to. Detected by

reference to apparently abstract properties of the system’s computational description using hidden Markov models, these memories are tied to human-scale and group level properties that form the basis of the system’s self-understanding.

In addition to providing a novel means for the description of the boundaries and distinctions between conflict and cooperation, our model allows us to pin-point the transitions between these two coarse-grained computational states. In the second part of our paper, we use this information to study the influence of administrative actions, individual users, and exogenous events over this society’s history. We find that top-down actions have only limited effectiveness in triggering a transition in either direction.

1. Methods

Our study here is based on the time-series of edits made on sixty-two of the most-edited pages on Wikipedia. This sample was chosen because each page then provided us with sufficient data that we could model each page independently; we thus were able to avoid making the (strong) assumption that editing practices were uniform. While our sample is by definition exceptional, and includes articles with tens of thousands of individual edits, it also covers a wide range of different topics and themes including (1) biographies of figures in both politics and entertainment, living and dead, from Genghis Khan to George W. Bush, Paul McCartney to Britney Spears (30% of sample); (2) other non-biographical pages associated with the arts and entertainment (16% of sample); (3) pages associated with countries, such as Cuba, Argentina, and the United States (18% of sample); (4) natural events ranging from the sinking of the Titanic to Hurricane Katrina and (5) political events such as World War II and the 2006 War in Lebanon (13% of sample); (6) technology topics such as the iPhone (9% of sample); and (7) religious topics such as the Catholic Church and Islam (6% of sample). See Appendix A; past work on Wikipedia editing practices have ranged from the histories of individual articles [10,11] to the full Wikipedia dataset [12]; ours is of a similar scale and selectivity to a number of other conflict studies [13–15], and the article list overlaps with two prior studies [1,16].

1.1. Tracking Conflict through Page Reverts

When editors interact on Wikipedia, they can disagree on what to do and how to behave. The resulting conflicts that emerge can be tracked by reference to “reverts”. Reverts are when one user takes a page to a previously-seen state, effectively undoing and discarding the work of others. For each page in our sample, the edit time-series was coarse-grained into “revert” (R) and non-revert (C) actions as described in Ref. [1]. We define a revert as an edit that takes the page to a previously-seen state (as detected by the MD5 page hash provided by the Wikipedia API), but exclude the relatively small number of cases where a user’s revert only undoes work contributed by that user herself (*i.e.*, we exclude self-reverts). More details on these methods, and additional robustness checks, of our methods are described in Ref. [16].

Reverts are noisy, imperfect signals of conflict; research has shown how more qualitative, interpretive, “thicker” [17] measures of conflict among users are strongly associated with these easily-tracked reverts [18], which can degenerate into edit wars [13]. Researchers have shown that reverts play a central role in marking both conflict and controversy [19–21], and that a great deal of information is encoded in reverts, including patterns of alliances among editors [10]. Reverts play a major role in the

social norms of Wikipedia itself [22], including highly central pages in the norm network such as the “three-revert rule” and strong norms against edit warring (see Ref. [16] for further discussion).

Reverts, however, are not equivalent to conflict. In general, reverts are best used to track the presence of task conflict [15], as opposed to other dimensions of conflict identified in the organizational behavior literature, such as affective conflict and process conflict [23,24]. While the task conflict represented by reverts is more ambiguous than other forms of conflict, it is still highly correlated with affective and relational conflict [25]; for example, the perception of rejection associated with being reverted is a significant factor in dissuading female contributors from participation [26]. At the same time, the tension between the need to change and to retain text can manifest itself in many different and more subtle ways [11,27,28]. There are multiple pathways for conflict to manifest itself that do not lead deterministically to reverts, an example being conflict associated with Wikipedia policies [12], and conflict that takes place in discussion on an article talk pages [29].

In short: while we rely on reverts in our analysis here, we emphasize that this is only one signal of conflict in Wikipedia, and that conflict itself is a far richer notion than the binary presence-absence data of our time series. The fact that reverts are only noisy, partial signals of conflict leads us to use a particular tool, the Hidden Markov Model. Hidden Markov Models are ideally suited to cases where a complex system leaves only a low-dimensional trace in an observable time-stream, allowing us to reconstruct, in part, some of the structure of the hidden, underlying process.

1.2. Hidden Markov Models

For each page, we have a binary time-series of Wikipedian conflict, consisting of Rs and Cs. Page-by-page, we approximate this time-series as a probabilistic finite state machine, more commonly known as a Hidden Markov Model (HMM, or machine). In an HMM, the observed behavior of the system—here, the registering of a new edit of type R or C—is (probabilistically) conditioned on the system’s position in a hidden, internal state space.

Our use of the HMM is conceptually simple. Each page is associated with a distinct machine. The evolution of the page is represented by the machine as it moves from one hidden, internal state to another from edit to edit. At each edit, and depending on the state the machine finds itself in, it will produce an observed edit behavior (probabilistically; here, R or C), and transition to a new internal state. The chance of going from one hidden state to another is dictated by a fixed probability. The origins of the HMM paradigm go back to Ref. [30]; for a recent, technical introduction to Hidden Markov Models, including the Baum-Welch algorithm used in this paper, see Ref. [31]; Fig. 1 provides an example from the *George_W._Bush* page.

In contrast to many simple time-series models, a Hidden Markov Model does not make a “no history” assumption: instead, the system memory is encoded by the position within the hidden state space. From the point of view of an external observer, this memory can extend arbitrarily far into the past. Strong theorems exist about the way in which this memory must, eventually, decay, but there is no “hard” cutoff where (for example) an observed event t steps in the past has zero influence when t is greater than some critical, finite T and one includes the more recent past as a predictor. For these reasons, an HMM is more general than the standard auto-regressive (AR [32]) models familiar from econometrics. AR models

condition the current state of the system on a fixed and finite number of steps of the observable sequence in the recent past; usually less than ten. By contrast, an HMM can encode arbitrarily long path dependencies; as we shall see, the correlation lengths our methods do detect in Wikipedia suggest that AR models would be very poor descriptions of the data indeed.

1.3. Fitting and Characterizing HMMs

We find the parameters for the HMM associated with the time-series on each page using the Expectation-Maximization algorithm (EM [30,31]); for each page, we run this hill-climbing algorithm 3,200 times with random initial seeds to study the ruggedness of the underlying likelihood landscape.

The one externally-fixed parameter in an HMM is the total number of hidden states; once this choice is made, the other parameters are determined by the data. Increasing the number of states means that the model is better able to predict the data. Using too many states, however, can mean that the model structure includes features that are due to random, one-off fluctuations in the data; this is called over-fitting. We use the Akaike Information Criterion (AIC [33]) to select the number of states for each page's model. Extensive tests, and comparison to the main competing model selection method, BIC [34], suggest that AIC is, for the systems we encounter here, preferred; see Appendix B for further discussion. For simplicity, and in keeping with usual practice, we work solely with the maximum-likelihood model rather than attempting to model the full distribution over distinct, but less-preferred, models.

1.4. Subspaces, Trapping Time, and Viterbi Reconstruction for HMMs

In general, as a system evolves in the HMM paradigm, it will spread out over its internal state space. A characteristic feature of this spreading out is the so-called relaxation time, which determines how long the system stays confined in a particular subspace of its possible internal states before returning to its characteristic stationary distribution. Once the best-fit HMM has been found, we can compute the relaxation time, τ , by reference to the second eigenvalue of the transition matrix, λ_2 ,

$$\tau = \frac{1}{1 - \lambda_2}. \quad (1)$$

Informally, and for the cases we consider here, when λ_2 is very close to unity (small spectral gap), the system can take a long time to return to the stationary distribution (see Appendix C for further discussion). This can be understood by writing down an approximation for the probability that the system will be found in a particular state at time t , $\vec{v}(t)$,

$$\begin{aligned} \vec{v}(t) &\propto v^{(1)} + \alpha v^{(2)} \lambda_2^t + \dots, \\ &\approx v^{(1)} + \alpha v^{(2)} e^{-t/(\tau - \frac{1}{2})} + \dots \end{aligned} \quad (2)$$

where $v^{(1)}$ is the stationary distribution of the chain, equal to the first eigenvector of the transition matrix, $v^{(2)}$ is the second eigenvector, and α depends upon the initial conditions. The remaining terms are exponentially suppressed relative to $v^{(2)}$; the approximation in the second line holds in the limit that λ_2 is close to unity.

The second eigenvector, $v^{(2)}$, describes the perturbation that takes the longest time to die away. It allows us to split the system into two subspaces based on the signs of its entries; we define all of the

states where $v^{(2)}$ is, say, positive, as “subspace one”, and the remaining states as “subspace two”. One set of states is associated with the positive $v^{(2)}$ values; the complementary set with those that are strictly negative. Informally, initial conditions that are weighted towards states solely in one of these two sets (all positive $v^{(2)}$ values or all negative) take the longest to decay to the stationary distribution. It is these two spaces that will define the system epochs.

The use of the sign structure of an HMM’s eigenvectors is a natural way to identify these subspaces; similar methods have been applied to, for example, the identification of metastable states in chemical reaction networks through what is called Perron-cluster cluster analysis (PCCA) [35–37]. Use of just the second eigenvector naturally splits the system into the two most significant subspaces. The sign structure of the third and higher eigenvectors can be used to further decompose the space in a hierarchical fashion, potentially subdividing these larger modules (see, *e.g.*, Ref. [38]). In this paper, we focus solely on the first division, and do not attempt to identify substructures within either of the two main clusters.

Once we know the parameters of the HMM, we can use Viterbi path reconstruction [39] to reconstruct the maximum-likelihood path through the state space. We then know, at any point in the time-series, where we are in the underlying computational state space.

