Bayesian Network: Tabu search algorithm

STATS 9864 - Fall 2023

Kyuson Lim
Aryan Rezanezhad

December 4^{th}, 2023

Department of Mathematics and Statistics
University of Western Ontario

Outline

Motivation

Tabu Algorithm: Example in Binary Sequences (Stack)

- Analyzed problem of a binary sequence (stack):

  1100101...

  in which each bit corresponds to the order assigned to the **courier**.
  - The value of **"1"** means that the courier undertakes the contract, the value of **"0"** that he **rejects** it.
- For example, the implementation of the 5 orders starts from the obtaining of the set of orders, described as a specified bit sequence, eg: **1 0 1 0 1**.
  - Check the **neighboring** solution: the algorithm checks all neighboring solutions (differing in one bit):

    10100 10111 10111 10001 11101 01101

  - **Check** Tabu List: algorithm checks whether the transition between data solutions, already been made.
  - If so, such a solution is skipped and then the best solution, according to the specified criterion, is chosen:

    11101

- **Add** to Tabu List: And adds it to a list of prohibited movements (TL).

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Motivation

- **Tabu Algorithm:** is a search heuristic designed for optimization problems (prposed by Glover, 1986), aiming to find the best possible solution by iteratively moving through potential solutions.

- **Forbidden Paths Concept:** It incorporates a *"tabu list"* to keep track of previously visited solutions, by restricting the possible local choices to subset. This list prevents the algorithm from revisiting the same solutions, helping to avoid getting stuck in local optima.

- **Adaptive Decision-Making:** Tabu Algorithm utilizes a flexible approach, accepting sub-optimal solutions when necessary.

- **Avoidance of Stagnation:** It addresses challenges faced by traditional optimization methods in complex, non-convex spaces.

Tabu Algorithm: Flow Chart

- Initialize with candidate solutions (greedy approach).
- **Tabu list:** set of rules and banned solutions used to filter which solutions will be admitted to the neighborhood.
- **Neighborhood structure:** Specify how to generate neighboring solutions. This involves defining the set of allowable moves
- Selected attributes in solutions recently visited are labelled "**tabu-active**". Solutions that contain tabu-active elements are banned.
- **Tabu tenure**: the number of iterations that a move is declared being in tabu.
- **First in First out (FIFO):** first-in and first-out in queque stack.

