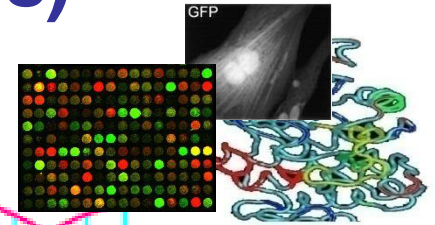
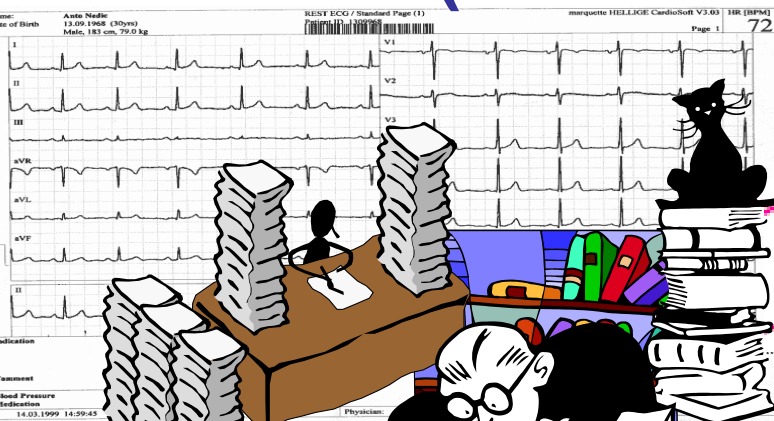


Machine Learning & Probabilistic Graphical Models (with *Scientific Applications*)



MQPEEGTGWLLELLSEVQLQQYFLRLRDDLN
VTRLSHFEYVKNEDLEKIGMGRPGQRRLWEAV
KRRKALCKRKSWSKVFSGKRLEAEFPPHHS
QSTFRKTSAPGGPAGEGPLQSLTCLIGEKDLRL
LEKLGDSFGVVRRGEWDAPSGKTVSVAVKC
LKPVDLSQPEAMDDFIREVNAMHSLDHRNLIR
LYGVVLTTPPMKMVTELAPLGSLLDRLRKHQG
HFLGTLSTRYAVQVAEGMGYLESKR

Hagit Shatkay

Computational Biomedicine & Machine Learning Lab

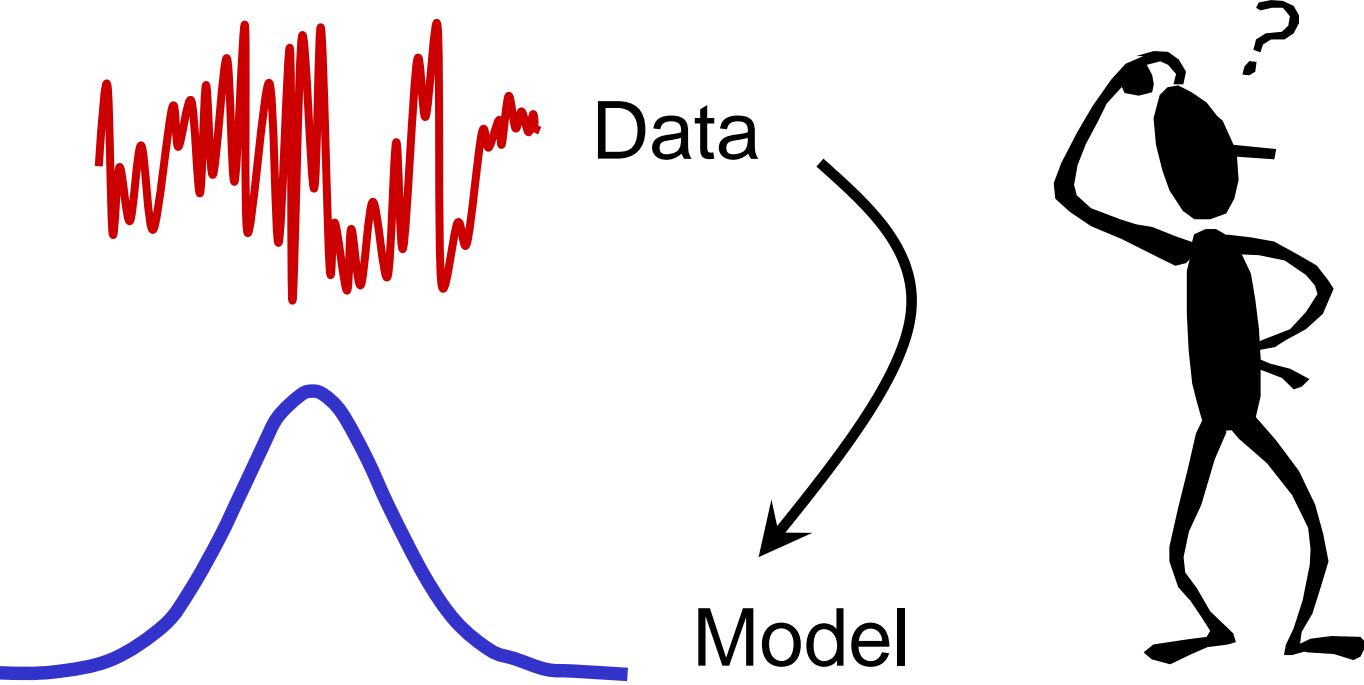
Dept. of Computer and Information Sciences, Dept. of Biomed. Eng.

