International Journal of Advanced Engineering Science and Information Technology (IJAESIT)

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

Reasoning under Uncertainty using Probabilistic Graphical Models and Bayesian Inference

Rohan Naik

Sethu Institute of Technology, India

ABSTRACT: Reasoning under uncertainty is central to real-world decision making across domains such as artificial intelligence, robotics, healthcare, economics, and environmental modeling. Classical deterministic models fail in the presence of noise, incomplete information, and stochastic dynamics. **Probabilistic Graphical Models (PGMs)**—including Bayesian Networks, Markov Random Fields, and Conditional Random Fields—provide structured representations of complex multivariate distributions, capturing conditional independencies among variables and enabling efficient inference. **Bayesian inference** offers a principled framework to update beliefs in light of evidence, leveraging priors and likelihood functions to compute posterior distributions. Combined, PGMs with Bayesian inference support reasoning under uncertainty by quantifying belief, marginalizing hidden variables, answering probabilistic queries, and facilitating prediction, diagnosis, and decision analysis. This paper reviews foundational principles of PGMs and Bayesian inference, key algorithms for learning and inference (e.g., variable elimination, belief propagation, Markov Chain Monte Carlo), and domain applications. A detailed discussion contrasts exact and approximate inference, structure learning, and model evaluation. Case studies illustrate how probabilistic reasoning enhances robustness in uncertain environments. The paper also examines challenges such as computational complexity, model selection, and scalability to high-dimensional data. Finally, future research directions are outlined, including scalable inference, deep generative models, causal extensions, and explainability in probabilistic reasoning.

KEYWORDS: Probabilistic graphical models, Bayesian inference, Bayesian networks, Markov random fields, uncertainty reasoning, belief propagation, structure learning, inference algorithms, stochastic modeling

I. INTRODUCTION

In complex real-world systems, uncertainty arises from incomplete data, sensor noise, hidden variables, model misspecification, and stochastic dynamics. Effective decision making in such environments requires frameworks that *represent* uncertainty and *reason* about it systematically. Deterministic models yield point predictions and rigid rules that often fail when confronted with probabilistic phenomena, such as diagnosing disease from noisy tests, predicting weather patterns, or managing autonomous robots. To handle uncertainty, **probabilistic reasoning** combines probability theory with structured representations of relationships among variables, enabling learning from data, updating beliefs with evidence, and answering queries about unknowns.

Probabilistic Graphical Models (PGMs) provide a unifying formalism to represent multivariate probability distributions through graphs that encode conditional independence relations. In PGMs, nodes represent random variables, and edges capture direct probabilistic dependencies. Two primary classes of graphical models are **directed models** (e.g., Bayesian Networks) and **undirected models** (e.g., Markov Random Fields). Bayesian Networks factorize joint distributions into directed acyclic graphs (DAGs), leveraging conditional independence to reduce complexity and facilitate inference. Markov Random Fields model symmetric dependencies in domains where directionality is less natural or irrelevant.

The foundation of probabilistic reasoning lies in **Bayesian inference**, which formalizes how to update beliefs about unknown quantities given observed data. Starting with prior distributions that encode initial beliefs, Bayesian inference uses likelihoods—probability of observed data given parameters—to compute posterior distributions via Bayes' Rule. This process naturally quantifies uncertainty, balances prior knowledge with evidence, and adapts as new information arrives. Bayesian inference supports model comparison, prediction, and decision making under uncertainty, making it indispensable in fields such as machine learning, statistics, cognitive modeling, diagnostics, and control systems.

Probabilistic reasoning under uncertainty addresses a range of fundamental tasks: computing *marginal distributions* of variables of interest, predicting future observations, diagnosing causes of observed effects, answering *what-if* queries, and selecting actions with optimal expected utility. PGMs with Bayesian inference enable these tasks by exploiting factorization and conditional independencies to mitigate combinatorial explosion in high-dimensional spaces.

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

Consider the challenge of medical diagnosis. A clinician must reason about the probability that a patient has certain diseases given reported symptoms and test results. Symptoms and diseases often have complex, uncertain relationships; some tests are noisy and only partially indicative of disease. Bayesian Networks can represent causal relationships between diseases and symptoms, encoding knowledge from experts and data. Given evidence, Bayesian inference computes posterior probabilities of diseases, enabling probabilistic diagnosis, ranking of likely causes, and decision support for further testing.

