**Universitat Rovira i Virgili**

# Task 1 - Criteria Modeling

## Logic Scoring of Preferences

**Author**:
Kacper Poniatowski

February 2025

# Contents

# 1 Introduction

## 1.1 Purpose of Report

This task introduces the LSP method to understand how to define utility functions to measure the suitability of a set of criteria. For this, we utilised a tool developed at the ITAKA research group of University Rovira i Virgili.

## 1.2 Overview of Task

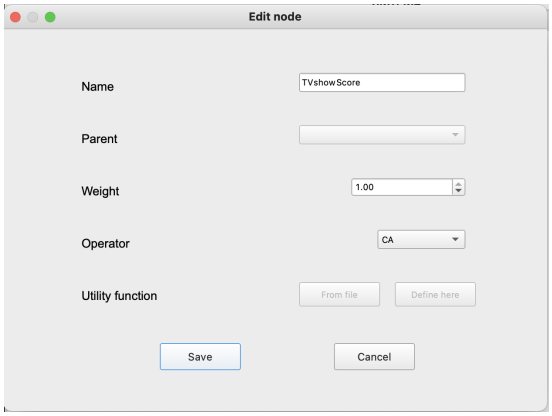
The goal of this task is to design a simplified version of a decision support system to evaluate and select television shows. A criteria tree is created within the LSP software to model the criteria hierarchy. Utility functions are defined to map the feature variables to performance scores. Alternatives are entered to evaluate their suitability against the decision makers' preferences. Lastly, the results are analysed in this report.

# 2 Criteria Hierarchy Definition

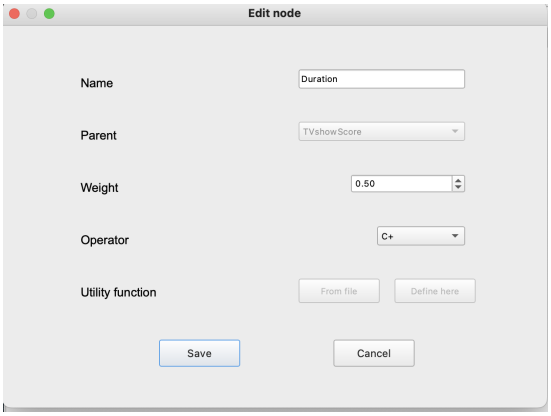

Figure 1: Criteria hierarchy tree used in the decision support system.

As shown in Figure 1, the hierarchical structure defines the overall evaluation criteria.
The nodes are as follows:

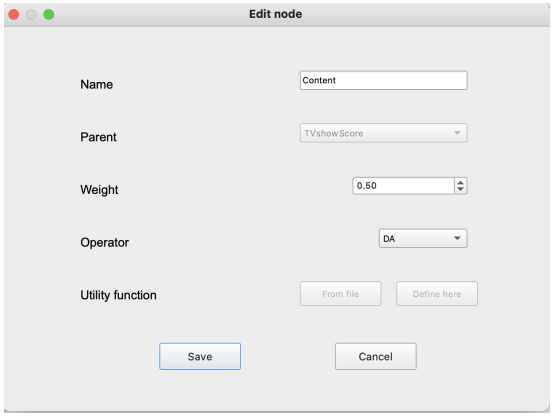
- **TVShowScore:** Represents the overall utility score indicating the suitability of the TV show.
- **Duration:** Evaluates the time-related aspects of the show, combining factors such as episode length and number of episodes.
- **EpisodeLength:** Measures the average duration of each episode, reflecting the show's pacing and structure.
- **NumberEpisodes:** Assesses the total number of episodes, capturing the show's completeness and content depth.
- **Content:** Represents the qualitative aspects of the show, focusing on genre-specific features.
- **Action:** Evaluates the level of action or dynamism.
- **SciFi:** Measures the extent of science-fiction elements present.
- **Comedy:** Assesses the proportion of comedic content.




(a) TVshowScore criterion



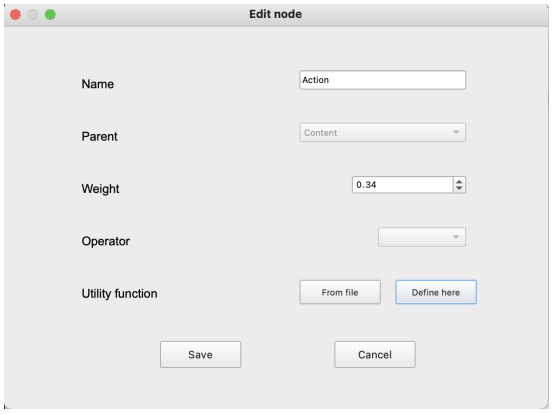

(b) Duration criterion



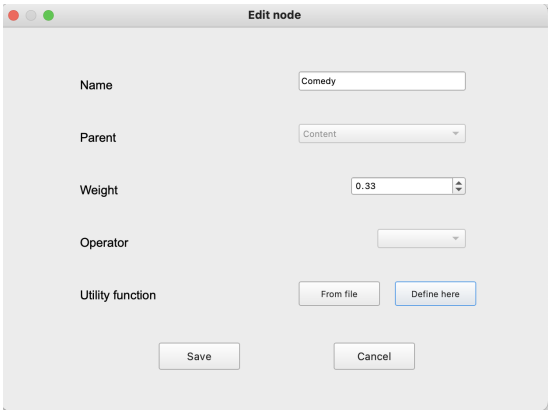

(c) Content criterion

Figure 2: Details of each of the non-leaf criteria.

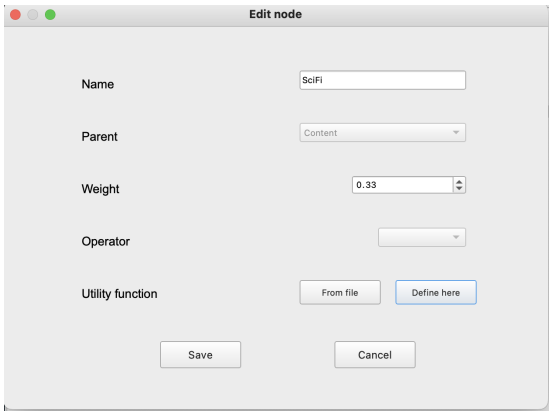
Figure 2 displays the details of each of the non-leaf criteria.



