Stochastic And-Or Grammars: A Unified Framework and Logic Perspective

Kewei Tu (ShanghaiTech University, China)

Representation Framework

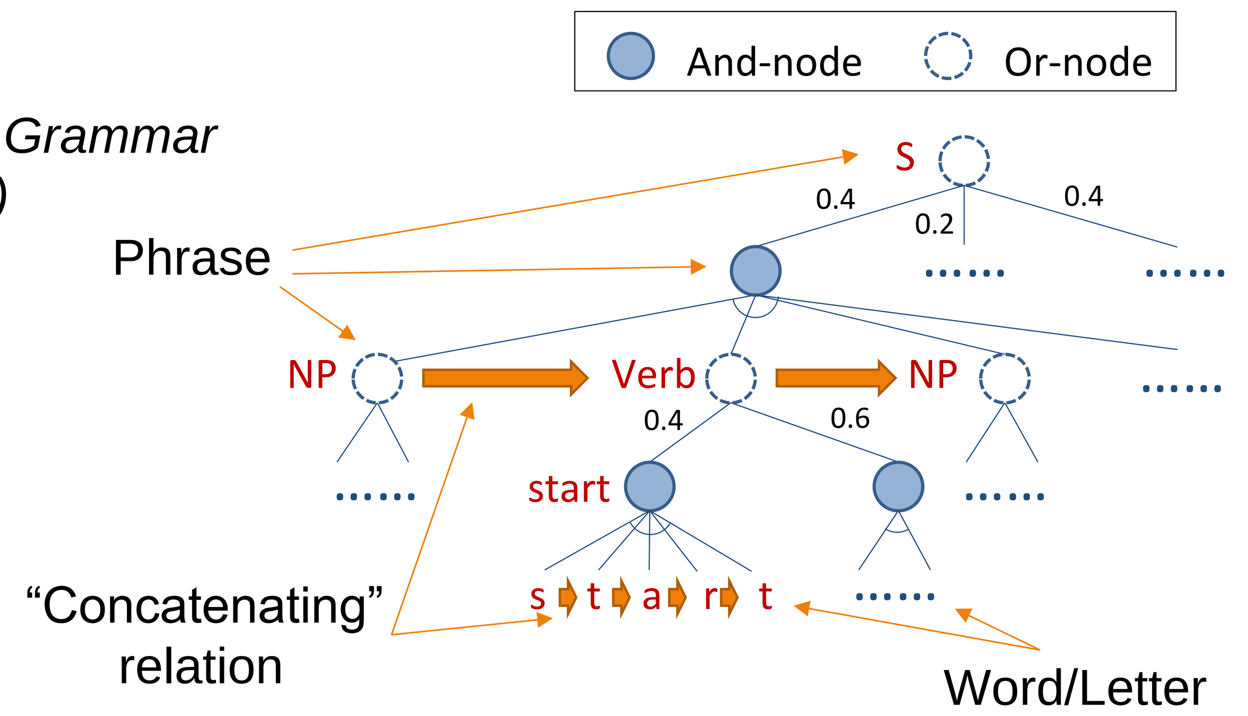
Stochastic And-Or grammars (AOG) are an extension of stochastic grammars of language, originally proposed to model images [Zhu & Mumford, 2006]. We propose a unified representation framework of stochastic AOGs that is agnostic to the type of the data being modeled.

Definition. A *stochastic context-free AOG* is a 5-tuple $\langle \Sigma, N, S, \theta, R \rangle$

