Learning Semantic Maps with Topological Spatial Relations Using Graph-Structured Sum-Product Networks

Kaiyu ZHENG, Andrzej PRONOBIS, Rajesh RAO

UNIVERSITY of WASHINGTON

I. Overview

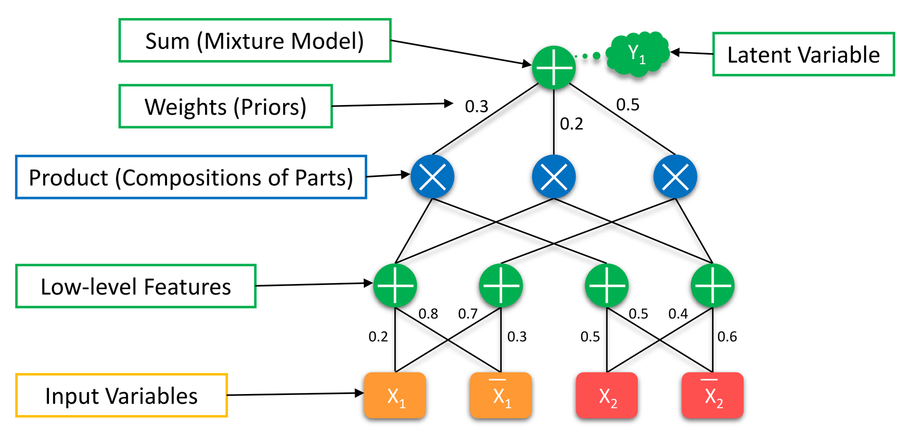
Motivation

- Real-world graph-structured data:
  - Complex, noisy, and dynamic (of varying size)
  - Example: topological graphs built from robot sensory data
- Yet, traditional structured-prediction:
  - Places strict constraints on variable interactions
  - Requires fixed number of variables
  - Requires static global structure

Contributions

- Graph-Structured Sum-Product Networks (**GraphSPNs**):
  - Learn **deep probabilistic models of graph-structured data**
  - Capture complex, noisy variable dependencies
  - Handle dynamic graphs with varying number of variables
  - Leverage Sum-Product Networks (SPNs)
- Learned models of **global semantic maps with topological spatial relations**
  - Disambiguate uncertain semantics based on noisy spatial relations
  - Infer semantic descriptions for unexplored places
  - Detect novel environment structure

II. Sum-Product Networks (SPNs)

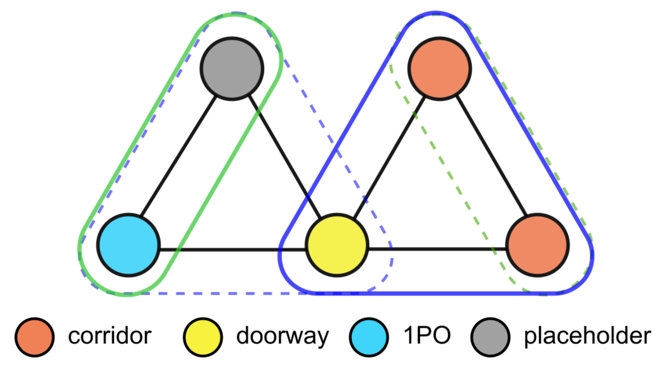


- New **deep probabilistic** architecture with solid theoretical foundations (Poon&Domingos UAI'11)
- Can be viewed as: deep architecture and graphical model
- Learn conditional or joint distributions
- **Tractable** partition function, exact inference
- **Structure semantics**: hierarchical mixture of parts