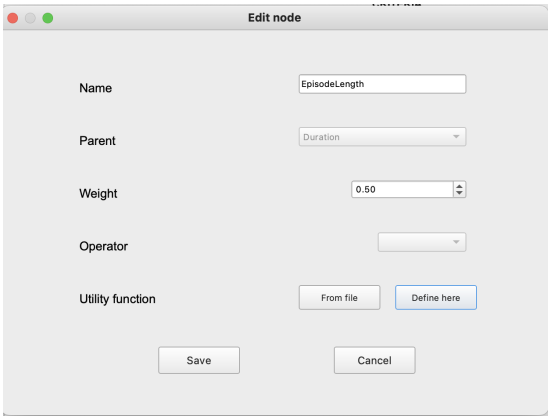

(a) Action criterion



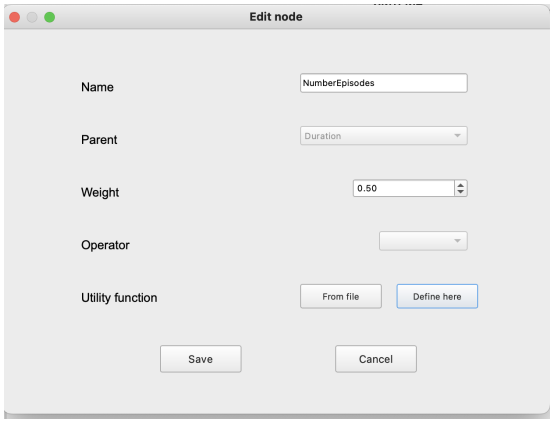

(b) Comedy criterion




(c) SciFi criterion




(d) EpisodeLength criterion




(e) NumberEpisodes criterion

Figure 3: Details of each of the leaf criteria.

Figure 3 displays the details of each of the leaf criteria.
The weights set for the criterion are discussed in section 4.

# 3 Utility Functions for Leaf Criteria

## 3.1 EpisodeLength

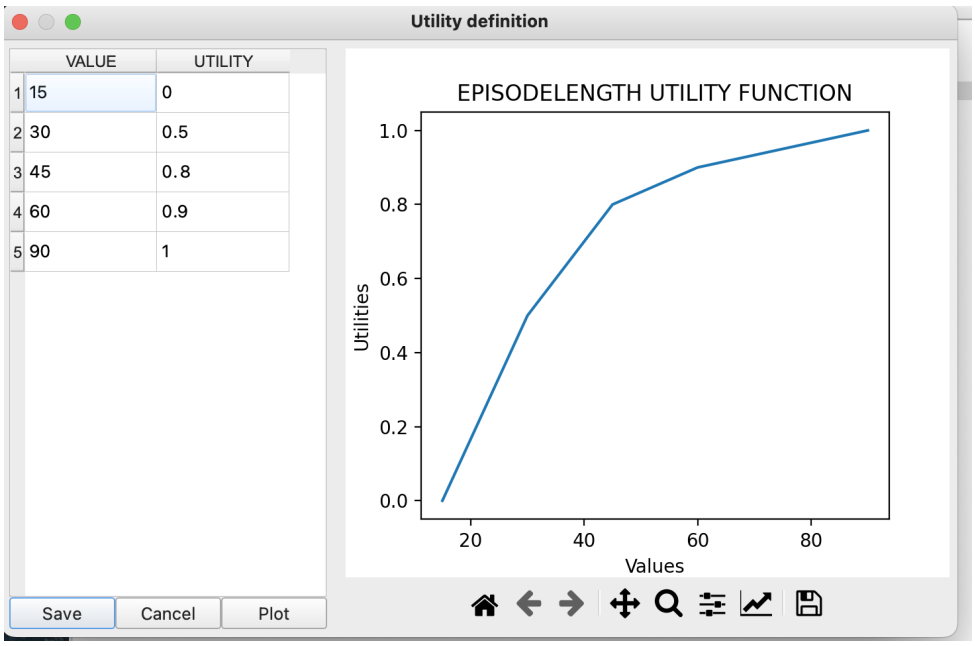



Figure 4: EpisodeLength utility function.

Figure Figure 4 displays the utility function for the EpisodeLength criteria.

- **Range:** 15-90 minutes.
- **Rationale:** Longer episode lengths are more desirable than short length episodes. When episodes are less than 30 or 40 minutes long, they struggle to captivate me and fail to develop any plots meaningfully due to the time constraints. Around the 40 minute mark these problems start to be alleviated, which is reflected in the increasing utility. As the episode length approaches 90 minutes, the function tapers off near its maximum utility.

