

Code Clones: Good, Bad, or Ugly?

Maarit Harsu, Tampere University of Technology, Finland

Tibor Bakota, University of Szeged, Hungary

István Siket, University of Szeged, Hungary

Kai Koskimies, Tampere University of Technology, Finland

Tarja Systä, Tampere University of Technology, Finland

Outline

- Introduction
- Background
- Different clone types (and their non-clone counterparts)
- Coupling in cloned and non-cloned situations
- Conclusions

Introduction

- Code cloning (copy-paste programming, duplicated code) is typically considered a bad convention, especially from the point of view of maintenance
 - problems in finding and changing the copied code
- Code cloning research
 - clone detection (to remove the clones)
 - linking clones together (to make changing easier)
 - classification of clones
 - reasons for cloning
 - situations where cloning is not so bad

Background

- Quality assessment of five systems from three Finnish embedded system manufacturers, from the area of machine industry
 - Inverse correspondence between coupling (CBO*) and clones
 - See the table on the next slide
- Searching for the explanation for the correspondence
 - Replacing clones with method calls may increase coupling
 - Different clone types and their non-clone counterparts

**CBO: coupling between object classes*

Coupling and clones

	System A	System B	System C	System D	System E
LOC	22,305	164,559	121,747	341,216	59,740
NCL	134	1,498	582	919	502
CBO-index	-8.85	-7.6	-6.15	-3.74	1.17
NOI-index		-7.97	-4.67	-2.56	1.39
CC (%)	32.7	16.9	11.44	9.94	7.47
CI		1505	741	2700	248
CI/LOC		0.009	0.006	0.008	0.004
Max-DCFLOC	307	507	110	206	60