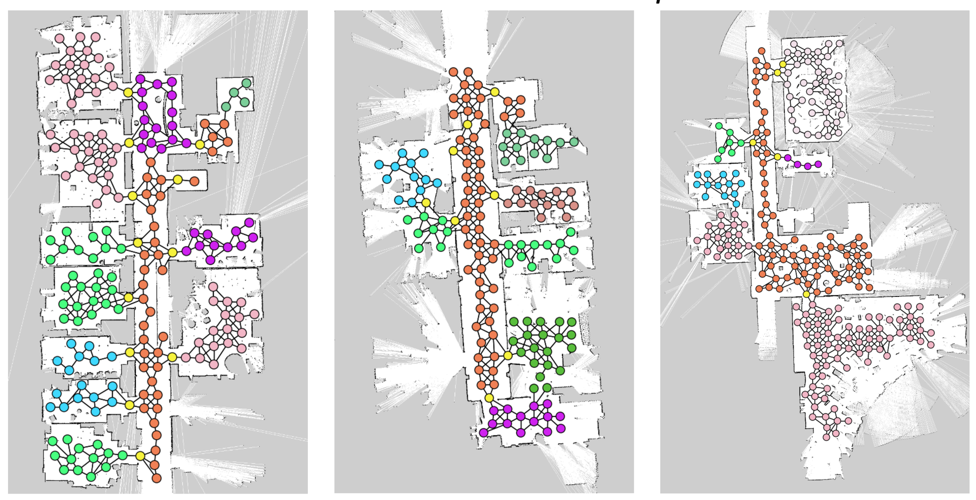
III. Semantic Maps

- Topological graphs anchoring local semantic information
- Dynamic: expands during world exploration
- **Nodes** represent places
  – with local semantic evidence
- **Placeholders** are unexplored places
  – with no evidence
- **Edges** indicate navigability & spatial relations

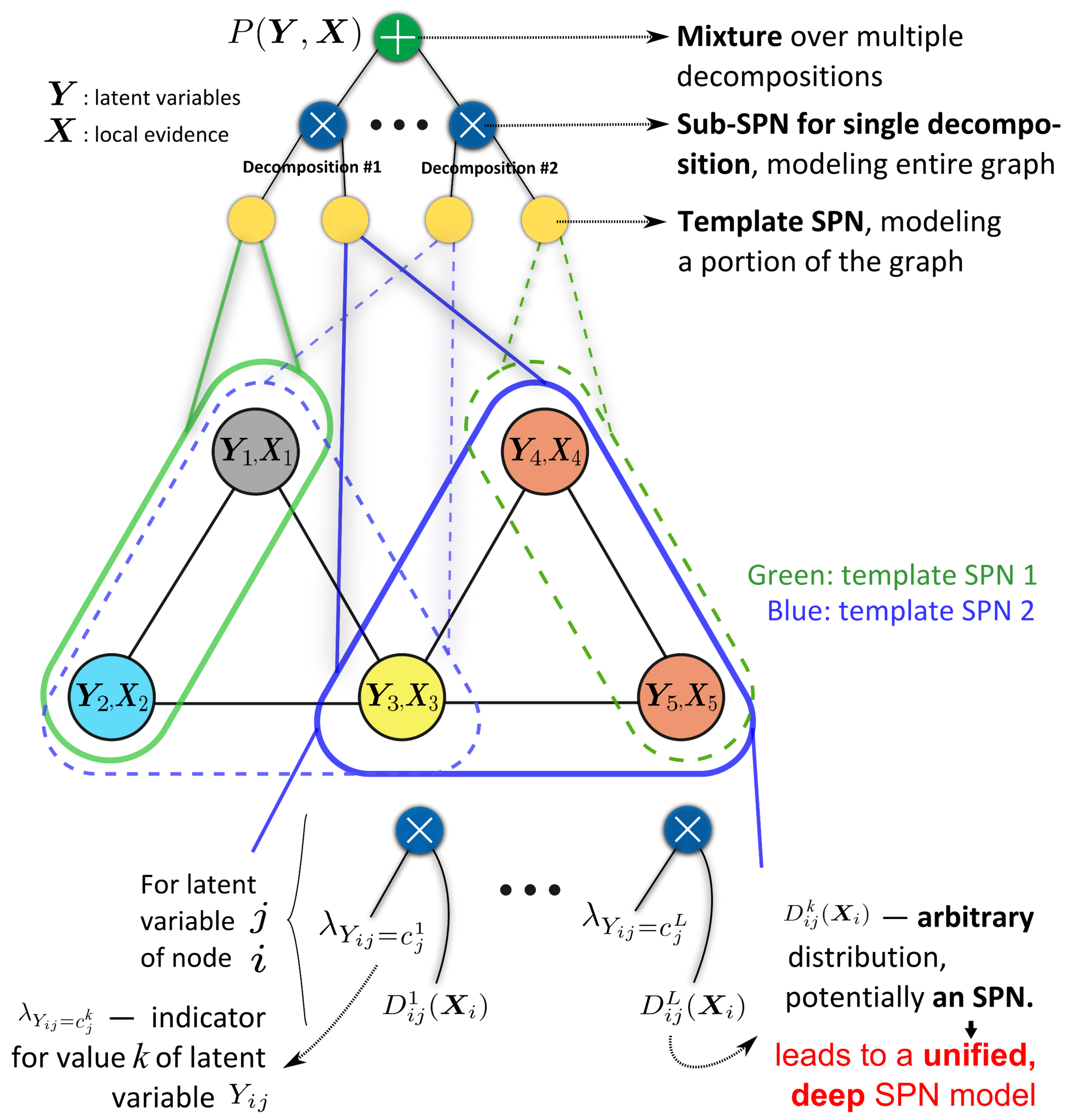
Example for a doorway connecting 2 rooms