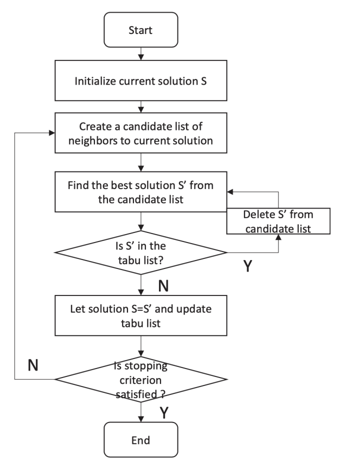


Figure 1. Flowchart of tabu search algorithm

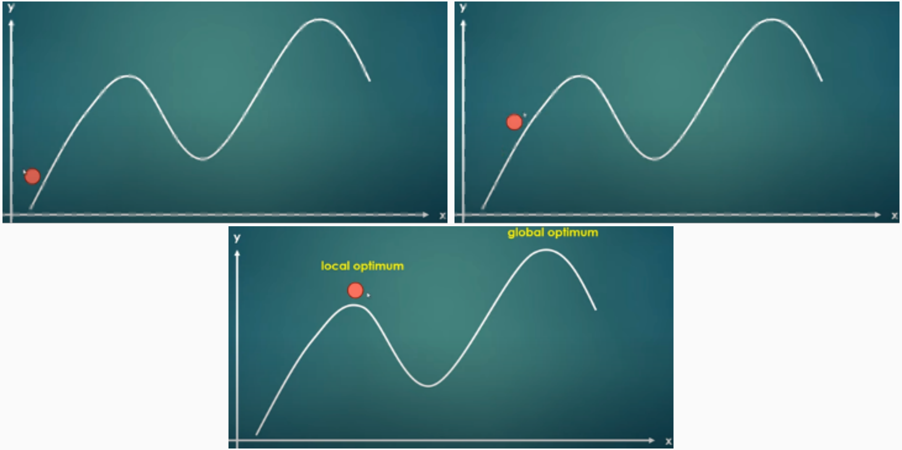
Introduction

Tabu Search

- Local search methods have a tendency to become stuck in suboptimal regions!!! Hill Climbing Algorithm
- Worsening moves can be accepted if no improving move is available (Stuck at a local minimum/maximum)
- Prohibitions ("Tabu") are introduced to discourage the search from coming back to previously visited solutions.
- Have to use a data structure to store prohibited points. We have to check according to this data structure whether a move is allowed or not

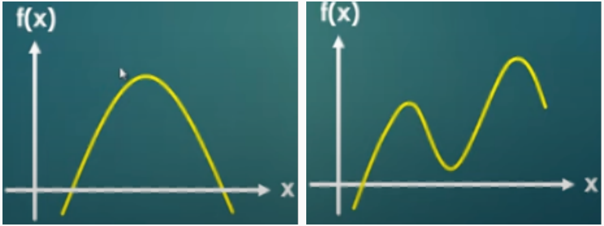
Limitations of Hill Climbing in Problem Solving

What's the problem with hill climbing?

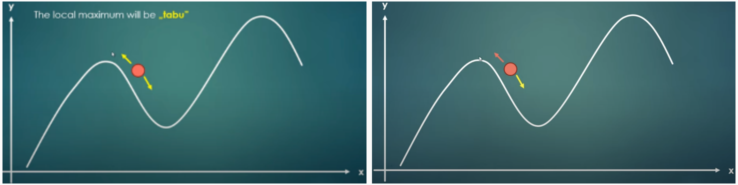


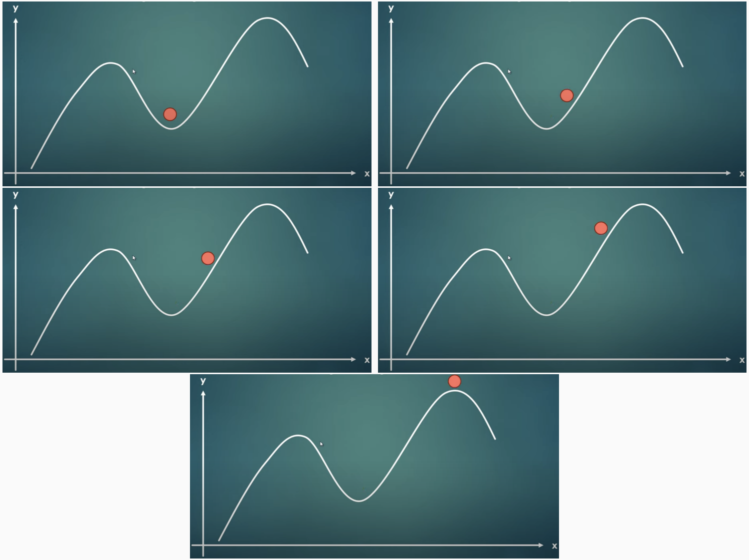
Convex and Non-Convex Search Spaces

- Effectiveness in Convex Search Spaces
  1. In convex search spaces, hill climbing excels by following the gradient direction.
  2. Successfully converges to the global maximum in scenarios like convex functions.
- Challenges in Non-Convex Search Spaces
  1. In non-convex search spaces, hill climbing faces limitations.
  2. Tends to converge to local maxima, necessitating heuristic or metaheuristic approaches for global optimization.