This allows us to associate a internal state of the HMM to each step in the time-series, and thus to each edit on the page itself. We can also as pinpoint when a transition from one subspace to the other occurs. When a system switches subspaces, there can be some flickering, with rapid shifting back and forth between subspaces before the system settles down. We only count a transition from one subspace to another when the system remains in the new subspace for more than ten time-steps.

Given this definition, we can then study time spent within each subspace; we define this empirical quantity as the *trapping time*, $\tilde{\tau}$; per the definition of τ and the Levin-Peres-Wilmer Theorem (see Appendix C), we expect the trapping time to be of order τ . We report both the trapping time and the relaxation time in our results here. The trapping time is the more empirical quantity, since it describes the actual behavior of the system: what state the system was actually in at a particular time, and how long the system was trapped in one subspace or the other during its evolution; we can report $\tilde{\tau}$ for each subspace separately or just consider the average. Meanwhile, the relaxation time is the more theoretical quantity, characterize the HMM in isolation, in terms of the generic properties of the kinds of time-series it tends to generate.

In order to test the extent to which τ (and $\tilde{\tau}$) are driven by system-wide patterns of linking between internal states (rather than just, for example, generic sparseness, or an overall tendency for any state to link to only a small number of other states), we consider a null model for τ where we shuffle the entries of each state’s probability vector. This scrambles the overall structure machine while keeping the list of transition probabilities for each state constant.

On the simplest level, subspaces can be defined by their average levels of conflict—the fraction of time they lead to the emission of a revert symbol, R. We track this using the revert ratio: the fraction of reverts in the higher-conflict subspace divided by the fraction of reverts in the lower conflict subspace.

Subspaces are defined by more than just their levels of reverting. What distinguishes them in a deeper sense are the relationships between their hidden internal states, which make some sequences more likely than others. System response to a revert will be different, and lead to different characteristic futures, in one subspace, compared to the other.

We characterize subspaces by reference to the relative frequency of motifs: short sequences of system behaviors. We consider the distribution of motifs when the system is in each subspace; we then rank motifs by reference to their partial-KL from the mixture distribution [40]. A motif i is characteristic of a subspace if the quantity $p_i \log p_i / m_i$ is large, where p_i is the probability of the motif in the subspace, and m_i is the probability of the motif overall. Partial-KL provides an information theoretic measure of the extent to which a particular motif is a signal of the underlying subspace [41].

1.5. Causes of State Transitions

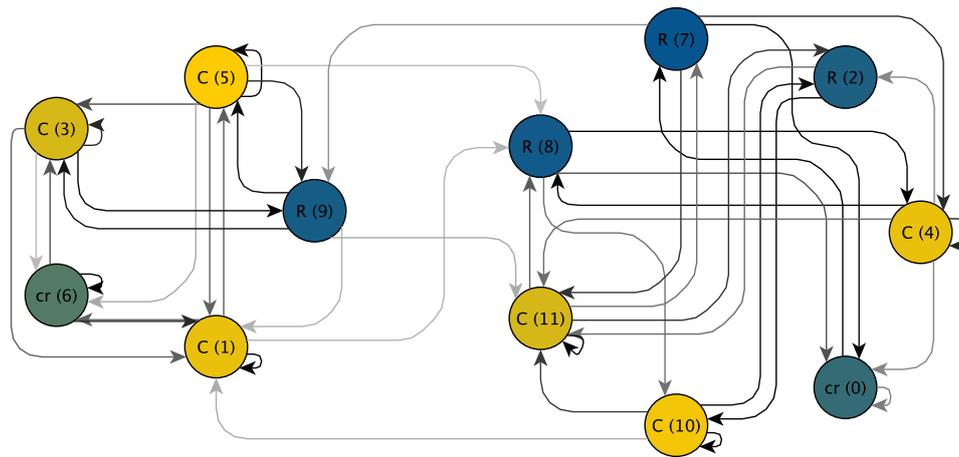
Given an identification of transition points, we return to the original time-series to catalog features of the system at these critical points, and to determine potential causes for the transition. We consider three types of potential causes. In all three cases, we look for associations between these potential causes, and transitions between subspaces. An association is defined as a potential cause occurring within a particular time window of a transition event.

First, we track *page protection events*, points at which administrators changed the access permissions of a particular page to either prevent or allow editing by different user classes. Records of page protection events are reliably logged beginning on 10 November 2003. Page protection is a crucial, hidden element of the system's conflict management, and has a significant impact on the composition of a page's editing population [42]. We expect page protection events to be the main source of top-down control that can switch the system between trapping subspaces. We characterize a protection event as “hard” when it leads to a restriction on who may edit the page; “soft”, when it releases the page from a restriction. The bulk of protection events shift pages between “anyone can edit” and “semi-protected”; the latter restricts edits to users who have been registered for a sufficient number of days, and have made a sufficient number of edits on other pages.

Second, we track *anti-social user events*. Transitions between subspaces may, potentially, be induced by unusually anti-social users who instigate self-sustaining conflict by—for example—publically violating norms of interaction. Rather than define an externally-imposed standard that may not reflect the reality of online interaction, we define anti-sociality by reference to community norms, as expressed by so-called user blocking events. Wikipedia administrators are able to block users from editing for a period of time. Receiving a block is a signal that one has (in the opinion of at least one administrator) violated a community norm.

We define a user as anti-social in the context of a particular page if that user has a blocking rate (number of blocks per total number of edits) higher than 95% of randomly-selected editors on the same page. For any transition, we check to see if the user with the most edits within a certain window meets this anti-social criterion; if so, we consider this a potential explanation of the transition. This provides a quantitative measure of the extent of internal opposition to dominant norms.

Finally, we track *major external events*; transition dates associated with major events concerning the topic of the page itself, and defined as significant increases in news coverage for the article subject.



Page Behavior	. . . CRCRRRCCCCRRRCCCC	RCRCRCRCRCRCRCRRRR . . .
Internal State	. . . 593939555559395555	848484848484848480707 . . .

Figure 1. Top panel: Hidden Markov Model for cooperation and conflict on the George W. Bush page. States are labelled by the probability of emitting each of the two output symbols; C ($P(C) > 0.8$, lighter yellow), R ($P(R) > 0.8$, darker blue) or cr (otherwise). Edge weights show transition probabilities; the lightest lines, connecting the two subspaces, correspond to probabilities of order 10^{-4} . Despite its complexity, the system is dominated by a transition logic that, on short timescales, confines the system to one of two separate modules with a high density of internal transitions. **Bottom panel:** below the diagram, as an example of our method, we show the Viterbi reconstruction in the neighborhood of one of the sixteen between-subspace transitions in our data. As the system wanders, during this period, among internal states 5, 9, 3, 8 and 4, it emits the symbols C and R probabilistically depending on which particular state it's in. On November 8th, 2004 at 9:11 UTC, long runs of cooperation and conflict gave way to more rapid-fire conflict and vandal-repair as the system crossed from the left module to the right, via the bridge between state 5 and state 8.

To do this, we rely on two well-curated, public databases—the New York Times archives, and the Guardian API¹—and track sudden increases in the density of news coverage. We quantify news spikes by reference to the rate of articles in a four-day window around the position in question, and define a major external event as one that leads to a fluctuation in this ratio over and above the null rate at 95% confidence. We search for news articles using the full text of the article's title.

Time-series data, including both raw data and the coarse-grained R-C series, resultant best-fit HMMs, and code for both EM estimation of Markov Chains and for Viterbi reconstruction are available online as the package SFIHMM.²

2. Results

¹ <http://developer.nytimes.com>; <http://open-platform.theguardian.com>

² See <http://bit.ly/sfihmm>.

For all sixty-two pages, the EM algorithm converged and we were able to find best-fit hidden Markov models. For the majority of pages (56 of 62), model selection preferred models with at least six internal states.

The network of transitions between states were sparse; any internal state had significantly non-zero transition probabilities to only a small number of others. States themselves usually had near-deterministic emission rules; a particular state usually had near-unity probability of emitting one of the two possible R or C symbols. Informally, this implies the existence of a variety of interaction motifs, particular patterns of Rs and Cs, at the multi-symbol level.

An example of the underlying computation process for a page is shown in Fig. 1, for the most-edited page in Wikipedia, that associated with George W. Bush. The sparseness of the internal state connectivity can be seen in the small number of high-probability transitions for each node; the determinism of the internal states from the fact that nearly all states are strongly biased, at at least 80-20, towards one or other of the two symbols. Also clearly visible is the modular structure of the hidden system; most transitions occur within one of two clearly separated subsets of the states. Over the course of nearly fifteen years and 45,448 edits, the page switched between these two subspaces only sixteen times; we show the time-series surrounding one switch, which happened on November 8th, 2004. Distinct, repeated motifs can be seen, informally, across the transition; we discuss these more below.

Table 1 then shows the recovered machines for the top ten pages by number of edits; the layout for each of the machines is dictated by the recovered modules, with states that are in the same module grouped together visually.

2.1. Epoch Detection

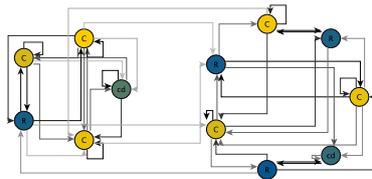
The distribution of relaxation times, τ , is shown in Fig. 2. For all sixty-two pages, τ is exceptionally long. The average relaxation time in our sample is 698 steps, corresponding to a λ_2 of roughly $1 - 10^{-3}$; the median is 287 steps. The effect size is large; the median relaxation time for the null model is only 12 steps, and for any particular page, the observed relaxation time is on average 50 times longer than the null expectation. In the majority of cases (52 of 62) there is significant evidence that global machine structure, in addition to transition sparseness, is leading to these long relaxation times.

