Modeling Relational Event Dynamics with statnet

Carter T. Butts
Department of Sociology and
Institute for Mathematical Behavioral Sciences
University of California, Irvine

Christopher S. Marcum
RAND Corporation

2012 Sunbelt Conference, Redondo Beach, CA

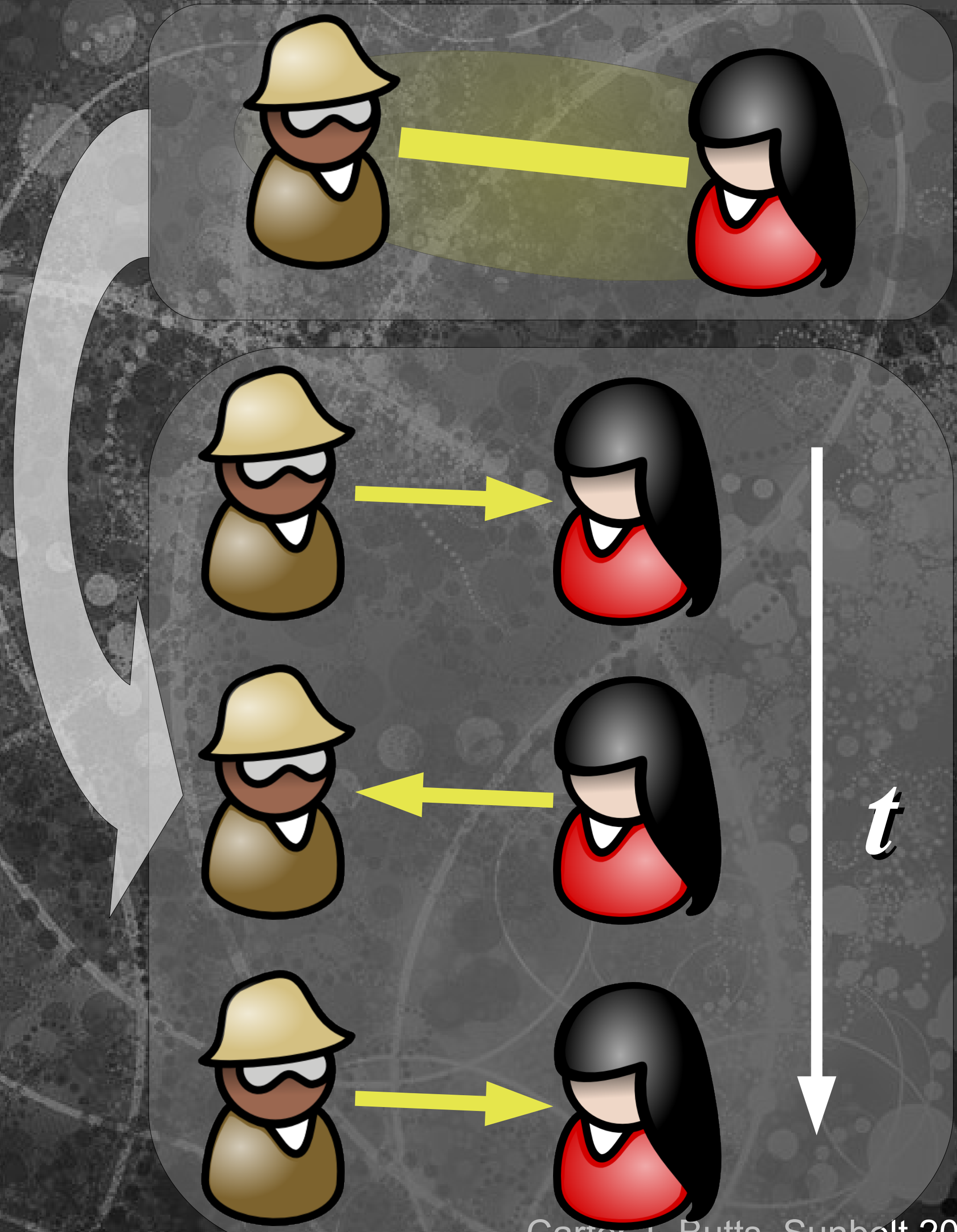
This work was supported by NSF awards IIS-0331707 and CMS-064257,
DOD ONR award N00014-8-1-1015, and NIH 1R01HD068395-01.

Overview

- Content in a nutshell
  - Introduction to the use of relational event models (Butts, 2006; 2008) for the modeling interaction dynamics
  - Why this approach?
    - Fairly general
    - Principled basis for inference (estimation, model comparison, etc.) from actually existing data
    - Utilizes well-understood formalisms (event history analysis, multinomial logit)
- This workshop:
  - Introduction to modeling approach
  - Dyadic relational event models
  - Egocentric relational event models
  - Modeling complex event sequences

Unpacking Networks: From Relationships to Action

- Conventional network paradigm: focus on temporally extensive relationships
- Powerful approach, but not always ideal
- Sometimes, we are interested in the social action that lies beneath the relationships....