Real-world semantic maps



IV. GraphSPNs



V. Experimental Setup

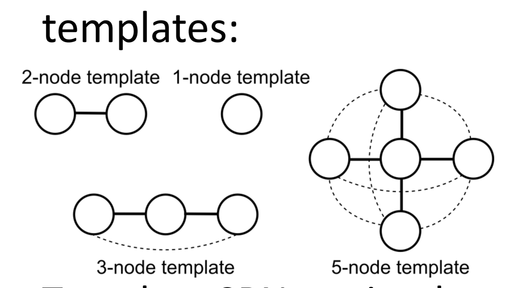
Dataset

- **99 topological graphs** on **11 floors** of **3 buildings** in different cities:
  - Freiburg, Germany
  - Saarbrücken, Germany
  - Stockholm, Sweden
- **10 semantic classes** per place
- Each node associated with one latent variable Y_i (semantic class)
- Introducing **noise**:
  - **20%** of nodes associated with incorrect evidence
  - Varying levels of uncertainty about semantic information
- $D_i^k(\boldsymbol{X}_i)$ defined over **single hypothetical binary observation** x_i (assume observed):

$$D_i^k(X_i) = \begin{cases} \alpha_i^k & X_i = x_i \\ 1-\alpha_i^k & X_i = \bar{x}_i \end{cases}$$

GraphSPN

- Graphs partitioned using templates:

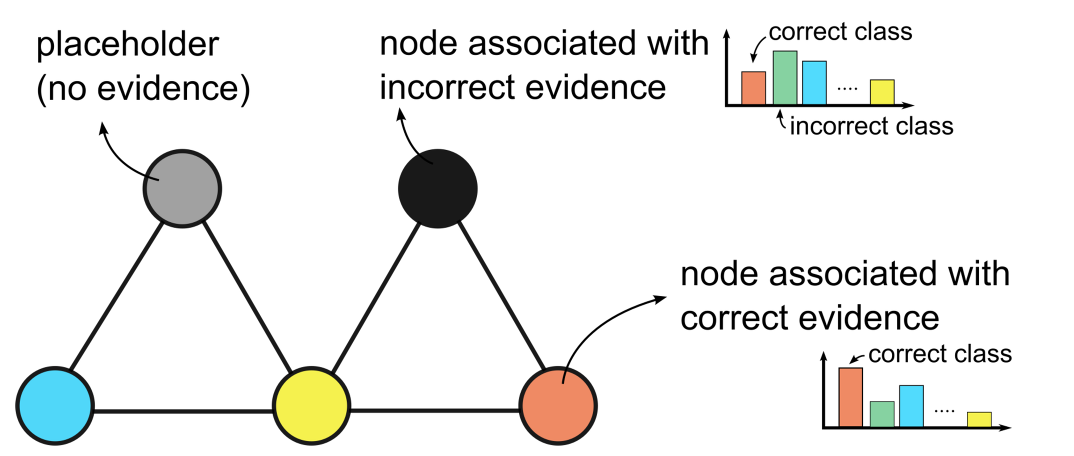


- Template SPNs trained on corresponding sub-graphs
- 5 decompositions used

Makov Random Fields

- MRF structure follows graph structure
- MRF-2: pairwise potentials
- MRF-3: 3-variable potentials
- Local evidence: $\phi_i(Y_i = c^k) = \alpha_i^k$

Example of noisified semantic map



Experimental Procedure

- **Learning**: all graphs from two buildings
- **Testing**: graphs from remaining building with various levels of noise
- **Marginal inference**: $\text{argmax}_k P(Y_i = c^k | \boldsymbol{X} = \boldsymbol{x})$

VI. Experimental Results

Experiments

***#1:* Disambiguate Semantic Info**

- Noisified graphs
- No placeholders
- Accuracy = percent of correctly classified nodes in the graphs

