Argumentation-Based Recommendations: Fantastic Explanations and How to Find Them

Antonio Rago, Oana Cocarascu and Francesca Toni

Imperial College London

explAIn Workshop

25th April 2018

Presentation Overview

1. Background
   - Recommender Systems
2. Research Summary
   - Aspect-Item Recommender Systems
   - Predicted Rating Calculations
   - Argumentation Explanations
3. Conclusions

Background – Recommender Systems

- Two main types of methods:
  - Content-based filtering (operating on information about users and their tastes).
  - Collaborative filtering (looking at similar users and their preferences).

- Common methods:
  - Latent factor models (based on matrix factorization).
  - Nearest neighbour models between items or users.

- The Netflix Prize has shown that matrix factorization models are superior to NN models.

- 4 of the desirable features:
  - Transparency - explaining how systems work and showing how they predict ratings.
  - Scrutability - allowing feedback based on these explanations.
  - Trust - correcting the systems based on user feedback.
  - Effectiveness - increasing the systems' accuracy with regards to users' preferences.

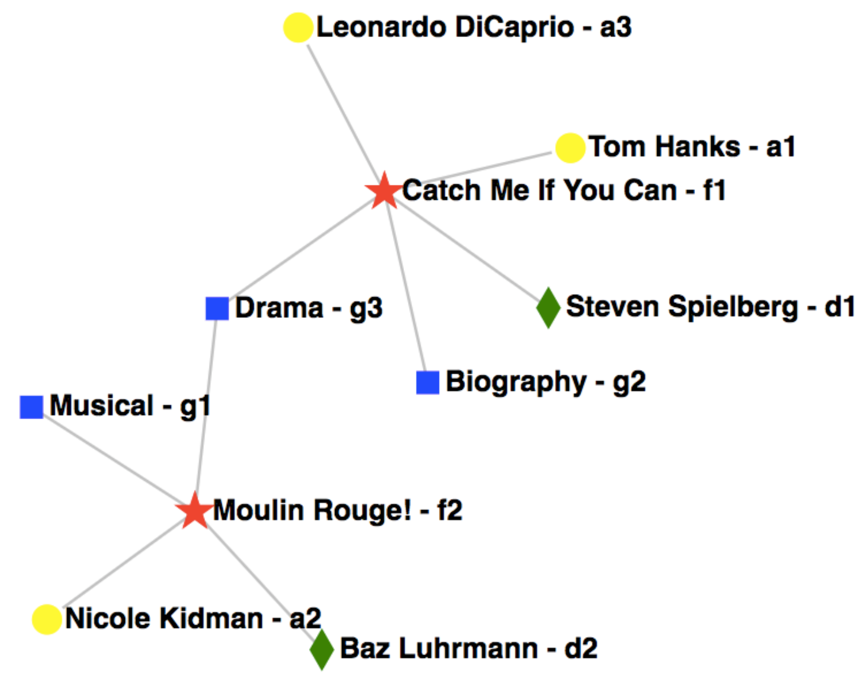
- Our method incorporates these features using *Argumentation-Based Explanations*.

Research Summary – Aspect-Item Recommender Systems

- We define a hybrid recommender system using an *Aspect-Item framework* (A-I):

Definition 1 An *Aspect-Item framework* (A-I) is a 6-tuple $\langle \mathcal{I}, \mathcal{A}, \mathcal{T}, \mathcal{L}, \mathcal{U}, \mathcal{R} \rangle$ such that:

- $\mathcal{I}$ is a finite, non-empty set of *items*;
- $\mathcal{A}$ is a finite, non-empty set of *aspects* and $\mathcal{T}$ is a finite, non-empty set of *types*, where each aspect in $\mathcal{A}$ has a unique type in $\mathcal{T}$; for any $t \in \mathcal{T}$, we use $\mathcal{A}_t$ to denote $\{a \in \mathcal{A} | \text{ the type of } a \text{ is } t\}$;
- the sets $\mathcal{I}$ and $\mathcal{A}$ are pairwise disjoint; we use $\mathcal{X}$ to denote $\mathcal{I} \cup \mathcal{A}$, and refer to it as the set of *item-aspects*;
- $\mathcal{L} \subseteq (\mathcal{I} \times \mathcal{A})$ is a symmetrical binary relation;
- $\mathcal{U}$ is a finite, non-empty set of *users*;
- $\mathcal{R}$ is a partial function of *ratings* such that $\mathcal{R} : \mathcal{U} \times \mathcal{X} \to [-1, 1]$.