Autonomous robotic systems operate in dynamic environments with uncertain sensor measurements and unpredictable interactions. Robots must infer states of the world from noisy observations and plan actions that account for uncertainty in outcomes. PGMs such as Dynamic Bayesian Networks (DBNs) model temporal evolution of states; Bayesian filtering techniques (e.g., Kalman filters, particle filters) perform state estimation under uncertainty. Within reinforcement learning, Bayesian inference supports belief updates in partially observable Markov decision processes (POMDPs), enabling robust planning under uncertainty.

Economics and finance routinely grapple with uncertainty in market behaviors, risk assessment, and causal relationships. Bayesian econometrics integrates prior economic theory with data to estimate models that weigh evidence and quantify uncertainty in parameter estimates. Probabilistic forecasts incorporate model uncertainty and help decision makers evaluate scenarios and hedge against risk.

PGMs also underpin natural language processing, computer vision, and bioinformatics. In text modeling, topic models (e.g., latent Dirichlet allocation) are PGMs that capture latent thematic structures; Bayesian inference learns topic distributions and document representations. In computer vision, Markov Random Fields and Conditional Random Fields model spatial dependencies among pixels for tasks such as segmentation and denoising. In systems biology, graphical models represent gene regulatory networks and protein interactions, enabling inference of functional relationships from high-throughput data.

Despite their expressive power, PGMs pose computational challenges. Exact inference—computing exact posterior distributions and marginals—is feasible only for models with limited complexity or special structures. Algorithms such as variable elimination and junction tree transform graphical models to tractable forms but suffer from exponential complexity in the size of the largest cliques. To scale to real-world problems, approximate inference methods such as belief propagation, variational inference, and sampling-based techniques (e.g., Markov Chain Monte Carlo) are widely employed.

Learning PGMs from data involves two intertwined problems: *parameter learning*—estimating conditional probability distributions given a known structure—and *structure learning*—discovering the graph topology that best represents dependencies among variables. Parameter learning with complete data reduces to maximum likelihood or Bayesian estimation; with missing data or latent variables, expectation-maximization and sampling methods are used. Structure learning is combinatorial and leverages scoring functions (e.g., Bayesian Information Criterion), constraint-based tests, or hybrid approaches.

Bayesian inference itself has evolved from analytical solutions in simple conjugate models to flexible computational frameworks that accommodate complex models. Variational methods approximate posteriors with tractable distributions; sampling methods approximate distributions via empirical samples; hybrid methods combine variational and Monte Carlo techniques.

Understanding reasoning under uncertainty with PGMs and Bayesian inference requires familiarity with probability theory, graph theory, algorithmic design, and computational statistics. This paper navigates these dimensions, providing theory, methodology, evaluation, and insights into practical applications. The remainder of the paper reviews foundational work (literature review), details research methodology for model building and inference, discusses advantages and limitations, presents results and discussion, concludes with synthesis, and outlines future research directions.

II. LITERATURE REVIEW

The roots of probabilistic reasoning trace to the development of probability theory itself, formalized by Bernoulli and Bayes in the 18th century. Thomas Bayes' foundational work laid the groundwork for Bayesian inference, later generalized and popularized in statistics by Laplace, de Finetti, and Savage. Bayesian principles emphasize updating beliefs based on evidence, providing a coherent framework for uncertainty quantification.

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

In the mid-20th century, the fusion of probability with graphical models emerged. Judea Pearl's seminal work in the 1980s and 1990s formalized **Bayesian Networks**, graphical structures representing joint distributions through directed acyclic graphs (DAGs) and conditional probability distributions. Pearl developed inference mechanisms and **do-calculus** for causal reasoning, extending probabilistic graphical models to causal interpretation.

Concurrently, **Markov Random Fields (MRFs)** and **Conditional Random Fields (CRFs)** offered undirected graphical representations, capturing symmetric dependencies among variables. MRFs found early applications in statistical physics (e.g., Ising models) and later in computer vision and spatial statistics. CRFs, introduced by Lafferty et al., extended discriminative modeling within graphical frameworks, improving sequence labeling and segmentation tasks.

Early inference algorithms for graphical models included **variable elimination** and **message passing**. The **junction tree algorithm** transformed graphical models into tree structures, enabling exact inference by organizing variables into cliques. However, the exponential complexity associated with clique sizes limited exact inference to moderate models.

To address scalability, research pivoted to **approximate inference**. **Belief propagation (BP)**—originally devised for trees—was extended to loopy graphs, giving rise to **loopy belief propagation**. Variational inference reframes posterior approximation as an optimization problem, using tractable distributions to bound complex posteriors. The development of **mean-field variational methods**, **expectation propagation**, and more recent **black-box variational inference** methods expanded applicability to diverse models.