Using Viterbi reconstruction, we can infer the actual transitions that occur on the pages. Coarse-graining subspace transitions to remove flickering, we find a total of 1,390 transition events over the sixty-two pages. The average trapping time, page-by-page, is 1,201 steps, or 190 days. Trapping times are longer than relaxation times. Epoch switching events in our sample do not appear to be concentrated at any special range of time in the last fifteen years; the distribution of switching times is indistinguishable, in a Kolmogorov-Smirnov test, from the distribution of edit times in general.

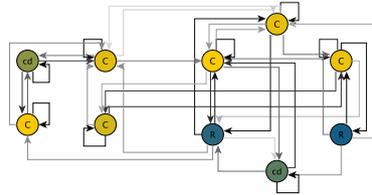
The two subspaces have distinct levels of expected reverts. We refer to the subspace with higher levels of reverting as “high-conflict” (type one), and the one with lower levels of reverting as “low-conflict” (type two). We emphasize that reverts track only one feature of conflict, and that the lower conflict subspace not only shows significant amounts of reverting, but also that (as we shall see), the lower conflict subspace may have more examples of norm-violating conflict. We suggest use of the phrase “type-one conflict” rather than “conflict associated with being in the higher-conflict subspace” (and similarly for “type two

Table 1. Hidden Markov Models, and derived parameters, for cooperation and conflict on the ten most-edited pages on Wikipedia. Editing patterns are characterized by high levels of determinism, and long-timescale trapping in distinct higher and lower-conflict spaces.

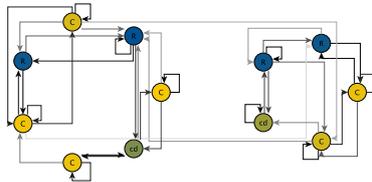
George_W._Bush
 τ : 1451 steps
 High/Low Conflict Ratio: 1.83
 Median edit-time (minutes):
 High: 3.53; Low: 15.85
 Average trapping time (months):
 High: 3.71; Low: 13.63



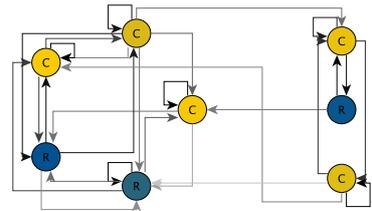
United_States
 τ : 360 steps
 High/Low Conflict Ratio: 3.61
 Median edit-time (minutes):
 High: 13.48; Low: 14.78
 Average trapping time (months):
 High: 5.78; Low: 4.2



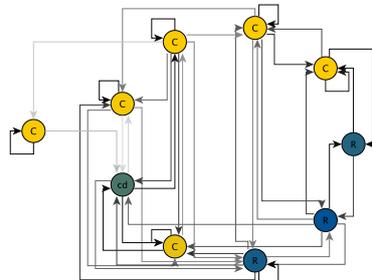
Wikipedia
 τ : 3695 steps
 High/Low Conflict Ratio: 1.75
 Median edit-time (minutes):
 High: 2.58; Low: 19.6
 Average trapping time (months):
 High: 7.28; Low: 45.8



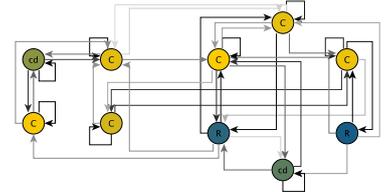
Michael_Jackson
 τ : 227 steps
 High/Low Conflict Ratio: 2.09
 Median edit-time (minutes):
 High: 7.75; Low: 8.42
 Average trapping time (months):
 High: 0.99; Low: 7.63



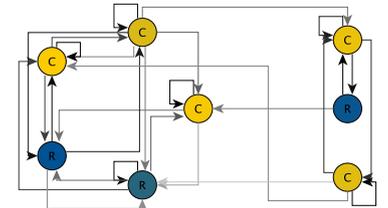
Catholic_Church
 τ : 2357 steps
 High/Low Conflict Ratio: 18.09
 Median edit-time (minutes):
 High: 13.72; Low: 4.63
 Average trapping time (months):
 High: 61.45; Low: 2.55



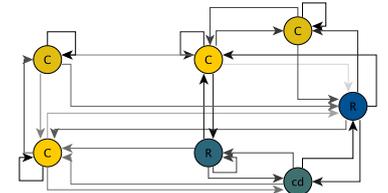
United_States
 τ : 360 steps
 High/Low Conflict Ratio: 3.61
 Median edit-time (minutes):
 High: 13.48; Low: 14.78
 Average trapping time (months):
 High: 5.78; Low: 4.2



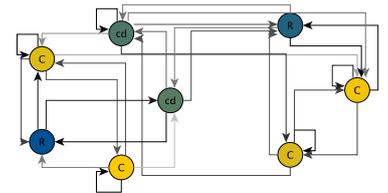
Michael_Jackson
 τ : 227 steps
 High/Low Conflict Ratio: 2.09
 Median edit-time (minutes):
 High: 7.75; Low: 8.42
 Average trapping time (months):
 High: 0.99; Low: 7.63



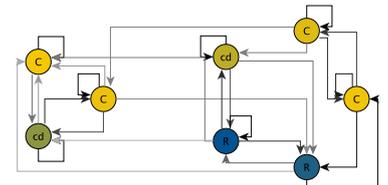
Barack_Obama
 τ : 53 steps
 High/Low Conflict Ratio: 2.31
 Median edit-time (minutes):
 High: 17.58; Low: 7.73
 Average trapping time (months):
 High: 1.82; Low: 0.46



Global_warming
 τ : 277 steps
 High/Low Conflict Ratio: 2.23
 Median edit-time (minutes):
 High: 9.72; Low: 19.45
 Average trapping time (months):
 High: 2.27; Low: 7.55



Islam
 τ : 287 steps
 High/Low Conflict Ratio: 2.67
 Median edit-time (minutes):
 High: 12.02; Low: 18.13
 Average trapping time (months):
 High: 3.16; Low: 7.87



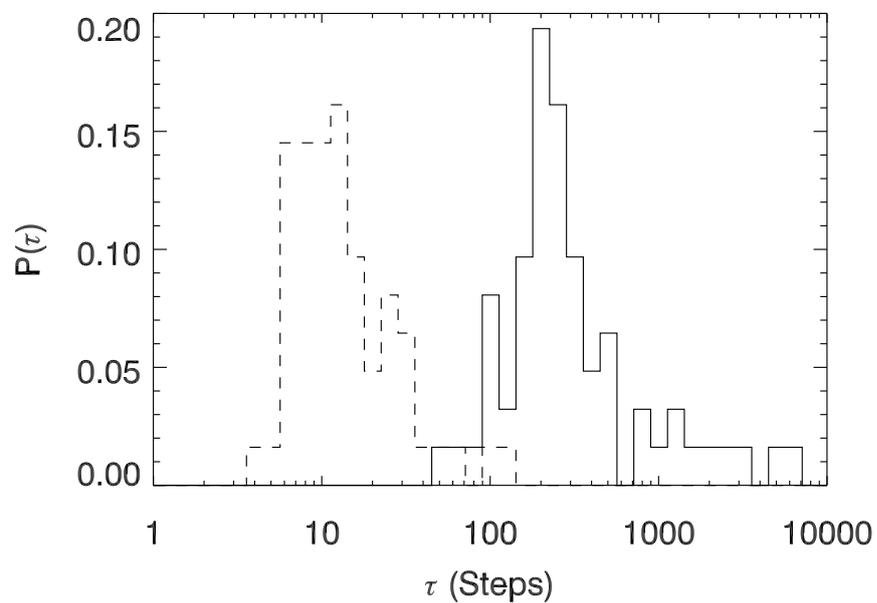


Figure 2. Relaxation time, τ , for the sixty-two pages in our sample (solid line). Times are exceptionally long, an average of 698 steps, and on average a factor of 50 times longer than expected for Markov models with similar sparseness (dotted line). The longest trapping times are for the pages associated with the Gaza War, the Russo-Georgian War, and the page describing Wikipedia itself.

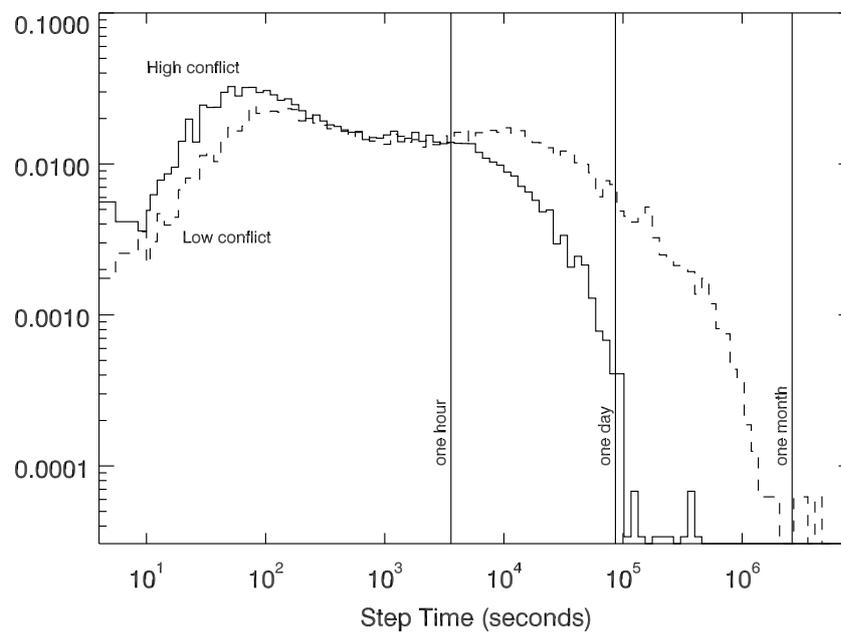


Figure 3. Distribution of edit rates for type one (high conflict) and type two (low conflict) subspaces. Shown here, as an example, the spacing between edits on the George W. Bush page. When the system is in the high-conflict subspace, edits occur once every 212 seconds (median; 3.5 minutes); in the low-conflict subspace, once every 951 seconds, or every 16 minutes. When the system is in the high-conflict subspace, users almost never wait more than a day to take action.

conflict”) for simplicity. While we use this language informally, we refer the reader to our discussion above regarding the limitations of reverts as tracer of more elaborate notions of conflict.