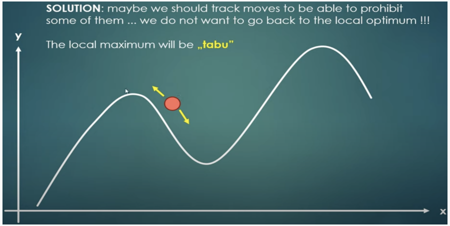
- **Objective: Finding the Global Optimum**
  1. Focus on seeking the global optimum rather than settling for a local optimum.
  2. Recognize the challenge when the gradient directs towards local maxima.
- **Challenge of Gradient-Driven Local Maximum**
  1. Illustrate the tendency of the gradient to lead back to local maxima.
  2. Acknowledge the issue of being stuck due to diminishing function values in each step.
- **Solution: Tabooing Local Maxima Moves**
  1. Propose a strategy of tracking moves to prevent revisiting local optima.
  2. Introduce the concept of tabooing, where certain moves are prohibited to explore alternative paths towards global maxima.

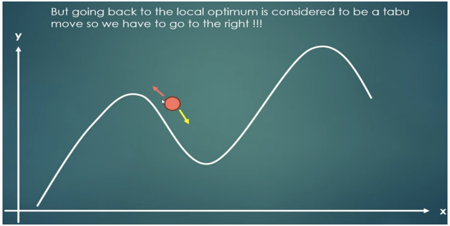



y
x

Tabu Tenure

- Tabu Tenure
  1. When a move is made tabu it is added to the so called tabu list with a certain value: this value is the "Tabu Tenure"
  2. With each iteration the tabu tenure is decremented by one.
  3. When the tabu tenure of a certain move is 0 $\rightarrow$ the certain move can be accepted.





Aspiration Criteria

- Exception: *Sometime we allow tabu moves*.
- Aspiration criteria allows a tabu move to be selected based on certain **constraints**.
  - For example, the given move allows a new global best solution AND all allowed moves are worse than actual one.
  - In these situations the move is accepted/its tabu tenure is renewed of course.
- Allowing a move even if it is tabu and if it results in a solution with an objective value better than that of the current best-known solution.

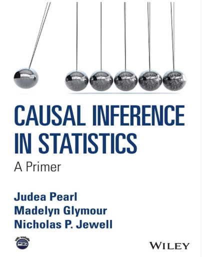
Objective and Goal: Tabu implementation in AI

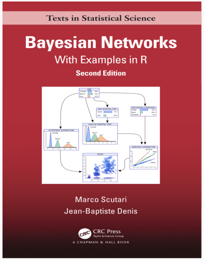
- Navigating Uncertainties with Bayesian Networks
  - **Motivation**: Explorer guided by Tabu Algorithm faces new challenges-imagine not just paths but a dynamic landscape of uncertainties.
  - **Goal**: Encountering unpredictable probabilities and causal inference, Bayesian Networks act as a compass, offering real-time predictions and adaptive strategies.
- Visual Representation
  - **Objective**: Explorer wields Tabu Algorithm for physical navigation and a Bayesian Network compass for uncertainties.
- Problem solution
  - In this optimization, the collaboration between Tabu Algorithm and Bayesian Networks transforms exploration-unlocking paths in categorical data analysis, adapting to uncertainties.

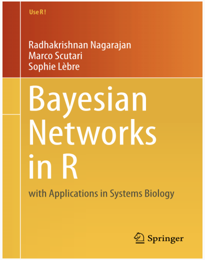
Implementation: Bayesian Network

Bayesian Network: probabilistic graphical model

- BNs, proposed by Pearl Judea, reflecting potential relationships among influencing factors, and conditional probability distribution tables (CPT), which demonstrate correlations between variables.
  - The **purpose** is to use Tabu algorithm in search for the best causal relationships.
  - The network structure of the graph reflects the relationships, and the relationship is built to represent how the occurrence of one variable influences the probability distribution of another.
  - Used for tasks in bio-statistical and epidemiological dataset for inference, reasoning, causal modeling, decision making under uncertainty and automated insight and prediction.



