Mining for medical relations in research articles

Identification of relations

By Olof Nordengren and Vilhelm Lundqvist

Contents

- 1. Introduction & Background
- 2. Relations, NLP background
- 3. Method the algorithm
- 4. Results
- 5. Problems, improvements

Our role in the BioNLP Project

We process abstracts to extract relations using NLP rules

Anna and Eric find pieces of the puzzle - we connect the pieces

Text mining finds and combines knowledge fragments in medical literature



www.aitslab.org

Relations in biomedical texts

Example abstract marked by Sonja, where colors: disease, protein, cell-death term, interaction, drug

We are interested in the interactions, when one of the agents is a cell-death related term or protein

Example: <mark>Hsp70</mark> inhibits cell death

Mechanisms underlying cancer cell death caused by inhibitors of subcellular **Hsp70** proteins have been elucidated. An inhibitor of **Hsp70**, apoptozole (Az), is mainly translocated into **lysosomes** of cancer cells where it induces **lysosomal membrane permeabilization**, thereby promoting lysosome-mediated apoptosis. Additionally, Az impairs autophagy in cancer cells owing to its ability to **disrupt thelysosomal function**. However, the Az-triphenylphosphonium conjugate, Az-TPP-O3, localizes mainly to mitochondria of cancer cells where it inhibits the mortalin-**p5** interaction and induces mitochondrial outer membrane permeabilization, consequently leading to mitochondria-mediated apoptosis. Unlike Az, Az-TPP-O3 does not have an effect on autophagy in cancer cells. Collectively, the findings indicate that inhibitors of lysosomal **Hsp70** and mitochon ndrial mortalin enhance cancer cell death via distinctively different mechanisms. Additionally, the findings arising from this effort demonstrate that studies aimed at determining subcellular locations and functions of small-molecule modulators provide a deeper understanding of their modes of action in cells.

Project resources

- ~20 000 000 abstracts from PubMed
- Identified Named Entities from Anna & Eric
- List of interaction keywords from Sonja

Purpose of relation extraction

Apply NLP rules to the abstract data to build an annotated dataset

Hannes will train a model based on the dataset

Background - Dependency Graphs

Break down a sentence into dependency relations - extended grammar

Each word has exactly 1 head, the result is a graph that can be traversed

Example:



Background - Noun Chunks

Used noun chunks instead of single words to gain more relevant information

Includes modifying words and compounds along with the main noun (called the root)



Background - The "A affects B" relation

Focus on the most common relation structure: nominal subject - keyword - direct object

Both the nsubj-chunk and dobj-chunk point to the same interaction



Our algorithm - Overview



Our algorithm - Noun chunks

Our algorithm - Noun chunk roots

Mechanisms underlying cancer cell <u>death</u> caused by <u>inhibitors</u> of subcellular Hsp70 <u>proteins</u> have been elucidated. An <u>inhibitor</u> of <u>Hsp70</u>, <u>apoptozole</u> (Az), is mainly translocated into <u>lysosomes</u> of cancer <u>cells</u> where <u>it</u> induces lysosomal membrane <u>permeabilization</u>, thereby promoting lysosome-mediated <u>apoptosis</u>. Additionally, <u>Az</u> impairs <u>autophagy</u> in cancer <u>cells</u> owing to its <u>ability</u> to disrupt the lysosomal <u>function</u>.

Our algorithm - Root heads

Mechanisms underlying cancer cell <u>death</u> caused <mark>by inhibitors</mark> of subcellular Hsp70 <u>proteins</u> have been <mark>elucidated</mark>. An <u>inhibitor</u> of <u>Hsp70</u>, <u>apoptozole</u> (Az), is mainly <mark>translocated into <u>Iysosomes</u> of</mark> cancer <u>cells</u> where <u>it</u> induces Iysosomal membrane <u>permeabilization</u>, thereby promoting Iysosome-mediated <u>apoptosis</u>. Additionally, <u>Az impairs autophagy</u> in cancer <u>cells</u> owing <mark>to</mark> its <u>ability</u> to <u>disrupt</u> the Iysosomal <u>function</u>.

