

The Relative Power of Synchronization Primitives


BROWN
Maurice Herlihy
CS176
Fall 2003

Wait-Free Implementation

- Every method call completes in finite number of steps
- Implies no mutual exclusion



Wait-Free Constructions

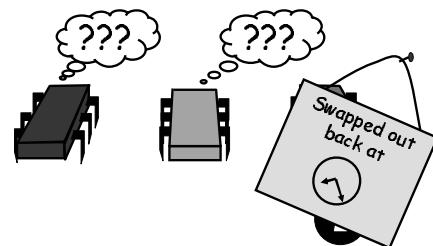
- Wait-free atomic registers
 - From safe registers
- Two-threaded FIFO queue
 - From atomic registers
 - And indirectly from safe registers

Rationale

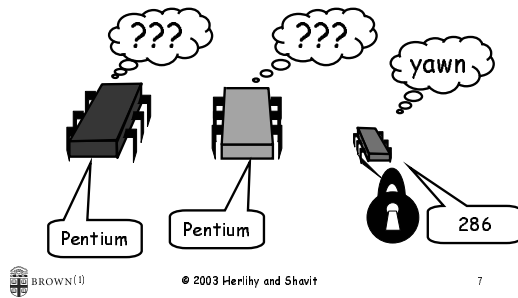
- We wanted atomic registers to implement mutual exclusion
- So we couldn't use mutual exclusion to implement atomic registers
- But wait, there's more!

Why is Mutual Exclusion so wrong?

Asynchronous Interrupts



Heterogeneous processors

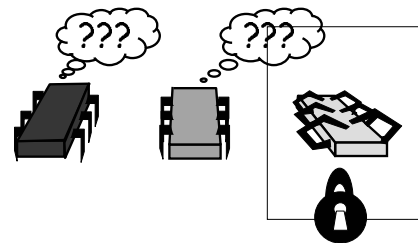


BROWN(1)

© 2003 Herlihy and Shavit

7

Fault-tolerance



BROWN(2)

© 2003 Herlihy and Shavit

8

Basic Questions

- Wait-Free Synchronization might be a good idea in principle
- But how do you do it
 - Systematically?
 - Correctly?
 - Efficiently?

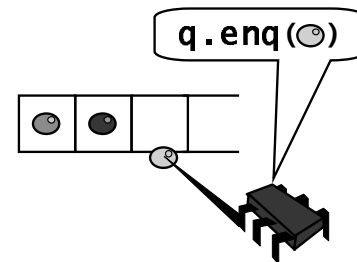


BROWN

© 2003 Herlihy and Shavit

9

FIFO Queue: Enqueue Method

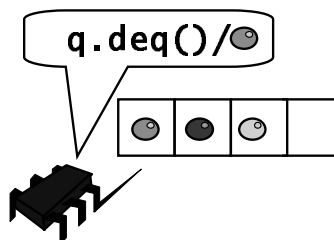


BROWN

© 2003 Herlihy and Shavit

10

FIFO Queue: Dequeue Method



BROWN

© 2003 Herlihy and Shavit

11

Two-Thread Wait-Free Queue

```
public class LockFreeQueue {
    int head = 0, tail = 0;
    Item[QSIZE] items;
    public void enq(Item x) {
        while (tail - head == QSIZE) {};
        items[tail % QSIZE] = x; tail++;
    }
    public Item deq() {
        while (tail == head) {}
        Item item = items[head % QSIZE];
        head++; return item;
    }
}
```



BROWN

© 2003 Herlihy and Shavit

12

Two-Thread Wait-Free Queue

```
public class LockFreeQueue {
    int head = 0, tail = 0;
    Item[QSIZE] items;
    public void enq(Item x) {
        while (tail-head == QSIZE) {};
        items[tail % QSIZE] = x; tail++;
    }
    public Item deq() { Put object in queue
        while (tail == head) {}
        Item item = items[head % QSIZE];
        head++; return item;
    }
}
```



BROWN

© 2003 Herlihy and Shavit

13

Two-Thread Wait-Free Queue

```
public class LockFreeQueue {
    int head = 0, tail = 0;
    Item[QSIZE] items;
    public void enq(Item x) {
        while (tail-head == QSIZE) {};
        items[tail % QSIZE] = x; tail++;
    }
    public Item deq() { Increment tail
        while (tail == head) {} counter
        Item item = items[head % QSIZE];
        head++; return item;
    }
}
```

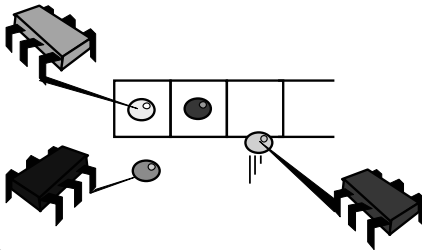


BROWN

© 2003 Herlihy and Shavit

14

What About Multiple Dequeuers?

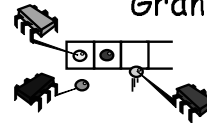


BROWN

© 2003 Herlihy and Shavit

15

Grand Challenge



- Implement a FIFO queue **Only new aspect**
 - Wait-free
 - Linearizable
 - From atomic read-write registers
 - **Multiple dequeuers**



BROWN

© 2003 Herlihy and Shavit

16

Consensus

- While you are ruminating on the grand challenge...
- We will give you another puzzle
 - Consensus
 - Pretty important ...