LOC: lines of code

NCL: number of classes

NOI: number of outgoing invocations

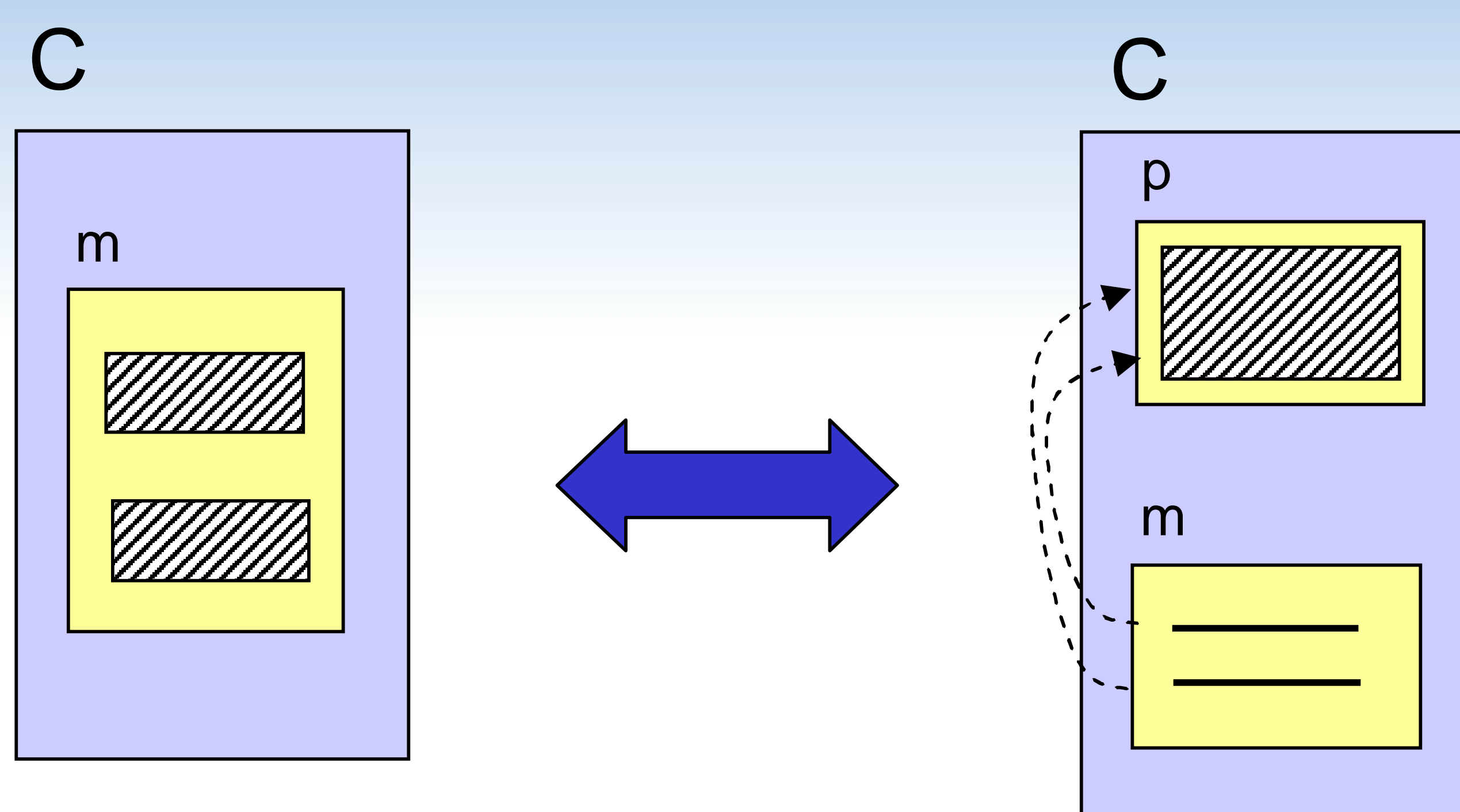
