



Financial Events Recognition in Web News for Algorithmic Trading

Frederik Hogenboom

fhogenboom@ese.eur.nl

Erasmus University Rotterdam
PO Box 1738, NL-3000 DR
Rotterdam, the Netherlands

October 18, 2012



ERASMUS UNIVERSITEIT ROTTERDAM

9th International Workshop on Web Information Systems Modeling (WISM 2012)

Introduction (1)

- Trading in financial markets: from human brokers to computer programs
- **Algorithmic trading**: *the use of computer programs for entering trade orders with algorithms deciding aspects like timing, price, and quantity of an order*
- Trading algorithms are more efficient:
 - Lower latency
 - Larger volume
 - Higher market coverage degree



Introduction (2)

- Usually, trading algorithms implemented in business tools make use of numerical inputs:
 - Stock prices
 - Trading volumes
 - Averages
 - ...
- News is not a common input
- However ... financial markets are extremely sensitive to breaking news!

News & Financial Markets

Steve Jobs resigns from Apple, Cook becomes CEO

Apple stock price falls on news of Steve Jobs resignation

(The Guardian) - Apple's CEO Steve Jobs returned in

Google buys Motorola Mobility for \$12.5B

(VentureBeat) - This morning, Google announced that it will buy Motorola Mobility — Moto's mobile device arm — for \$12.5 billion. **Google will acquire Motorola Mobility** for \$40 per share in cash, a 63 percent premium over the company's Friday closing price. Google says it will run Motorola Mobility as a separate business. Motorola spun off its business into two divisions last year, Mobility and Solutions (the data and telecom portion), as a response to declining profits.

Google shares were down around 1.5 percent, while **Motorola Mobility's stock jumped 57 percent**. The company says Motorola Android phones won't be receiving any special treatment as a consequence of the deal — but that's a tough nut to swallow, since Google often plays favorites.

have soared from





News & Algorithmic Trading (1)

- Large news companies (e.g., Reuters, Dow Jones) started to provide product services offering **tagged** news items
- Tagging facilitates machine interpretation if used in a Semantic Web context
- However, tags are coarse-grained and provide general information about:
 - Company
 - Topic
 - Industry
 - ...



News & Algorithmic Trading (2)

- Most current annotations:
 - Are based on titles and abstracts, instead of the full content
 - Are not linked to ontologies
- We would rather like to incorporate a fine-grained, ontology-based annotation, which allows for the identification of financial events:
 - Acquisition
 - Stock splits
 - Dividend announcements
 - ...

Events

- Event:
 - Complex combination of relations linked to a set of empirical observations from texts
 - Can be defined as:
 - <subject> <predicate> e.g., <Person> <Resigns>
 - <subject> <predicate> <object> e.g., <Company> <Buys> <Company>
- Common event domains:
 - Finance
 - Politics
 - Environment
 - Medical

Proposal

- Financial Events Recognition in News for Algorithmic Trading (FERNAT) framework:
 - Financial ontology
 - Lexico-semantic event extraction rules
 - Ontology updates
 - Incorporate events into trading algorithms
- Challenges:
 - Identify relevant news information
 - Use information in the most effective way
 - Perform tasks **timely** and **accurately**



FERNAT Framework

- Pipeline for financial events extraction: news messages are parsed to interpretable tokens
- Patterns that identify (extract) financial events are matched on tokens
- Events are used for ontology updating
- Approved events are used in decision making, i.e., trading in financial markets