University of Delaware

Brown CS (PhD – ML, Robotics) → NIH/NCBI (PostDoc) →
Celera Genomics → Queen's, Kingston, ON → U. of Delaware



Learning Models from Data (*Model Fitting*)



Caution:

Machine Learning  **(Deep) Neural Networks**



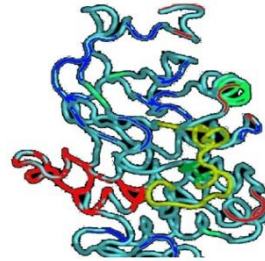
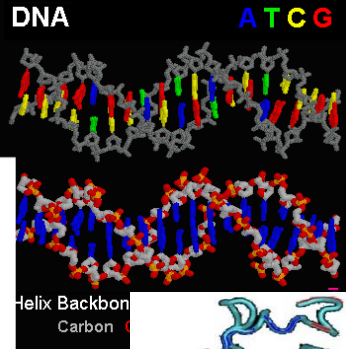
Data



MQPEEGTG
 VKNEE
 FSGKR
 EKDLR
 DVLSQ

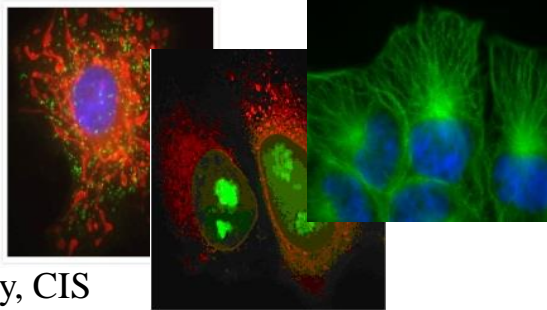
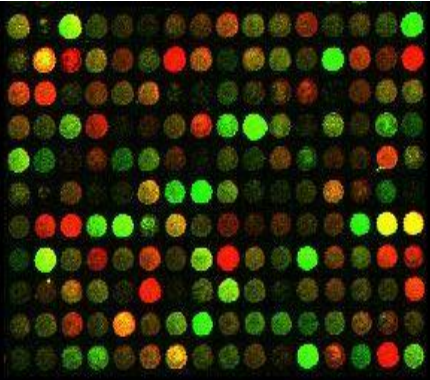
FGVVRRGEV
 EVNAMHSLI