## 3.2 NumberEpisodes

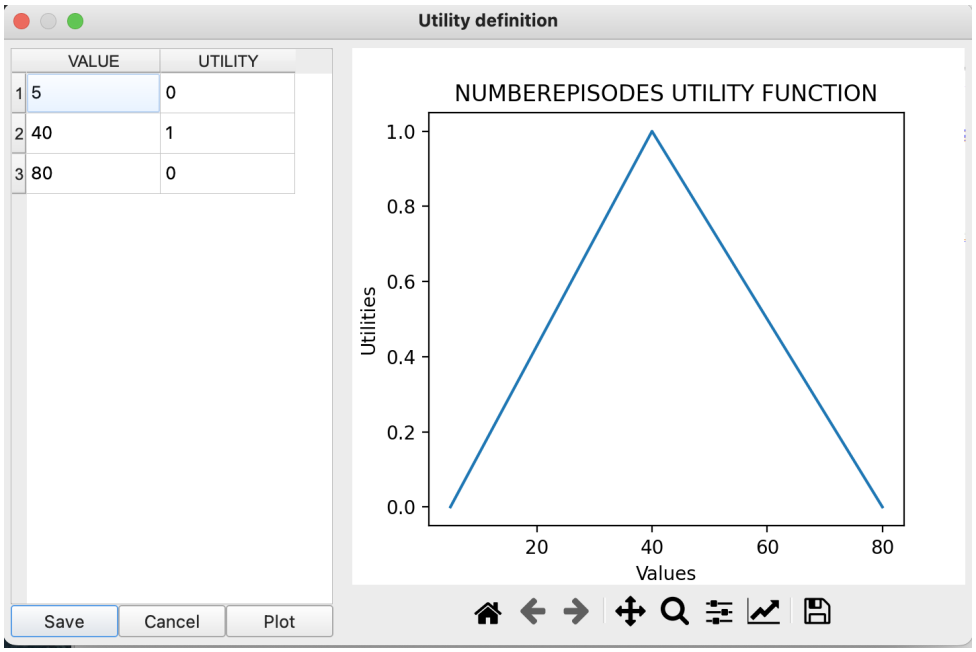



Figure 5: NumberEpisodes utility function.

Figure Figure 5 displays the utility function for the NumberEpisodes criteria.

- **Range:** 5-80 episodes.
- **Rationale:** Minimum utility begins at 5 episodes, increases linearly until maximum utility is reached at 40 episodes and proceeds to decrease linearly until minimum utility is reached once again at 80 episodes. In my opinion, 40 episodes represent the sweet spot for a TV show—approximately equivalent to 5 seasons—which strikes a great balance between quantity and quality.

## 3.3 Action

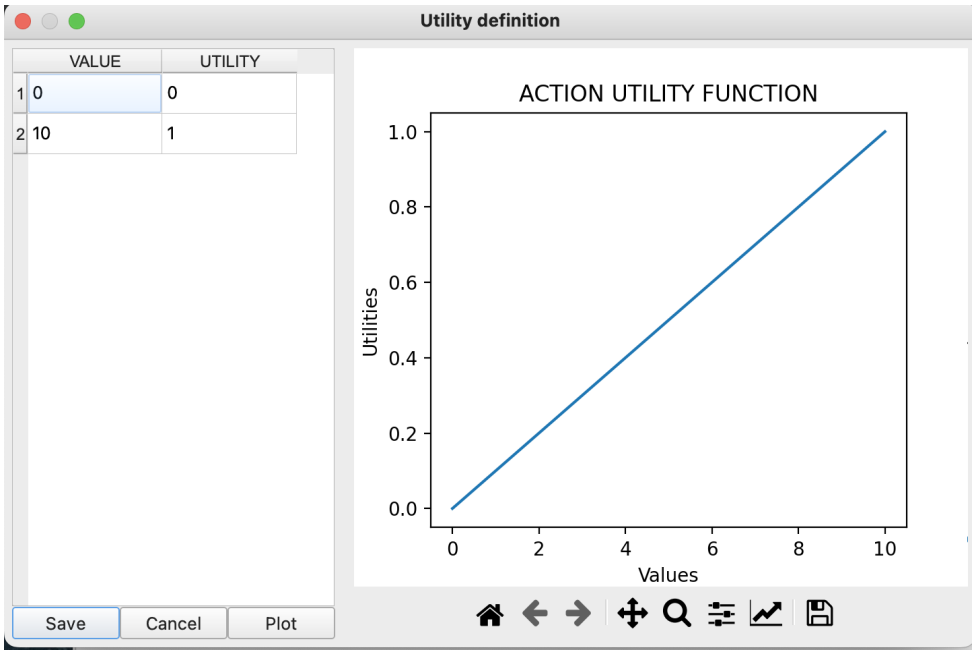



Figure 6: Action utility function.

Figure Figure 6 displays the utility function for the Action criteria.

- **Range:** 0-10 (0: no action, 10: full of action).
- **Rationale:** Since the action genre is one of my personal favourites, a TV show with no action achieves a minimum utility score (0), while an action-packed TV show scores the maximum utility score (10). The score increases linearly with the amount of action in the show.

### 3.4 SciFi

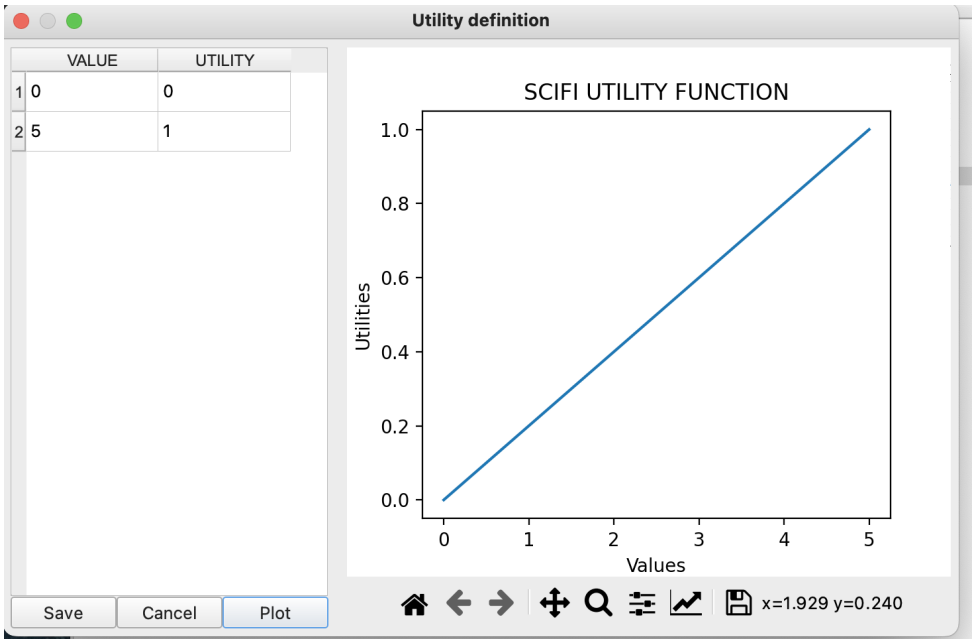



Figure 7: SciFi utility function.

Figure Figure 7 displays the utility function for the SciFi criteria.