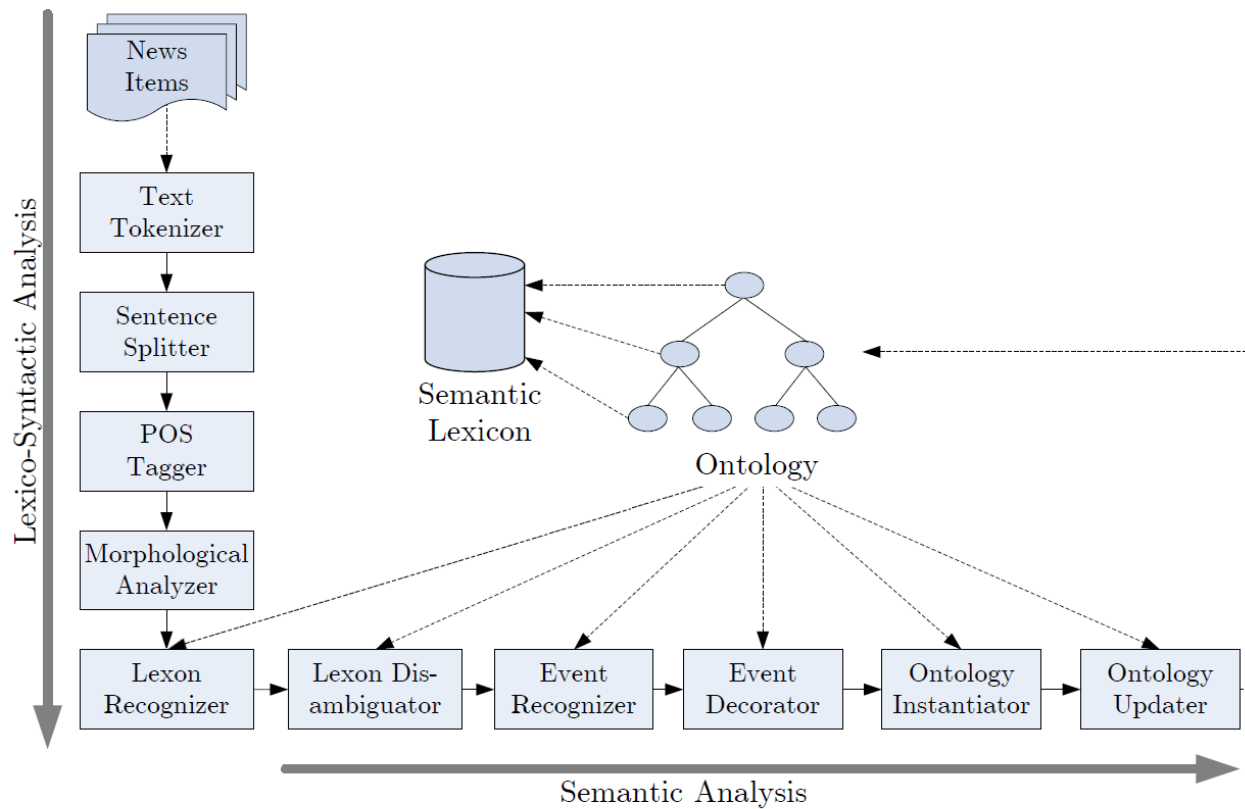
Processing Pipeline (1)

- Types of analysis:
 - Lexico-syntactic
 - Semantic
- Cornerstone is a domain ontology for financial events and their related facts:
 - Expert view of the financial world at a certain moment in time
 - Concepts are anchored to synsets from a semantic lexicon





Processing Pipeline (2)





Event Pattern Language (1)

- Various pattern-languages for:
 - News processing frameworks (e.g., PlanetOnto)
 - General purpose frameworks (e.g., CAFETIERE, KIM, etc.)
- Language types:
 - Lexico-syntactic
 - Lexico-semantic
- However:
 - Insufficient use of semantics (if any)
 - Limited syntax
 - Cumbersome in use

Event Pattern Language (2)