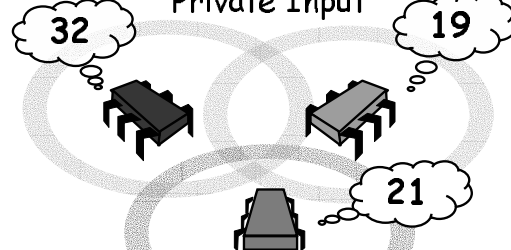


BROWN

© 2003 Herlihy and Shavit

17

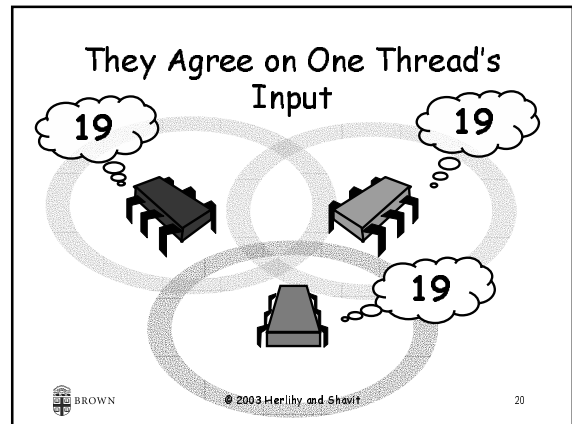
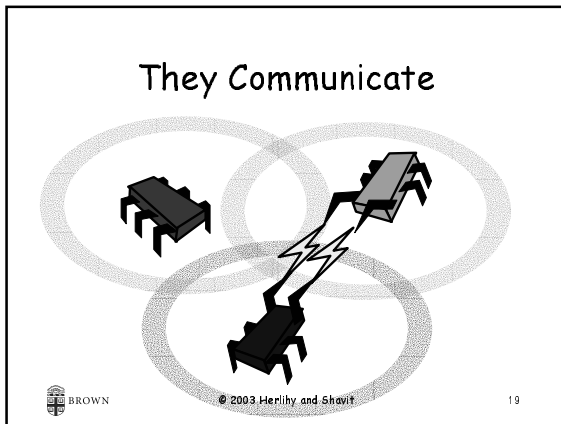
Consensus: Each Thread has a Private Input



BROWN

© 2003 Herlihy and Shavit

18



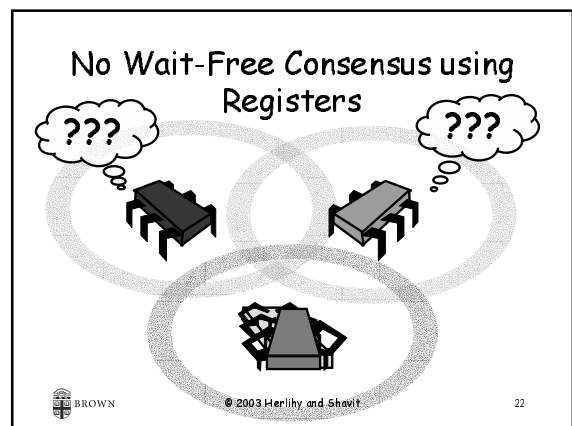
Formally: Consensus

Consistent: all threads decide the same value

Valid: the common decision value is some thread's input

Wait-free: each thread decides after a finite number of steps

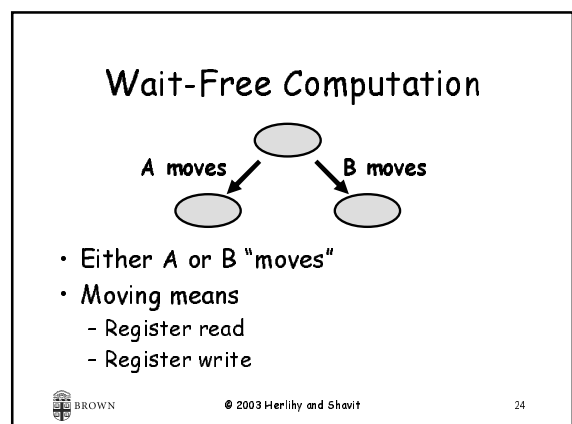
The slide includes a BROWN logo, copyright notice, and slide number 21.

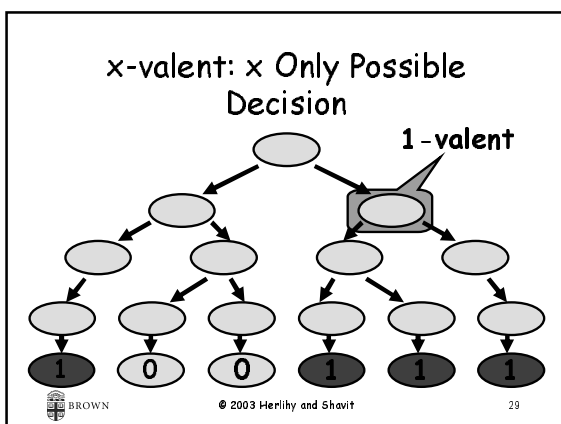
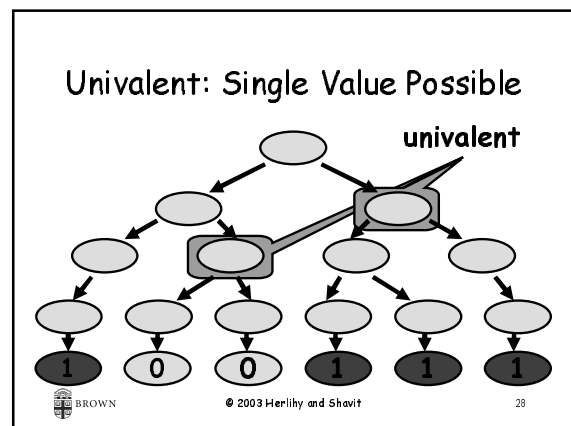
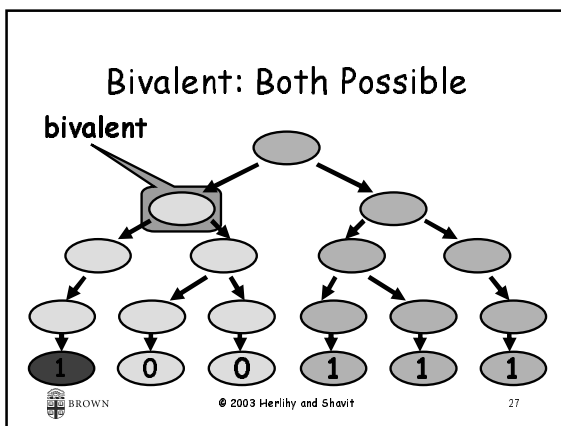
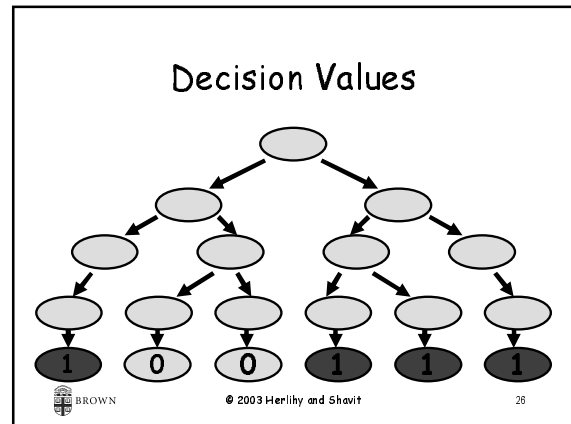
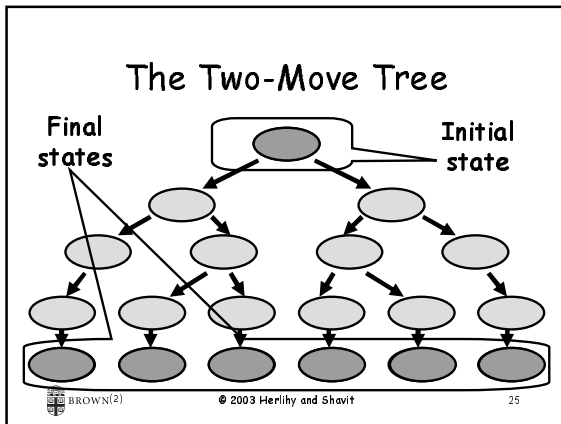


