Defining a Catalog of Programming Anti-Patterns for Concurrent Java

Jeremy S. Bradbury, Kevin Jalbert

Software Quality Research Group
Faculty of Science (Computer Science)
University of Ontario Institute of Technology
Oshawa • Ontario • Canada
jeremy.bradbury@uoit.ca, kevin.jalbert@mycampus.uoit.ca

SPAQu'09 • October 25, 2009





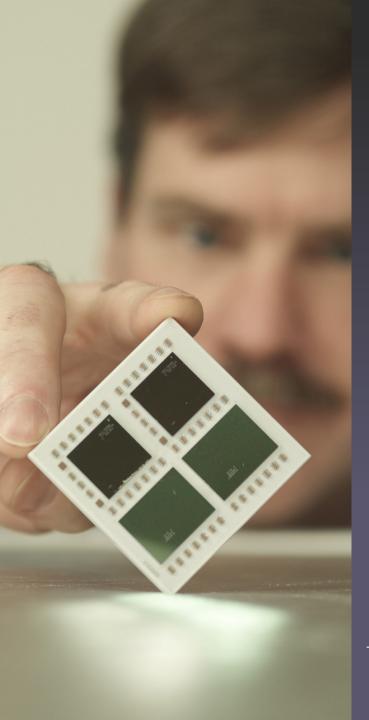
Leveraging the ful power of multicore processors demands new tools and new thinking from the software industry.

Software and the Concurrency Revolution

bg thing" and "the way of the finare," be fire the part able to agree at the partial feature has a size of the partial read with the comparer architecture for unsuppressed with the manifestor potential, while the partial read with the partial

...humans are quickly overwhelmed by concurrency and find it much more difficult to reason about concurrent than sequential code. Even careful people miss possible interleavings...

- Herb Sutter & James Larus, Microsoft [SL05]



In the future applications will need to be **concurrent** to fully exploit CPU throughput gains [Sut05]

I conjecture that most multithreaded general purpose application are so full of concurrency bugs that - as multicore architectures become commonplace - these bugs will begin to show up as system failures.

- Edward A. Lee [Lee06]



Java Concurrency

- Java concurrency is built around the notion of multithreaded programs
 - sleep(), yield(), join() can affect the status of a thread
- Access to shared data supported primarily through synchronized methods and blocks
- Synchronization blocks can be used in combination with implicit monitor locks
 - Monitor methods: wait(), notify(), notifyAll()

Java Concurrency

- Explicit Lock: same semantics as the implicit monitor locks plus additional functionality such as timeouts during lock acquisition.
- Semaphore: Maintains a set of permits that restrict the number of threads accessing a resource.
- Latch: Allows threads to wait until other threads complete a set of operations.
- Barrier: A point at which threads from a set wait until all other threads reach that point.
- Exchanger: Allows two threads to exchange objects at a given synchronization point.

```
import java.io.FileOutputStream;
import java.io.FileNotFoundException;
import java.io.IOException;
/**
  * Created by IntelliJ IDEA.
  * User: amit rotstein I.D: 037698867
  * Date: Oct 17, 2003
  * Time: 1:02:13 PM
  * To change this template use Options | File Templates.
  * Modified by J.S. Bradbury (Feb. 2007)
  */
public class Bug implements Runnable {
   static int Num_Of_Seats_Sold = 0;
   int Maximum_Capacity, Num_of_tickets_issued;
   boolean StopSales = false;
   Thread threadArr [];
   public Bug (int size, int cushion) {
       Num of tickets issued = size;
       Maximum Capacity = Num of tickets issued - cushion;
        threadArr = new Thread [Num_of_tickets_issued];
        //starting the selling of the tickets:
        for (int i = 0;
        i < Num_of_tickets_issued; i ++) {</pre>
            System.out.println ("Creating seller thread # " + i);
            threadArr [i] = new Thread (this);
            threadArr [i].start ();
            // "make the sale !!!"
   /**
       * the selling post:
      * making the sale & checking if limit was reached ( and updating "StopSales" ),
   public void run () {
        if (StopSales == false) {
            System.out.println ("Ticket sold");
            Num_Of_Seats_Sold ++;
        synchronized (this) {
            if (Num_Of_Seats_Sold == Maximum_Capacity) {
                System.out.println ("Maximum capacity reach - no ticket sold");
                StopSales = true;
                // updating
            if (Num Of Seats Sold > Maximum Capacity) throw new RuntimeException ("bug
found - oversold seats!!");
       }}
```

```
import java.io.FileOutputStream;
import java.io.FileNotFoundException;
import java.io.IOException;
/**
  * Created by IntelliJ IDEA.
  * User: amit rotstein I.D: 037698867
  * Date: Oct 17, 2003
  * Time: 1:02:13 PM
  * To change this template use Options | File Templates.
  * Modified by J.S. Bradbury (Feb. 2007)
  */
public class Bug implements Runnable {
   static int Num_Of_Seats_Sold = 0;
   int Maximum_Capacity, Num_of_tickets_issued;
   boolean StopSales = false;
   Thread threadArr []:
   public Bug (int size, int cushion) {
       Num of tickets issued = size;
      Maximum Capacity = Num of tickets issued - cushion;
       threadArr = new Thread [Num of tickets issued]:
       //st /**
       for
                   * the selling post:
       i <
                   * making the sale & checking if limit was reached ( and updating "StopSales" ),
                   */
               public void run () {
                    if (StopSales == false) {
                         System.out.println ("Ticket sold");
                         Num Of Seats Sold ++;
   /**
      * the
                    synchronized (this) {
      * mak
      */
                         if (Num_Of_Seats_Sold == Maximum_Capacity) {
   public v
                              System.out.println ("Maximum capacity reach - no ticket sold");
       if (
                              StopSales = true;
                              // updating
                         }
       sync
                         if (Num_Of_Seats_Sold > Maximum_Capacity) throw new RuntimeException ("bug
           found - oversold seats!!"):
                    }}
found – over
       }}
```