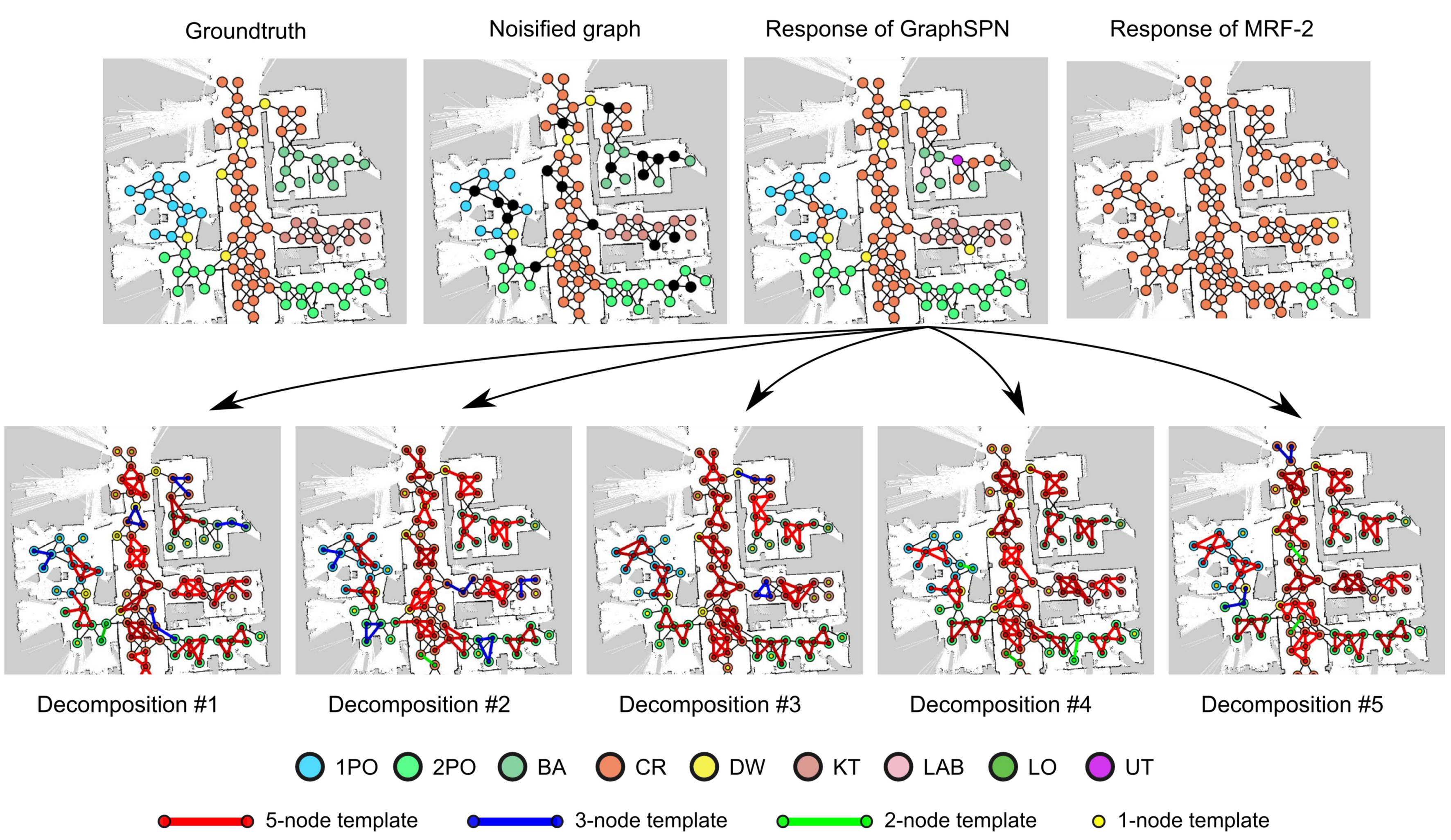
***#2:* Infer Placeholder Semantics**

- Noisified graphs
- With placeholders
- Accuracy = percent of correctly classified placeholders in the graphs

