THE DIFFICULTY OF DETERMINING EXACT EQUALITY OF FLOATING POINT QUANTITIES

Consider the following statement:

```
if (Math.sqrt(x) == y)
        System.out.println ("yes");
else
        System.out.println("no");
```

Its meaning is clear: if $Math.sqrt(x)$ and y are equal, display "yes," otherwise display "no." (Method $sqrt$ in Java's $Math$ class computes $\sqrt{}$, which is returned as type **double**.) Though this may seem straightforward, the meaning of "$==$" is problematic. The computer will print "yes" only if the values of $Math.sqrt(x)$ and y are <u>exactly</u> equal. Unfortunately, a computer cannot always perform calculations and comparisons exactly. (You may have noticed this in your experience with hand calculators.) First, computers cannot always store floating point quantities exactly. For example, the floating point value 5.0 may be represented internally as 4.999999 or 5.000001. (This is never the case with integers, which are always stored exactly.) Second, computers can only add, subtract, multiply, and divide. All other operations, such as $\sqrt{}$, for example, must be reduced to these basic operations. Thus, if x has the value 25 and y has the value 5 in the statement above, $sqrt(x)$ may not turn out to be exactly 5 since the $\sqrt{}$ operation performed by the $sqrt$ method can be only approximated using the basic operations. Then, the condition in the statement will evaluate to *false* and "no" will be incorrectly displayed.

Here is what can be done to avoid this. $sqrt(25)$ may not be computed as precisely 5, but it will be close, say 4.999999. It will lie with a certain <u>margin of error</u> of the exact value. Instead of demanding exact equality, the program should test whether the two quantities, $Math.sqrt(x)$ and y, lie within a specified margin of error of each other. So, the statement above should be written

```
if (Math.abs(Math.sqrt(x) - y) < .00001)
        System.out.println ("yes");
else
        System.out.println("no");
```

The margin of error is .00001. The assumption is that if the two quantities are within .00001 of each other, any difference is probably due to the inaccuracy of representing and manipulating floating point quantities. In algebraic notation, the test is written $|\sqrt{x} - y| < .00001$. Absolute value $|\ |$ (method abs shown above) is needed because it's not known which quantity might be slightly larger.

Not all conditions involving equality need to be complicated by such considerations. If all values are integers and only the four operations mentioned earlier are used, strict equality poses no problem. (Note that the mod operator "%" is just a form of division.)

(OVER)

The size of the margin of error is determined by several factors: the number of digits maintained and the method of computation used by the particular computer. .00001 will probably suffice for your needs, though some applications may require a larger value.

Note that the comparison operators <=, >=, and != require similar adjustments when applied to floating point quantities.