Slice-based Formal Specification Measures

Mapping Coupling and Cohesion Measures to Formal Z

Andreas Bollin
Software Engineering & Soft Computing
Alpen-Adria Universität Klagenfurt
Austria

Content

- 1. Motivation
  - Short status-quo, Evolution of Formal Specifications
  - Why looking at measures?
- 2. Basics
  - Dependencies and Slices
  - Quality measures
- 3. Assessing the quality of Z specifications
  - Sensitivity
  - Correlations
  - Longitudinal Study

Motivation (1/2)

- Specification and Verification is taught at AAU Klagenfurt in the 5th semester (computer-science curriculum).
  - Typical students' quotes:
    - "Formal specifications are too complex"
    - "Is there Eclipse support?"
  - Approx. 10-15% of them then select the course "Systematic SW development" in their 6th semester

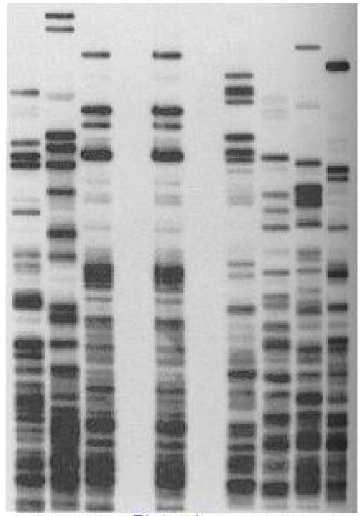
- Several projects show that formal specification languages are not "Write-Once Languages"

  Requirement changes lead to changes, as the specification is only beneficial when kept up-to-date

➔ Formal Specifications evolve

Motivation (2/2)

- **How to detect such a (supposed) evolution?**
  - There are a lot of papers and articles dealing with programming measures, but hardly any for formal specifications
  - Most of the measures around are quantity/size-based
- **What is needed?**
  - Fingerprint (quantitative and qualitative formal specification measures)
  - Answers to the question: How complex are different versions of a specification? What about their quality?



- **This contribution reports on our experiences**
  1. Looked for a set of quality measures and analyzed their expressiveness
  2. Collected sample Z specifications (11.182 lines, 4.614 LSC, 222 p.) and used them to analyze the measures' correlations
  3. Examined the WSDL Specification (16.641 lines, 1.413 LSC, 116 p.) and its 139 versions in the CVS

Software Engineering SESC and Soft Computing
UNIVERSITÄT KLAGENFURT
Institut für Informatik-Systeme

Basics | Dependencies and Slices

- **The bases for the calculation of the measures are**
  - **Dependencies ("Data" and "Control") and specification slices**

  - **For Z: approach is going back to the works of Oda and Araki (1993), Chang and Richardson (1994), and Bollin (2004)**
    - **Syntactical approximation to a semantical analysis**
    - **Predicates that are part of the post-condition of a schema are control dependent on predicates that are part of the pre-condition**
    - **Has been refined and implemented in a Java prototype called ViZ**
  - **To avoid confusion with the Z precondition calculation operation**
    - **Introduced the notion of a specification prime**
    - **Primes correspond to the predicates and declarations of a specification**
    - **Their equivalent in programming languages would be "source instructions"**

Basics | Quality (1/3)

- Quality is quite often measured by the dual properties of coupling and cohesion
  - Coupling
    - Number of local info flow entering (fan-in) and leaving (fan-out) a procedure, so degree of mutual interdependence
    - ➔ Can be calculated via slicing [Harman, 97]
  - Cohesion
    - Intra-modular functional relatedness of a SW module
    - ➔ Weiser introduced tightness, overlap, coverage, parallelism, clustering as indicators
- The underlying idea: use of slice profiles
  - A **Slice Profile (SP)** is the set of the slices for all possible points of interest, its union is called **Slice Union (SU)**
  - In our case, the points of interest are the primes representing predicates in a Z schema

Basics | Quality (2/3)

❑ The values for cohesion are calculated as follows:

Measure	*Definition*	*Description*
Tightness ($\tau(\Psi, \psi)$)	$\frac{\lvert SP_{int}(\Psi,\psi)\rvert}{CC(\psi)}$	Tightness measures the number of primes in every slice.
MinCoverage ($Cov_{min}(\Psi, \psi)$)	$\frac{1}{CC(\psi)} \mid SP_{i-min} \mid$	MinCoverage is the ratio of the smallest slice SP_i in the slice profile $SP(\Psi, \psi)$ to the length of ψ.
Coverage ($Cov(\Psi, \psi)$)	$\frac{1}{n}\sum_{i=1}^{n} \frac{\lvert SP_i\rvert}{CC(\psi)}$	Coverage compares the length of all possible specification slices SP_i ($SP_i \in SP(\Psi, \psi)$) to the length of ψ.
MaxCoverage ($Cov_{max}(\Psi, \psi)$)	$\frac{1}{CC(\psi)} \mid SP_{i-max} \mid$	MaxCoverage is the ratio of the largest slice SP_i in the slice profile $SP(\Psi, \psi)$ to the length of ψ.
Overlap ($O(\Psi, \psi)$)	$\frac{1}{n}\sum_{i=1}^{n} \frac{\lvert SP_{int}(\Psi,\psi)\rvert}{\lvert SP_i\rvert}$	Overlap measures how many primes are common to all possible specification slices SP_i ($SP_i \in SP(\Psi, \psi)$).

SP contains n slices for a given schema, SP_i is one slice out of the slice profile, SP_{int} is the intersection of all the slices in the slice profile SP, $SP_{i\text{-}min}$ is the smallest slice in the slice profile, $SP_{i\text{-}max}$ is the smallest slice in the slice profile, CC counts the number of primes

Basics | Quality (3/3)

❑ The values for coupling are defined as follows:

Measure	*Definition*	*Description*
Inter-Schema Flow – $F(\psi_s, \psi_d)$	$\frac{\lvert(SU(\psi_d) \cap \psi_s)\rvert}{\lvert\psi_s\rvert}$	Inter-Schema flow F measures the number of primes of the slices in ψ_d that are in ψ_s.
Inter-Schema Coupling – $C(\psi_s, \psi_d)$	$\frac{F(\psi_s,\psi_d)\ \lvert\psi_s\rvert + F(\psi_d,\psi_s)\ \lvert\psi_d\rvert}{\lvert\psi_s\rvert + \lvert\psi_d\rvert}$	Inter-Schema coupling C computes the normalized ratio of the flow in both directions.
Schema Coupling – $\chi(\psi_i)$	$\frac{\sum_{j=1}^{n} C(\psi_i,\psi_j)\ \lvert\psi_j\rvert}{\sum_{j=1}^{n} \lvert\psi_j\rvert}$	Schema Coupling χ is the weighted measure of inter-schema coupling of ψ_i and all n other schemata.

➔ **Up to now we defined measures for a specification's quality (coupling, cohesion). Now the following steps are missing**

- **Demonstration that the measures are meaningful and sensitive**
- **Demonstration that the measures are not just proxies for size measures**
- **Demonstration of the use of the measures**

Assessment | Sensitivity (1/3)

- **Step 1 (Sensitivity): It is checked by looking at sample operations and their influence on the measures**
  - **Adding a precondition-prime. This should increase cohesion values as the internal semantic connections are extended**
  - **Adding a postcondition-prime**
    - **When it is not related to all other primes then new "trains of thought" are introduced, decreasing the value for coupling**
    - **When it is related to all other primes then it extends existing "trains of thought", so the values for coupling should increase**
  - **Coupling is checked by a specification that contains structurally equivalent operation schemata**
    - **Raising the number of relations should increase the value for coupling**
    - **Removing relations between the schemata should decrease the value**

