Applications of FrameNet and ASRL

Collin F. Baker, ICSI

FrameNet Tutorial

May 31, 2015

FrameNet Brasil

- Housed at Universidade Federal de Juiz de Fora, Minas Gerais, Brazil

- PIs: Maria Margarida Martins Salomão and Tiago Timponi Torrent

FrameNet Brasil

- Three projects:
  - Building Constructicon for Brazilian Portuguese
  - Copa 2014 (2014 Soccer World Cup)
  - Olympia (2016 Summer Olympics)

FrameNet Brasil: Constructicon

- A repertoire of syntactic constructions, structures whose recognition, treatment and interpretation are key for Natural Language Processing

- Starting with 2 types of constructions:

  - *Para* + infinitive

  - Multiword quantifiers as example of constructions

    - comp. ling. "multiword expression" = Cog. Ling. "construction"?

    - clear division between work of lexicon and work of construction

    - full integration of both in parsing, NLP

    - Comparative conx. in EN, BR PT, SE (collab. with Swedish FN)

FrameNet Brasil: Integrating metaphor

- model productive metaphors in Brazilian Portuguese, or import them from other databases (MetaNet)

- describe metaphorical conx., e.g., the Inceptive Aspect Construction:

Maria	rompeu	a chorar
Maria	break.PAST.3SG	to cry
Maria	broke	into tears

- create metonymy links

- infer the relations between schemas and the more specific frames via the network of relations

Copa do Mundo 2014

(Salomão et al. 2011, 2013)

Copa do Mundo 2014

(Salomão et al. 2011, 2013)

- Developed a frame-based trilingual (**Portuguese – Spanish – English**) electronic dictionary

Copa do Mundo 2014

(Salomão et al. 2011, 2013)

- Developed a frame-based trilingual (**Portuguese – Spanish – English**) electronic dictionary
- Covering the domains of **soccer, the World Cup, and general tourism**

Copa do Mundo 2014

(Salomão et al. 2011, 2013)

- Developed a frame-based trilingual (**Portuguese – Spanish – English**) electronic dictionary
- Covering the domains of **soccer, the World Cup, and general tourism**
- Search by word, frame name, or sentence containing a frame-evoking word in any language

Copa do Mundo 2014

(Salomão et al. 2011, 2013)

- Developed a frame-based trilingual (**Portuguese – Spanish – English**) electronic dictionary
- Covering the domains of **soccer, the World Cup, and general tourism**
- Search by word, frame name, or sentence containing a frame-evoking word in any language
- Returns frame description, LUs in all 3 languages, sometimes image or video

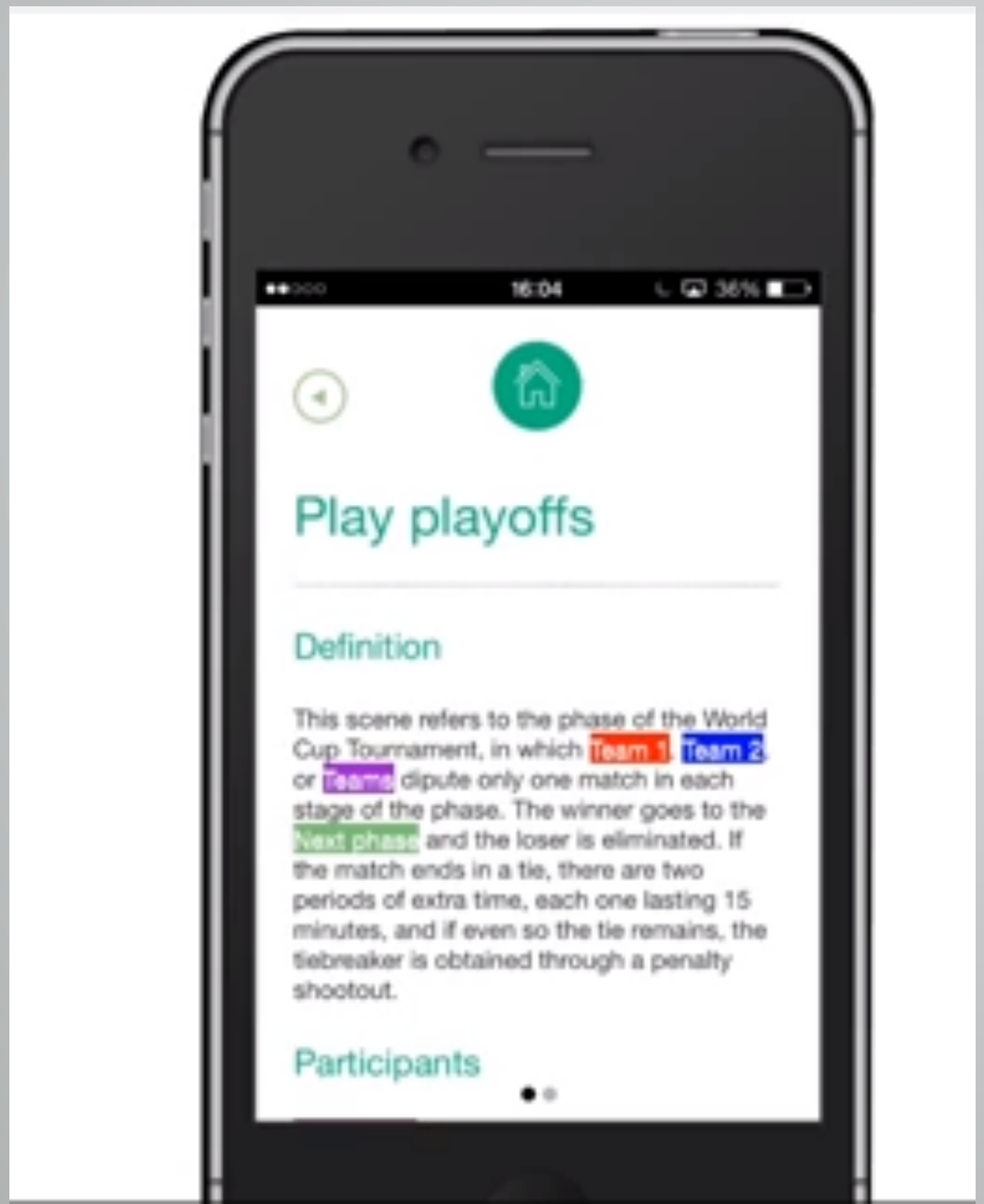
Copa do Mundo 2014

(Salomão et al. 2011, 2013)

- Developed a frame-based trilingual (**Portuguese – Spanish – English**) electronic dictionary
- Covering the domains of **soccer, the World Cup, and general tourism**
- Search by word, frame name, or sentence containing a frame-evoking word in any language
- Returns frame description, LUs in all 3 languages, sometimes image or video
- Website was used by tourists, tourism professionals, journalists and staff of 2014 World Cup in Brazil.