On average, the higher conflict subspace is roughly 2.5 times more revert-prone than the “low-conflict” subspace. Residency times are also slightly longer in the high-conflict state (207 days vs. 149 days). When the system is trapped in the higher conflict subspace, we find that editing usually, though not always, accelerates: the spacing between edits declines. In the mean, across all pages, users edit on average 2.6 times faster when in the higher conflict subspace; 46 of the 62 pages show this acceleration (high conflict faster than low conflict). The differences in edit rate are driven in part by a long-tail of long-wait times; if we consider not the average time between edits, but the median time between edits, the difference shrinks, and only 37 of 62 pages show acceleration. An example of the distribution of edit timescales is shown in Fig. 3, for the case of the George W. Bush page; when the system is in the high-conflict subspace, edits occur almost every three minutes, compared to once every 16 minutes in the low-conflict subspace. Other differences appear; 37% of edits are made by unregistered users (“IP addresses”) in the high-conflict subspace, compared to 21% of edits in the low-conflict subspace.

As can be seen in the sample Viterbi reconstruction in Fig. 1, the transition from one state to the other is characterized by more than just an increase in revert-rate. High- and low-conflict subspaces are characterized by very different interaction styles. Table 2 shows the characteristic motifs over the full set of 62 pages. As expected, the low-conflict subspaces are distinguished by runs of cooperation; however, they are also, unexpectedly, distinguished by runs of reverts. When reverts follow each other, they are usually part of an edit war [18], involving a small number of users repeatedly undoing each other’s work in a tit-for-tat pattern.

Conversely, vandalism (when a user inserts patently offensive or nonsense content into a page) followed by a (pro-social) revert is one of the characteristics of the high-conflict subspace; the vandal appears as a C move, and the pro-social repair as R. CR events are roughly 2.7 times as common when the system is in the high-conflict subspace, and edit comments that report vandalism (using the keywords “vandal” or the shorthand “rvv”) are 1.9 times as common.

Not all CR motifs are simple vandalism, however. The underlying CR pattern is simply a rejected proposal; repeated CRs in a timeseries is a signature of repeated rejection of different proposals. As an example, we take a (randomly chosen) short period when the Hillary Clinton page is in the high-conflict subspace.³ Of the eighty edits of that epoch, there are twelve CR pairs; six of them revert simple vandalism (introduction of sexual slurs), while five involve rejection of content that was contributed by a single editor apparently acting in good faith, and one is ambiguous.

CR motifs are a significant source of the higher levels of system conflict: while 28% of all edits in the high-conflict subspace are reverts, only 8% of reverts remain once CR motifs are dropped. Removing CR motifs, however, increases the relative rates of conflict; the high-conflict subspace has 2.7 times more reverts than the low-conflict subspace when CR pairs are removed, compared to 2.5 in the base condition. When comparing conflict rates between the subspaces, the residue is worse than the mixture.

³ See https://en.wikipedia.org/w/index.php?title=Hillary_Clinton&offset=20070217204656&limit=87&action=history; the system entered the high-conflict subspace on 8 February 2007, 09:21 UTC, and left on 17 February 2007, 20:14 UTC.

Table 2. Characteristic motifs of the higher (type one) and lower (type two) conflict subspaces across all 62 pages, ranked by partial-KL. The low-conflict subspace is characterized by long runs of cooperation, but also by long runs of reversion. Conversely, the high-conflict subspace is characterized by more rapid patterns of alternation between R and C moves.

Size	Type One (High-Conflict) Motifs	Type Two (Low-Conflict) Motifs
2	CR, RC	CC, RR
3	CRC, RCR, RCC, CCR	CCC, RRR, RRC, CRR
4	RCRC, CRRC, CRCC, CCRC, RCCR	CCCC, RRRR, RRCC, RCRR, CRRC
5	CRCRC, RCRCR, RCRCC, CCRCR, CRCCR	CCCCC, RRCRR, RRRRR, CRRRR, RRRRC

Transitions are not accompanied by significant population shifts. On average, 20.2% of the users that have appeared in the 100 edits just prior to a transition point also appear in the 100 edits just following. This is, in fact, slightly higher than the persistence we expect across an arbitrary point (17.7%) in our data. Transitions to the high (or low) conflict subspace are not associated with unusual rates of turnover.

2.2. Drivers of Conflict Transitions

For the three event types that may be causal drivers of the transition to or from the high-conflict state—page protection, anti-social user, and external events—there are three potentially relevant questions:

1. How many events are there, and what fraction are associated with a transition? (Effectiveness)
2. What fraction of transitions are associated with an event? (Explanatory Power)
3. For those transitions that we can associate with an event, what fraction have the expected effect? (Valence)

Item 1 we consider a measure of event *effectiveness*; an event type that more often leads to a transition is more effective. Item 2 we consider the extent to which an event type can *explain* an observed transition; if more transitions are associated with such an event type, that type is the more influential in explaining state switching. Item 3, *valence*, measures the extent to which the event has the expected effect; when an administrator locks a page, for example, it is natural to assume that the goal is to push the system out of a high-conflict state. Measures of valence only make sense when an event type is effective or has explanatory power.

These measures should be kept distinct, just as they are, for example, in the testing of a new drug. Taking a drug may correlate with a change in the patient’s symptoms (apparent effect); a patient’s symptoms may fluctuate independently of her use of the drug (low explanatory power); the drug may increase, rather than reduce, the severity of the symptoms (wrong valence).

2.2.1. Page Protection Events

Administrators can exercise direct control over who edits a page. These page protection events are the primary mechanism for authorities to influence the patterns of editing on the page; they are a blunt and top-down instrument that restricts editing on a particular page to increasingly smaller populations (“hard”

protection), or, conversely, opens it up to the wider community (“soft” protection events). Releasing a page from protection can occur by explicit action or by default, when a prior restriction is given an expiration date.

Page protection is mostly *ineffective* in inducing a transition. Of the 1,545 protection events in our data, only 136 (8.8%) occur within ten edits of a transition event; while the effect is statistically significant (35, or 2.3% expected; $p < 10^{-3}$), the effect size is small.

Page protection *can explain some*, but far from all, transition events. There are 1,387 transition events that we can, potentially, associate with a page protection;⁴ thus, at most, about 9.8% of observed transitions can be associated with page protections (2.5% expected; $p < 10^{-3}$).

Norms on the encyclopedia itself urge the use of page protection as a response to conflict. We thus might expect protections to lead to a transition to the low-conflict state; and, conversely, that weakening of protection to lead to a transition to the high-conflict state. This happens 80% and 72% of the time, respectively. Interventions are not usually successful in changing a page state, but when they do, their effects *do have the correct valence*. Page protection appears to operate in the direction administrators expect.

Increasing the window increases the number of transitions we can potentially associate with a protection event. When we enlarge the window size to one hundred edits, 397 of the transition events, or 29%, are now associated with a page protection event. The signal is statistically significant ($p < 10^{-3}$), but the false detection rate is higher (240, or 17% expected) and the valence less certain (65% rather than 80%). Post-selecting for the optimal window provides an upper bound: at most, only about 11% of transitions can be reliably explained by page protection events.

Top-down control thus has at best only weak effects on the transition between high- and low-conflict subspace. This is surprising, given the association between vandal-like CR patterns, and the high-conflict subspace. “Locking down” a page may lighten the burden for pro-social users who police vandalism, but the high-conflict subspace describes more than just vandalism, and excluding users—even large classes of them—only rarely induces lasting effects on long-term patterns of cooperation and conflict. Conversely, opening up a page rarely leads to an emergence of new conflict. Most top-down actions are unsuccessful in causing a page transition; most transitions can not be explained by reference to top-down action. Comfortingly, however, when associations can be found, they do have the correct valence. It is unusual for a hard protection event, for example, to be associated with a transition to the high-conflict subspace.

2.2.2. Anti-Social User Events

A transition to the high-conflict state could be occasioned by the appearance of particularly anti-social users. Recall that anti-social users are defined by having an average blocking rate higher than 95% of users who have appeared on the page. Anti-social users are said to “dominate” a transition when they make the plurality of edits within the window in question.

⁴ This is less than the total number of transitions, because we do not have reliable records of protection events prior to 10 November 2003.

There are a total of 710 transitions to the high-conflict state in our data. Of those transitions, 212 are associated, within 10 steps, with the dominance of an anti-social user, compared to an expected 208 in the null. We do not find above-null evidence of anti-social users triggering a transition ($p > 0.1$). Anti-social user events *do not explain* the transitions we see in the data. Expanding the range to 100 steps leaves the results unchanged: 255 transitions can be associated with an anti-social user event, compared to 263 in the null. Despite a significant focus on the management of problematic users within the encyclopedia [9,22], their appearance at a particular point does not appear to be a proximate cause of conflict.

