Computer Programs That Write Themselves

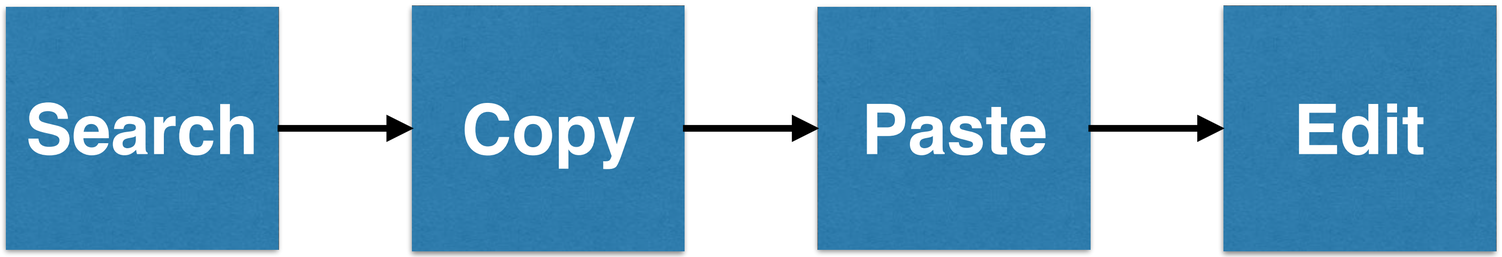
Ruben Martins
rubenm@cs.cmu.edu

17-355/17-665/17-819 Program Analysis
April 16, 2019

Carnegie Mellon University

Traditional Approach

What happens when you need to implement a new module?



Tedious and error prone approach!

Motivation

Motivation

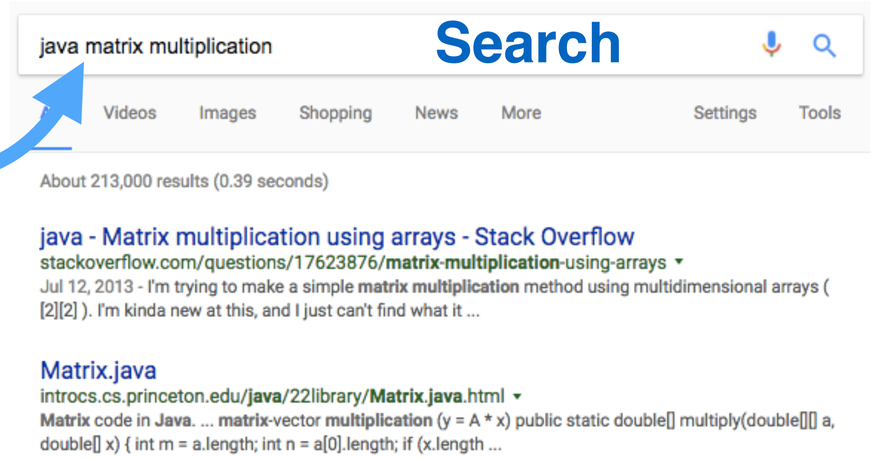
Description: **Matrix multiplication**

```
void multiply(final Vector<Vector<Double>> first,
              final Vector<Vector<Double>> second,
              Vector<Vector<Double>> res)
```

Motivation

Description: Matrix multiplication

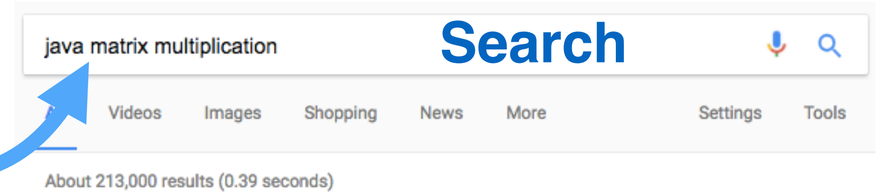
```
void multiply(final Vector<Vector<Double>> first,
              final Vector<Vector<Double>> second,
              Vector<Vector<Double>> res)
```



Motivation

Description: Matrix multiplication

```
void multiply(final Vector<Vector<Double>> first,
              final Vector<Vector<Double>> second,
              Vector<Vector<Double>> res)
```



Java. Matrix multiplication.

Tested with matrices of different size.

Available Code

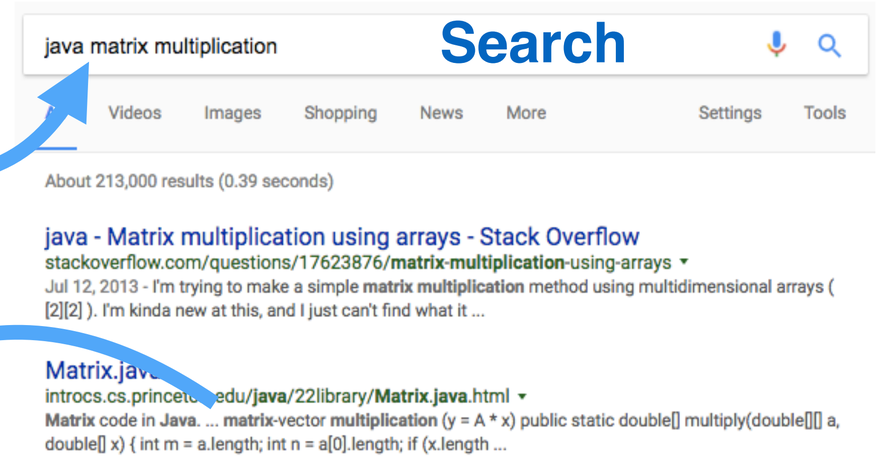
```
public class Matrix {

/**
 * Matrix multiplication method.
 * @param m1 Multiplicand
 * @param m2 Multiplier
 * @return Product
 */
   public static double[][] multiplyByMatrix(double[][] m1, double[][] m2) {
       int m1ColLength = m1[0].length; // m1 columns length
       int m2RowLength = m2.length;    // m2 rows length
       if(m1ColLength != m2RowLength) return null; // matrix multiplication is not possible
       int mRRowLength = m1.length;    // m result rows length
       int mRColLength = m2[0].length; // m result columns length
       double[][] mResult = new double[mRRowLength][mRColLength];
       for(int i = 0; i < mRRowLength; i++) {         // rows from m1
           for(int j = 0; j < mRColLength; j++) {     // columns from m2
               for(int k = 0; k < m1ColLength; k++) { // columns from m1
                   mResult[i][j] += m1[i][k] * m2[k][j];
               }
           }
       }
       return mResult;
   }
```

```
double[][] multiplyByMatrix(
        double[][] m1, double[][] m2)
```

Motivation

Description: **Matrix multiplication**

```
void multiply(final Vector<Vector<Double>> first,
              final Vector<Vector<Double>> second,
              Vector<Vector<Double>> res)
```



Java. Matrix multiplication.

Tested with matrices of different size.

Available Code

```
public class Matrix {

/**
 * Matrix multiplication method.
 * @param m1 Multiplicand
 * @param m2 Multiplier
 * @return Product
 */
   public static double[][] multiplyByMatrix(double[][] m1, double[][] m2) {
       int m1ColLength = m1[0].length; // m1 columns length
       int m2RowLength = m2.length;    // m2 rows length
       if(m1ColLength != m2RowLength) return null; // matrix multiplication is not possible
       int mRRowLength = m1.length;    // m result rows length
       int mRColLength = m2[0].length; // m result columns length
       double[][] mResult = new double[mRRowLength][mRColLength];
       for(int i = 0; i < mRRowLength; i++) {         // rows from m1
           for(int j = 0; j < mRColLength; j++) {     // columns from m2
               for(int k = 0; k < m1ColLength; k++) { // columns from m1
                   mResult[i][j] += m1[i][k] * m2[k][j];
               }
           }
       }
       return mResult;
   }
```

```
double[][] multiplyByMatrix(
        double[][] m1, double[][] m2)
```

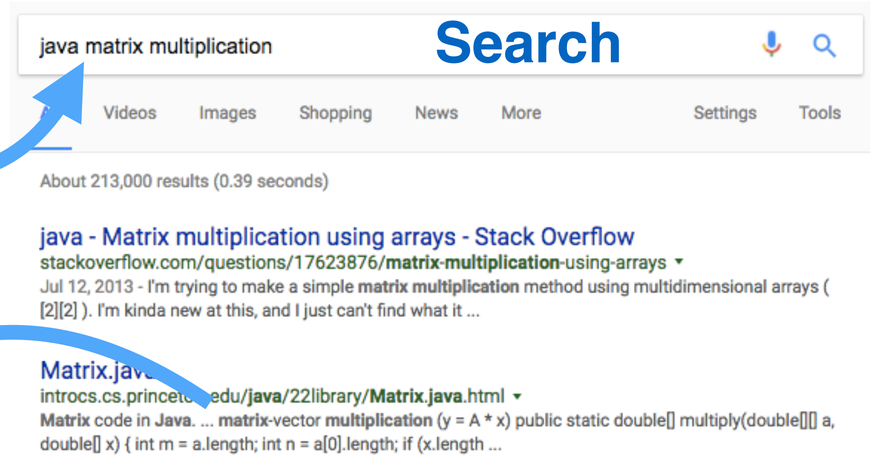
Edit

Desired Code

Motivation

Description: Matrix multiplication

```
void multiply(final Vector<Vector<Double>> first,
              final Vector<Vector<Double>> second,
              Vector<Vector<Double>> res)
```



Java. Matrix multiplication.

Tested with matrices of different size.

Available Code

```
public class Matrix {

/**
 * Matrix multiplication method.
 * @param m1 Multiplicand
 * @param m2 Multiplier
 * @return Product
 */
   public static double[][] multiplyByMatrix(double[][] m1, double[][] m2) {
       int m1ColLength = m1[0].length; // m1 columns length
       int m2RowLength = m2.length;    // m2 rows length
       if(m1ColLength != m2RowLength) return null; // matrix multiplication is not possible
       int mRRowLength = m1.length;    // m result rows length
       int mRColLength = m2[0].length; // m result columns length
       double[][] mResult = new double[mRRowLength][mRColLength];
       for(int i = 0; i < mRRowLength; i++) {         // rows from m1
           for(int j = 0; j < mRColLength; j++) {     // columns from m2
               for(int k = 0; k < m1ColLength; k++) { // columns from m1
                   mResult[i][j] += m1[i][k] * m2[k][j];
               }
           }
       }
       return mResult;
   }
```

```
double[][] multiplyByMatrix(
        double[][] m1, double[][] m2)
```

Edit

Desired Code

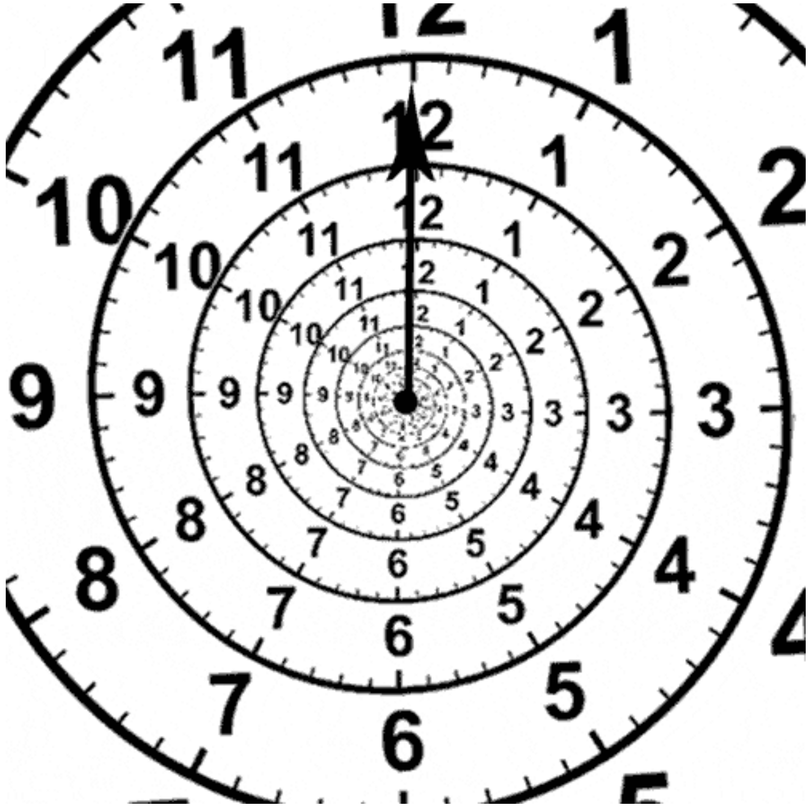
Long time spent on search
Significant effort required to edit!

Traditional Approach

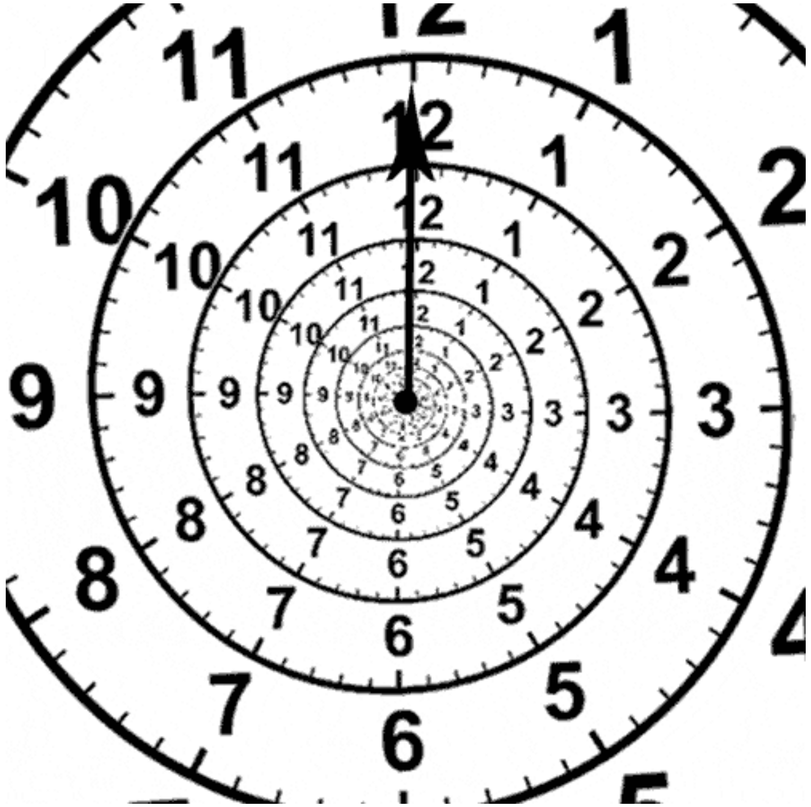
Available Code ≠ **Desired Code**

Significant effort required to edit!

10 minutes later…



10 minutes later…



Traditional Approach

Available Code + **Improper Edit** = **Buggy Code**

Bugs are easily integrated while editing!

Program Synthesis

- What if the program could write itself?

Goal

Description

Desired Code

Goal

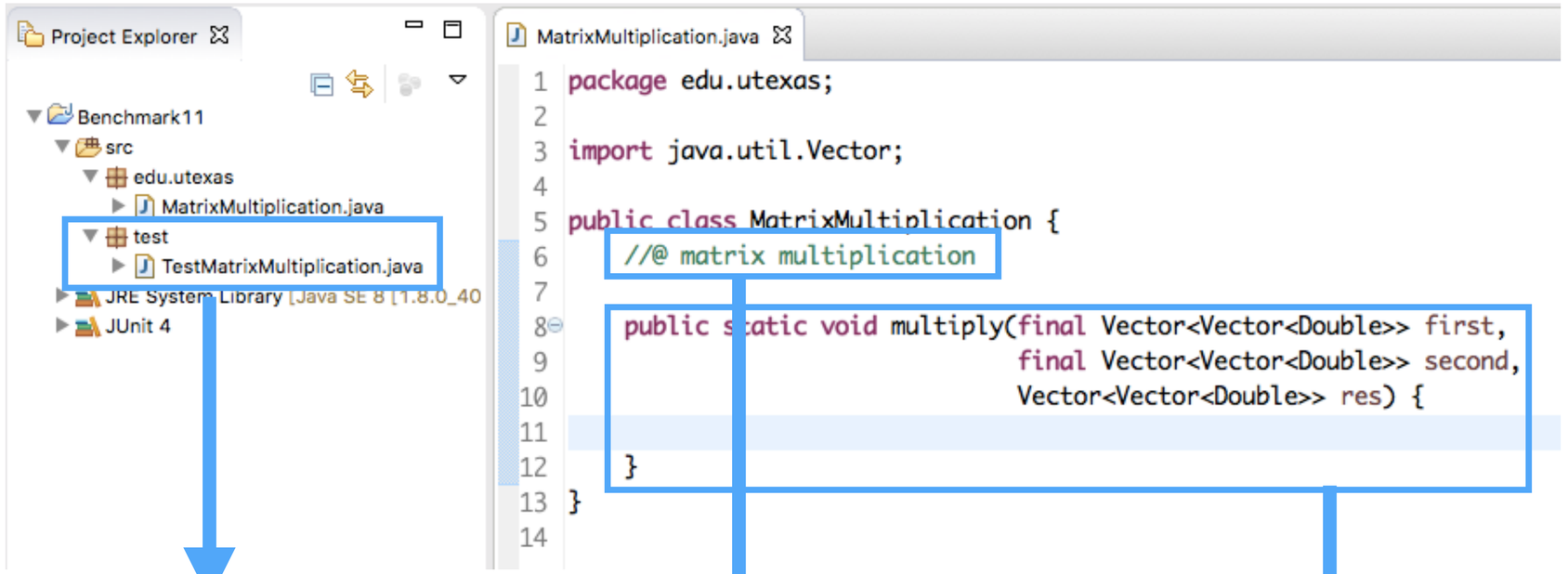
Description + 

= **Desired Code**

Goal

- Search and reuse code **automatically**
- Method description
  1. Natural language description
  2. Method signature
  3. Test suite