Example for schema cohesion values

$SP(\psi)$			Specifications	Measures
		—	*Test1* $n, n' : \mathbb{N}$ $m, m' : \mathbb{N}$ ――― $n' = n + 1$	$\#SP_{int}(\psi) = 1$ $\tau(\psi) = 1.00$ $Cov_{min}(\psi) = 1.00$ $Cov(\psi) = 1.00$ $Cov_{max}(\psi) = 1.00$ $O(\psi) = 1.00$
	—	—	*Test2* $n, n' : \mathbb{N}$ $m, m' : \mathbb{N}$ ――― $n' = n + 1$ $m' = m + 1$	$\#SP_{int}(\psi) = 0$ $\tau(\psi) = 0.00$ $Cov_{min}(\psi) = 0.50$ $Cov(\psi) = 0.50$ $Cov_{max}(\psi) = 0.50$ $O(\psi) = 0.00$
	— —	— —	*Test3* $n, n' : \mathbb{N}$ $m, m' : \mathbb{N}$ $delta? : \mathbb{N}$ ――― $delta? > 0$ $n' = n + delta?$ $m' = m - delta?$	$\#SP_{int}(\psi) = 1$ $\tau(\psi) = 0.33$ $Cov_{min}(\psi) = 0.67$ $Cov(\psi) = 0.67$ $Cov_{max}(\psi) = 0.67$ $O(\psi) = 0.50$
	— — —	— — —	*Test4* $n, n' : \mathbb{N}$ $m, m' : \mathbb{N}$ $delta? : \mathbb{N}$ $set? : \mathbb{P}\mathbb{N}$ ――― $delta? > 0$ $set? \neq \varnothing$ $n' = n + delta?$ $m' = m - delta?$	$\#SP_{int}(\psi) = 2$ $\tau(\psi) = 0.50$ $Cov_{min}(\psi) = 0.75$ $Cov(\psi) = 0.75$ $Cov_{max}(\psi) = 0.75$ $O(\psi) = 0.67$
— — —	— — —	— — —	*Test5* $n, n' : \mathbb{N}$ $m, m' : \mathbb{N}$ $delta? : \mathbb{N}$ $set? : \mathbb{P}\mathbb{N}$ $p, p' : \mathbb{N}$ ――― $delta? > 0$ $set? \neq \varnothing$ $n' = n + delta?$ $m' = m - delta?$ $p' = p + delta?$	$\#SP_{int}(\psi) = 2$ $\tau(\psi) = 0.40$ $Cov_{min}(\psi) = 0.60$ $Cov(\psi) = 0.60$ $Cov_{max}(\psi) = 0.60$ $O(\psi) = 0.33$

Assessment | Sensitivity (3/3)

- The analysis of the different situations shows that
  - The values for coupling and cohesion are sensitive to modifications of the specification
  - The values behave as expected

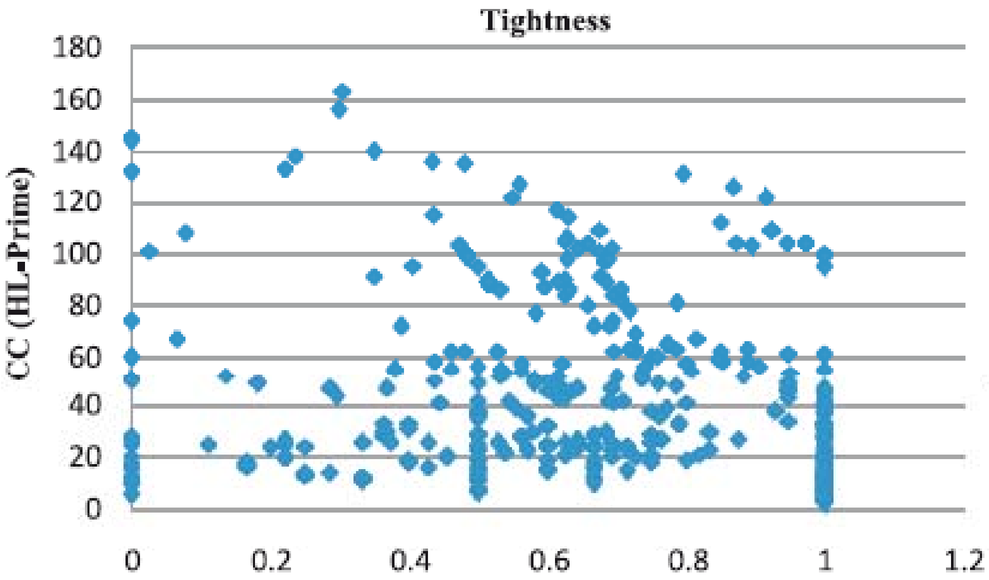
- However, there are some limitations
  - When the specifications are very dense (textbook examples) then the slices are as big as the specification. This effect diminishes when the specifications are getting larger
  - Coupling is not sensitive to changes that add additional dependencies between primes
  - Slicing only works fine when the specifications are “well-formed” (which means that the syntactical approximation really works)

