RUTGERS

Rutgers Business School
Newark and New Brunswick

Automating the process of taxonomy creation and comparison of taxonomy structures

Vasundhara Chakraborty, Miklos Vasarhelyi

Nineteenth Annual Strategic and Emerging Technologies Research Workshop

San Francisco, CA, USA, July 31, 2010.

Outline

- Research Questions
- Motivation
- Contribution
- Related work
- Methodology
- Results
- Conclusion

Research Questions

- RQ1: What method should be used to create a taxonomy automatically using historical data from financial statements?
- RQ2: What are the structural differences between the official XBRL pension footnote taxonomy and a pension footnote taxonomy created by the proposed method?
- RQ3: Is tagging of pension footnote data more effective using tags produced from the alternate method as compared to the tags from the official XBRL taxonomy?

Motivation

Examples of variations
9.PENSION, POSTRETIREMENT AND OTHER EMPLOYEE BENEFIT PLANS
Note 21—Employee Benefit Plans
Pension and Postretirement Plans
Note 6—Retirement Plans
Note 11. Employee Benefit Plans •*Defined Benefit Pension Plans and Postretirement Plans* (as a separate section under Employee Benefit plans)
w. Retirement-Related Benefits
EMPLOYEE BENEFIT PLANS
Defined Contribution Pension Plans
Defined Benefit Pension Plans
Albertson's defined benefit Plan
Shaw's defined benefit plan
14. Employee Benefit Plans •*Defined Benefit Pension Plan* •*Defined Contribution Plans* •*Postretirement Benefits*

Motivation

- Bovee, Kogan, Nelson, Srivastava, Vasarhelyi (2005) proposes that historical data should be used for taxonomy creation.

- Bovee, Ettredge, Srivastava, Vasarhelyi (2002) raises questions about how well a taxonomy represents a firms' preferred reporting practices.

- Manually using historical data to create taxonomies can be laborious. Automating it, even to a certain extent, can reduce the complexities.

- As regulations change, firms must report differently. Therefore, taxonomies would also need to be updated often. Automation can reduce complexities.

Contribution

- Develop a method to partially automate the taxonomy creation process.
- Create a generic tool for applying the methodology.
- Show structural differences between the XBRL taxonomy and taxonomy using historical data.
- Demonstrate a use of the tool for other exploratory research in accounting.

Related work

Information retrieval literature

- Salton(1989)
- Chen (1992)
- Chen (1994)
- Chen (1995)
- Crouch (1988), Crouch & Yang (1999)
- Chuang (2005)