- **Range:** 0-5 (0: complete reality-based, 5: complete science-fiction).
- **Rationale:** Similarly to the action genre, science fiction is one of my favourites, therefore the same rationale applies for the utility function in this criteria.

### 3.5 Comedy

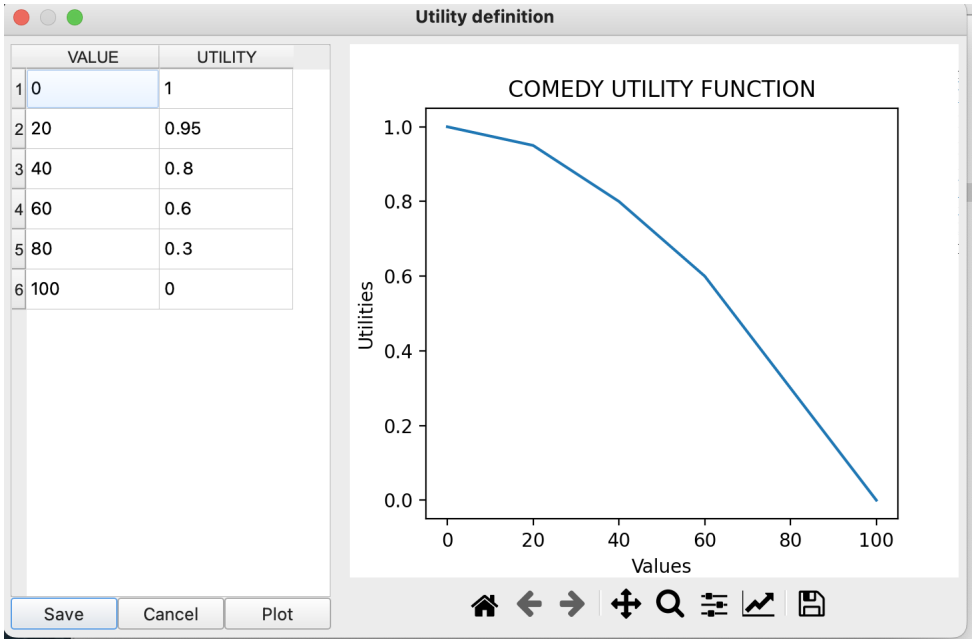



Figure 8: Comedy utility function.

Figure Figure 8 displays the utility function for the Comedy criteria.

- **Range:** 0-100 (0: completely drama-based, 100: completely comedy-based).
- **Rationale:** The comedy genre does not score favorably in my personal preferences, so a TV show with 0% comedy receives the maximum utility score, while a show that is entirely comedy scores the minimum. I chose a curved (non-linear) relationship because while a few comedic elements can be refreshing, an excessive amount of comedy quickly becomes a turn-off for me.

## 4 Weights and Aggregation Operators

### 4.1 Weight Distributions

| Node | Weight |
|---|---|
| TVshowScore (Root) | 1.0 (implicit) |
| Duration | 0.5 |
| EpisodeLength | 0.5 |
| NumberEpisodes | 0.5 |
| Content | 0.5 |
| Action | 0.34 |
| SciFi | 0.33 |
| Comedy | 0.33 |

Table 1: Equal weights assigned at each level of the criteria hierarchy.

Table 1 displays the weights assigned for each of the criteria (leaf and non-leaf). As stated in the task brief, equal weights are considered at each level in the hierarchy, thus implicitly combining to a total score of 1.0 at the root. Action has been assigned a weight of 0.34 as opposed to 0.33 to ensure the weights total to 1.0 for that level. The LSP software allows inputs only up to 2 decimal places.

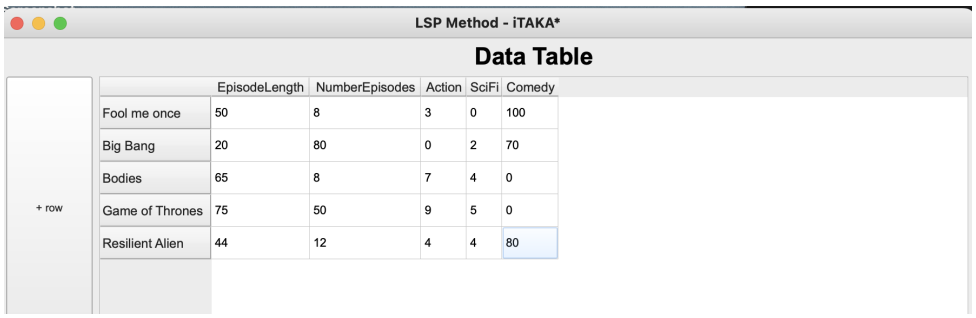
### 4.2 Aggregation Operators

| Node | Aggregation Operator | Justification |
|---|---|---|
| TVShowScore | CA | The CA (Compensatory Average) operator is used at the top level to balance the contributions of both Duration and Content, allowing a lower score in one area to be partially offset by a higher score in the other. [1] |
| Duration | C+ | The C+ operator (partially conjunctive) is chosen for Duration because both EpisodeLength and NumberEpisodes must meet a certain threshold to yield a high overall score, emphasizing that both factors are important. [1] |
| Content | DA | The DA (Disjunctive Aggregation) operator is applied to Content, as it allows for compensation among the content subcriteria (Action, SciFi, Comedy); a strong performance in one can help mitigate a lower performance in another. [1] |

Table 2: Aggregation operators and their justification for each intermediate node in the LSP hierarchy.

## 5 Evaluating Alternatives

### 5.1 Data Entry




| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy |
|---|---|---|---|---|---|
| Fool me once | 50 | 8 | 3 | 0 | 100 |
| Big Bang | 20 | 80 | 0 | 2 | 70 |
| Bodies | 65 | 8 | 7 | 4 | 0 |
| Game of Thrones | 75 | 50 | 9 | 5 | 0 |
| Resilient Alien | 44 | 12 | 4 | 4 | 80 |

Figure 9: Alternatives in tabular form.