Actions and Relational Events

- **Action: discrete event in which one entity emits a behavior directed at one or more entities in its environment**
  - **Useful "atomic unit" of human activity**
  - **Represent formally by relational events**
- **Relational event:** $a=(i,j,k,t)$
  - $i \in \mathcal{S}$: **"Sender" of event** a; $s(a)=i$
  - $j \in \mathcal{R}$: **"Receiver" of event** a; $r(a)=j$
  - $k \in \mathcal{C}$: **"Action type" ("category") for event** a; $c(a)=k$
  - $t \in \mathbb{R}$: **"Time" of event** a; $\tau(a)=t$

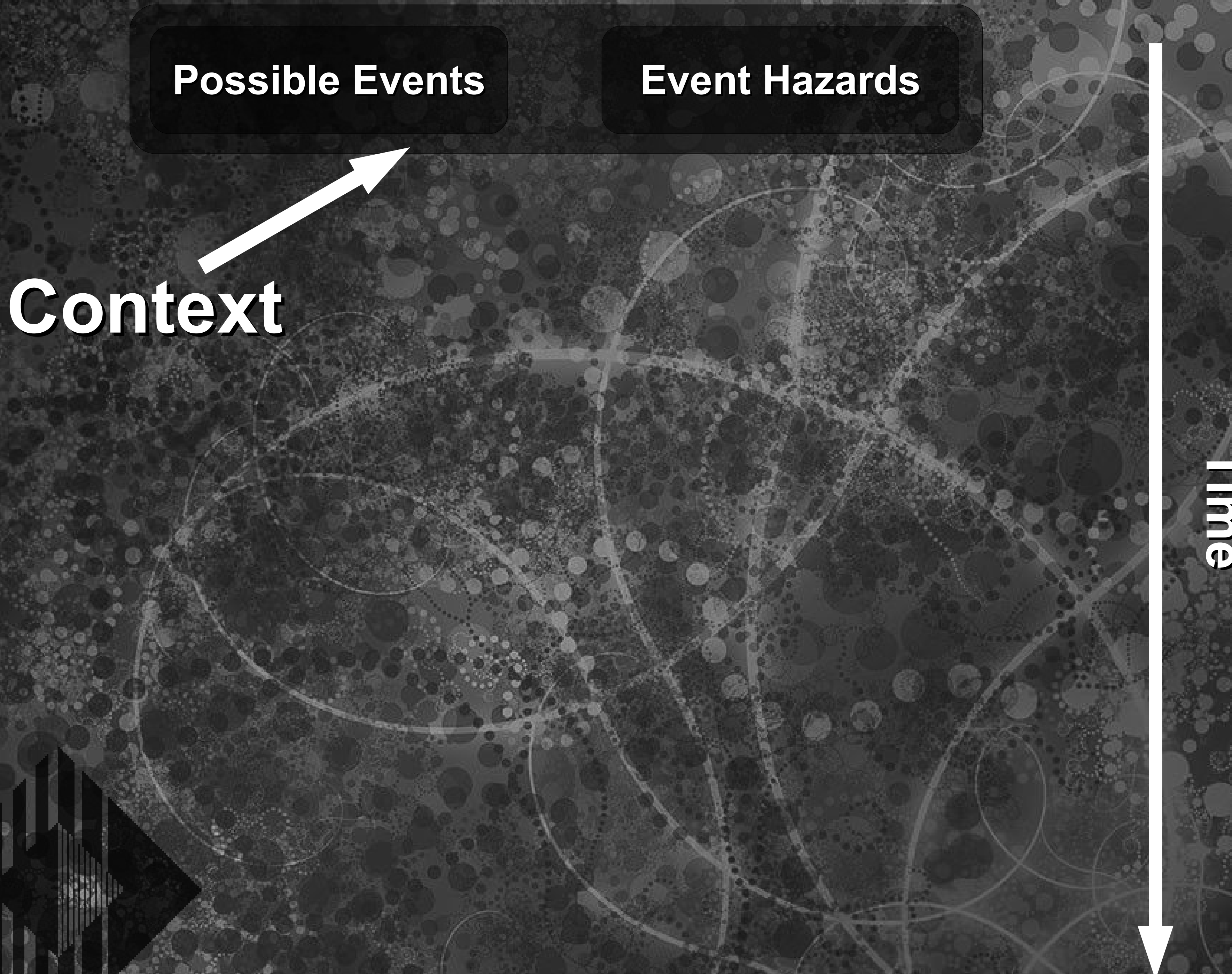
Events in Context