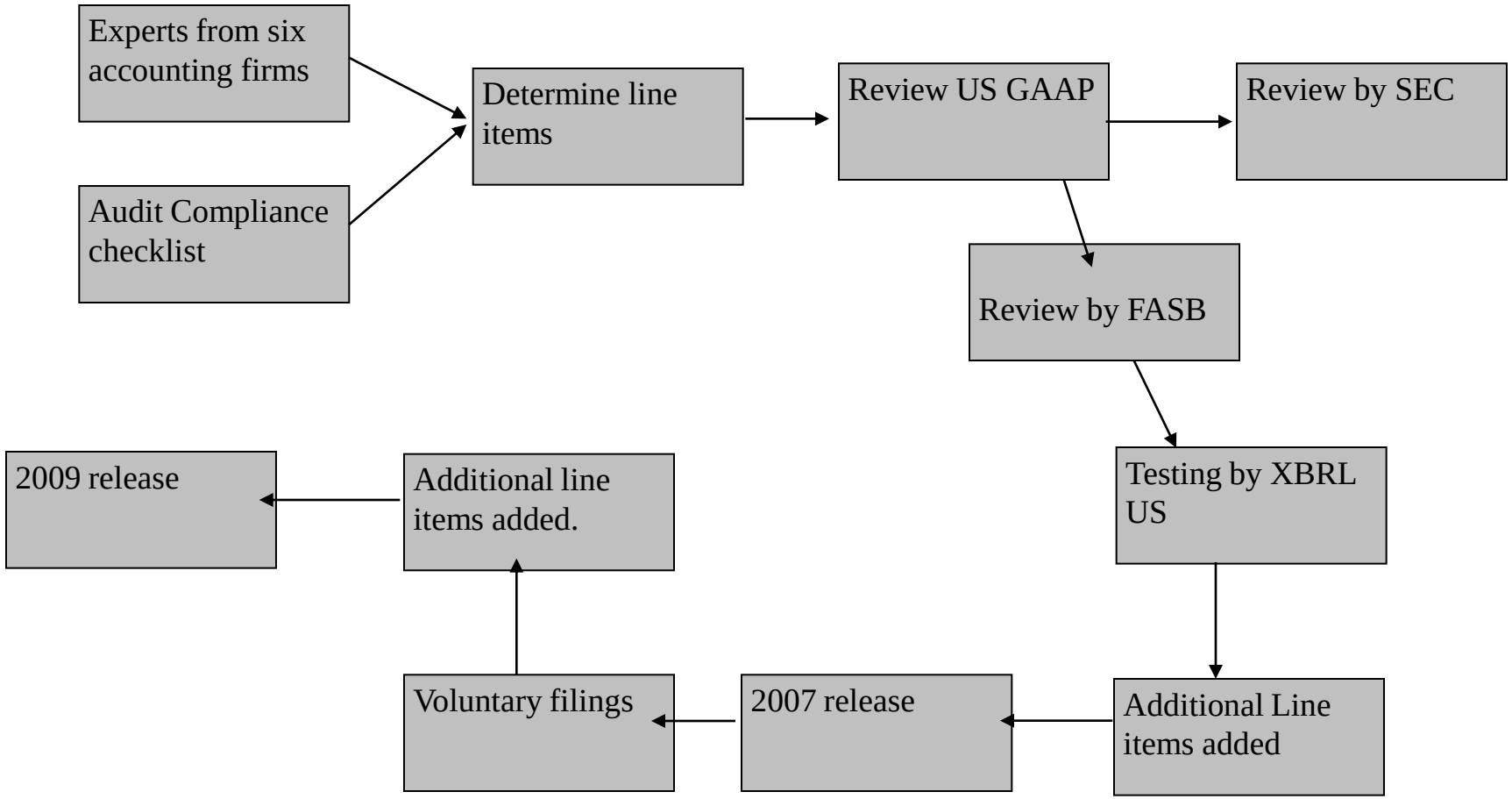
Accounting

- Wu and Gangolly(2000)
- Fisher(2004)
- Garnsey(2006)

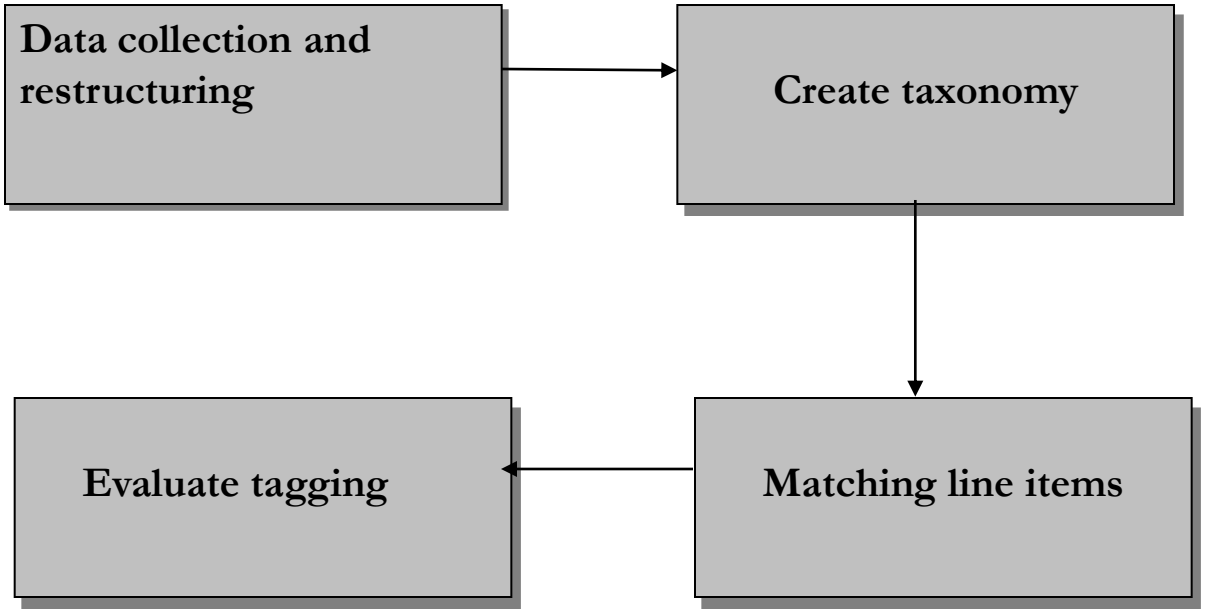
Data:

- 10K statements of 120 companies (randomly picked from a list of Fortune 500)
- 80 were used as the training dataset.
- 40 were used as the test dataset.

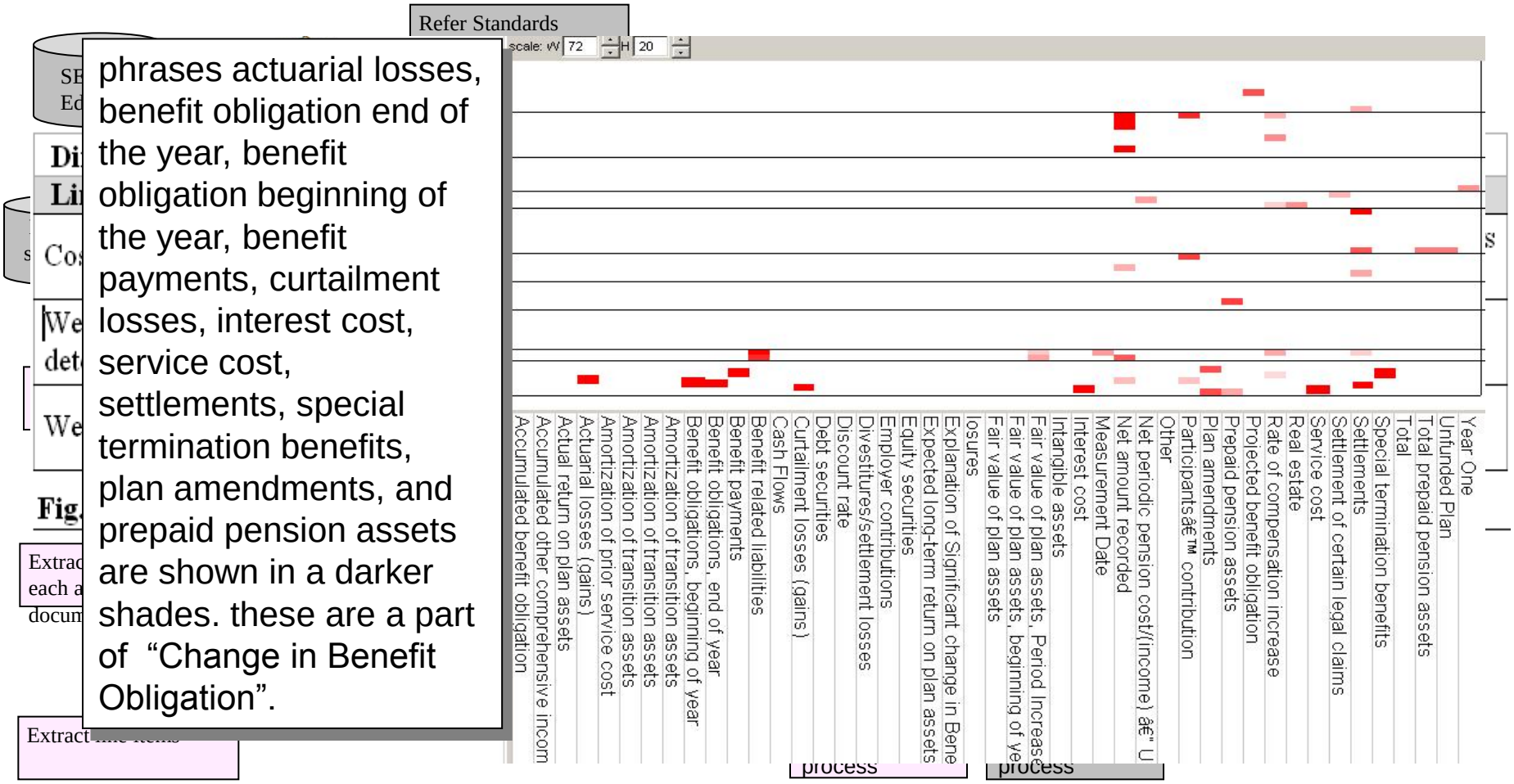
Taxonomy generation process followed by XBRL US