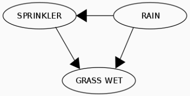




Bayesian Network- Example

- A simple example of chain-type Bayesian network is the **first order Markov chain**.
- Variables: R (Rain), G (Grass wet), S (Sprinkler) $\rightarrow$ T/F (XOR)
- Ex. $G \rightarrow T|R \rightarrow T$ or $G \rightarrow T$, the event of the grass being wet depends on both rain and artificial irrigation.

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

$$\text{Factorization: } P(R, S, G) = P(R)P(S|R)P(G|R, S)$$

- Causal Markov Condition: Ignoring effects, all relevant probabilistic information about a variable is contained in direct causes.

Tabu algorithm- Bayesian network

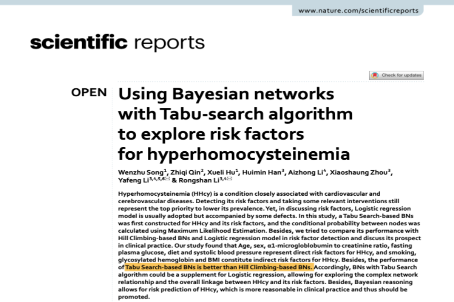
- **bnlearn**: is designed to provide a flexible simulation suite for methodological research and effective and scalable data analysis tools
- tabu in scoring method:

```
tabu(x, start = NULL, whitelist = NULL, blacklist = NULL,
    score = NULL, debug = FALSE, tabu = 10, max.tabu = tabu,
    max.iter = Inf, maxp = Inf, optimized = TRUE)
```

  - tabu a positive integer number, the length of the tabu list used in the tabu function (default:10).
  - max.tabu: a positive integer number, the iterations tabu search can perform without improving the best network score (default:10).
  - score: a character string, the label of the network score to be used in the algorithm. If none is specified, the default score is the Bayesian Information Criterion for both discrete and continuous data sets.

Tabu algorithm- Bayesian network Example

- By the paper below, the Tabu Search-based BNs is better than Hill Climbing-based BNs.
- The structure learning of BNs is carried out by using the "tabu" function in the package "bnlearn" in R.
    - The ROC (receiver-operating characteristic) curve with AUC (area under the curve) for accuracy values have demonstrated that the Tabu is much better performed.


www.nature.com/scientificreports

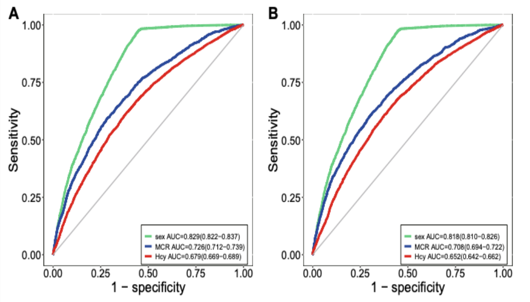
scientific reports

Check for updates

OPEN **Using Bayesian networks with Tabu-search algorithm to explore risk factors for hyperhomocysteinemia**

Wenzhu Song[1], Zhiqi Qin[2], Xueli Hu[1], Huimin Han[3], Aizhong Li[4], Xiaoshaung Zhou[3], Yafeng Li[3,4,5,6] & Rongshan Li[3,4]

Hyperhomocysteinemia (HHcy) is a condition closely associated with cardiovascular and cerebrovascular diseases. Detecting its risk factors and taking some relevant interventions still represent the top priority to lower its prevalence. Yet, in discussing risk factors, Logistic regression model is usually adopted but accompanied by some defects. In this study, a Tabu Search-based BNs was first constructed for HHcy and its risk factors, and the conditional probability between nodes was calculated using Maximum Likelihood Estimation. Besides, we tried to compare its performance with Hill Climbing-based BNs and Logistic regression model in risk factor detection and discuss its prospect in clinical practice. Our study found that Age, sex, α1-microgloblobumin to creatinine ratio, fasting plasma glucose, diet and systolic blood pressure represent direct risk factors for HHcy, and smoking, glycosylated hemoglobin and BMI constitute indirect risk factors for HHcy. Besides, the performance of Tabu Search-based BNs is better than Hill Climbing-based BNs. Accordingly, BNs with Tabu Search algorithm could be a supplement for Logistic regression, allowing for exploring the complex network relationship and the overall linkage between HHcy and its risk factors. Besides, Bayesian reasoning allows for risk prediction of HHcy, which is more reasonable in clinical practice and thus should be promoted.