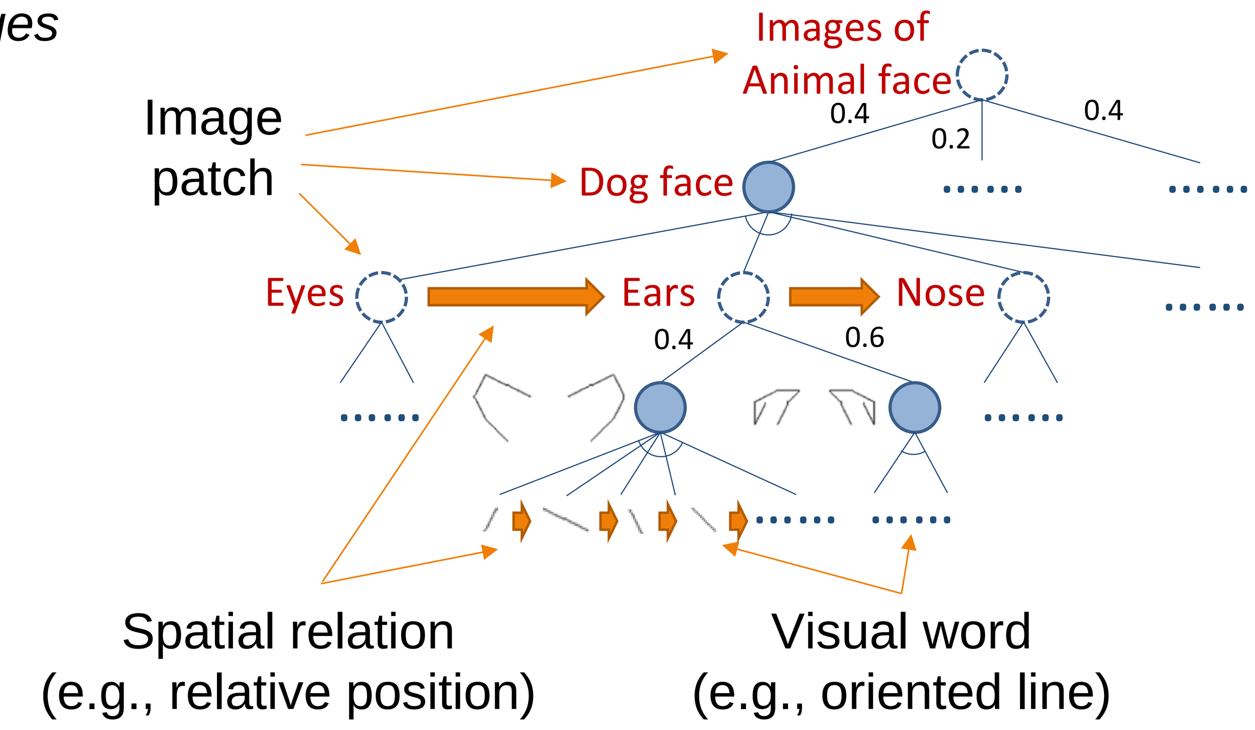
- Σ is a set of terminal nodes representing atomic patterns
- N is a set of nonterminal nodes representing high-level patterns
  - Two disjoint subsets: And-nodes, Or-nodes
- $S \in N$ is a start symbol representing a complete pattern
- θ is a function that maps an instance of a terminal/nonterminal x to a parameter θ_x
- R is a set of production rules
  - And-rule $\langle r, t, f \rangle$
    $r: A \rightarrow \{x_1, x_2, \dots, x_n\}$ for some $n \geq 2$
    $t(\theta_{x_1}, \theta_{x_2}, \dots, \theta_{x_n})$ is a parameter relation between child nodes
    $\theta_A = f(\theta_{x_1}, \theta_{x_2}, \dots, \theta_{x_n})$ is a parameter function
  - Or-rule $\langle r, p \rangle$
    $r:\ O \rightarrow x$
    p is the probability of O producing x

