

# Using Layout Information for Spreadsheet-Visualization

Sabine Hipfl  
Institut für Informatik-Systeme  
Universität Klagenfurt  
shipfl@edu.uni-klu.ac.at

## Contents

---

- Introduction
  - Main Idea
- Layout Information
  - 4 Heuristics
- Summary

## Introduction (1)

---

- extending the approach of **semantic classes** [MC, 2002]
- semantic classes are an abstraction of a spreadsheet
- based on the similarity of formulas
  - **logical equivalence classes** [CHM, 2002]
- content of cells and spatial situations are taken into account and
- it is checked, whether the semantic content is of repetitive nature
  
- user-parametrization is necessary
  - whether semantically related cells are spread out
    - column-wise,
    - row-wise or
    - block-wise
  - cut-off parameter
  - equivalence classes

## Introduction (2)

---

- parametrization can be a hurdle as
  - deep knowledge about the approach is necessary
  - it might not be suitable for the untrained user
  
  - the orientation can be indicated only for the whole spreadsheet
  - semantically related blocks having different size will not be identified in their full extent
- **using layout information for identifying semantically related cells**

# Main Idea

## 3 kinds of layout information

- labels
  - formatting of the content of cells
  - lines and frames
- labels have the highest impact

### Core:

1. cells containing numbers or formulas are assigned to labels  
→ **layout areas**
2. the cells of the layout areas are checked for semantic classes

# 5 Kinds of Labels

1. **running numbers**
2. labels interpreted as ordinal numbers
3. counters
4. labels with complete identity
5. labels with partial identity

	(mbd)	(mmb)	(mmb)
1	57,53	21,00	21,00
2	115,07	42,00	63,00
3	172,60	63,00	126,00
4	172,60	63,00	189,00
5	172,60	63,00	252,00
6	172,60	63,00	315,00
7	172,60	63,00	378,00
8	167,05	60,97	438,97

	A	B	C	D	E
Z		2000	2001	2002	2003
3	Households with TV	3.200.000	3.258.018	3.316.044	3.374.070
4					
5	Market Shares				
6	Cable	78.0%	72.0%	68.0%	62.0%
7	Satellite Dish	12.0%	14.0%	16.0%	18.0%
8	Interactive TV	2.0%	7.0%	10.0%	15.0%
9	Broadcast TV	8.0%	7.0%	6.0%	5.0%
10	TOTAL	100.0%	100.0%	100.0%	100.0%

## 5 Kinds of Labels

1. running numbers
2. **labels interpreted as ordinal numbers**
3. counters
4. labels with complete identity
5. labels with partial identity

	A	B	C	D	E	F
10		July	August	September	October	November
11	Net Sales	\$156,333	\$161,556	\$165,148	\$167,006	\$168,341
12	COGS	95,363	98,549	100,740	101,874	102,688
13	Gross Profit	60,970	63,007	64,408	65,132	65,653
14	G&A Expenses	11,569	11,955	12,221	12,358	12,457
15	Depreciation	20,000	20,000	20,000	20,000	20,000
16	EBIT	29,401	31,052	32,187	32,774	33,196

EuSprIG 2004

7/25

## 5 Kinds of Labels

1. running numbers
2. labels interpreted as ordinal numbers
3. **counters**
4. labels with complete identity
5. labels with partial identity

	A	B	C
1	Task Number	Lead/ Lag	Duration Optimistic
2	Start (S)		
3	Task 1		10
4	Task 2	-5	15
5	Task 3		21
6	Task 4	10	15
7	Task 5		13
8	Task 6		30
9	Task 7	-5	20
10	Task 8	5	10
11	Task 9	20	11
12	Task 10		23
13	Task 11	5	22

Year 1	Year 2	Year 3	Year 4	Year 5
18.367	20.203	22.224	24.446	26.891
15.096	16.606	18.267	20.093	22.102
3.270	3.597	3.957	4.353	4.788

EuSprIG 2004

8/25

## 5 Kinds of Labels

---

1. running numbers
2. labels interpreted as ordinal numbers
3. counters
4. **labels with complete identity**
5. labels with partial identity

	D	E	F	G	H	I	J	K
7	Budget	Actual	Diff.	% Sales	Budget	Actual	Diff.	% Sales
8	5.857	6.044	187		11.715	10.661	-1.054	
9	5.325	6.309	984	100,0%	10.651	10.685	34	100,0%
10	1.787	2.669	882	42,3%	3.574	4.494	921	42,1%

EuSpRIG 2004

9/25

## 5 Kinds of Labels

---

1. running numbers
2. labels interpreted as ordinal numbers
3. counters
4. labels with complete identity
5. **labels with partial identity**

	A	B	C
1	<i>Quarter</i>	<i>Starting Sales</i>	<i>Growth</i>
2	First Qtr 2000	\$12.200.000	3,00%
3	Second Qtr 2000	\$14.000.000	2,50%
4	Third Qtr 2000	\$9.500.000	3,50%

EuSpRIG 2004

10/25

## 2 Kinds of Assignment

---

### 1. **geometrical assignment**

### 2. semantical assignment

of a cell to a label, if:

- the label is above or on the left side of the cell,
- there is no other label between the cell and the label and
- the cell itself is not a label.