CC: clone coverage

CI: clone instances

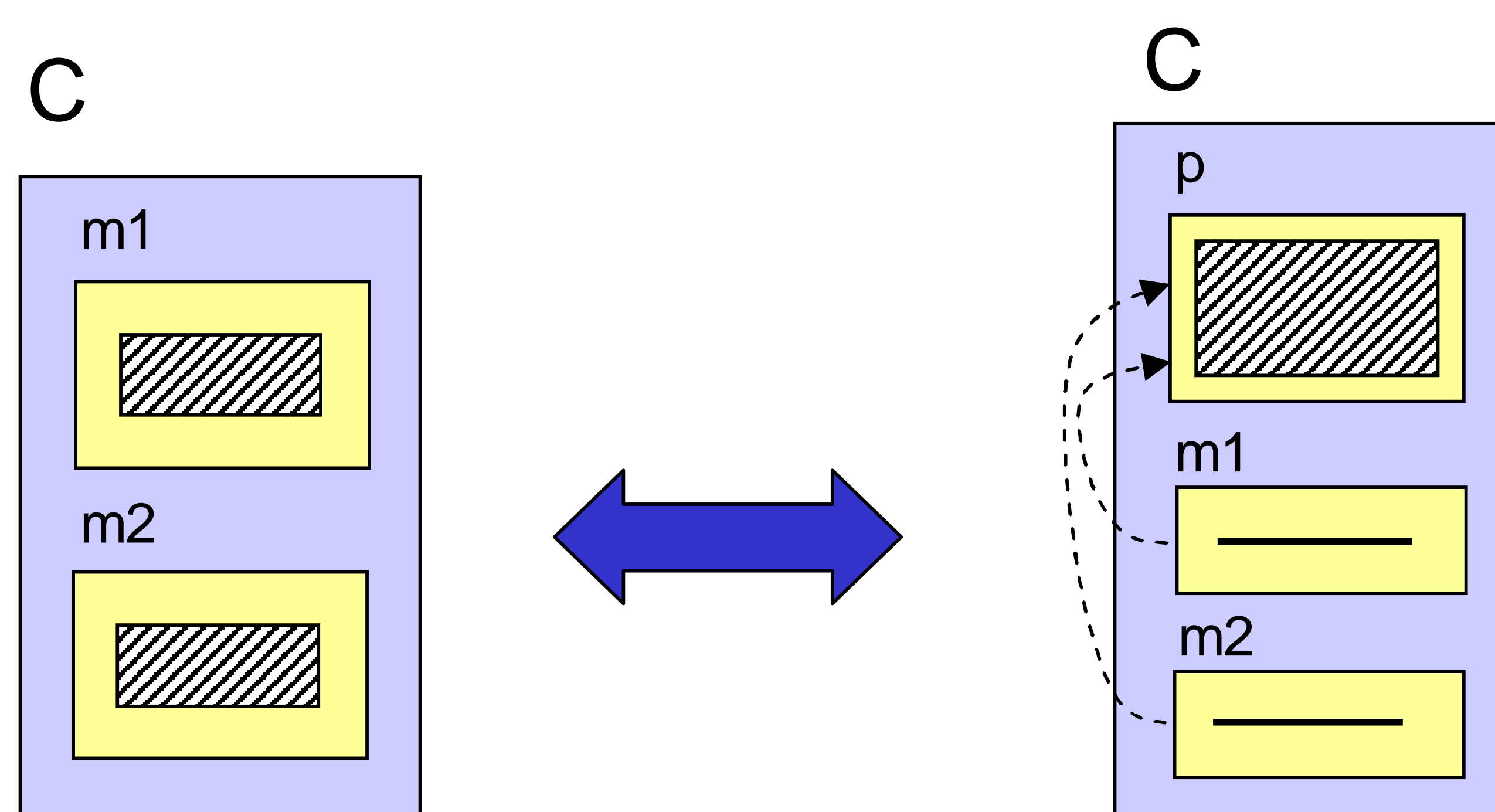
DCFLOC: lines of code for code duplications

