CAP 6412 Advanced Computer Vision


Boqing Gong
March 15, 2016
Last week: Spring Break

- **ECCV**
  - One of the top conferences on Computer Vision
  - Every other year (2016, 2014, 2012, ...)
  - On par with CVPR (every year), ICCV (every other year, 2015, 2013, ...)
  - ~2000 submissions, ~22% acceptance rate

- **Double-blind peer review**
  - Conference attendees vote for Program Chairs (PCs)
  - PCs select Area Chairs (ACs) and distribute papers to ACs
  - ACs assign each of their papers to at least three Reviewers
  - Reviewers read and evaluate papers
  - Authors respond to reviewers
  - Reviewers read author responses and discuss the papers
  - ACs meet together and decide whether to accept or to reject papers (orals, posters)
Last week: Spring Break

• A typical program of CVPR/ECCV/ICCV
  • Sunday: Tutorials
  • Monday—Thursday: Oral presentations, spotlights, poster presentations
  • Friday—Saturday: Workshops
<table>
<thead>
<tr>
<th>Date</th>
<th>Reading</th>
<th>Authors</th>
<th>Title</th>
<th>Conference/Publication Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next</td>
<td></td>
<td>Tuesday (03/22)</td>
<td>&amp; Secondary papers</td>
<td></td>
</tr>
</tbody>
</table>
Assignment 9: Due on 03/22, 12pm


Next week: DAG-CNN & Transferability

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Topic</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niladri Basu Bal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mert Ozerdem</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What’s next

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
<th>Volunteer for class Tuesday</th>
<th>Volunteer for class Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 2</td>
<td>CNN &amp; object recognition</td>
<td>Dustin Morley</td>
<td>Jason Tiller</td>
</tr>
<tr>
<td>Week 3</td>
<td>CNN &amp; object localization</td>
<td>Samer Iskander</td>
<td>Syed Ahmed</td>
</tr>
<tr>
<td>Week 4</td>
<td>CNN &amp; transfer learning</td>
<td>Harish Raviprakash</td>
<td>Karan Daei-Mojdehi</td>
</tr>
<tr>
<td>Week 5</td>
<td>CNN &amp; segmentation, super-resolution</td>
<td>Jose Sanchez (Super-resolution)</td>
<td>Goran Igi</td>
</tr>
<tr>
<td>Week 6</td>
<td>CNN &amp; videos (optical flow, pose)</td>
<td>Abdullah Jamal</td>
<td>Amar Kelu Nair (Optical Flow)</td>
</tr>
<tr>
<td>Week 7</td>
<td>Image captioning</td>
<td>Suhas Nithyanandappa (Image captioning)</td>
<td>Nandakishore Puttashamachar</td>
</tr>
<tr>
<td>Week 8</td>
<td>Visual question answering</td>
<td>Javier Lores</td>
<td>Aisha Urooj</td>
</tr>
<tr>
<td>Week 9</td>
<td>Attention model, aligning books with movies</td>
<td>Fareeha Irfan (aligning books w movies)</td>
<td>shrayas somashekar</td>
</tr>
<tr>
<td>Week 10</td>
<td>DAG-CNN &amp; transferability</td>
<td>Niladri Basu Bal</td>
<td>Mert Ozerdem</td>
</tr>
<tr>
<td>Week 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What’s next

• Sign up for volunteer presentations at
  • https://docs.google.com/spreadsheets/d/1DxMQ_RVMx8BLmc5gi51dtXI2ZJzktR8EzJJZkVpGW4/edit#gid=0

• Suggest papers you would like to read, share, challenge
Today

• Administrivia
• Recurrent Neural Networks (RNNs) (II)
• OCR in the wild, by Aisha
(Discrete-time) RNN

- Three time steps and beyond

\[
x_1, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots, x_T
\]
\[
h_t = \sigma \left( W^{(hh)} h_{t-1} + W^{(hx)} x_t \right)
\]
\[
\hat{y}_t = \text{softmax} \left( W^{(S)} h_t \right)
\]

Image credits: Richard Socher
(Discrete-time) RNN

- Three time steps and beyond
- A layered feedforward net
- Tied weights for different time steps
- Conditioning (memorizing?) on all previous input
- Cheap to save memory in RAM