Copa do Mundo

https://youtu.be/Zi7vElBzU68

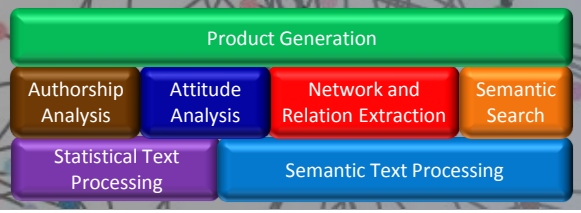


FRAMENET BRASIL
COPA DO MUNDO

Olympics version with frames for all sports coming soon!

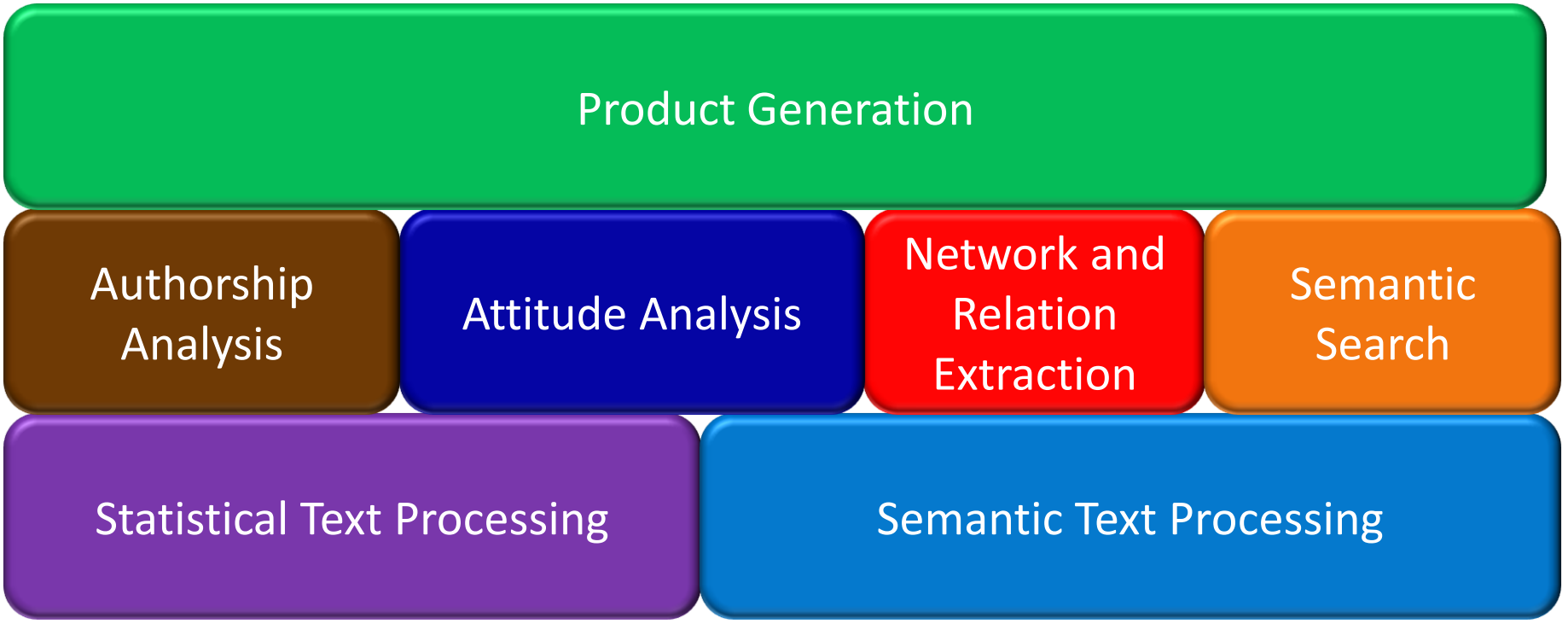
Decisive Analytics Corporation

- Long-term collaboration with FN through a series of subcontracts, e.g. our current work on
    - Spatial relations
        - Negation, tense, mood and aspect
- E.g. of their products:
    - Network extraction
    - Attitude analysis
    - Semantic search



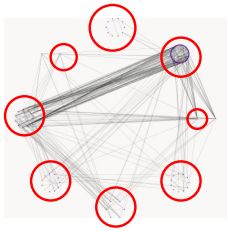
Technology Profile

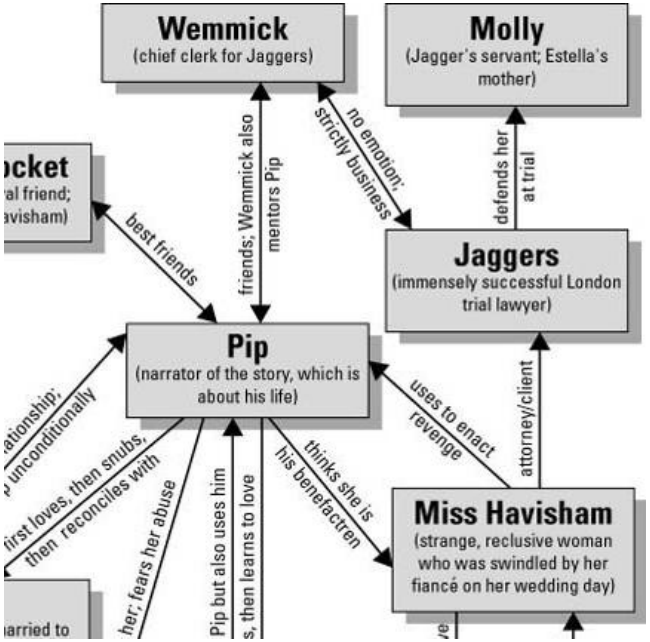
DECISIVE ANALYTICS
Analytic Products Division



Network Extraction and Refinement

- Network Extraction
  - *Disentangling the hairball of networks discovered in text*





Cliffs Notes (2013) "Great Expectations Character Map".

