MASS-CAUSAL

Causal Discovery and Modeling of Signal Pathways with Mass Cytometry Data

WORKSHOP PROGRAMME

Sunday 26/4			
9:00	9:10	**Ioannis Tsamardinos**	Welcome
9:10	9:50	**Angelika Schmidt**	The human immune system - Tregs take center stage.
9:50	10:30	**Angelika Schmidt**	Signaling in the immune system: Activation and suppression of T cell receptor signaling.
10:30	**11:00**		***Coffee break***
11:00	11:40	**Szabolcs Elias**	T cell subsets and T cell differentiation
11:40	12:20	**Karen Sachs**	Introduction to Mass Cytometry
12:20	13:00	**George Papoutsoglou**	Modeling signaling pathways with Ordinary Differential Equations
13:00	**14:00**		***Lunch break***
14:00	14:40	**Yannis Pantazis**	Modeling signaling pathways with Stochastic Differential Equations
14:40	15:20	**Gordon Ball**	Finding Mechanistic Hypotheses in ODE Models
15:20	**15:50**		***Coffee break***
15:50	16:30	**Takis Benos**	BD2K effort
16:30	17:00	**Round table event**	Data Sharing
Monday 27/4			
9:00	9:55	**Alain Hauser**	Basics of DAG-based causal models
9:55	10:35	**Jonas Peters Eleni Sgouritsa**	Causal discovery beyond conditional independencies (part 1)
10:35	**11:05**		***Coffee break***
11:05	11:45	**Jonas Peters Eleni Sgouritsa**	Causal discovery beyond conditional independencies (part 2)
11:45	12:25	**Jan Lemeire**	Faithfulness in causality
12:25	13:20	**Joris Mooij**	Causal discovery from cytometry data: what are the remaining challenges?
13:20	**14:20**		***Lunch break***
14:20	15:00	**Karen Sachs**	Probabilistic modeling: Methods and Challenges
15:00	15:40	**Ioannis Tsamardinos**	Integrative causal analysis, CAUSALPATH overview
15:40	**16:10**		***Coffee break***
16:10	17:00	**Round table event**	Software sharing and design
Tuesday 28/4			
9:00	9:40	**Alessio Farcomeni**	Variable selection under constraints with extensions to path analysis
9:40	10:20	**Lizzie Silver**	Scaling up the GES and FCI algorithms, using background knowledge, and more
10:20	**10:50**		***Coffee break***
10:50	11:30	**David Gomez-Cabrero**	Data management approaches to signal pathways
11:30	12:10	**Round table event**	Brainstorming: causal discovery with signal pathways
12:10	12:50	**Round table event**	Brainstorming: organizing collaborations, events, work dissemination
12:50	**13:50**		***Lunch break***

Abstracts

Sunday 26/4

The human immune system - Tregs take center stage.	After a short introduction about the immune system in general, the focus will be on T cells, a type of immune cells, - on a layman level. We will discuss how these T cells are activated, as well as how they are regulated by so-called regulatory T cells (Tregs). Tregs have emerged in the last years as indispensable players in the immune system which are needed to prevent overreactions of the immune system, such as autoimmune disease and allergies, and Tregs took center stage in recent breakthroughs of immune therapy. We will introduce the basics of Tregs and Treg-targeted therapies, and elaborate briefly on our ongoing work dealing with the generation of Tregs in cell culture.
Signaling in the immune system: Activation and suppression of T cell receptor signaling.	We will introduce general concepts and major pathways of cellular signaling using immune cell signaling as an example. We will then detail the signaling pathways involved in T cell receptor (TCR) signaling, which lead to activation of a T cell and its effector functions. We will show data on how Tregs can rapidly inhibit TCR signaling in conventional T cells, which consequently down-modulates the immune effector response. As an outlook, we will touch upon ongoing and future work aiming to decipher mechanisms of Treg-mediated suppression of TCR signaling.
T cell subsets and T cell differentiation	TBA
Introduction to Mass Cytometry	I will describe how CyTOF works, discuss logistics and compare to classic (fluorescence-based) flow cytometry. I'll also discuss some of the computational challenges, use cases and existing computational tools.
Traditional approaches to signal pathway induction using Ordinary Differential Equations	Typically, to develop an understanding on how signaling networks operate, dynamic models of ODEs using mass action kinetics are employed. During this talk I will present alternative descriptions of protein-protein interaction using ODEs, how these have been successfully employed for explaining a biological system's behavior and the latest advances towards de-novo network reconstruction.
Modeling signaling pathways with Stochastic Differential Equations	Stochastic phenomena in biochemical reaction networks can crucially affect the time evolution and the properties of metabolic, regulatory and signaling pathways. In this talk, I will present various modeling approaches that take into account the intrinsic stochasticity of reaction networks.
Finding Mechanistic Hypotheses in ODE Models	We examine how the properties of a complex ODE model can be explored to find an appropriate level at which to describe functionality and hence identify different mechanistic hypotheses about how the system evolves, using the example of a model of atherosclerosis onset.
BD2K effort	TBA

Monday 27/4

Basics of DAG-based causal models	A widespread framework for modelling causality is based on DAGs, directed acyclic graphs, whose arrows represent cause-effect-relationships. In this talk, the possibilities and limitations of this framework are discussed, and important classes of estimators are presented.
Causal discovery beyond conditional independences	Independence-based methods (such as IC, PC or FCI) exploit conditional independences to infer (parts of) the underlying causal structure. In this talk, we present alternative ideas for causal discovery that are based on additive noise models, "independence" of cause and mechanism or invariant prediction among different environments. Parts of this talk will focus on the two-variable case, in which there are no (conditional) independences. This talk is meant to be a tutorial, no prior knowledge is required.
Faithfulness in causality	Conditional independencies identified in the data are powerful evidence about the causal structure of the underlying phenomena, and as such have given rise to new algorithms to learn the causal structure from observational data. Faithfulness is an assumption that makes this task easy: all independencies are assumed to come from the structure. However, this assumption is not holding in practice.
Causal discovery from cytometry data: what are the remaining challenges?	I will discuss some of the challenges of causal discovery and causal prediction within the biological domain, focussing on protein expression data. In particular, I will present an analysis of the flow cytometry data set from Sachs et al. (2005) as a case study. I expect that many observations generalize to mass cytometry data. The good news: theory and algorithms are not sophisticated enough yet to deal properly with these types of data sets, so there is still interesting

	work to do. The bad news: many of our usual simplifying assumptions (no feedback, faithfulness, perfect interventions, only observational data, causal sufficiency, linearity) do not seem to hold in this domain, so one may wonder whether this application domain could be too ambitious (for now).
Probabilistic modeling: Methods and Challenges	Since our original effort in use of probabilistic models for reconstruction of signaling regulation from single cell proteomic data (in 2005), advances have been made by ourselves and others towards improving reconstruction accuracy. I will discuss my perspective on some of the remaining challenges, and describe algorithms that address some of these challenges.
Integrative causal analysis, CAUSALPATH overview	TBA

Tuesday 28/4

Variable selection under constraints with extensions to path analysis	A Bayesian method for consistent model selection in regression under constraints is presented. Its extension to systems of related regression models briefly discussed.
Scaling up the GES and FCI algorithms, using background knowledge, and more	Three topics: (1) Joe Ramsey has scaled up the GES algorithm so it can handle >1M variables, and FCI so it can handle >1K. (2) Preliminary work on generalizing GES to use background information and handle latent variables. (3) Brief mentions of other projects at CMU: learning from undersampled time series; learning endogenous latent structures; multiple indicator model search.
Data management approaches to signal pathways	TBA