Image credits: Richard Socher
Detour: Hidden Markov Model

• A probabilistic model of sequences

\[ x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \]

\[ y_1 \rightarrow y_2 \rightarrow y_3 \rightarrow y_4 \]

• Emission probability: \( P(y_i \mid x_i) \)
• Transition probability: \( P(x_i \mid x_{i-1}) \)
• Initial probability: \( P(x_0) \)

Image credits: Erik Sudderth
Detour: Hidden Markov Model

• Useful for modeling sequences
• Discrete hidden states, which satisfy Markov assumption

• Inference and learning (optional)
  • Evaluation: forward probability
  • Decoding: forward-backward algorithm, Viterbi decoding
  • Learning: EM algorithm (Baum-Welch)

Image credits: Erik Sudderth
Begin: Slides from Geoffrey Hinton
Hidden Markov Models (computer scientists love them!)

- Hidden Markov Models have a discrete one-of-N hidden state. Transitions between states are stochastic and controlled by a transition matrix. The outputs produced by a state are stochastic.
  - We cannot be sure which state produced a given output. So the state is “hidden”.
  - It is easy to represent a probability distribution across N states with N numbers.

- To predict the next output we need to infer the probability distribution over hidden states.
  - HMMs have efficient algorithms for inference and learning.
A fundamental limitation of HMMs

• Consider what happens when a hidden Markov model generates data.
  • At each time step it must select one of its hidden states. So with N hidden states it can only remember log(N) bits about what it generated so far.

• Consider the information that the first half of an utterance contains about the second half:
  • The syntax needs to fit (e.g. number and tense agreement).
  • The semantics needs to fit. The intonation needs to fit.
  • The accent, rate, volume, and vocal tract characteristics must all fit.

• All these aspects combined could be 100 bits of information that the first half of an utterance needs to convey to the second half. $2^{100}$ is big!
Recurrent neural networks

• RNNs are very powerful, because they combine two properties:
  • Distributed hidden state that allows them to store a lot of information about the past efficiently.
  • Non-linear dynamics that allows them to update their hidden state in complicated ways.
• With enough neurons and time, RNNs can compute anything that can be computed by your computer.
Do generative models need to be stochastic?

• Linear dynamical systems and hidden Markov models are stochastic models.
  • But the posterior probability distribution over their hidden states given the observed data so far is a deterministic function of the data.

• Recurrent neural networks are deterministic.
  • So think of the hidden state of an RNN as the equivalent of the deterministic probability distribution over hidden states in a linear dynamical system or hidden Markov model.
Recurrent neural networks

• What kinds of behaviour can RNNs exhibit?
  • They can oscillate. **Good for motor control?**
  • They can settle to point attractors. **Good for retrieving memories?**
  • They can behave chaotically. **Bad for information processing?**
  • RNNs could potentially learn to implement lots of small programs that each capture a nugget of knowledge and run in parallel, interacting to produce very complicated effects.

• But the computational power of RNNs makes them very hard to train.
  • For many years we could not exploit the computational power of RNNs despite some heroic efforts (e.g. Tony Robinson’s speech recognizer).
End: Slides from Geoffrey Hinton
Today

• Administrivia
• Recurrent Neural Networks (RNNs) (II)
• OCR in the wild, by Aisha
Upload slides before or after class

• See “Paper Presentation” on UCF webcourse

• Sharing your slides
  • Refer to the originals sources of images, figures, etc. in your slides
  • Convert them to a PDF file
  • Upload the PDF file to “Paper Presentation” after your presentation
Aligning Books and Movies: Towards Story-like Visual Explanations by Watching Movies and Reading Books

Yukun Zhu, Ryan Kiros, Richard Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, Sanja Fidler
Motivation

Books provide rich fine-grained information:

- **Appearance**: how a character, an object or a scene looks like
- **High-level semantics**: what someone is thinking, feeling and how these states evolve through a story

**Aim**: Align books with its movie releases in order to provide rich descriptive explanations for visual content that go semantically far beyond the captions available in current datasets.
Overview of Evaluations in this work

- Evaluate similarity between sentences, trained by a corpus of Books
- Similarity between Movie clips and sentences in Books, via learnt embedding
- Weaving “context” of sentences into similarity evaluation using a 3-layer CNN
- Timeline wise alignment between books and movie using Conditional Random Field
Demonstrated Results