Hunter

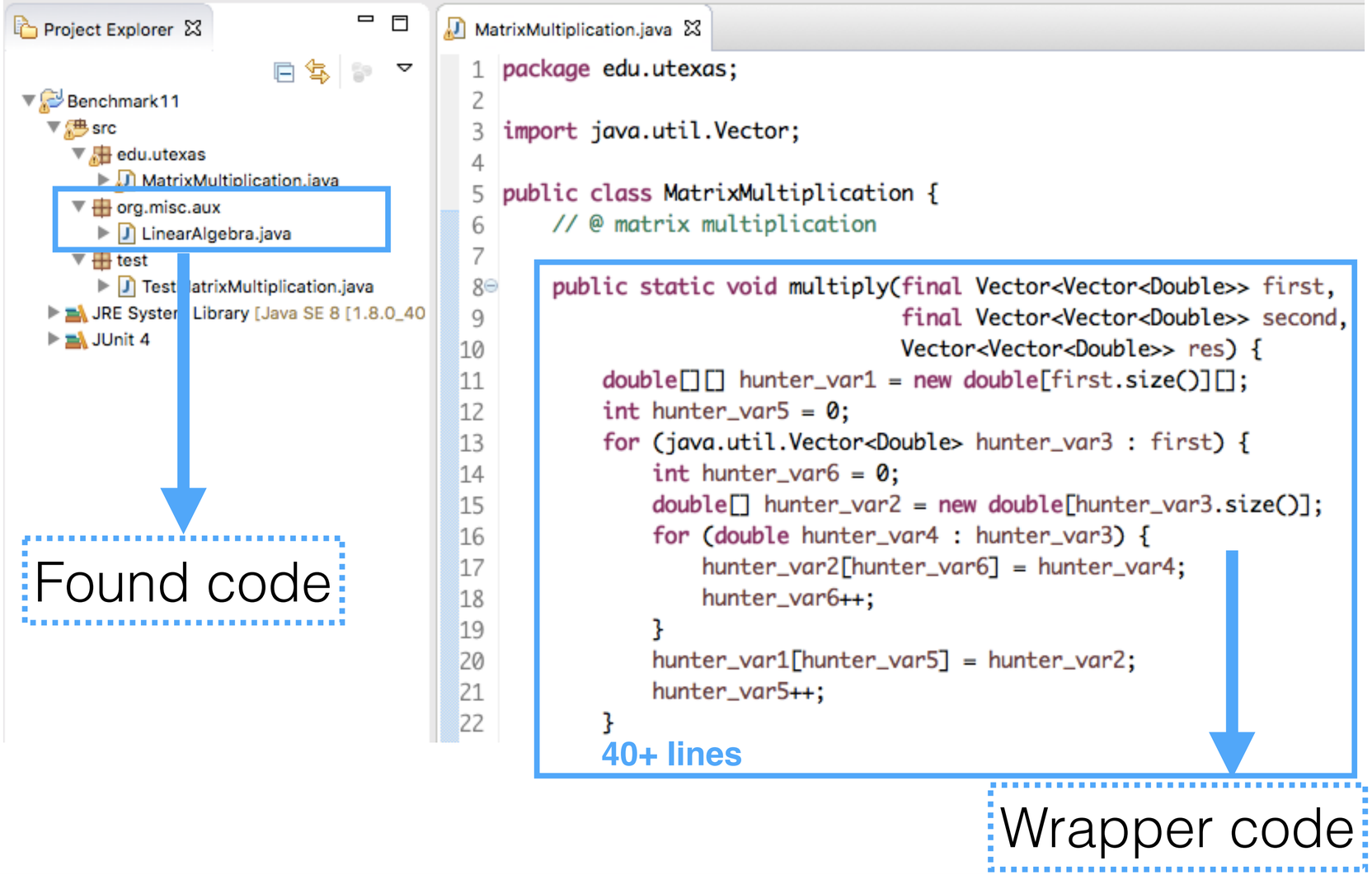


Test Suite

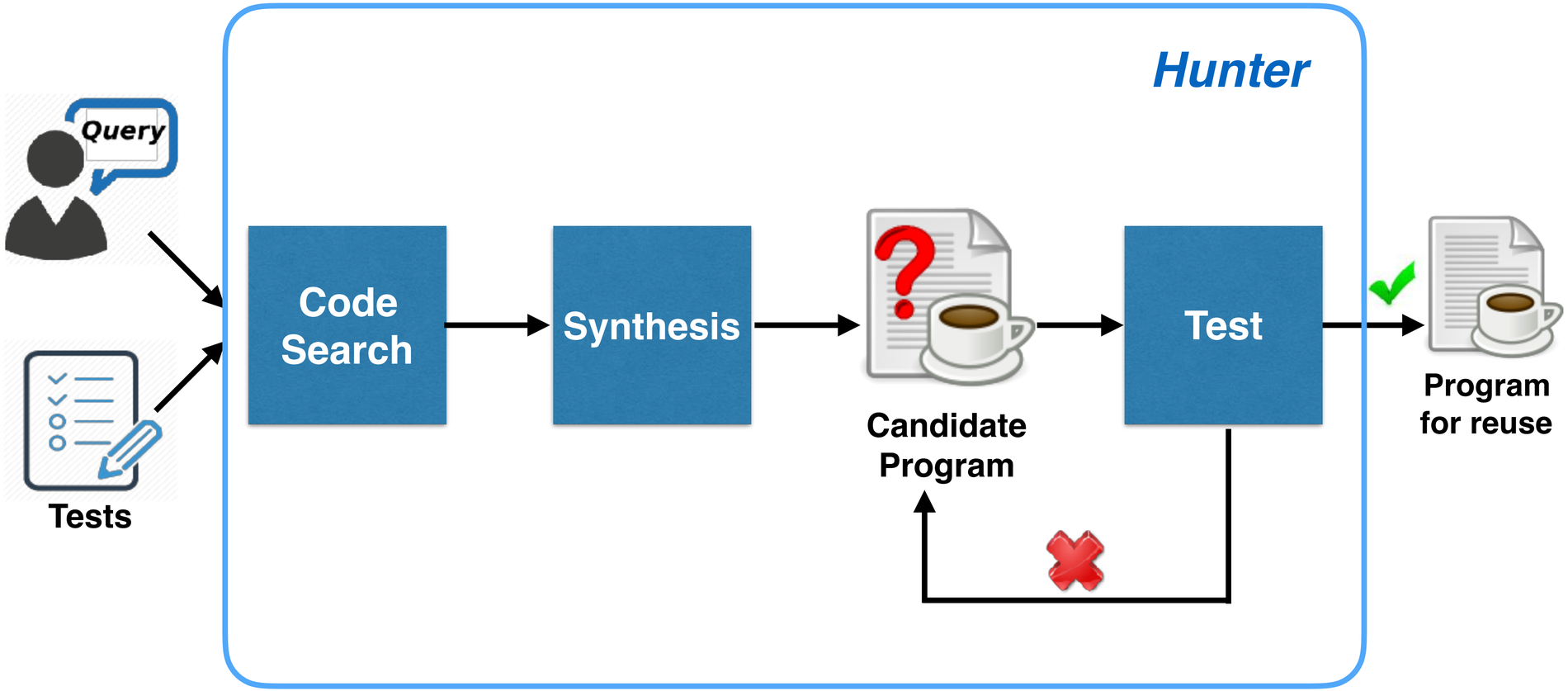
Natural language description

Method Signature

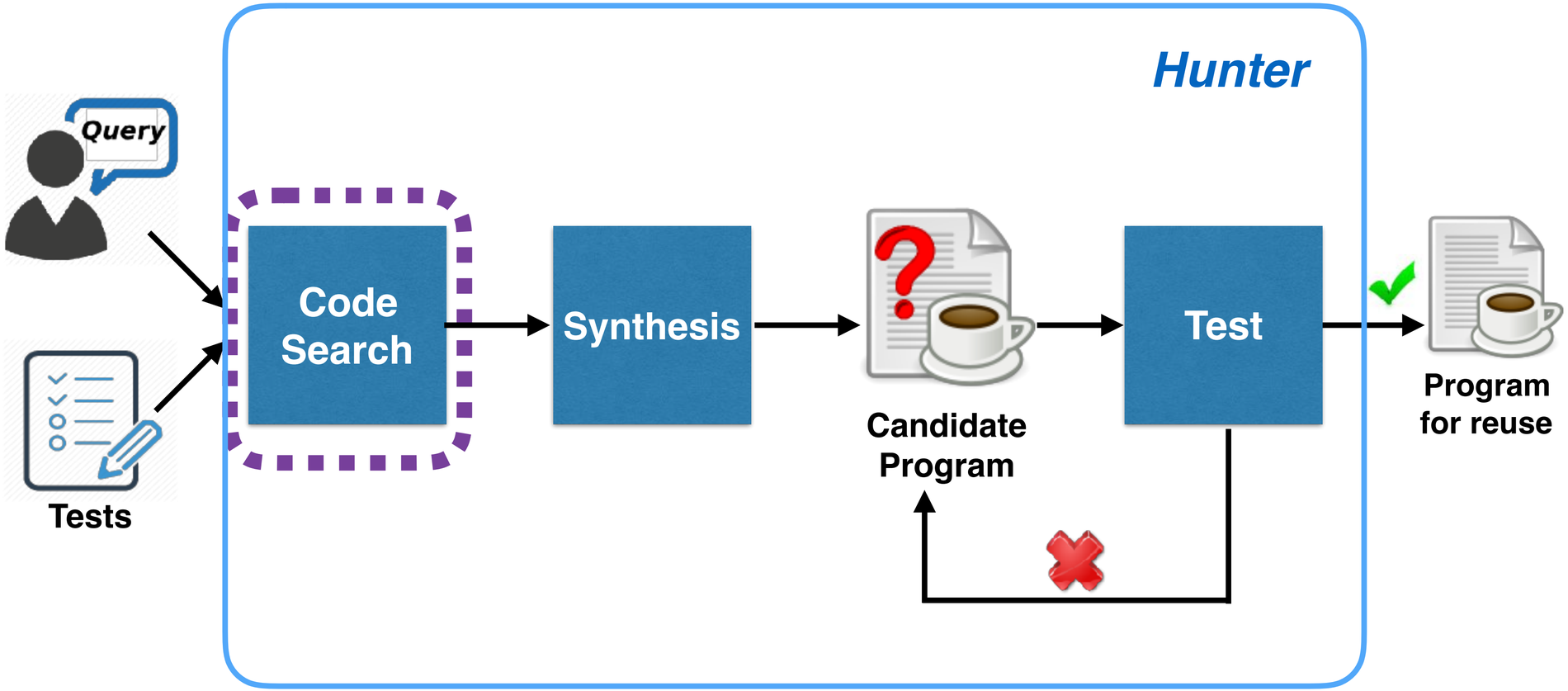
Hunter


Project Explorer
Benchmark11
src
edu.utexas
MatrixMultiplication.java
org.misc.aux
LinearAlgebra.java
test
Test MatrixMultiplication.java
JRE System Library [Java SE 8 [1.8.0_40
JUnit 4
MatrixMultiplication.java
1 package edu.utexas;
2
3 import java.util.Vector;
4
5 public class MatrixMultiplication {
6 // @ matrix multiplication
7
8 public static void multiply(final Vector<Vector<Double>> first,
9 final Vector<Vector<Double>> second,
10 Vector<Vector<Double>> res) {
11 double[][] hunter_var1 = new double[first.size()][];
12 int hunter_var5 = 0;
13 for (java.util.Vector<Double> hunter_var3 : first) {
14 int hunter_var6 = 0;
15 double[] hunter_var2 = new double[hunter_var3.size()];
16 for (double hunter_var4 : hunter_var3) {
17 hunter_var2[hunter_var6] = hunter_var4;
18 hunter_var6++;
19 }
20 hunter_var1[hunter_var5] = hunter_var2;
21 hunter_var5++;
22 }
40+ lines
Found code
Wrapper code

