

# Group 5 - Module 1: AI Problem Solving

---

## Group Members:

Nils Dunlop, 20010127-2359, Applied Data Science, e-mail: gusdunlni@student.gu.se (9 hours)

Francisco Alejandro Erazo Piza, 19930613-9214, Applied Data Science, e-mail: guserafr@student.gu.se (9 hours)

We hereby declare that we have both actively participated in solving every exercise. All solutions are entirely our own work, without having taken part of other solutions." (This is independent and additional to any declaration that you may encounter in the electronic submission system.)

## Assignment 1

---

### Problem 1: Predict the Temperature

---

#### Characterization

- **Function:** This problem can be modeled as a regression problem where the goal is to predict a continuous value (temperature) based on other variables (like date, time, historical temperature data and geographical location).
- **Dynamic System:** Weather is inherently dynamic, with changes over time being significant. Therefore, this problem also fits into the dynamic system model where the current state (current weather conditions) influences future states (future temperatures).
- **Probable Data Characteristics:**
  - Temporal data is crucial as historical weather patterns greatly inform predictions.
  - Geographical variables are also key as weather patterns vary significantly by location.
  - External factors like climate change and seasonal patterns might influence long-term predictions.

#### Possible Solutions

##### Statistics and Machine Learning:

- **Time Series Analysis:** Utilize statistical methods like Autoregressive Integrated Moving Average (ARIMA) Prediction Model for short-term predictions (such as the next weekend). ARIMA is a statistical model that uses time series data to predict future trends and would be effective for short-term predictions. Another option is the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, which is an extension of ARIMA that includes seasonal components.
- **Machine Learning Models:** Employ models like Random Forests or Neural Network (especially LSTM - Long Short Term Memory networks for capturing temporal dependencies), which can handle multiple variables and are effective for both short-term and long-term predictions.

#### Suggested Solution

Given the complexity of weather patterns, a combination of time series analysis and machine learning might offer the most robust solution. For next weekend's prediction, focus on short-term patterns and recent data. For next year's prediction, incorporate long-term trends and potential climatic shifts.

## Problem 2: Bingo Lottery Problem

---

### Characterization

- **Constraint Satisfaction Problem:** The goal of selecting a set of numbers to ensure exactly 120 winning tickets with no ticket winning more than once can be modeled as a constraint satisfaction problem. The constraints include the number of winning tickets and the condition that each ticket can only win once.
- **Optimization Problem:** This problem can also be framed as an optimization problem where the objective is to minimize the difference from the desired 120 winning tickets while holding true to the constraints.
- **Discrete Algorithmic Problem:** Furthermore, the problem fits into a discrete algorithmic framework given the finite set of numbers (1 to 25) and a fixed number of tickets (1000).
- **Probable Data Characteristics::**
  - The randomness in the distribution of numbers across tickets.
  - The need for balance in the draw to avoid multiple wins on a single ticket.
  - A finite and discrete set of numbers and tickets.

### Possible Solutions

- **Constraint Satisfaction Problem:**
  - **Constraint programming:** can effectively address the bingo lottery problem by defining variables and constraints, such as ensuring a single win per ticket and aiming for 120 winners. It involves selecting a set of numbers that meet these conditions and using constraint solvers to efficiently find solutions, focusing on optimal distribution of numbers across the tickets to maintain balance and fairness.
- **Optimization Problem:**
  - **Linear Programming:** could be used to find an optimal set of numbers that maximizes the number of winners (up to 120) while ensuring that no ticket wins more than once.
- **Discrete Algorithmic Problem:**
  - **Backtracking:** Algorithms such as backtracking could be applied especially for searching through combinations of numbers to find a set that meets the criteria.

### Suggested Solution

Given the specificity of the problem, constraint programming or optimization techniques might offer the most optimal approach. These methods can explicitly incorporate the problem's constraints and seek solutions that satisfy them.