- Movie/Book Alignment
- Describing a particular movie shot via its corresponding explanation in Book
- Given a Movie, retrieve its corresponding Book from a corpus of Books
- Caption a Movie clip w/ a paragraph from “any” Book
- Caption MS-COCO images using paragraph from Books
Related Work

- Early work on movie-to-text alignment include dynamic time warping for aligning movies to scripts with the help of subtitles [5, 4].
- Sankar et al. [28] further developed a system which identified sets of visual and audio features to align movies and scripts without making use of the subtitles.
- Aligning plot synopses to shots in the TV series for story-based content retrieval. This work adopts a similarity function between sentences in plot synopses and shots based on person identities and keywords in subtitles.
Challenges

Plot synopsis closely follow the storyline of movies, books are more verbose and might vary in the storyline from their movie release.

- **Parallel to our work:** [BOOK2MOVIE]
  - Tapaswi et al. aims to align scenes in movies to chapters in the book.
    - Coarse approach: Operates on chapter-level
    - Dataset evaluates on 90 scene-chapter correspondences,
    - Matches the presence of characters in a scene to those in a chapter
    - Uses hand-crafted similarity measures between sentences in the subtitles and dialogs in the books
  - This paper:
    - Our approach: sentence/paragraph level.
    - Dataset draws 2,070 shot-to-sentences alignments.
    - We remove character names to learn semantics and context better
    - CNN learnt similarity function
Dataset and Ground Truth

Annotators find correspondences:

- Mark **exact time** in movie to **line number** of beginning of the matched sentence.
  - If a shot is longer: Indicate time of ending.
  - If description in more lines: Indicate the last line.

- Tag alignment: ‘visual’, ‘dialogue’ or ‘audio’ match

**Total**: 11 movie/book pairs, 2070 correspondences
Training Sentence Similarity model


Overview

Sentence $i$ is encoded. Conditioning on this the model tries to reconstruct the sentence before and after it.

Motivation

Sentences that have similar surrounding context are likely to be both semantically and syntactically similar.
“Thought Vectors”

The term is popularized by Geoffrey Hinton (Google).

word vector:

- represents the word’s meaning, relative to others (context)
- linked by grammar

thought vector:

- represents a thought, relative to others
- linked by chain of reasoning

Goal: Feed enough data (thoughts) to NN, enabling it to mimic those thoughts
Sentence Embedding Model using Skip Thought

**Training Data:** Corpus of 11,038 books from the web

| he drove down the street off into the distance . | he started the car , left the parking lot and merged onto the highway a few miles down the road .  
| | he shut the door and watched the taxi drive off .  
| | she watched the lights flicker through the trees as the men drove toward the road .  
| | he jogged down the stairs , through the small lobby , through the door and into the street .  
| | a messy business to be sure , but necessary to achieve a fine and noble end .  
| | they saw their only goal as survival and logically planned a strategy to achieve it .  
| | there would be far fewer casualties and far less destruction .  
| | the outcome was the lisbon treaty .  
| the most effective way to end the battle . |  

Table 3: Qualitative results from the sentence embedding model. For each query sentence on the left, we retrieve the 4 nearest neighbor sentences (by inner product) chosen from books the model has not seen before.
Sequence to Sequence Learning

• Given a tuple \((s_{i-1}, s_i, s_{i+1})\) of contiguous sentences:
  – the sentence \(s_i\) is encoded using LSTM.
  – the sentence \(s_i\) attempts to reconstruct the previous sentence and next sentence \(s_{i+1}\).

• The input is the sentence triplet:
  – I got back home.
  – I could see the cat on the steps.
  – This was strange.
Encoder

\(w_i^1, \ldots, w_i^N\) be the words in sentence \(s_i\)  
\(N\) is the number of words

At each time step, the encoder produces a hidden state \(h^t_i\).

\[
\begin{align*}
    r^t &= \sigma(W_r x^t + U_r h^{t-1}) \\
    z^t &= \sigma(W_z x^t + U_z h^{t-1}) \\
    \bar{h}^t &= \tanh(W x^t + U(r^t \odot h^{t-1})) \\
    h^t &= (1 - z^t) \odot h^{t-1} + z^t \odot \bar{h}^t
\end{align*}
\]

\(\bar{h}^t\) is the proposed state update at time \(t\)