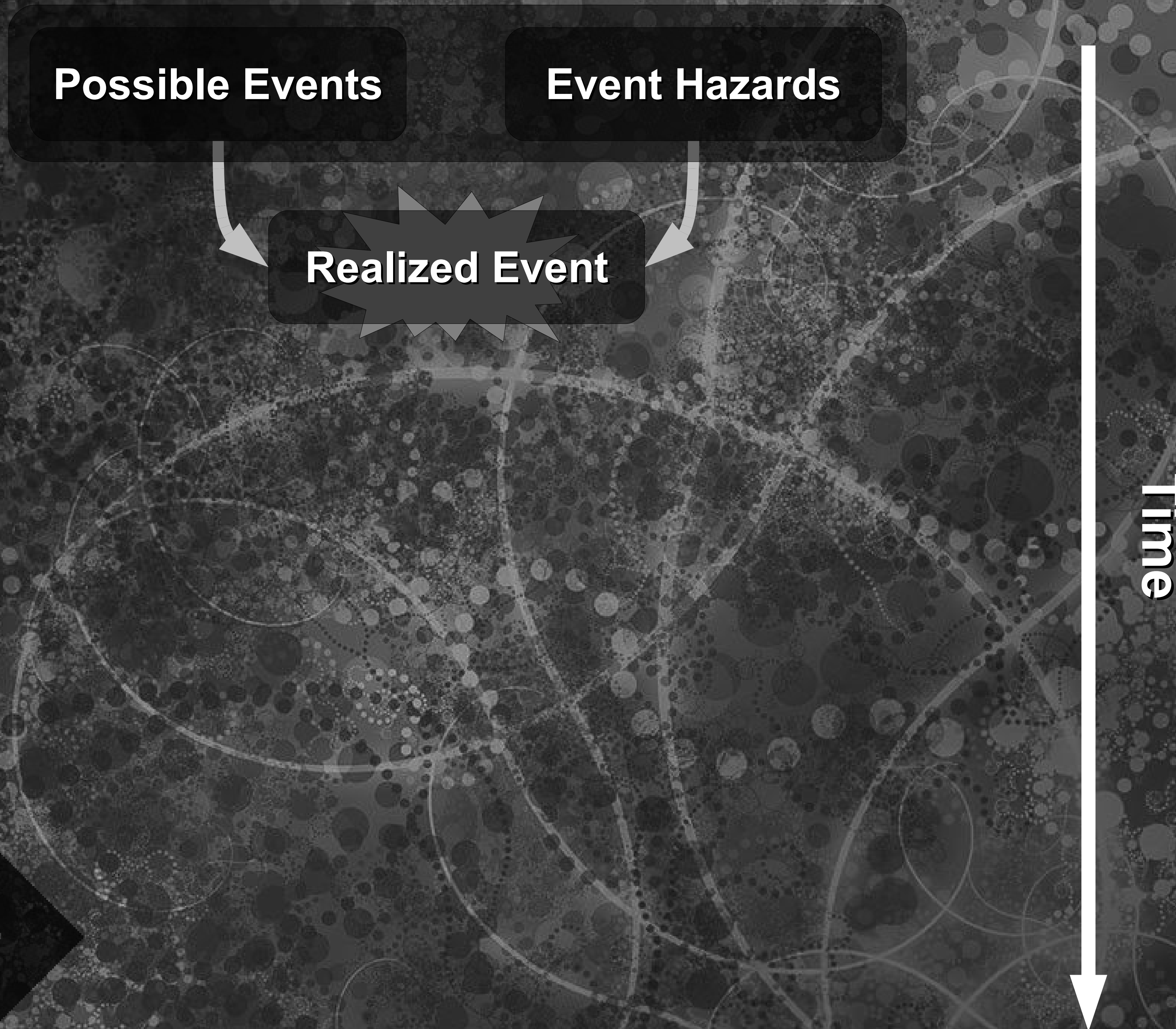
- **Multiple actions form an event history,** $A_t = \{a_i : \tau(a_i) \leq t\}$
  - **Take** a_0: $\tau(a_0)=0$ **as "null action",** $\tau(a_i) \geq 0$
  - **Possible actions at** t **given by** $\mathbb{A}(A_t) \subseteq \mathcal{S} \times \mathcal{R} \times \mathcal{C}$
    - **Forms support for next action**
    - **Assume here that** $\mathbb{A}(A_t)$ **finite, constant between actions; may be fixed, but need not be**
- **Goal: model** A_t
  - **Treat actions as events in continuous time**
  - **Hazards depend upon past history, covariates**