Our algorithm - Nsubj or Dobj dependency

Mechanisms underlying cancer cell death caused by inhibitors of subcellular Hsp70 proteins have been elucidated. An inhibitor of Hsp70, apoptozole (Az), is mainly translocated into lysosomes of cancer cells where it induces lysosomal membrane permeabilization, thereby promoting lysosome-mediated apoptosis. Additionally, **Az impairs autophagy** in cancer cells owing to its ability to disrupt the lysosomal function.

dobi

nsubi

Our algorithm - Prepositions



Our algorithm - Filter by relevant terms

Our algorithm - Finished, pass to Hannes

Preliminary Results

Recall TP / (TP + FN)	3/(3+41) = 6.8%
Precision TP / (TP + FP)	3 / (3 + 0) = 100%
F1-score 2*Rec.*Prec. / (Rec. + Prec.)	12.8%

[1] ADP - reptilase clot retracti corpus/out json pubmed19n0001.xml.txt.docria [2] AMP PMID=30556 [1] AP duration 1] ATP Hypothalamic regulation of prolactin secretion in animals (mammals) and man appears to be similar, and no significant differences have yet been demonstrated. The hypothalamus contains neurotransmitters and 1] ATPase polypeptides that can either inhibit or stimulate prolactin release, although the predominant influence under 11 ATPase activities of soluble basal conditions is to inhibit prolactin release. Thus pituitary stalk section or placement of lesions in the basal tuberal region of the hypothalamus results in increased prolactin release and sometimes in initiation 1] Acid hydrolysis of the urine of lactation. Among agents in the hypothalamus that can inhibit prolactin release, the most important 1] Acid production by carboh appear to be an as yet unidentified polypeptide prolactin release inhibiting factor (PIF) and dopamine. There 1] Acrosin from cock sperm is some evidence that dopamine may account for most, if not all, of the prolactin release inhibiting activity of 11 Actinomycin D the hypothalamus. Agents that increase dopamine activity, i.e. L-dopa, monoamine oxidase inhibitors, etc., 1] Added oxidized nicotinamic depress prolactin release. Acetylcholine also can inhibit prolactin release, but it appears to act via the [1] Addition of 3-amino-1,2,4-t catecholamines. Of the agents in the hypothalamus that stimulate prolactin release, the most important 11 Addition of lead in vitro appear to be an as yet uncharacterized polypeptide prolactin releasing factor (PRF), thyrotropin releasing 1] Addition of many oxidizable hormone (TRH) and serotonin. TRH is as effective in releasing prolactin as in releasing TSH, but under 1] Administration of E2 to OV most physiological states. TSH and prolactin release do not occur together. Serotonin and its precursors. 11 Administration of hydrocorl tryptophan and 5-hydroxytryptophan, are powerful releasors of prolactin and have been shown to be [1] Administration of nerve grc involved in some physiological states in which prolactin is released, i.e. during suckling, stress, etc. Other A 11- agents in the hypothalamus that can stimulate prolactin release include GABA and some prostaglandins. inhibit but these have not vet been shown to be involved in physiological control of prolactin secretion. Exteroceptive stimuli that alter prolactin release act through the CNS and hypothalamus, but some hormones and drugs also can act directly on the pituitary to promote or depress prolactin release. Abstract contains 0 other detected relations.

Identified problems

Statements of no relation ("... since LSD did not increase the DOPA accumulation...")

Coreferences ("A diphosphonate (EHDP) [...] was given to [...] volunteers for 28 days. <mark>It caused</mark> a significant increase in mean Pi and P50 in both healthy and diabetic subjects")

Complex relations, for example passive relations: "We found that active dopamine (DA) uptake was inhibited by S1694."

Include more interaction keywords

Thank you for watching our presentation!

Any questions?

By Olof Nordengren and Vilhelm Lundqvist