	D	E	F	G
7	Budget	Actual	Diff.	% Sales
8	5.857	6.044	187	
9	5.325	6.309	984	100,0%
10	1.787	2.669	882	42,3%

## 2 Kinds of Assignment

---

### 1. geometrical assignment

### 2. **semantical assignment**

of a cell to a label, if:

- the label and the content of the cell are having the same formatting information,
- there is no other cell between the cell and the label containing another kind of formatting information and
- the cell must be geometrically assignable to the label.

	A	B	C	D	E
6		Budget			
7		2001	%Sales	Budget	Actual
8	Overtun	70.289		5.857	6.044
9	Sales	63.903	100,0%	5.325	6.309

## 4 Heuristics

---

1. Assignment based on Running Numbers and Counters
2. Assignment based on Complete and Partial Identity
3. Assignment based on Semantics
4. Spatial Assignment

### Main steps

- Identify labels,
- geometrically assign cells to the identified labels and
- group them into equivalence classes.

## Heuristic 1

---

### **Assignment based on running numbers and counters**

- identification of
  - running numbers
  - labels interpretable as ordinal numbers
  - counters
- geometrical assignment of cells
  - layout areas
- subsequently
  - grouping into semantic classes
  - heuristic 2 is applied

## Example – Heuristic 1

	A	B	C	D	E
		4,50	YEAR 1	YEAR 2	YEAR 3
1 Inflation Rate (%)		4,50			
2 Sale Value - Unit.		80,00 €	83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold		2.500.000	2.625.000	2.756.250	2.894.063
5 Sales Revenue		200.000.000,00 €	219.450.000,00 €	240.791.512,50 €	264.208.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	-	180.000.000,00 €	199.056.416,67 €	207.037.910,42 €	215.379.616,39 €
8 Variable costs	-	105.000.000,00 €	104.238.750,00 €	108.929.493,75 €	113.831.320,97 €
9 Fixed costs	-	70.000.000,00 €	73.150.000,00 €	76.441.750,00 €	79.881.628,75 €
10 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
11 Gross Profit		20.000.000,00 €	20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	-	18.000.000,00 €	31.160.000,00 €	33.179.000,00 €	35.349.000,00 €
13 Several Expenses	-	18.000.000,00 €	18.810.000,00 €	19.655.750,00 €	20.541.042,00 €
14 Interests of Loans*	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
15 Net Operating Profit		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	-	300.000,00 €	1.614.812,50 €	86.190,31 €	2.021.980,61 €
18 Net Profit		1.700.000,00 €	9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation		5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	-	- €	- €	- €	48.642.274,00 €
22 Residual Value	-	- €	- €	- €	- €
23 Cash flow		6.700.000,00 €	24.866.062,50 €	35.678.328,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value		51.197.159,31 €			
26 Present value - Cash Flows of the Old Plant	-	38.211.000,00 €			
27 Cash flow increment		12.986.159,31 €			

EuSprIG 2004

15/25

## Example – Heuristic 1

	A	B	C	D	E
		4,50	YEAR 1	YEAR 2	YEAR 3
1 Inflation Rate (%)		4,50			
2 Sale Value - Unit.		80,00 €	83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold		2.500.000	2.625.000	2.756.250	2.894.063
5 Sales Revenue		200.000.000,00 €	219.450.000,00 €	240.791.512,50 €	264.208.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	-	180.000.000,00 €	199.056.416,67 €	207.037.910,42 €	215.379.616,39 €
8 Variable costs	-	105.000.000,00 €	104.238.750,00 €	108.929.493,75 €	113.831.320,97 €
9 Fixed costs	-	70.000.000,00 €	73.150.000,00 €	76.441.750,00 €	79.881.628,75 €
10 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
11 Gross Profit		20.000.000,00 €	20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	-	18.000.000,00 €	31.160.000,00 €	33.179.000,00 €	35.349.000,00 €
13 Several Expenses	-	18.000.000,00 €	18.810.000,00 €	19.655.750,00 €	20.541.042,00 €
14 Interests of Loans*	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
15 Net Operating Profit		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	-	300.000,00 €	1.614.812,50 €	86.190,31 €	2.021.980,61 €
18 Net Profit		1.700.000,00 €	9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation		5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	-	- €	- €	- €	48.642.274,00 €
22 Residual Value	-	- €	- €	- €	- €
23 Cash flow		6.700.000,00 €	24.866.062,50 €	35.678.328,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value		51.197.159,31 €			
26 Present value - Cash Flows of the Old Plant	-	38.211.000,00 €			
27 Cash flow increment		12.986.159,31 €			