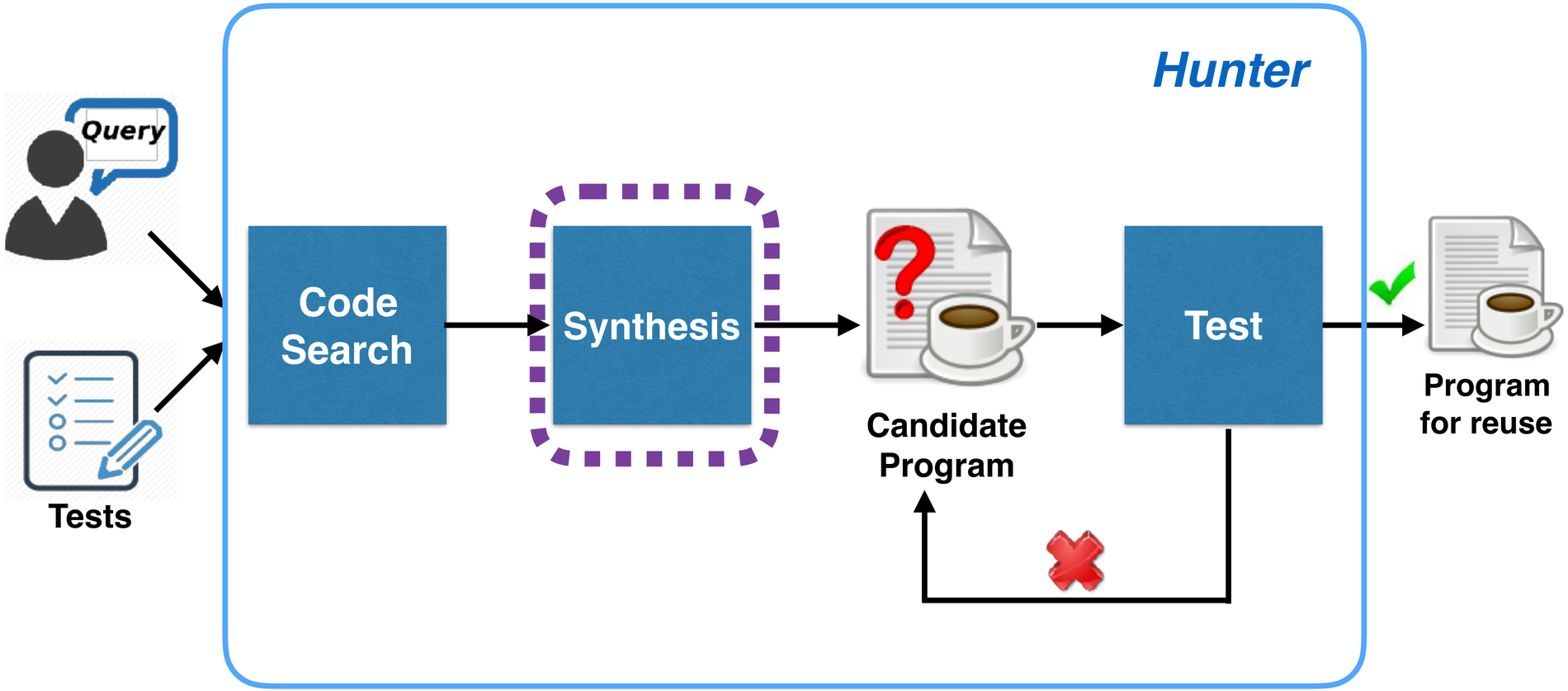
Hunter architecture



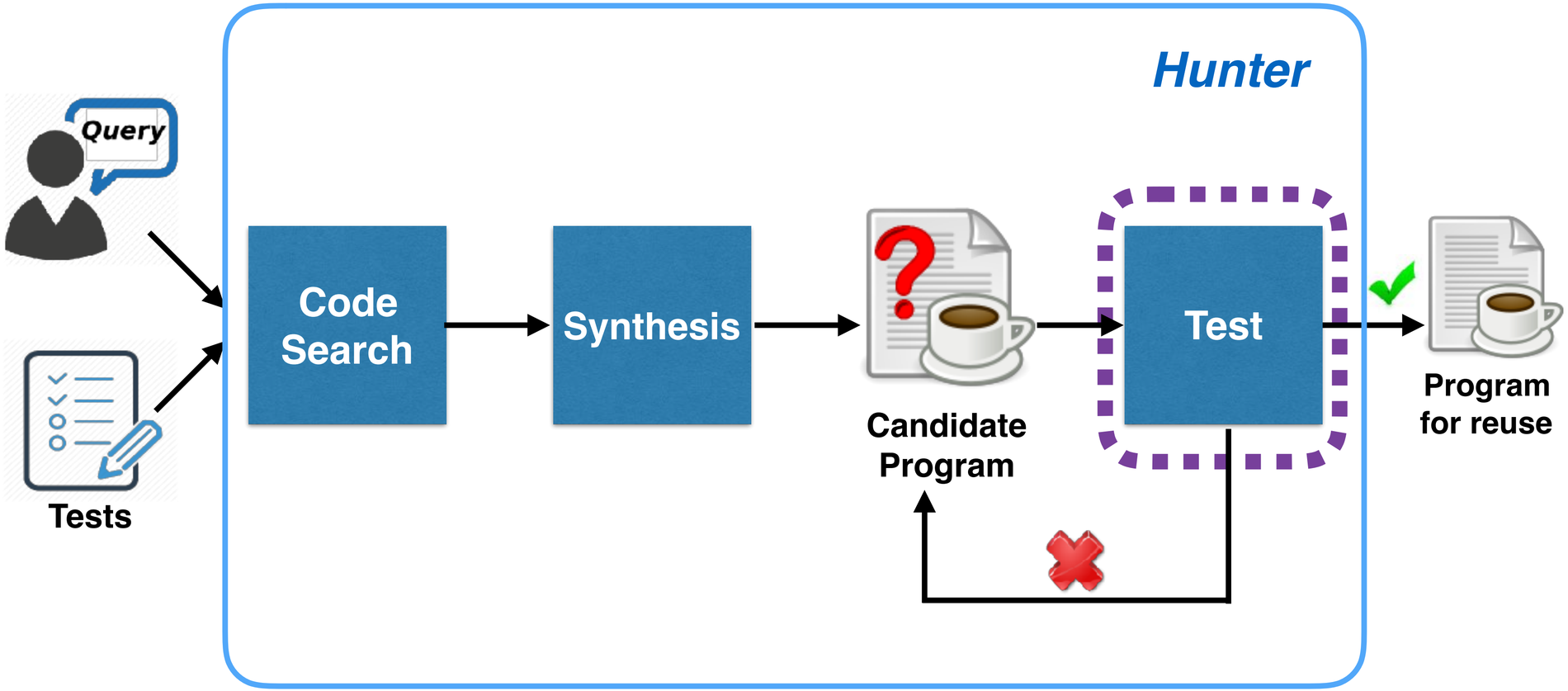
Hunter architecture



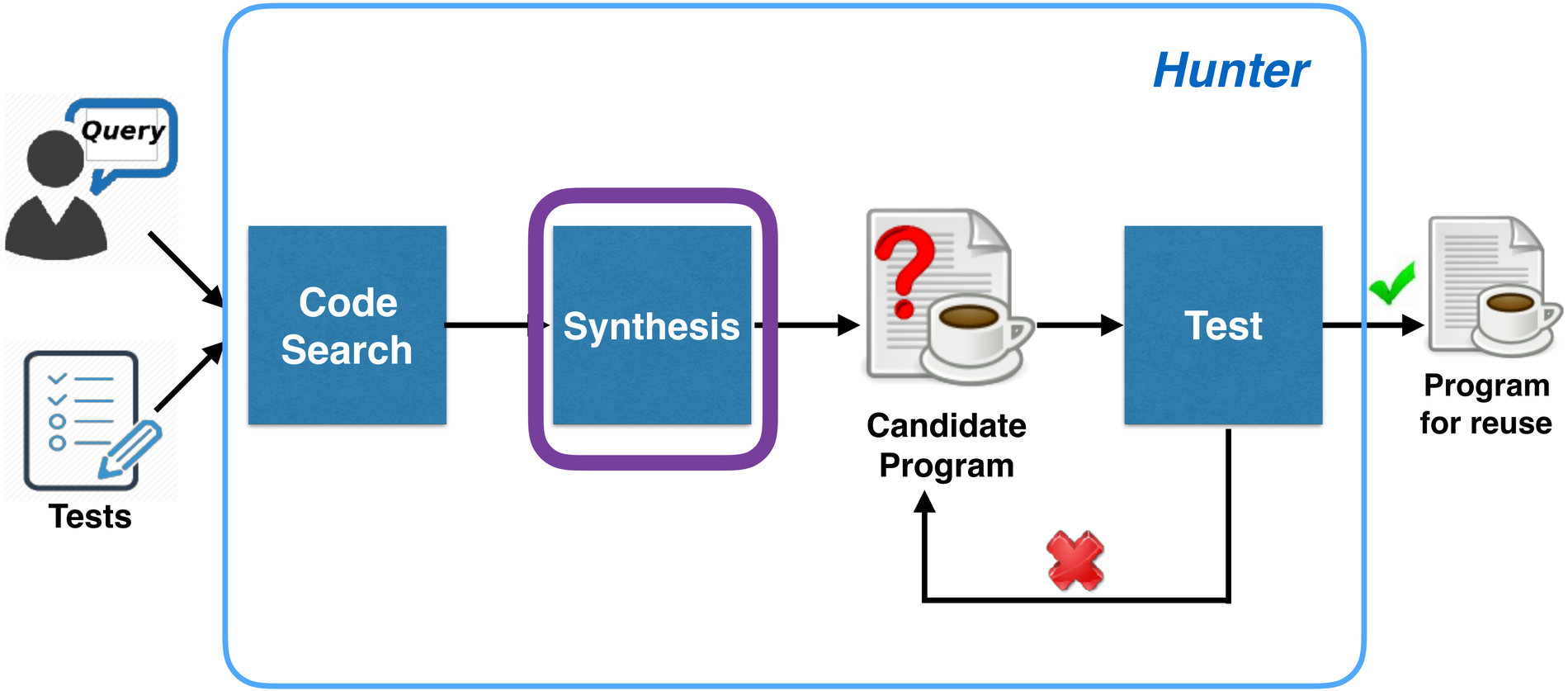
Hunter architecture



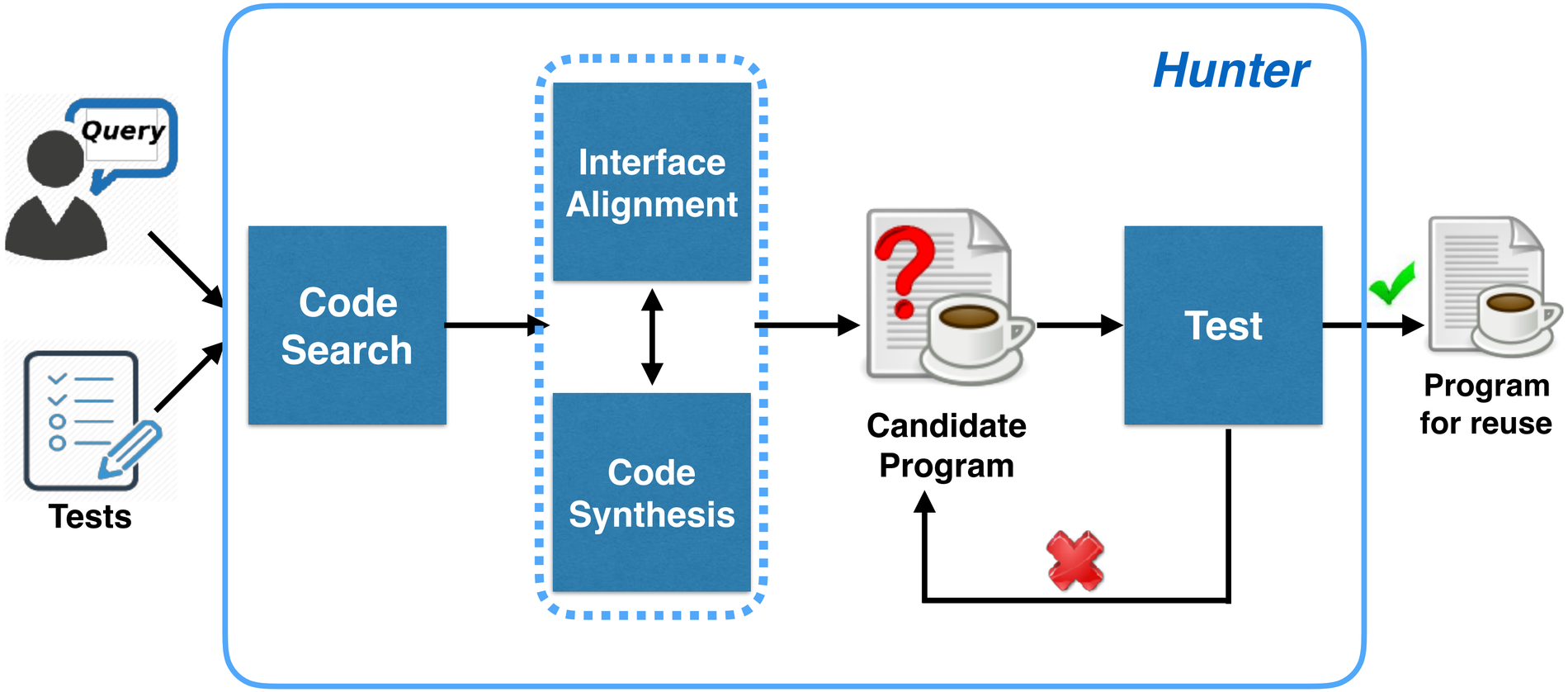
Hunter architecture



Hunter architecture

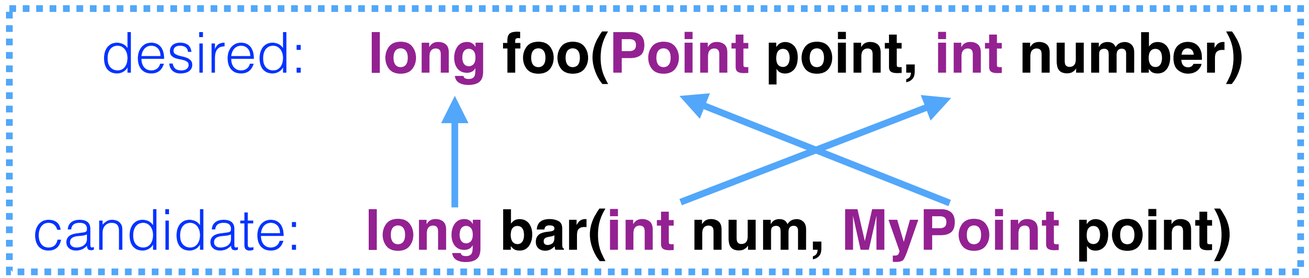


Hunter architecture



Interface alignment

- Generate reuse scheme by mapping arguments and return value from candidates to the desired signature



Code synthesis

- Synthesize **wrapper** to invoke candidates properly
- Handle type conversions by building succinct graph representation and performing reachability analysis

Code synthesis

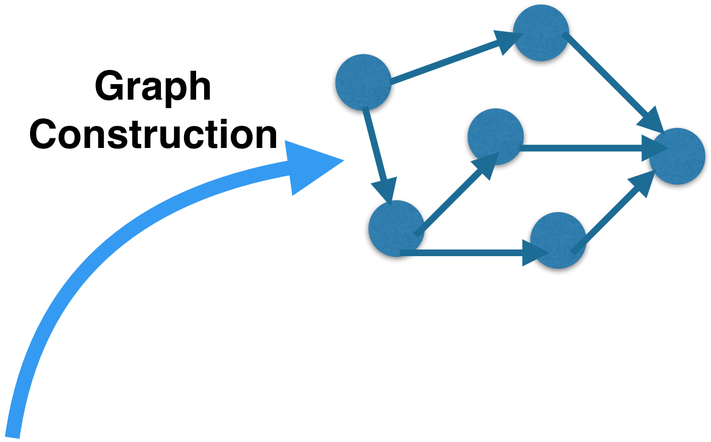
- Synthesize **wrapper** to invoke candidates properly
- Handle type conversions by building succinct graph representation and performing reachability analysis

```
class Point {
    int x; int y;
    Point(int _x, int _y);
}
class MyPoint {
    int x; int y;
    int getX();
    int getY();
}
```

Code synthesis

- Synthesize **wrapper** to invoke candidates properly
- Handle type conversions by building succinct graph representation and performing reachability analysis

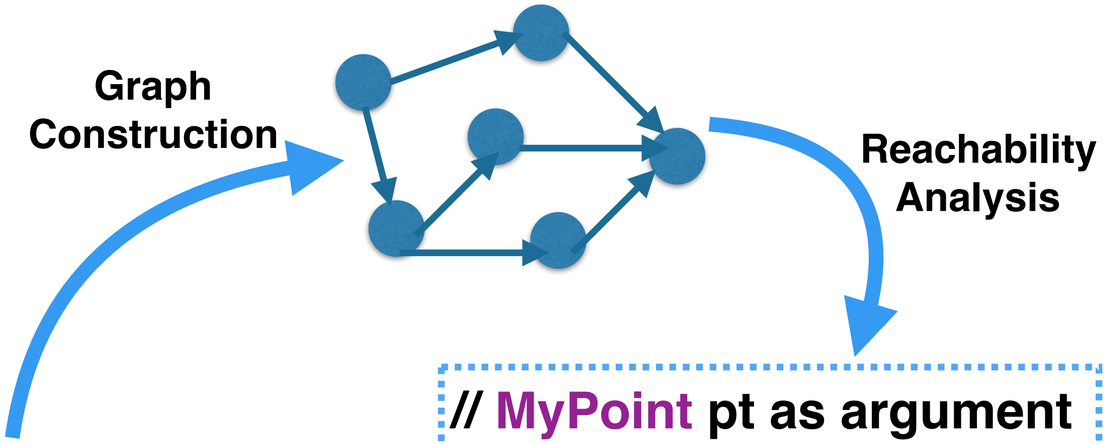
```
class Point {
    int x; int y;
    Point(int _x, int _y);
}
class MyPoint {
    int x; int y;
    int getX();
    int getY();
}
```



Code synthesis

- Synthesize **wrapper** to invoke candidates properly
- Handle type conversions by building succinct graph representation and performing reachability analysis

```
class Point {
    int x; int y;
    Point(int _x, int _y);
}
class MyPoint {
    int x; int y;
    int getX();
    int getY();
}
```




```
// MyPoint pt as argument
int x = pt.getX();
int y = pt.getY();
Point p = new Point(x, y);
```

Who can Program Synthesis help?

Who can Program Synthesis help?

- Are we trying to replace programmers? **No!**
  - We want to make programmers life easier
  - Automating tedious and repetitive tasks

Who can Program Synthesis help?

- Are we trying to replace programmers? **No!**
  - We want to make programmers life easier
  - Automating tedious and repetitive tasks
- 99% of computer users **cannot program**!
  - They struggle with simple repetitive tasks
  - Help non-CS people to automate their daily tasks

Outline

Code Reuse

FSE'16

Complex Java APIs

POPL'17

Data Wrangling

PLDI'17

Learning in Synthesis

PLDI'18

Outline

Code Reuse

FSE'16

Complex Java APIs

POPL'17

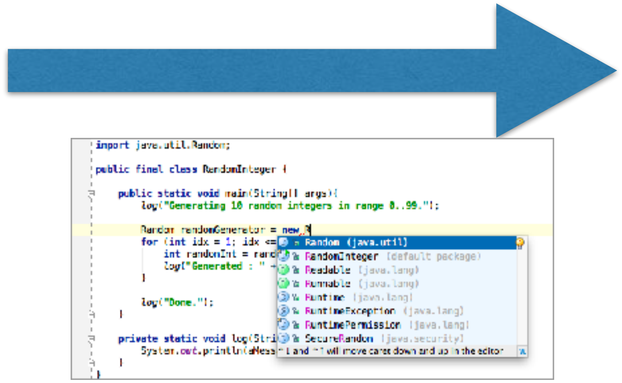
Data Wrangling

PLDI'17

Learning in Synthesis

PLDI'18

API exploration can be tedious!



```
import java.util.Random;

public final class RandomInteger {

    public static void main(String[] args){
        log("Generating 10 random integers in range 0..99.");

        Random randomGenerator = new R
        for (int idx = 1; idx <=
            int randomInt = rand
            log("Generated : "
        }

        log("Done.");
    }

    private static void log(Stri
        System.out.println(Mess
    }
}
```

```
Random (java.util)
RandomInteger (default package)
Readable (java.lang)
Runnable (java.lang)
Runtime (java.lang)
RuntimeException (java.lang)
RuntimePermission (java.lang)
SecureRandom (java.security)
```

Send HTTP request
Compute GCD
Rotate an image

Programmers spend a lot of effort
learning APIs!

Synthesizing programs with APIs

- Automatically synthesize an HTTP POST request:

```
import com.mashape.unirest.*;

String sendHttpPost(String url, String body) {









}
```

Synthesizing programs with APIs

- Automatically synthesize an HTTP POST request:

```
import com.mashape.unirest.*;

String sendHttpPost(String url, String body) {

  HttpRequestWithBody req = post(url);
  RequestBodyEntity ent = req.body(body);
  BaseRequest breq = ent;
  HttpResponse resp = breq.asString();
  String result = resp.getStringBody();
  return result;

}
```

Demo

- A demo of **SyPet** is available at:

https://utopia-group.github.io/sypet/assets/files/sypet-demo.m4v

How to find the correct program?

How to find the correct program?

Use **Petri net reachability** analysis to look for well-typed programs of the desired type

How to find the correct program?

Use **Petri net reachability** analysis to look for well-typed programs of the desired type

- Model relationships between components using Petri nets

How to find the correct program?

Use **Petri net reachability** analysis to look for well-typed programs of the desired type

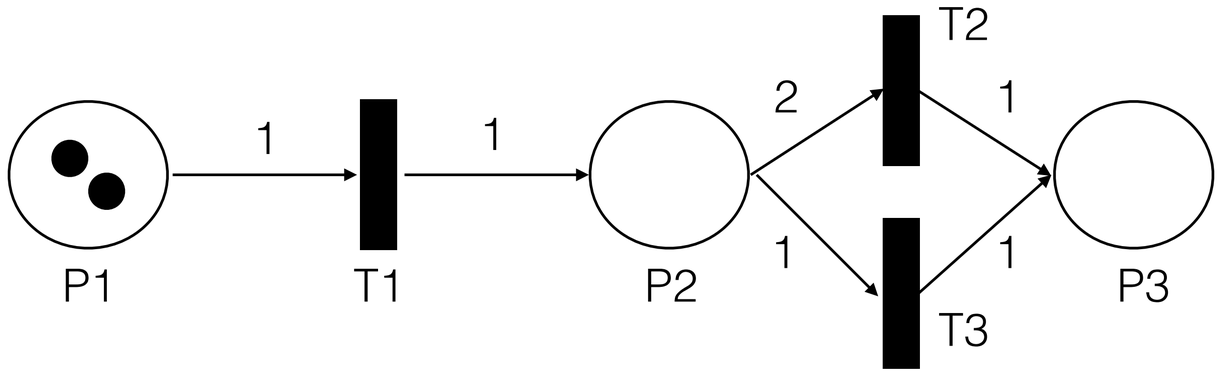
- Model relationships between components using Petri nets
- Use type signature of desired method to mark initial and target configurations

How to find the correct program?