Clone types 1 and 2 (with their non-clone counterparts)

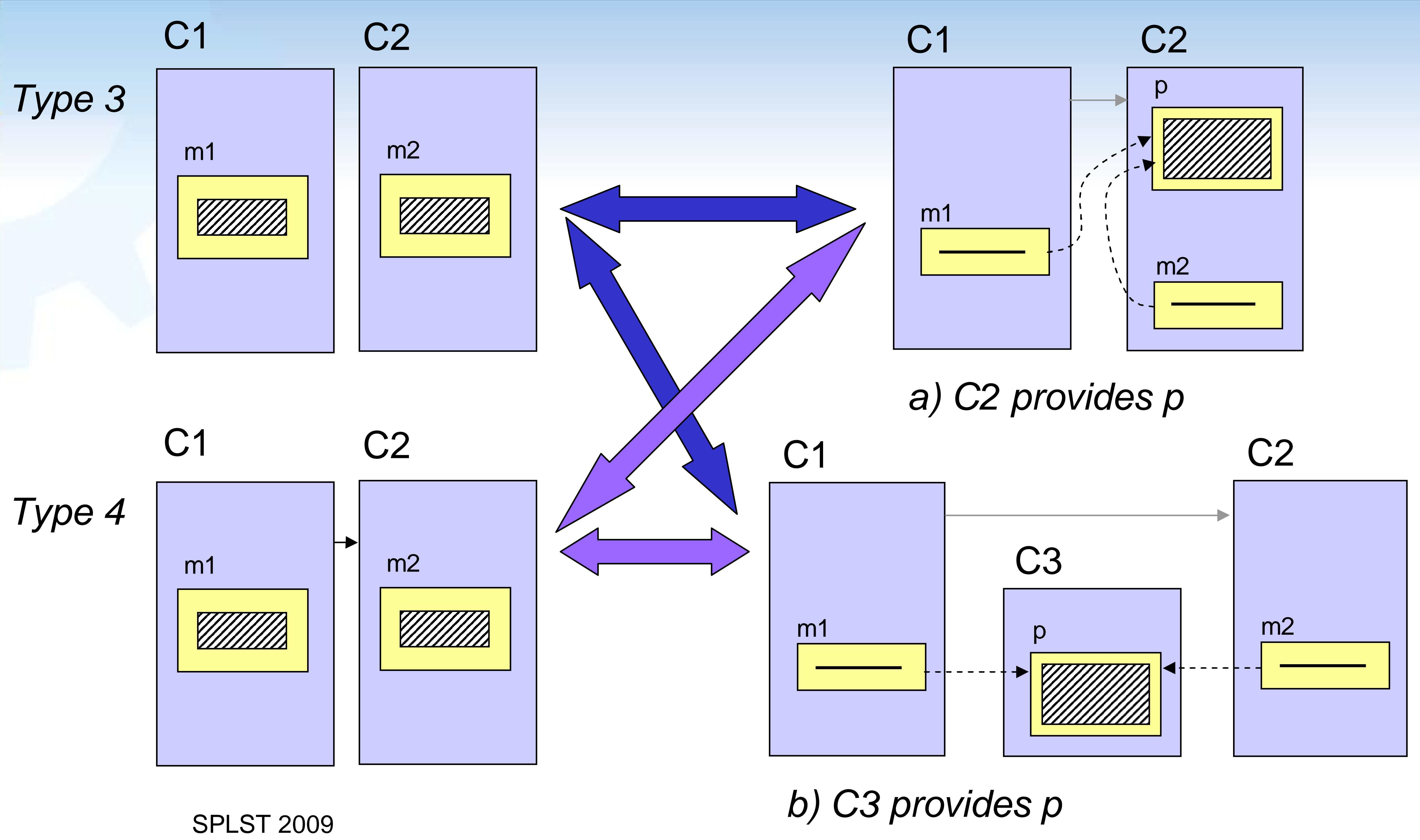
Type 1



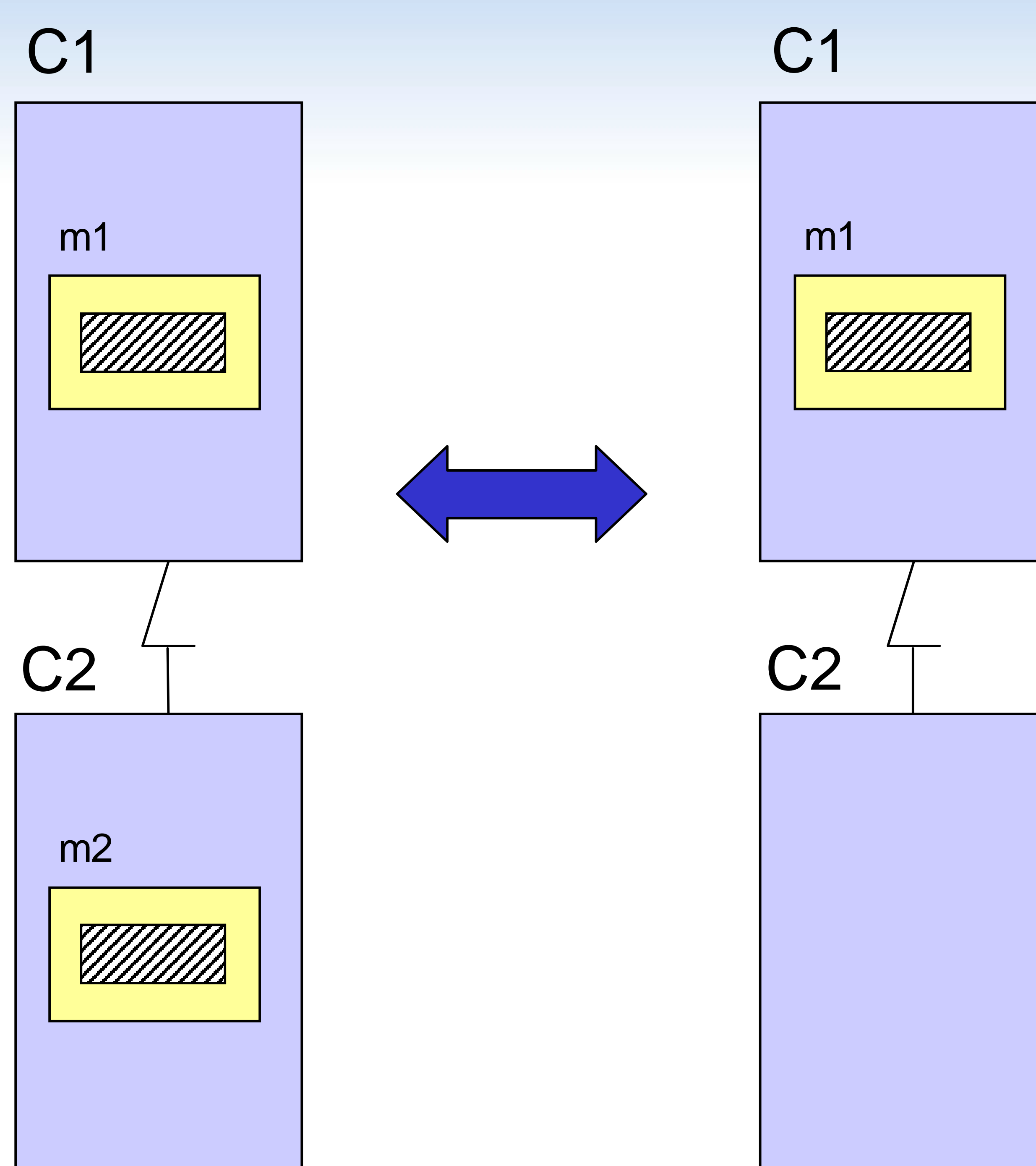
Type 2



Clone types 3 and 4 (with their non-clone counterparts)

