Inducing Semantic Micro-Clusters from Deep Multi-View Representations of Novels

Lea Frermann[†], György Szarvas[‡]
l.frermann@ed.ac.uk, szarvasg@amazon.de
[†]University of Edinburgh [‡]Amazon Development Center Germany GmbH

Introduction

What should I read next?

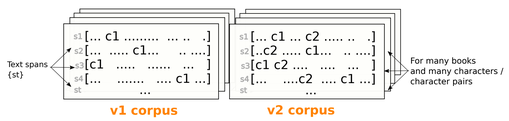
- Humans have multi-faceted book representations and preferences, e.g.,
  characters 'I want a book about space pirates'
  relations 'I want a book about a secret baby romance'
- Preferences are reflected by expert-provided book labels in the Amazon catalog reflect this
- Even fine-grained tags like 'secret baby romance'
- However, exhaustive and arbitrarily fine-grained manual tagging is prohibitive

Can we automatically induce **structured** and **interpretable** book representation that are useful in downstream tasks, e.g., content-based **recommendation**?

Contributions

1. Deep, interpretable multi-view plot representations of book plots
2. A general multi-view loss function to learn *distinct* views
3. A large-scale, empirical, task-oriented evaluation of plot representations

MVPlot Input



v1 s_t 20-word text spans mentioning **only** c1
v2 s_t 20-word text spans mentioning **only** c1 and c2
$\{s_t\}$ chronologically ordered $\{s_1, s_2, \ldots, s_T\}$

MVPlot Output

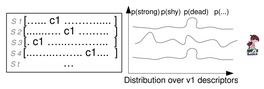
1. Global **property descriptors v1**, e.g.,

   smart clever knowledge pretty small slim blonde lonely desparate sad

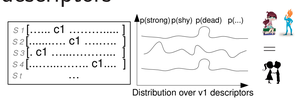
2. Global **relation descriptors v2**, e.g.,

   murder chase prosecute hate dislike fear rage dearest love like attract

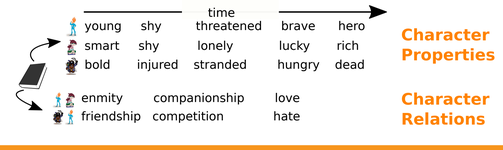
3. **Property trajectories** $\rightarrow \forall b \forall c$: sequence of probability distributions over **v1** descriptors



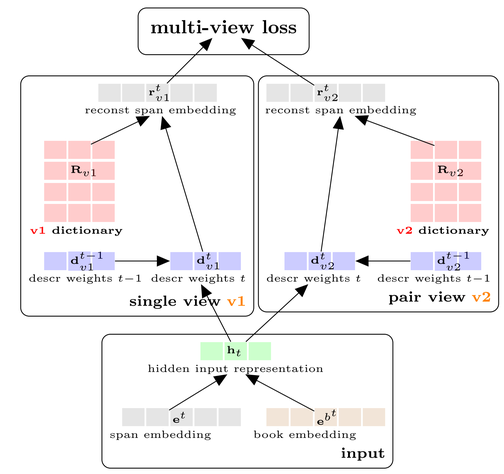
4. **Relationship trajectories** $\rightarrow \forall b \forall c1, c2$: sequence of distributions over **v2** descriptors



MVPlot Overview



MVPlot – Model Architecture



The Multi-View Loss Function: Hinge Loss with Contrastive Max-margin estimation

1) Within each view, efficiently represent and **accurately reconstruct** the input

2) Across views, encode relevant and **distinct aspects**

$$J(\theta) = \sum_t \beta \sum_{n \in N} \underbrace{max(0, 1 - \mathbf{r}^t_{v1}\mathbf{e}^t_{v1} + \mathbf{r}^t_{v1}\mathbf{e}^{n_i}_{v1})}_{\text{maximize margin between true input and negative samples (same level)}} + (1-\beta) \underbrace{max(0, 1 - \mathbf{e}^t_{v_1}\mathbf{r}^t_{v1} + \mathbf{e}^t_{v_1}\mathbf{r}^t_{v2})}_{\text{maximize margin between true input and reconstruction from other level's dictionary}} + \lambda \underbrace{||\mathbf{R}_{v1}\mathbf{R}^T_{v1} - \mathbf{I}||}_{\text{Regularizer}}$$