This high null-expectation rate (35%) implies that a significant fraction of edits—at least on high-traffic pages—are made by users who have received significant numbers of administratively-imposed sanctions compared to their total number of edits. The simple appearance of this kind of user, however, is ineffective in inducing a transition. They seem to play many roles in the system, including roles consistent with the system remaining in the low-conflict subspace, perhaps due to preference change and dynamical learning of social norms [43].

Anti-social users *do not explain* transition events and they are *ineffective* in inducing a transition.

2.2.3. Major External Events

Not all pages are associated with a topic sufficiently newsworthy that external events can be tracked; once we eliminate pages where less than 10% of months have an article during the page lifetime, we are left with a sample of 59 pages.

Epoch transitions associated with external events include George W. Bush's election win in November 2004, Bob Dylan's first number one hit in thirty years in September 2006, and Barack Obama's election win in November 2008 and his inauguration in January 2009. However, many significant news events are not associated with epoch transitions, including the death of Michael Jackson in 2009 (which left other significant traces on the page [18]), and John Kerry's election loss in November 2004. Our sample of most-edited pages on Wikipedia includes a page on the 2006 Lebanon War, begun during the war itself. Previous work on breaking news collaborations has found distinctive editing patterns associated with events of this form [44]; in the Lebanon case, the two transitions (to high-conflict, on 9 October 2006; and back to low-conflict, on 13 August 2013), occur after the war itself has concluded.

There are a total of 1,367 transitions that we could potentially associate with newsworthy events. A total of 146 of these transitions are associated (10.6%) with such an event. We expect, on average, 126 events in the null case (9.2%). While there is weak evidence for some above-null association ($p \approx 0.03 < 0.05$), the effect size appears to be small: less than 2% of all transition events can be reliably associated with outlier news-coverage.

Unusual spikes of real world news coverage *do not explain* transition events and, because of the lack of association, we can also assert that they are *ineffective* in inducing a transition.

3. Discussion

The results described here provide novel evidence for the existence of an epoch-like structure of conflict and cooperation on Wikipedia, distinguished by behavioral motifs. When we model page-editing

as a finite state machine, we find that pages can be trapped in one or another subset of their computational subspaces, often for hundreds of days at a time.

At the most coarse-grained level, this trapping places the system in either a low or high-conflict state. In the high conflict state, we also see an acceleration of activity by a factor of two or more. Subspaces are characterized by more than just overall levels of conflict: they have distinct motifs of interaction, including tit-for-tat (repeated reverting) in the low-conflict subspace, and propose-reject (C then R) and strong signals of vandal repair in the high-conflict subspace.

The epoch structure revealed by our finite state analysis is defined by far more than just the density of reverts. As can be seen in the example time-series Fig. 1, the low-conflict subspace can have long runs of reverting and edit warring. Low-conflict, in other words, does not mean zero conflict, and may even mean norm violation—repeated reverting is likely to violate the “three revert rule” (3RR) that prohibits a single user from making three reverts to a page in a row.

In game theoretic accounts of behavior on Wikipedia, reverts can be likened to “defect” moves, and repeated reverts by different users something we might expect from a tit-for-tat like strategy under noise [16]. The association of long strings of reverts with the low-conflict subspace fits with accounts that put tit-for-tat at the heart of successful resolution of collective action problems from asocial beginnings [45]. The fact that they violate the Wikipedia-specific 3RR norm suggests that this more ancient tit-for-tat strategy may still play an important functional role in managing conflict. Individuals may be banding together: past work on Wikipedia conflict has examined the influence of edit wars that draw in multiple mutually-reverting subgroups, and these more structured interactions may characterize the low-conflict subspace. In the case of edits to the Terri Schiavo page, for example, a network analysis finds three groups of antagonists: one group of administrators, and two groups associated with a particular point of view [10].

The patterns we describe here extend over thousands of edits; our epoch-structure is on a longer timescale than the laws that describe short-term, repeated cooperation on timescales of minutes and hours [1]. It is consistent with findings of the importance of long-term system memory in the bursty structure of edit wars [13,29,46–48]. Our work goes beyond previous studies to explicitly construct an approximation to the full social grammar of system conflict [49]. Since the number of users far exceeds the number of states, each state is an irreducible coarse-graining of an (implicit) fine-grained, complete account of the mental states of hundreds of individuals acting in a context set by the text on the page itself.

A natural language for the discussion of the transition between conflict and cooperation comes from the critical transitions literature. This work draws on tools from catastrophe theory, critical bifurcations, and phase transitions [50,51]. They focus on the unusual and potentially chaotic dynamics that emerge from the interaction of large numbers of heterogeneously-coupled units. These accounts are often mechanism-neutral and focus on abstract properties of a time-series of events. In the literature on critical transitions, for example, one looks for properties such as spatial correlation [52], critical slowing-down [53], or flickering [54], and can apply their logic to clinical depression just as well as ecosystem collapse [55]. In its focus on discrete states, rather than continuous fields, our work is complementary to this tradition. It can be thought of as the extension of linguistics to the construction of collectively-implemented social grammars [49].

We have focused on the binary classification revert/non-revert. Recent work [2] has studied the motifs associated more fine-grained classifications of Wikipedia editing, focusing in particular on the nature of the user making the edit (for example: is this the users first edit on Wikipedia, the first on this article, the first in this “session”?). In a mixed strategy that combines quantitative and qualitative analysis, they looked at short timescale motifs of these more complex patterns associated with content co-production “routines” [56,57]. Our work suggests that extending this analysis, using the hidden Markov model tools presented here, to longer timescales may well reveal correspondingly longer-timescale epochs, which may or may not be nested within the conflict and cooperation structure we find here. Indeed, a parallel analysis of different timeseries would us to determine the extent to which different cultural practices coexist and interact: the how the strategies people use to improve content vary in the presence of the strategies people use to defuse conflict. Indeed, one of the most influential essays in Wikipedia’s norm network explicitly combines the two: the so-called “BOLD, revert, discuss cycle”⁵, which explicitly links content improvement and content reverting [22].

Our work also connects to the now-extensive literature in organizational behavior on Wikipedia and computer-supported collaborative work (CSCW). Our findings on the dynamic switching between patterns of interaction is consistent with work that has emphasised the different ways that Wiki-like organizations can deal with resource flows that lead to either generative or constrained tensions [58]. Studies in this literature have emphasised the role of the community structure [14,59], and describe systems like Wikipedia as a knowledge conversation [60] where the interaction between participants, and between participants and the wider context of content on the encyclopedia [61–63], rather than just between participant and content, determines the system’s evolution over time.

We have chosen a particularly simple computational model for our system—formally, a system at the base of the Chomsky hierarchy [64]. A finite-state machine can, in short runs, approximate more sophisticated grammars, but the pumping lemma (and its simple probabilistic extension; see Refs. [1,65]) means that, on sufficiently long timescales, it must fail. Indeed, our epoch-like structure already suggests the existence of long-range memory: repeated CR motifs, for example, in one month, correlate with the presence of CR motifs many months later. Examination of the detailed structure of our machines may suggest priors for higher-order grammars, such as the context-free languages that allow for nested pairing.

Within the field of timeseries analysis itself, a open question concerns the relationship between these simple finite-state models and the epsilon machines [66–68]; while hidden Markov models generically imply infinite-state epsilon machines, recent progress has suggested ways to lossily compress these representations to a finite system [69]. These results currently suggest that, because of the exceedingly long relaxation times (small spectral gaps) seen in Wikipedia, estimating some of the quantities relevant to the computational mechanics paradigm is impossible even for the very longest sequences ($\sim 10^4$ data points) we can observe. However, other recent work in this tradition suggests that, with a sufficiently explicit phenomenological model, estimation of quantities such as excess entropy may be possible [70]. Our current belief is that discovery of these very small spectral gaps suggests that a finite-state machine is only an approximate model, and that the true nature of memory storage and processing is likely to be

⁵ https://en.wikipedia.org/wiki/Wikipedia:BOLD,_revert,_discuss_cycle

more interesting yet. Its description may demand novel mathematical structures beyond the simple cases currently found in the quantitative literature.

We examined three *prima facie* plausible mechanisms for switching the system between these two states: top-down administrative actions, the appearance of unusually anti-social editors, and major external events in the real world. In only one case—administrator action—did we find statistically significant associations compatible with a causal role for state-space switching.

One of the characteristics of the high-conflict subspace are motifs often (though not always) associated with vandalism—the repeated CR pairs. Part of the explanation of the transition to the high-conflict subspace ought to involve new sources of this particular behavior. Efforts to exclude anonymous IP addresses, or registered but untrusted users, suggests that administrators are able to recognize sources of vandalism and intervene to prevent them. However, while vandalism is more commonly associated with the higher-conflict type one state, the majority of edits do not show signatures of vandalism, and locking down a page is largely ineffective in shifting the system to the lower-conflict type two state. Moreover, transitions in (or out) of the higher-conflict state are not associated with higher user turnover, and so can not be explained by an unexpected influx of new users.

Top-down effects are not the only thing that can drive pages between higher and lower conflict states: the effects of multiple user-user and user-page interactions can conspire to leave long-term traces on the page, as shown by the mathematical model of Refs. [71,72]. In this model, user behavior is dictated by a set of parameters that describe when a user will attempt to alter the page or, conversely, adapt to the page as it stands and, via discussion, to the opinions of others. When the user population is fixed, this process drives the system to consensus. However, when user turnover is sufficiently high (and the tolerance of different opinions sufficiently low), the system can be driven into a state of permanent conflict. Between these two phases is a line of critical coexistence, where—paralleling what we see in our analysis here—periods of consensus are interrupted by stretches of war.