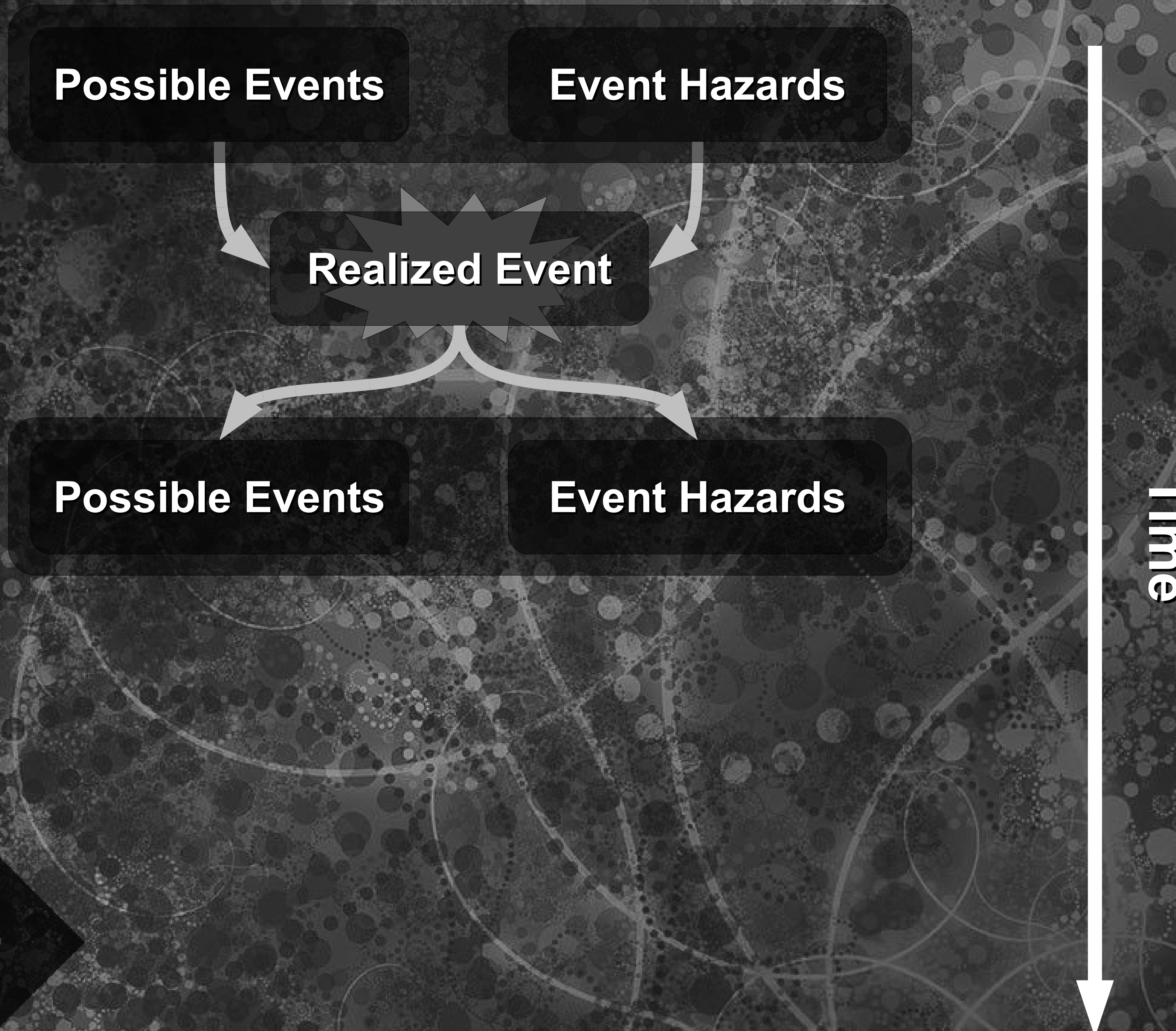
Possible Events

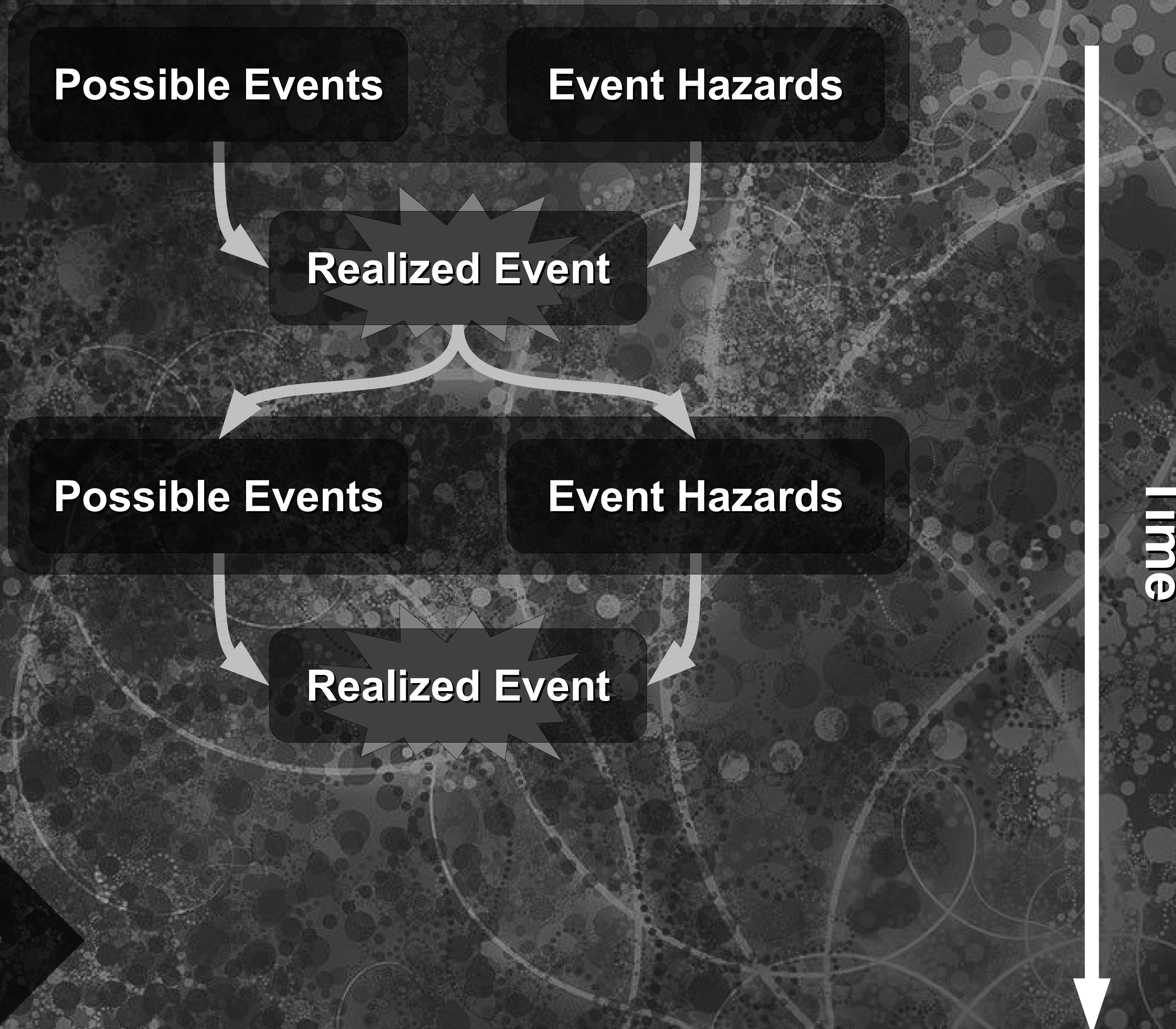
Event Hazards

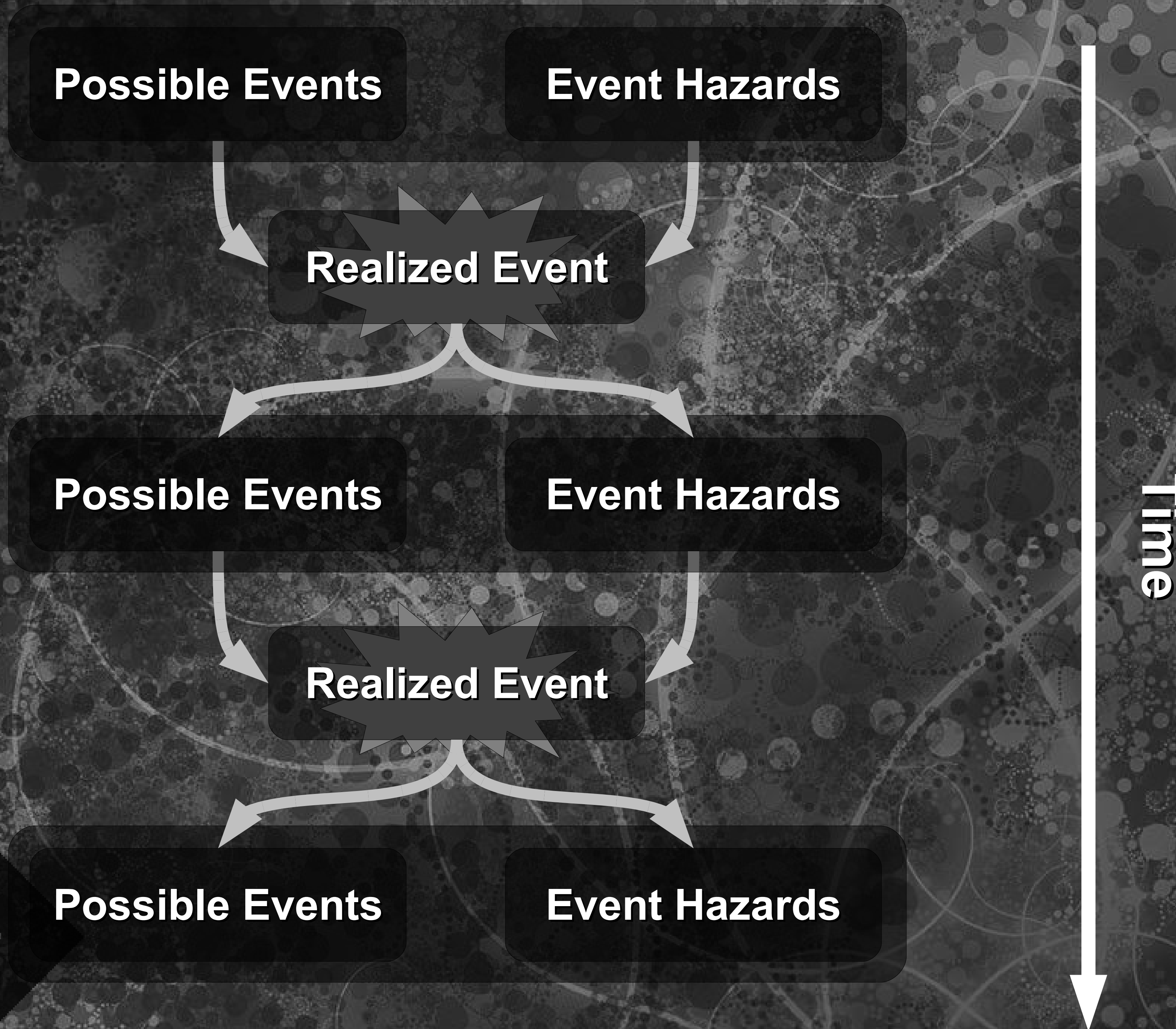
Time


Possible Events
Event Hazards
Context
Time


Possible Events
Event Hazards
Realized Event
Time


Possible Events
Event Hazards
Realized Event
Possible Events
Event Hazards
Time


Possible Events
Event Hazards
Realized Event
Possible Events
Event Hazards
Realized Event
Time