EuSprIG 2004

16/25



## Heuristic 2

### Assignment based on complete and partial identity

- identification of
  - labels with complete identity
  - labels with partial identity
- geometrical assignment of cells
  - layout areas
- subsequently
  - grouping into semantic classes
  - heuristic 3 is applied

EuSprIG 2004

17/25

## Example – Heuristic 2

	A	B	C	D	E
1 Inflation Rate (%)		4,50	YEAR 1	YEAR 2	YEAR 3
2 Sale Value - Unit.		80,00 €	83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold		2.500.000	2.625.000	2.756.250	2.894.063
5 Sales Revenue		200.000.000,00 €	219.450.000,00 €	240.791.512,50 €	264.208.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	-	180.000.000,00 €	199.065.416,67 €	207.037.910,42 €	215.379.616,39 €
8 Variable costs	-	105.000.000,00 €	104.238.750,00 €	108.929.493,75 €	113.831.320,97 €
9 Fixed costs	-	70.000.000,00 €	73.150.000,00 €	76.441.750,00 €	79.881.628,75 €
10 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
11 Gross Profit		20.000.000,00 €	20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	-	18.000.000,00 €	31.160.000,00 €	33.179.000,00 €	35.349.000,00 €
13 Several Expenses	-	18.000.000,00 €	18.810.000,00 €	19.665.750,00 €	20.541.042,00 €
14 Interests of Loans*	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
15 Net Operating Profit		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	-	300.000,00 €	1.614.812,50 €	86.190,31 €	2.021.980,61 €
18 Net Profit		1.700.000,00 €	9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation		5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	-	- €	- €	- €	48.642.274,00 €
22 Residual Value	-	- €	- €	- €	- €
23 Cash flow		6.700.000,00 €	24.866.062,50 €	35.678.328,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value		51.197.159,31 €			
26 Present value - Cash Flows of the Old Plant	-	38.211.000,00 €			
27 Cash flow increment		12.986.159,31 €			

EuSprIG 2004

18/25

## Example – Heuristic 2

	A	B	C	D	E
1 Inflation Rate (%)		4,50	YEAR 1	YEAR 2	YEAR 3
2 Sale Value - Unit.		80,00 €	83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold		2.500.000	2.625.000	2.756.250	2.894.063
5 Sales Revenue		200.000.000,00 €	219.450.000,00 €	240.791.512,50 €	264.206.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	-	180.000.000,00 €	199.055.416,67 €	207.037.910,42 €	215.379.616,39 €
8 Variable costs	-	105.000.000,00 €	104.238.750,00 €	108.929.493,75 €	113.831.320,97 €
9 Fixed costs	-	70.000.000,00 €	73.150.000,00 €	76.441.750,00 €	79.881.628,75 €
10 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
11 Gross Profit		20.000.000,00 €	20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	-	18.000.000,00 €	31.160.000,00 €	33.179.000,00 €	35.349.000,00 €
13 Severals Expenses	-	18.000.000,00 €	18.810.000,00 €	19.655.750,00 €	20.541.042,00 €
14 Interests of Loans*	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
15 Net Operating Profit		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	-	300.000,00 €	1.614.812,50 €	86.190,31 €	2.021.980,61 €
18 Net Profit		1.700.000,00 €	9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	-	- €	- €	- €	48.642.274,00 €
22 Residual Value	-	- €	- €	- €	- €
23 Cash flow		6.700.000,00 €	24.866.062,50 €	35.678.328,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value		51.197.159,31 €			
26 Present value - Cash Flows of the Old Plant	-	38.211.000,00 €			
27 Cash flow Increment		12.986.159,31 €			

EuSpRIG 2004

19/25

## Heuristic 3

### Assignment based on semantics

- for each label, not yet grouped to layout areas
  - identify semantically assignable cells and
  - geometrically assign them
  - layout areas
  
- subsequently
  - grouping into logical areas [ACM,2000]
  - heuristic 4 is applied