Use **Petri net reachability** analysis to look for well-typed programs of the desired type

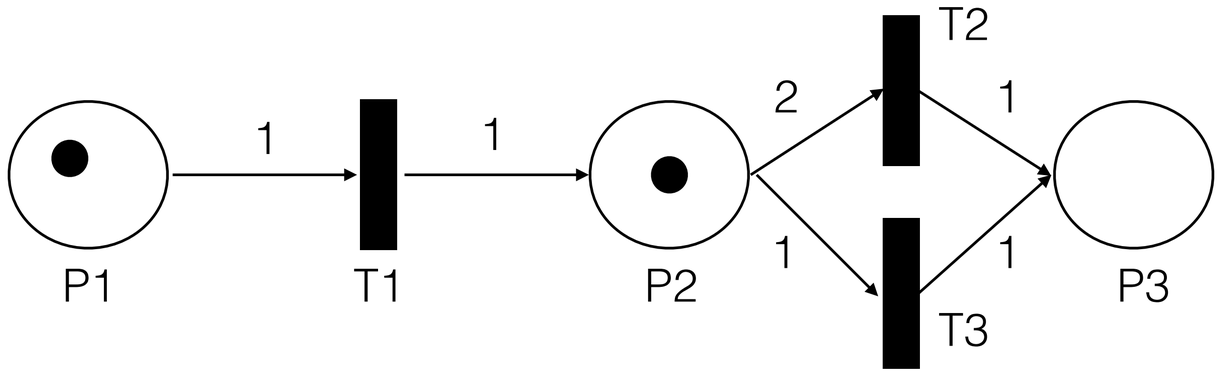
- Model relationships between components using Petri nets
- Use type signature of desired method to mark initial and target configurations
- Perform reachability analysis to find valid sequences of method calls

Petri nets in a nutshell



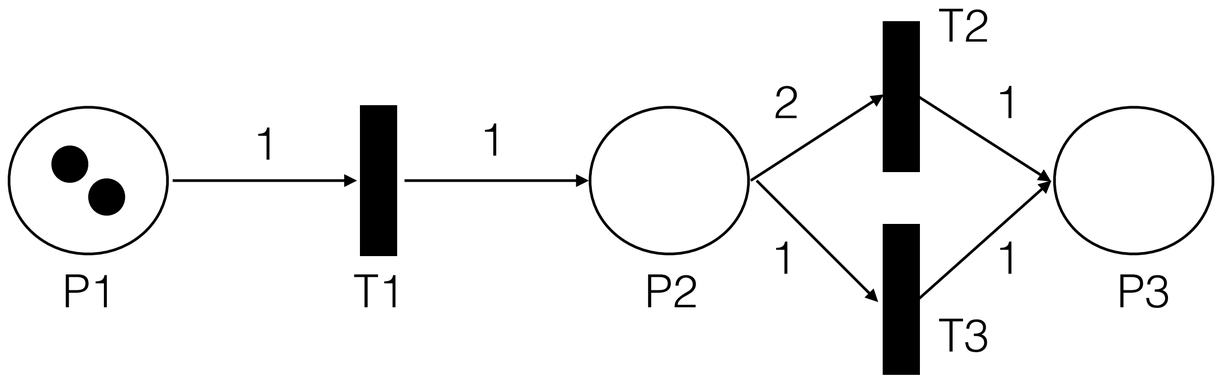
- Petri net is a generalized graph with two kinds of nodes: **places** and **transitions**
- Each place contains zero or more tokens; edges are labeled with a number (of tokens)
- A transition T can fire if, for each edge (p,T) with label n, place p contains at least n tokens
- Firing a transition T consumes (resp. produces) the indicated number of tokens at the source (resp. target) nodes

Petri nets in a nutshell



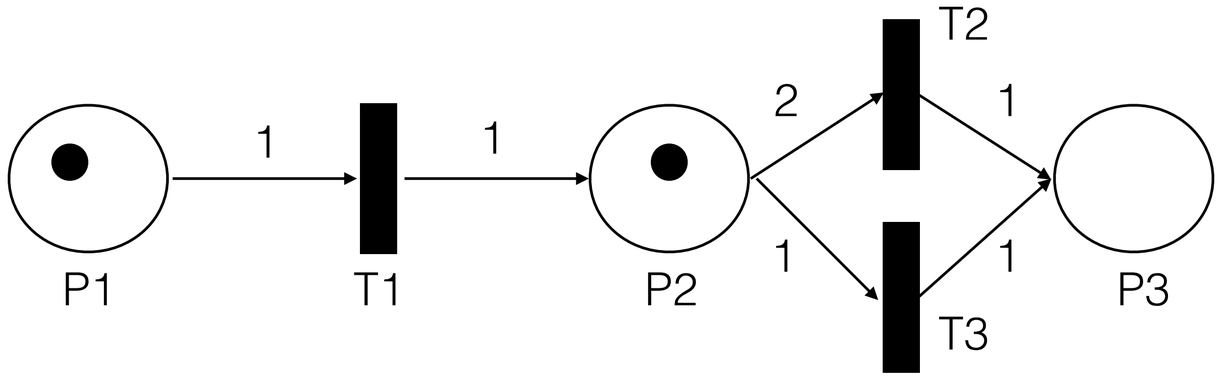
- Petri net is a generalized graph with two kinds of nodes: **places** and **transitions**
- Each place contains zero or more tokens; edges are labeled with a number (of tokens)
- A transition T can fire if, for each edge (p,T) with label n, place p contains at least n tokens
- Firing a transition T consumes (resp. produces) the indicated number of tokens at the source (resp. target) nodes

Reachability problem in Petri nets



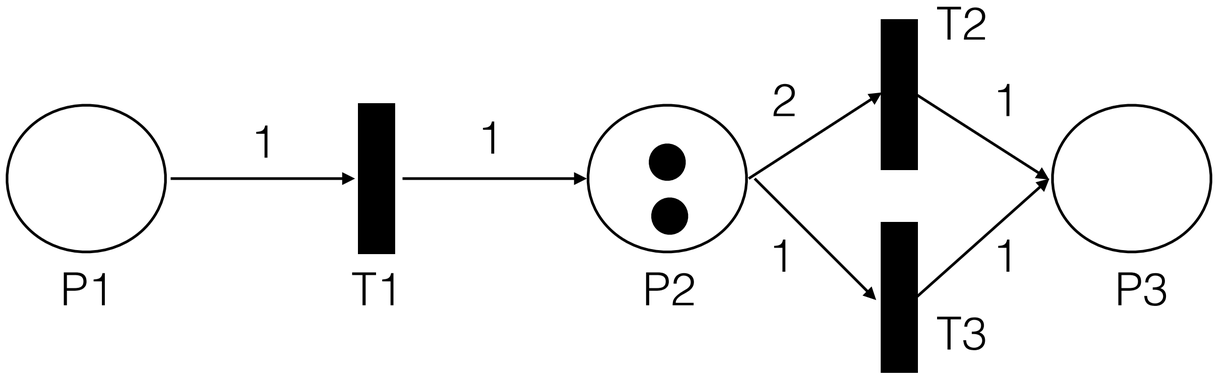
- Reachability problem: Given a Petri net with initial marking M and a target marking M′, is it possible to obtain M’ by firing a sequence of transitions?
- Example: Consider marking M′ : [P1→0,P2→0,P3→1].
- This marking is reachable, and accepting run is T1,T1,T2.

Reachability problem in Petri nets



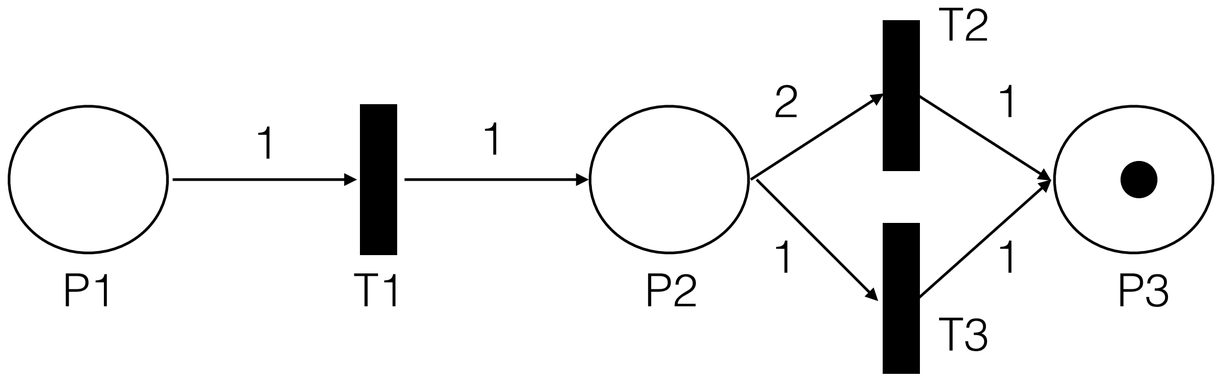
- Reachability problem: Given a Petri net with initial marking M and a target marking M′, is it possible to obtain M’ by firing a sequence of transitions?
- Example: Consider marking M′ : [P1→0,P2→0,P3→1].
- This marking is reachable, and accepting run is T1,T1,T2.

Reachability problem in Petri nets



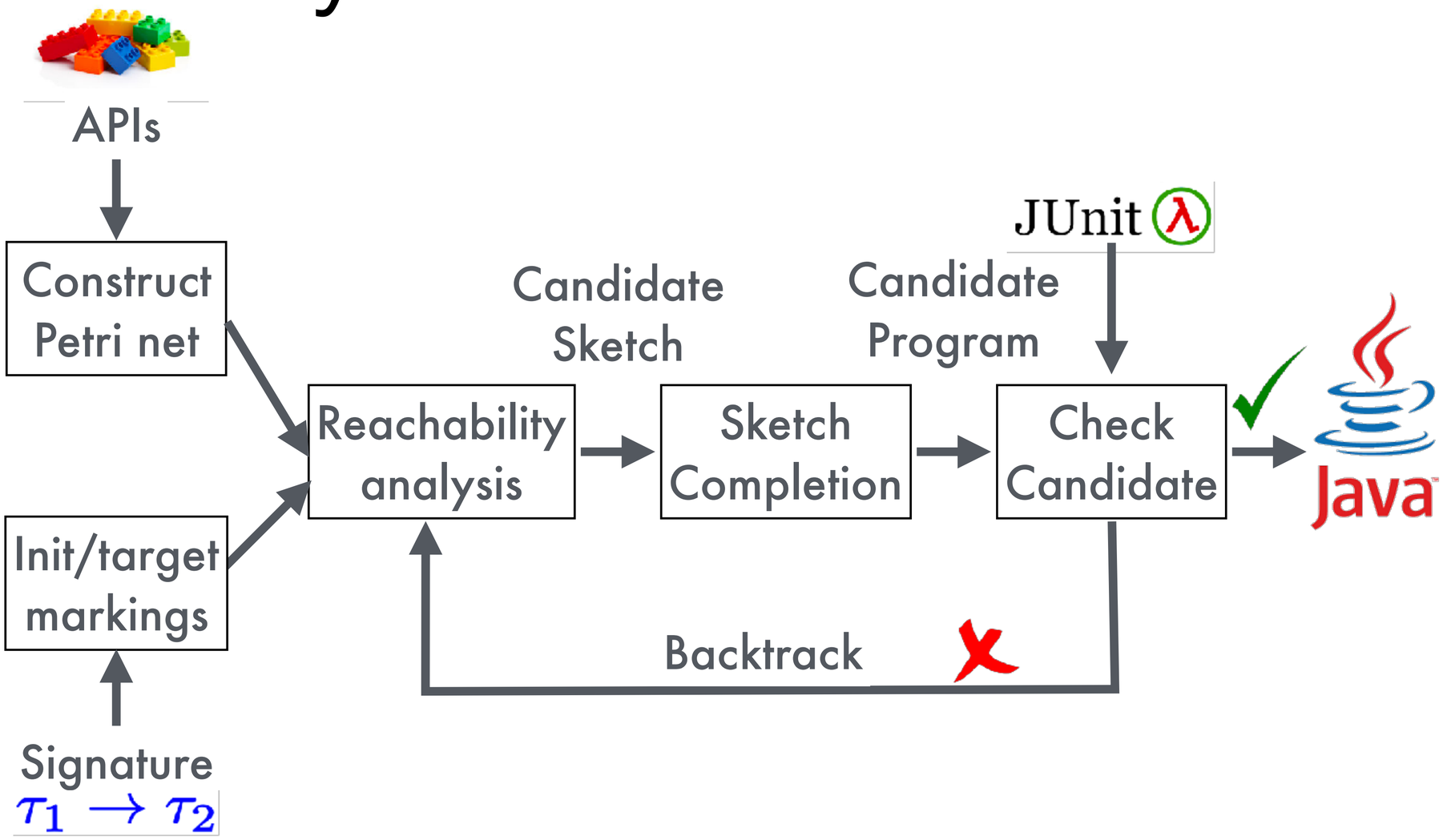
- Reachability problem: Given a Petri net with initial marking M and a target marking M′, is it possible to obtain M’ by firing a sequence of transitions?
- Example: Consider marking M′ : [P1→0,P2→0,P3→1].
- This marking is reachable, and accepting run is T1,T1,T2.

Reachability problem in Petri nets

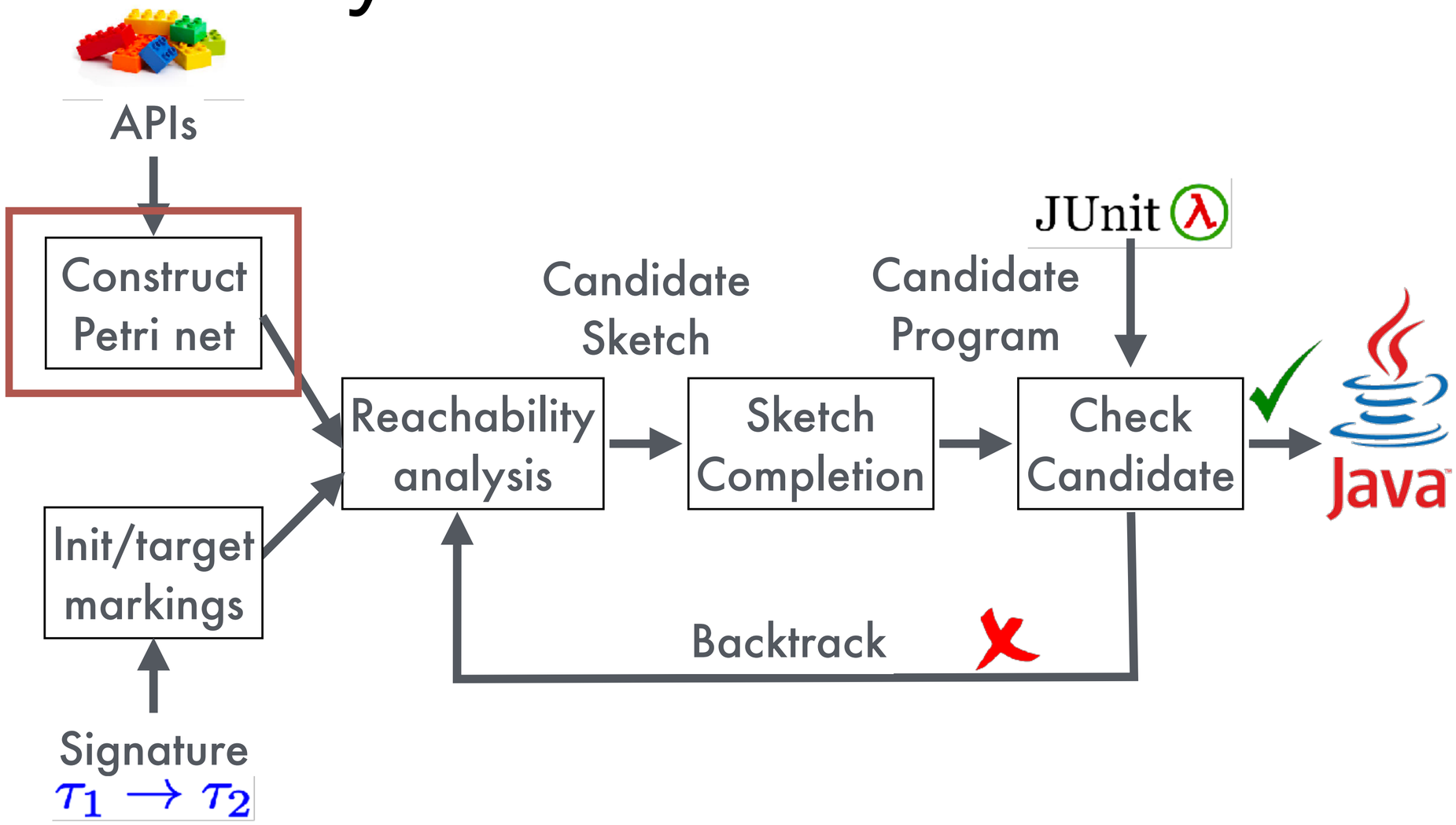


- Reachability problem: Given a Petri net with initial marking M and a target marking M′, is it possible to obtain M’ by firing a sequence of transitions?
- Example: Consider marking M′ : [P1→0,P2→0,P3→1].
- This marking is reachable, and accepting run is T1,T1,T2.

SyPet architecture

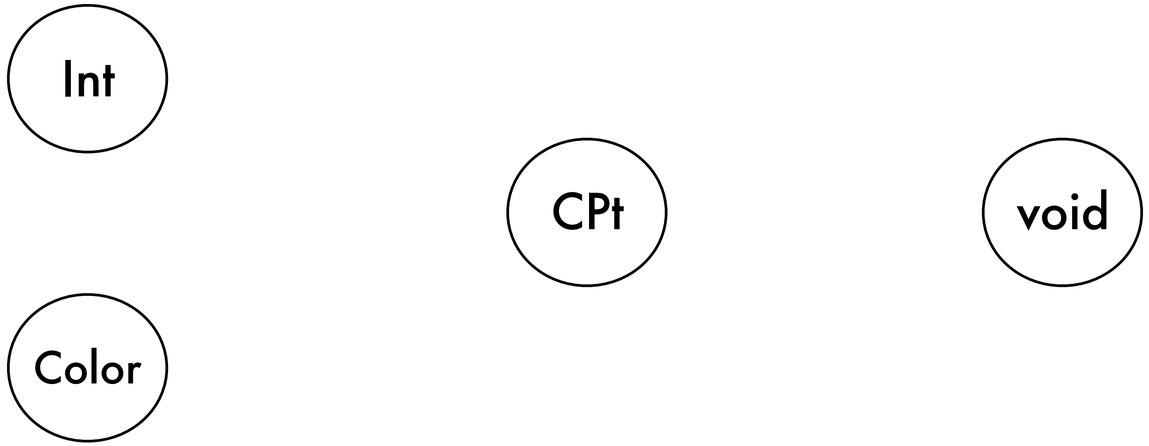


SyPet architecture



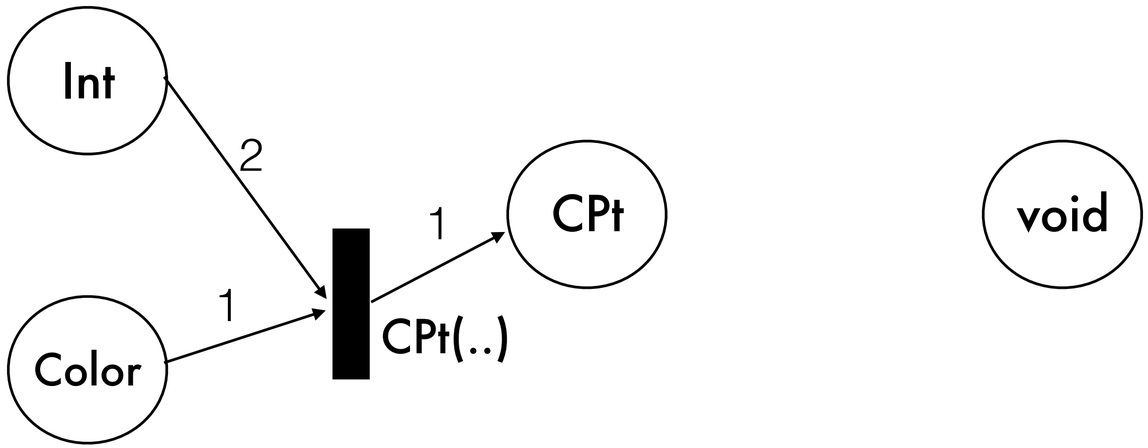
Petri net construction

```
class CPt {
  CPt(Int x, Int y, Color c);
  Int getX();
  void setColor(Color c);
  ...
}
```