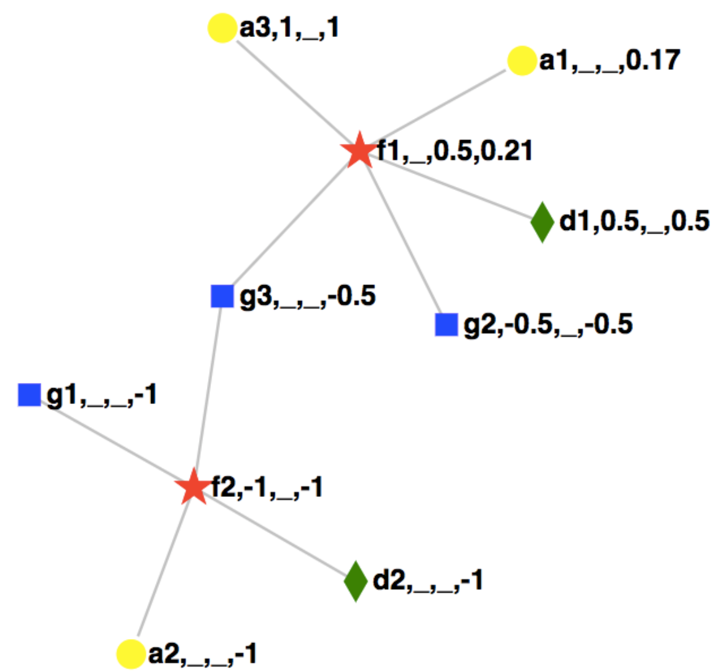


- This allows us to calculate predicted ratings each item-aspect for the user based on their ratings and similar users' ratings on item-aspects.

Research Summary – Predicted Rating Calculations

- Predicted ratings, based on user's and similar users' ratings, propagate through the graph.
  - Averaging techniques for each item-aspect's and type's effect on a predicted rating.
  - Unique weighting parameters for user similarities and preferences for each user .
- Unrated items with the highest predicted ratings are then recommended.
- Our method performs competitively when its **accuracy** is compared with ML techniques:

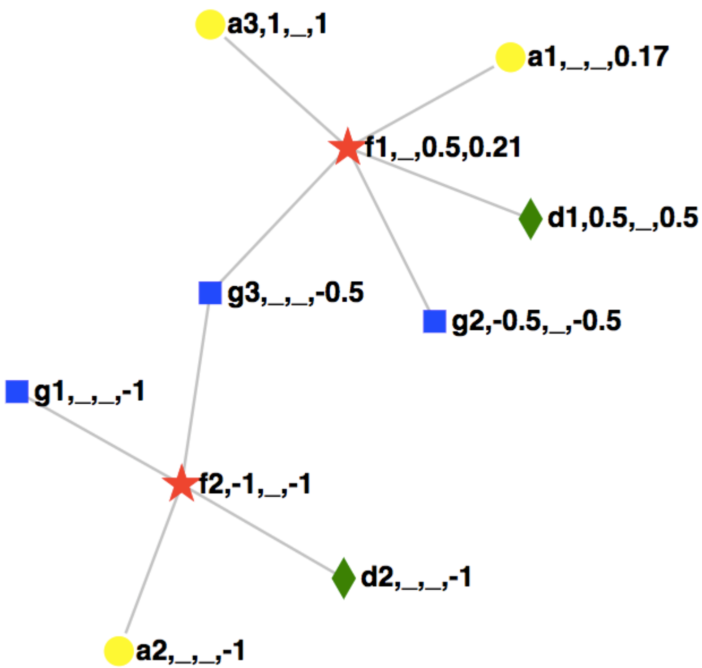
Model	Min #movies training set/ #movies 'cold-start'			
	10/5	20/5	20/7	20/10
Co-clustering	0.834	0.841	0.851	0.867
KNN	0.855	0.857	0.859	0.866
KNN with z score	0.855	0.853	0.864	0.875
NMF	0.837	0.842	0.853	0.861
Slope one	0.862	0.860	0.872	0.882
SVD	0.859	0.863	0.873	0.878
A-I model	**0.949**	**0.940**	**0.933**	**0.934**