These simple mathematical models provide an intriguing way for explaining our results. The bistability we observe empirically may be the consequence of an underlying process of belief revision when augmented by the presence of a common resource for information sharing. The question then becomes: what drives Wikipedia to the critical line separating these two phases?

While we are able to rule out spikes in news coverage as deterministic exogenous drivers, one of the key limitations of our work here is our inability to rule out more complicated interactions with the larger world. Long-term system memory may reside in users, in the text of the page, or well beyond the bounds of the encyclopedia. In the end, Wikipedia is not an isolated system. Its internal logic is coupled to the wider world, one rich in its own bursty and autocorrelated structure [73–77], and there is no clear division between the two [78,79]. It is this joint process we observe and must analyse: a small and shifting fragment of the internet as a whole.

4. Conclusion

Conflict is endemic to social life. Rather than exclude conflict altogether, however, most societies attempt to manage the ways in which conflict and cooperation interact. The co-existence of conflict and cooperation is a basic theme of studies of intra-group conflict in the biological sciences [80,81].

Quarreling and fighting are not simply unstructured forms of letting off steam. They are complex phenomena that make significant cognitive demands on individuals [82,83] and that result in epochs that decouple from day-night cycles and far outlast any particular clash [84,85]. Wikipedia, and IT-mediated social participation systems more generally, are unusual in the extent to which their underlying system norms foreground, and even valorize, conflict during co-production. In their broader structures, however the phenomena we see here would be familiar to researchers studying regularities in conflict across the biological world.

Finite-state modeling of Wikipedia's social grammar provides new insight into how individual-level complexity leaves a trace on the system as a whole. In a large-scale, virtual environment, with many tens of thousands of participants drawn from the wider internet, we find that conflict retains a complex logical structure, with context-dependent system memory that means that any particular event must be understood in a wider context that extends many months back, and may help define the system's future years hence. The dynamics of this memory and its associated patterns of conflict, can not be captured by simple top-down accounts.

Our results indicate new points of contact for the mathematical, social, and biological sciences. By measuring the essentially computational properties of a social world—its spectral gap or the logic of its subspaces—we gain new insight into the ways in which our species can trap itself in the very patterns it creates.

Acknowledgements

I thank Bradi Heaberlin (Indiana University), Taha Yasseri (Oxford Internet Institute), Brian Keegan (Harvard Business School), Cosma Shalizi (Carnegie Mellon University) and my three anonymous referees for comments on early versions of this manuscript, and the Global Brain Institute of Vrije Universiteit Brussel, Belgium at which this work was presented. I thank Nate Metheny (Santa Fe Institute) for the construction of machines Ganesha, Saraswati, and Laxmi that made the calculations in this paper possible. This work was supported in part by National Science Foundation Grant EF-1137929.

A. Articles in our analysis

The number of edits in each page's time-series is listed in parentheses after the article title, along with a simple classification: George_W._Bush (45448; biography, politician), United_States (33725; geography), Wikipedia (32592; technology), Michael_Jackson (27587; biography, entertainment), Catholic_Church (24813; religion), Barack_Obama (23889; biography, politician), World_War_II (23173; event, political), Global_warming (20003; science), 2006_Lebanon_War (19972; event, political), Islam (18523; religion), Canada (18150; geography), Eminem (18066; biography, entertainment), September_11_attacks (17564; event), Paul_McCartney (16973; biography, entertainment), Israel (16790; geographic), Hurricane_Katrina (16753; event), Xbox_360 (16753; technology), Pink_Floyd (16037; biography, entertainment), Iraq_War (15891; event), Blackout_(Britney_Spears_album) (15832; entertainment), Turkey (15663; geography), Super_Smash_Bros._Brawl (15432; technology), World_War_I (15292; event), Gaza_War (14920; event), Lost_(TV_series) (14897; entertainment), Blink-182 (14789; entertainment), Scientology (14727; religion), John_Kerry (14307; biography, political),

Heroes_(TV_series) (14223; entertainment), Australia (14186; geography), China (14023; geography), Bob_Dylan (13916; biography, entertainment), Neighbours (13547; entertainment), The_Holocaust (13346; event), Atheism (13295; religion), Hilary_Duff (13222; biography, entertainment), Mexico (13213; geography), The_Dark_Knight_(film) (13025; entertainment), France (12800; geographic), John_F._Kennedy (12788; biography, politician), Lindsay_Lohan (12757; biography, entertainment), Girls'_Generation (12746; entertainment), Argentina (12745; geography), Virginia_Tech_massacre (12682; event), RMS_Titanic (12451; event), Russo-Georgian_War (12365; event), Homosexuality (12170; science), Circumcision (12149; religion, science), Hillary_Rodham_Clinton (11981; biography, politician), Star_Trek (11919; entertainment), Shakira (11712; biography, entertainment), Sweden (11666; geography), New_Zealand (11639; geography), Paris_Hilton (11635; biography, entertainment), Wizards_of_Waverly_Place (11520; entertainment), Genghis_Khan (11410; biography, politician), Cuba (11390; geography), Linux (11316; technology), Che_Guevara (11250; biography, politician), Golf (11141; entertainment), iPhone (11085; technology), God (10731; religion).

B. Choosing the Number of States in an HMM

Table B.1. Using AIC and BIC to choose the number of states in a Hidden Markov model fit. Here we take an actual model from our data (the 8-state best fit model for the `God` page), use that model to generate a new time-series of equal length (10,731 samples), and attempt to fit a new model, using either AIC or BIC to select the preferred number of states in a manner similar to Refs. [86,87]. The table lists the fraction of the time this process led to a preferred machine of each size, for the two different penalties. Both AIC and BIC tend to *underestimate* model complexity; in general, BIC performs worse, significantly underestimating the true number of states. AIC performs better, recovering the correct number of states nearly half the time.

Number of states:	1	2	3	4	5	6	7	Truth 8	9	10
AIC	0.0%	0.0%	0.0%	0.0%	0.0%	37.5%	7.2%	45.8%	9.3%	0.0%
BIC	0.0%	0.0%	0.0%	20.8%	54.1%	23.9%	1.0%	0.0%	0.0%	0.0%

Choosing the number of states to include in an HMM is an example of a model selection problem—one is selecting between models with different numbers of states. In general, the larger the HMM, the better the data can be fit: when does one stop improving the fit because the model is becoming “too complex”? A generic solution to model selection is cross-validation: one fits the model using a subset of the data (the “training set”), and sees how well the model performs when predicting out of sample (the “test set” or “hold-out” set).

Cross-validation uses the phenomenon of over-fitting to determine when a model is too complex: if a model has too many parameters, it will overfit to the training set, finding “patterns” that are really due to coincidence. These patterns will fail to hold in the test set, and will degrade performance on the fit; this degradation can be measured, and one stops increasing model complexity when performance on the test

set first starts to decline. Standard cross-validation techniques work best when the data is independently sampled—*i.e.*, when it is possible to construct test set that is uncorrelated with the training set conditional on the underlying model. As increasing levels of correlation appear in the data, the construction of an appropriate training set becomes difficult.

When, as is the case for HMMs, the model is Bayesian, another method is possible: the likelihood penalty. To use a likelihood penalty, you fit the model to all of the data, and note the posterior log-likelihood. You then apply a penalty, reducing the log-likelihood depending on features related to the complexity of the model, including, usually, the total number of parameters. After applying this penalty, it is usually the case that one particular model—often not the most complex one—maximizes the penalized log-likelihood, and this is the one considered preferred. Numerous likelihood penalties exist, including the Bayesian Information Criterion (BIC [34]), the Bayesian Evidence (introduced in Ref. [88]; used on Wikipedia data in Refs. [1,16]), and the Akaike Information Criterion (AIC; introduced in Ref. [33]).

The existence of both long and short-range correlations in data fit using an HMM makes the use of cross-validation and hold-out techniques difficult. Ref. [86] propose two methods for cross-validation on HMMs that attempt to compensate for the failure of independence. Both methods have difficulties, and work well only on a subset of HMMs, where correlations are not particularly long-range, and the transition matrix is not too sparse. Because we find both conditions are violated in our data, we thus do not attempt cross-validation tests. Ref. [86] also consider the AIC and BIC methods; they find that both work well in recovering the true size of an HMM used to generate simulated data. (We are not aware of work that has tested the use of Bayesian Evidence on simulated data, and defer this interesting, but involved, question to later work.)

In general, Ref. [86] finds that, when it fails, BIC tends to slightly underestimate, and AIC to slightly overestimate, the true number of states, although both criteria work well. Work by Ref. [87] confirms this result. However, both papers consider regimes that do not directly apply here, and both papers found cases where (for example) BIC significantly underestimated model complexity. In addition, both papers consider problems with an order of magnitude less data than we have, and true model HMM sizes much smaller than ten. In the main text, we use the AIC penalty, a common choice made across the biological and signal processing communities, and strongly argued for on general grounds by Ref. [89] when the goal is minimizing prediction loss.

To validate our choice, and, in particular, to determine whether AIC or BIC produces more valid results for the regimes relevant here, we did an in-depth test with a particular model—the eight-state HMM associated with the `God` page. We took the derived HMM for this page, and used it to generate 96 simulated datasets of equal length to the original. We then ran our fitting code on each of these datasets, and compared what happened when we used of the AIC and the BIC criteria to select the preferred number of states. Consistent with Ref. [86,87], we found that BIC tended to underestimate model complexity, choosing machines that significantly smaller than the true number, and in fact never recovering the true system size. We found, by contrast, that AIC worked better; like BIC, it still often underestimated model complexity, but did so by smaller amounts. Conversely, a small fraction of the time (less than 10%) it preferred a model one state more complex than reality.