\(z^t\) is the update gate,

\(r^t\) is the reset gate
Decoder

\[ \begin{align*}
  r^t &= \sigma(W_r^d x^{t-1} + U_r^d h^{t-1} + C_r h_i) \\
  z^t &= \sigma(W_z^d x^{t-1} + U_z^d h^{t-1} + C_z h_i) \\
  \bar{h}^t &= \tanh(W^d x^{t-1} + U^d (r^t \odot h^{t-1}) + C h_i) \\
  h_{i+1}^t &= (1 - z^t) \odot h^{t-1} + z^t \odot \bar{h}^t
\end{align*} \]
**Encoder**

\[
\begin{align*}
    r^t &= \sigma(W_r x^t + U_r h^{t-1}) \\
    z^t &= \sigma(W_z x^t + U_z h^{t-1}) \\
    \bar{h}^t &= \tanh(W x^t + U (r^t \odot h^{t-1})) \\
    h^t &= (1 - z^t) \odot h^{t-1} + z^t \odot \bar{h}^t
\end{align*}
\]

**Decoder**

\[
\begin{align*}
    r^t &= \sigma(W'^d x^{t-1} + U'^d h^{t-1} + C_r h_i) \\
    z^t &= \sigma(W'_z x^{t-1} + U'_z h^{t-1} + C_z h_i) \\
    \bar{h}^t &= \tanh(W'^d x^{t-1} + U'^d (r^t \odot h^{t-1}) + C h_i) \\
    h_{i+1}^t &= (1 - z^t) \odot h^{t-1} + z^t \odot \bar{h}^t
\end{align*}
\]
Learning Objective

• We are given a tuple \((s_{i-1}, s_i, s_{i+1})\) of contiguous sentences.

• **Objective:** The sum of the log-probabilities for the next and previous sentences conditioned on the encoder representation:

\[
\sum_t \log P(w_{i+1}^t | w_i^{<t}, h_i) + \sum_t \log P(w_{i-1}^t | w_i^{<t}, h_i)
\]

Forward sentence  
Previous sentence
Aligning Book w/ Movie

Inspired from the Multimodal Neural Model in Kiros et al.

**Our approach:** Embedding learnt

**Training Data**

- Each **Clip Description:**
  - *Descriptive Video Service* dataset used to learn embedding
  - 94 movies, 54000 described clips
  - Pre-processing: Replace names with token ‘someone’

- Each **Movie Clip** vector:
  - mean-pooled features (GoogLeNet and hybrid-CNN) across each frame in the clip
Vector Representation of description using LSTM:

\[
\begin{align*}
    i^t &= \sigma(W_{xi}x^t + W_{hi}m^{t-1} + W_{ci}c^{t-1}) \\
    f^t &= \sigma(W_{xf}x^t + W_{hf}m^{t-1} + W_{cf}c^{t-1}) \\
    a^t &= \tanh(W_{xc}x^t + W_{hc}m^{t-1}) \\
    c^t &= f^t \odot c^{t-1} + i^t \odot a^t \\
    o^t &= \sigma(W_{xo}x^t + W_{ho}m^{t-1} + W_{co}c^t) \\
    m^t &= o^t \odot \tanh(c^t)
\end{align*}
\]

The states \((i^t, f^t, c^t, o^t, m^t)\) correspond to input, forget, cell, output and memory vectors for embedding word \(x^t\) of sentence at time \(t\).

\[m^N = m\] is vector representation of sentence of length \(N\).
Aligning Book w/ Movie

Let \( q \) be a movie clip vector and its embedding

\[
\mathbf{v} = \mathbf{W}_I q
\]

Scoring Function: \( s(m, v) = m \cdot v \)

Optimize pairwise ranking loss:

\[
\min_{\theta} \sum_m \sum_k \max\{0, \alpha - s(m, v) + s(m, v_k)\} + \sum_v \sum_k \max\{0, \alpha - s(v, m) + s(v, m_k)\},
\]

where, \( m_k \) is non-descriptive vector for embedding \( v \) and \( v_k \) contrastive clip vector for sentence vector \( m \)

Model trained with *Stochastic Gradient Descent w/o momentum*
Context Aware Similarity

*Local-level ambiguity:*

Despite being a dark novel, *Gone Girl* has 15 instances of "I love you". Match not isolated from surrounding context.