Possible Events
Event Hazards
Realized Event
Possible Events
Event Hazards
Realized Event
Possible Events
Event Hazards
Time

Carter T. Butts, Sunbelt 2012

Event Model Likelihood: Piecewise Exponential Case

- **Natural simplifying assumption: actions arise as conditionally independent, Poisson-like events with piecewise constant rates**
  - **Intuition: hazard of each possible event is *locally* constant, given complete event history up to that point**
    - **Waiting times conditionally exponentially distributed**
    - **Rates *can* change when events transpire, but not otherwise**
      - **Compare to related assumption in Cox prop. hazards model**
    - **Possible events likewise change only when something happens**
- **Can use to derive event likelihood**
  - **Let $M=|A_t|$, $\tau_i=\tau(a_i)$, w/hazard function $\lambda_{a_i A_k \theta}=\lambda(a_i, A_k, \theta)$; then**

$$p(A_t|\theta)=\left[\prod_{i=1}^{M}\left(\lambda_{a_i A_{\tau_{i-1}} \theta}\prod_{a'\in\mathbb{A}(A_{\tau_i})}\exp\left(-\lambda_{a' A_{\tau_{i-1}} \theta}\left[\tau_i-\tau_{i-1}\right]\right)\right)\right]\left[\prod_{a'\in\mathbb{A}(A_t)}\exp\left(-\lambda_{a' A_t \theta}\left[t-\tau_M\right]\right)\right]$$

The Problem of Uncertain Event Timing

- **Likelihood of an event sequence depends on the detailed history**
  - **Problem: exact timing is generally uncertain for many data sources (e.g., transcripts), though order is known**
  - **What if we only have (temporally) ordinal data?**
- **Stochastic process theory to the rescue!**
  - **Thm: Let $X_1,\ldots,X_n$ be independent exponential r.v. w/rate parameters $\lambda_1,\ldots,\lambda_n$. Then the probability that $x_i=\min\{x_1,\ldots,x_n\}$ is $\lambda_i/(\lambda_1+\ldots+\lambda_n)$.**
  - **Implication: likelihood of ordinal data is a product of multinomial likelihoods**
    - **Identifies rate function up to a constant factor**

Event Model Likelihood: Ordinal Timing Case

- **Using the above, we may write the likelihood of an event sequence A_t as follows:**

$$p(A_t|\theta) = \prod_{i=1}^{M} \left[\frac{\lambda_{a_i A_{\tau_{i-1}} \theta}}{\sum_{a' \in \mathbb{A}(A_{\tau_i})} \lambda_{a_i A_{\tau_{i-1}} \theta}} \right]$$

- **Dynamics governed by rate function, λ**

$$\lambda_{a A_t \theta} = \begin{cases} \exp\left(\lambda_0 + \theta^T u\left(s(a), r(a), c(a), A_t, X_a\right)\right) & a \in \mathbb{A}(A_t) \\ 0 & a \notin \mathbb{A}(A_t) \end{cases}$$

  - **Where λ_0 is an arbitrary constant, $\theta \in \mathbb{R}^p$ is a parameter vector, and u: $(i,j,A_t,X) \rightarrow \mathbb{R}^p$ is a vector of statistics**

Interpreting the Parameters

- In general, each unit change in u_i multiplies the hazard of an associated event by $\exp(\theta_i)$
  - For ordinal time case, unit difference in u_i adds unit of θ_i to log odds of a vs a'
- Connection to multinomial choice models
  - Let $\mathbb{A}_i(A_t)$ be the set of possible actions for sender i at time t. Then, conditional on no other event occurring before i acts, the probability that i's next action is a is given by

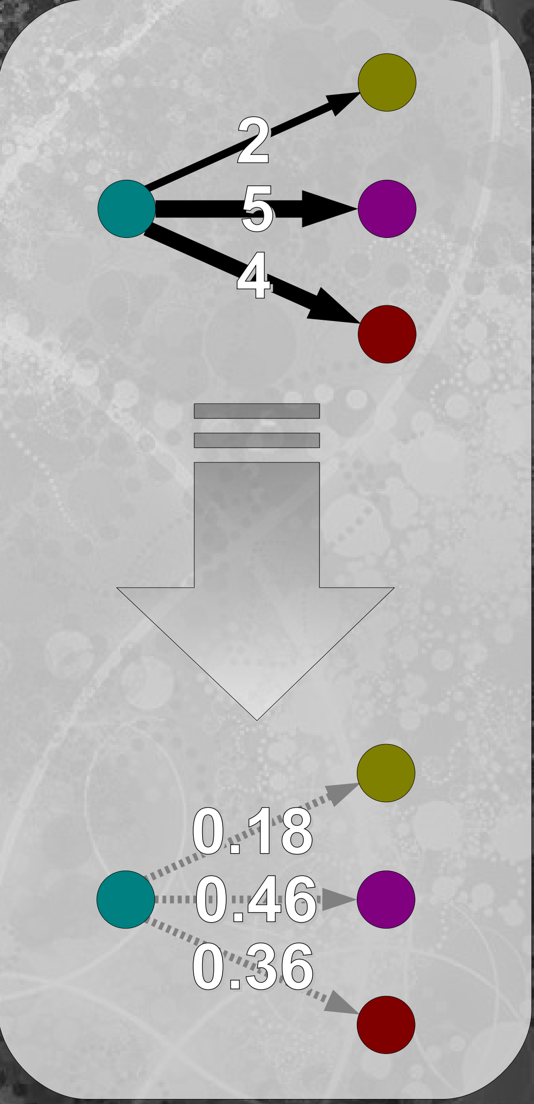
$$p(a|\theta)=\frac{\exp\left[\theta^T u\left(i,r(a),c(a),A_t,X_a\right)\right]}{\sum_{a'\in\mathbb{A}_i\left(A_t\right)}\exp\left[\theta^T u\left(i,r(a'),c(a'),A_t,X_{a'}\right)\right]}$$