## Problem 3: Public Transport Departure Forecast

---

### Characterization

- **Regression:** This problem can be modeled as a regression task where the goal is to predict a continuous value: the departure time of the next vehicle.
- **Time Series Prediction:** Given that this involves predicting future events based on historical data the problem fits into the time series prediction category.
- **Dynamic System:** As a result of the public transport being influenced by external factors such as traffic, weather, and other events, the problem also fits into the dynamic system model.
- **Probable Data Characteristics:**
  - Variability in traffic and operational conditions.
  - Historical patterns of departures and delays.
  - Real-time data from vehicles and traffic systems.

### Possible Solutions

- **Statistics and Machine Learning:**
  - **Historical Data Analysis:** Use historical departure and arrival times to predict future schedules.
  - **Machine Learning Models:** Implement models like Random Forest or Long Short-Term Memory networks which can process sequential data effectively.
- **Real-Time Data Integration:**
  - Integrate real-time data from GPS trackers on vehicles, traffic reports, and other relevant sources to adjust predictions dynamically.
- **Simulation:**
  - Simulate traffic conditions and public transport operations to predict departure times under various scenarios.

### Suggested Solution

An effective approach for predicting public transport departure times involves a blend of analyzing historical data, integrating real-time information, and utilizing advanced predictive models. The selection of the most suitable method comes down to the data's availability and quality as well as the unique operational dynamics of the transport system.

# Problem 4: Film Festival Problem

---

## Characterization

- **Constraint Satisfaction Problem:** The primary challenge is to create a schedule that satisfies multiple constraints (film timings, individual preferences, overlapping shows).
- **Optimization Problem:** Maximizing the number of high-priority films in each individual's schedule.
- **Probable Data Characteristics:**
  - Constraints include show timings, individual availability, and venue capacity.
  - Different individuals will have different priorities for the films they wish to see.
  - The same film being shown multiple times adds complexity to the scheduling.

## Possible Solutions

- **Discrete Algorithm Design:**
  - **Graph Algorithms:** Representing the scheduling problem as a graph problem, where nodes are film screenings and edges represent the possibility of attending consecutive screenings.
  - **Backtracking Algorithms:** Useful in exploring possible schedules and backtracking when constraints are violated.
- **Optimization:**
  - **Integer Linear Programming:** Formulating the problem as an ILP where the objective is to maximize the number of high-priority films in the schedule while adhering to constraints.
  - **Heuristic Algorithms:** Such as genetic algorithms or simulated annealing, to find good (though not necessarily optimal) solutions in a reasonable amount of time.
- **Constraint Programming:**
  - The primary approach for this problem. A constraint solver can be used to ensure that all the constraints (show timings, individual preferences) are met.

## Suggested Solution

For the Film Festival Problem, the best method seems to be using both constraint programming and optimization. The main idea is to set up the problem so we can easily find solutions that fit the film schedules and what people like. We can use heuristic methods to find good enough solutions quickly, which is really helpful when there are a lot of different things to consider.

# Problem 5: Product Rating in Consumer Test

---

## Characterization

- **Multivariate Analysis:** Assessing dishwashers involves considering multiple variables such as cleaning efficiency, energy consumption, noise level, durability, and user-friendliness.
- **Qualitative and Quantitative Assessment:** The problem involves both objective measurements (like energy use) and subjective evaluations (like user satisfaction).
- **Probable Data Characteristics:**
  - **Standardization of Metrics:** Ensuring that all products are tested and evaluated under similar conditions and using the same criteria.
  - **Weight Assignment:** Different features might have different levels of importance to consumers.
  - **Consumer Preferences:** Understanding and reflecting the preferences and priorities of the target consumer group.

## Possible Solutions

- **Statistics and Machine Learning:**
  - **Descriptive Statistics:** Using mean, median, standard deviation, etc. for quantifiable aspects of the dishwashers.
  - **Factor Analysis:** To identify underlying variables (factors) that explain the pattern of correlations among the multiple observed variables.
- **Prescriptive Modelling:**
  - **Weighted Scoring System:** Assign weights to different features based on their importance. For example, energy efficiency might be more important than the variety of wash cycles.
  - **Score Normalization:** Normalize scores across different tests to make them comparable.
- **Simulation:**
  - **Monte Carlo Simulation:** Could be used for uncertainty analysis, especially if there's variability in the test results.
- **Consumer Feedback Analysis:**
  - **Sentiment Analysis:** If consumer reviews are part of the data, sentiment analysis can be employed to gauge user satisfaction.