Proof Strategy

- Assume otherwise
- Reason about the properties of any such protocol
- Derive a contradiction
- Quod Erat Demonstrandum

The slide includes a BROWN logo, copyright notice, and slide number 23.





Summary

- Wait-free computation is a tree
- Bivalent system states
 - Outcome not fixed
- Univalent states
 - Outcome is fixed
 - May not be "known" yet
- 1-Valent and 0-Valent states

BROWN

© 2003 Herlihy and Shavit

30

Claim

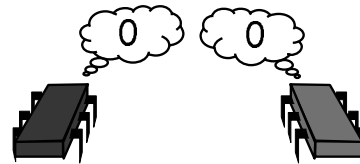
- Some initial state is bivalent
- Outcome depends on
 - Chance
 - Whim of the scheduler
- Multiprocessor gods do play dice ...



© 2003 Herlihy and Shavit

31

Univalent Initial State



All executions must decide 0



© 2003 Herlihy and Shavit

32 (1)

Univalent Initial State



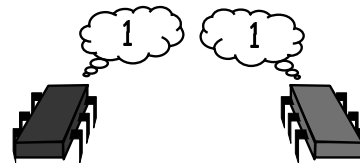
Including this solo execution by A



© 2003 Herlihy and Shavit

33 (1)

Univalent Initial State



All executions must decide 1



© 2003 Herlihy and Shavit

34 (1)

Univalent Initial State



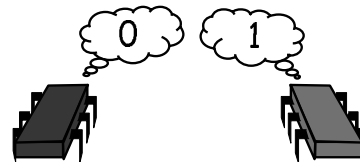
Including this solo execution by B



© 2003 Herlihy and Shavit

35 (1)

Univalent Initial State?



Imagine all executions deciding alike...

(cue in the John Lennon...)



© 2003 Herlihy and Shavit

36

Univalent Initial State?

Including this solo execution by A
which we know decides 0

BROWN(2) © 2003 Herlihy and Shavit 37

Univalent Initial State?

Including this solo execution by B
which we know decides 1

BROWN(2) © 2003 Herlihy and Shavit 38

Uh-Oh

- Solo execution by A must decide 0
- Solo execution by B must decide 1

BROWN © 2003 Herlihy and Shavit 39

Uh-Oh

How univalent is that?
(QED)

- Solo execution by A must decide 0
- Solo execution by B must decide 1

BROWN © 2003 Herlihy and Shavit 40

Critical States

critical

0-valent 1-valent

BROWN(3) © 2003 Herlihy and Shavit 41(3)

From a Critical State

c

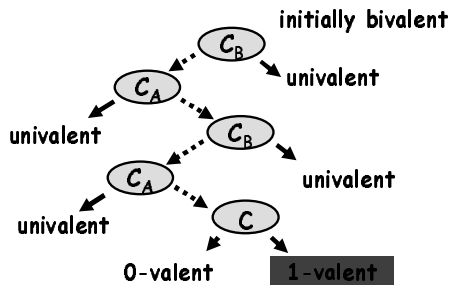
0-valent 1-valent

If A goes first, protocol decides 0

If B goes first, protocol decides 1

BROWN © 2003 Herlihy and Shavit 42

Reaching Critical State



Critical States

- Starting from a bivalent initial state
- The protocol can reach a critical state
 - Otherwise we could stay bivalent forever
 - And the protocol is not wait-free

Model Dependency

- So far, memory-independent!
- True for
 - Registers
 - Message-passing
 - Carrier pigeons
 - Any kind of asynchronous computation

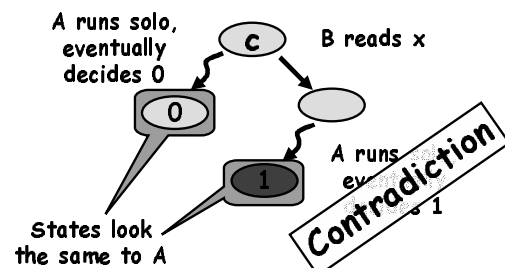
What are the Threads Doing?