V
 S
 M
 Y
 WAHKPEDRPTFVALRDFLLEAQPDMRALQDEEPDKLHIQ

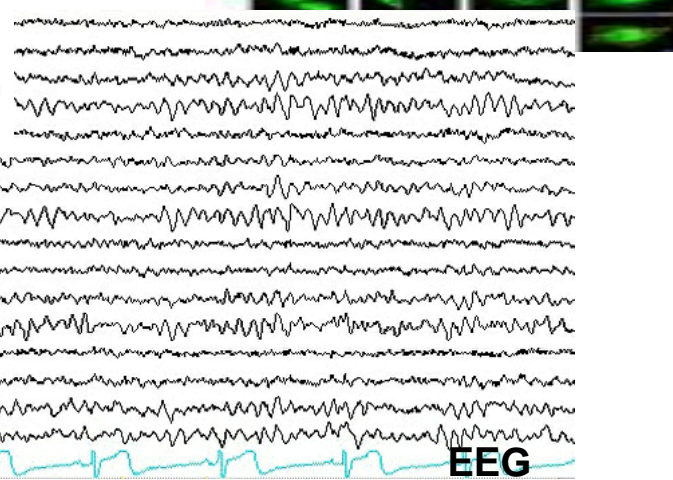


NRTLTCVGPFPNRNVVTSVAGLS
 RHCWGFPDYNVMVQCWAHK
 MRALQDFEEDPKLHIQMNDV

F3-C3
 C3-P3
 P3-D1
 FP2-F4
 F4-C4



<http://www.mcl.tulane.edu/cw/proc/efn.jpg>



Models (*Probabilistic Graphical Models*)

Probabilistic : Encoding Uncertainty via Conditional Probability

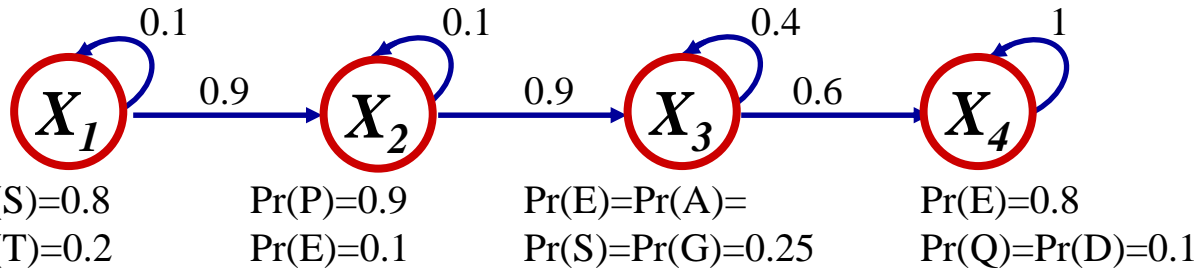
$$\Pr(\mathbf{X} = \mathbf{e}_i \mid \mathbf{Y} = \text{Observations} + \text{Knowledge})$$

Graphical :

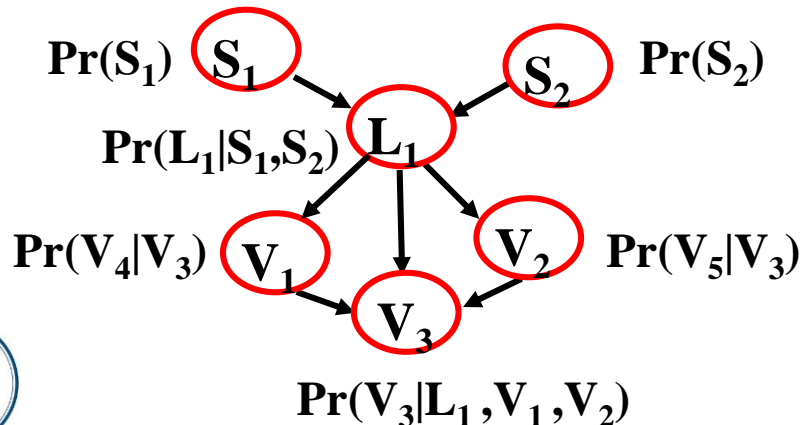
Nodes: Random Variables

Edges: Dependencies

Missing Edges: Independencies!



Hidden Markov Model



Bayesian Network

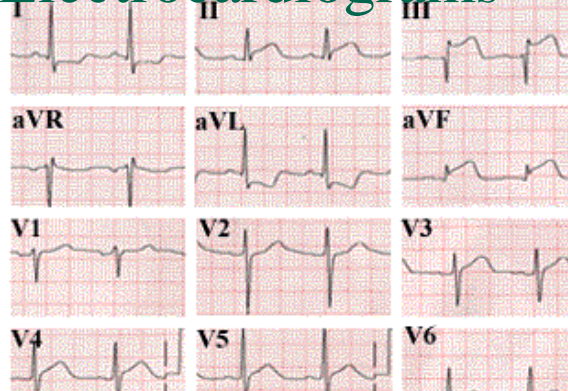


Current Projects: Prediction of Heart Disease (HCM)

