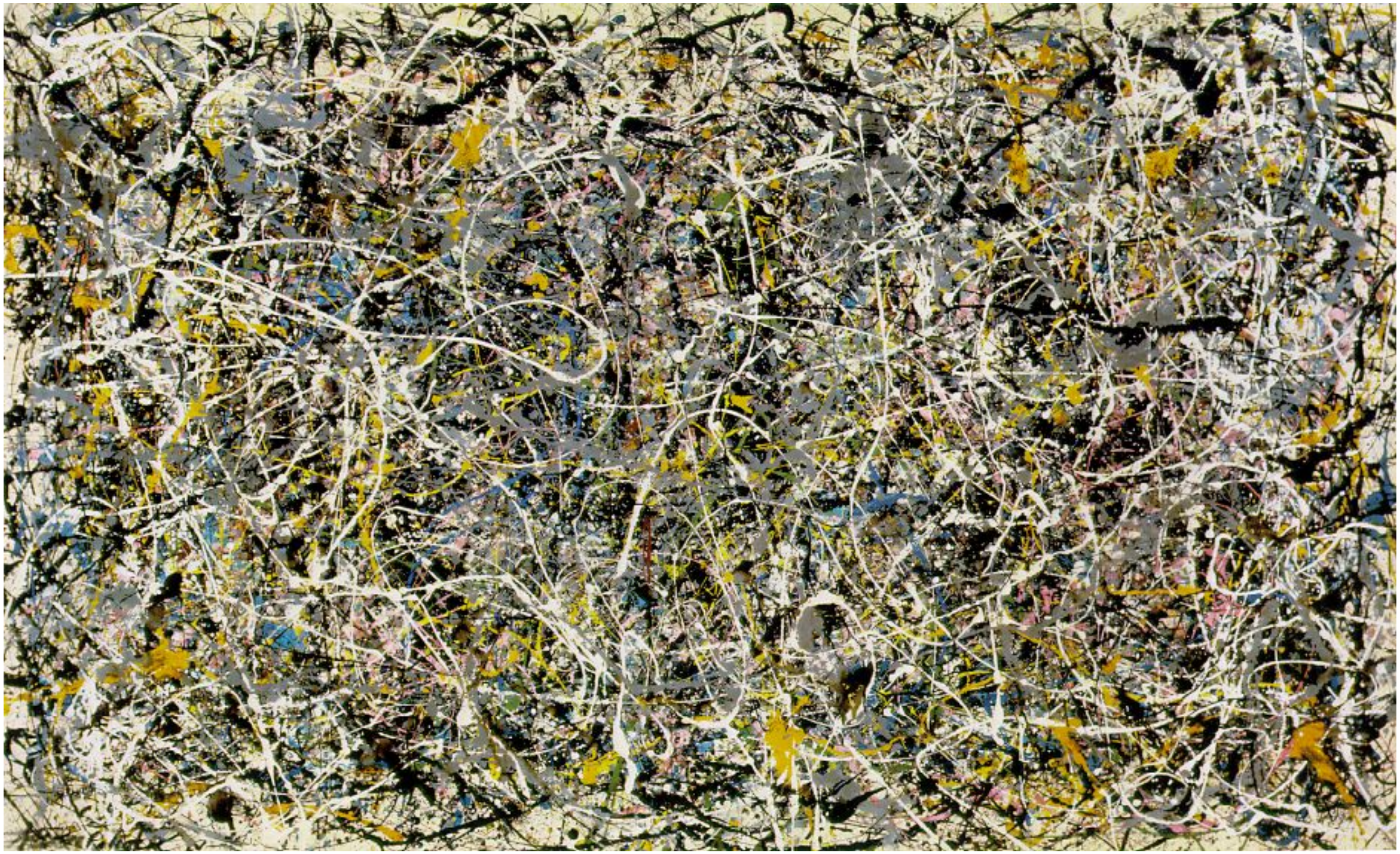
Making Sense of a Complex World

Ken Kreutz-Delgado

Professor of Intelligent Systems & Machine Learning

ECE Department JSOE UCSD

“In the beginning … the Earth was without form … ” (Genesis, KJV)



(Jackson Pollack)