Figure 9 displays the alternative set to be evaluated against decision-makers' preferences (that is, my preferences). These alternatives are provided according to the task brief.

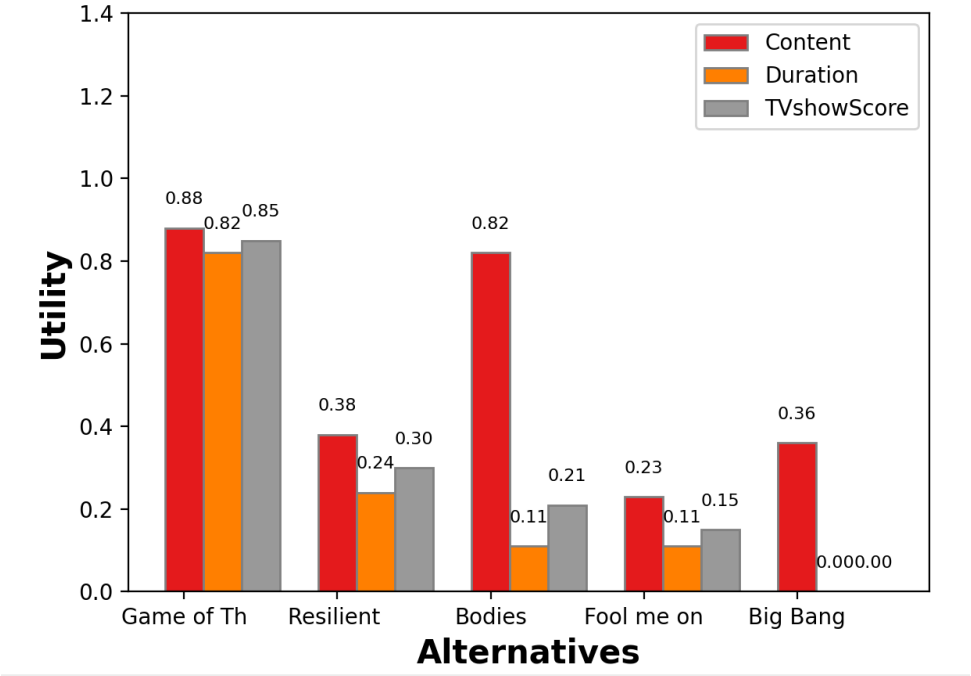
### 5.2 Performance Evaluation

LSP Method - iTAKA*

**Performance Table**

| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy | Content | Duration | TVshowScore |
|---|---|---|---|---|---|---|---|---|
| Fool me once | 0.83 | 0.09 | 0.3 | 0 | 0 | 0.23 | 0.11 | 0.15 |
| Big Bang | 0.17 | 0 | 0 | 0.2 | 0.45 | 0.36 | 0 | 0 |
| Bodies | 0.92 | 0.09 | 0.7 | 0.4 | 1 | 0.82 | 0.11 | 0.21 |
| Game of Thrones | 0.95 | 0.75 | 0.9 | 0.5 | 1 | 0.88 | 0.82 | 0.85 |
| Resilient Alien | 0.78 | 0.2 | 0.4 | 0.4 | 0.3 | 0.38 | 0.24 | 0.3 |

(a) Performance scores for each of the criterion.




(b) Plot of performance scores for each alternative.

Figure 10: Performance metrics for each alternative.

Figure 10 displays the utility scores for each alternative in tabular form and also as a graph. The best performing alternative is Game of Thrones with an overall score of 0.85. The worst performing alternative is Big Bang, with an overall score of 0. The remaining three alternatives all scored relatively poorly; Fool me once scored 0.15, Bodies scored 0.21, and Resilient Alien scored 0.3.

These overall scores are in accordance with my preferences as outlined in the previous steps. To further back up the solidity of these scores, I can confirm that Game of Thrones is one of my favorite TV shows and that I did not enjoy The Big Bang Theory (assuming this is the referenced show).

## 6 Analysis of Aggregation and Polarity

Polarity refers to the balance between the conjunctivity and disjunctivity of the aggregation. When an aggregation is conjunctive, all criteria must be satisfied to achieve a high overall score. Conversely, when an aggregation is disjunctive, a high score in one criteria can offset poor scores in other criteria.

Table 3 describes each of the aggregation operators utilised. These figures and descriptions have been sourced from the lecture slides provided to us. [1]

| Node | Operator | Polarity Description | And-ness | Or-ness |
|---|---|---|---|---|
| TVShowScore | CA | Medium Conjunctive | 75% | 25% |
| Duration | C+ | Strong Conjunctive | 87% | 13% |
| Content | DA | Medium Disjunctive | 25% | 75% |

Table 3: Aggregation operators with corresponding polarity and and-ness/or-ness split.

Choosing CA for TVshowScore provides a minor balance between the contributions of Duration and Content; a low score in one can be partially compensated by a high score in the other but ideally to achieve a high overall score, both Duration and Content should achieve a high score. Choosing C+ for Duration demands high performance in both EpisodeLength and NumberEpisodes as the balance is heavily favouring conjunctivity. A low score in either significantly reduces the overall Duration score. Choosing DA for Content allows a strong performance in one aspect to offset a lower score in another, resulting in a more forgiving aggregation.