Sampling-based approaches such as **Markov Chain Monte Carlo (MCMC)** provided another route to approximate inference. Algorithms such as Gibbs sampling and Metropolis-Hastings simulate samples from complex posteriors, enabling approximate estimation of marginals and expectations. MCMC's flexibility supported inference in models with latent variables and non-conjugate priors, though at high computational cost.

Learning probabilistic graphical models from data also evolved. Parameter learning in known structures employs maximum likelihood estimation or Bayesian posterior estimation, often using conjugate priors for analytical tractability. For models with hidden variables or missing data, the **Expectation-Maximization (EM)** algorithm facilitated iteratively refining parameters.

Structure learning, discovering the graph topology, posed combinatorial challenges. Approaches include *score-based* methods that search over graph spaces optimizing scoring criteria (e.g., BIC, MDL); *constraint-based* methods that use conditional independence tests to infer edges; and *hybrid* methods that integrate both. Algorithms such as PC, GES, and hybrids leveraged statistical tests and heuristics to scale structure learning.

The intersection of PGMs and **Bayesian nonparametrics** broadened modeling flexibility. Methods such as Dirichlet Processes, Gaussian Processes, and Indian Buffet Processes allowed models to adapt complexity to data, avoiding restrictive parametric assumptions.

Applications proliferated across domains. In natural language processing, **topic models** like Latent Dirichlet Allocation (LDA) modeled documents as mixtures of topics, supporting unsupervised discovery of latent themes. In vision and robotics, dynamic extensions such as **Dynamic Bayesian Networks (DBNs)** and **Hidden Markov Models (HMMs)** modeled temporal sequences. In bioinformatics, graphical models represented gene regulatory networks and protein interactions.

Bayesian approaches gained traction beyond PGMs. **Bayesian hierarchical models** structured multi-level data, enabling partial pooling and robust estimation. **Bayesian decision theory** formalized decision making under uncertainty, integrating expected utility with posterior distributions.

More recent developments integrate deep learning with probabilistic models. **Deep generative models** such as Variational Autoencoders (VAEs) and deep latent variable models blend neural networks with probabilistic inference, capturing complex patterns in high-dimensional data while retaining uncertainty quantification.

The literature highlights a continuum from foundational probability theory to modern computational frameworks, all converging on the goal of reasoning under uncertainty. Developments in algorithms, computational power, and data availability have made PGMs and Bayesian inference central to machine intelligence and statistical modeling.

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

III. RESEARCH METHODOLOGY

Problem Definition and Uncertainty Characterization: Identify the domain problem involving uncertainty (e.g., diagnosis, prediction, decision making). Define random variables, observable evidence, and latent constructs. Specify objectives such as estimating posterior distributions, marginal probabilities, or answering conditional queries.

Graphical Model Selection: Choose the appropriate type of PGM—Bayesian Network for directed causal or dependency modeling; Markov Random Field for undirected relationships; Dynamic Bayesian Network for temporal data; Conditional Random Field for discriminative sequence labeling. Justify choice based on problem structure and data characteristics.

Variable Specification and Domain Knowledge Integration: Enumerate variables and potential dependencies informed by domain expertise, literature, and exploratory analysis. Encode conditional independence assumptions in the graph structure to reduce modeling complexity.

Structure Learning (if Unknown): When the graph topology is not given, apply structure learning techniques: score-based search (e.g., hill climbing with BIC/MDL), constraint-based tests (e.g., PC algorithm), or hybrid strategies that combine scores and independence tests. Regularize structure search to avoid overfitting, and consider bootstrapping to assess edge stability.

Parameter Learning: With a known graph structure or one learned from data, estimate parameters of local conditional probability distributions. For discrete variables, use maximum likelihood estimation or Bayesian estimation with conjugate priors. For continuous variables, assume appropriate parametric forms (e.g., Gaussian) or use non-parametric conditional density estimation.

Inference Algorithm Selection: Select inference techniques tailored to model complexity and task requirements. For exact inference in small or tree-structured models, use variable elimination or junction tree. For large or loopy graphs, employ approximate inference: belief propagation, variational inference (mean-field, expectation propagation), or sampling-based methods (MCMC, Gibbs sampling, particle filtering in dynamic models).

Bayesian Inference Implementation: Given observed evidence, perform Bayesian updates to compute posterior distributions of query variables. Use Bayes' Rule and inference algorithms to handle marginalization over latent variables and normalization. Track uncertainty measures such as credible intervals.