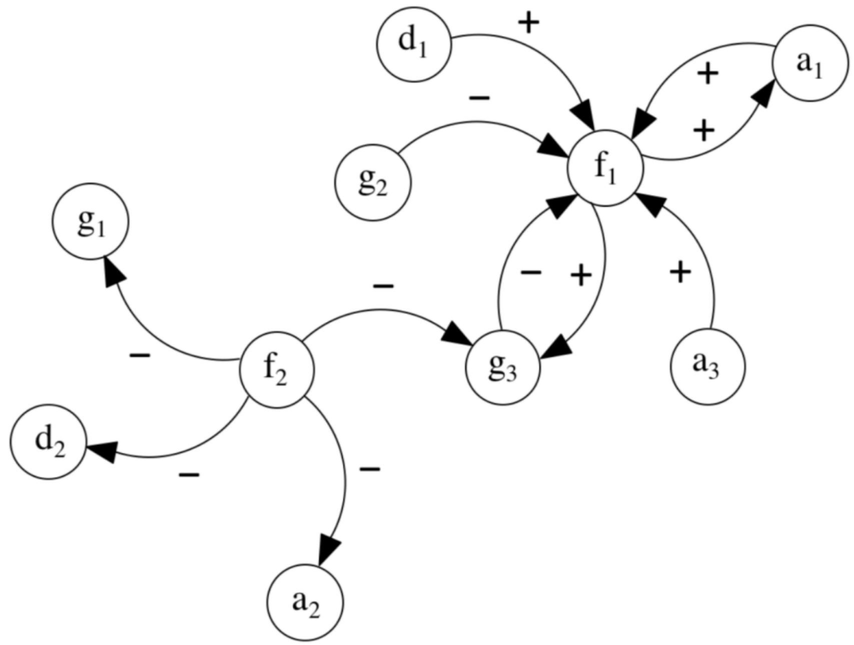


Key: id, user rating, average similar user rating, predicted rating

Research Summary – Argumentation Explanations

- A-I recommender systems allow argumentation readings of recommendations:





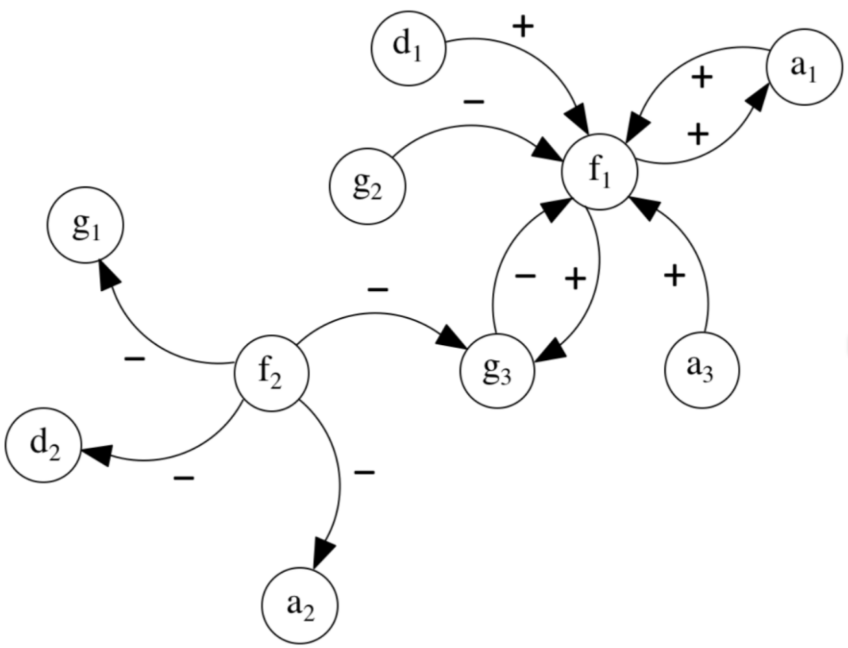
- Item-aspects are treated as arguments (that the user (dis)likes that item-aspect).

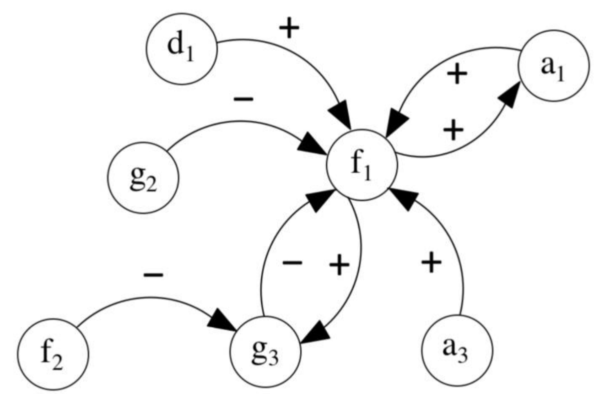
- The relations between arguments depends on user ratings for direction and (predicted) ratings for polarity.