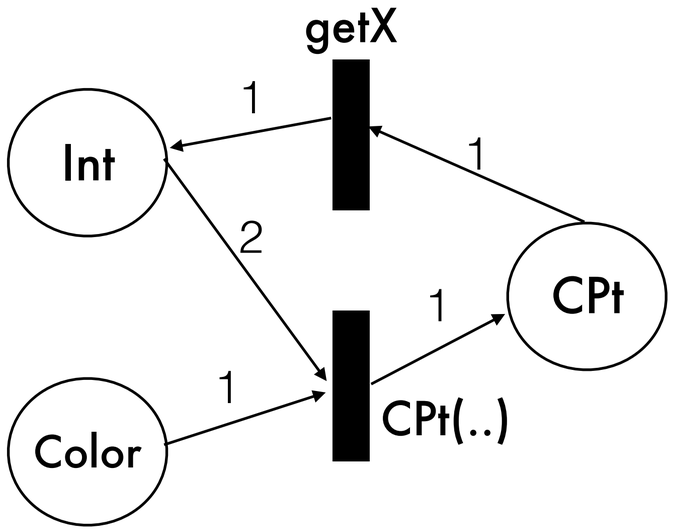
Petri net construction

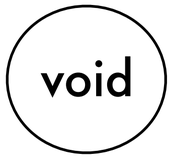
```
class CPt {
  CPt(Int x, Int y, Color c);
  Int getX();
  void setColor(Color c);
  ...
}
```



Petri net construction

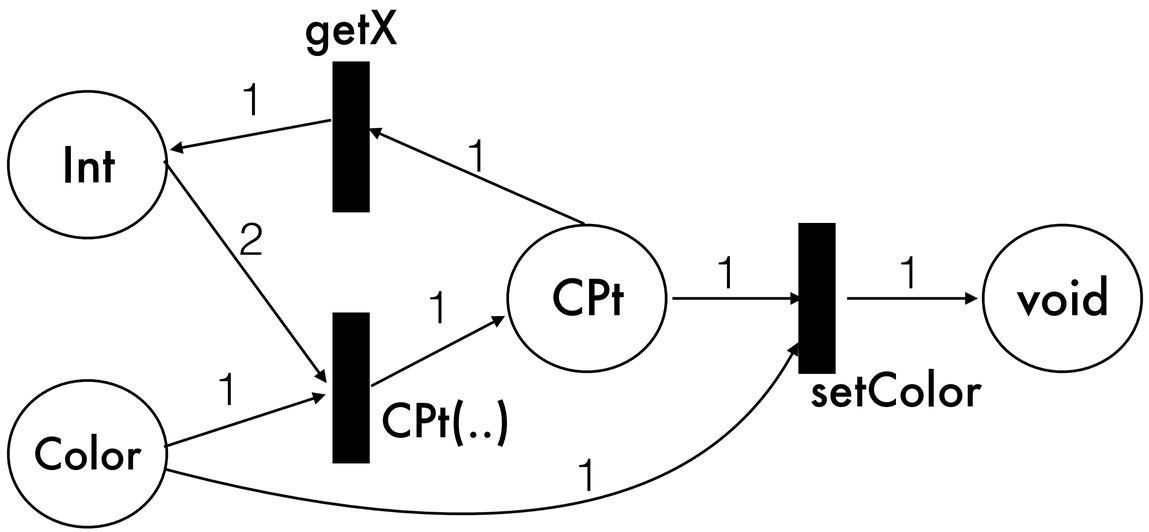
```
class CPt {
  CPt(Int x, Int y, Color c);
  Int getX();
  void setColor(Color c);
  ...
}
```



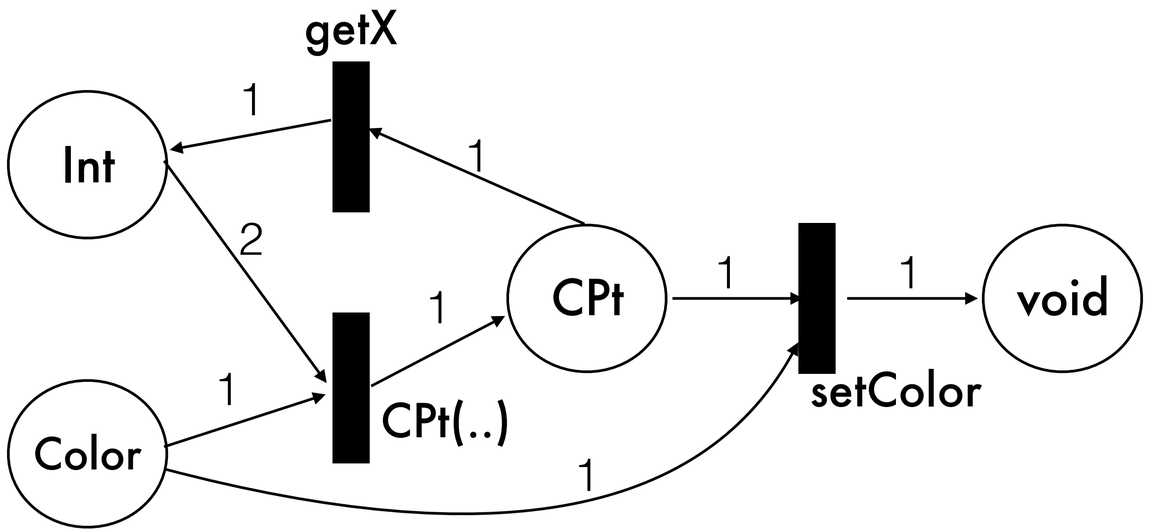


Petri net construction

```
class CPt {
  CPt(Int x, Int y, Color c);
  Int getX();
  void setColor(Color c);
  ...
}
```

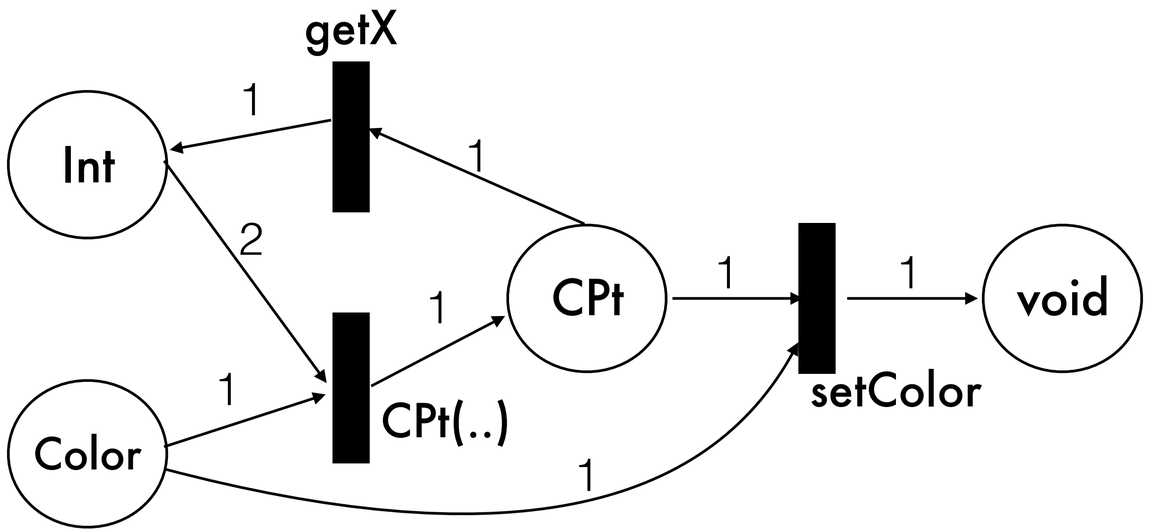


Clone transitions



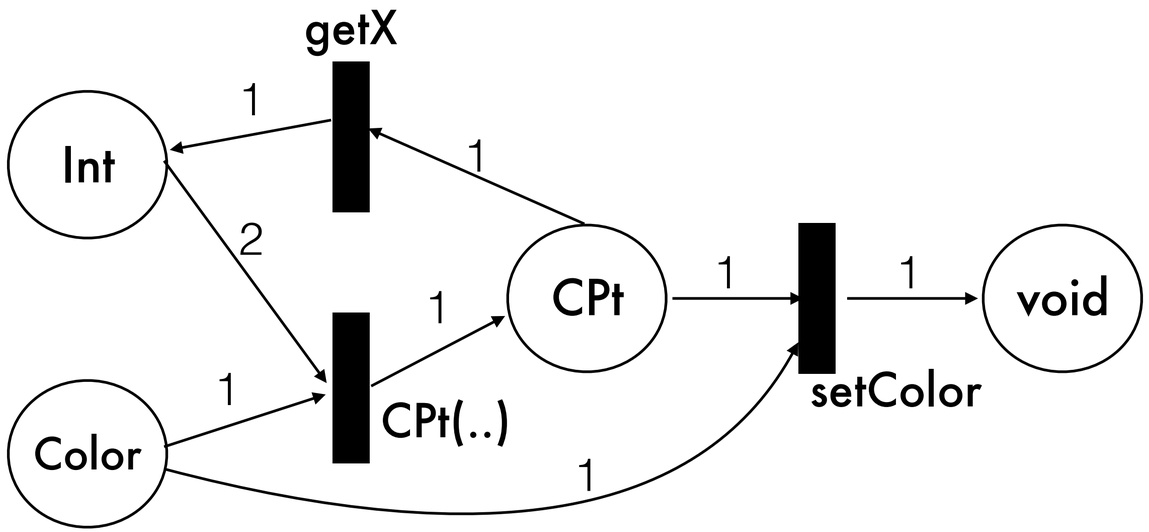
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects



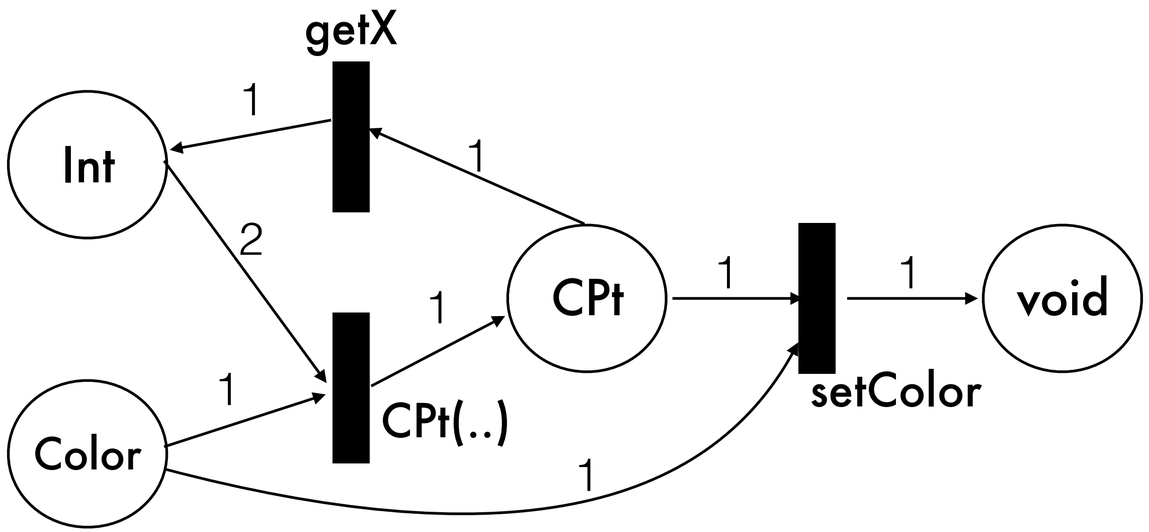
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!



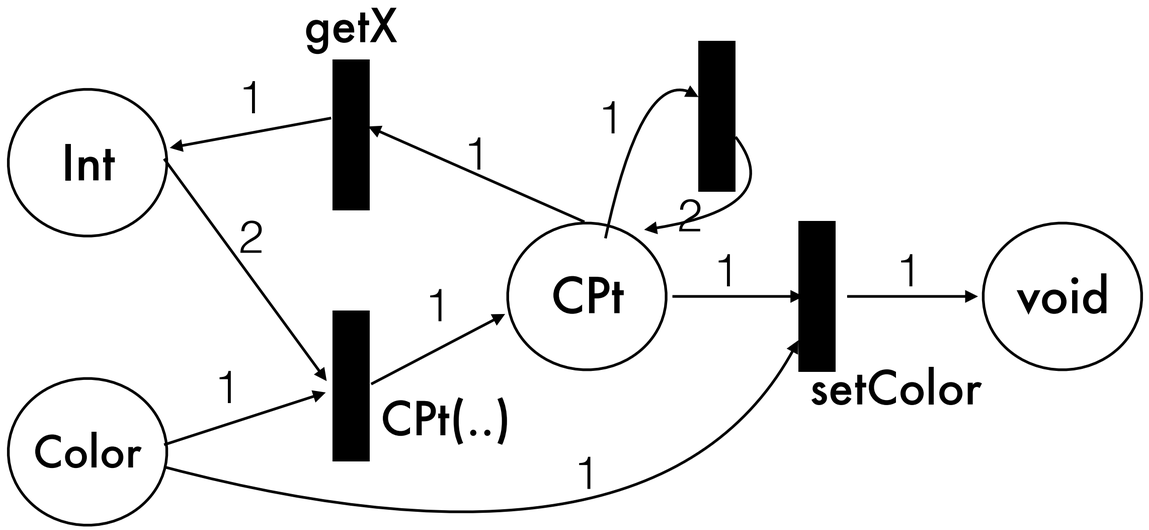
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!
- Therefore, augment Petri net model with **clone transitions**



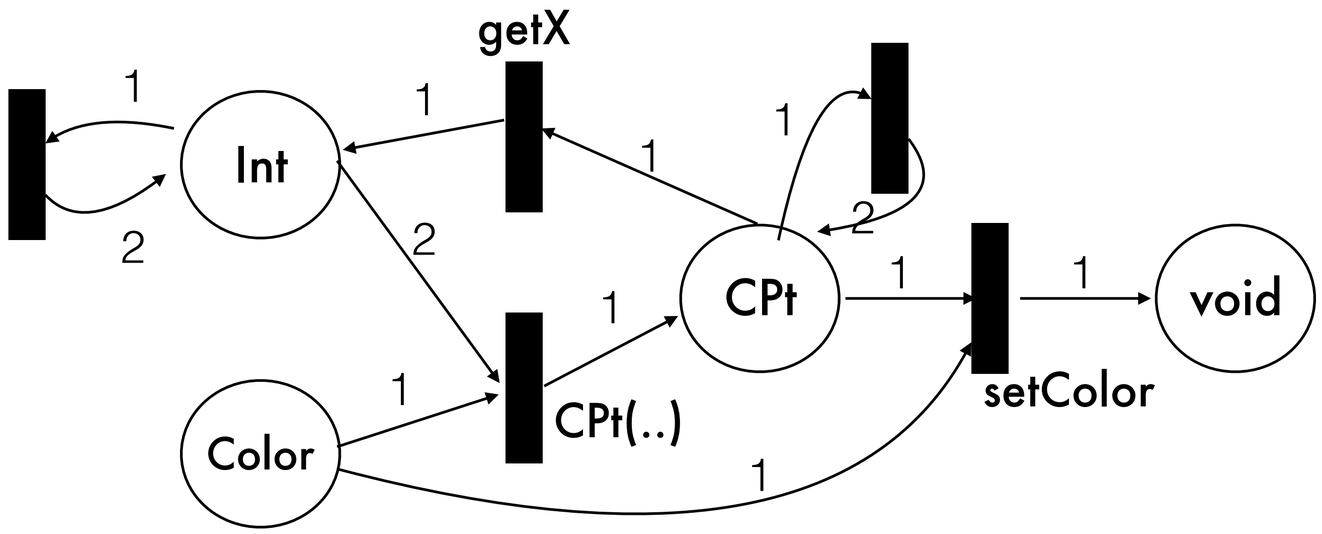
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!
- Therefore, augment Petri net model with **clone transitions**



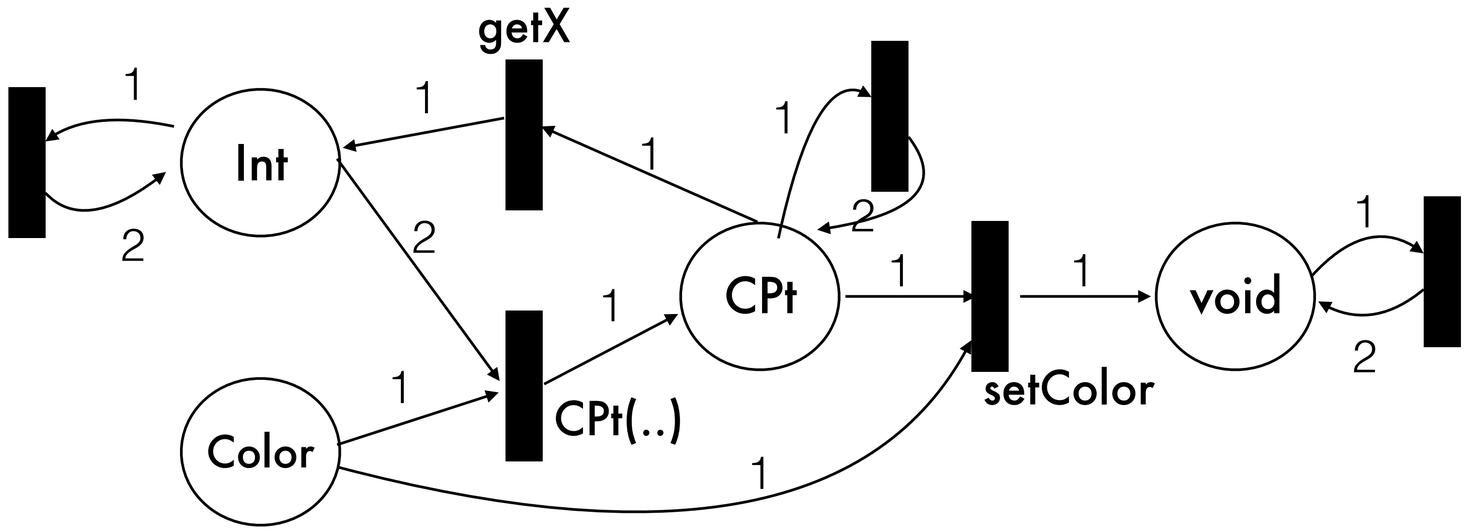
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!
- Therefore, augment Petri net model with **clone transitions**