- Reads and/or writes
- To same/different registers

Possible Interactions

	A reads x	A reads y	A writes x	A writes y
x.read()	?	?	?	?
y.read()	?	?	?	?
x.write()	?	?	?	?
y.write()	?	?	?	?

Reading Registers



Possible Interactions

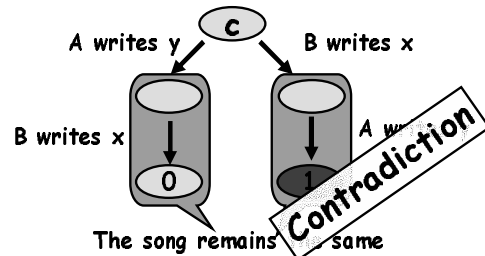
	x.read()	y.read()	x.write()	y.write()
x.read()	no	no	no	no
y.read()	no	no	no	no
x.write()	no	no	?	?
y.write()	no	no	?	?



© 2003 Herlihy and Shavit

49

Writing Distinct Registers



© 2003 Herlihy and Shavit

50

Possible Interactions

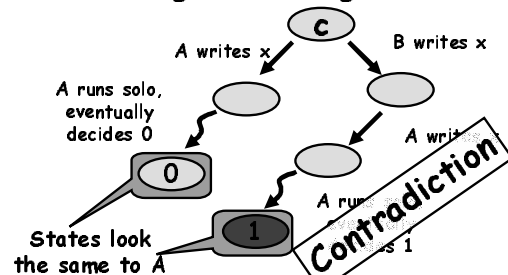
	x.read()	y.read()	x.write()	y.write()
x.read()	no	no	no	no
y.read()	no	no	no	no
x.write()	no	no	?	no
y.write()	no	no	no	?



© 2003 Herlihy and Shavit

51

Writing Same Registers



© 2003 Herlihy and Shavit

52

That's All, Folks!

	x.read()	y.read()	x.write()	y.write()
x.read()	no	no	no	no
y.read()	no	no	no	no
x.write()	no	no	no	no
y.write()	no	no	no	no



© 2003 Herlihy and Shavit

53

QED

Atomic Registers Can't Do Consensus

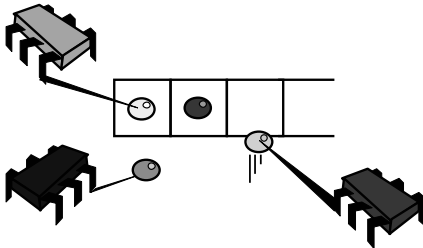
- If protocol exists
 - It has a bivalent initial state
 - Leading to a critical state
- What's up with the critical state?
 - Case analysis for each pair of methods



© 2003 Herlihy and Shavit

54

What Does Consensus have to do with Concurrent Objects?



Java Jargon Watch

- Define *Consensus* as an abstract class
- We implement some methods
- Leave you to do the rest ...

Consensus Object

```
abstract class Consensus {
    private Object[] proposed =
        new Object[N];

    public void propose(Object value) {
        proposed[Thread.myIndex()] = value;
    }

    abstract public Object decide();
}
```

Consensus Object

```
abstract class Consensus {
    private Object[] proposed =
        new Object[N];

    public void propose(Object value) {
        proposed[Thread.myIndex()] = value;
    }

    abstract public Obj
```

Each thread's proposed value

Consensus Object

Propose a value

```
sensus {
    private Object[] proposed =
        new Object[N];

    public void propose(Object value) {
        proposed[Thread.myIndex()] = value;
    }

    abstract public Object decide();
}
```

Consensus Object

Decide a value: abstract method means subclass does the heavy lifting (real work)

```
public void propose(Object value) {
    proposed[Thread.myIndex()] = value;
}

abstract public Object decide();
}
```

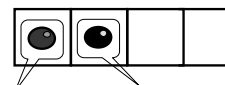
Can FIFO Queue Implement Consensus?



FIFO Consensus



announce array



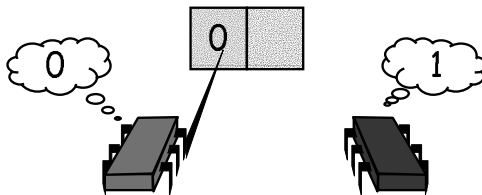
FIFO Queue
with red and
black balls



© 2003 Herlihy and Shavit

62

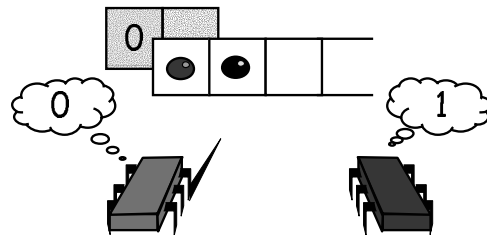
Protocol: Write Value to Array



© 2003 Herlihy and Shavit

63

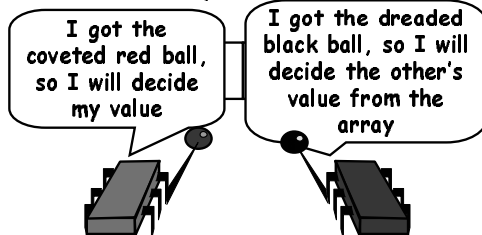
Protocol: Take Next Item from Queue



© 2003 Herlihy and Shavit

64

Protocol: Take Next Item from Queue



© 2003 Herlihy and Shavit

65

Consensus Using FIFO Queue

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    public QueueConsensus() {
        queue = new Queue();
        queue.enq(Ball.RED);
        queue.enq(Ball.BLACK);
    }
    ...
}
```

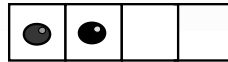


© 2003 Herlihy and Shavit

66

Initialize Queue

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    public QueueConsensus() {
        this.queue = new Queue();
        this.queue.enq(Ball.RED);
        this.queue.enq(Ball.BLACK);
    }
    ...
}
```



© 2003 Herlihy and Shavit

67

Who Won?