# 7 Additional Scenarios

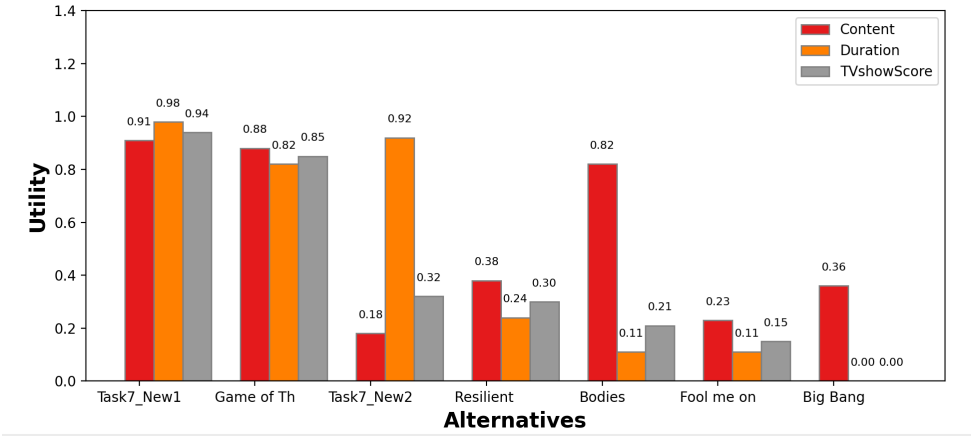
| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy |
|---|---|---|---|---|---|
| Fool me once | 50 | 8 | 3 | 0 | 100 |
| Big Bang | 20 | 80 | 0 | 2 | 70 |
| Bodies | 65 | 8 | 7 | 4 | 0 |
| Game of Thrones | 75 | 50 | 9 | 5 | 0 |
| Resilient Alien | 44 | 12 | 4 | 4 | 80 |
| Task7_New1 | 80 | 40 | 10 | 10 | 100 |
| Task7_New2 | 90 | 35 | 2 | 2 | 95 |

Figure 11: Alternatives including 2 new additional scenarios.

Figure 11 displays the set of alternatives including the two new additional scenarios, as outlined in the task brief.

| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy | Content | Duration | TVshowScore |
|---|---|---|---|---|---|---|---|---|
| Fool me once | 0.83 | 0.09 | 0.3 | 0 | 0 | 0.23 | 0.11 | 0.15 |
| Big Bang | 0.17 | 0 | 0 | 0.2 | 0.45 | 0.36 | 0 | 0 |
| Bodies | 0.92 | 0.09 | 0.7 | 0.4 | 1 | 0.82 | 0.11 | 0.21 |
| Game of Thrones | 0.95 | 0.75 | 0.9 | 0.5 | 1 | 0.88 | 0.82 | 0.85 |
| Resilient Alien | 0.78 | 0.2 | 0.4 | 0.4 | 0.3 | 0.38 | 0.24 | 0.3 |
| Task7_New1 | 0.97 | 1 | 1 | 1 | 0 | 0.91 | 0.98 | 0.94 |
| Task7_New2 | 1 | 0.86 | 0.2 | 0.2 | 0.07 | 0.18 | 0.92 | 0.32 |

(a) Performance scores for each of the alternatives.




(b) Plot of performance scores for each alternative.

Figure 12: Performance metrics for each alternative including the new scenarios.

Figure 12 displays the performance scores of each of the alternatives, including the additional two scenarios.

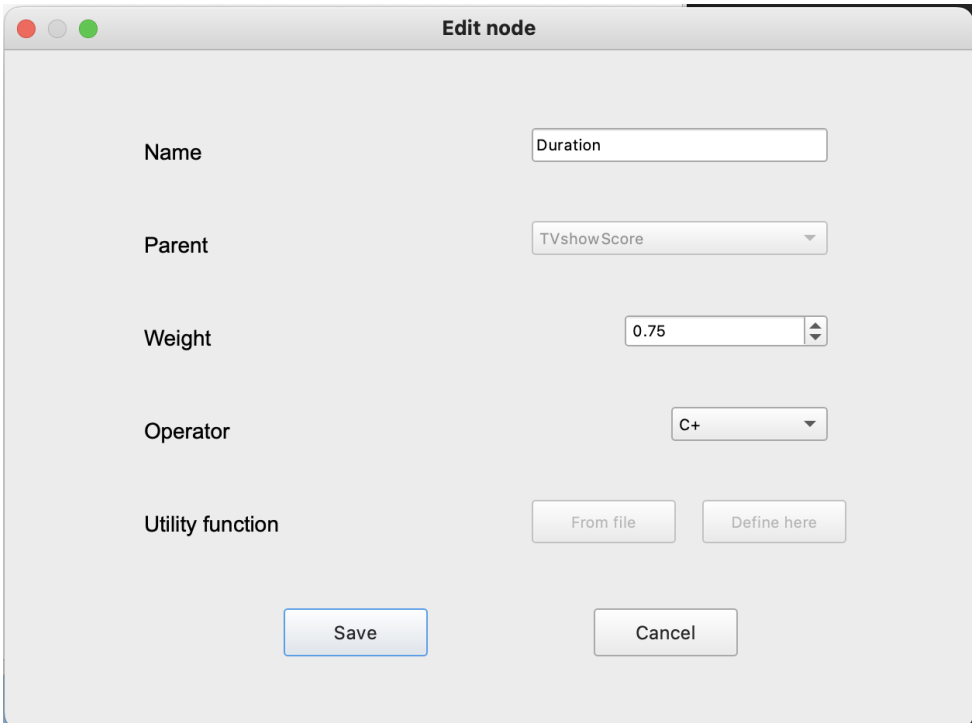
- `Task7_New1`: Satisfies preferences in both Duration and Content, but scores poorly in Comedy - thus satisfying the constraint regarding one poor performing elementary criterion.
- `Task7_New2`: Strong performance in Duration but poor performance in Content, as specified in the task brief. To achieve these scores, both EpisodeLength and NumberEpisodes are near the maximum utility score. SciFi is scored at 2 out of 5, Action is scored at 2 out of 10, and Comedy scored at 100 out of 100. These figures result in very poor performance in Content.

After the addition of these new scenarios, `Task7_New1` emerges as the best-performing alternative. This is unsurprising, as both Duration and Content are well satisfied, leading to strong conjunctive aggregation performance at TVShowScore. Despite performing strongly in Duration, `Task7_New2` is a poor choice due to its weak performance in Content, which results in poor conjunctive aggregation performance at TVShowScore.