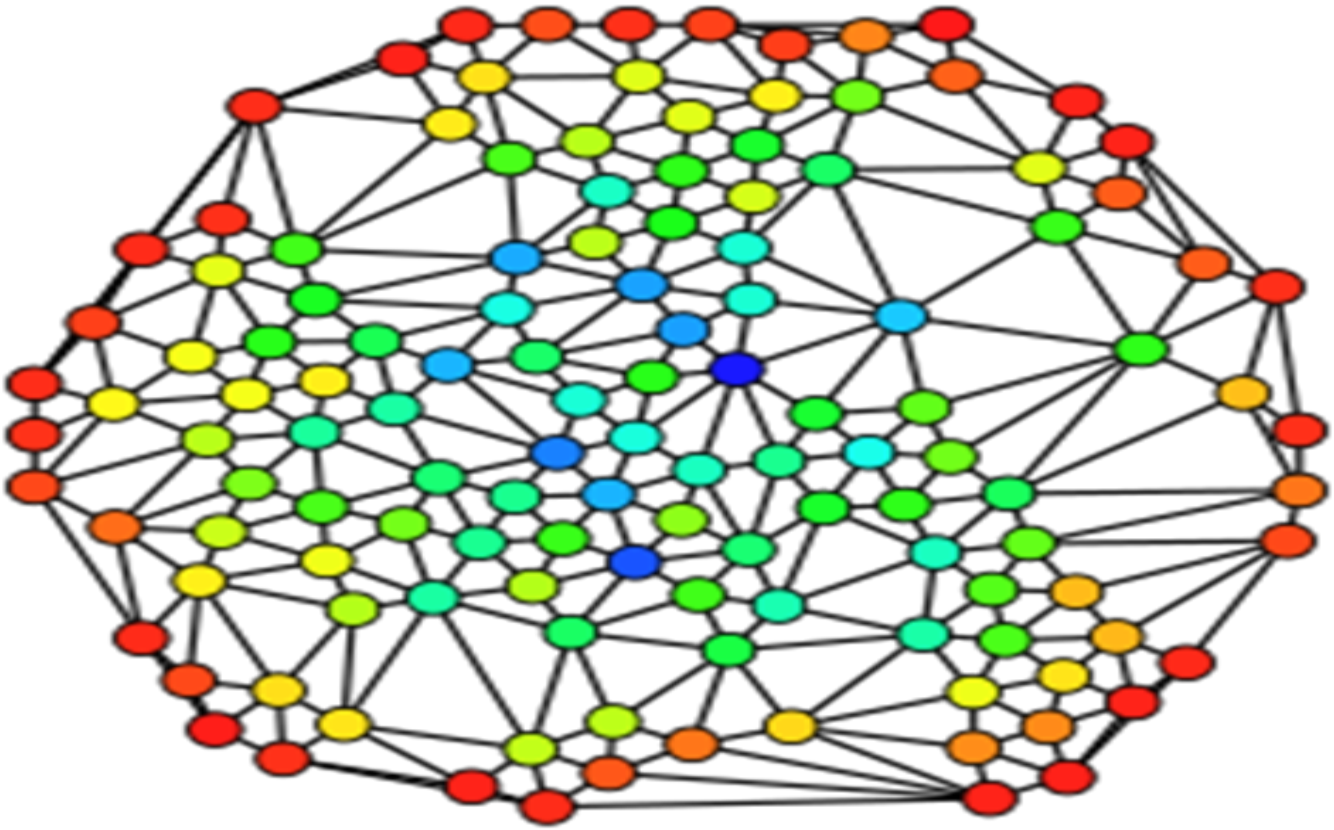
The World *is* Complex – <u>but Structured</u>

Organisms extract "meaning" and discern "order" in the world.

Living things are the end result of an immensely numerous succession of ancestors that have survived and evolved to successfully discern and exploit structure in the world.

There must be structure and regularities in the world that over evolutionary time we have become "attuned to" – an "*evolutionary a priori*" [Wuketits 1990].

The World – How to discern structure and manage its complexity?