Motivation

DECISIVE ANALYTICS
Analytic Products Division

- Gathering everything we've been told about how entities (Person, Organization, Location) are connected in the world
- Avoiding a abstraction to one-type-of-node/one-type-of-edge
- Avoiding the "hairball" problem

Network and Relation Extraction

DECISIVE ANALYTICS
Analytic Products Division

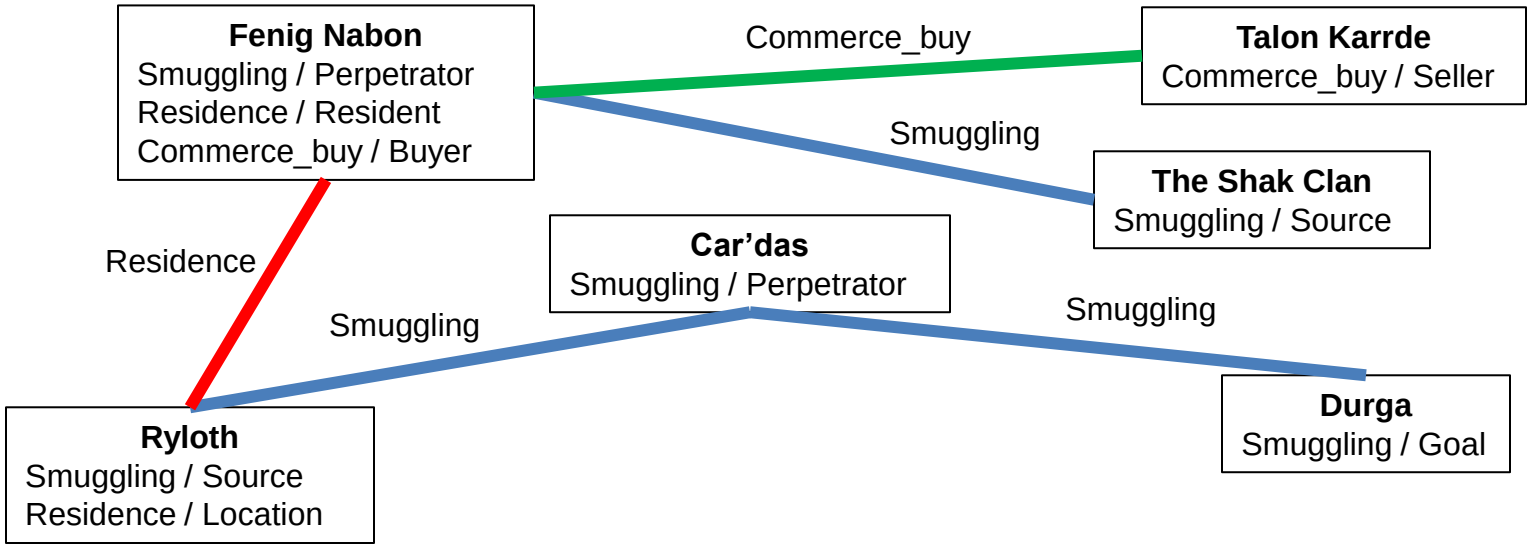
Source Text:

Fenig Nabon smuggled the undocumented containers from the Shak Clan.
A source reported that Fenig Nabon bought a large quantity of merchandise form Talon Karrde.
The Car'das smuggled a troupe of dancers form Ryloth to Durga.
Nabon resides in Ryloth.

Semantic Labeling:

Labeled sentence	Frame	Roles
Fenig Nabon smuggled the undocumented containers from the Shak Clan.	smuggled →Smuggling	Perpetrator Goods Source
A source reported that Fenig Nabon bought a large quantity of merchandise from Talon Karrde.	bought →Commerce_buy	Buyer Goods Seller
The Car'das smuggled a troupe of dancers from Ryloth to Durga.	smuggled →Smuggling	Perpetrator Goods Source Goal
Nabon resides in Ryloth.	resides →Residence	Resident Location

Network:

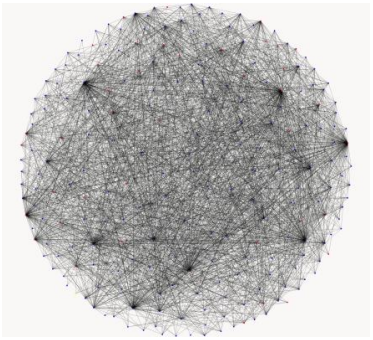


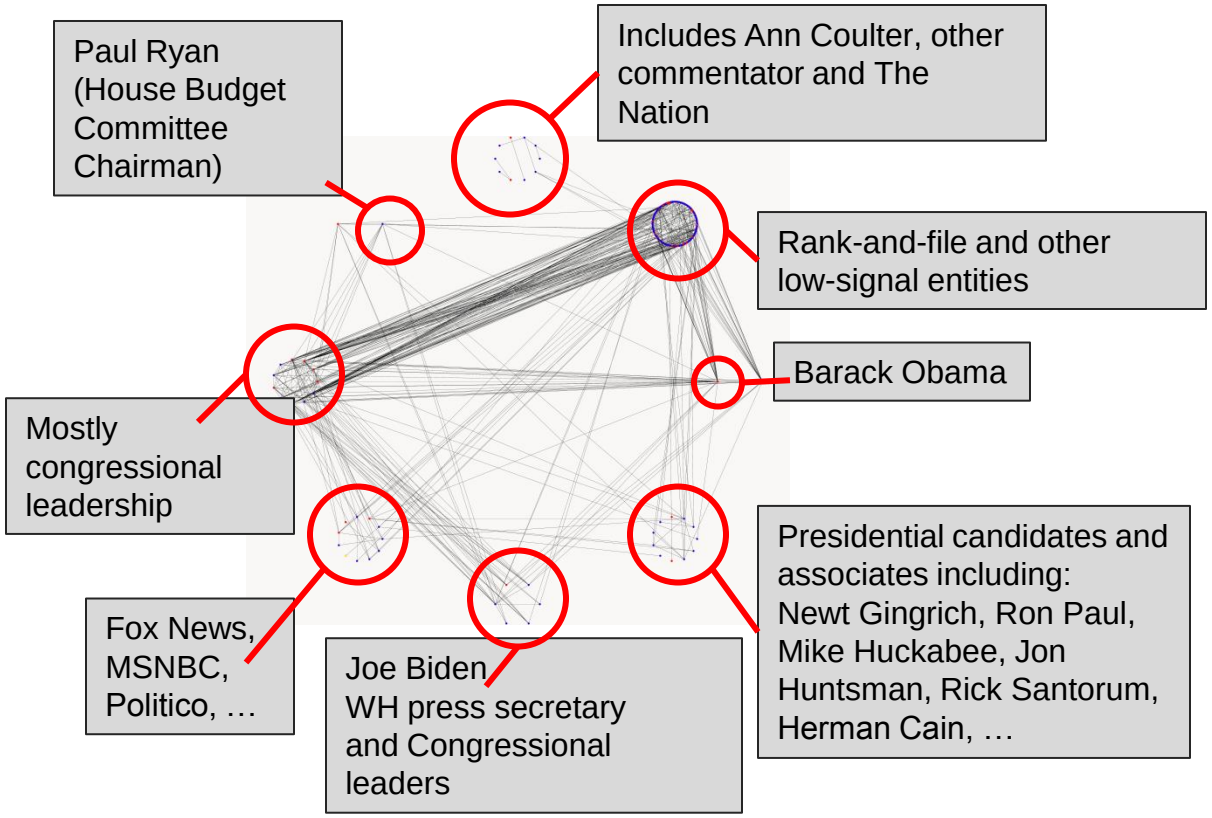
Organization Discovery

DECISIVE ANALYTICS
Analytic Products Division

Relational modeling algorithms can identify the latent organization within a large, complex graph (e.g. solve the hairball problem)

Network automatically generated from corpus of articles about debt-ceiling debate





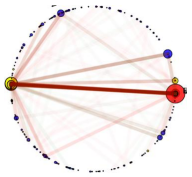
Network Extraction: Conclusion

DECISIVE ANALYTICS
Analytic Products Division

- An approach for going from frame labeled data to an entity network
- Filtering can focus analysis
- Relational modeling can reorganize a network into meaningful clusters based on frame data

Attitude Analysis

- Attitude Analysis
  - *Identifying* Factions *in Connected Entities*

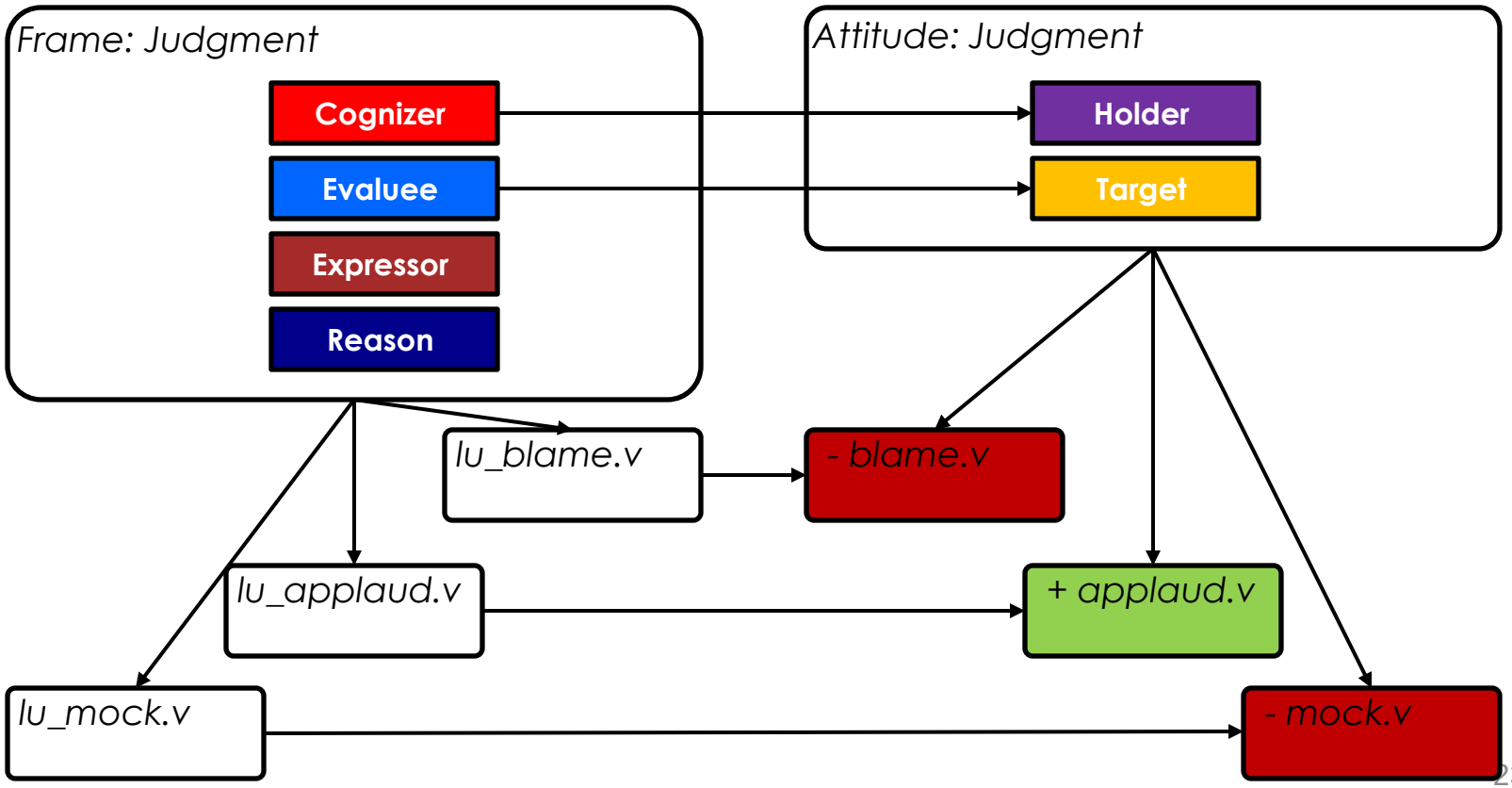


Stuart Barr (2008). “Fighting Topis”. *Stuart Barr flickr photostream*.