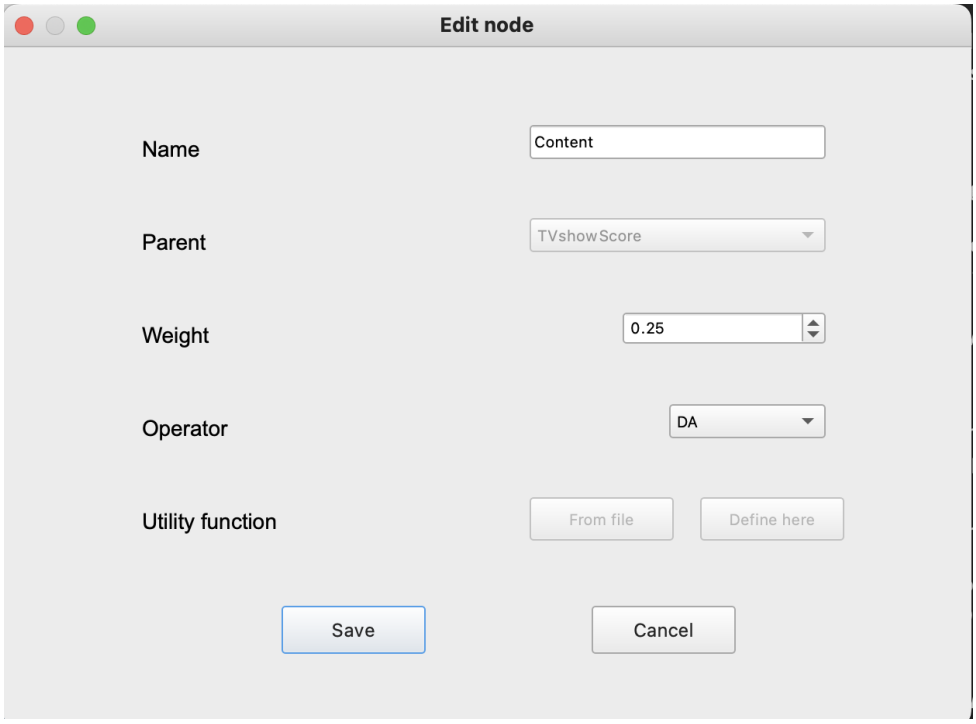
# 8 Sensitivity: Changing Weights or Operators

## 8.1 Adjusting Weights

To demonstrate the effect of modifying importance of criterion, the weight of Duration was increased to 0.75, thus reducing the weight of Content to 0.25 to satisfy the requirement of weights totaling to 1. Figure 13 displays these adjusted weights.




(a) Updated duration criterion.

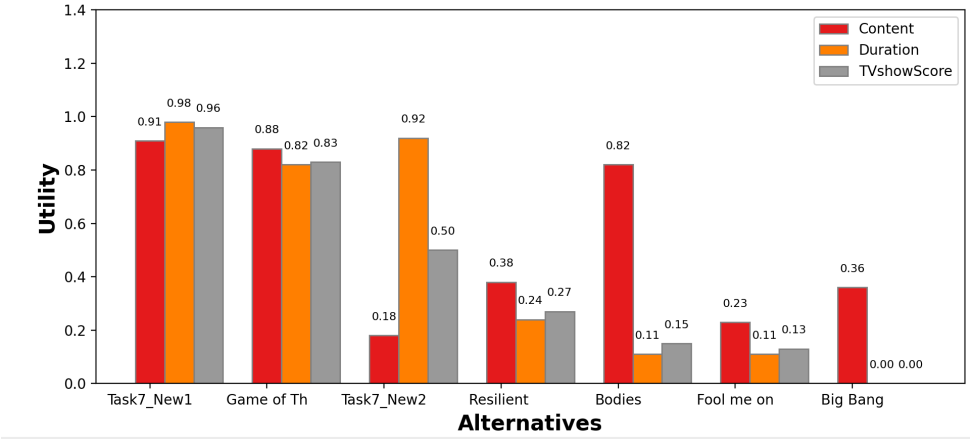



(b) Updated content criterion.

Figure 13: Updated weightings of both intermediary nodes, Content and Duration.

| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy | Content | Duration | TVshowScore |
|---|---|---|---|---|---|---|---|---|
| Fool me once | 0.83 | 0.09 | 0.3 | 0 | 0 | 0.23 | 0.11 | 0.13 |
| Big Bang | 0.17 | 0 | 0 | 0.2 | 0.45 | 0.36 | 0 | 0 |
| Bodies | 0.92 | 0.09 | 0.7 | 0.4 | 1 | 0.82 | 0.11 | 0.15 |
| Game of Thrones | 0.95 | 0.75 | 0.9 | 0.5 | 1 | 0.88 | 0.82 | 0.83 |
| Resilient Alien | 0.78 | 0.2 | 0.4 | 0.4 | 0.3 | 0.38 | 0.24 | 0.27 |
| Task7_New1 | 0.97 | 1 | 1 | 1 | 0 | 0.91 | 0.98 | 0.96 |
| Task7_New2 | 1 | 0.86 | 0.2 | 0.2 | 0.07 | 0.18 | 0.92 | 0.5 |

(a) Updated performance scores.




(b) Updated performance score graph.

Figure 14: Updated performance scores and corresponding graph.

Figure 14 displays the updated scores after increasing the weight of Duration (and consequently reducing the weight of Content). In some cases, the scores changed only marginally due to the balance between Content and Duration. The alternative with the most significant change is `Task7_New2`, which saw a significant increase in score due to the re-weighting in favor of Duration.

## 8.2 Adjusting Aggregation Operators

Using different polarities for aggregation operators can have profound differences in performance results. This is demonstrated by adjusting the polarity of the Duration aggregator from C+ to D+. This change essentially switches the conjunctivity and disjunctivity of the aggregator, turning Duration disjunctive.