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    ...
    public decide() {
        Ball ball = this.queue.deq();
        if (ball == Ball.RED)
            return proposed[i];
        else
            return proposed[j];
    }
}
```



© 2003 Herlihy and Shavit

68

Who Won?

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    ...
    public decide() {
        Ball ball = this.queue.deq();
        if (ball == Ball.RED)
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Race to dequeue
first queue item



© 2003 Herlihy and Shavit

69

Who Won?

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    ...
    public decide() {
        Ball ball = this.queue.deq();
        if (ball == Ball.RED)
            return proposed[i];
        else
            return proposed[j];
    }
}
```

I win if I was first



© 2003 Herlihy and Shavit

70

Who Won?

```
public class QueueConsensus
    extends Consensus {
    private Queue queue;
    ...
    public decide() {
        Ball ball = this.queue.deq();
        if (ball == Ball.RED)
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Other thread wins if
I was second



© 2003 Herlihy and Shavit

71

Why does this Work?

- If one thread gets the red ball
- Then the other gets the black ball
- Winner decides her own value
- Loser can find winner's value in array
 - Because threads write array
 - Before dequeuing from queue



© 2003 Herlihy and Shavit

72

Theorem

- We can solve 2-thread consensus using only
 - A two-dequeuer queue, and
 - Some atomic registers



© 2003 Herlihy and Shavit

73

Implications

- Given
 - A consensus protocol from queue and registers
- Assume there exists
 - A queue implementation from atomic registers
- Substitution yields:
 - A wait-free consensus protocol from atomic registers

contradiction



© 2003 Herlihy and Shavit

74(1)

Corollary

- It is impossible to implement
 - a two-dequeuer wait-free FIFO queue
 - from read/write memory.



© 2003 Herlihy and Shavit

75

Consensus Numbers

- An object X has consensus number n
 - If it can be used to solve n -thread consensus
 - Taking any number of instances of X
 - together with atomic read/write registers
 - and implement n -thread consensus
 - But not $(n+1)$ -thread consensus



© 2003 Herlihy and Shavit

76

Consensus Numbers

- Theorem
 - Atomic read/write registers have consensus number 1
- Theorem
 - Multi-dequeuer FIFO queues have consensus number at least 2



© 2003 Herlihy and Shavit

77

Consensus Numbers Measure Synchronization Power

- Theorem
 - If you can implement X from Y
 - And X has consensus number c
 - Then Y has consensus number at least c



© 2003 Herlihy and Shavit

78

Synchronization Speed Limit

- Conversely
 - If X has consensus number n
 - And Y has consensus number m
 - Then there is no way to construct a wait-free implementation of X by Y
- This theorem will be very useful
 - Unforeseen practical implications!

Theoretical
Caveat: Certain
weird exceptions
exist



© 2003 Herlihy and Shavit

79

Earlier Grand Challenge

- Snapshot means
 - Write any array element
 - Read multiple array elements atomically
- What about
 - Write multiple array elements atomically
 - Scan any array elements
- Call this problem multiple assignment



© 2003 Herlihy and Shavit

80

Multiple Assignment Theorem

- Atomic registers cannot implement multiple assignment
- Weird or what?
 - Single write/multiple read OK
 - Multi write/multiple read impossible



© 2003 Herlihy and Shavit

81 (1)

Proof Strategy

- If we can write to 2/3 array elements
 - We can solve 2-consensus
 - Impossible with atomic registers
- Therefore
 - Cannot implement multiple assignment with atomic registers



© 2003 Herlihy and Shavit

82 (1)

Proof Strategy

- Take a 3-element array
 - A writes atomically to slots 0 and 1
 - B writes atomically to slots 1 and 2
 - Any thread can scan any set of locations



© 2003 Herlihy and Shavit

83 (1)

Double Assignment Interface

```
interface Assign2 {
    public void assign(int i1, int v1,
                     int i2, int v2);
    public int read(int i);
}
```



© 2003 Herlihy and Shavit

84 (4)

Double Assignment Interface

```
interface Assign2 {
    public void assign(int i1, int v1,
                      int i2, int v2);
    public int read(int i);
}
```

Atomically assign
value[i₁] = v₁
value[i₂] = v₂



BROWN(4)

© 2003 Herlihy and Shavit

85(4)

Double Assignment Interface