- Identify and exploit relationships and connections (graph structure)
- Determine utility (importance), likelihoods & causal effects (probabilities)

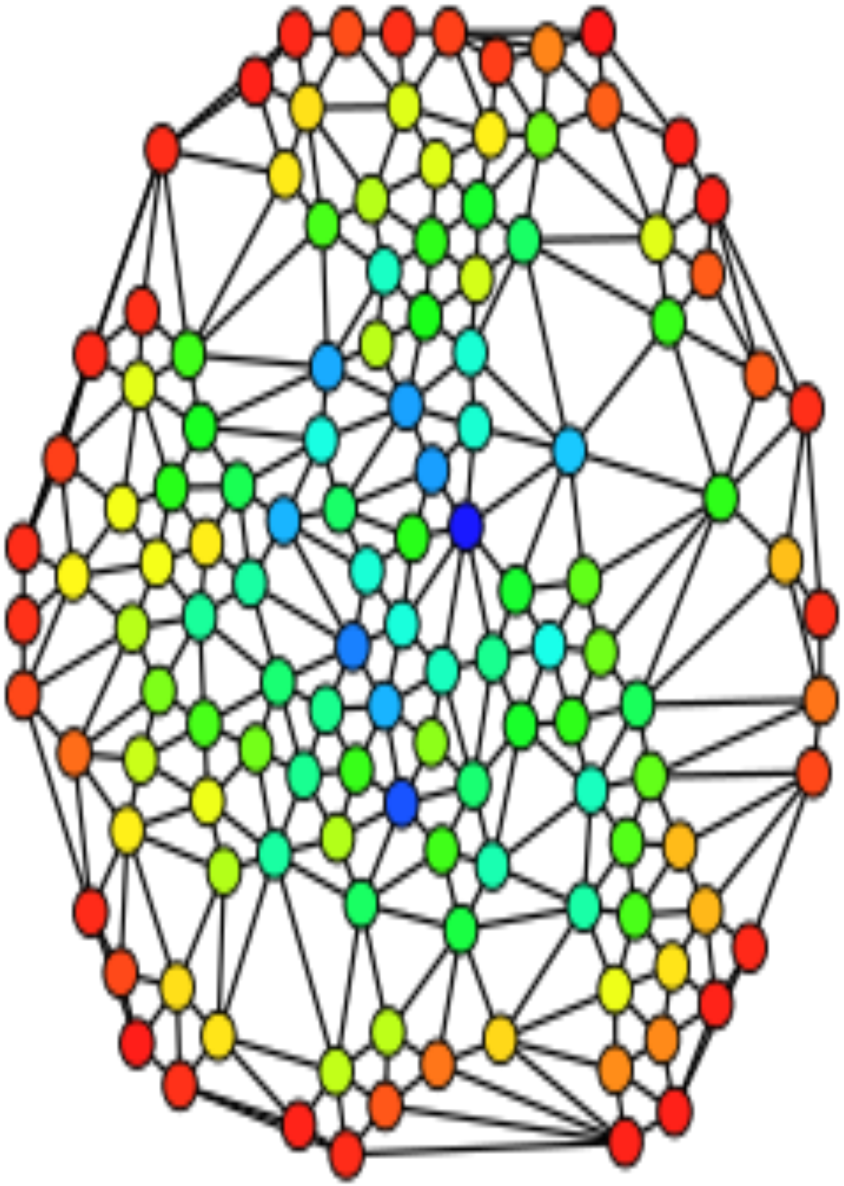
Encoding Information About the World

- Code information either extensionally or Intensionally
  - Intension versus Extension

- Extensional coding explicitly lists *instantiated events, things, and facts* about the world
  - Difficult to add information in a consistent manner
  - Difficult to reason with (uses logic-based reasoning).
  - Number of instantiated facts to encode is huge

- Intensional coding encodes *relationships and possible states of affairs* (propositions) about events, things, facts and dependencies
  - Easy to expand
  - Easy to visualize graphically
  - Encodes facts about the world *implicitly* not explicitly

Complexity of a Fully Interrelated World



The world can be modeled as interrelated "things + attributes" that occur or co-occur with certain probabilities.

Thus we need to learn what "things" exist and their "states", singly and collectively. We can think of a "thing" in a given "state" as denoting a situation ***k***, where x_k ***= 1*** or ***-1*** depending on that situation either being the case or not being the case.