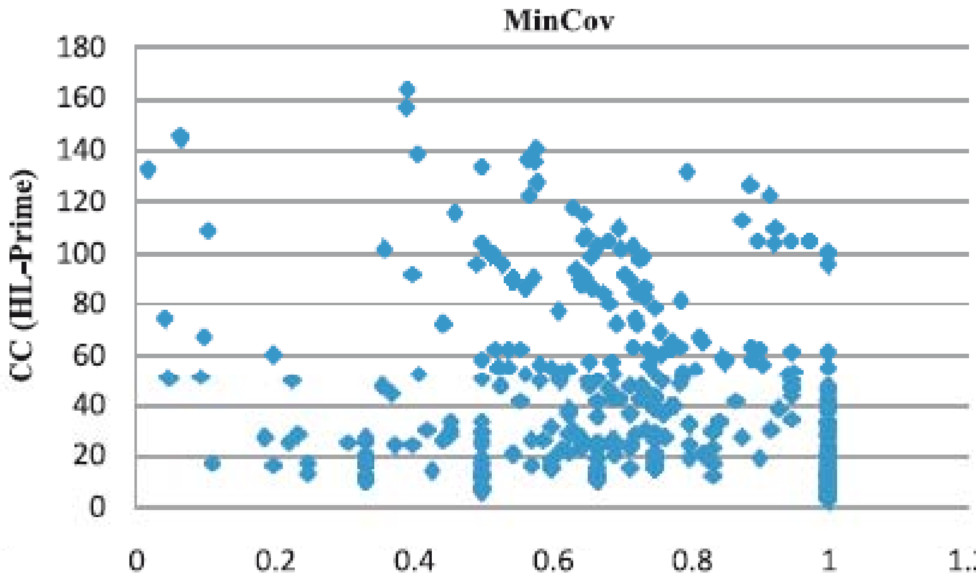
Universität Klagenfurt
Institut für Informatik-Systeme

Assessment | Correlations (1/4)

- **Step 2 (Correlations): The experimental subjects are**
  - **A collection of freely available Z specifications**
  - **11.186 lines of specification text, 613 operation schemata**
  - **Approx. 6000 slices have been calculated**

- **1st step: Looking at the relation between size and quality measures**





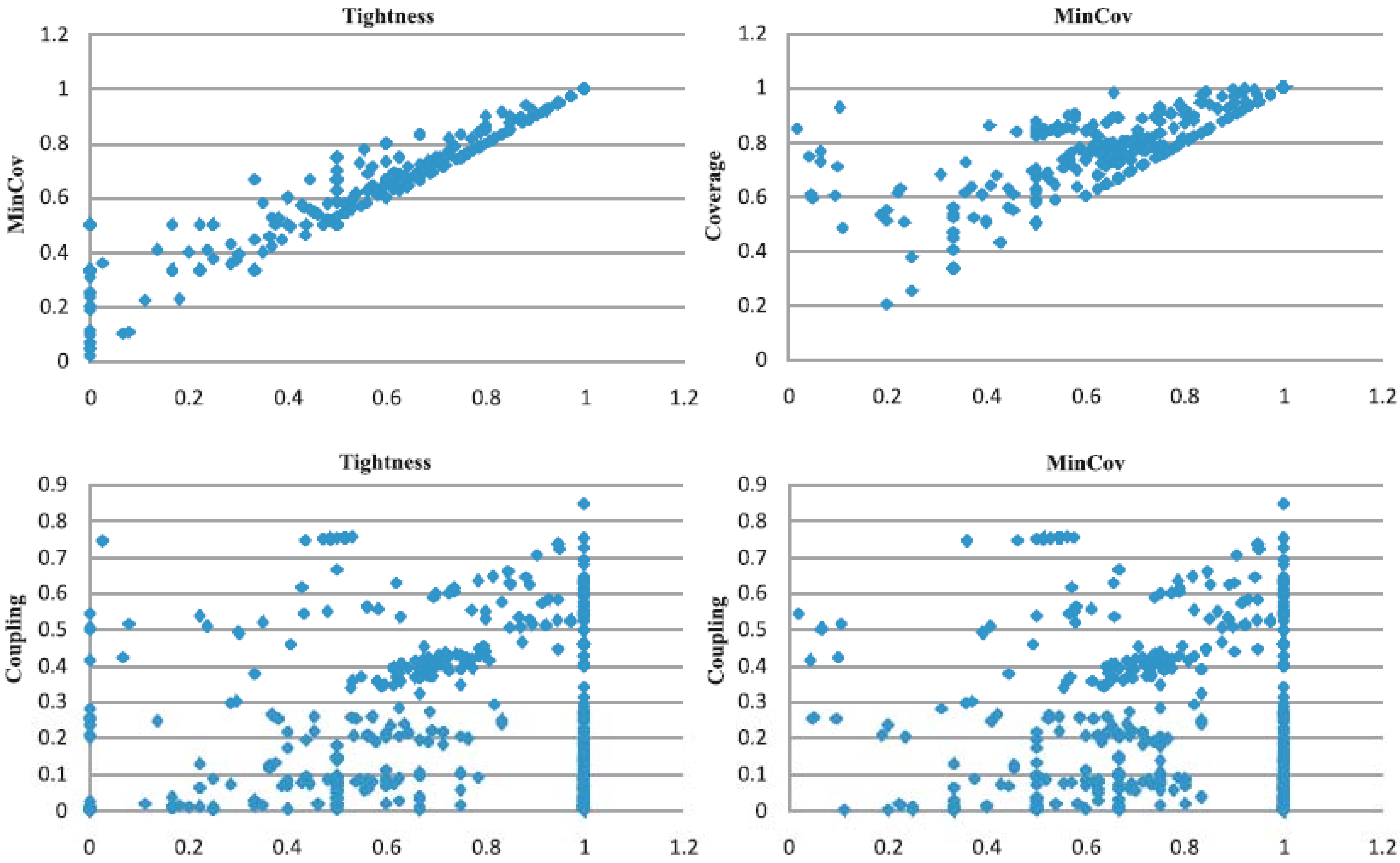
Assessment | Correlations (2/4)