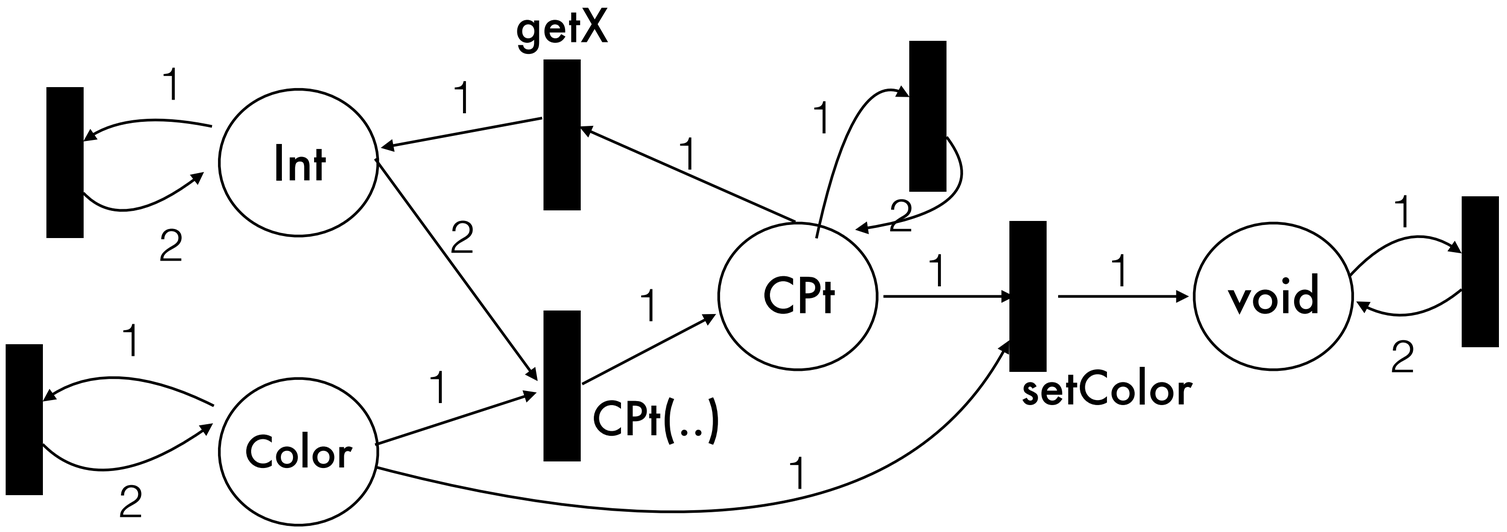
Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!
- Therefore, augment Petri net model with **clone transitions**



Clone transitions

- Our construction so far views objects as“resources”– every method “consumes” and “produces” objects
- But in conventional languages, we can reuse objects!
- Therefore, augment Petri net model with **clone transitions**



SyPet architecture

APIs

Construct Petri net

Init/target markings

Signature $\tau_1 \rightarrow \tau_2$

Reachability analysis

Candidate Sketch

Sketch Completion

Candidate Program

JUnit

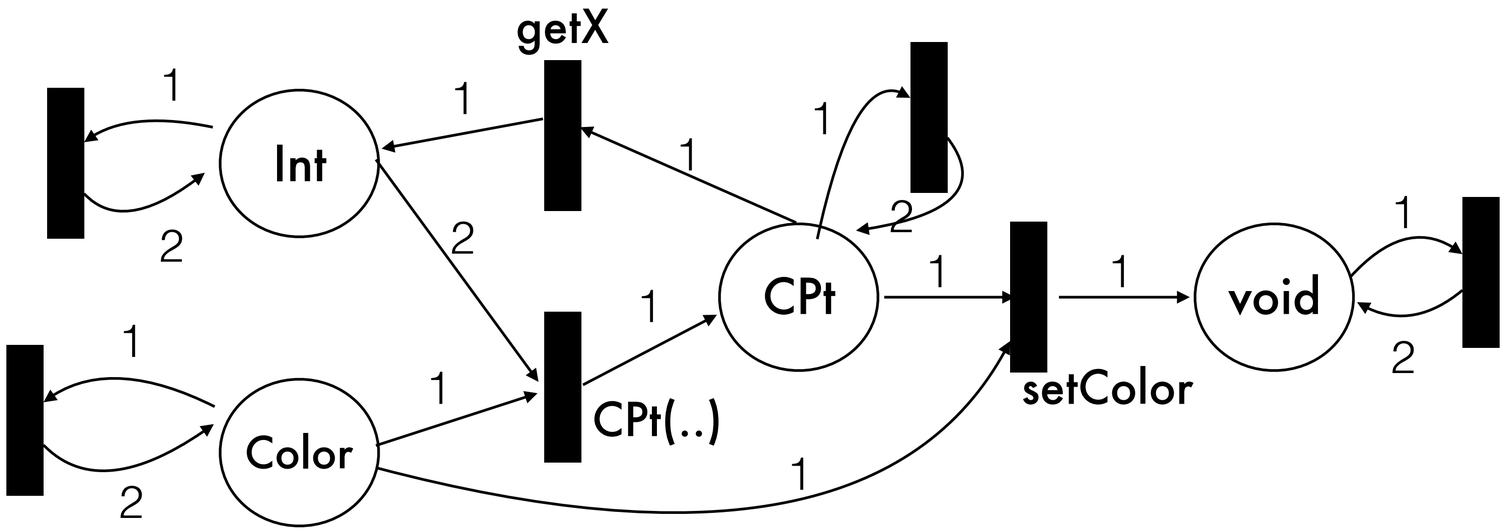
Check Candidate

Java

Backtrack

Initial and target markings

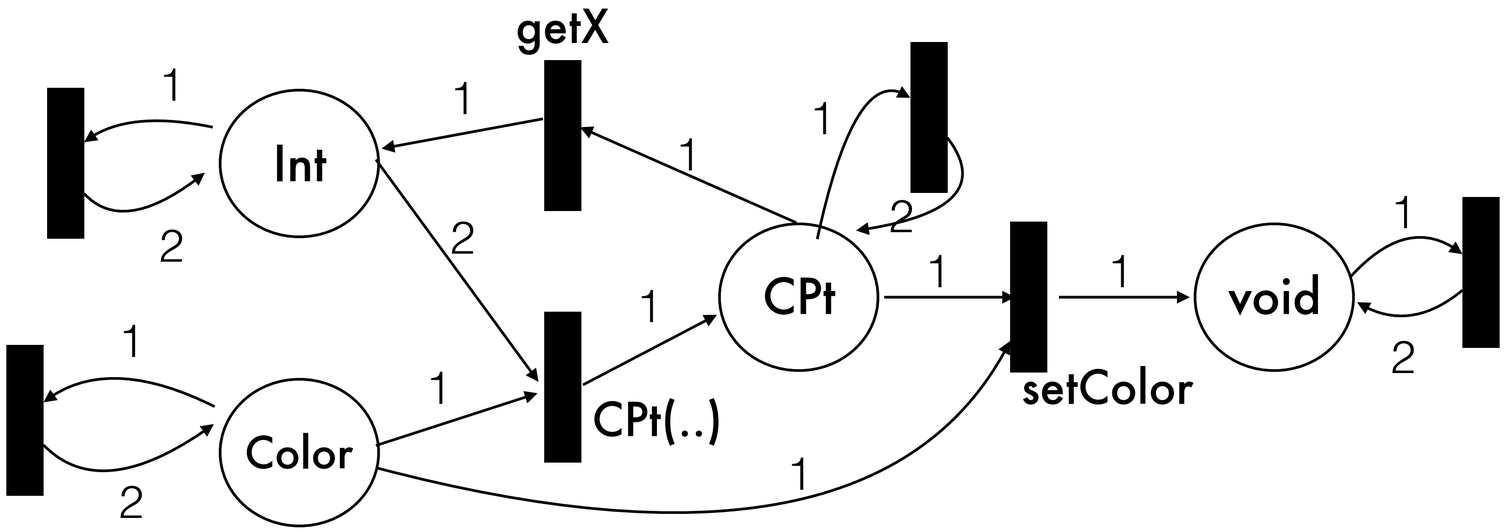
Use signature to determine initial and target markings of Petri net



Initial and target markings

Use signature to determine initial and target markings of Petri net

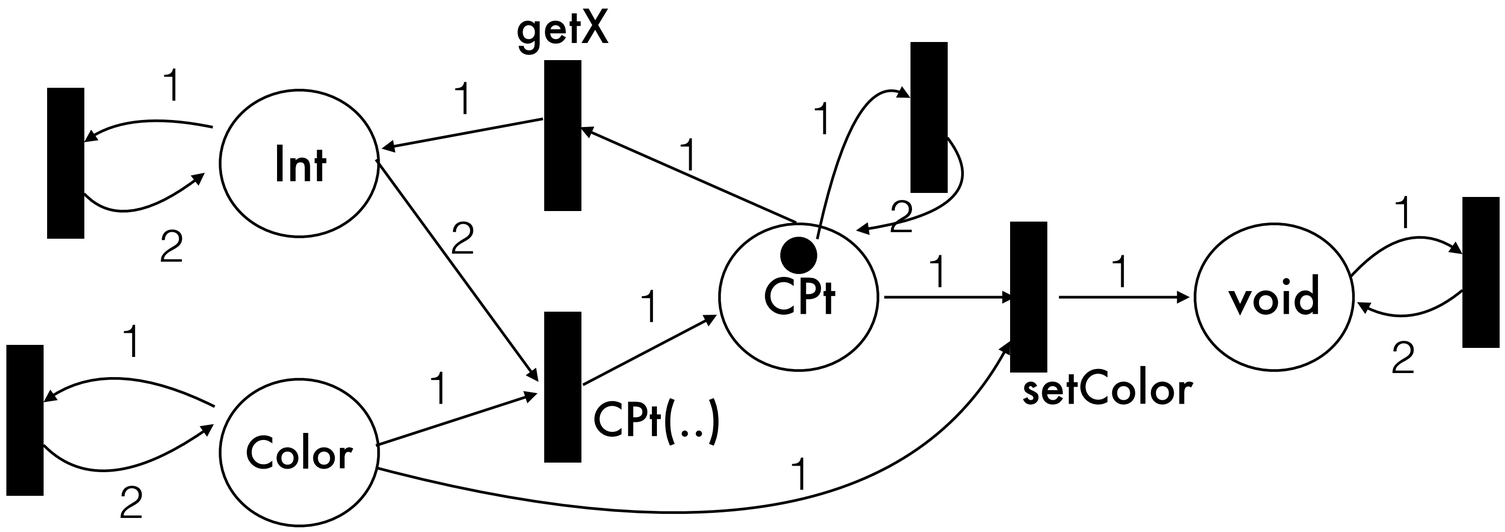
```
CPt shift (CPt p, Int shiftX, Int shiftY)
```



Initial and target markings

Use signature to determine initial and target markings of Petri net

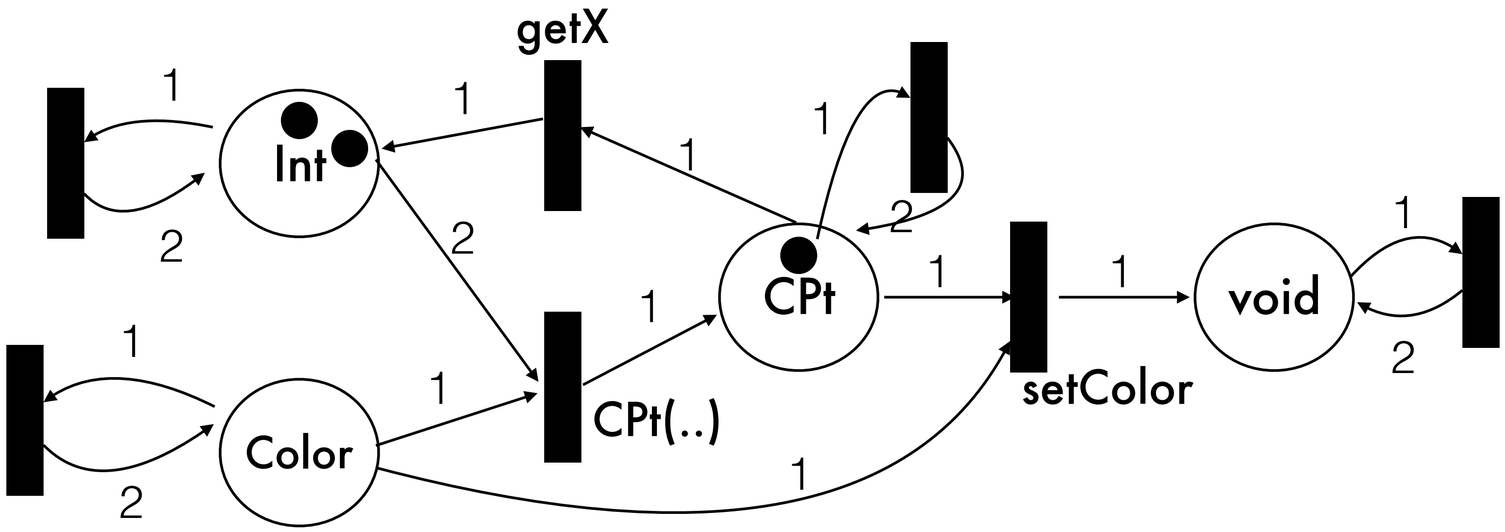
```
CPt shift (CPt p, Int shiftX, Int shiftY)
```



Initial and target markings

Use signature to determine initial and target markings of Petri net

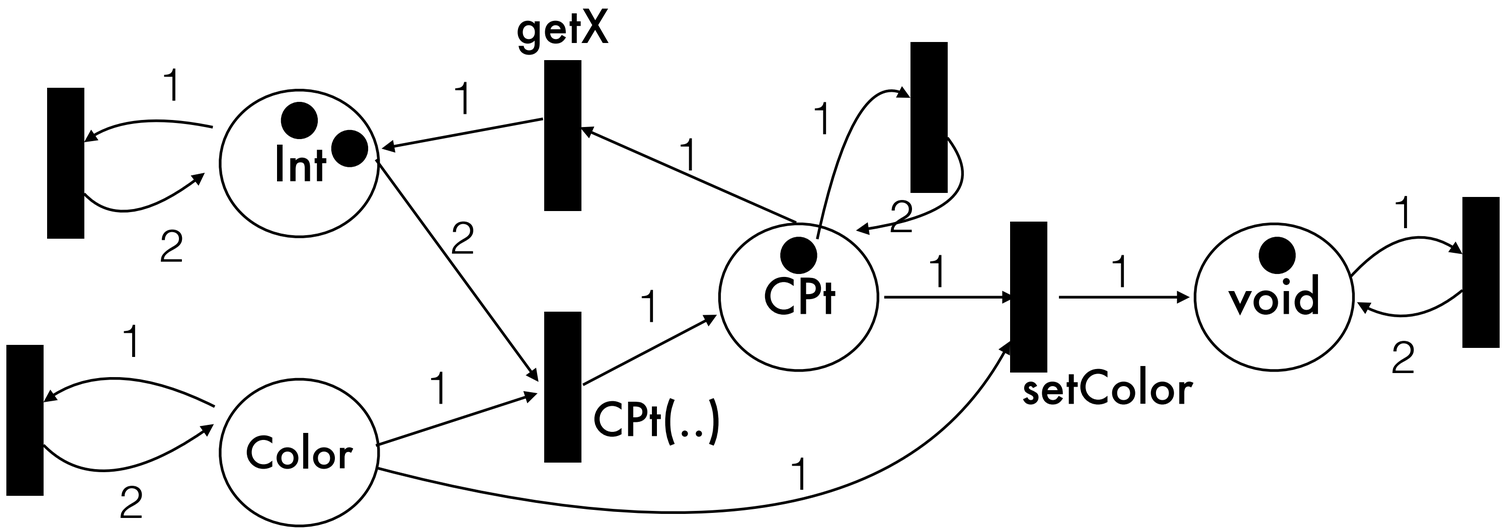
```
CPt shift (CPt p, Int shiftX, Int shiftY)
```



Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

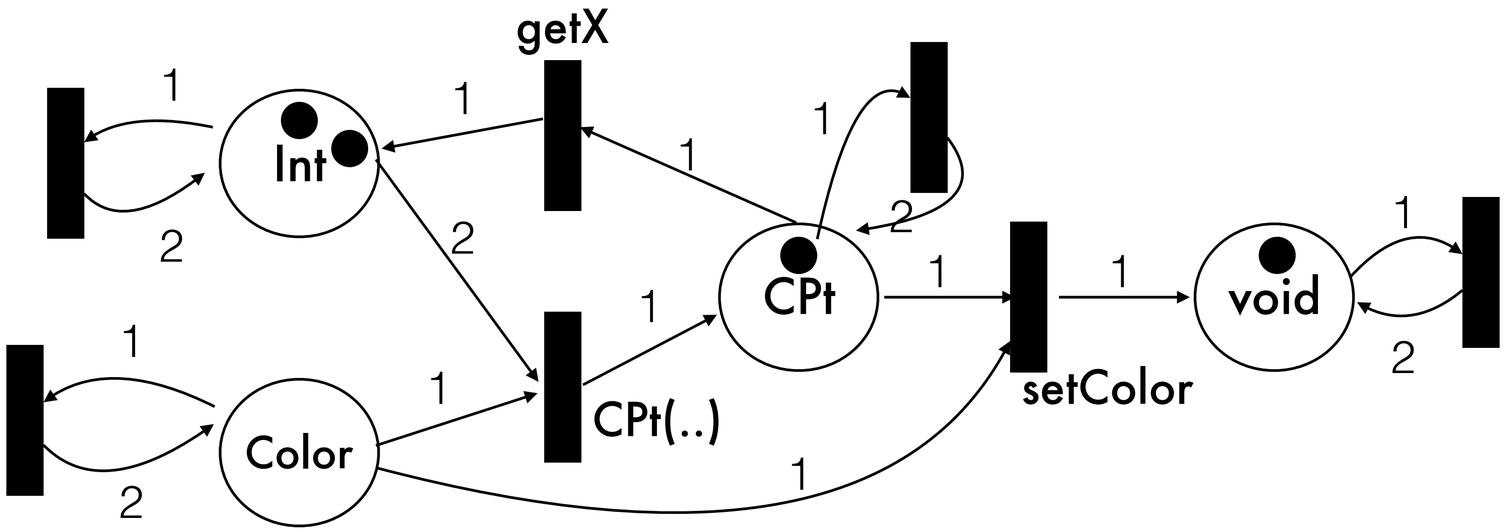


Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

Target marking:



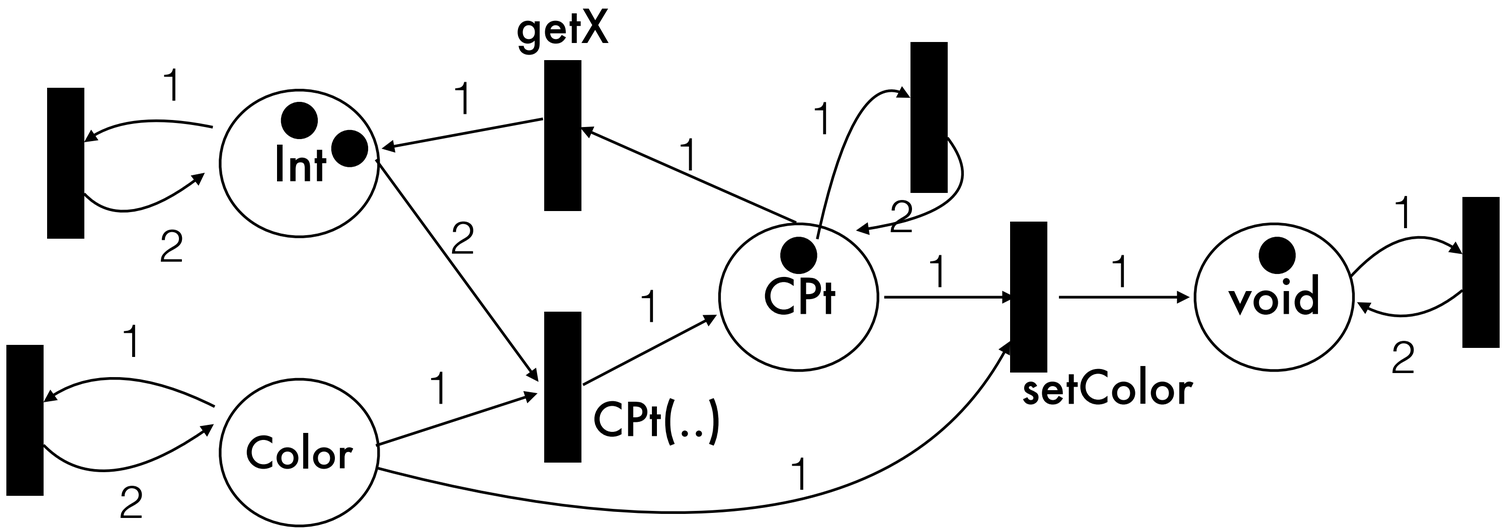
Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

Target marking:

Cpt = 1



Initial and target markings

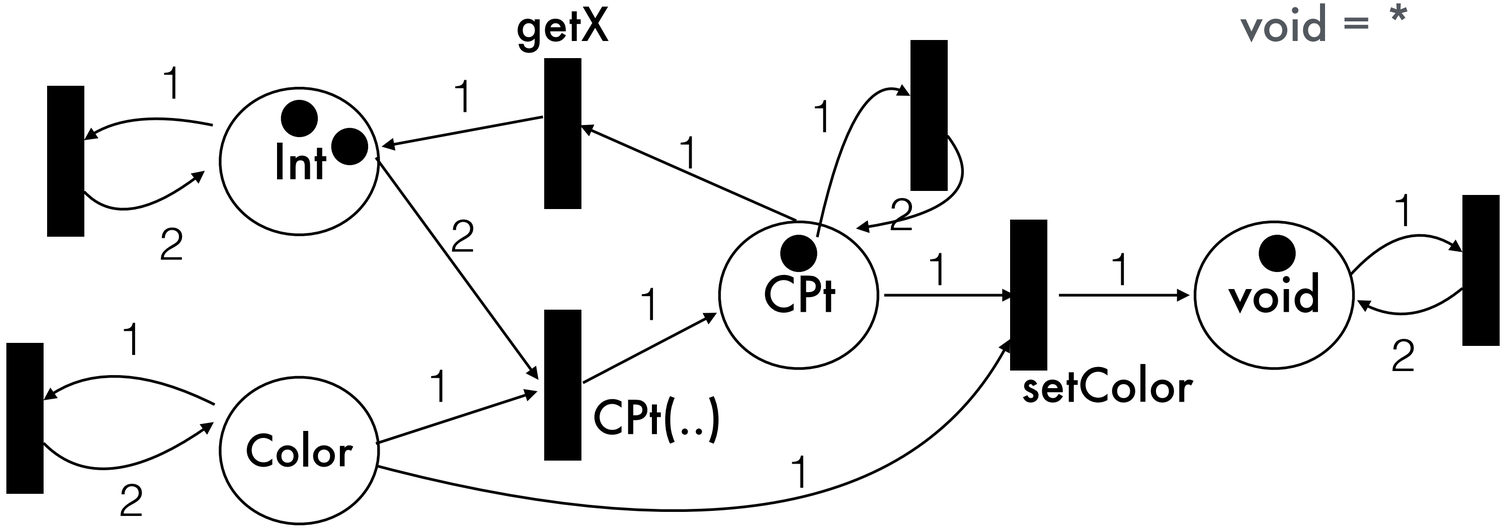
Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

Target marking:

Cpt = 1

void = *



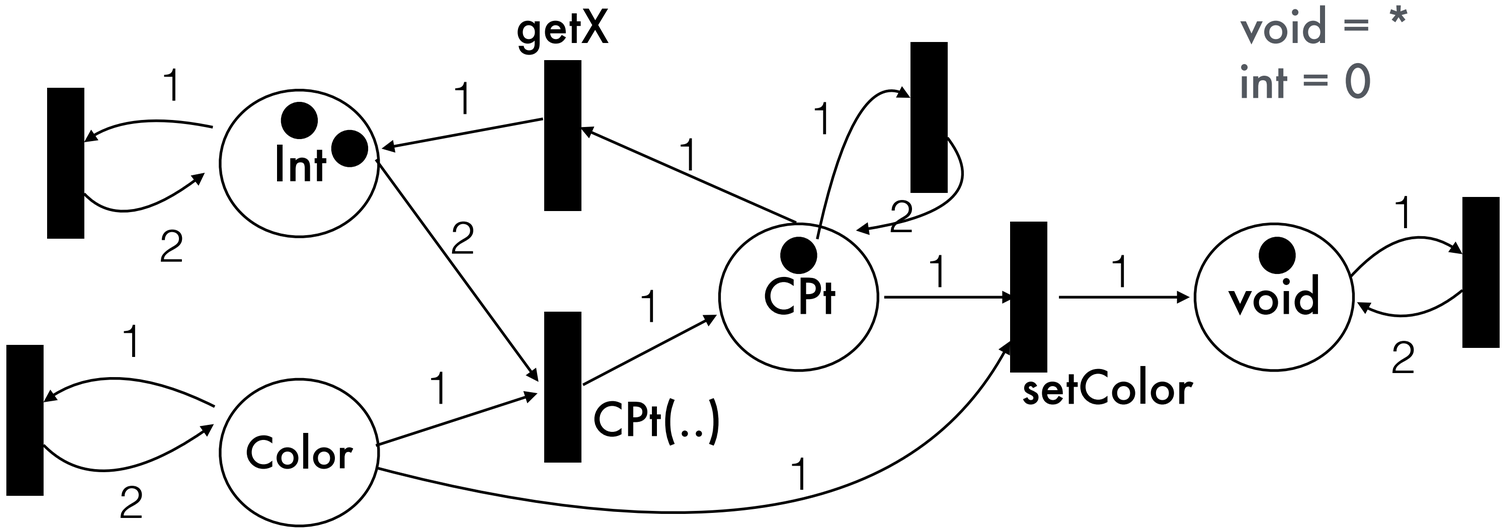
Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

Target marking:

Cpt = 1
void = *
int = 0



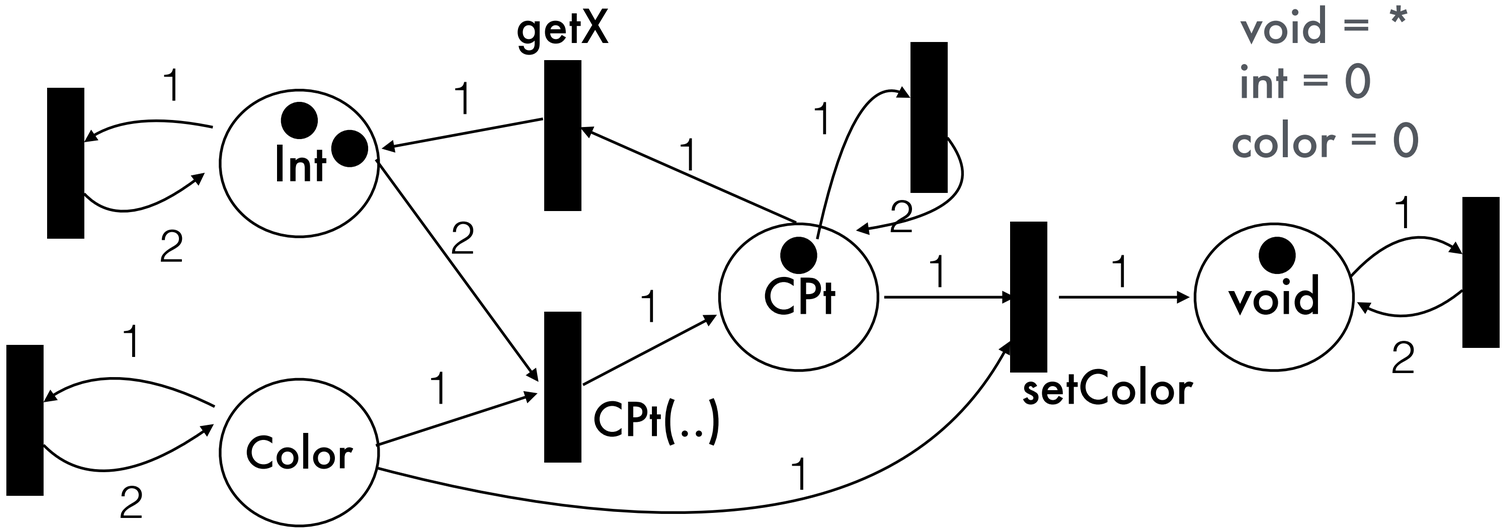
Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```

Target marking:

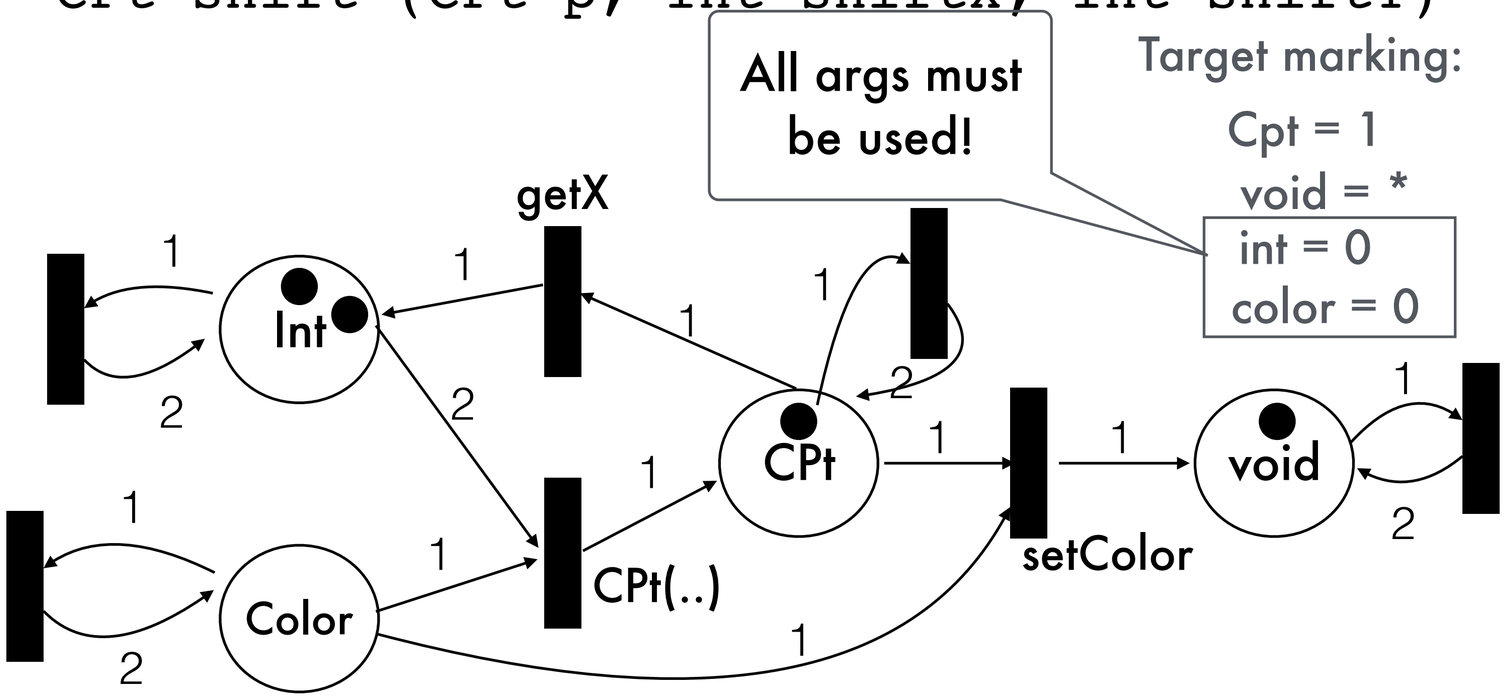
Cpt = 1
void = *
int = 0
color = 0



Initial and target markings

Use signature to determine initial and target markings of Petri net

```
CPt shift (CPt p, Int shiftX, Int shiftY)
```



Exercise 1: Building a Petri Net

```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

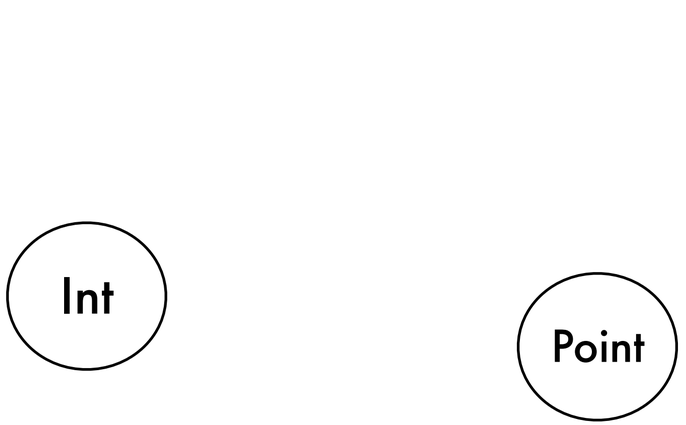
```
class MyPoint {
  MyPoint(int x, int y);
  int getX();
  int getY();
}
```

- Build a petri net with the classes Point and MyPoint:
  - Hint: What are the places (i.e., types)?
  - Hint: What are the transitions (i.e., methods)?
  - Hint: Don’t forget the clone edges!

Exercise 1: Solution

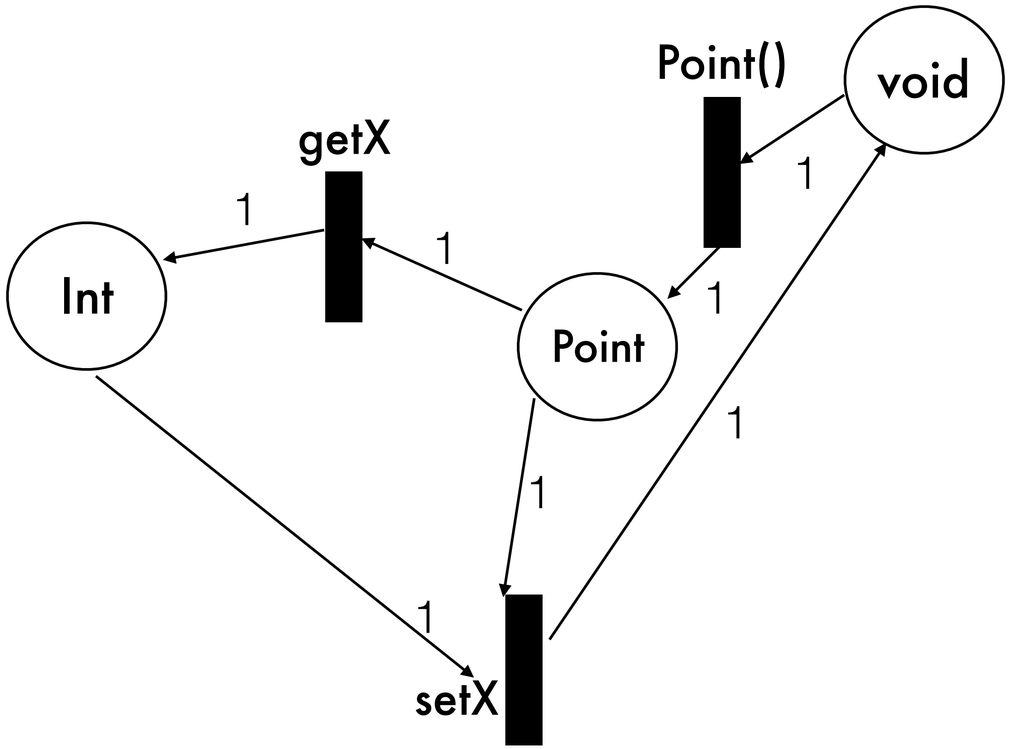
```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