Examples of Induced Descriptors

Property Descriptors v1

laugh scream laughing yell joke cringe disgrace embarrassment hate cursing
snug fleece warm comfortable wet blanket flannel cozy comfort roomy

Relation Descriptors v2

love loving lovely dear sweetest dearest thank darling congratulation hello
associate assistant senior chairman executive leadership vice director liaison

An Empirical and Scalable Evaluation Framework

Genre	Amazon Catalog Tags – Character Types
Mystery	British Detectives; Private Investigators; Female Protagonists;
Romance	Cowboys; Criminals & Outlaws; Doctors; Royalty & Aristocrats
SciFi	AIs; Aliens; Clones; Corporations; Mutants; Pirates; Psychics

- expert (publisher) provided gold standard tags
- assume: books that share a tag are indeed similar
- evaluate local neighborhoods in model space of 10,000 books

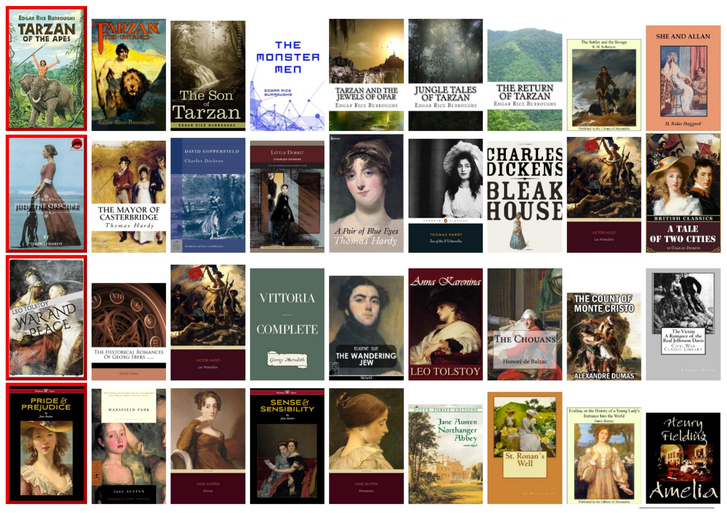
Book Similarity in Model Space

$\forall b, \forall \{c\} \in b$ c : averaged property trajectory
$\forall b, \forall \{c1, c2\} \in b$ p : averaged relation trajectory
Align two books $b1, b2$ s.th. Euclidean distance is minimized

We can cluster books based on different views and ask

- Which books have similar character types (v1)?
- Which books have similar character pair relations(v2)?
- Which books are similar overall (v1,v2)?

Examples of Local Book Neighborhoods

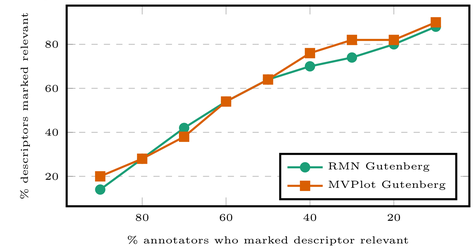


Data

	Amazon	Gutenberg
# books	10,000	3,500
# v1 sequences	91,511	45,182
# v2 sequences	70,156	60,493

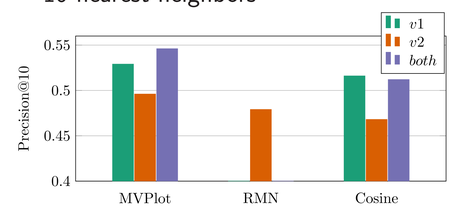
Descriptor Quality – Results

- RMN lyyer et al (2016)
- Data: Gutenberg, **v2** descriptors
- MTurk: "Do words describe a relation?" y / n



Book Neighborhood Quality – Results

- RMN lyyer et al (2016), Cosine baseline
- Data: Amazon
- 500 / 10K reference books
  10 nearest neighbors



Discussion & Conclusions

General Deep Multi-view Learning

- arbitrary number of views
- arbitrary data given view-specific corpora (e.g., social sciences)

Scalable Extrinsic Evaluation

- large set of diverse gold book tags
- adaptable to book properties of interest

RMN: Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber, and Hal Daumé III. *Feuding families and former friends: Unsupervised learning for dynamic fictional relationships.* NAACL 2016.
MVPlot: Lea Frermann, György Szarvas. *Inducing Semantic Micro-Clusters from Deep Multi-View Representations of Novels.* EMNLP 2017.

http://www.frermann.de/