- 1st step (contd.): A closer look at size and quality measures

Conceptual Complexity Correlation (n=613)						
	Pearson		Spearman		Kendall	
Measure	*R*	*p*	*R*	*p*	*R*	*p*
Tightness	-.466	.000	-.416	.000	-.308	.000
MinCov	-.540	.000	-.456	.000	-.342	.000
Coverage	-.216	.000	-.297	.000	-.199	.000
MaxCov	0.118	.002	-.120	.001	-.067	.012
Overlap	-.528	.000	-.475	.000	-.362	.000
Coupling	0.605	.000	0.726	.000	0.524	.000

Assessment | Correlations (3/4)

❑ 2nd step: Coupling and Cohesion Measures



Assessment | Correlations (4/4)

❑ 2nd step: Coupling and Cohesion Measures (contd.)

Metric Comparison (n=613)								
			Pearson		Spearman		Kendall	
Sig.	*Measure 1*	*Measure 2*	*R*	*p*	*R*	*p*	*R*	*p*
Strong	*Tightness*	*MinCov*	0.970	.000	0.977	.000	0.910	.000
	Tightness	*Coverage*	0.845	.000	0.888	.000	0.734	.000
	Tightness	*Overlap*	0.913	.000	0.886	.000	0.796	.000
	MinCov	*Coverage*	0.837	.000	0.901	.000	0.751	.000
	MinCov	*Overlap*	0.849	.000	0.834	.000	0.726	.000
	Coverage	*MaxCov*	0.859	.000	0.828	.000	0.709	.000
Moderate	*Tightness*	*MaxCov*	0.525	.000	0.557	.000	0.449	.000
	MinCov	*MaxCov*	0.514	.000	0.589	.000	0.478	.000
	Coverage	*Overlap*	0.613	.000	0.628	.000	0.502	.000
Weak	*Tightness*	*Coupling*	-.101	.006	-.100	.007	-.055	.000
	MinCov	*Coupling*	-.181	.000	-.158	.000	-.101	.000
	Coverage	*Coupling*	0.150	.000	0.044	.138	0.044	.058
	MaxCov	*Coupling*	0.333	.000	0.190	.000	0.171	.000
	Overlap	*Coupling*	-.173	.000	-.213	.000	-.308	.000
	MaxCov	*Overlap*	0.263	.000	0.234	.000	0.181	.000

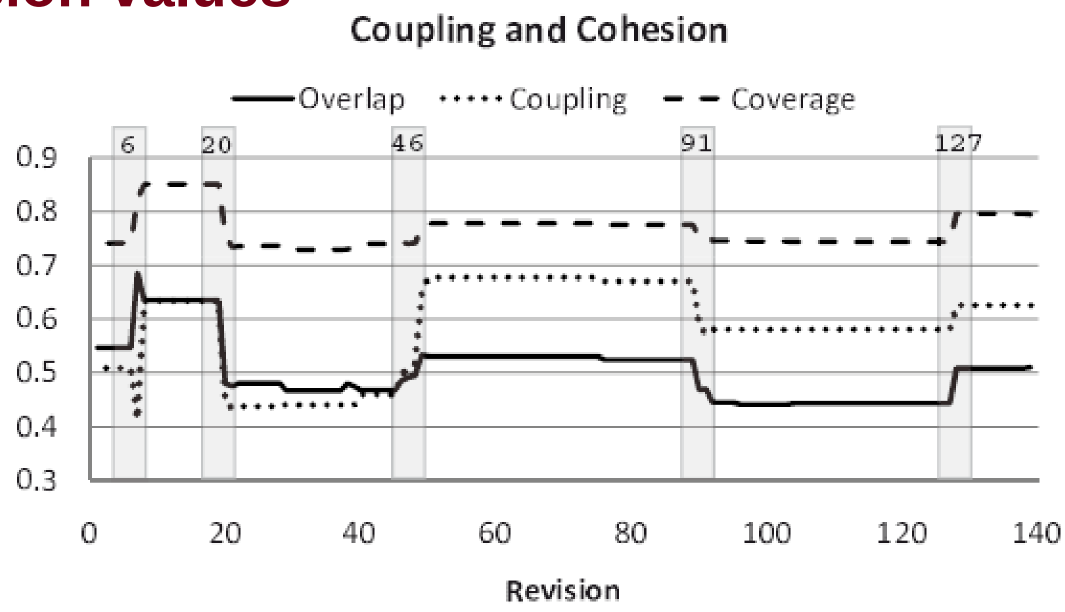
Assessment | Evolution (1/4)