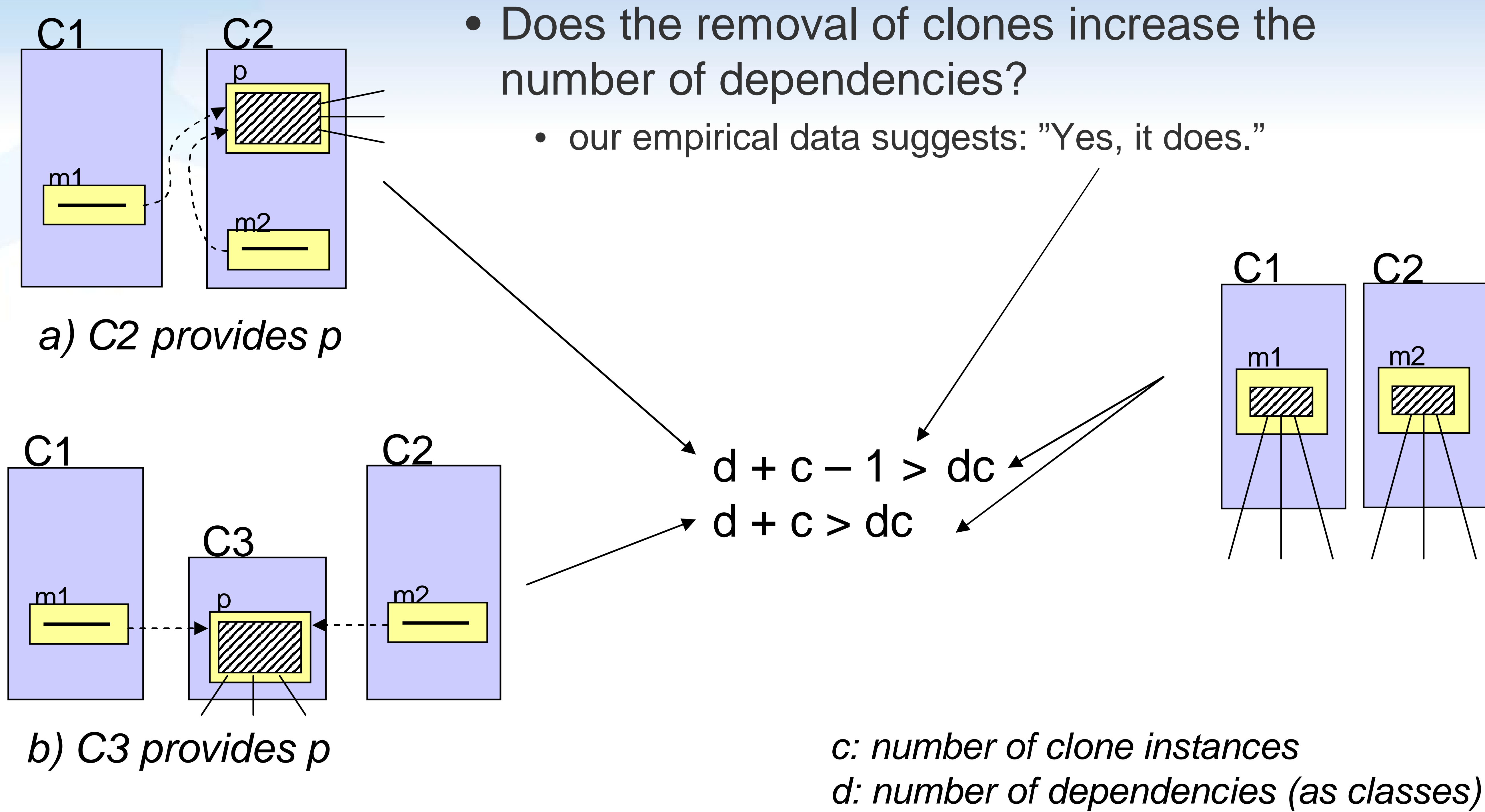


Clone type 5 (with its non-clone counterpart)



Coupling (number of dependencies)