A pattern is defined as something that "...describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice" [GHJV95].

An anti-pattern defines a recurring <u>bad</u> design solution [Mey06].

Related Work

- Prior to J2SE 5.0, Farchi, Nir, and Ur developed a bug
- pattern taxonomy for Java concurrency [FNU03].
 - The bug patterns are based on common mistakes programmers make when developing concurrent code in practice.
- This taxonomy has been extended to included the java.util.concurrent library [BCD06].

Related Work

- An anti-pattern catalog for Java multithreaded software has already been developed [HAT+04]
 - primarily contains design anti-patterns related to efficiency, quality and style...
- Our work focuses on the identification of anti-patterns based on bugs and includes anti-patterns related to the correctness of the program.
 - Therefore, we believe that both catalogs are complementary

Defining Concurrency Anti-Patterns

- pattern name: the anti-pattern name is based on the corresponding bug's name.
- problem: the problem describes the corresponding bug that is being addressed.
- context: the context in which the problem often occurs.
- solution: the solution describes general steps that can be taken to correct the anti-pattern.

Pattern name	Problem	Context	Solution
The	A pattern in which "two or more concurrent	Trying to use	Use synchronization to protect both
Interference	threads access a shared variable and when at	operations	write and read access to shared
anti-pattern."	least one access is a write, and the threads use	Involving shared	variables.
	no explicit mechanism to prevent the access	data without	
	from being simultaneous." [17]. The Interference	protecting the	
	bug pattern can also be generalized from	access to the	
	classic data race interference to include high level data races" which deal "with accesses	shared data.	
	to sets of fields which are related and should be		
The deadlock	accessed atomically" [18].	Tudos to protect	Domesia unnacessari
	"a situation where two or more processes are	Trying to protect	Remove unnecessary
anti-pattern.**	unable to proceed because each is waiting for one of the others to do something in a deadlock	access to operations	synchronization if possible. Remove unnecessary nested
	cycle For example, this occurs when a thread	Involving shared	synchronization if possible.
	holds a lock that another thread desires and	data.	Ensure nested synchronization
	vice-versa" [17].	Colle.	always occurs in the same order.
Starvation	This bug occurs when their is a failure to	Trying to use	When available use fairness
anti-pattern.+	"allocate CPU time to a thread. This may be	concurrency	parameter for concurrent
una pattorni.	due to scheduling policies" [5]. For example,	Independent of	mechanisms like semaphores. This
	an unfair lock acquisition scheme might cause a	scheduling	will ensure that no thread can
	thread never to be scheduled.	policies.	unfairly acquire semaphore permits.
Resource	"A group of threads together hold all of a finite	Trying to	One solution is to consider allocating
exhaustion	number of resources. One of them needs	optimize a	additional resources. Another
anti-pattern.+	additional resources but no other thread gives	concurrent	solution is to limit all threads' access
	one up"[5].	program by	to resources.
		limiting	
		resources.	
Incorrect	This pattern occurs when there is an incorrect	Trying to protect	Correct the count to the appropriate
count	initialization in a barrier for the number of	access to	value.
Initialization	parties that must be waiting for the barrier to	operations	
anti-pattern.+	trip, or an incorrect initialization of the number	Involving shared	
	of threads required to complete some action in	data.	
	a latch, or an incorrect initialization of the		
	number of permits in a semaphore.		