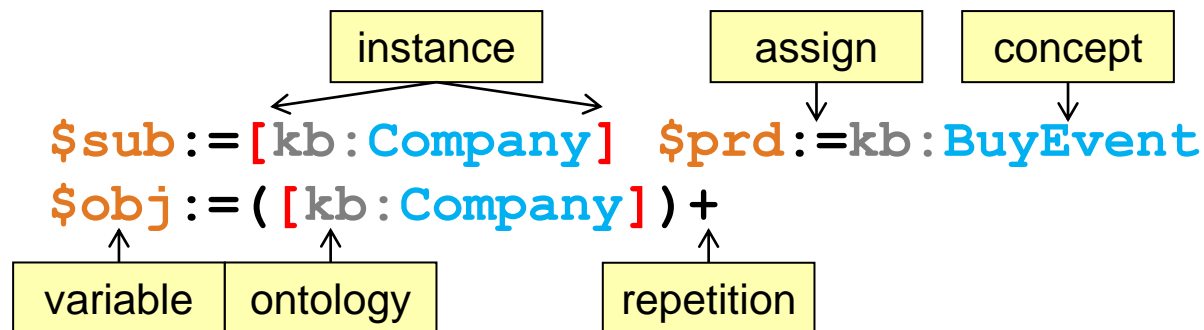
- Semantic Web:
 - Collection of technologies that express content meta-data
 - Offers means to help machines understand human-created data on the Web
- Ontologies:
 - Can be used to store domain-specific knowledge in the form of concepts (classes + instances)
 - Also contain inter-concept relations





Event Pattern Language (3)

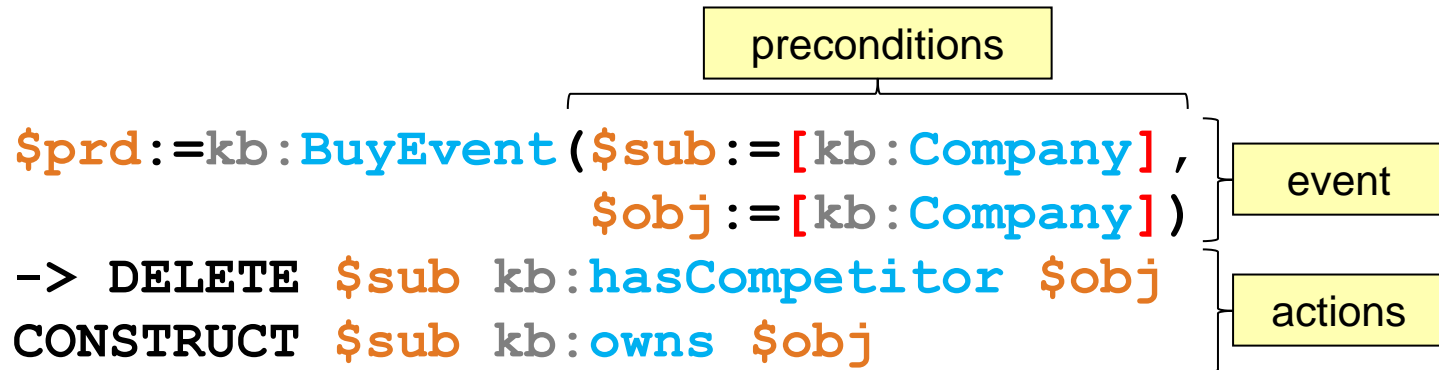
- Example lexico-semantic pattern that mines texts for company acquisitions:



- **Scope:**
 - Sentence
 - Tokens allowed in between $\$sub$, $\$prd$, and $\$obj$

Ontology Updates (1)

- Example update statement:



- Based on SPARQL/Update



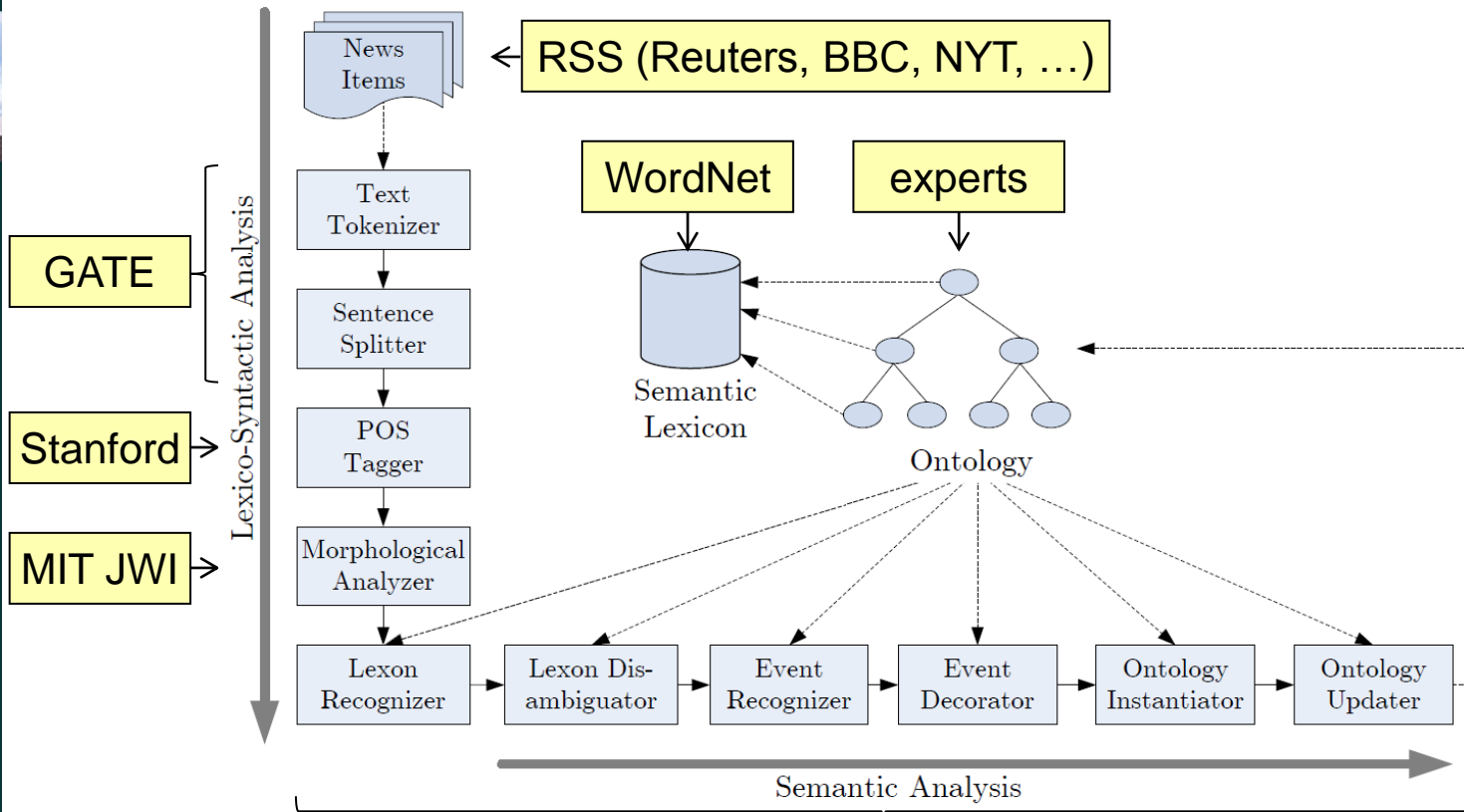
Ontology Updates (2)

- Ontology Update Language (OUL):
 - Automatic updating mechanism from active databases: event triggers
 - Uses event-condition-action model, which executes SPARQL/Update statements
- We have extended OUL to OULx:
 - Prefixes and negation operator
 - Execution models:
 - Immediate / deferred updating
 - Chaining of updates through internal triggering (rippling)
 - Looping
 - Execution of multiple matching handlers

Implementation (1)

- Processing pipeline architecture based on GATE
- Most lexico-syntactic components can be implemented by using existing software
- Individual semantic components have been implemented as proof-of-concept
- Currently, no final implementation of the full framework yet
- Implementation will be finalized Q4 2012 – Q1 2013

Implementation (2)



Legend:

- Name Processing Unit
- Information Flow
- UsedBy Relationship

ongoing research

Implementation (3)