Therefore we can model the world graphically and The state of the world probabilistically $p(x_1, \cdots, x_n)$

However general, a fully connected world is too complex to handle. If there are *n* situations In the world then the number of independent probability values to specify is $\mathbf{2^n - 1}$

For example **if n = 300**, then the number of probability values to specify is $\mathbf{2^{300}}$ **»** $\mathbf{10^{90}}$, a value larger than the number of electrons, protons, and neutrons estimated to exist In the entire known universe …

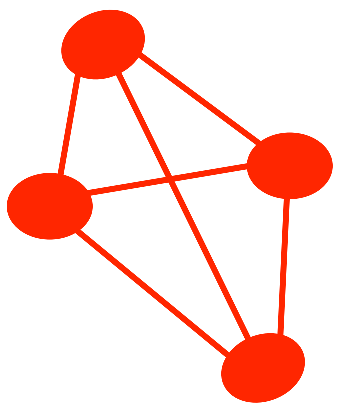
The World <u>must</u> have Exploitable Structure

- "Things" have causal and influential interactions
  - Probabilistic relationships are often *Markovian*
- Directed interactions, generally limited or localized
  - Interconnectivity is limited, directed and structured

This suggests *"…that the fundamental structure of human knowledge can be represented by dependency graphs and that mental tracing of links in these graphs are the basic steps in querying and updating that knowledge"* [Pearl 1986].

This shows why conditional Independences are Important

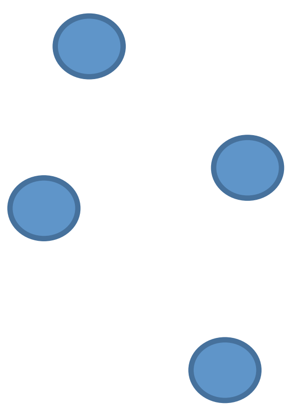
E.g: n = 100 and nodes all Boolean (0-1)



Fully dependent/connected world:

2^{100} » 10^{30} probability values

This extreme is too complex!



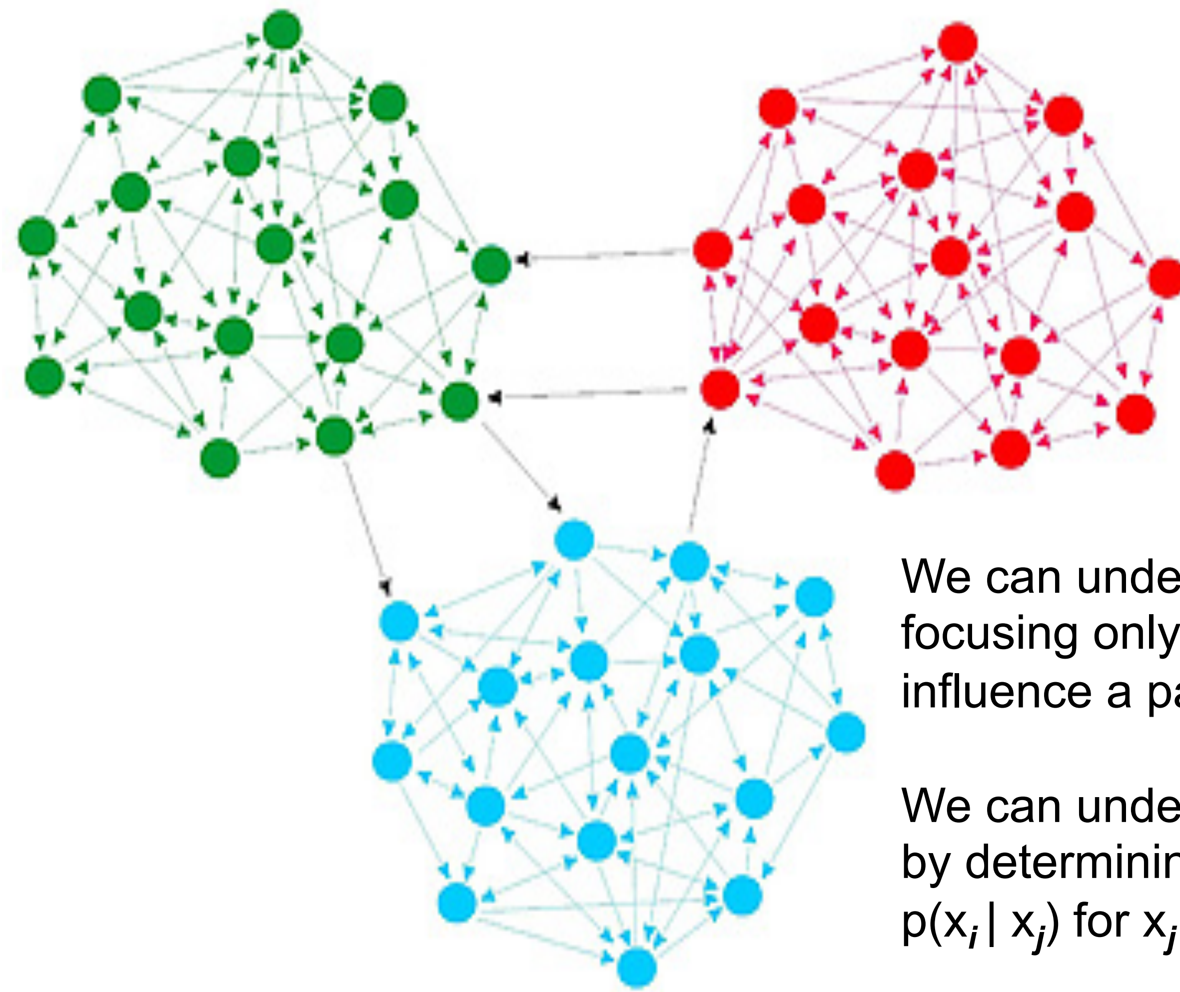
Fully independent/disconnected world:

100 – 1 = 99 probability values

This extreme is too simple!

The middle ground of “sparse” connectivity and exploiting conditional dependencies can be just right.

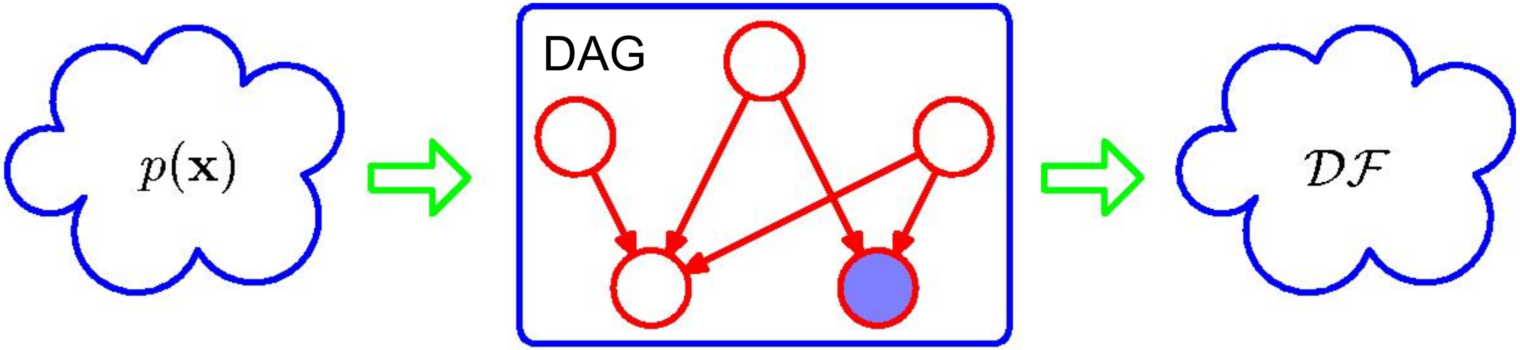
Example: DGs & Markovian Structure



We can understand a ***Directed Graph*** by focusing only on the nodes that casually influence a particular node of interest.

We can understand its ***Markovian Structure*** by determining the ***transition probabilities*** $p(x_i \mid x_j)$ for x_j given x_i

PGMs as "Distribution Filters"