(Learning: **Supervised** + **Unsupervised**)

With Drs. T. Abraham & R. Abraham et al (JHU & UCSF)

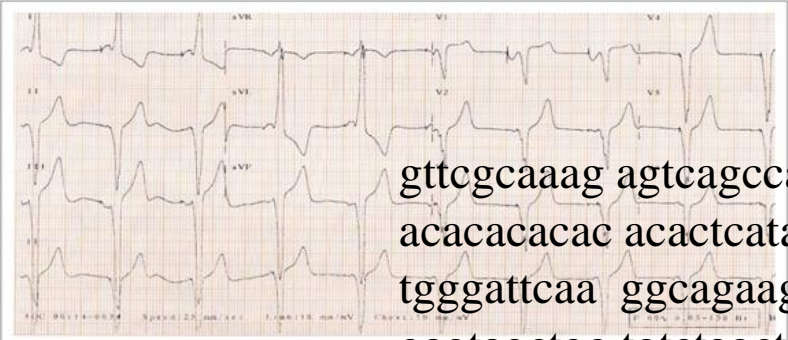
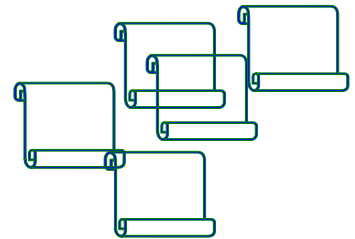
Electrocardiograms



Echo-cardiograms



Questionnaires



gttcgcaaag agtcagcca c gtgtgtacac
 acacacacac acactcata g acataatggg attagttcca
 tgggattcaa ggcagaagt caagtctgtt cactgttgc aggaaatagg
 gagtagetgg tgtetagetat tcattttttg ctctgtttg ctctgtcta gaatttctag
 atctcggtec cagtttggg accctgcttg ttgtctgatg atgtaggtac aaagcctgat
 ggaggggaaa ggtgatgadc aagttctcag gaggtcaggt tcaactcttg ctctgttctg
 ggaggggcac atttacacca gatgcggagt ccgctttgat acctgctgca
 ggagctggga ttctcagga aaccaggggt tggctctgagg ttagagatcc

Fig. 1 - Sinus rhythm, left atrial overload; note the initial QRS complex following alterations in the ventricular repolarization.