```
interface Assign2 {
    public void assign(int i1, int v1,
                      int i2, int v2);
    public int read(int i);
}
```

Return i-th value

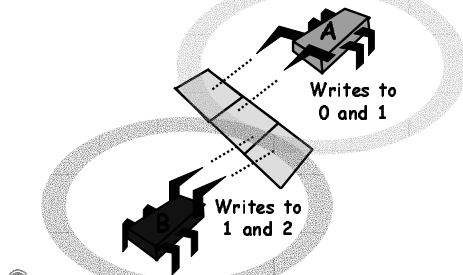


BROWN(4)

© 2003 Herlihy and Shavit

86(4)

Initially

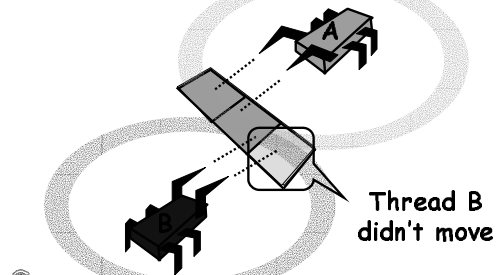


BROWN

© 2003 Herlihy and Shavit

87

Thread A wins if

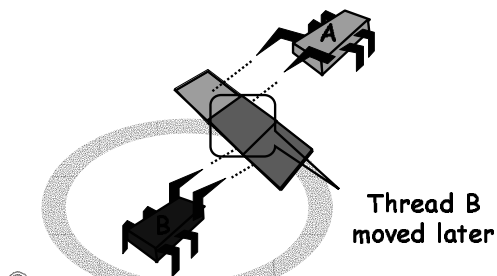


BROWN

© 2003 Herlihy and Shavit

88 (1)

Thread A wins if

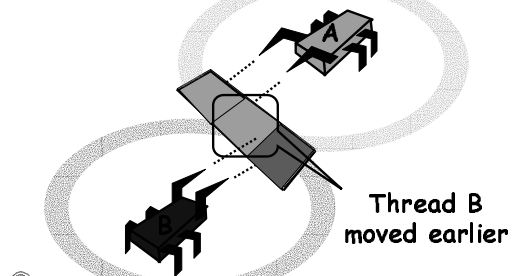


BROWN

© 2003 Herlihy and Shavit

89 (1)

Thread A loses if



BROWN

© 2003 Herlihy and Shavit

90 (1)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```



BROWN[4]

© 2003 Herlihy and Shavit

91(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Three slots
initialized to
EMPTY



BROWN[4]

© 2003 Herlihy and Shavit

92(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Assign id 0 to 0,1
(or id 1 to 1,2)



BROWN[4]

© 2003 Herlihy and Shavit

93(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Read the register my
thread didn't assign



BROWN[4]

© 2003 Herlihy and Shavit

94(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Other thread didn't
move, so I win



BROWN[4]

© 2003 Herlihy and Shavit

95(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Other thread moved
later, so I win



BROWN[4]

© 2003 Herlihy and Shavit

96(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

OK, I win.



BROWN(4)

© 2003 Herlihy and Shavit

97(4)

Multi-Consensus Code

```
class MultiConsensus extends Consensus{
    Assign2 a = new Assign2(3, EMPTY);
    public Object decide() {
        a.assign(i, i, i+1, i);
        int other = a.read((j+1) % 3);
        if (other==EMPTY||other==a.read(j))
            return proposed[i];
        else
            return proposed[j];
    }
}
```

Other thread moved first, so I lose



BROWN(4)

© 2003 Herlihy and Shavit

98(4)

Summary

- If a thread can assign atomically to 2 out of 3 array locations
- Then we can solve 2-consensus
- Therefore
 - No wait-free multi-assignment
 - From read/write registers



BROWN

© 2003 Herlihy and Shavit

99

Read-Modify-Write Objects

- Method call
 - Returns object's prior value x
 - Replaces x with mumble(x)



BROWN

© 2003 Herlihy and Shavit

100

Read-Modify-Write

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
    getAndMumble() {
        int prior = this.value;
        this.value = mumble(this.value);
        return prior;
    }
}
```



BROWN(1)

© 2003 Herlihy and Shavit

101

Read-Modify-Write

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
    getAndMumble() {
        int prior = this.value;
        this.value = mumble(this.value);
        return prior;
    }
}
```



BROWN(1)

© 2003 Herlihy and Shavit

102

Read-Modify-Write

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
    getAndMumble() {
        int prior = this.value;
        this.value = mumble(this.value);
        return prior;
    }
}
```

Return prior value



BROWN()

© 2003 Herlihy and Shavit

103

Read-Modify-Write

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
    getAndMumble() {
        int prior = this.value;
        this.value = mumble(this.value);
        return prior;
    }
}
```

Apply function to current value



BROWN()

© 2003 Herlihy and Shavit

104

RMW Everywhere!

- Most synchronization instructions
 - are RMW methods
- The rest
 - Can be trivially transformed into RMW methods



BROWN

© 2003 Herlihy and Shavit

105

Example: Read

```
public abstract class RMWRegister {
    private int value;

    public void synchronized read() {
        int prior = this.value;
        this.value = this.value;
        return prior;
    }
}
```



BROWN()

© 2003 Herlihy and Shavit

106

Example: Read

```
public abstract class RMW {
    private int value;

    public void synchronized read() {
        int prior = this.value;
        this.value = this.value;
        return prior;
    }
}
```

Apply $f(v)=v$, the identity function



BROWN()

© 2003 Herlihy and Shavit

107

Example: getAndSet

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
    getAndSet(int v) {
        int prior = this.value;
        this.value = v;
        return prior;
    }
    ...
}
```



BROWN()

© 2003 Herlihy and Shavit

108

Example: getAndSet

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
        getAndSet(int v) {
        int prior = this.value;
        this.value = v;
        return prior;
    }
    ...
}
```

$F(x)=v$ is constant function



BROWN[1]

© 2003 Herlihy and Shavit

109

getAndIncrement

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
        getAndIncrement() {
        int prior = this.value;
        this.value = this.value + 1;
        return prior;
    }
    ...
}
```



BROWN[1]

© 2003 Herlihy and Shavit

110

getAndIncrement

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
        getAndIncrement() {
        int prior = this.value;
        this.value = this.value + 1;
        return prior;
    }
    ...
}
```

$F(x) = x+1$



BROWN[1]

© 2003 Herlihy and Shavit

111

getAndAdd

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
        getAndAdd(int a) {
        int prior = this.value;
        this.value = this.value + a;
        return prior;
    }
    ...
}
```



BROWN[1]

© 2003 Herlihy and Shavit

112

Example: getAndAdd

```
public abstract class RMWRegister {
    private int value;

    public void synchronized
        getAndAdd(int a) {
        int prior = this.value;
        this.value = this.value + a;
        return prior;
    }
    ...
}
```

$F(x) = x+a$



BROWN[1]

© 2003 Herlihy and Shavit

113

compareAndSet

```
public abstract class RMWRegister {
    private int value;
    public boolean synchronized
        compareAndSet(int expected,
                      int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    }
    ...
}
```



BROWN[1]

© 2003 Herlihy and Shavit

114

compareAndSet

```
public abstract class RMWRegister {
    private int value;
    public boolean synchronized
        compareAndSet(int expected,
                      int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
```

If value is expected, ...