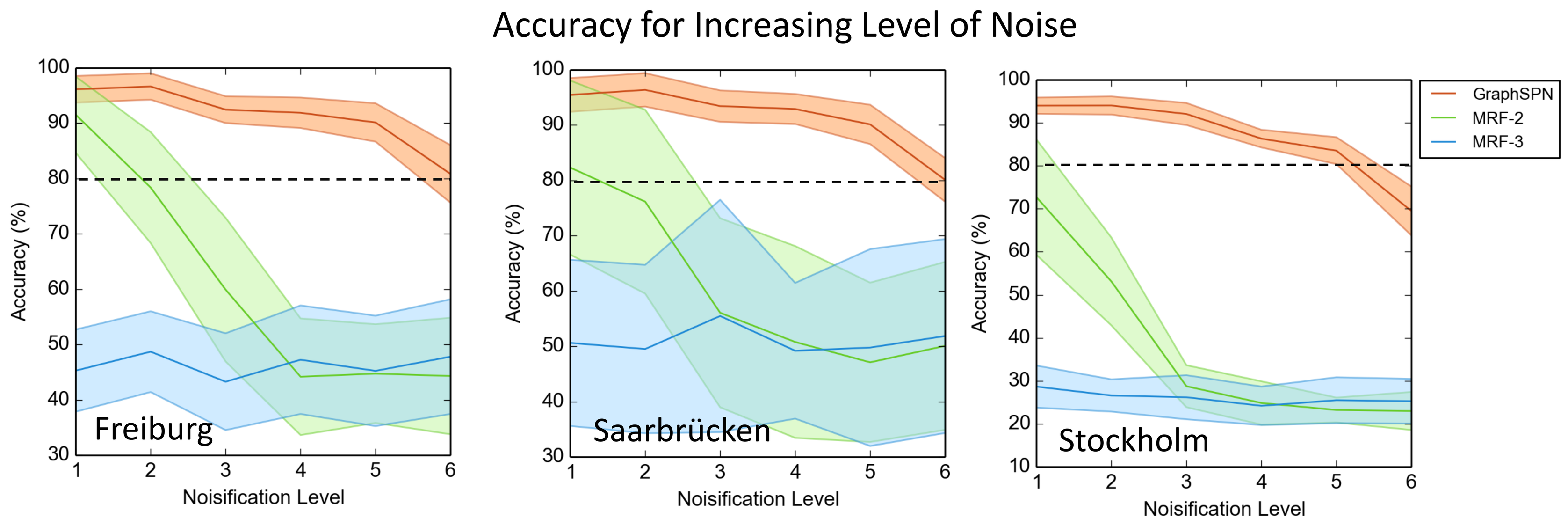
***#3:* Novel Structure Detection**

- No noisification ($\boldsymbol{X} = \emptyset$)
- No placeholders
- Simulating world structure changes by swapping evidence
  - DW and CR, CR and 1PO **(novel)**
  - 1PO and 2PO **(normal)**
- Structure is novel if: $P(\boldsymbol{Y} = \boldsymbol{y}) < threshold$
- Novelty detection by thresholding likelihood normalized by graph size

#1 Disambiguating Semantic Info

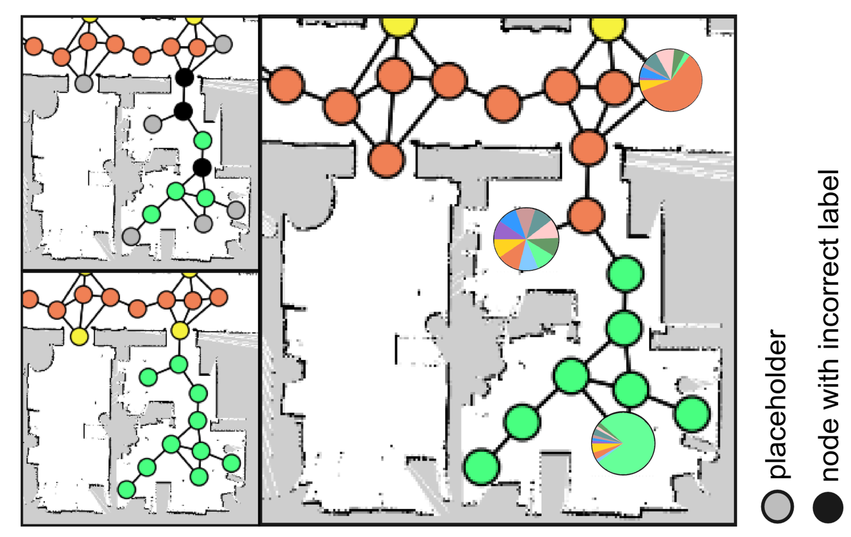


Accuracy for Increasing Level of Noise



#2 Infer placeholder semantics

Marginal Inference Over Placeholder Class



Accuracy for Different Levels of Noise

GraphSPN

NL	Freiburg	Saarbrücken	Stockholm
2	67.58%(+/-10.42)	78.15%(+/-9.95)	67.57%(+/-11.11)
5	40.59%(+/-12.22)	55.18%(+/-19.67)	37.56%(+/-10.44)

MRF-2

NL	Freiburg	Saarbrücken	Stockholm
2	28.32%(+/-7.53)	39.85%(+/-19.42)	12.44%(+/-3.46)
5	24.23%(+/-11.40)	30.58%(+/-5.57)	10.04%(+/-2.59)

MRF-3

NL	Freiburg	Saarbrücken	Stockholm
2	28.71%(+/-5.43)	31.94%(+/-5.26)	10.11%(+/-0.51)
5	18.02%(+/-7.49)	28.86%(+/-6.16)	8.96%(+/-1.19)

#3 Novel structure detection

ROC curve