Imperial College London

Research Summary – Argumentation Explanations

- *Argumentation explanations* can then be extracted.
- The argumentation explanation for f_1 is the subgraph in which all nodes have a path to f_1:



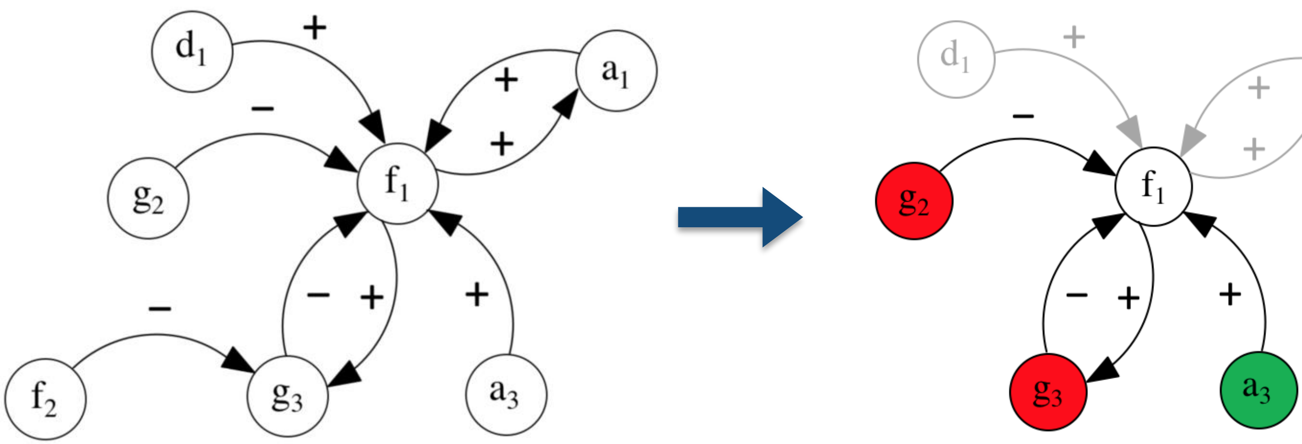


- Contains item-aspects which affected f_1's predicted rating and therefore its recommendation.

Research Summary – Argumentation Explanations

- Explanations allow users to interact with recommendations and provide feedback, e.g.:

User: “I did not enjoy Catch Me If You Can, why did you recommend it to me?”



Aspects with biggest positive and negative effects on f_1’s predicted rating highlighted

- The user’s positive rating on a_3 has therefore had the biggest effect. The response may be:

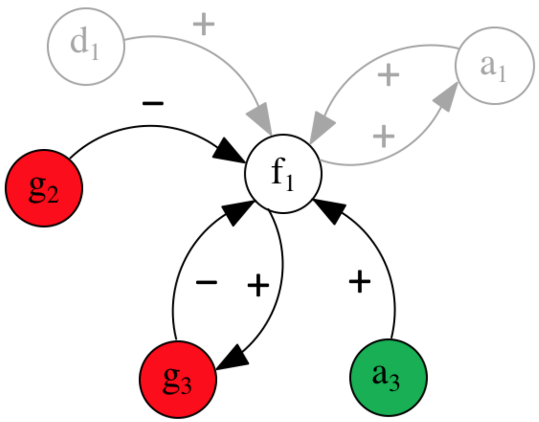
User: “I don’t care about the actors in a film, consider the actors in a film less.”

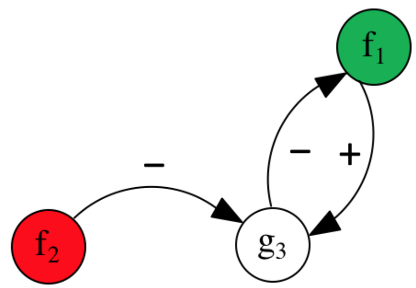
- This reduces the user’s unique constant for actors in a film, reduces f_1’s predicted rating and (we posit) improves the recommender system’s accuracy.

Research Summary – Argumentation Explanations

- Similarly, since the user hasn't given a rating to g_3, they could ask for reasoning on it:

User: "Why do you think I don't like the genre Drama?"





Items with biggest positive and negative effects on g_3's predicted rating highlighted

- Similar users' (positive, overall) ratings on f_1 have increased g_3's predicted rating:

User: "The users who rate Catch Me If You Can positively are not similar to me."

- The system also allows ratings to be changed, e.g. f_2 could be rated lower in this case.
- Reducing either the users' similarity or f_2's rating is guaranteed to reduce f_1's predicted rating.

Conclusions

- We have present a method for incorporating quantitative argumentation to recommender systems:
  - Aspect-Item frameworks provide a method for a hybrid recommender system.
  - Using a simple algorithm for calculating predicted ratings, the A-I Recommender Systems perform competitively with traditional ML methods.
  - A-I frameworks admit argumentation readings of recommendations in the form of argumentation frameworks.
  - These argumentation frameworks can be used as the underlying structure to provide visual and linguistic explanations of recommendations to users.
  - The explanations allow feedback to be provided by users, adjusting parameters in the system which further improve the accuracy of the system for each unique user.

Thank You

Any Questions?