Fitting Relational Event Models

- Given A_t and u, how do we estimate θ?
  - Parameters interpretable as logged rate multipliers (in u)
- We have $p(A_t|\theta)$, so can conduct likelihood-based inference
  - Find MLE $\theta^* = \arg\max_\theta p(A_t|\theta)$, e.g., using a variant Newton-Rapheson or other method
  - Can also proceed in a Bayesian manner
    - Posit $p(\theta)$, work with $p(\theta|A_t) \propto p(A_t|\theta)p(\theta)$
  - Some computational challenges when $|\mathbb{A}|$ is large; tricks like MC quadrature needed to deal with sum of rates across support

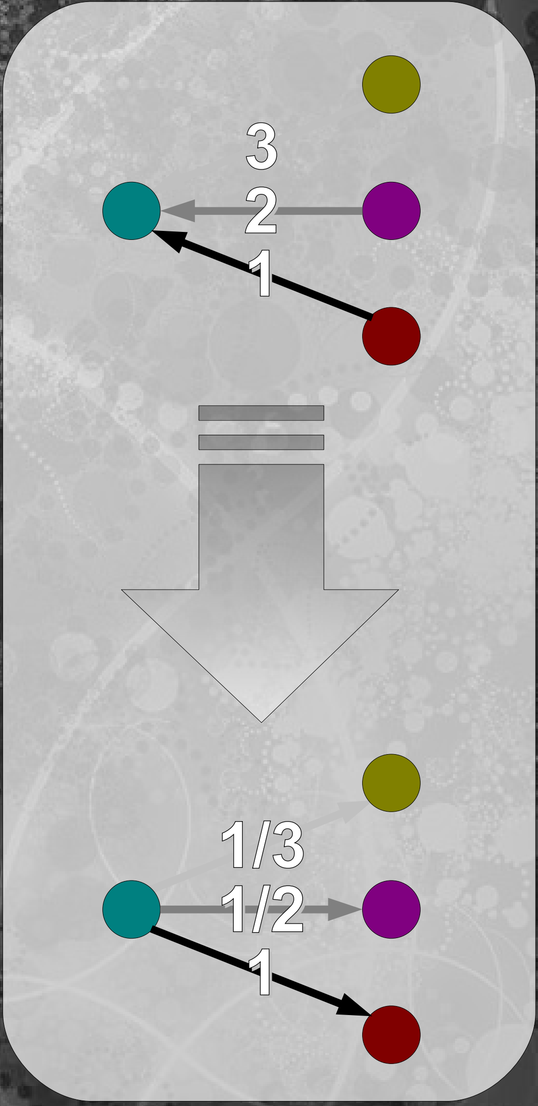
Persistence Effects

- Inertia-like effect: past contacts may tend to become future contacts
  - Unobserved relational heterogeneity
  - Availability to memory
  - (Compare w/autocorrelation terms in an AR process)
- Simple implementation: fraction of previous contacts as predictor
  - Log-rate of (i,j) contact adjusted by $\theta d_{ij}/d_i$



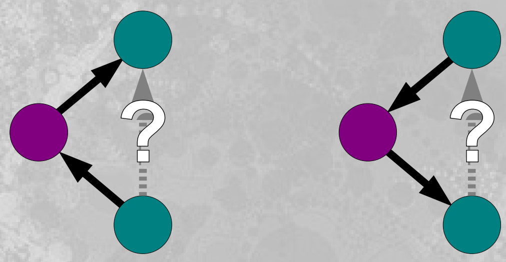
Recency/Ordering Effects

- Ordering of past contact potentially affects future contact
  - Reciprocity norms
  - Recency effects (salience)
- Simple parameterization: dyadic contact ordering effect
  - Previous incoming contacts ranked
    - Non-contacts treated as rank ∞
  - Log-rate of outgoing (i,j) contact adjusted by $\theta(1/\mathrm{rank}_{ji})$

