Rafal Urbaniak

Research Scientist, Basis Research Institute, New York, USA

Selected programming skills

(1) Bayesian and causal probabilistic programming (ChiRho and Pyro), (2) Causal dynamical systems modeling (ChiRho Dynamical), (3) Causal explanation module (ChiRho Explainable), (4) Data science stack (numpy, scipy, pandas, scikit-learn, matplotlib, seaborn, jupyter).

(1) Bayesian statistics (STAN, JAGS, BUGS, rethinking, rjags, runjags, BESTmcmc), (2) Bayesian networks (bnlearn, gRain), (3) discrete data models (countreg, vcd, vcdExtra, car, MASS), (4) data wrangling and visualization (tidyverse).

Selected open source contributions

Collab-creatures

Leading role in the development of a package that supports the use of Bayesian and causal inference with Pyro and ChiRho in the study of animal collaborative behavior. The current version focuses on foraging animals and the probabilistic identification of foraging strategies. (contribution metrics)

Explainable causal reasoning (ChiRho module)

Leading role in the development of this **core functionality in an open source programming language** developed by Basis. The package provides a systematic, unified approach to causal explanation computations by means of a single generic program transformation that can be applied to any arbitrary causal program. A range of causal explanation queries can be modeled in terms of probabilistic queries in the expanded model. (contribution metrics)

Cities: evaluating policy transfer via similarity analysis and causal inference

Modeling backend lead building the causal inferential backend for Polis.basis.ai, an open source inferential tool for local policymakers, built by Basis as part of the Opportunity Project cooperation with the Census Open Innovation Labs. (contribution metrics)

A Bayesian approach to uncertainty in word embedding bias estimation

A new Bayesian method of **measuring bias in language models**, contrasted with those employing single-number summaries. A related paper published in *Computational Linguistics*. (contribution metrics)

Probabilistic coherence measures over Bayesian networks

Algorithms for calculating the main existing **coherence** measures over Bayesian networks, with **a new method** essentially relying on the causal structure, with a resolution of multiple counterexamples to the existing. A related paper published in *Artificial Intelligence and Law*. (contribution metrics)

Legal Probabilism Bayesian networks in R

R implementation of **Bayesian network methods for criminal evidence evaluation**, based on our work with Marcello Di Bello on the Legal Probabilism entry in the Stanford Encyclopedia of Philosophy. 72 commits resulting in `5,794++ 2,235--` lines.

Namespotting: Username toxicity and actual toxic behavior on Reddit

In cooperation with Samurai Labs, we use algorithmic detection and Bayesian statistical methods to **analyze two large data streams (329k users) of Reddit content** to study the **correlations between username toxicity** (of various types, such as offensive or sexually explicit) **and their online toxic behavior** (personal attacks, sexual harassment among others). Results published in *Computers in Human Behavior*

Short-term impact of personal attacks on Reddit user activity

In cooperation with Samurai Labs, we **tracked 148,317 users and identified personal attacks among 182,528 posts** and comments using their high precision software. I **analyzed the data** from three perspectives: (i) classical statistical methods, (ii) Bayesian estimation, and (iii) model-theoretic analysis with hurdle and zero-inflated models. They agree: **despite the popular view that attacks make a communication channel more lively, personal attacks decrease the victims' activity and increase churn**. Results published in *Computers in Human Behavior*

Education

PhD in Logic and Philosophy of Mathematics
University of Calgary, 2008

Positions held

Postdoctoral Fellow (Research Foundation Flanders)
Associate Professor (University of Gdansk)
Instructor of record (University of Calgary)

Selected grants & awards

Kosciuszko Fellow, Northeastern University, Boston
Bednarowski Trust Fellow, University of Edinburgh
PI, National Science Centre OPUS project
PI, National Science Centre SONATA BIS project
Trinity College LRHub Visiting Fellow, Dublin
Visiting Fellow, Banaras Hindu University
British Academy Visiting Fellow, Bristol University

Selected academic activities

Taught more than 10 university courses related to probability and statistical programming.
Published 3 books and ca. 35 research papers.
Gave more than 30 invited lectures (e.g. University of Oxford, Paris I Panthéon-Sorbonne, University of Edinburgh, University of Canterbury, Kyoto University, Keio University, Nagoya University, University of Turin).
Awarded 13 major research grants and awards.
Organized 15 international conferences organized.
Referee for 18 journals.