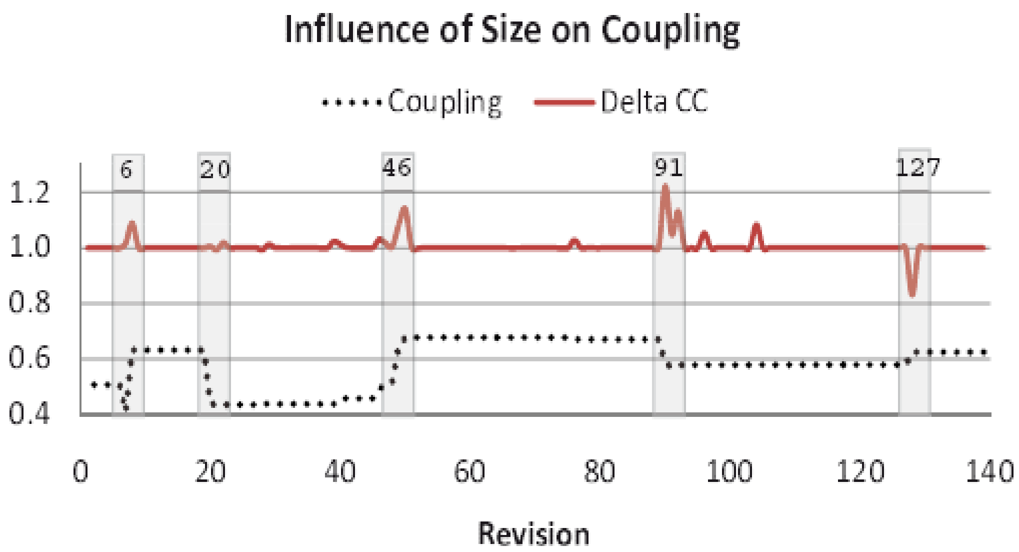
- Step 3 (Applicability): Z Specification of the Web Service Definition Language (WSDL)
  - Version 1.0 released 2002 (no Z specification), Version 2.0 released 2006, development started in 2004 with a Z specification
  - Up to 2006 exactly 139 revisions have been checked in (47 of them relevant, >80.000 Primes)
- Some observations from the CVS log:
  - Till Rev. 1.6: structural changes and finally a model extension
  - Revision 1.20: Last calls, then refactoring, bug fixing
  - Revisions 1.28/1.29: Refactoring, last calls
  - Revisions 1.46ff: Several extensions to the model
  - Revision 1.77: Maintenance action
  - Revisions 1.91ff: Massive extensions of model
  - Revisions 1.96, 1.104: Simplification of model
  - Revisions 1.127ff: Change requests, massive refactoring

