The Same Thing – Only Different: Classification of Movies by their Story Types

Chang Liu, Armin Shmilovici, Mark Last

Department of Software and Information Systems Engineering

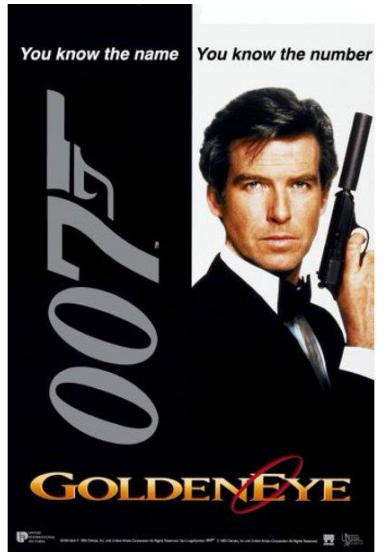
Ben-Gurion University of the Negev (Israel)

Text2Story'23 Workshop, Dublin

(Republic of Ireland), 2-April-2023

Introduction to Story-Types

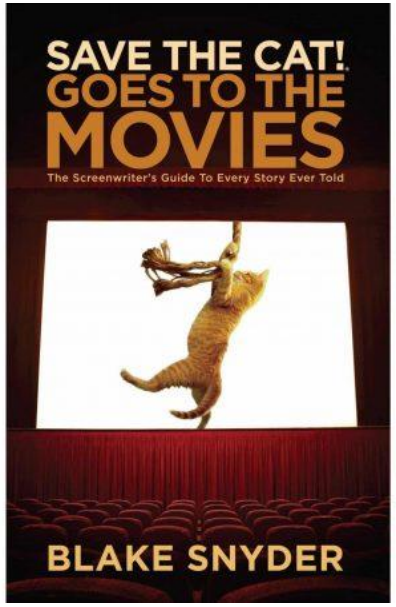
- Johnny English and 007 James Bond movies: both present a superficial story about a fictitious British secret agent who fights an evil secret organization that threatens the safety of the world. What is the difference?
  - Johnny English - The character is innocent about his clumsiness and ridiculed by the establishment in which he operates – A *Fool Triumphant* type of story
  - 007 James Bond – A character with a superhero-like fighting skills - A *Superhero* type of story
- The Hollywood cliché is of the studio executive telling the script-writer: “Give me the same thing... only different!”
- *Story types depict the development of movie stories in terms of the protagonist’s character traits and the motivations that drive him in facing his challenges.*



Motivation

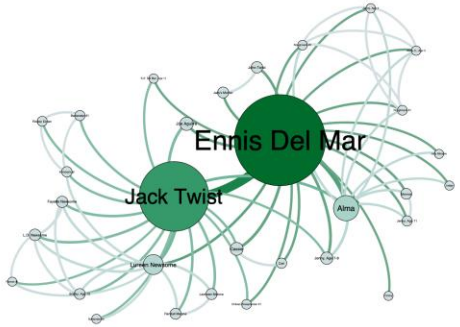
- Our Goal - Understand and recognize the narrative/story in movies
- Movies typically contain a relatively simple story
- *Our first objective* is to provide a labeled dataset of movies to facilitate the use of supervised machine learning algorithms for the problem of story types classification. https://github.com/llafcode/The-Movie-Narrative-Dataset-MND
- *Our second objective* is to provide a lightweight solution for the challenging task of story type classification, with the use of relatively simple methodology and features.
- *The original contributions* to the domain of computational narrative understanding in movies are two-fold:
  - a) We introduce the first benchmark dataset for the problem of story type classification that will be released to the research community; and
  - b) We demonstrate that the story type of a movie can be automatically detected using some relatively simple movie features.

The 10 Story-Types of Blake Snyder



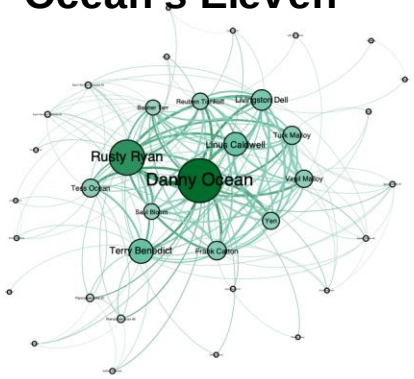
- Story types: *Monster in the House, Golden Fleece, Out of the Bottle, Dude with a Problem, Rites of Passage, Buddy Love, Whydunit, The Fool Triumphant, Institutionalized, Superhero*
- Story types ≠ genres (e.g., *Adventure, Horror, Comedy, Romantic)*
- **Hypothesis 1: Most movies are like that**
- **Hypothesis 2: Even non-experts can identify**
- **We would like to Classify of Movies by their Story Types**
- *We assume that simple features are sufficient to for classifying story-types*
- *e.g., a Buddy Love's character network will have two main highly connected nodes, while a Golden Fleece's character network will have many connected nodes.*

Brokeback Mountain



Buddy Love: An inadequate hero must rise above an extremely difficult situation to be with a uniquely unlikely partner who is the only one capable of bringing him peace.