How this was actually done

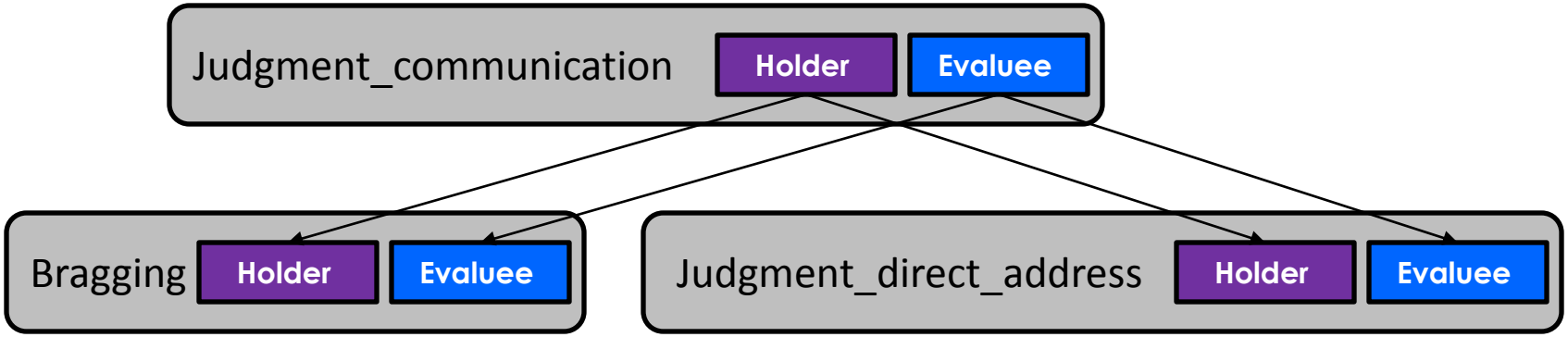
DECISIVE ANALYTICS
Analytic Products Division

- Start with a semi-manual process mapping FrameNet to Attitudes



The mapping process

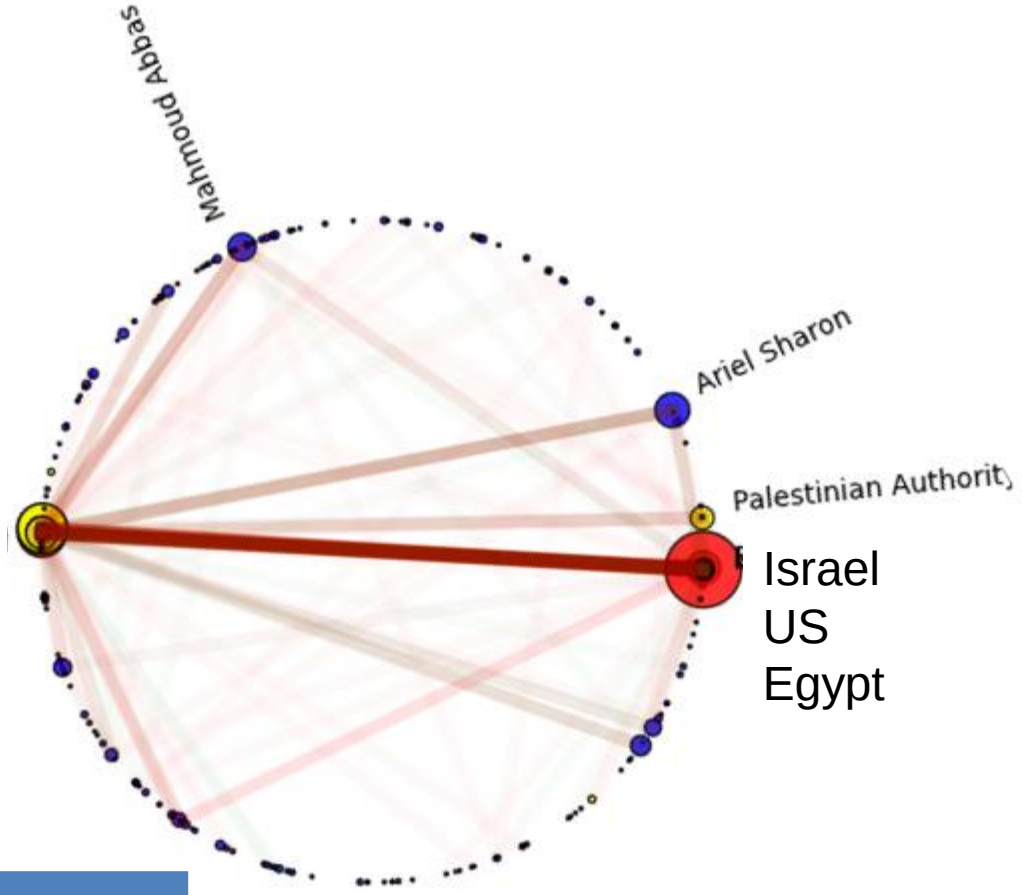
- A semi-manual mapping process
  - Start with FrameNet and Polarity Lexicon
  4. Inheritance links to pre-populate Holder/Target mapping for Frame Elements



How do different people feel about each other?