Hermes News Portal v2.0

Home Original graph Search graph Results Recommendations Import news Rule Editor Users Evaluation

Rules Editor Annotation Validation Manual Annotations Evaluation

Rule Groups

- CEO Discovery
- Product Discovery
- Shares Discovery
- Competitor Discovery
- Profit Discovery
- Loss Discovery
- Partner Discovery
- Subsidiary Discovery
- President Discovery
- Sales Discovery

Rules

- Company CEO Person
- Person Company CEO
- Person CEO IN Company
- Person Wildcard CEO Wildcard Company
- Company Wildcard CEO Wildcard Person
- Company Wildcard Person Become CEO

+ Add X Remove Up Down

Editor

Rule Name: Company CEO Person

Rule Description:

```
($sub, kb:hasCEO, $obj) :- $sub=[kb:Company] ('\s' | ',')? kb:CEO ', '? $obj=[kb:Person]
```

Output

Create Concept Rule
Create Relation Rule
Insert Concept
Insert Orth
Insert POS
Insert Literal
Combine
Insert Range
Validate Syntax
Save Rule

+ Add Edit Name X Remove
Up Down

Implementation (4)



The screenshot shows the Hermes News Portal v2.0 interface. The top navigation bar includes tabs for Home, Original graph, Search graph, Results, Recommendations, Import news, Rule Editor, Users, and Evaluation. Below this, there are buttons for Rules Editor, Annotation Validation, Manual Annotations, and Evaluation. A control bar contains 'Run!' (checked), 'Reclassify' (unchecked), 'Keep History for Evaluation' (unchecked), 'Save Annotations' (checked), and 'Clear All Annotations' (checked).

The main content area is divided into two sections: 'Patterns Found' and 'News Items'.

Patterns Found

Subject	Relation	Object	Times Found	Valid
kb:Google	kb:hasCEO	kb:Eric_Schmidt	2	<input type="checkbox"/>
kb:Disney	kb:hasCEO	kb:Robert_A_Iger	1	<input type="checkbox"/>
kb:Gameloft	kb:hasCEO	kb:Michel_Guillemot	1	<input type="checkbox"/>
kb:BP	kb:hasCEO	kb:Tony_Hayward	1	<input type="checkbox"/>
kb:Alcatel-Lucent	kb:hasCEO	kb:Ben_Verwaayen	1	<input type="checkbox"/>
kb:Viacom	kb:hasCEO	kb:Philippe_P_Dauman	1	<input type="checkbox"/>
kb:NTR	kb:hasCEO	kb:Jim_Barry	1	<input type="checkbox"/>
kb:Coke	kb:hasCEO	kb:Muhtar_Kent	1	<input type="checkbox"/>
kb:Picarro	kb:hasCEO	kb:Woelk	1	<input type="checkbox"/>
kb:Microsoft	kb:hasCEO	kb:Steven_A_Ballmer	1	<input type="checkbox"/>
kb:National_Arts_Council	kb:hasCEO	kb:Benson_Puah	1	<input type="checkbox"/>
kb:British_Airways	kb:hasCEO	kb:Willie_Walsh	1	<input type="checkbox"/>
kb:Royal_Dutch_Shell	kb:hasCEO	kb:Peter_Voser	1	<input type="checkbox"/>
kb:Ford	kb:hasCEO	kb:Jacques_Nasser	1	<input type="checkbox"/>
kb:Merck	kb:hasCEO	kb:Richard_T_Clark	1	<input type="checkbox"/>
kb:BP	kb:hasCEO	kb:Anthony_B_Hayward	1	<input type="checkbox"/>

News Items

Past Clouds Future of Europe's New Antitrust Chief
Wed Mar 03 02:26:22 CET 2010

The news item text is partially visible and contains the following content:

...to the sales on the job of the antitrust chief for the European Union and Monday the independence of the... clouding his new role, in which he must keep companies and countries from breaching rules on market-rigging and government subsidies. Mr. Almunia's predecessor as antitrust commissioner, Neelie Kroes of the Netherlands, had a reputation for bargaining long and hard with powerful people like **Steven A. Ballmer**, chief executive of **Microsoft**. At the height of the financial crisis, she insisted that rules on competition be respected as governments bailed out banks, auto companies and other businesses. Giants like the Dutch bank ING and the Royal Bank of Scotland were required to make significant divestitures. Now, the question is whether Mr. Almunia will be up to his new job's challenges, including the protectionist tendencies rising in conjunction with unemployment in many European countries and the potential for trans-Atlantic tensions, especially as companies like Google continue to expand globally. His spokeswoman, Amelia Torres, said Mr. Almunia would not grant interviews so early in his nearly five-year term as competition commissioner. Last Monday, in his first public appearance in the role, Mr. Almunia

Implementation (5)



The screenshot shows the 'Hermes News Portal v2.1.1' interface. The 'Rule Editor' tab is active, displaying the configuration for a rule named 'AddCompetitor'. The interface includes a menu bar with options like 'Home', 'Original graph', 'Search graph', 'Import news', 'Rule Editor', 'Results', 'Users', and 'Evaluation'. Below the menu, there are buttons for 'Extraction Rules Editor', 'Update Rules Editor', 'Extraction Validation', 'Update Validation', 'News', and 'Updates'. A 'Save' button and a checked 'Compile handlers' checkbox are also visible.

The main configuration area for the 'AddCompetitor' rule is divided into several sections:

- Name:** 'AddCompetitor'
- Description:** (empty field)
- Prefixes:** 'kb: <http://www.hermes.com/knowledgebase.owl#>'
- Request:** A dropdown menu set to 'Add', a dropdown menu set to 'Not uni...', and a text field containing '?companyX kb:hasCompetitor ?companyY'.
- Condition:** A tree view showing a 'Conditions' folder containing a 'Not' folder, which in turn contains a 'Contains(?companyX kb:hasCompetitor ?companyY)' condition.
- Actions:** A tree view showing an 'Actions' folder containing three actions: 'Apply request', 'Insert(?companyY kb:hasCompetitor ?companyX)', and 'Feedback(?companyX became a competitor of ?companyY)'.

Each section has a set of control buttons: '+ Add / Edit', 'X Remove', 'Up', and 'Down'.

At the bottom, a 'Console' window displays the message: 'Handlers compiled succesfully!' in green text.

Using Events for Trading

- We aim to improve returns of technical trading rules:
 - Assign impact factors to events
 - Determine trading signals (buy, hold, sell)
 - Combine signal with common signals
- Rules are constructed using genetic programming:
 - No ad-hoc specifications of rules
 - Evolutionary algorithm
 - Optimization on returns (various horizons)
- Ideally, optimal rules should include news signals



Using Events for Value-at-Risk

- Value-at-Risk (**VaR**) is a threshold value, such that the probability of future returns exceeding this threshold is at a given confidence level
- Infrequent events cause a distortion in the calculation of VaR using historical data, because there is a deviation from the general trend
- Solution: cleaning data used for VaR calculation by smoothing out the deviations associated with certain (infrequent) events





Considerations

- Considering news involves an updated knowledge base, likely resulting in more informed and accurate trading decisions and VaR calculations
- Using ontologies enables reasoning and should hence make trading rules and VaR predictions more advanced
- Speed issues:
 - Improved accuracy at the cost of speed, as our employed technologies are rather heavy
 - Lags between publication of news and reaction in the stock market could be substantial enough to cover for the increase in processing time
 - Separation between computationally intensive event recognition and algorithmic trading



Conclusions

- We have proposed the FERNAT framework:
 - News processing pipeline
 - Outputs are applied to algorithmic trading
- Implementation:
 - Pipeline can be constructed using existing software
 - Specialized (semantic) components under development
 - Event pattern language and ontology update language
- Future work:
 - Wrap up implementation
 - Application for algorithmic trading