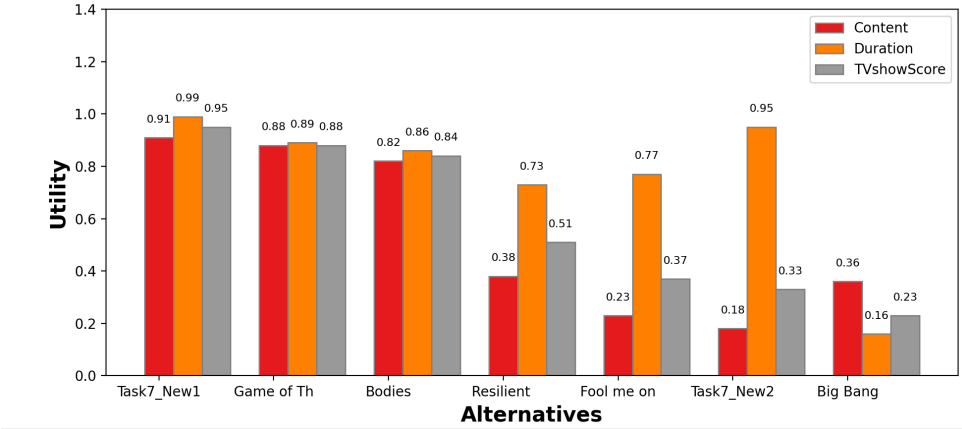
| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy | Content | Duration | TVshowScore |
|---|---|---|---|---|---|---|---|---|
| Fool me once | 0.83 | 0.09 | 0.3 | 0 | 0 | 0.23 | 0.77 | 0.37 |
| Big Bang | 0.17 | 0 | 0 | 0.2 | 0.45 | 0.36 | 0.16 | 0.23 |
| Bodies | 0.92 | 0.09 | 0.7 | 0.4 | 1 | 0.82 | 0.86 | 0.84 |
| Game of Thrones | 0.95 | 0.75 | 0.9 | 0.5 | 1 | 0.88 | 0.89 | 0.88 |
| Resilient Alien | 0.78 | 0.2 | 0.4 | 0.4 | 0.3 | 0.38 | 0.73 | 0.51 |
| Task7_New1 | 0.97 | 1 | 1 | 1 | 0 | 0.91 | 0.99 | 0.95 |
| Task7_New2 | 1 | 0.86 | 0.2 | 0.2 | 0.07 | 0.18 | 0.95 | 0.33 |

(a) Performance scores using D+ for Duration.

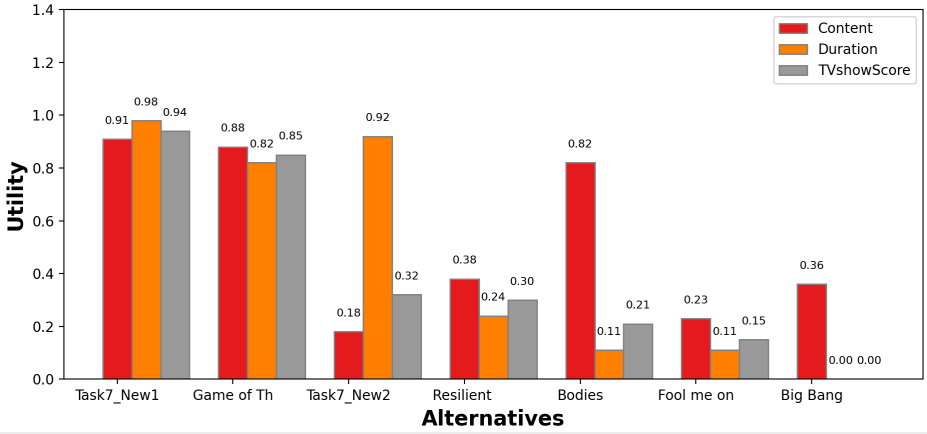
| | EpisodeLength | NumberEpisodes | Action | SciFi | Comedy | Content | Duration | TVshowScore |
|---|---|---|---|---|---|---|---|---|
| Fool me once | 0.83 | 0.09 | 0.3 | 0 | 0 | 0.23 | 0.11 | 0.15 |
| Big Bang | 0.17 | 0 | 0 | 0.2 | 0.45 | 0.36 | 0 | 0 |
| Bodies | 0.92 | 0.09 | 0.7 | 0.4 | 1 | 0.82 | 0.11 | 0.21 |
| Game of Thrones | 0.95 | 0.75 | 0.9 | 0.5 | 1 | 0.88 | 0.82 | 0.85 |
| Resilient Alien | 0.78 | 0.2 | 0.4 | 0.4 | 0.3 | 0.38 | 0.24 | 0.3 |
| Task7_New1 | 0.97 | 1 | 1 | 1 | 0 | 0.91 | 0.98 | 0.94 |
| Task7_New2 | 1 | 0.86 | 0.2 | 0.2 | 0.07 | 0.18 | 0.92 | 0.32 |

(b) Performance scores using C+ for Duration.

Figure 15: Comparison of scores using different polarities for Duration aggregation.




(a) Performance scores graph using D+ for Duration.




(b) Performance scores graph using D+ for Duration.

Figure 16: Comparison of score graph using different polarities for Duration aggregation.

Figure 16 and Figure 15 display the results of switching from a strongly conjunctive operator (C+) to a strongly disjunctive operator (D+) at the Duration node. Under the conjunctive setting, both *EpisodeLength* and *NumberEpisodes* must be high to achieve a strong score in *Duration*. In contrast, the disjunctive operator allows a high score in either criterion to compensate for a low score in the other, thereby boosting the overall *Duration* score even if one of the sub-criteria is lacking. This shift in polarity demonstrates how the model can become more forgiving or more demanding, depending on the decision maker's preference for balancing (OR-like) versus simultaneously satisfying (AND-like) sub-criteria.

# 9 Conclusion

In this task, we utilised the LSP method to build a decision support system for TV show selection. We defined a hierarchy of criteria, created utility functions to represent personal preferences, and specified weights and aggregation operators at each level. The analysis showed that choosing different operators (e.g., C+, DA, CA) and adjusting their polarities can significantly impact the final performance scores by focusing on either conjunctive or disjunctive behaviors. Additionally, changing the relative importance of criteria can also influence which alternatives achieve higher scores.

Overall, this exercise highlights the flexibility and power of multi-criteria decision support systems in tailoring the evaluation process to accurately reflect individual or organisational preferences.

## References

[1] Professor Aida Valls. Lecture slides on aggregation operators. Lecture slides, PMCDSS Course, 2025. Provided during lectures.