Understanding Biological Research Documents using a Neural Network

Varun Mittal
Motivation: Assist researchers in literature search by annotating entities and establishing a relation type between them.
Challenges faced by researchers

1. Exponentially increasing number of articles.
   Reference: http://blogs.discovermagazine.com/neuroskeptic/

2. Literature search

3. Analyzing articles
   Image Credit: Google.com
Importance of annotations in Biology

Reference: http://string-db.org/cgi/network.pl
Current state of art - Statistical method
Challenges in the current approach

<table>
<thead>
<tr>
<th>Different word, same meaning (Synonyms)</th>
<th>Different expression, common meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein names</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Different notations</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Same word, different content</td>
</tr>
<tr>
<td>Different notations</td>
<td></td>
</tr>
<tr>
<td>5mg/kg of cyclosporine per day</td>
<td></td>
</tr>
<tr>
<td>Cyclosporine 5mg/kg per day</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Learning rate - research time</td>
</tr>
<tr>
<td></td>
<td>Learning rate - machine learning</td>
</tr>
</tbody>
</table>
Proposed state of Art: Suggest relations between entities

Note:
EV1 : Gene
CXC4: Protein
Methodology

A saxophone is used for jazz.

A/DT saxophone/NN is/VBZ used/VBN for/IN jazz/NN ./.

Extract Part of Speech Tags from example

Collect samples and train a neural network
Summary

PROBLEM

Stimulation of glioma cell motility by expression, proteolysis, and release of the L1 neural cell recognition molecule

Background

Glioma cells are particularly motile and can travel difficulty through the base of the brain, apparently without adhering to the normal anatomy. The neural cells are hypothesized to be involved in the development and progression of gliomas. The hypothesis is that these cells are not only a stimulator of tumor growth but also a source of potential therapeutic targets. The hypothesis is supported by recent studies showing that glioma cells are highly motile and can travel through the brain without adhering to the normal anatomy.

Current Challenges

- Low penetration of glioma cells into the brain
- High invasiveness of glioma cells
- Difficulty in targeting glioma cells for therapeutic purposes

Proposed Solution

- Development of new therapeutic strategies targeting glioma cells
- Identification of potential therapeutic targets in glioma cells
- Development of new technologies for targeted delivery of therapeutic agents to glioma cells

Diagram

- EV1
- CXC4
- HasDependency
- IsA
- Gene
- Protein

Legend

- cell
- gene
- protein
Acknowledgment

Dr. Matthew Tolentino (Committee Chair) - metolent@uw.edu

Dr. Raghavi Sakpal - rsakpal@uw.edu

Dr. Ka Yee Yeung - kayee@uw.edu

Dr. Ling-Hong Hung - (lhhung@uw.edu)

Dr. Kunal Mukerjee (kunalm@microsoft.com)