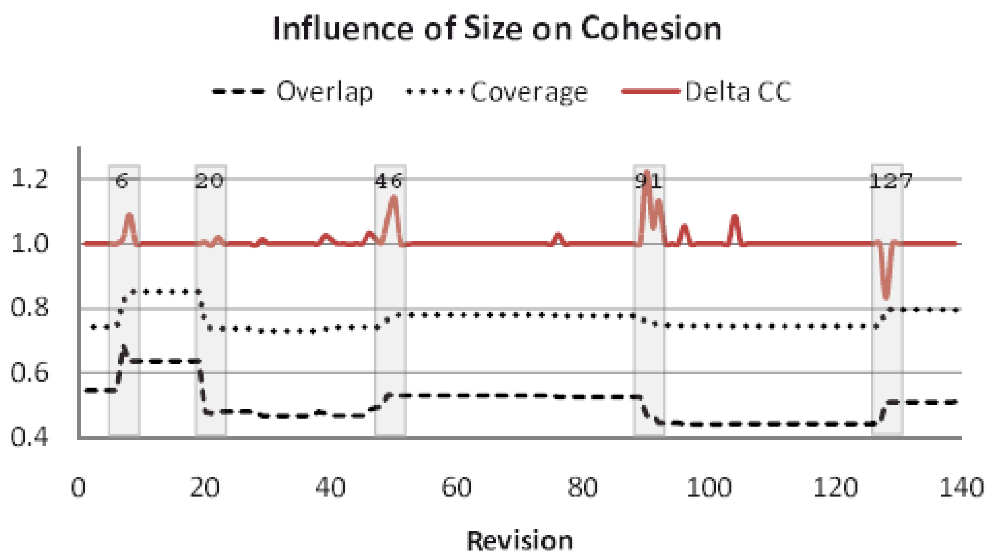
Assessment | Evolution (2/4)

Influence on coupling and cohesion values

- Are sensitive to changes of the specification
  - Coupling increases a bit
  - Cohesion decreases a bit
- At rev. 1.6: Coupling and cohesion are not related!
- The size of the modification is not the dominant factor







Assessment | Evolution (3/4)

❑ How to measure the evolution of formal specifications?

- Idea: An "ideal" specification operation contains a "single thought" per schema
- One approach might be to look at the **size of the schema (CC)** and the size of the **slice intersection** $|SP_{int}|$
- Progress towards such an "ideal" specification would appear as a convergence between schema size and |SLint|.

❑ "Divergence" and "Deterioration"

- A divergence indicates some modification (deterioration?)
- The divergence $\delta(\Psi)$ is defined as the average schema size in respect to the average size of the slice intersections
- It makes more sense to look at the difference in "deterioration" between consecutive versions of the specification. This is expressed by the notion of the relative deterioration $\rho(\Psi_n)$

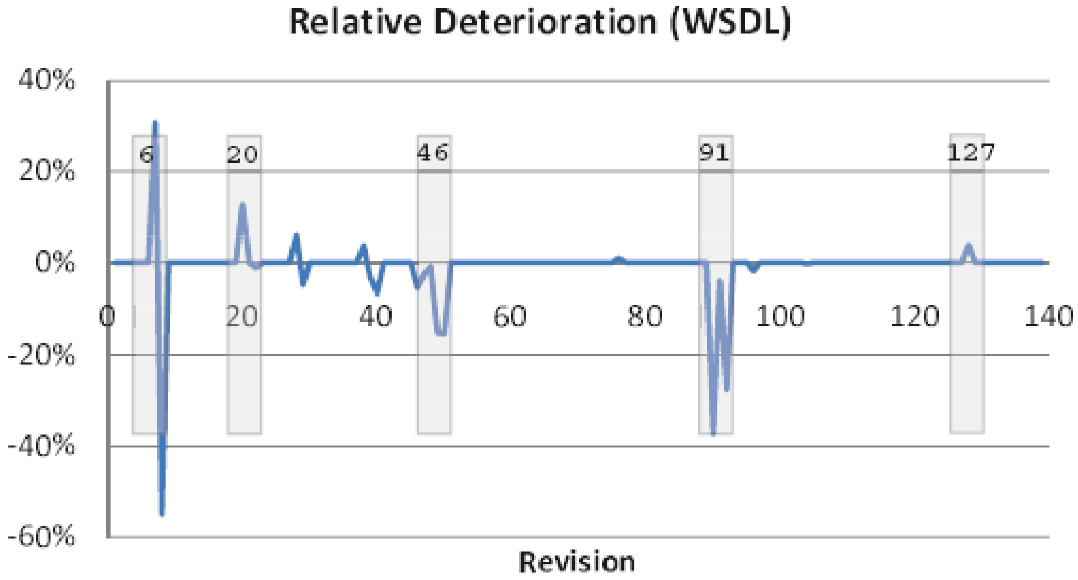
$$\delta(\Psi) = \frac{\sum_{i=1}^{n} CC(\psi_i) - |\, SP_{int}(\Psi, \psi_i)\, |}{n}$$