BROWN[]

© 2003 Herlihy and Shavit

115

compareAndSet

```
public abstract class RMWRegister {
    private int value;
    public boolean synchronized
        compareAndSet(int expected,
                      int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
```

... replace it



BROWN[]

© 2003 Herlihy and Shavit

116

compareAndSet

```
public abstract class RMWRegister {
    private int value;
    public boolean synchronized
        compareAndSet(int expected,
                      int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
```

Report success



BROWN[]

© 2003 Herlihy and Shavit

117

compareAndSet

```
public abstract class RMWRegister {
    private int value;
    public boolean synchronized
        compareAndSet(int expected,
                      int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
```

Otherwise report failure



BROWN[]

© 2003 Herlihy and Shavit

118

Definition

- A RMW method
 - With function mumble(x)
 - is non-trivial if there exists a value v
 - Such that $v \neq \text{mumble}(v)$
- Read() is trivial
- getAndIncrement() is non-trivial



BROWN

© 2003 Herlihy and Shavit

119

Par Example

- Identity(x)=x
 - is trivial
- getAndIncrement(x) = x+1
 - is non-trivial



BROWN

© 2003 Herlihy and Shavit

120

Theorem

- Any non-trivial RMW object has consensus number at least 2
- No wait-free implementation of RMW registers from atomic registers
- Hardware RMW instructions not just a convenience



© 2003 Herlihy and Shavit

121

Reminder

- Subclasses of Consensus have
 - **propose(x)** method
 - which just stores x into this.announce[i]
 - Built-in method
 - **decide()** method
 - which determines winning value
 - Customized, class-specific method



© 2003 Herlihy and Shavit

122

Proof

```
public class RMWConsensus
    implements Consensus {
    private RMWRegister r = v;

    public Object decide() {
        if (r.getAndMumble() == v)
            return this.announce[i];
        else
            return this.announce[j];
    }
}
```



© 2003 Herlihy and Shavit

123

Proof

```
public class RMWConsensus
    implements Consensus {
    private RMWRegister r = v;

    public Object decide() {
        if (r.getAndMumble() == v)
            return this.announce[i];
        else
            return this.announce[j];
    }
}
```

Initialized to v



© 2003 Herlihy and Shavit

124

Proof

```
public class RMWConsensus
    implements Consensus {
    private RMWRegister r = v;

    public Object decide() {
        if (r.getAndMumble() == v)
            return this.announce[i];
        else
            return this.announce[j];
    }
}
```

Am I first?



© 2003 Herlihy and Shavit

125

Proof

```
public class RMWConsensus
    implements Consensus {
    private RMWRegister r = v;

    public Object decide() {
        if (r.getAndMumble() == v)
            return this.announce[i];
        else
            return this.announce[j];
    }
}
```

Yes, return my input



© 2003 Herlihy and Shavit

126

Proof

```
public class RMWConsensus
    implements Consensus {
    private RMWRegister r;

    public Object decide() {
        if (r.getAndMumble() == v)
            return this.announce[i];
        else
            return this.announce[j];
    }
}
```

No, return other's input



BROWN(4)

© 2003 Herlihy and Shavit

127

Proof

- We have displayed
 - A two-thread consensus protocol
 - Using any non-trivial RMW object



BROWN

© 2003 Herlihy and Shavit

128

Interfering RMW

- Let F be a set of functions such that for all f_i and f_j , either
 - Commute: $f_i(f_j(v)) = f_j(f_i(v))$
 - Overwrite: $f_i(f_j(v)) = f_i(v)$
- Claim: Any such set of RMW objects has consensus number exactly 2



BROWN

© 2003 Herlihy and Shavit

129

Examples

- Test-and-Set $f(v)=1$
Overwrite $f_i(f_j(v)) = f_i(v)$
- Swap $f(v,x)=x$
Overwrite $f_i(f_j(v)) = f_i(v)$
- Fetch-and-inc $f(v)=v+1$
Commutate $f_i(f_j(v)) = f_j(f_i(v))$