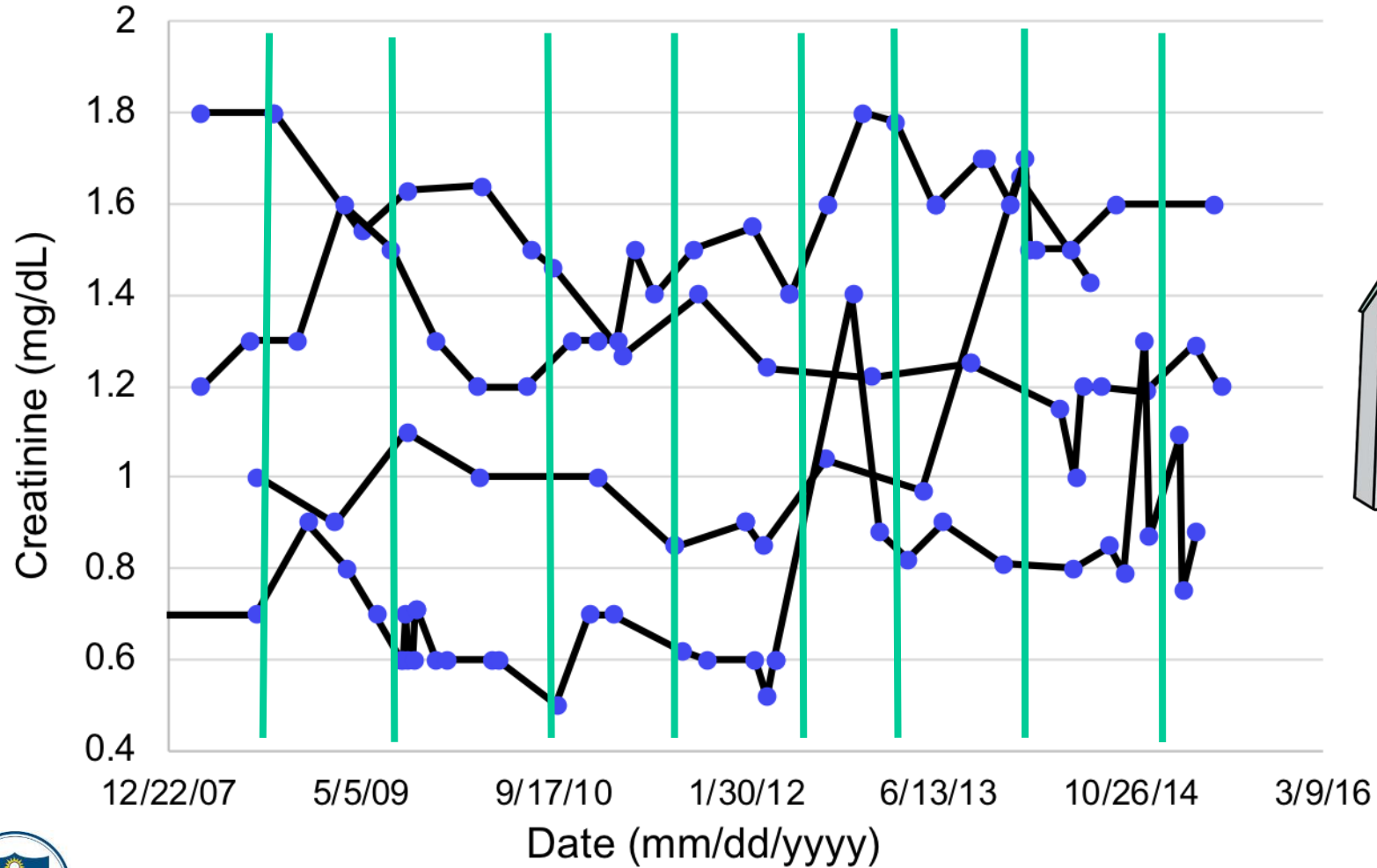


Predicting hospitalization from temporal data

(Learning: **Supervised** + **Unsupervised**)

With Dr. Claudine Jurkovitz et al (Christiana Care)

Creatinine over Time



Computationally Glancing at (Biomedical) Images (Learning: *Supervised* + *Unsupervised*)

Using *Text* and **IMAGE** to find **EVIDENCE**:
Relevant Info about genes proteins or drug-interactions..

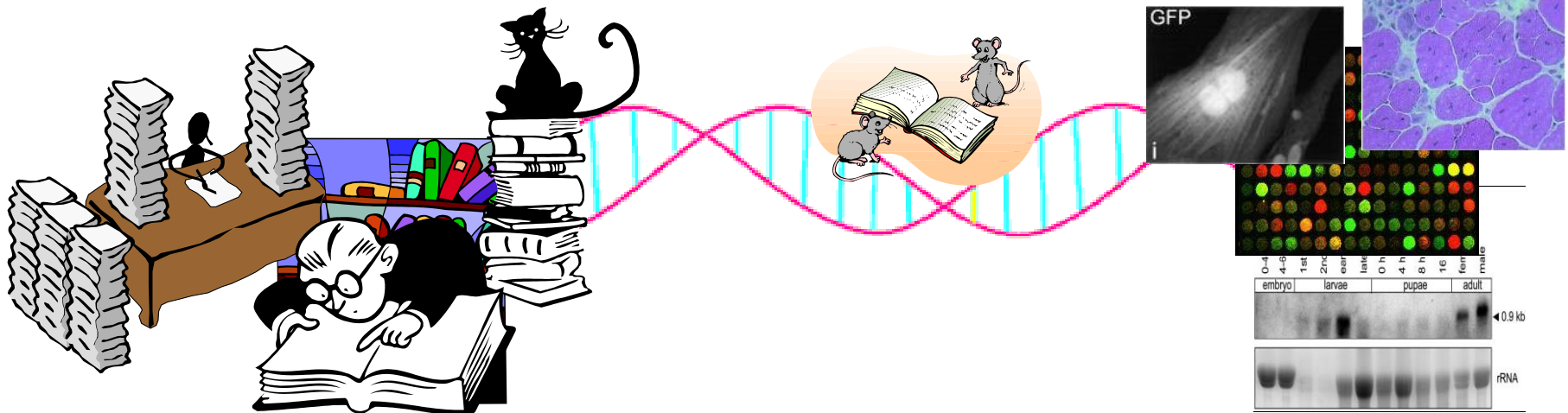
Martin Ringwald, Cynthia Smith, Judy Blake (Jackson Labs)