DECISIVE ANALYTICS
Analytic Products Division

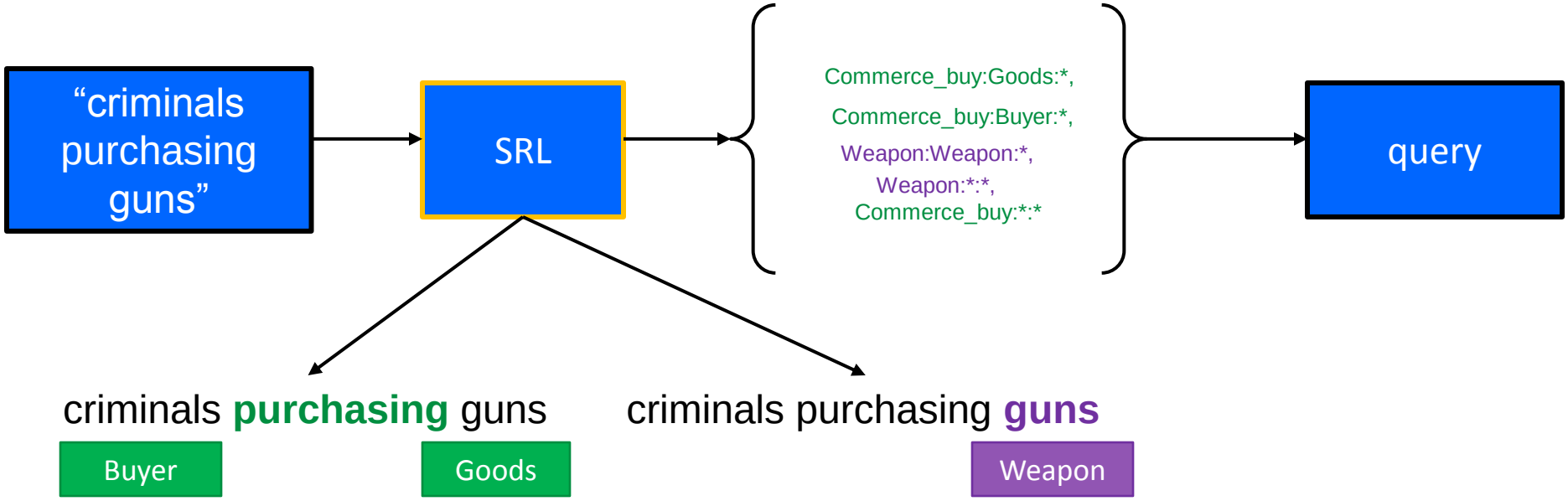
Palestine Fateh-Hamas
Palestinian
Palestinians
Palestinian Authority Interior
Gaza
West Bank
Israeli



Measure	Fraction	Accuracy
Membership	10/12	83.3%
Implied Relations	25/35	71.4%

Automatic Query Generation

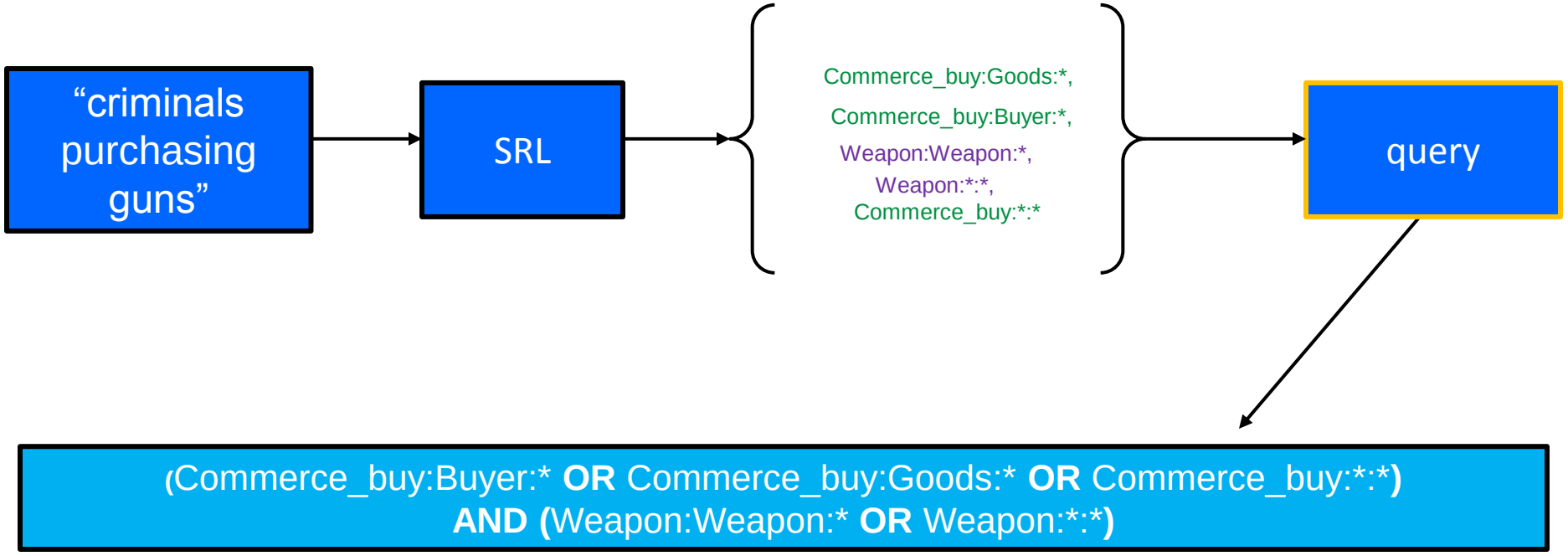
Take plain text input and generate a FrameNet-based query.



Automatic Query Generation

Decisive Analytics
Analytic Products Division

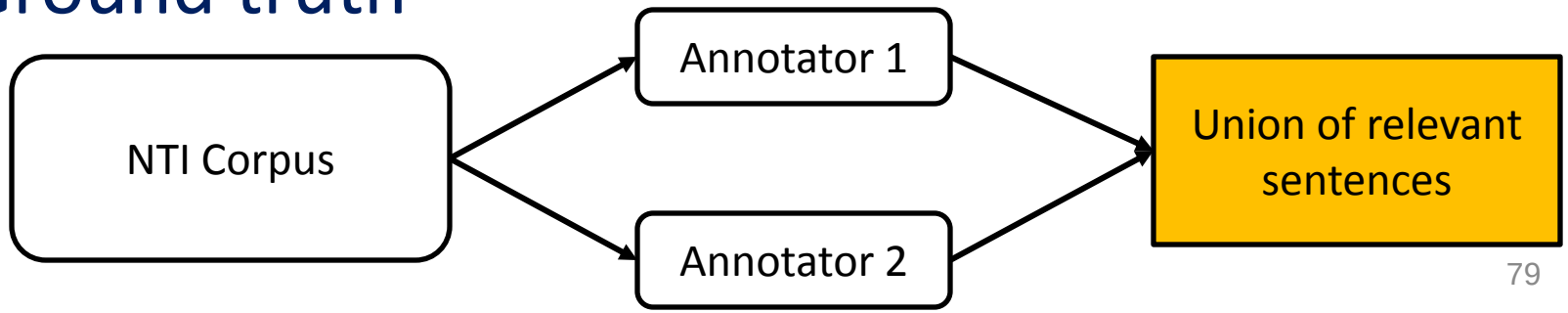
Take plain text input and generate a FrameNet-based query.