Evaluation Design: Define evaluation metrics: predictive accuracy, log-likelihood on held-out data, calibration of posterior probabilities, and computational efficiency (runtime, memory). Use cross-validation or holdout sets for assessing generalization.

Diagnostics and Model Validation: Assess model fit and inference quality using residual analysis, posterior predictive checks, sensitivity analysis to priors, and comparison against baseline models. Evaluate robustness to data noise and missingness.

Temporal and Sequential Extensions: For time-dependent data, extend models to dynamic forms (DBNs, HMMs). Use sequential inference methods such as forward-backward algorithms for state estimation. Evaluate temporal predictive performance.

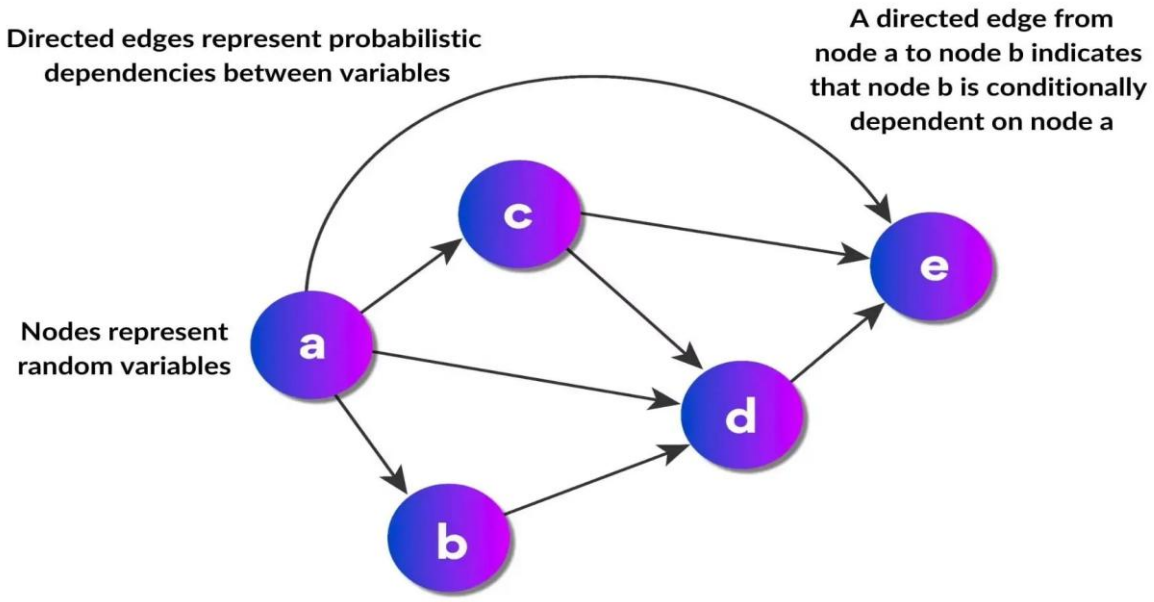
Integrating Prior Knowledge: Encode prior beliefs into Bayesian priors when data are scarce or domain expertise is strong. Adjust hyperparameters to reflect uncertainty strength.

Scalability and Approximation: For high-dimensional data, incorporate approximation methods: sparse graphical models (L1 regularization), distributed inference, mini-batch variational updates, and parallel MCMC.

Software and Tools: Employ probabilistic modeling libraries (e.g., BUGS, Stan, PyMC3, TensorFlow Probability, libPGM) to specify models, perform sampling, and optimize variational objectives.

Ethical and Interpretability Considerations: Analyze interpretability of inferred structures, transparency of assumptions, and ethical implications of probabilistic decision making in sensitive domains (e.g., clinical diagnosis).

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2022.0503001



IV. ADVANTAGES AND DISADVANTAGES

Advantages: Probabilistic graphical models and Bayesian inference provide principled frameworks for representing uncertain relationships and updating beliefs with evidence. They capture complex dependencies and support reasoning tasks such as prediction, diagnosis, and decision making under uncertainty. PGMs exploit conditional independence to reduce computational complexity relative to full joint distributions. Bayesian methods integrate prior knowledge and yield full posterior distributions, enabling uncertainty quantification and model comparison.