Examples

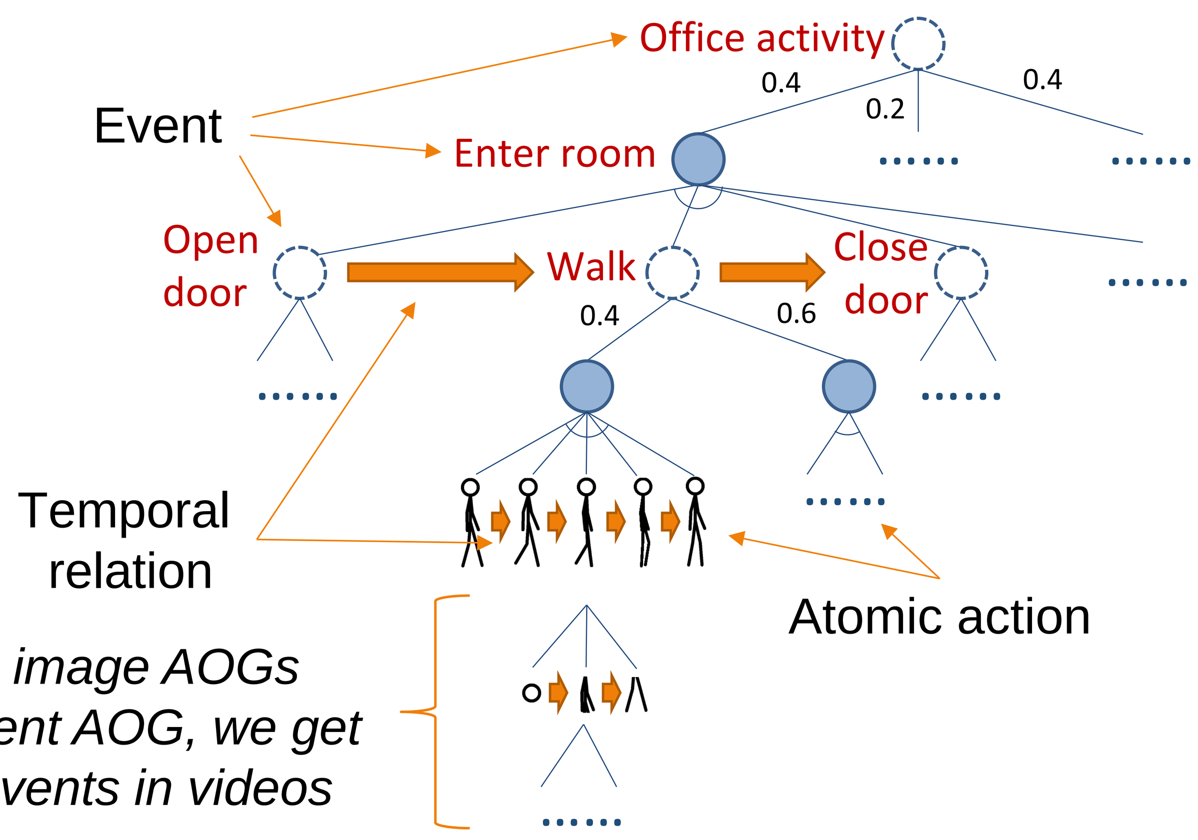
Context-free Grammar (of language)