EuSpRIG 2004

20/25

## Example – Heuristic 3

	A	B	C	D	E
			YEAR 1	YEAR 2	YEAR 3
1 Inflation Rate (%)		4,50			
2 Sale Value - Unit.		80,00 €	83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold		2.500.000	2.625.000	2.756.250	2.894.063
5 Sales Revenue		200.000.000,00 €	219.450.000,00 €	240.791.512,50 €	264.208.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	-	180.000.000,00 €	199.055.416,67 €	207.037.910,42 €	215.379.616,39 €
8 Variable costs	-	105.000.000,00 €	104.238.750,00 €	108.929.493,75 €	113.831.320,97 €
9 Fixed costs	-	70.000.000,00 €	73.150.000,00 €	76.441.750,00 €	79.881.628,75 €
10 Depreciation	-	5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
11 Gross Profit		20.000.000,00 €	20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	-	18.000.000,00 €	31.160.000,00 €	33.179.000,00 €	35.349.000,00 €
13 Several Expenses	-	18.000.000,00 €	18.810.000,00 €	19.655.750,00 €	20.541.042,00 €
14 Interests of Loans*	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
15 Net Operating Profit		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax		2.000.000,00 €	10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	-	300.000,00 €	1.614.812,50 €	86.190,31 €	2.021.980,61 €
18 Net Profit		1.700.000,00 €	9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation		5.000.000,00 €	21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	-	- €	12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	-	- €	- €	- €	48.642.274,00 €
22 Residual Value	-	- €	- €	- €	- €
23 Cash flow		6.700.000,00 €	24.866.062,50 €	35.678.328,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value		51.197.159,31 €			
26 Present value - Cash Flows of the Old Plant	-	38.211.000,00 €			
27 Cash flow Increment		12.986.159,31 €			

EuSpRIG 2004

21/25

## Heuristic 4

### Spatial Assignment

- identify cells, which are not yet assigned and
- geometrically assign them
  - layout areas
- no further grouping is done as
  - no statement on semantical relatedness is provided

EuSpRIG 2004

22/25

## Example – Heuristic 4

	A	B	C	D	E
			YEAR 1	YEAR 2	YEAR 3
1 Inflation Rate (%)		4,50			
2 Sale Value - Unit.	80,00 €		83,60 €	87,36 €	91,29 €
3 Real growth of Sales (%)		5,00			
4 Units Sold	2.500.000		2.625.000	2.756.250	2.894.063
5 Sales Revenue	200.000.000,00 €		219.450.000,00 €	240.791.512,50 €	264.208.487,09 €
6 Decrease of Variable Costs (%)		5,00			
7 Total costs	- 180.000.000,00 €		- 199.055.416,67 €	- 207.037.910,42 €	- 215.379.616,39 €
8 Variable costs	- 105.000.000,00 €		- 104.238.750,00 €	- 108.929.493,75 €	- 113.831.320,97 €
9 Fixed costs	- 70.000.000,00 €		- 73.150.000,00 €	- 76.441.750,00 €	- 79.881.628,75 €
10 Depreciation	- 5.000.000,00 €		- 21.666.666,67 €	- 21.666.666,67 €	- 21.666.666,67 €
11 Gross Profit	20.000.000,00 €		20.394.583,33 €	33.753.602,08 €	48.828.870,71 €
12 Operating Expenses	- 18.000.000,00 €		- 31.160.000,00 €	- 33.179.000,00 €	- 35.349.000,00 €
13 Severe Expenses	- 18.000.000,00 €		- 18.810.000,00 €	- 19.655.750,00 €	- 20.541.042,00 €
14 Interests of Loans*	- €		- 12.350.000,00 €	- 13.523.250,00 €	- 14.807.958,00 €
15 Net Operating Profit	2.000.000,00 €		10.765.416,67 €	574.602,08 €	13.479.870,71 €
16 Earnings Before Income Tax	2.000.000,00 €		10.765.416,67 €	574.602,08 €	13.479.870,71 €
17 Income Tax (15%)	- 300.000,00 €		- 1.614.812,50 €	- 86.190,31 €	- 2.021.980,61 €
18 Net Profit	1.700.000,00 €		9.150.604,17 €	488.411,77 €	11.457.890,10 €
19 Depreciation	5.000.000,00 €		21.666.666,67 €	21.666.666,67 €	21.666.666,67 €
20 Interests of Loans	- €		12.350.000,00 €	13.523.250,00 €	14.807.958,00 €
21 Financing (payment)	- €		- €	- €	- 48.642.274,00 €
22 Residual Value	- €		- €	- €	- €
23 Cash flow	6.700.000,00 €		24.866.062,50 €	35.678.326,44 €	709.759,23 €
24 Cash flow to Present Value			22.503.223,98 €	29.219.981,93 €	526.046,60 €
25 Cash flow Sum - Present Value	51.197.159,31 €				
26 Present value - Cash Flows of the Old Plant	- 38.211.000,00 €				
27 Cash flow increment	12.986.159,31 €				

EuSpRIG 2004

23/25

## Summary

- The user is relieved to be highly experienced in a specific visualization technique as
  1. explicit parametrization is avoided
  2. implicit identification of the orientation of blocks of cells is possible
  3. the orientation is identified for each block individually
  4. semantically related blocks of cells with different sizes can be identified

EuSpRIG 2004

24/25



---

# ?Questions?