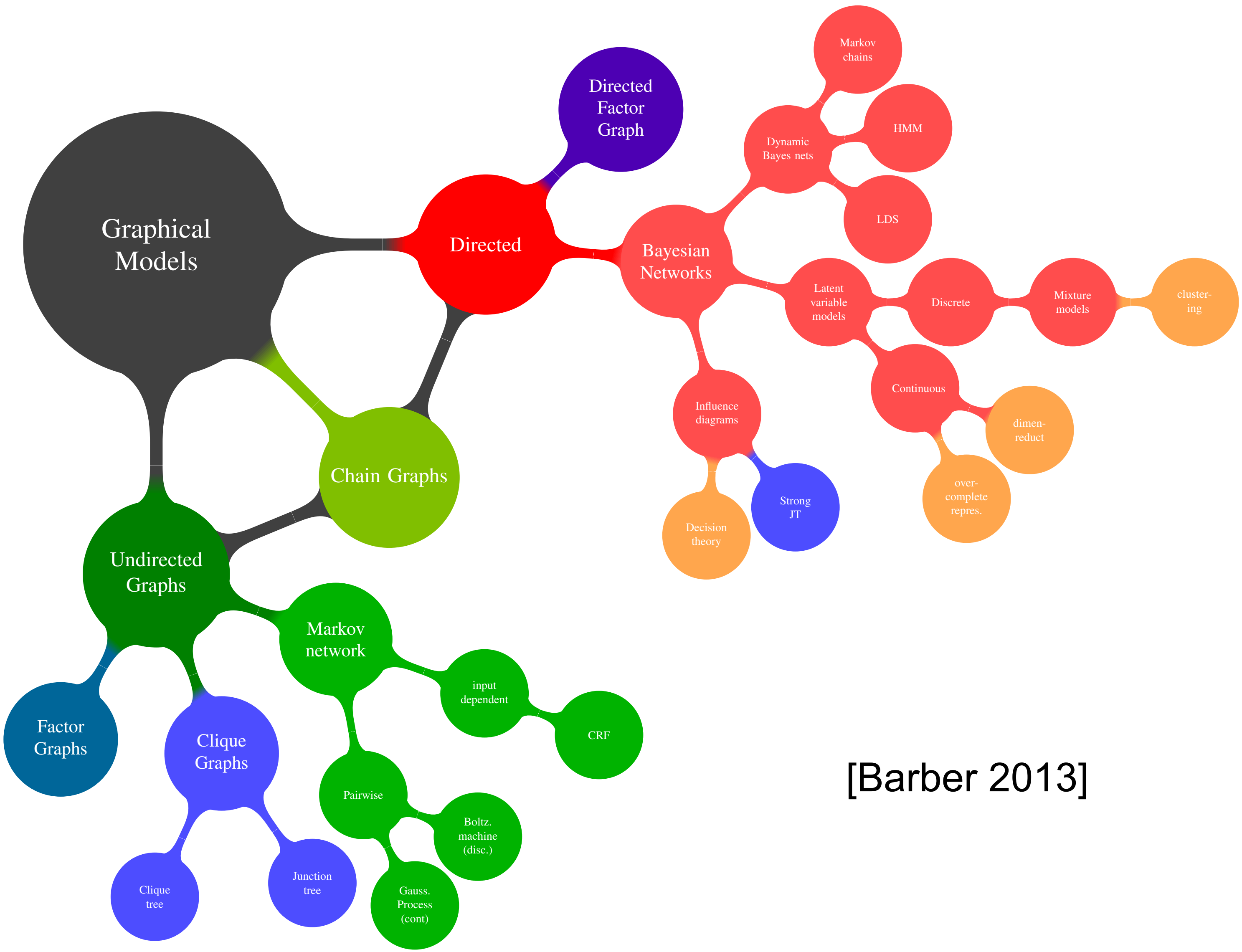
[Bishop 2006]

Probabilistic Graphical Models (PGMs) impose strong constraints which explains why using them can be effective, even if the node conditional probabilities are not accurately or precisely known, as long as the node conditional probabilities are qualitatively and comparatively reasonable [Pearl 1986,1988].

> "This suggests that the notions of dependence and conditional dependence are more basic to human reasoning than are the numerical values attached to probability judgments" [Pearl 1986].

Directed <u>Acyclic</u> Graphs (DAGs) are among the most tractable PGMs, but still have limitations. E.g., they do <u>**not**</u> admit feedback (aka *reentry* or *reverberation*).

Taxonomy of PGMs



[Barber 2013]

References

[Barber 2012] D. Barber, ***Bayesian Reasoning & Machine Learning***, CUP, 2012.

[Bishop 2006] C. Bishop, ***Pattern Recognition and Machine Learning***, Springer, 2006.

[Koller 2009] D. Koller & N. Friedman, ***Probabilistic Graphical Models***, MIT, 2009.

[Murphy 2012] K. Murphy, ***Machine Learning: a Probabilistic Perspective***, MIT, 2012.

[Pearl 1986] J. Pearl, **"Fusion, Propagation, and Structuring in Belief Networks",** *Artificial Intelligence*, **29**:241-288.

[Pearl 1988] J. Pearl, ***Probabilistic Reasoning in Intelligent Systems,*** Morgan-Kaufmann, 1988.

[Russell 2003] S. Russell & P. Norvig, ***Artificial Intelligence: a Modern Approach***, 2e, Prentice-Hall, 2003.

[Wuketits 1990] F. Wuketits, ***Evolutionary Epistemology and its Implications for Humankind,*** SUNY Press, 1990.