To compute **dialogue similarity:**

- **BLEU**: to find near identical
- **Tf-idf** *(term frequency-inverse document frequency)*: find duplicates but weighting down less frequent words
- **Skip-thought**: to find semantically similar paraphrased sentences
Context Aware Similarity

To obtain **Similarity**, we take into account:

- Individual similarity measures
- Fixed context window, in movie and book

Stack a set of $M$ similarity measures into a tensor $S(i, j, m)$, where

- $i$: indices of sentences in the subtitle
- $j$: in the book
- $m$: individual similarity measures
Context Aware Similarity

M = 9 similarities measures used:

- Visual and sentence embedding
- BLEU1-5
- tf-idf
- A uniform prior

To predict a combined score\( (i, j) = f(S(I, J,M)) \) at each location \((i, j)\) based on all similarity measures:

**3-layer Convoluted Neural Network**, with ReLU nonlinearity and dropout.

Cross-entropy optimized over training with Adam’s Algorithm
Figure 3: Our CNN for context-aware similarity computation. It has 3 conv. layers and a sigmoid layer on top.
Global Movie/Book Alignment

Most scenes follow a *timeline*.

Dynamic time warping not suitable since storyline can have crossings in time.
Global Movie/Book Alignment

Movie/Book alignment modelled as inference in a Conditional Random Field

- Each node $y_i$: alignment (shot w/ subtitle, sentence in the book)
- State space: set of all sentences in the book.
- **CRF energy:**

$$- \log p(x,y; \omega) = \sum_{i=1}^{K} \omega_u \phi_u(y_i) + \sum_{i=1}^{K} \sum_{j \in N(i)} \omega_p \psi_p(y_i, y_j)$$

- $K$: number of nodes (shots)
- $N(i)$: the left and right neighbor of $y_i$
- $\phi_u(\cdot)$ unary potential: output of CNN
- $\psi_p(\cdot)$ pairwise potentials measured by

$$\psi_p(y_i, y_j) = \frac{(d_s(y_i, y_j) - d_b(y_i, y_j))^2}{(d_s(y_i, y_j) - d_b(y_i, y_j))^2 + \sigma^2}$$

- $d_s(y_i, y_j)$ time span between two neighbouring sentences in the subtitle
- $d_b(y_i, y_j)$ distance of their state space in the book
- $\sigma^2$ is a robustness parameter to avoid punishing giant leaps too harsh

Pairwise potential to sure state consistency and incorporating long silence in the movie

$$\psi_q(y_i, y_j) = \frac{(d_b(y_i, y_j))^2}{(d_b(y_i, y_j))^2 + \sigma^2}$$
Evaluation

**Model:** CNN + CRF  
**Dataset:** 11 Books/Movies  
**Training Data:** 1 Book/Movie *Gone Girl*  
**Test Data:** Remaining 10 Movies

Recalled paragraph/shot considered Ground Truth, if:

- Paragraph at most 3 paragraphs away  
- Shot was at most 5 subtitles away

**Average Precision** reported at multiple alignment thresholds
## Evaluation: Movie/Book Alignment

<table>
<thead>
<tr>
<th>Movie/Book</th>
<th>UNI AP</th>
<th>UNI Recall</th>
<th>SVM AP</th>
<th>SVM Recall</th>
<th>1 layer CNN w/o one feature</th>
<th>CNN-3</th>
<th>CRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight Club</td>
<td>1.22</td>
<td>0.73</td>
<td>0.45</td>
<td>0.41</td>
<td>0.50</td>
<td>1.95</td>
<td>5.17</td>
</tr>
<tr>
<td>The Green Mile</td>
<td>14.05</td>
<td>10.38</td>
<td>12.26</td>
<td>12.74</td>
<td>11.79</td>
<td>17.92</td>
<td>19.81</td>
</tr>
<tr>
<td>Harry Potter and the Sorcerers Stone</td>
<td>0.00</td>
<td>0.00</td>
<td>6.92</td>
<td>10.12</td>
<td>9.83</td>
<td>28.80</td>
<td>27.60</td>
</tr>
<tr>
<td>American Psycho</td>
<td>0.00</td>
<td>0.00</td>
<td>8.09</td>
<td>8.18</td>
<td>7.84</td>
<td>74.13</td>
<td>78.23</td>
</tr>
<tr>
<td>One Flew Over the Cuckoo Nest</td>
<td>5.68</td>
<td>1.01</td>
<td>8.14</td>
<td>6.27</td>
<td>8.49</td>
<td>14.83</td>
<td>21.13</td>
</tr>
<tr>
<td>Shawshank Redemption</td>
<td>0.00</td>
<td>0.00</td>
<td>8.60</td>
<td>8.89</td>
<td>7.99</td>
<td>19.33</td>
<td>19.96</td>
</tr>
<tr>
<td>The Firm</td>
<td>0.05</td>
<td>1.38</td>
<td>7.91</td>
<td>8.66</td>
<td>6.22</td>
<td>18.34</td>
<td>20.74</td>
</tr>
<tr>
<td>Brokeback Mountain</td>
<td>2.36</td>
<td>27.00</td>
<td>16.55</td>
<td>17.82</td>
<td>15.16</td>
<td>31.80</td>
<td>30.58</td>
</tr>
<tr>
<td>The Road</td>
<td>1.12</td>
<td>41.90</td>
<td>6.58</td>
<td>7.83</td>
<td>5.11</td>
<td>19.80</td>
<td>19.58</td>
</tr>
<tr>
<td>No Country for Old Men</td>
<td>0.00</td>
<td>0.00</td>
<td>9.00</td>
<td>9.39</td>
<td>9.40</td>
<td>28.75</td>
<td>30.45</td>
</tr>
<tr>
<td>Mean Recall</td>
<td>3.88</td>
<td>38.01</td>
<td>52.66</td>
<td>52.95</td>
<td>47.07</td>
<td>66.77</td>
<td>69.10</td>
</tr>
<tr>
<td>AP</td>
<td>0.40</td>
<td>10.97</td>
<td>9.62</td>
<td>9.88</td>
<td>5.94</td>
<td>22.51</td>
<td>23.17</td>
</tr>
</tbody>
</table>

Table 4: Performance of our model for the movies in our dataset under different settings and metrics.
"Certainly, Mr. Cheswick. A vote is now before the group. Will a show of hands be adequate, Mr. McMurphy, or are you going to insist on a secret ballot?" "I want to see the hands. I want to see the hands that don't go up, too."

"Everyone in favor of changing the television time to the afternoon, raise his hand."

I took the envelope and left the rock where Andy had left it, and Andy's friend before him.

Dear Red, if you're reading this, then you're out. One way or another, you're out. And if you've followed along this far, you might be willing to come a little further. I think you remember the name of the town, don't you? I could use a good man to help me get my project on wheels. Meantime, have a drink on me-and do think it over. I will be keeping an eye out for you. Remember that hope is a good thing, Red, maybe the best of things, and no good thing ever dies. I will be hoping that this letter finds you, and finds you well.

Your friend, Peter Stevens I didn't read that letter in the field.
"Yes, Patrick?" She reenters the office trying to downplay her eagerness.

"Would you like to accompany me to dinner?" I ask, still staring at the crossword, gingerly erasing the m in one of the many meals I've filled the puzzle with. "That is, if you're not... doing anything."

"Oh no," she answers too quickly and then, I think, realizing this quickness, says, "I have no plans."

---

You need any help?

American Psycho

He realized he must be in the hospital wing. He was lying in a bed with white linen sheets, and next to him was a table piled high with what looked like half the candy shop.

"Tokens from your friends and admirers," said Dumbledore, beaming. "What happened down in the dungeons between you and Professor Quirrell is a complete secret, so, naturally, the whole school knows. I believe your friends Misters Fred and George Weasley were responsible for trying to send you a toilet seat. No doubt they thought it would amuse you. Madam Pomfrey, however, felt it might not be very hygienic, and confiscated it."
## Evaluation: Book Retrieval