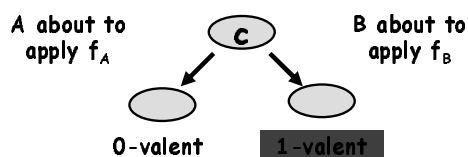


BROWN

© 2003 Herlihy and Shavit

130

Meanwhile Back at the Critical State

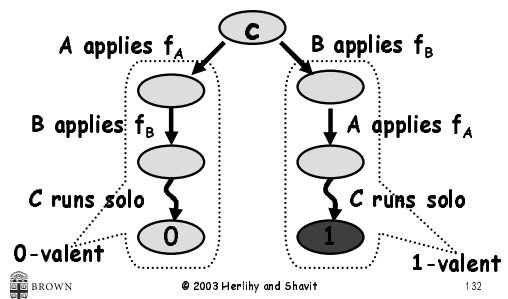


BROWN

© 2003 Herlihy and Shavit

131

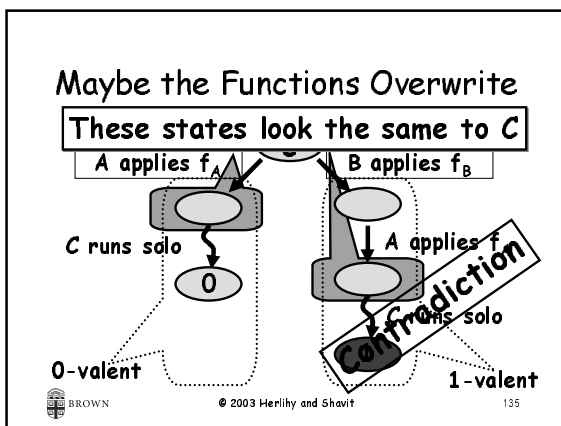
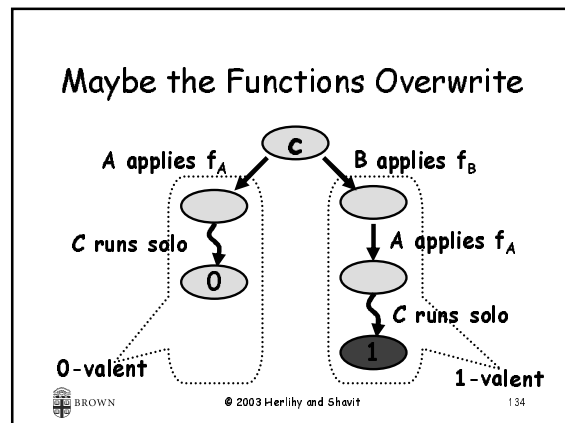
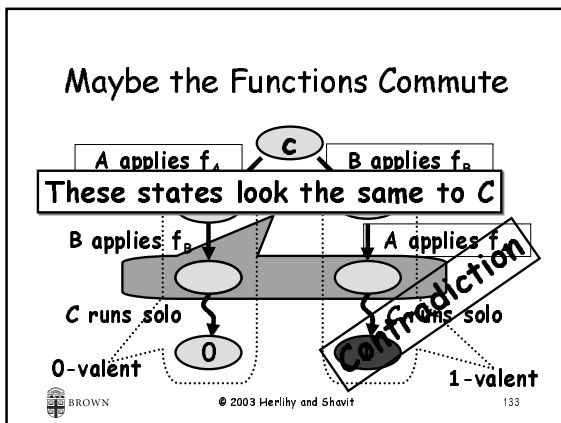
Maybe the Functions Commute



BROWN

© 2003 Herlihy and Shavit

132



- ### Impact
- Many early machines provided these "weak" RMW instructions
 - Test-and-set (IBM 360)
 - Fetch-and-add (NYU Ultracomputer)
 - Swap (Original SPARCs)
 - We now understand their limitations
 - But why do we want consensus anyway?
- BROWN © 2003 Herlihy and Shavit 136

compareAndSet

```

public abstract class RMWRegister {
    private int value;
    public boolean synchronized
    compareAndSet(int expected,
                  int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
  
```

BROWN © 2003 Herlihy and Shavit 137

compareAndSet

```

public abstract class RMWRegister {
    private int value;
    public boolean synchronized
    compareAndSet(int expected,
                  int update) {
        int prior = this.value;
        if (this.value==expected) {
            this.value = update; return true;
        }
        return false;
    } ... }
  
```

replace value if expected, ...

BROWN © 2003 Herlihy and Shavit 138

compareAndSet Has ∞ Consensus Number

```
public class RMWConsensus {
    implements Consensus {
    private AtomicInteger r =
        new AtomicInteger(-1);

    public Object decide() {
        r.compareAndSet(-1,i);
        return this.announce[r.get()];
    }
}
```



BROWN[4]

© 2003 Herlihy and Shavit

139

compareAndSet Has ∞ Consensus Number

```
public class RMWConsensus {
    implements Consensus {
    private AtomicInteger r =
        new AtomicInteger(-1);

    public Object decide() {
        r.compareAndSet(-1,i);
        return this.announce[r.get()];
    }
}
```

Initialized to -1



BROWN[4]

© 2003 Herlihy and Shavit

140

compareAndSet Has ∞ Consensus Number

```
public class RMWConsensus {
    implements Consensus {
    private AtomicInteger r =
        new AtomicInteger(-1);

    public Object decide() {
        r.compareAndSet(-1,i);
        return this.announce[r.get()];
    }
}
```

try to swap in my id



BROWN[4]

© 2003 Herlihy and Shavit

141

compareAndSet Has ∞ Consensus Number

```
public class RMWConsensus {
    implements Consensus {
    private AtomicInteger r =
        new AtomicInteger(-1);

    public Object decide() {
        r.compareAndSet(-1,i);
        return this.announce[r.get()];
    }
}
```

Decide winner's preference



BROWN[4]

© 2003 Herlihy and Shavit

142

The Consensus Hierarchy

1 Read/Write Registers, Snapshots...

2 getAndSet, getAndIncrement, ...

.

∞ compareAndSet, ...



BROWN

© 2003 Herlihy and Shavit

143

Multiple Assignment

- Atomic k-assignment
- Solves consensus for $2k-2$ threads
- Every even consensus number has an object (can be extended to odd numbers)



BROWN

© 2003 Herlihy and Shavit

144

Lock-Free Implementations

- Infinitely often some method call completes in a finite number of steps
- Pragmatic approach
- Implies no mutual exclusion



BROWN(2)

© 2003 Herlihy and Shavit

145

Lock-Free Implementations

- Lock-free consensus is just as impossible
- Lock-free = Wait-free for finite executions
- *All the results we presented hold for lock-free algorithms also.*



BROWN(2)

© 2003 Herlihy and Shavit

146

There is More: Universality

- Consensus is universal
- From n-thread consensus
 - Wait-free/Lock-free
 - Linearizable
 - n-threaded
 - Implementation
 - Of any sequentially specified object



BROWN

© 2003 Herlihy and Shavit

147

The Relative Power of Synchronization Methods

Nir Shavit
Multiprocessor Synchronization
Spring 2003



BROWN

Notes For The Relative Power of Synchronization Methods

- Students had a lot of questions during lecture so I added a lot of slides...
- Added lock-freedom in the end, especially since we will talk about it when doing universal stuff. It needs more lock-free stuff since it became major later
- What about robustness, gotta say somewhere that reduction theorem works only for deterministic data structures
- I updated many slides but hav't listed which yet, sorry
- Added slide for getting to CS
- Added slide for using only two registers



BROWN

© 2003 Herlihy and Shavit

149