## Suggested Solution

In a consumer test for dishwashers, using a mix of weighted scoring for various features, descriptive statistics for the numbers, and maybe sentiment analysis for customer feedback could work well. The main goal is to create a clear, fair, and detailed system that really shows how well the dishwashers perform and how much consumers like them.

# Problem 6: Constraint Satisfaction and Constraint Programming

---

**Explain the main concepts of constraint satisfaction problems and constraint programming.**

## Constraint Satisfaction Problems

Constraint Satisfaction Problems are mathematical problems defined as a set of objects whose state must satisfy a number of constraints or limitations. CSPs are widely used in computer science to model problems in fields such as artificial intelligence, operations research and computer programming.

### Key Concepts

- **Variables:** In Constraint Satisfaction Problems, variables can be anything from numbers, colors, objects, etc. with each variable having a range of possible values (domain).
- **Constraints:** These are rules that define which combinations of values are allowed. For example in a scheduling problem, a constraint might be that two events cannot occur at the same time.
- **Solutions:** A solution to a Constraint Satisfaction Problem is an assignment of values to all variables that satisfies all constraints.

## Solving Constraint Satisfaction Problems

- **Search Algorithms:** Often involve systematically exploring combinations of variable assignments.
- **Backtracking:** A common search technique in Constraint Satisfaction Problems, where the algorithm tries possible values for a variable and backtracks when a variable has no legal values left to assign.
- **Constraint Propagation:** Reducing the search space by deducing the possible values of a variable based on the constraints.

## Constraint Programming

Constraint Programming is a paradigm for solving Constraint Satisfaction Problems that involves programming languages, techniques and tools specifically designed for these problems.

### Key Concepts

- **Representation of Variables and Domains:** Constraint Programming involves representing variables and their possible values.
- **Encoding of Constraints:** Constraints are expressed in a way that the Constraint Programming system can understand and manipulate.
- **Search Mechanism:** Constraint Programming systems provide mechanisms to explore the solution space, including backtracking and heuristics.

## Constraint Programming Techniques

- **Propagation Algorithms:** Used to reduce the domains of variables by eliminating values that cannot be part of a solution.
- **Global Constraints:** Constraints that cover a number of variables and provide efficient ways to check a constraint and reduce domains.

## Generalize constraint programming and constraint propagation to numbers.

- **Numerical Constraints:** Instead of just true/false, constraints can involve numerical values, such as " $x + y < 20$ ". This is common in optimization problems.
- **Use of Numbers in Constraints:** In CP, you can replace true/false with numbers to represent different states or weights. For instance, instead of just saying whether a path is available or not you could assign a cost to each path.

In conclusion, CSPs and CP offer versatile frameworks for modeling and solving complex problems where multiple variables and constraints are involved. The extension of these concepts to numerical domains broadens their applicability making them suitable for a wide range of problems in various fields.

### References:

- Wikipedia Contributors (2024). Constraint programming. [online] Wikipedia. Available at: [https://en.wikipedia.org/wiki/Constraint\\_programming](https://en.wikipedia.org/wiki/Constraint_programming) [Accessed 23 Jan. 2024].
- Wikipedia Contributors (2024). Constraint satisfaction problem. [online] Wikipedia. Available at: [https://en.wikipedia.org/wiki/Constraint\\_satisfaction\\_problem](https://en.wikipedia.org/wiki/Constraint_satisfaction_problem) [Accessed 23 Jan. 2024].

# Self Check

---

- Have you answered all questions to the best of your ability?  
Yes, we have.
- Is all the required information on the front page, is the file name correct etc.?  
Indeed, all the required information on the front page has been included.
- Anything else you can easily check? (details, terminology, arguments, clearly stated answers etc.?)  
We have checked, and everything looks good.

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js