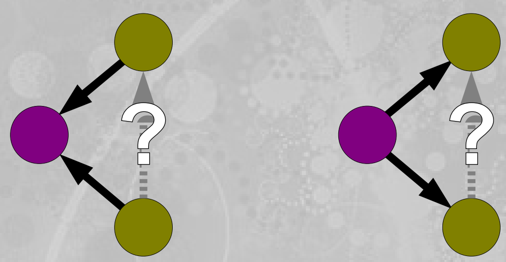


Triadic/Clustering Effects

- Can also control for endogenous triadic mechanisms
  - Two-path effects
    - Past outbound two-path flows lead to/inhibit direct contact (transitivity)
    - Past inbound two-path flows lead to/inhibit direct contact (cyclicity)
  - Shared partner effects
    - Past outbound shared partners lead to/inhibit direct contact (common reference)
    - Past inbound shared partners lead to/inhibit direct contact (common contact)



Two-Path Effects

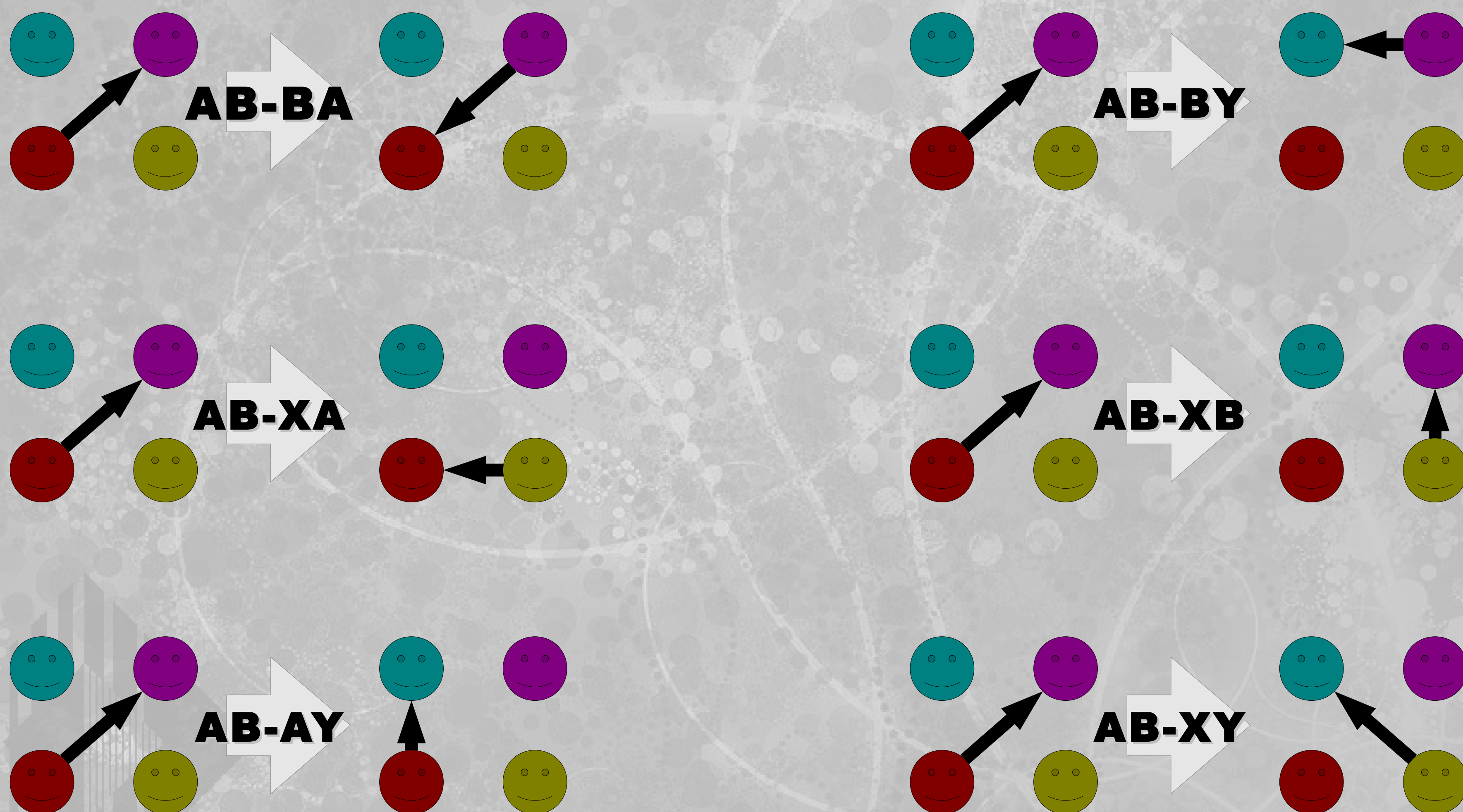


Shared Partner Effects

Participation Shifts

- Proposal of Gibson (2003) for studying conversational dynamics
  - Classify actors into *senders*, *receivers*, and *bystanders*
  - When roles change, a *participation shift* ("P-shift") is said to occur
  - Study conversational dynamics by examining the incidence of P-shifts
- P-shift typology
  - For dyadic communication, 6 possible P-shifts; allowing indefinite targets expands set to 13
  - Can compute observed, potential shifts given an event sequence

Dyadic P-Shifts, Illustrated



Preferential Attachment

- Past interactive activity affects tendency to receive action
  - E.g., emergent coordination roles
    - Exposure-based saliency ("who's out there?")
    - Practice/specialization (efficiency)
- Implement via past total degree effect on hazard of receipt
  - Fraction of all past calls due to *i* as effect for all *j* to *i* events