Pattern name	Problem	Context	Solution
Nonatomic	"an operation that "looks" like one operation in	Trying to	Use the volatile keyword when using
operations	one programmer model (e.g., the source code	perform an	64-bit variables.
assumed to	level of the programming language) but actually	operation on a	
be atomic	consists of several unprotected operations at	shared data	
anti-pattern.*	the lower abstraction levels" [8].	variable	
		atomically.	
Two-state	"Sometimes a sequence of operations needs to	Trying to protect	Combine the multiple critical regions
access bug	be protected but the programmer wrongly	access to	Into one critical region.
anti-pattern.*	assumes that separately protecting each	operations	
	operation is enough" [8].	Involving shared	
		data.	
Wrong lock or	"A code segment is protected by a lock but	Trying to protect	Identify all accesses to shared data
no lock bug	other threads do not obtain the same lock	access to	and use the same lock object to
anti-pattern.*	instance when executing. Either these other	operations	protect these critical regions. This
	threads do not obtain a lock at all or they obtain	Involving shared	may involve added a new lock or
	some lock other than the one used by the code	data.	replacing incorrect locks with the
Bauchte	segment" [8].	Today to	correct one.
Double-	"When an object is initialized, the thread local	Trying to	Use locks to synchronize all access
checked lock	copy of the objects field is initialized but not all	Initialize shared variables	to the object or use volatile. Do not
anti-pattern.*	object fields are necessarily written to the heap.		perform lazy initialization on shared
	This might cause the object to be partially	without using	objects.
The sleep()	initialized while its reference is not null* [8]. "The programmer assumes that a child thread	protection.	"The correct solution would be for
anti-pattern.*	should be faster than the parent thread in order	Trying to coordinate	the parent thread to use the join()
anu-pattern.	that its results be available to the parent thread	threads based	method to explicitly wait for the child
	when it decides to advance. Therefore, the	on assumptions	thread* [8].
	programmer sometimes adds an 'appropriate'	regarding	uneau [o].
	sleep() to the parent thread. However, the	thread timing.	
	parent thread may still be quicker in some	anous anning.	
	environment."[8].		
Missing or	This pattern generalizes the losing a notify bug	Trying to	In the case of a notify signal, "One
nonexistent	pattern to all signals. The losing a notify bug is	coordinate	way of avoiding this bug pattern is to
signals	defined as occurring "If a notify() is executed	threads based	repeatedly execute the notify()
anti-pattern.+	before its corresponding wait(), the notity() has	on assumptions	operation until a condition stating
	no effect and is "lost" the programmer	regarding	that the notify() was received
	implicitly assumes that the wait() operation will	thread timing.	occurs*[8]. Use concurrent
	occur before any of the corresponding notify()		mechanisms such as barriers and
	operations" [8]. Another example of this		join() to prevent thread timing
	problem can occur at a barrier. If an await()		Issues. Analogous solutions exist for
	from one thread never occurs then all of threads		other signals.
	at the barrier may be stuck waiting.		
Notify Instead	If a notify() is executed instead of notifyAli()	Trying to	Replace notify() with notifyAli().
of notify all	then threads with some of its corresponding	coordinate threads.	
anti-pattern.**	walt() calls will not be notified [16].		Ensure that every lock() acquisition
A "blocking" critical	"A thread is assumed to eventually return control but it never does" [8].	Using locks to try and protect	has a corresponding unlock().
section	control but it riever does [0].	access to	If it is possible to throw an exception
anti-pattern.*		operations	Inside a critical region the unlock()
and pattorn.		Involving shared	must be placed in a finally block.
		data.	The finally block will be executed
		- Control	regardless if the exception is thrown.
			regulations if the exception is thrown.