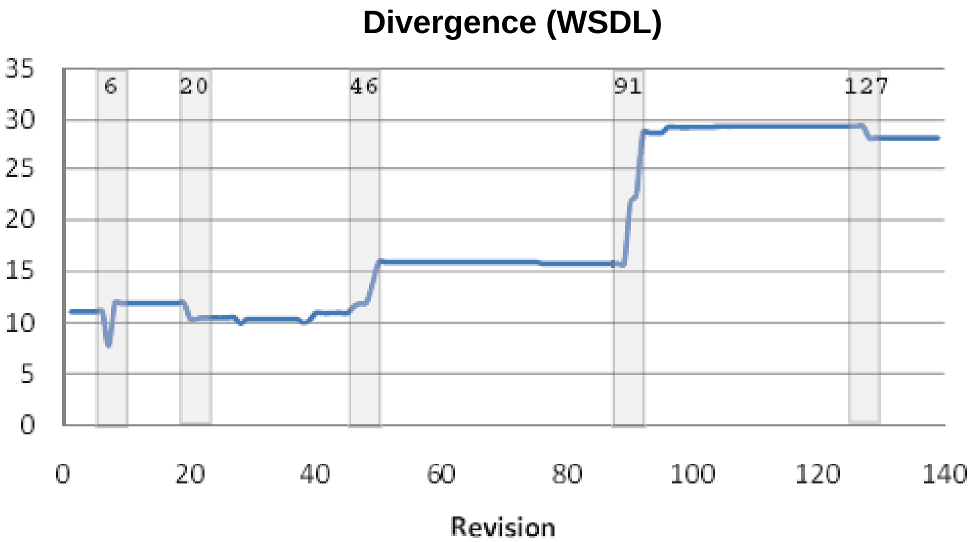
$$\rho(\Psi_n) = 1 - \frac{\delta(\Psi_{n-1})}{\delta(\Psi_n)}$$

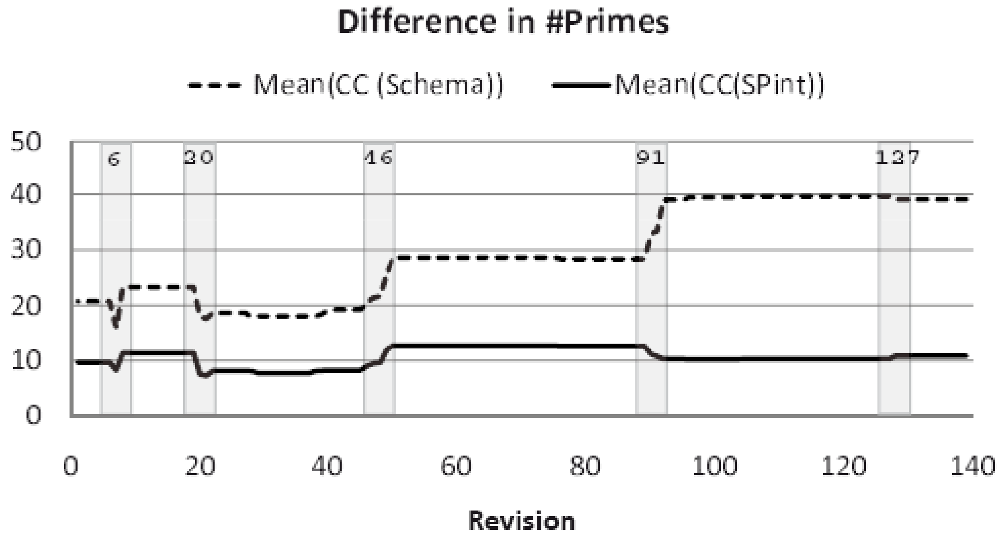
Assessment | Evolution (4/4)

Divergence and Deterioration

- Divergence alone is not very expressive → Rel. Deterioration
- A change request typically has 3 phases (good to be seen between revs.1.20 and 1.46):
  - Refactoring of actual version
  - Adding/Removing/Modifying "functionality"
  - Adjusting the documentation







Summary

- This contribution presented an approach of mapping slice-based measures to formal Z specifications
- Pros
  - Coupling and cohesion measures (based on slice profiles) can be reasonably mapped to formal Z specifications. They are not just proxies for size-based specification measures
  - The measures are sensitive to changes of the specification and “react” quite similar to the corresponding values of programming languages
  - Could be used as a fingerprint for specifications, but also in observing the effects of maintenance operations
- Cons
  - The empirical basis is still very small, additional data would be helpful (baseline, …)
  - The syntactical approximation (for slice generation) might be a stumbling block

Summary

Thank you …

Questions / Hints ?