Xiangying Jiang, Pengyuan Li, Gongbo Zhang (Udel)

Prof. Liz Marai and her group (University of Ill. , Chicago)

Prof. Cahndra Kambhamettu and Vinit Singh (UDel)

Profs. Luis Rocha (Indiana University) and Lang Li (Ohio State)





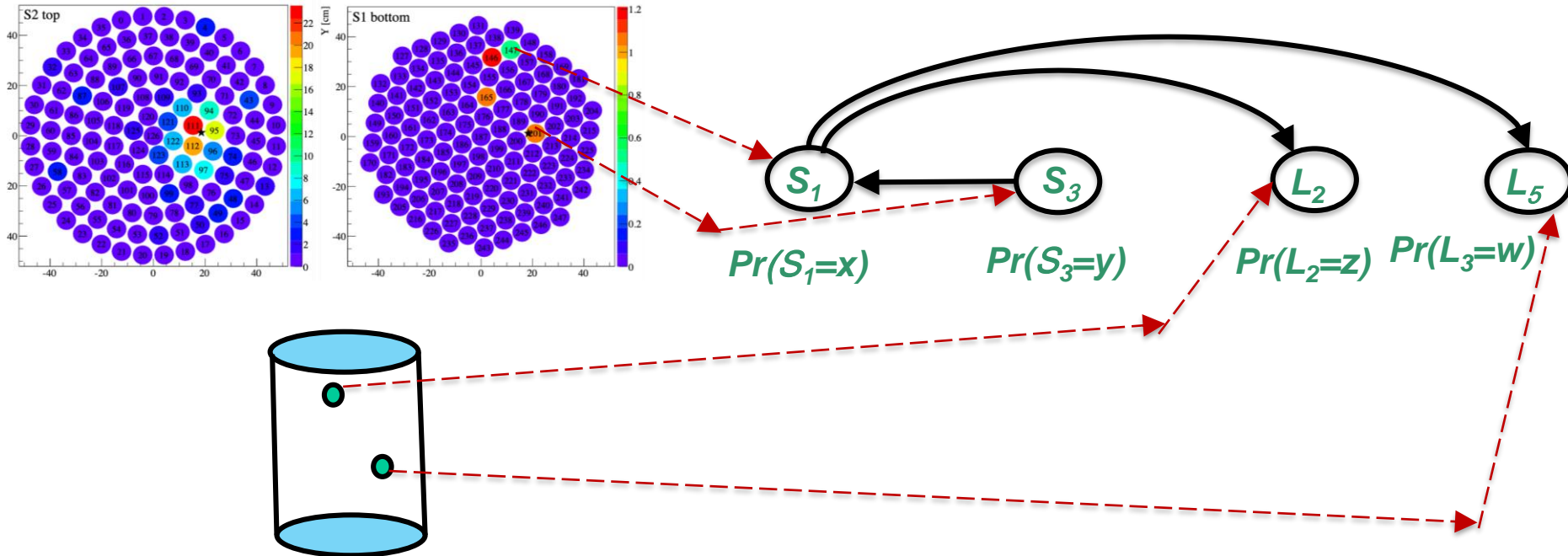
Science-enhanced Machine Learning for Astro-physics Event Detection



With Christopher Tunnell
& Waheed Bajwa

Probabilistic modeling of sensors, events & relationships (Learning: *Quasi Supervised*) **New**

(Sparse and Constrained) Graphical Models



Nodes: Random variables: Sensors and Areas in the Tank

Edges: Indicate interdependency among Sensors/Locations/Events

Introducing Domain Knowledge via:

Priors; Distributions; Interdependency constraints



Thank You

shatkay@cis.udel.edu

<http://www.cis.udel.edu/~shatkay>