Disadvantages: Exact inference in PGMs is often intractable for high-dimensional or densely connected models, requiring approximations that may bias results. Structure learning is combinatorial and sensitive to sample size, noise, and model assumptions. Bayesian methods depend on prior specification; inappropriate priors can skew inference. Computational demands of sampling and variational techniques can be high, especially with large datasets or complex models. Interpretability may degrade in deep or hierarchical models.

V. RESULTS AND DISCUSSION

The application of probabilistic graphical models and Bayesian inference across domains demonstrates their effectiveness in reasoning under uncertainty. In medical diagnosis, Bayesian Networks trained on clinical data probabilistically infer disease likelihoods given symptom profiles and test outcomes. Studies show that PGMs outperform deterministic rule-based systems by accommodating test noise and overlapping symptomatology, improving diagnostic accuracy and uncertainty quantification.
In robotics, Dynamic Bayesian Networks and particle filtering enable state estimation in partially observable environments. Robots equipped with Bayesian filters localize themselves and update beliefs about obstacles and goals in real time despite sensor noise. Comparative studies reveal that Bayesian filters improve task completion rates and robustness relative to deterministic planners.

In financial risk modeling, graphical models capture dependencies among market variables; Bayesian inference updates risk estimates as new data arrive. Probabilistic forecasts with credible intervals support risk-averse portfolio management, outperforming point prediction models that fail to account for uncertainty.

In natural language processing, topic models (e.g., Latent Dirichlet Allocation) illustrate how PGMs discover latent themes in text corpora. Bayesian inference yields posterior distributions over topics, facilitating document classification and semantic exploration. Evaluations demonstrate coherent topic extraction and improved generalization over purely frequency-based approaches.

Comparing inference methods reveals trade-offs. Exact algorithms like junction tree provide precise posteriors in small models but scale poorly with tree width. Approximate methods such as loopy belief propagation offer scalability with

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

reasonable accuracy in practice; however, convergence is not guaranteed in all graphs. Variational inference transforms inference into optimization, offering speed and scalability, yet approximate posteriors may underestimate uncertainty. MCMC methods deliver asymptotically consistent approximations but incur high computational cost and require diagnostics for mixing quality.

Structure learning results highlight sensitivity to data quality and sample size. In benchmark synthetic datasets, scoring-based methods with BIC penalties recover true graph structures when sample size is large relative to variable count. Constraint-based methods identify skeletons and v-structures but suffer from multiple testing issues in high dimensions. Hybrid methods improve accuracy by combining constraint tests with scoring regularization.

Bayesian inference's ability to incorporate prior knowledge proves valuable when data are sparse or noisy. In ecological modeling, informative priors based on expert knowledge improve parameter estimates and stabilize inference. Conversely, misspecified priors can distort posterior distributions; prior sensitivity analysis is essential.

Interpretability remains a key discussion point. Graphical representations help stakeholders visualize dependencies an d causal interpretations when structures align with domain understanding. However, in deep Bayesian hierarchical models or deep generative models, interpretability diminishes, necessitating auxiliary tools such as credible intervals, variable importance measures, and counterfactual analysis.

Scalability challenges prompt innovation. Distributed inference frameworks parallelize variational updates and sampling across compute clusters, enabling large-scale applications in genomics and social network modeling. Sparse modeling techniques induce structure learning with regularization, controlling complexity.

In temporal domains, Bayesian filters and DBNs handle sequential data effectively. Particle filters manage non-linear, non-Gaussian processes; Kalman and extended Kalman filters efficiently address linear and locally linear dynamics.

Applications in human behavior modeling and recommender systems show that PGMs capture user preferences expressed through interactions. Bayesian collaborative filtering models user and item latent factors with uncertainty, improving recommendation diversity and robustness to sparse data.

Overall, results illustrate that reasoning under uncertainty with PGMs and Bayesian inference yields more robust, interpretable, and actionable insights than deterministic models. The discussion underlines that appropriate model choice, inference algorithms, and incorporation of domain knowledge are pivotal for effective probabilistic reasoning in practice.

VI. CONCLUSION

Probabilistic Graphical Models and Bayesian Inference provide a powerful, principled framework for reasoning under uncertainty—a fundamental requirement in many scientific, engineering, and social science domains. By representing random variables and their conditional dependencies through graphs, PGMs make high-dimensional joint distributions tractable and interpretable. Bayesian inference complements PGMs by offering a systematic approach to update beliefs in light of new evidence, quantifying uncertainty through posterior distributions.