Moving around the FrameNet links

DECISIVE ANALYTICS
Analytic Products Division

- Hypothetical scenario
  - Draw links between countries that have exchanged resources (technology, money, supplies, people, etc.) with respect to chemical, nuclear, and biological warfare
- Dataset
  - Texts from the Nuclear Threat Initiative website
- Ground truth



Moving around the FrameNet links

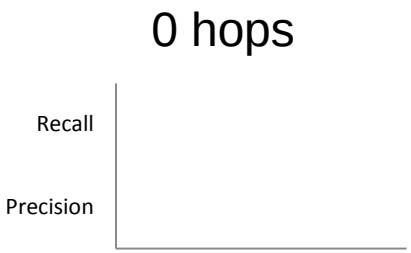
- Ground truth

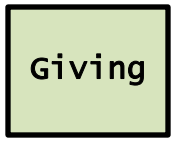
China is a key supplier of CW equipment and technologies to Iran according to US Department of Defense.

North Korea also allegedly provided more than 10 scientists to work on the Libyan missile program .

Other defensive material purchased by Iran includes respirators from Spain, protective gear from South Korea, and atropine autoinjectors from the Netherlands.

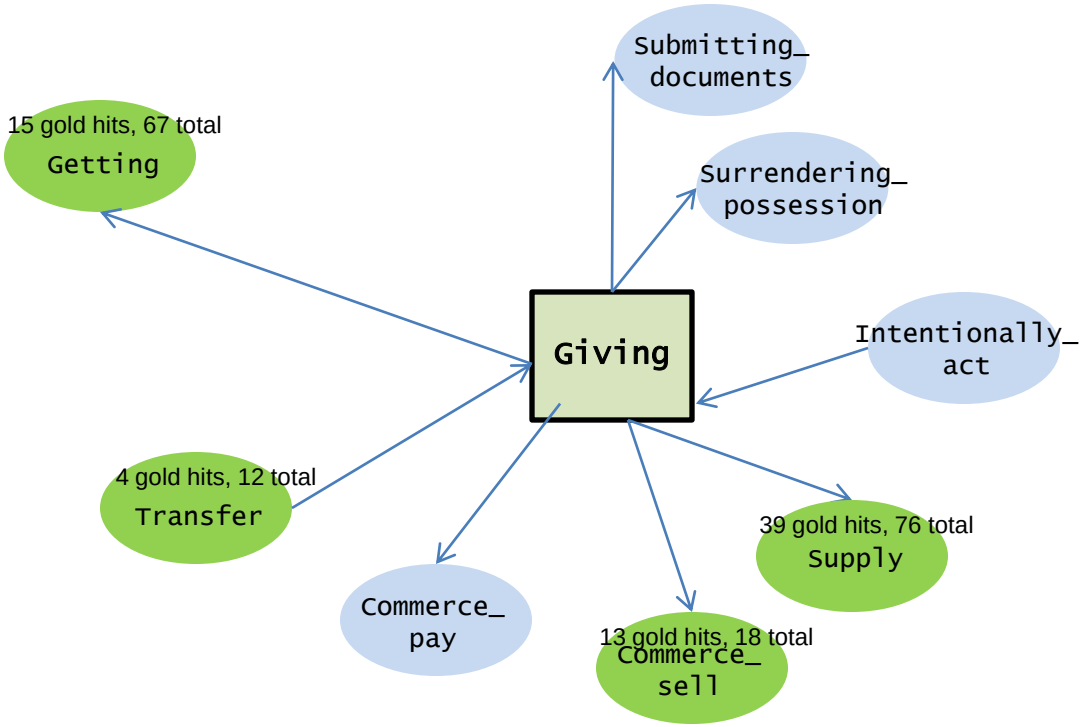
Moving around the FrameNet links

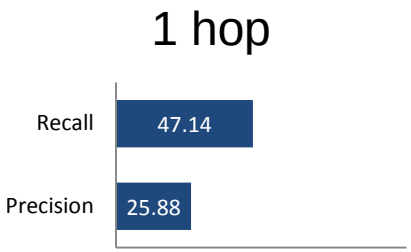




The `Giving` frame is central to the idea of resource exchange; however, a frame search for `Giving` with no hops does not find any relevant results, so the recall and precision metrics are both 0.