<table>
<thead>
<tr>
<th>MOVIE</th>
<th>BOOKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight Club</td>
<td>100.0</td>
</tr>
<tr>
<td>Green Mile</td>
<td>100.0</td>
</tr>
<tr>
<td>Harry Potter</td>
<td>100.0</td>
</tr>
<tr>
<td>American Psy.</td>
<td>100.0</td>
</tr>
<tr>
<td>Shawshank</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BOOKS</th>
<th>MOVIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight Club</td>
<td>45.4</td>
</tr>
<tr>
<td>No Country</td>
<td>45.2</td>
</tr>
<tr>
<td>One Flew</td>
<td>45.1</td>
</tr>
<tr>
<td>The Road</td>
<td>43.6</td>
</tr>
<tr>
<td>The Firm</td>
<td>43.0</td>
</tr>
<tr>
<td>American Psy.</td>
<td>42.7</td>
</tr>
<tr>
<td>One Flew</td>
<td>40.1</td>
</tr>
<tr>
<td>Stephen King</td>
<td>39.6</td>
</tr>
<tr>
<td>The Firm</td>
<td>38.9</td>
</tr>
<tr>
<td>The Firm</td>
<td>38.0</td>
</tr>
<tr>
<td>The Road</td>
<td>36.7</td>
</tr>
<tr>
<td>The Firm</td>
<td>39.7</td>
</tr>
<tr>
<td>One Flew</td>
<td>39.5</td>
</tr>
<tr>
<td>Stephen King</td>
<td>39.1</td>
</tr>
<tr>
<td>The Firm</td>
<td>39.0</td>
</tr>
<tr>
<td>The Road</td>
<td>38.7</td>
</tr>
<tr>
<td>The Firm</td>
<td>55.5</td>
</tr>
<tr>
<td>One Flew</td>
<td>54.9</td>
</tr>
<tr>
<td>Stephen King</td>
<td>53.5</td>
</tr>
<tr>
<td>The Firm</td>
<td>53.1</td>
</tr>
<tr>
<td>The Road</td>
<td>52.6</td>
</tr>
<tr>
<td>No Country</td>
<td>51.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BOOKS</th>
<th>MOVIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight Club</td>
<td>84.0</td>
</tr>
<tr>
<td>No Country</td>
<td>62.0</td>
</tr>
<tr>
<td>One Flew</td>
<td>61.4</td>
</tr>
<tr>
<td>The Road</td>
<td>60.9</td>
</tr>
<tr>
<td>The Firm</td>
<td>59.1</td>
</tr>
<tr>
<td>The Road</td>
<td>58.0</td>
</tr>
<tr>
<td>The Firm</td>
<td>60.8</td>
</tr>
<tr>
<td>The Road</td>
<td>79.1</td>
</tr>
<tr>
<td>The Firm</td>
<td>77.8</td>
</tr>
<tr>
<td>The Road</td>
<td>76.9</td>
</tr>
</tbody>
</table>
Evaluation: Describe Shot via other Books

American Psycho
I have your license.
I know who you are. I know where you live. I'm keeping your license, and I'm going to check on you, mister Raymond K. Hessel. In three months, and then in six months, and then in a year, and if you aren't back in school on your way to being a veterinarian, you will be dead.
You didn't say anything.

[00:13:29:00:13:33] Lady, if you don't shut your fucking mouth, I will kill you.

Fight Club
You didn't say anything.
Get out of here, and do your little life, but remember I'm watching you, Raymond Hessel, and I'd rather kill you than see you working a shit job for just enough money to buy cheese and watch television.
Now, I'm going to walk away so don't turn around.

Harry Potter and the Sorcerer's Stone
I'm warning you now, boy.

[00:05:46:00:05:48] Bane Chronicles-2
She has graciously allowed me into her confidence. Magnus could read between the lines. Axel didn't kiss and tell, which made him only more attractive.
The escape is to be made on Sunday. Alex went on, "The plan is simple, but exacting. We have arranged it so the guards have seen certain people leaving by certain exits at certain times. On..."

[02:06:23:02:06:26] - I'm sorry I didn't tell you, Rachel. - No, No, Bruce...

Batman Begins
I'm gonna give you a sedative. You'll wake up back at home.

[01:38:41:01:38:44]
the club was a little emptier than i would have expected for the late afternoon, and the bartender, in red waistcoat and bowtie, was busy wiping down his counter, replacing peanuts and putting out new coasters. a television with the latest la liga news was hung in an upper corner, and behind him, rows of bottles were reflected in a giant bar mirror. above the stools, a pergola-type overhead structure held rows of wine glasses. it was a classy place, with ferns in the corner, and not the kind of bar to which i was accustomed. my places usually had a more ... relaxed feel.

he felt like an idiot for yelling at the child, but his frustration and trepidation was getting the better of him. he glanced toward the shadowed hall and quickly nodded toward melissa before making his way forward. he came across more children sitting upon a couch in the living room. they watched him, but did n’t move and did n’t speak. his skin started to feel like hundreds of tiny spiders were running up and down it and he hurried on.
Thank you.