Overview of the proposed method



Automating the process of taxonomy creation and comparison of taxonomy structures



phrases actuarial losses, benefit obligation end of the year, benefit obligation beginning of the year, benefit payments, curtailment losses, interest cost, service cost, settlements, special termination benefits, plan amendments, and prepaid pension assets are shown in a darker shades. these are a part of "Change in Benefit Obligation".

Details of the taxonomy creation process

Performance evaluation of the parsing module

	Pension header			Section identifier			Line item parsing			Overall		
	No of occurrences	Number correctly identified	Success Rate (%)	No of occurrences	Number correctly identified	Success Rate (%)	No of occurrences	Number correctly identified	Success Rate (%)	No of occurrences	Number correctly identified	Success Rate (%)
Training dataset	78	78	100	1890	1834	97	21000	20580	98	22968	22492	97.92
Test dataset	40	40	100	1060	996	94	15400	14784	96	16500	15820	95.8

Comparison of taxonomies

Only in the official XBRL taxonomy

Only from historical data

5. Information for pension plans with an accumulated benefit obligation
- Projected benefit obligation
- Accumulated benefit obligation
- Fair value of plan assets

6. Weighted-average asset allocation of the pension and postretirement plans
- Equity securities
- Debt securities
- Real estate
- Long duration bonds
- U.S Stocks
- International stocks
- Emerging markets stocks and bonds
- Alternative investments
- Other
- Total

Required items

Found under Information on Plan assets in XBRL taxonomy

FAS 132r(a) requires more granular reporting

7. Information for pension plans with an accumulated benefit obligation in excess of plan assets (BOPA)
- Projected benefit obligation
- Accumulated benefit obligation
- Accumulated postretirement benefit obligation(APBO)
- Fair value of plan assets
- ABO less fair value of plan assets

Not used in post FAS 158

Comparison of taxonomies

Only in the official XBRL taxonomy

Only from historical data

14. Accumulated Benefit Obligation

15. Accumulated other comprehensive income, before tax

- *Net Gains (losses), before tax*
- *net prior service cost(credit) before tax*
- *Net transition assets(Obligations), before tax*
- *Minimum pension liability, before tax*
- *Total*

16. Amounts Amortized from Accumulated Other Comprehensive Income (Loss) in next Fiscal year

- Amortization of net gains(losses)
- Amortization of net Prior service cost(credit)
- Amortization of net Transition Asset(Obligation)
- Total

17. Pension plans with a benefit obligation in excess of plan assets

- Aggregate Benefit Obligation
- Aggregate Fair value of Plan assets

20. Estimated Future employer contributions in Next Fiscal Year

21. Alternative Methods to Amortize Prior Service Amounts

22. Alternative Methods to Amortize net gains and losses

23. Method to Determine Vested Benefit Obligation

24. Description of any Substantive Commitment Used as Basis for Accounting for Benefit Obligation

As part of ChBO or CHPlanAssets

Old way before ABO or PBO was used

Comparison of taxonomies

Only in the official XBRL taxonomy

Only from historical data

25. Special Termination Benefits during Period

- *Description of Event Resulting in Special or Contractual Termination benefits recognized during period*
- *Cost of providing Special termination benefits*

25a.Plan Amendment

- *Description*
- *Effect on Accumulated Benefit Obligation*
- *Effect on Net Periodic Benefit Cost*

As part of Change in Benefit obligation

Special events based on materiality

25b. Explanation of Significant change in Benefit Obligations or Plan assets not apparent from other disclosures