Because our main concern is to minimize overfitting and the introduction of fictitious structure, without losing too much of the actual structure in the process, AIC’s slight tendency to underestimate model complexity is not a major concern. A small fraction of the time (roughly 10%), AIC preferred a machine that was one state larger; however, robustness tests show that none of our main conclusions depend upon the exact number of states being that chosen by the AIC criterion. We recover the same conclusions, for example, if we arbitrarily fix the number of states equal to twelve. It is worth noting that we do not believe the “true” model of Wikipedian conflict is itself a finite-state machine [1]; *i.e.*, the fundamental problem is finding the best approximation, rather than locating the correct model within a known class.

C. Relaxation Time, Mixing Time, Decay Time, Trapping Time,

By the Levin-Peres-Wilmer Theorem [90], the relaxation time, τ , defined in Eq. 1 above, provides an upper bound on the mixing time, $\tau'(\epsilon)$, the maximum time it takes an arbitrary initial condition to be no further than a small distance ϵ from the stationary distribution, where distance is defined as the maximum absolute value difference for any state occupation probability. In particular, we have

$$(\tau - 1) \log \left(\frac{1}{2\epsilon} \right) \leq \tau'(\epsilon) \leq \tau \log \left(\frac{1}{\epsilon \pi_{\min}} \right), \quad (3)$$

where π_{\min} is the smallest value in the stationary probability distribution; of characteristic order 10^{-1} in the HMMs considered in this paper. These relationships hold when the HMM is “reversible” and “irreducible”; in empirical work such as that presented here those conditions are almost always satisfied. For example, they hold when states in an HMM have a small self-loop probability—no matter how small—and a directed path exists between any two states; both of which is true by default when using a Dirichlet prior, and is true by inspection in our EM fits. This justifies a useful and intuitive interpretation of τ as a measure of how quickly an arbitrary initial condition converges to the average, as well as suggesting more sophisticated measures that take into account the allowable level of deviation (ϵ), and inhomogeneities in the stationary distribution itself (π_{\min}).

The relaxation time τ is related to another natural quantity, the decay time of the second eigenvector. The time constant for decay, τ_d , is related to λ_2 as

$$\lambda_2^t = e^{-t/\tau_d}, \quad (4)$$

which implies that

$$\tau_d = -\frac{1}{\ln \lambda_2}. \quad (5)$$

On doing a Laurent series expansion around λ_2 equal to unity, we find

$$\tau_d = \frac{1}{1 - \lambda_2} - \frac{1}{2} + \mathcal{O}(\lambda_2 - 1) + \dots, \quad (6)$$

which implies that, to zeroth order in $\lambda_2 - 1$ (*i.e.*, when relaxation times are long), $\tau_d \approx \tau - 1/2$ as used in Eq. 2. All of the machines in this paper are in the regime where the relaxation time and decay time are nearly equivalent.

In contrast to relaxation, mixing, and decay time, trapping time is an empirically-measured quantity. To compute trapping time, you define a set of internal states of interest, and then use the Viterbi algorithm to

reconstruct the maximum-likelihood path through a particular time-series. Trapping time is then defined as the average length of time the system spends in the set of interest before it leaves. In this paper, we track trapping time for the two main subspaces as defined by the sign structure of the second eigenvector.

References

1. Simon DeDeo. Collective phenomena and non-finite state computation in a human social system. *PLoS ONE*, 8(10):e75818, 10 2013.
2. Brian C Keegan, Shakked Lev, and Ofer Arazy. Analyzing organizational routines in online knowledge collaborations: A case for sequence analysis in CSCW. *arXiv:1508.04819*, 2015. <http://arxiv.org/abs/1508.04819>. To appear, 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing. 27 February—2 March 2016, San Francisco, CA, USA.
3. F. Jelinek. Continuous speech recognition by statistical methods. *Proceedings of the IEEE*, 64(4):532–556, April 1976.
4. L. Bahl, J. Baker, P. Cohen, N. Dixon, F. Jelinek, R. Mercer, and H. Silverman. Preliminary results on the performance of a system for the automatic recognition of continuous speech. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '76.*, volume 1, pages 425–429, Apr 1976.
5. Steven J DeRose. Grammatical category disambiguation by statistical optimization. *Computational Linguistics*, 14(1):31–39, 1988.
6. Kenneth Ward Church. A stochastic parts program and noun phrase parser for unrestricted text. In *Proceedings of the Second Conference on Applied Natural Language Processing*, ANLC '88, pages 136–143, Stroudsburg, PA, USA, 1988. Association for Computational Linguistics.
7. Steven L. Salzberg, Arthur L. Delcher, Simon Kasif, and Owen White. Microbial gene identification using interpolated markov models. *Nucleic Acids Research*, 26(2):544–548, 1998.
8. David Darmon, Jared Sylvester, Michelle Girvan, and William Rand. Predictability of user behavior in social media: Bottom-up v. top-down modeling. In *2013 International Conference on Social Computing (SocialCom)*, pages 102–107. IEEE, 2013.
9. Joseph Michael Reagle. *Good faith collaboration: The culture of Wikipedia*. MIT Press, Cambridge, MA, USA, 2010.
10. Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. He says, she says: Conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 453–462, New York, NY, USA, 2007. ACM.
11. Gerald C Kane, Jeremiah Johnson, and Ann Majchrzak. Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities. *Management Science*, 60(12):3026–3048, 2014.
12. Travis Kriplean, Ivan Beschastnikh, David W. McDonald, and Scott A. Golder. Community, Consensus, Coercion, Control: Cs*W or How Policy Mediates Mass Participation. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work*, GROUP '07, pages 167–176, New York, NY, USA, 2007. ACM.
13. Fernanda B. Viégas, Martin Wattenberg, and Kushal Dave. Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 575–582, New York, NY, USA, 2004. ACM.
14. Gerald C. Kane. A multimethod study of information quality in wiki collaboration. *ACM Trans. Manage. Inf. Syst.*, 2(1):4:1–4:16, March 2011.

15. Ofer Arazy, Oded Nov, Raymond Patterson, and Lisa Yeo. Information quality in wikipedia: The effects of group composition and task conflict. *Journal of Management Information Systems*, 27(4):71–98, 2011.
16. Simon DeDeo. Group minds and the case of Wikipedia. *Human Computation*, 1(1):5–29, 2014. arXiv:1407.2210. <http://arxiv.org/abs/1407.2210>.
17. Clifford Geertz. Thick description: Toward an interpretive theory of culture. In Michael Martin and Lee C. McIntyre, editors, *Readings in the philosophy of social science*, pages 213–231. MIT Press, Cambridge, MA, 1994.
18. Taha Yasseri, Robert Sumi, András Rung, András Kornai, and János Kertész. Dynamics of conflicts in Wikipedia. *PLoS ONE*, 7(6):e38869, 2012.
19. Aniket Kittur, Ed H. Chi, and Bongwon Suh. What’s in Wikipedia?: Mapping topics and conflict using socially annotated category structure. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’09, pages 1509–1512, New York, NY, USA, 2009. ACM.
20. Ulrik Brandes and Jürgen Lerner. Visual analysis of controversy in user-generated encyclopedias? *Information Visualization*, 7(1):34–48, 2008.
21. Bongwon Suh, Ed H Chi, Bryan A Pendleton, and Aniket Kittur. Us vs. them: Understanding social dynamics in Wikipedia with revert graph visualizations. In *IEEE Symposium on Visual Analytics Science and Technology, 2007*, pages 163–170. IEEE, 2007.
22. Bradi Heaberlin and Simon DeDeo. The evolution of Wikipedia’s norm network. *Future Internet*, 8(2):14, 2016.
23. Pamela J. Hinds and Diane E. Bailey. Out of sight, out of sync: Understanding conflict in distributed teams. *Organization Science*, 14(6):615–632, 2003.
24. Karen A Jehn. A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly*, 42(3):530–557, 1997.
25. Aniket Kittur and Robert E. Kraut. Beyond Wikipedia: Coordination and conflict in online production groups. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW ’10, pages 215–224, New York, NY, USA, 2010. ACM.
26. Benjamin Collier and Julia Bear. Conflict, criticism, or confidence: An empirical examination of the gender gap in Wikipedia contributions. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW ’12, pages 383–392, New York, NY, USA, 2012. ACM.
27. Nicolas Auray, Céline Poudat, and Pascal Pons. Democratizing scientific vulgarization. the balance between cooperation and conflict in French Wikipedia. *Observatorio (OBS*) Journal*, 1(3), 2007.
28. Joseph M. Reagle Jr. “Be Nice”: Wikipedia norms for supportive communication. *New Review of Hypermedia and Multimedia*, 16(1-2):161–180, 2010.
29. Fernanda B Viégas, Martin Wattenberg, Jesse Kriss, and Frank Van Ham. Talk before you type: Coordination in Wikipedia. In *40th Annual Hawaii International Conference on System Sciences, 2007*, page 78. IEEE, 2007.
30. Leonard E. Baum and Ted Petrie. Statistical inference for probabilistic functions of finite state markov chains. *Ann. Math. Statist.*, 37(6):1554–1563, 12 1966.
31. William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes, 3rd edition: The art of scientific computing*. Cambridge University Press, 2007.