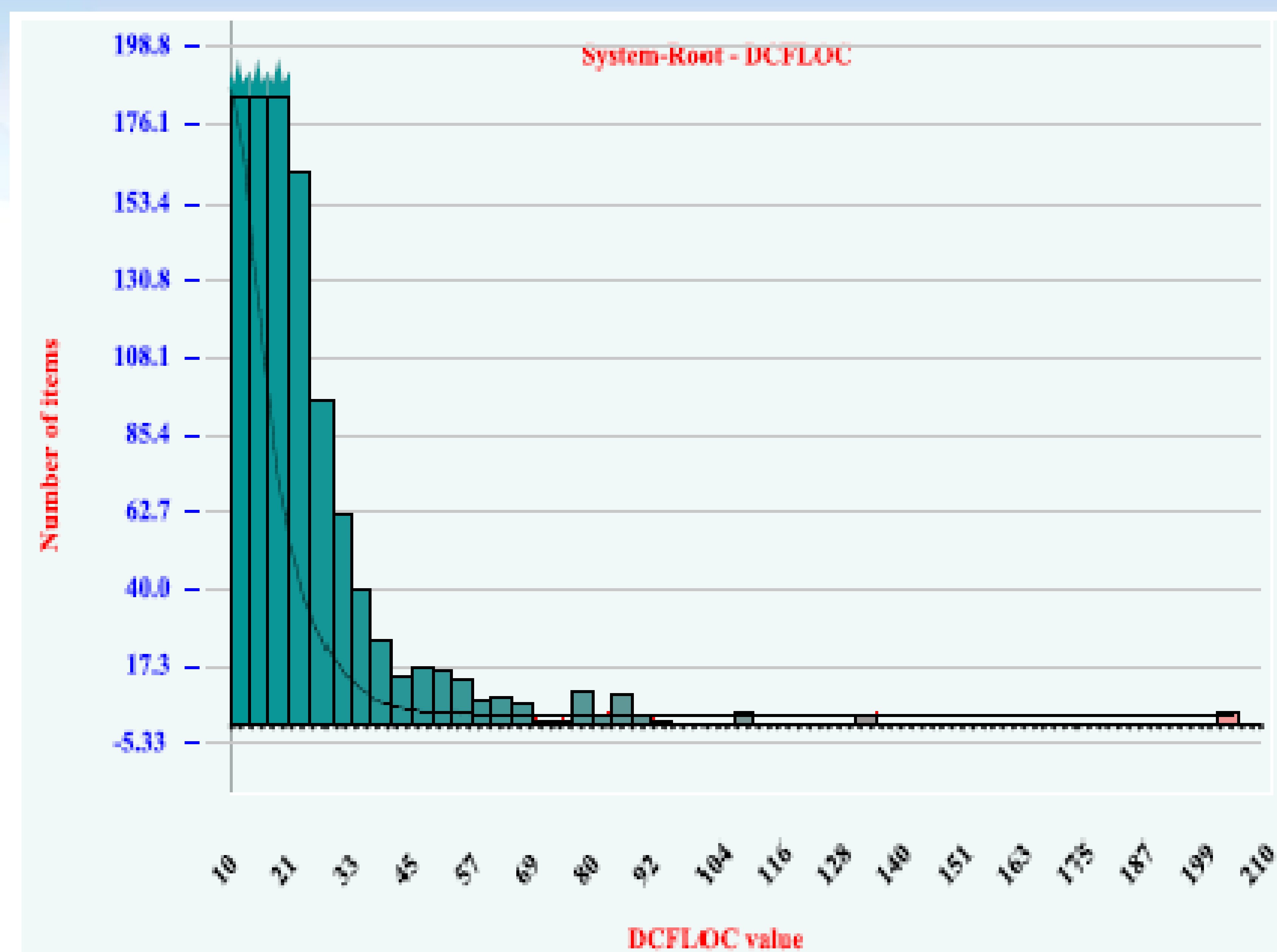
Clone type 3 as an example



Analysis (of the previous slide)

- What does it mean that d (number of dependencies) is very small
 - such clones have none or very few connections to other classes
 - they are most probably small
 - they are not very critical from the point of view of maintenance, because the changes have only local effects
- This reasoning is in line with our empirical data
 - see the histogram in the next slide

Empirical characteristics of the clones



SPLST 2009

Conclusions

- Possible explanations for "good" clones found
- More empirical evidence needed
- Trade-off between coupling and cloning

Questions?