26. Settlement and Curtailments

- *Description*
- *Effect on Accumulated Benefit Obligation*

As part of Change in Benefit obligation

27. Measurement Date

28. Pension plans with a Accumulated benefit obligation in excess of plan assets

- Aggregate Projected Benefit Obligation
- Aggregate Accumulated benefit obligation
- Aggregate Fair value of Plan assets

An application of the tool:

Comparison of Pension footnote reporting structure of Fortune 1000 companies between 2000 and 2010

The Company's pension plan weighted-average asset allocations at December 31, by asset category, are as follows:

Long duration bonds
U.S. stocks
International stocks
Emerging markets stocks and bonds
Alternative (private) investments
Total

A new section added probably due to FAS 132r(a) but Equity securities, Debt securities, Real estate, Other?

An application of the methodology and tool:

Comparison of Pension footnote reporting structure of Fortune 1000 companies between 2000 and 2010

Change in Benefit Obligation
Benefit obligation at January 1
Service cost
Interest cost
Plan participant contributions
Medicare PartD subsidy in 2010
Plan amendments
Actuarial (gain) loss
Acquisitions included in 2010
Divestitures included in 2010
Benefits paid
Curtailment
Recognition of termination benefits
Foreign currency exchange rate change
Benefit obligation at December 31
Accumulated benefit obligation portion of above at December 31

Fig. 20 Group of items added in a year

Fig. 21 New Terms added in a year

Change in Fair Value of Plan Assets
Fair value of plan assets at January 1
Acquisitions included in 2010
Divestitures included in 2010
Actual return on plan assets
Company contributions
Plan participant contributions
Medicare PartD subsidy included in 2010
Benefits paid
Foreign currency exchange rate change
Fair value of plan assets at December 31:

Similar terms added by different companies but in different sections

An application of the methodology and tool:

Comparison of Pension footnote reporting structure of Fortune 1000 companies between 2000 and 2010

Fig. 22 New Terms added in 2010 as per SFAS 158 requirements

Funded Status AS in 2010
Amounts Recognized in the Consolidated Balance Sheet at December 31
Noncurrent assets
Current liabilities
Noncurrent liabilities
Total recognized

New terms added in 2010 as per SFAS 158 requirements

An application of the methodology and tool:

Comparison of Pension footnote reporting structure of Fortune 1000 companies between 2000 and 2010

Fig. 24 New Terms added in in 2010

Severence Accrual

As a result of the 2008 business environment's impact on our operating and capital plans, a reduction in our overall employee work force occurred in 2009.

Beginning balance

Accruals

Benefit payments

Accrual reversals

Ending Balance

New terms added in 2010 due to economic downturn

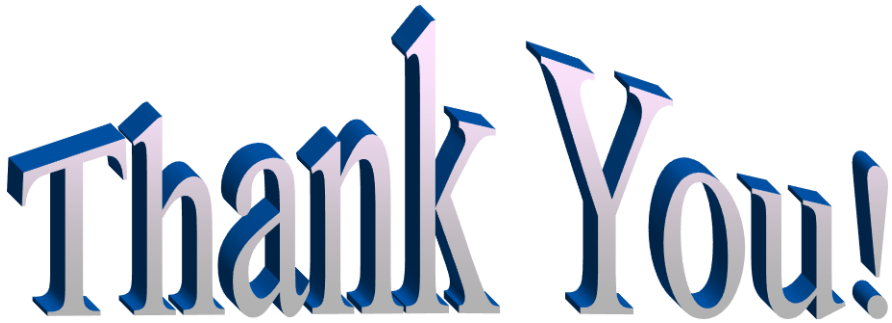
Conclusion

Findings

- We can use historical data to formalize and semi-automate the process of taxonomy creation.
- Comparison of taxonomies show that companies tend to aggregate whereas a more disaggregated structure is followed in the XBRL taxonomy.
- Some new terms or change in position of terms have been found in the historical data taxonomy compared to the XBRL taxonomy.

Limitations and future research

- Results obtained may represent the trends observed in the companies that were randomly chosen.
- Future research can be carried out to explore some of these pension footnote reporting trends of companies based on size or industries.


Thank You!