32. Terence C Mills. *Time series techniques for economists*. Cambridge University Press, Cambridge, UK, 1991.
33. Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
34. Gideon Schwarz et al. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
35. Peter Deuffhard, Michael Dellnitz, Oliver Junge, and Christof Schütte. Computation of essential molecular dynamics by subdivision techniques. In *Computational molecular dynamics: challenges, methods, ideas*, pages 98–115. Springer, 1999.
36. Frank Cordes, Marcus Weber, and Johannes Schmidt-Ehrenberg. *Metastable conformations via successive Perron-cluster cluster analysis of dihedrals*. Konrad-Zuse-Zentrum für Informationstechnik Berlin [ZIB], 2002.
37. Peter Deuffhard and Marcus Weber. Robust perron cluster analysis in conformation dynamics. *Linear algebra and its applications*, 398:161–184, 2005.
38. Frank Noé, Illia Horenko, Christof Schütte, and Jeremy C Smith. Hierarchical analysis of conformational dynamics in biomolecules: transition networks of metastable states. *The Journal of Chemical Physics*, 126(15):155102, 2007.
39. G David Forney Jr. The Viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278, 1973.
40. Sara Klingenstein, Tim Hitchcock, and Simon DeDeo. The civilizing process in London’s Old Bailey. *Proceedings of the National Academy of Sciences*, 111(26):9419–9424, 2014.
41. Simon DeDeo, Robert X. D. Hawkins, Sara Klingenstein, and Tim Hitchcock. Bootstrap methods for the empirical study of decision-making and information flows in social systems. *Entropy*, 15(6):2246–2276, 2013.
42. Benjamin Mako Hill and Aaron Shaw. Page protection: Another missing dimension of wikipedia research. In *Proceedings of the 11th International Symposium on Open Collaboration, OpenSym ’15*, pages 15:1–15:4, New York, NY, USA, 2015. ACM.
43. Simon DeDeo. Major transitions in political order. *arXiv:1512.03419*, 2015. <http://arxiv.org/abs/1512.03419>. To appear in *From Matter to Life: Information and Causality*, Cambridge University Press, edited by Sara Imari Walker, Paul C.W. Davies, George Ellis.
44. Brian Keegan, Darren Gergle, and Noshir Contractor. Staying in the loop: Structure and dynamics of wikipedia’s breaking news collaborations. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration, WikiSym ’12*, pages 1:1–1:10, New York, NY, USA, 2012. ACM.
45. R Axelrod and WD Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.
46. Róbert Sumi, Taha Yasserli, András Rung, András Kornai, and János Kertész. Characterization and prediction of Wikipedia edit wars. In *Proceedings of the ACM WebSci’11*, pages 1–3, Koblenz, Germany, June 2011.
47. Róbert Sumi, Taha Yasserli, András Rung, András Kornai, and János Kertész. Edit wars in Wikipedia. *arXiv:1107.3689*, 2011.
48. Taha Yasserli and János Kertész. Value production in a collaborative environment. *Journal of Statistical Physics*, 151(3-4):414–439, 2013.

49. Ray Jackendoff. *Language, consciousness, culture: Essays on mental structure*. MIT Press, Cambridge, MA, USA, 2007.
50. Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.
51. Steven J Lade and Thilo Gross. Early warning signals for critical transitions: a generalized modeling approach. *PLoS Computational Biology*, 8(2):e1002360, 2012.
52. Vasilis Dakos, Egbert H van Nes, Raúl Donangelo, Hugo Fort, and Marten Scheffer. Spatial correlation as leading indicator of catastrophic shifts. *Theoretical Ecology*, 3(3):163–174, 2010.
53. Vasilis Dakos, Marten Scheffer, Egbert H van Nes, Victor Brovkin, Vladimir Petoukhov, and Hermann Held. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105(38):14308–14312, 2008.
54. Rong Wang, John A Dearing, Peter G Langdon, Enlou Zhang, Xiangdong Yang, Vasilis Dakos, and Marten Scheffer. Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature*, 492(7429):419–422, 2012.
55. Ingrid A van de Leemput, Marieke Wichers, Angélique OJ Cramer, Denny Borsboom, Francis Tuerlinckx, Peter Kuppens, Egbert H van Nes, Wolfgang Viechtbauer, Erik J Giltay, Steven H Aggen, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1):87–92, 2014.
56. Martha S Feldman and Brian T Pentland. Reconceptualizing organizational routines as a source of flexibility and change. *Administrative science quarterly*, 48(1):94–118, 2003.
57. Brian T Pentland and Thorvald Hærem. Organizational routines as patterns of action: Implications for organizational behavior. *Annu. Rev. Organ. Psychol. Organ. Behav.*, 2(1):465–487, 2015.
58. S. Faraj, S.L. Jarvenpaa, and A. Majchrzak. Knowledge collaboration in online communities. *Organization Science*, 22(5):1224–1239, 2011.
59. Robert G. Fichman Gerald C. Kane. The shoemaker’s children: Using wikis for information systems teaching, research, and publication. *MIS Quarterly*, 33(1):1–17, 2009.
60. Ann Majchrzak, Samer Faraj, Gerald C. Kane, and Bijan Azad. The contradictory influence of social media affordances on online communal knowledge sharing. *Journal of Computer-Mediated Communication*, 19(1):38–55, 2013.
61. Sean Hansen, Nicholas Berente, and Kalle Lyytinen. Wikipedia, critical social theory, and the possibility of rational discourse. *The Information Society*, 25(1):38–59, 2009.
62. Sam Ransbotham and Gerald C Kane. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly*, 35(3):613, 2011.
63. S. Ransbotham, G.C. Kane, and N.H. Lurie. Network characteristics and the value of collaborative user-generated content. *Marketing Science*, 31(3):387–405, 2012.
64. Noam Chomsky. *Aspects of the theory of syntax*. MIT Press, Cambridge, MA, 1965.
65. M. O. Rabin and D. Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3(2):114–125, April 1959.
66. James P Crutchfield and Karl Young. Inferring statistical complexity. *Physical Review Letters*, 63(2):105, 1989.

67. James P Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, 75(1):11–54, 1994.
68. James P Crutchfield and Cosma Rohilla Shalizi. Thermodynamic depth of causal states: Objective complexity via minimal representations. *Physical review E*, 59(1):275, 1999.
69. Sarah E. Marzen and James P. Crutchfield. Predictive rate-distortion for infinite-order Markov processes. *Journal of Statistical Physics*, 163(6):1312–1338, 2016.
70. Sarah E. Marzen and James P. Crutchfield. Statistical signatures of structural organization: The case of long memory in renewal processes. *Physics Letters A*, 380(17):1517 – 1525, 2016.
71. János Török, Gerardo Iñiguez, Taha Yasseri, Maxi San Miguel, Kimmo Kaski, and János Kertész. Opinions, conflicts, and consensus: modeling social dynamics in a collaborative environment. *Physical review letters*, 110(8):088701, 2013.
72. Gerardo Iñiguez, János Török, Taha Yasseri, Kimmo Kaski, and János Kertész. Modeling social dynamics in a collaborative environment. *EPJ Data Science*, 3(1):1–20, 2014.
73. Albert-Laszlo Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 2005.
74. Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie Glance, and Matthew Hurst. *Patterns of Cascading Behavior in Large Blog Graphs*, chapter 60, pages 551–556. 2007.
75. Juliette Stehlé, Alain Barrat, and Ginestra Bianconi. Dynamical and bursty interactions in social networks. *Physical Review E*, 81(3):035101, 2010.
76. Zhao Zhi-Dan, Xia Hu, Shang Ming-Sheng, and Zhou Tao. Empirical analysis on the human dynamics of a large-scale short message communication system. *Chinese Physics Letters*, 28(6):068901, 2011.
77. Márton Karsai, Kimmo Kaski, Albert-László Barabási, and János Kertész. Universal features of correlated bursty behaviour. *Scientific reports*, 2, 2012.
78. Barry Wellman, Janet Salaff, Dimitrina Dimitrova, Laura Garton, Milena Gulia, and Caroline Haythornthwaite. Computer networks as social networks: Collaborative work, telework, and virtual community. *Annual Review of Sociology*, 22:213–238, 1996.
79. Barry Wellman. Computer networks as social networks. *Science*, 293(5537):2031–2034, 2001.
80. Jessica C Flack, David C Krakauer, and Frans BM de Waal. Robustness mechanisms in primate societies: a perturbation study. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568):1091–1099, 2005.
81. Jessica C Flack, Michelle Girvan, Frans BM De Waal, and David C Krakauer. Policing stabilizes construction of social niches in primates. *Nature*, 439(7075):426–429, 2006.
82. Simon DeDeo, David Krakauer, and Jessica Flack. Inductive game theory and the dynamics of animal conflict. *PLoS Computational Biology*, 6(5):e1000782, 2010.
83. Elizabeth A. Hobson and Simon DeDeo. Social feedback and the emergence of rank in animal society. *PLoS Computational Biology*, 11(9):e1004411, 09 2015.
84. Simon DeDeo, David Krakauer, and Jessica Flack. Evidence of strategic periodicities in collective conflict dynamics. *Journal of The Royal Society Interface*, 8(62):1260–1273, 2011.

85. Jessica C Flack. Multiple time-scales and the developmental dynamics of social systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597):1802–1810, 2012.
86. Gilles Celeux and Jean-Baptiste Durand. Selecting hidden Markov model state number with cross-validated likelihood. *Computational Statistics*, 23(4):541–564, 2008.
87. Silvia Bacci, Silvia Pandolfi, and Fulvia Pennoni. A comparison of some criteria for states selection in the latent Markov model for longitudinal data. *Advances in Data Analysis and Classification*, 8(2):125–145, 2014.
88. David JC MacKay. *Information theory, inference and learning algorithms*. Cambridge University Press, Cambridge, MA, USA, 2003.
89. Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, New York, NY, USA, 2003.
90. David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, USA, 2009. See Ch. 12; Theorem 12.3.