Pattern name	Problem	Context	Solution
Nonatomic	"an operation that "looks" like one operation in	Trying to	Use the volatile keyword when using
operations	one programmer model (e.g., the source code	perform an	64-bit variables.
assumed to	level of the programming language) but actually	operation on a	
be atomic	consists of several unprotected operations at	shared data	
anti-pattern.*	the lower abstraction levels" [8].	variable	
		atomically.	
Two-state	"Sometimes a sequence of operations needs to	Trying to protect	Combine the multiple critical regions
access bug	be protected but the programmer wrongly	access to	Into one critical region.
anti-pattern.*	assumes that separately protecting each	operations	
	operation is enough" [8].	involving shared	
Wrong lock or	"A code segment is protected by a lock but	Trying to protect	Identify all accesses to shared data
no lock bug	other threads do not obtain the same lock	access to	and use the same lock object to
anti-pattern.*	instance when executing. Either these other	operations	protect these critical regions. This
	threads do not obtain a lock at all or they obtain	Involving shared	may involve added a new lock or
	some lock other than the one used by the code	data.	replacing incorrect locks with the
Dauble	seament* (8)	Tuday tr	correct one
Double-	"When an object is initialized, the thread local	Trying to	Use locks to synchronize all access
checked lock anti-pattern.*	copy of the objects field is initialized but not all object fields are necessarily written to the heap.	Initialize shared variables	to the object or use volatile. Do not perform lazy initialization on shared
anu-patiern.			
	This might cause the object to be partially initialized while its reference is not null*[8].	without using protection.	objects.
The sleep()	"The programmer assumes that a child thread	Trying to	"The correct solution would be for
anti-pattern.*	should be faster than the parent thread in order	coordinate	the parent thread to use the join()
anu-pattern.	that its results be available to the parent thread	threads based	method to explicitly wait for the child
	when it decides to advance. Therefore, the	on assumptions	thread*[8].
	programmer sometimes adds an 'appropriate'	regarding	incas [o].
	sleep() to the parent thread. However, the	thread timing.	
	parent thread may still be quicker in some		
	environment."[8].		
Missing or	This pattern generalizes the losing a notify bug	Trying to	In the case of a notify signal, "One
nonexistent	pattern to all signals. The losing a notify bug is	coordinate	way of avoiding this bug pattern is to
signals	defined as occurring "If a notify() is executed	threads based	repeatedly execute the notify()
anti-pattern.+	before its corresponding wait(), the notity() has	on assumptions	operation until a condition stating
	no effect and is "lost" the programmer	regarding	that the notify() was received
	implicitly assumes that the wait() operation will	thread timing.	occurs*[8]. Use concurrent
	occur before any of the corresponding notify()		mechanisms such as barriers and
	operations" [8]. Another example of this		join() to prevent thread timing
	problem can occur at a barrier. If an await()		Issues. Analogous solutions exist for
	from one thread never occurs then all of threads		other signals.
	at the barrier may be stuck waiting.		
Notify Instead	If a notify() is executed instead of notifyAli()	Trying to	Replace notify() with notifyAll().
of notify all	then threads with some of its corresponding	coordinate	
anti-pattern.**	walt() calls will not be notified [16].	threads.	
A "blocking"	"A thread is assumed to eventually return	Using locks to	Ensure that every lock() acquisition
critical	control but it never does*[8].	try and protect	has a corresponding unlock().
section		access to	If it is possible to throw an exception
anti-pattern."		operations	Inside a critical region the unlock()
		Involving shared data.	must be placed in a finally block. The finally block will be executed
		udid.	regardless if the exception is thrown.
1		l	regardless if the exception is thrown.

Example: Wrong lock or no lock anti-pattern

Problem:

- A code segment is protected by a lock but other threads do not obtain the same lock instance when executing. Either these other threads do not obtain a lock at all or they obtain some lock other than the one used by the code segment.

 [FNU03]
- Context: Trying to protect access to operations involving shared data.
- Solution: Identify all accesses to shared data and use the same lock object to protect these critical regions.
 This may involve added a new lock or replacing incorrect locks with the correct one.

Example: Wrong lock or no lock anti-pattern

```
Object lock1 = new Object();
public void m1 () {
  <statement n1>
 //critical region
  <statement c1>
 synchronized (lock1) {
    <statement c2>
  <statement c3>
  <statement n2>
```