Ocean's Eleven



Golden Fleece: A driven hero must lead a group of allies to retrieve a prized possession through a perilous journey that was not what the hero expected.

The 10 Story Types

- **Monster in the House:** A hero is forced to save a trapped group of people from being killed by a monster he inadvertently unleashed. Examples: Jurassic World, Jaws.
- **Golden Fleece:** A driven hero must lead a group of allies to retrieve a prized possession through a perilous journey that was not what the hero expected. Examples: Avengers, Infinity War, Ocean 8.
- **Out of the Bottle:** A greedy hero must learn to undo a spell he initiated before it turns into a curse he cannot undo. Examples: Liar Liar, Big.
- **Dude with a Problem:** An unsuspecting hero must survive at all costs when he is dragged into a life or death situation he did not see coming and cannot escape. Examples: 1917, the Martian.
- **Rites of Passage:** A troubled hero's only way to overcome a growing life crisis is to defeat his worst enemy – himself. Examples: Brooklyn, Inside Out.
- **Buddy Love:** An inadequate hero must rise above an extremely difficult situation to be with a uniquely unlikely partner who is the only one capable of bringing him peace. Examples: E.T., Zootopia.
- **Whydunit:** A devoted hero must find the truth behind an intriguing mystery before he is swallowed by the darkness he desperately seeks to expose. Examples: Captain Marvel, Bladerunner, The Silence of the Lambs.
- **The Fool Triumphant:** An innocent hero whose only way to defeat the prejudices of a group is to change himself without losing what made him the group's target of contempt in the first place – his uniqueness. Example: Moneyball.
- **Institutionalized:** An outsider whose only way to save his individuality is by going against the many who wish to make him like them. Examples: American Sniper.
- **Superhero:** A uniquely special hero must defeat an opponent with stronger capabilities by using the same powers that disconnect him from the people he hopes to save. Examples: Iron Man, Taken.

Dataset Collection and Preparation

- 45 cinema movies (from the MovieGraph dataset)
- 119 distinct human annotators (screened out of 180 applicants)
- 9 movies had Gold-Standard labeling (by expert script-writers)
- 3 movies per annotator (including a Gold-Standard movie)
- Each movie had at least 5 distinct annotators
- At least half of the annotators agreed on the same story-type for 40 out of the 45 movies (poor agreement for movies with multiple story-lines)
- Agreement on 8 out of 9 Gold-Standard labeling
- Conclusion: we verified the two hypotheses regarding data collection:
  - (1) Most movies adhere fairly well to the general structure, described by the screenwriting book of Snyder
  - (2) Even non-experts can identify those story types after watching a movie.

A Machine Learning Experiment

- 45 cinema movies split into 5 balanced folds
- 17 movies with a single label, 16 movies with two labels, 12 movies with three labels
- Due to highly imbalanced story-types, we merges classes: 26 movies of the *Buddy Love* type, 19 of the *Dude with a Problem* type, 13 of the *Rites of Passage* type, and 22 assigned with the *Other* label
- Feature Engineering (e.g., Character appearance networks, scenes/minute)
- DT for decision tree and NB for Naive Baye

Method	Prec.	Recall	F1	Acc.	Selected Features
Buddy Love					
Baseline	0.58	1.00	0.73	0.58	1. avg. edge weights
DT/NB+L1	**0.67***	0.92	0.77	**0.69***	2. Std edge weights 3. scenes/min
Dude with a Problem					
Baseline	0.42	1.00	0.45	0.29	1. # key char (3^{rd} quat)
NB+L1	**0.58***	0.75	0.63	**0.67***	2. scenes/min 3. number of edges
Rites of Passage					
Baseline	0.29	1.00	0.45	0.29	1. avg. degree 2. avg. scene duration
DT+L1	**0.75***	0.63	0.63	**0.80***	3. scenes/min 4. duration
Other					
Baseline	0.49	1.00	0.66	0.49	1. # key char (full) 2. # key char (3^{rd} quat)
NB+L1	**0.63***	0.81	0.70	**0.67***	3. std. edge weights 4. avg. edge weights

Discussion

- F1-measure in the range of [0.63-0.77]
- A Small and imbalanced dataset with few features, yet the first one.
- *Preliminary results suggest that simple movie features can be used by machine learning algorithms to detect the abstract concepts of Story-Types.*
- *A progress to the goal of Narrative Understanding*
- May be used for movie recommendation systems