Moving around the FrameNet links

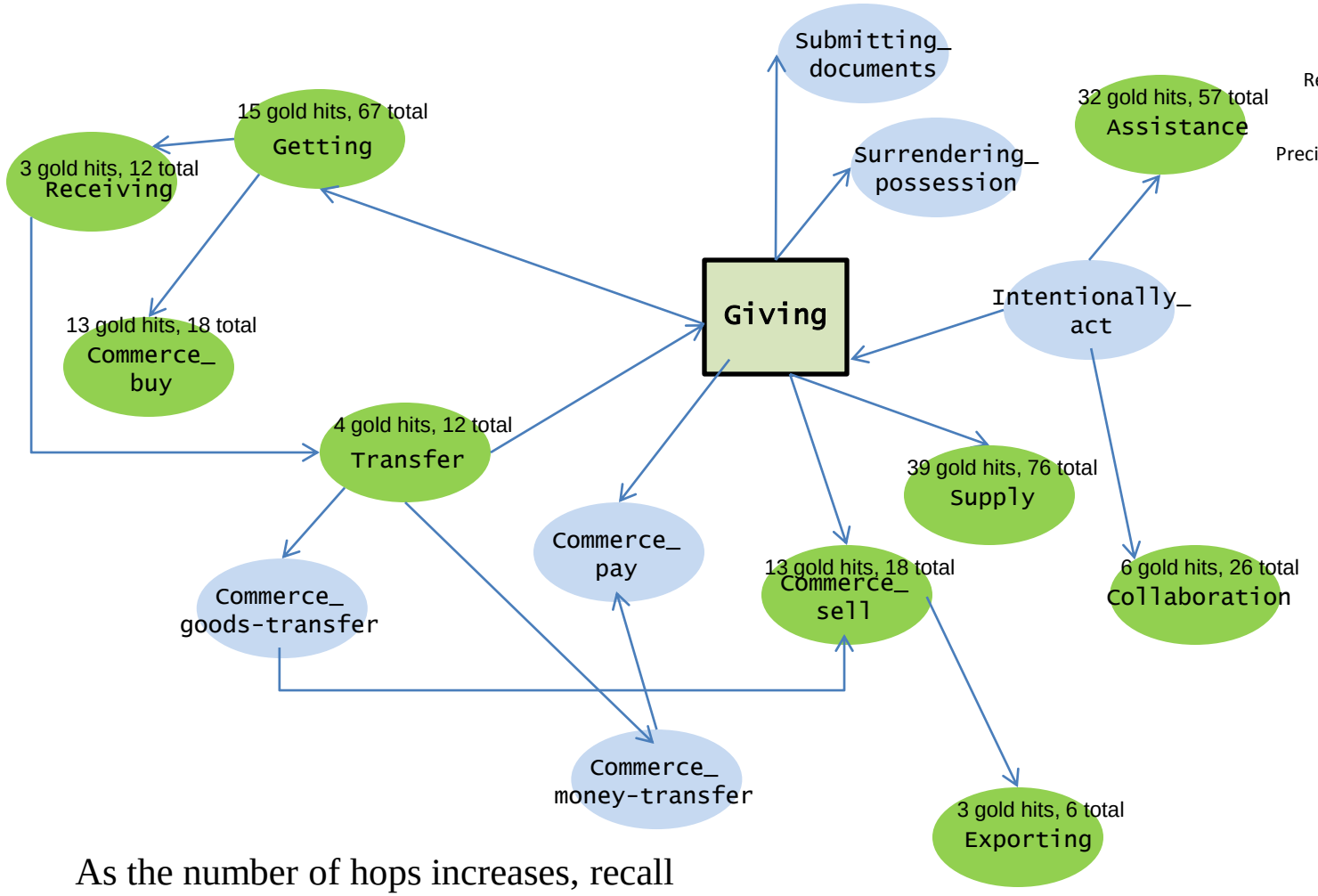


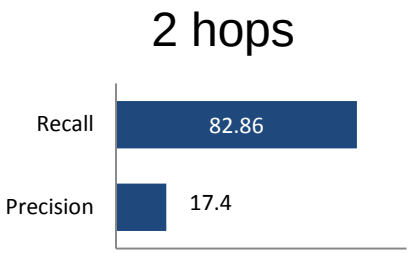


Searching over all frames 1 hop from the original `Giving` frame returns almost half of all the relevant results.

Moving around the FrameNet links

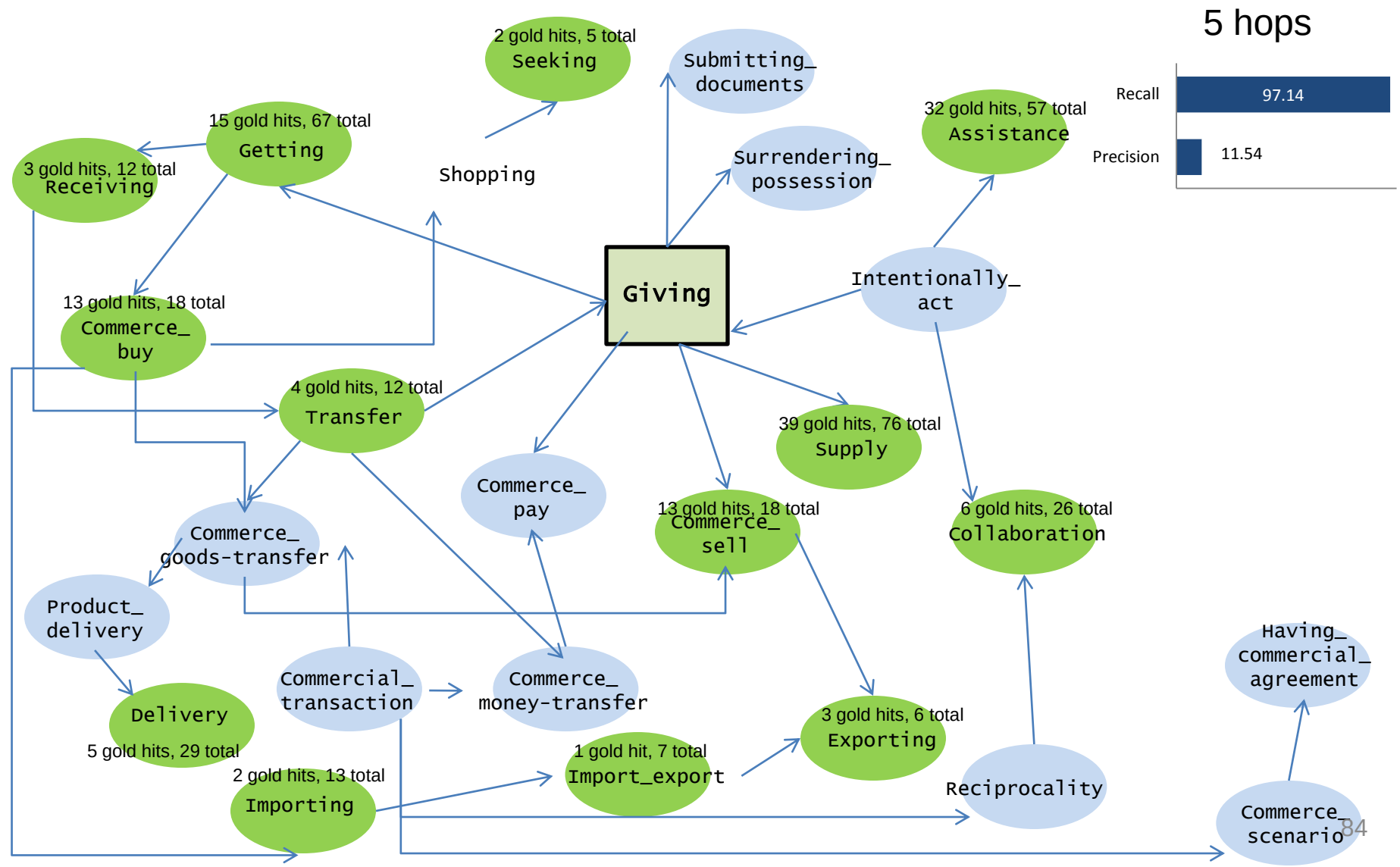
DECISIVE ANALYTICS
Analytic Products Division





As the number of hops increases, recall increases and precision decreases.

Moving around the FrameNet links



Semantic Search: Conclusion

- Ability to execute queries over frames, frame elements, and “terms”
- Terms can be abstracted too
- Queries may be generated from simple text
- Results come in several forms
- Many more features on the docket

Thank you!