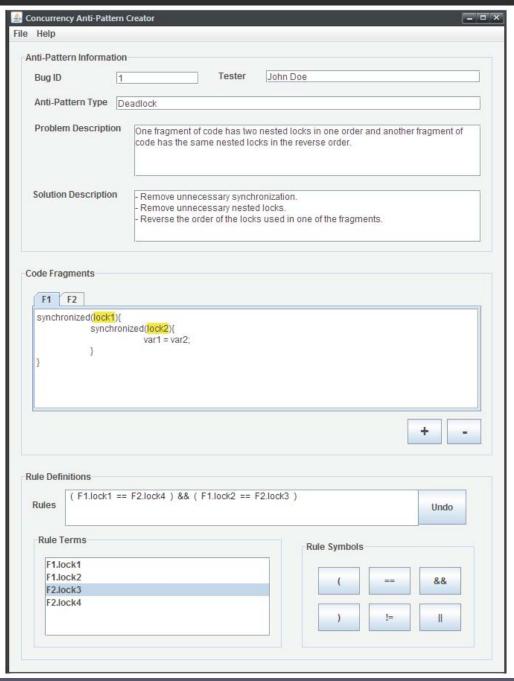
```
Object lock1 = new Object();
public void m1 () {
  <statement n1>
 synchronized (lock1) {
    //critical region
    <statement c1>
    <statement c2>
    <statement c3>
 <statement n2>
```

Detecting Concurrency Anti-Patterns

- The main motivation for our work on concurrency anti-patterns has been the automatic detection (and correction)
- We have developed a combined approach to static analysis and testing that uses our concurrency antipatterns to focus the testing effort