Data analysis

Dataset Exploration

- Retrieved from `https://www.kaggle.com/code/marcotucci/asthma-proj-bayesian-network`, 9 categorical variables (sex, age, urbanization, education, geographic area, allergy, smoke, sedentary), and 2755 samples in total.
- The **aim** of this project is to study the impact of some **factors** on bronchial asthma, using Tabu search algorithm in Bayesian network.

```
> head(dat)
     sex   age urbanization education geographic_area allergy smoke sedentary asthma
1   male adult          low       low   south/islands     yes   yes       yes    yes
2 female   old          low       low   south/islands     yes    no       yes    yes
3 female adult         high      high          centre      no    no       yes    yes
4   male adult       medium       low   south/islands     yes    no        no     no
5 female adult          low      high           north      no    no        no     no
6 female adult       medium      high           north      no   yes        no    yes
```

Dataset Exploration: BN (1)

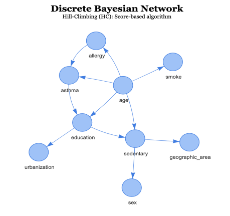
- A comparison of HC(Hill Climbing) and Tabu search algorithm:

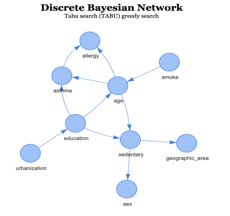
```r
## hill-climbing search
bn.learn.hc = hc(dat)

# plot
viewer(bn.learn.hc,
       bayesianNetwork.width = "100%",
       bayesianNetwork.height = "80vh",
       #bayesianNetwork.layout = "layout_with_sugiyama",
       bayesianNetwork.title="Discrete Bayesian Network",
       bayesianNetwork.subtitle = "Hill-Climbing (HC): Score-based algorithm"
       #bayesianNetwork.footer = "Fig. 1 - Layout with Sugiyama"
)

## tabu algorithm
bn.learn.hc = tabu(dat)

viewer(bn.learn.hc,
       bayesianNetwork.width = "100%",
       bayesianNetwork.height = "80vh",
       #bayesianNetwork.layout = "layout_with_sugiyama",
       bayesianNetwork.title="Discrete Bayesian Network",
       bayesianNetwork.subtitle = "Tabu search (TABU) greedy search"
       #bayesianNetwork.footer = "Fig. 1 - Layout with Sugiyama"
)

bn.fit(bn.learn.hc, data =dat)
```

```
> bn.learn.hc

  Bayesian network learned via Score-based methods

  model:
   [urbanization][smoke][education|urbanization]
   [age|education:smoke][sedentary|age:education]
   [asthma|age:education][sex|sedentary]
   [geographic_area|sedentary][allergy|age:asthma]
  nodes:                                 9
  arcs:                                  11
    undirected arcs:                     0
    directed arcs:                       11
  average markov blanket size:           2.67
  average neighbourhood size:            2.44
  average branching factor:              1.22

  learning algorithm:                    Tabu Search
  score:                                 BIC (disc.)
  penalization coefficient:              3.960586
  tests used in the learning procedure:  427
  optimized:                             TRUE
```

Dataset Exploration: BN (2)

- A comparison of HC(Hill Climbing) and Tabu search algorithm:





- **Causal discovery**: based on the network graph, the probabilities can be inferred ('cpquery').

```
Conditional probability table:

, , asthma = no

       age
allergy     adult       old     young
    no  0.8789290 0.8925144 0.8393782
    yes 0.1210710 0.1074856 0.1606218

, , asthma = yes

       age
allergy     adult       old     young
    no  0.4014733 0.6621622 0.4214876
    yes 0.5985267 0.3378378 0.5785124


  Parameters of node smoke (multinomial distribution)

Conditional probability table:
        no       yes
0.8424682 0.1575318
```

```
Conditional probability table:

, , asthma = no

       age
allergy     adult       old     young
    no  0.8789290 0.8925144 0.8393782
    yes 0.1210710 0.1074856 0.1606218

, , asthma = yes

       age
allergy     adult       old     young
    no  0.4014733 0.6621622 0.4214876
    yes 0.5985267 0.3378378 0.5785124


  Parameters of node smoke (multinomial distribution)

Conditional probability table:
        no       yes
0.8424682 0.1575318
```

```
Conditional probability table:

, , education = high

      age
asthma     adult       old     young
   no  0.6130884 0.6489796 0.6418919
   yes 0.3869116 0.3510204 0.3581081

, , education = low

      age
asthma     adult       old     young
   no  0.6120527 0.4559194 0.5903614
   yes 0.3879473 0.5440806 0.4096386
```

Conclusion: Causal Discovery (1)

- From **Tabu** algorithm potential causes of the form variable X causally influences variable Y:
- The causal relationship of asthma with respect to smoking and age:
  - Result of asthma, when age is young and smokers:
    - The probability of **having asthma** among individuals who are both young and smokers, is 0.36.
    - The probability of **not having asthma** among individuals who are both young and smokers, is 0.586.
  - Result of asthma, when age is old and smokers:
    - The probability of **having asthma** among individuals who are both old and smokers, is 0.47.
    - The probability of **not having asthma** among individuals who are both old and smokers, is 0.496.
  - Result of asthma, when age is adult and smokers:
    - The probability of **having asthma** among individuals who are both adult and smokers, is 0.385.
    - The probability of **not having asthma** among individuals who are both adult and smokers, is **0.645**.

Discussion

- **Advantage**:
  - **Efficient Heuristic Approach**: Tabu Search is often more computationally efficient than exact algorithms, making it suitable for large-scale problems such as the Bayesian Network.
  - **Ease of Implementation**: The basic principles of Tabu Search are relatively simple, making it accessible and easy to implement compared to some other optimization algorithms. This simplicity contributes to its popularity and applicability.
  - **Memory Mechanism**: The use of a Tabu List helps prevent revisiting the same solutions in the short term, promoting diversification to escape from local minima.
  - **Adaptability to Problem Dynamics**: Tabu Search can adapt to changes in the problem. By adjusting parameters or modifying the search strategy dynamically, the algorithm remains effective in dynamic environments.
- **Disadvantage**:
  - **Problem-Specific Tuning**: While Tabu Search is versatile, its effectiveness often relies on problem-specific tuning.
  - **Computational Intensity** for Large Problems: Computationally demanding with a large number of random variables as exponential growth of computational effort.
  - **No Guarantee** of Global Optimality: Like many heuristic optimization algorithms, Tabu Search does not guarantee finding the global optimum. It is possible for the algorithm to converge to a local minimum or plateau.

Biblibiography

References

- Song, W., Qin, Z., Hu, X., Han, H., Li, A., Zhou, X., ... & Li, R. (2023). Using Bayesian networks with Tabu-search algorithm to explore risk factors for hyperhomocysteinemia. Scientific Reports, 13(1), 1610.
- Scutari, M., & Denis, J. B. (2021). Bayesian networks: with examples in R. CRC press.
- Musial, K., Kotowska, J., Gornicka, D., & Burduk, A. (2017). Tabu search and greedy algorithm adaptation to logistic task. In Computer Information Systems and Industrial Management: 16th IFIP TC8 International Conference, CISIM 2017, Bialystok, Poland, June 16-18, 2017, Proceedings 16 (pp. 39-49). Springer International Publishing.

Thank you for the participation and understandings !