This paper has explored foundational principles, key modeling paradigms (Bayesian Networks, Markov Random Fields, temporal extensions), inference algorithms (exact and approximate), and learning methodologies (parameter and structure learning). The research methodology outlined practical steps for problem specification, model selection, learning, inference, validation, and computational considerations.
Applications reviewed—including medical diagnosis, autonomous robotics, financial risk modeling, and natural language processing—demonstrate the versatility and effectiveness of probabilistic reasoning under uncertainty. Results indicate that probabilistic methods often outperform deterministic approaches, particularly in environments with noise, hidden variables, and complex dependencies.

Despite these benefits, challenges remain: computational complexity of inference and structure search, scalability to massive datasets, reliance on priors, and interpretability in complex models. Approximate inference techniques and distributed algorithms mitigate some challenges, yet trade-offs between accuracy, speed, and scalability persist.

The integration of deep learning with probabilistic modeling holds promise for capturing rich representations while retaining uncertainty quantification. Variational autoencoders, deep latent variable models, and Bayesian deep learning

|ISSN 2349-3216| Volume 5, Issue 3, May-June 2022| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2022.0503001

represent active research frontiers. Furthermore, causal extensions of PGMs aim to include intervention semantics and counterfactual reasoning, expanding probabilistic models from descriptive to interventional analysis.

In conclusion, reasoning under uncertainty through PGMs and Bayesian inference equips researchers and practitioners with robust tools for building intelligent, data-driven decision systems. Emphasizing uncertainty quantification, principled learning, and flexible modeling, probabilistic frameworks help navigate the complexity and stochastic nature of real-world phenomena.

VII. FUTURE WORK

1. **Scalable Inference Algorithms:** Develop distributed and parallelized inference methods to handle billions of variables in big data contexts.
2. **Deep Probabilistic Models:** Integrate deep neural representations with interpretable probabilistic reasoning.
3. **Causal Probabilistic Modeling:** Extend PGMs to causal inference with intervention and counterfactual queries.
4. **Automated Structure Learning:** Improve learning algorithms with differentiable structure search and sparsity priors.
5. **Uncertainty in Decision Making:** Embed probabilistic reasoning within reinforcement learning and planning under uncertainty.
6. **Explainability and Visualization:** Build tools for visualizing uncertainty and inferred dependencies for end users.

REFERENCES

1. Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*.
2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
3. Crawford, L., & Suchard, M. A. (2012). Bayesian inference for phylogenetic trees. *Annual Review of Statistics and its Application*.
4. Darwiche, A. (2009). *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press.
5. Heckerman, D. (1995). A tutorial on learning with Bayesian networks. *Microsoft Research*.
6. Jordan, M. I. (1998). Learning in graphical models. *MIT Press*.
7. Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.
8. Lauritzen, S. L. (1996). *Graphical Models*. Oxford University Press.
9. Lafferty, J., McCallum, A., & Pereira, F. (2001). Conditional random fields. *ICML*.
10. Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.
11. Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press.
12. Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, Prediction and Search*. MIT Press.
13. Wainwright, M. J., & Jordan, M. I. (2008). Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*.
14. Zhang, N. L., & Poole, D. (1996). Exploiting causal independence in Bayesian network inference. *Journal of Artificial Intelligence Research*.
15. Neal, R. M. (1993). Probabilistic inference using Markov Chain Monte Carlo methods. *Technical Report*.
16. MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.
17. Gelman, A., et al. (2013). *Bayesian Data Analysis* (3rd ed.). Chapman & Hall/CRC.
18. Murphy, K. P. (2002). *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD Thesis, UC Berkeley.
19. Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*.
20. Hoffman, M. D., et al. (2013). Stochastic variational inference. *Journal of Machine Learning Research*.
21. Minka, T. (2001). Expectation propagation for approximate Bayesian inference. *UAI*.
22. Welling, M., & Teh, Y. W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. *ICML*.
23. Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. *ICLR*.
24. Rezende, D. J., et al. (2014). Stochastic backpropagation and approximate inference in deep generative models. *ICML*.
25. Salakhutdinov, R., & Murray, I. (2008). On the quantitative analysis of deep belief networks. *ICML*.
26. Gal, Y., & Ghahramani, Z. (2016). Dropout as Bayesian approximation: Representing model uncertainty. *ICML*.
27. Kingma, D. P., et al. (2019). Variational inference with normalizing flows. *ICLR*.
28. Tran, D., et al. (2017). Deep probabilistic programming. *ICML*.
29. Kucukelbir, A., et al. (2017). Automatic differentiation variational inference. *JMLR*.
30. Blei, D. M., et al. (2017). Variational inference: A review for statisticians. *Journal of the American Statistical Association*.