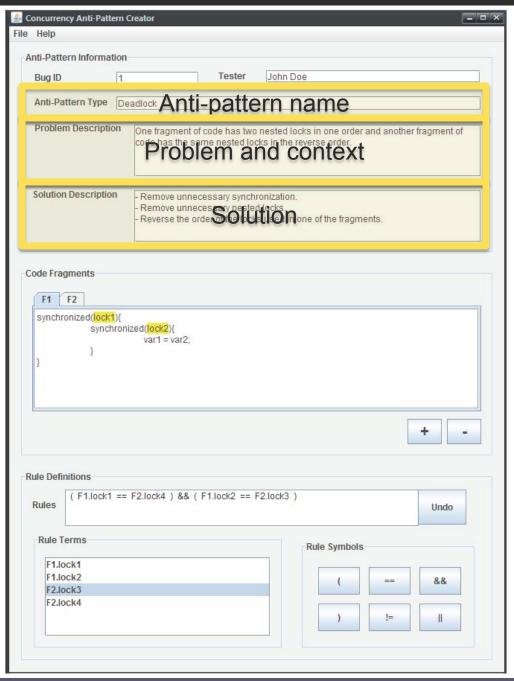
Concurrency Anti-Pattern Creator

 A tool for the creation and storage of concurrency antipatterns



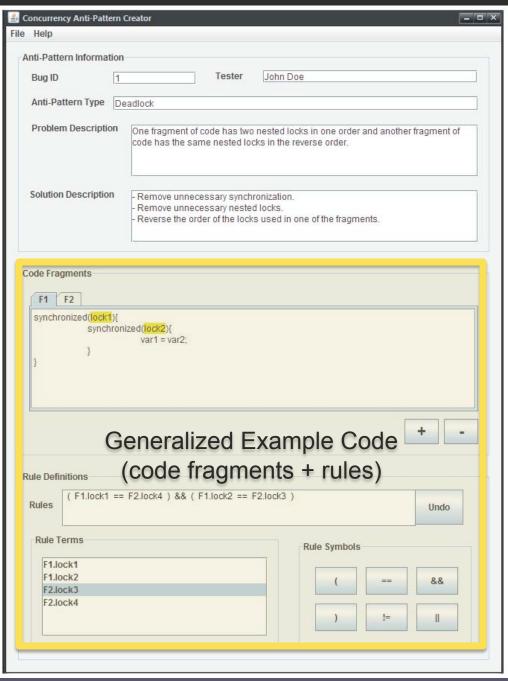
Concurrency Anti-Pattern Creator

 A tool for the creation and storage of concurrency antipatterns

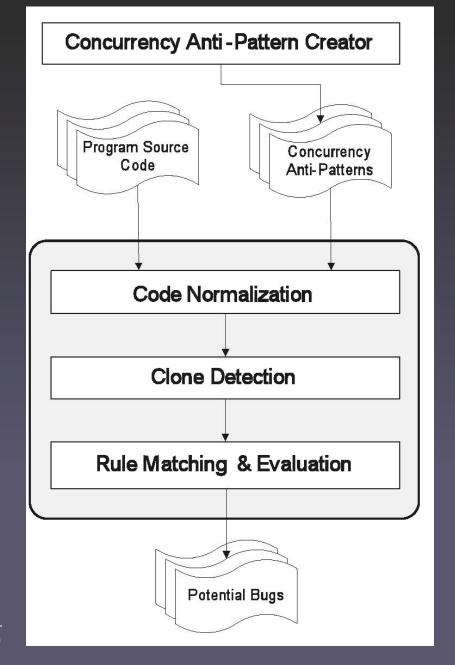


Concurrency Anti-Pattern Creator

 A tool for the creation and storage of concurrency antipatterns

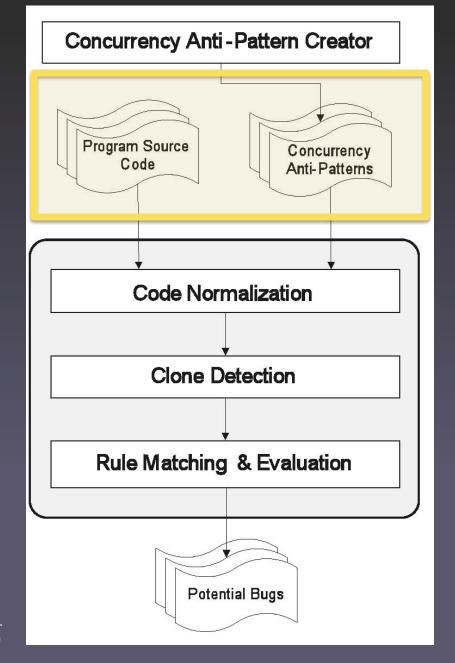


- A tool for the detecting concurrency antipatterns in Java programs
- Built using the pattern matching algorithm in the ConQAT* clone detection tool



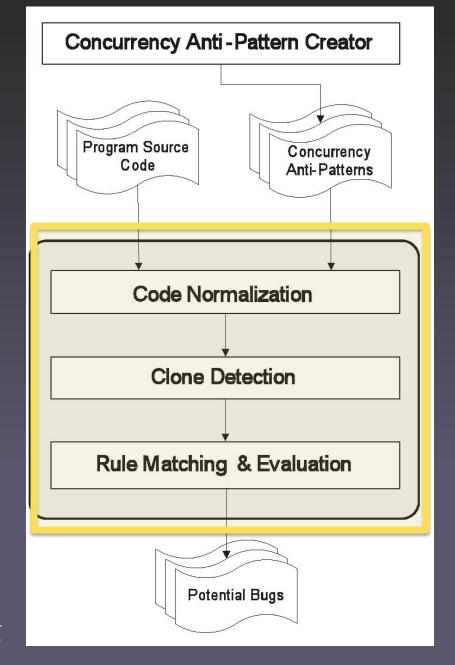
* ConQAT website (http://congat.cs.tum.edu/)

- A tool for the detecting concurrency antipatterns in Java programs
- Built using the pattern matching algorithm in the ConQAT* clone detection tool



* ConQAT website (http://congat.cs.tum.edu/)

- A tool for the detecting concurrency antipatterns in Java programs
- Built using the pattern matching algorithm in the ConQAT* clone detection tool



* ConQAT website (http://congat.cs.tum.edu/)

- A tool for the detecting concurrency antipatterns in Java programs
- Built using the pattern matching algorithm in the ConQAT* clone detection tool

* ConQAT website (http://conqat.cs.tum.edu/)

Concurrency Anti-Pattern Creator Program Source Concurrency Code Anti-Patterns **Code Normalization** Clone Detection Rule Matching & Evaluation Potential Bugs

Summary

- We have presented a catalog of 13 programming antipatterns for concurrent Java that are comprehensive with
- respect to:
 - the Java concurrency features
 - an existing concurrency bug pattern taxonomy
- Catalog available at:

http://svilab.science.uoit.ca/concurr-catalog/

 One important contribution of this work is that the catalog provides solutions – previous work has focused on enumerating different kinds of concurrency

Future Work

- We plan to conduct additional research on the benefits of the catalog with respect to static analysis and testing.
- We are interested in combining our work with more highlevel concurrency design patterns [GHJV95], [Lea00] and other concurency anti-patterns [HAT+04].

Defining a Catalog of Programming Anti-Patterns for Concurrent Java

Jeremy S. Bradbury, Kevin Jalbert

Software Quality Research Group
Faculty of Science (Computer Science)
University of Ontario Institute of Technology
Oshawa • Ontario • Canada
jeremy.bradbury@uoit.ca, kevin.jalbert@mycampus.uoit.ca

SPAQu'09 • October 25, 2009