AOG of Images



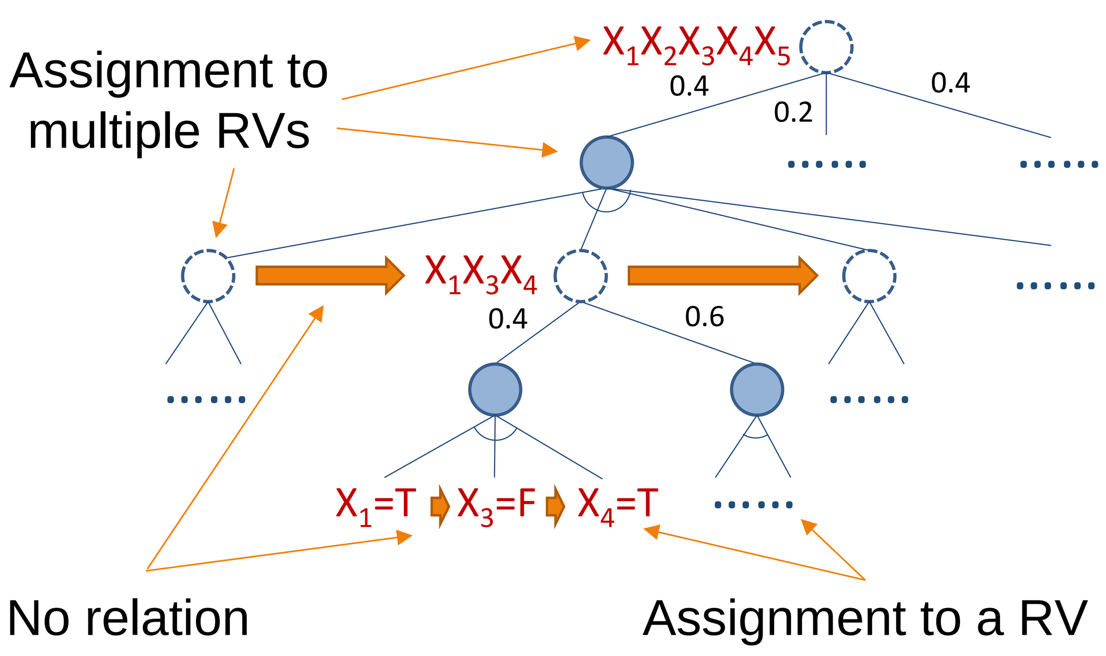
AOG of Events



By attaching image AOGs below an event AOG, we get an AOG of events in videos

Decomposable Sum-Product Network

(with additional constraints on the AOG structure)



Special cases of stochastic AOGs

Natural Language Processing	Computer Vision	Machine Learning
• Stochastic Context-free Grammars • Hidden Markov Models • Linear Context-free Rewriting Systems • Constraint-based Grammar Formalisms • etc.	• And-Or Graphs • Pictorial Structures Models • Deformable Part Models • Flexible Mixture-of-Parts Models • etc.	• Sum-Product Networks • Naïve Bayes • Biclustering • Thin Junction Trees • Mixtures of Trees • Latent Tree Models • etc.

Inference

The main inference problems:

- Infer the most likely compositional structure of a data sample (parsing)
- Compute the marginal probability of a data sample

Both are NP-hard! But we propose the following condition that is reasonable in many scenarios (e.g., in modeling text or images) and leads to tractable inference.

(Composition Sparsity) *For any data sample X, the number of valid compositions in X is polynomial in |X|.*

Exact parsing can be done by bottom-up dynamic programming. Algorithm sketch:

- Input: a stochastic context-free AOG in generalized Chomsky normal form, a data sample consisting of instances of terminal nodes
- Regard terminal instances as size 1 compositions
- For $i = 2, \dots, |X|$ and $j = 1, \dots, i-1$
  - For any pair of existing compositions of size j and $i-j$
    - If they can be combined using a production rule of the AOG, then store the new composition of size i and compute its probability (unless the composition already exists, then we shall update its probability)
- From the composition of size $|X|$ and root S, backtrack to construct the parse tree

Logic Interpretations

Stochastic AOGs can be seen as probabilistic models of relations and compositions. We provide logic interpretations of stochastic context-free AOGs to connect them to the field of statistical relational learning.

Interpretation as a subset of first-order probabilistic logic

There are two types of formulas in the logic:

And-rule:

$$\forall x \exists y_1, y_2, \dots, y_n, A(x) \rightarrow \bigwedge_{i=1}^{n} \big(B_i(y_i) \wedge R_i(x, y_i)\big) \wedge R_\theta\big(\theta(x), \theta(y_1), \theta(y_2), \dots, \theta(y_n)\big)$$

Or-rule:

$$\forall x, A(x) \rightarrow B(x)\ : p$$

(with additional constraints that for each true grounding of $A(x)$, among all the grounded Or-rules with $A(x)$ as the left-hand side, exactly one is true)

In our possible world semantics, each possible world contains exactly a parse tree structure in which each node is an object with one or more unary relations, each edge is a binary relation, and all the relations conform to the And-rules and Or-rules. The probability of a possible world is the product of probabilities of the Or-rules that are grounded in the parse tree.

This logic interpretation of stochastic context-free AOGs resembles **tractable Markov logic** in many aspects, implying a deep connection between the two and pointing to a potential research direction of investigating novel tractable statistical relational models by borrowing ideas from the stochastic grammar literature.

Interpretation as a stochastic logic program

And-rule:

$$1.0\colon\ a(X,P)\ \text{:-}\ b_1(X_1,P_1), b_2(X_2,P_2), \cdots, b_n(X_n,P_n), append([X_1,\dots,X_n],X), r_1(X,X_1), r_2(X,X_2), \cdots, r_n(X,X_n), r_\theta(P,P_1,\dots,P_n).$$

Or-rule with nonterminal right-hand side:

$$p\colon\ a(X,P)\ \text{:-}\ b(X,P).$$

Or-rule with terminal right-hand side:

$$p\colon\ a([t],[\dots]).$$

Goal of the program:

$$\text{:-}\ s(X,P).$$

Scan to read the paper