Exercise 1: Solution



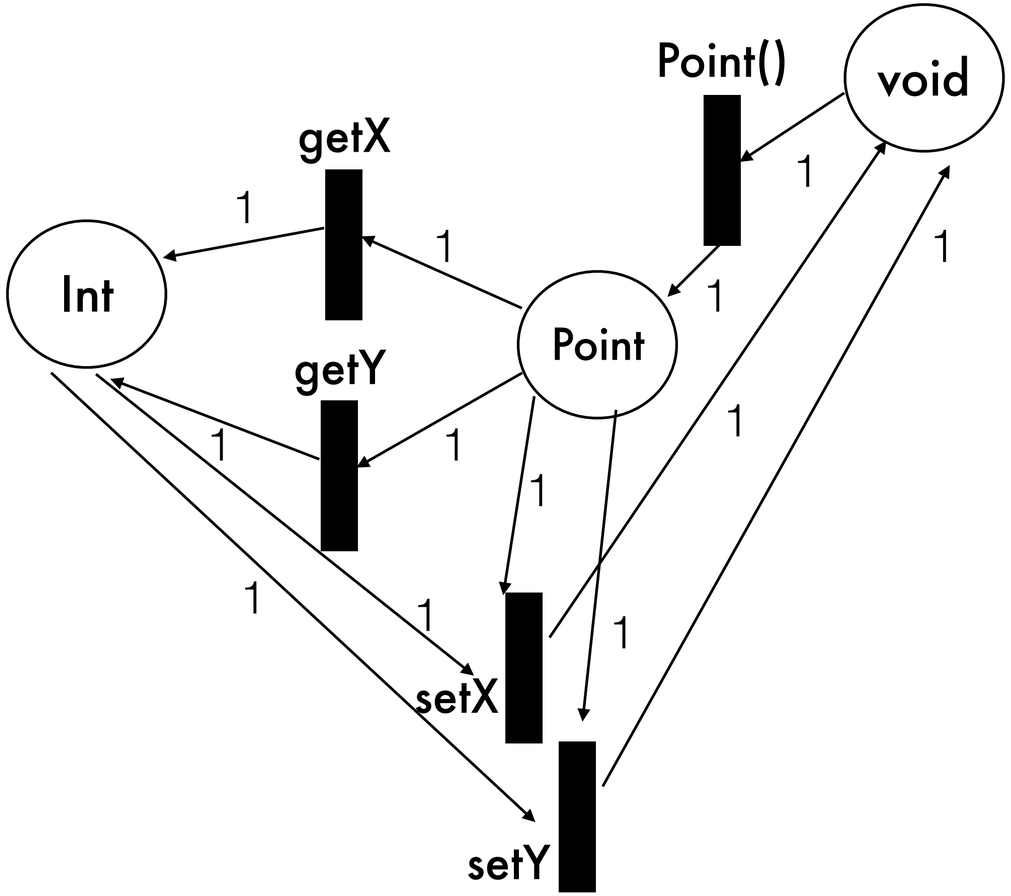

```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

Exercise 1: Solution



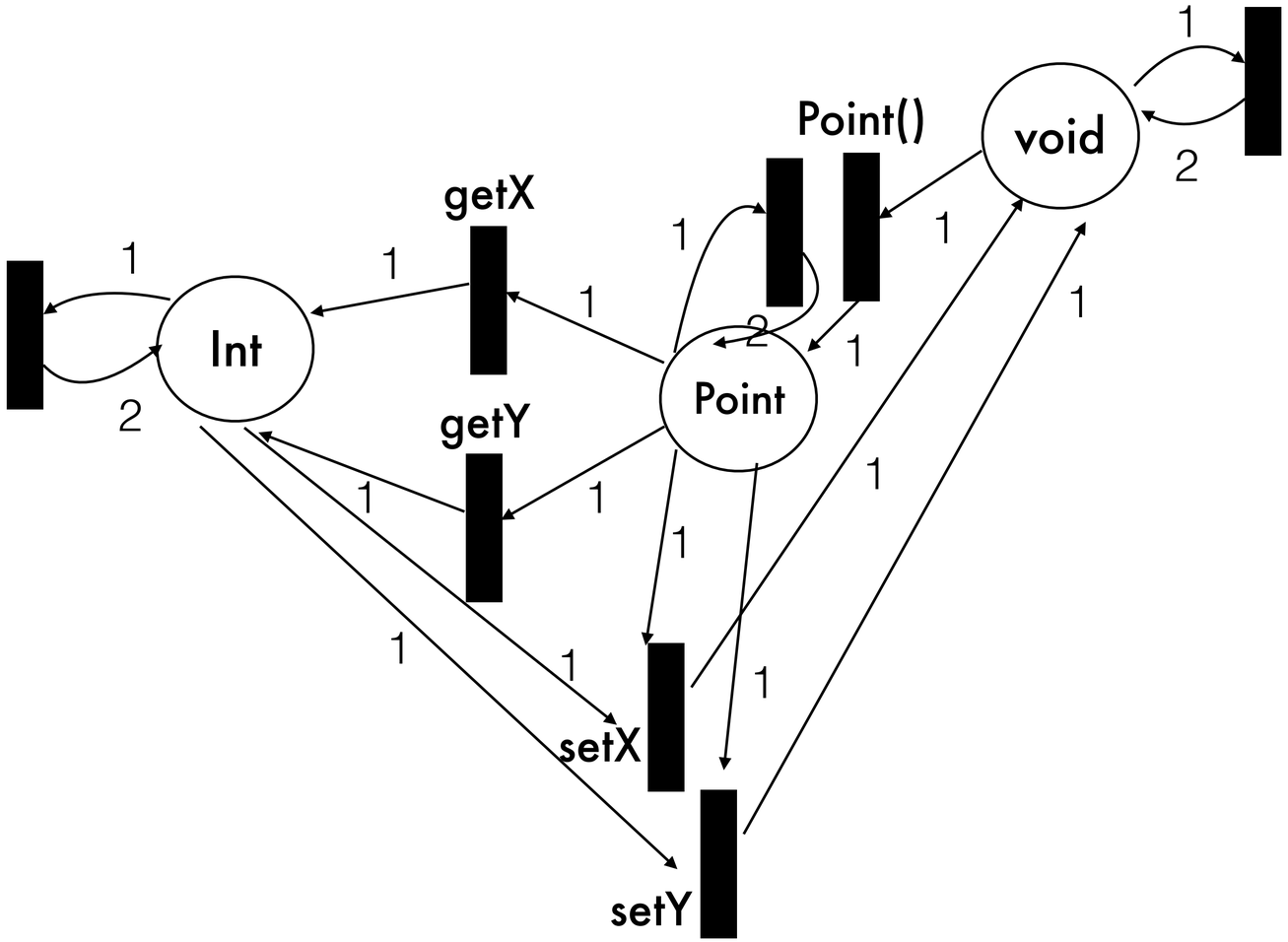

```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

Exercise 1: Solution




```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

Exercise 1: Solution

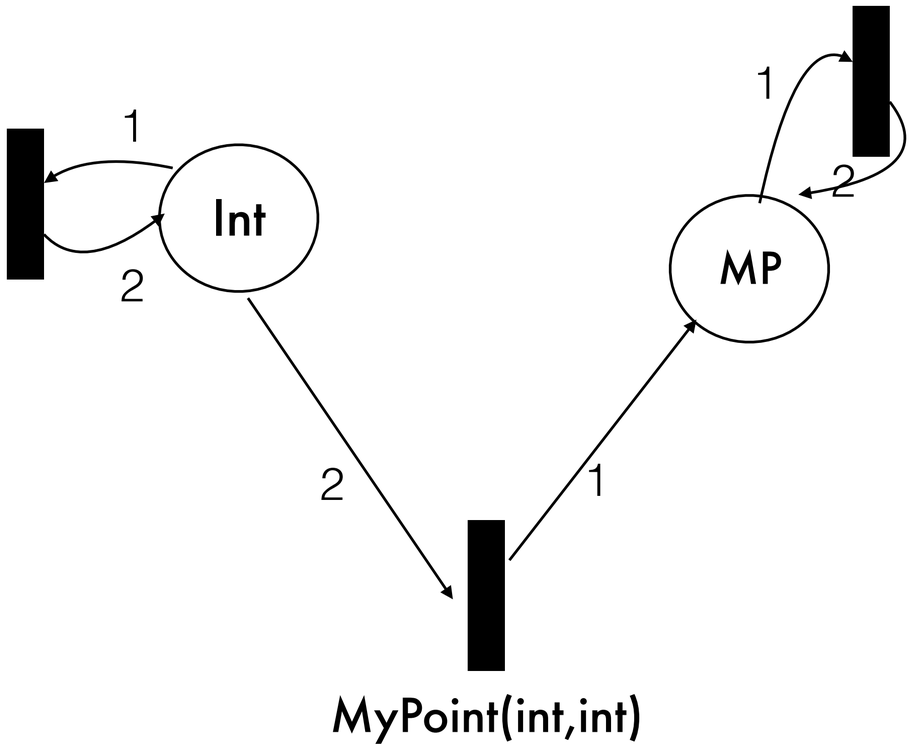



```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
 }
```

Exercise 1: Solution

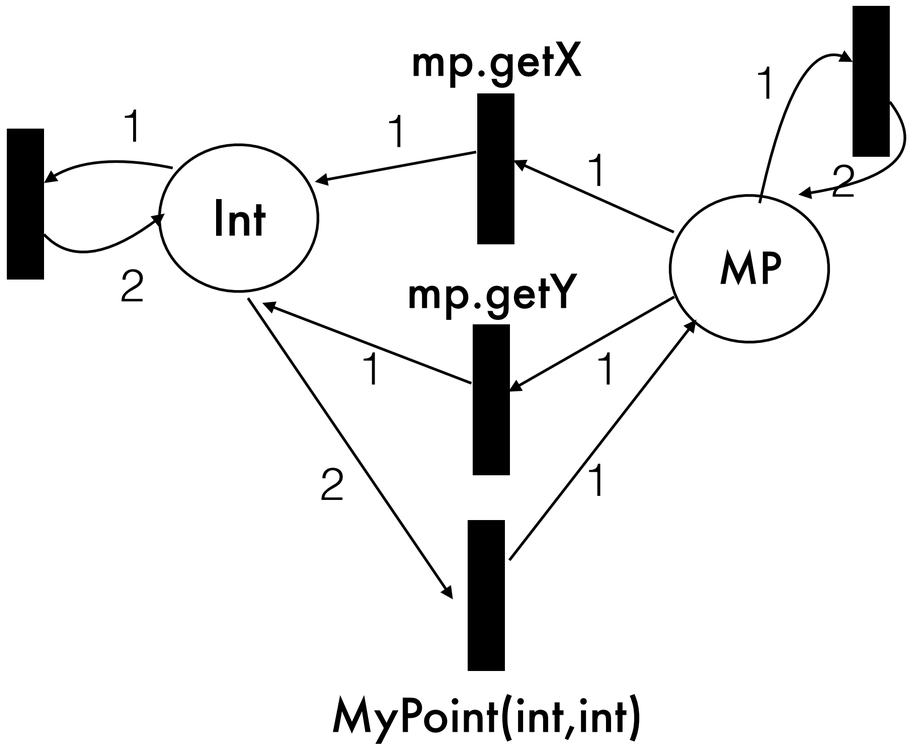
```
class MyPoint {
  MyPoint(int x, int y);
  int getX();
  int getY();
}
```

Exercise 1: Solution



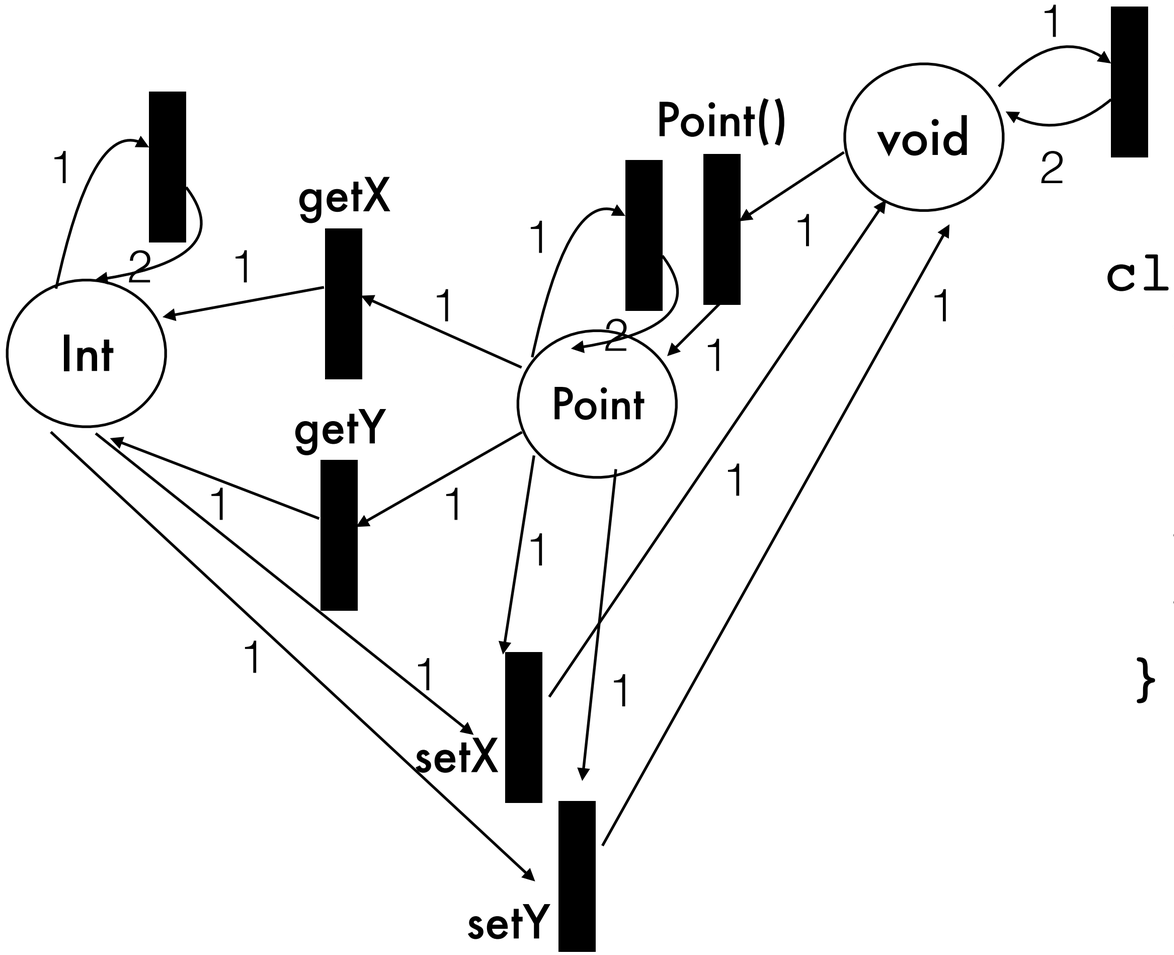

```
class MyPoint {
  MyPoint(int x, int y);
  int getX();
  int getY();
}
```

Exercise 1: Solution



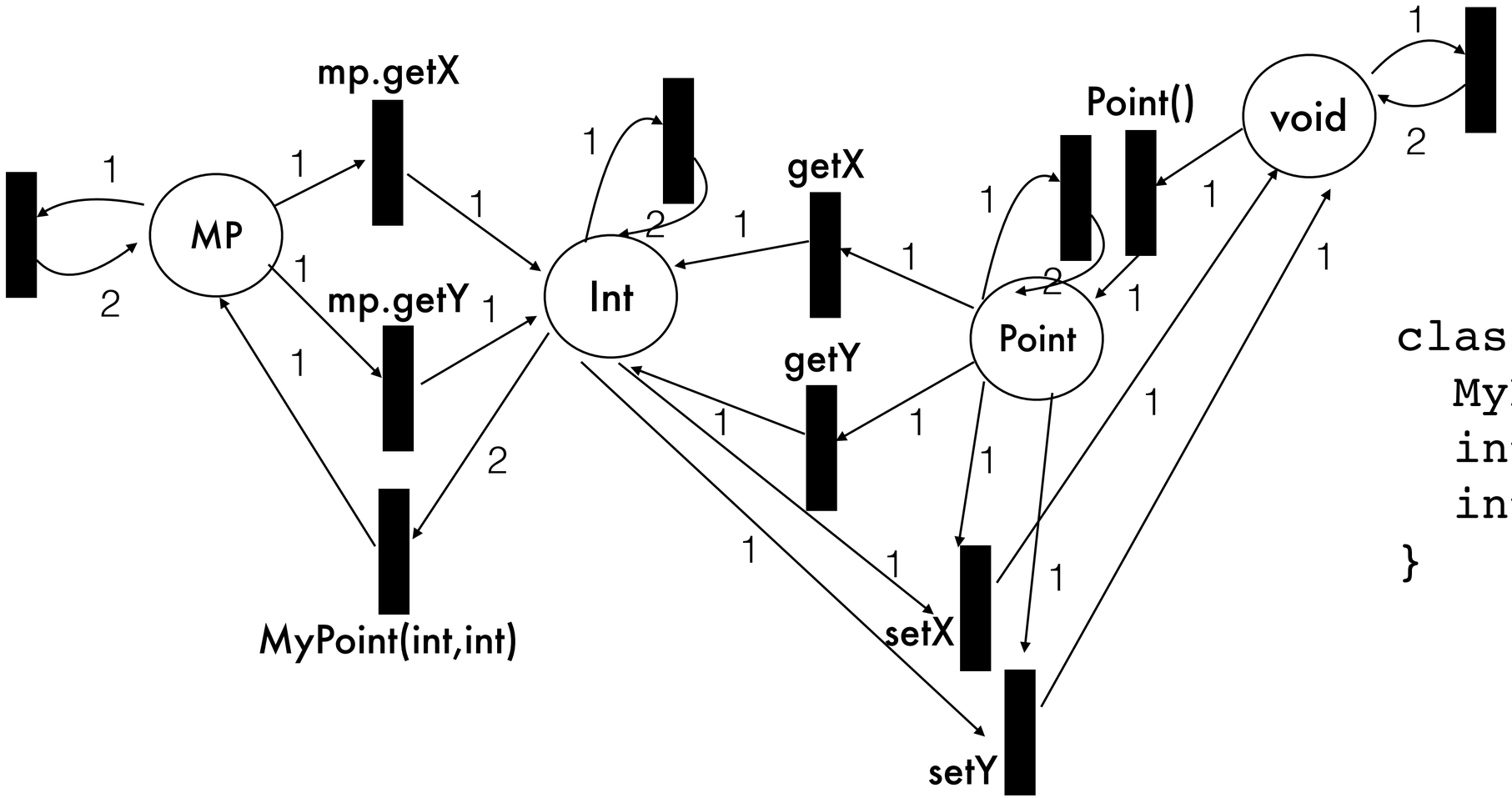

```
class MyPoint {
  MyPoint(int x, int y);
  int getX();
  int getY();
}
```

Exercise 1: Solution



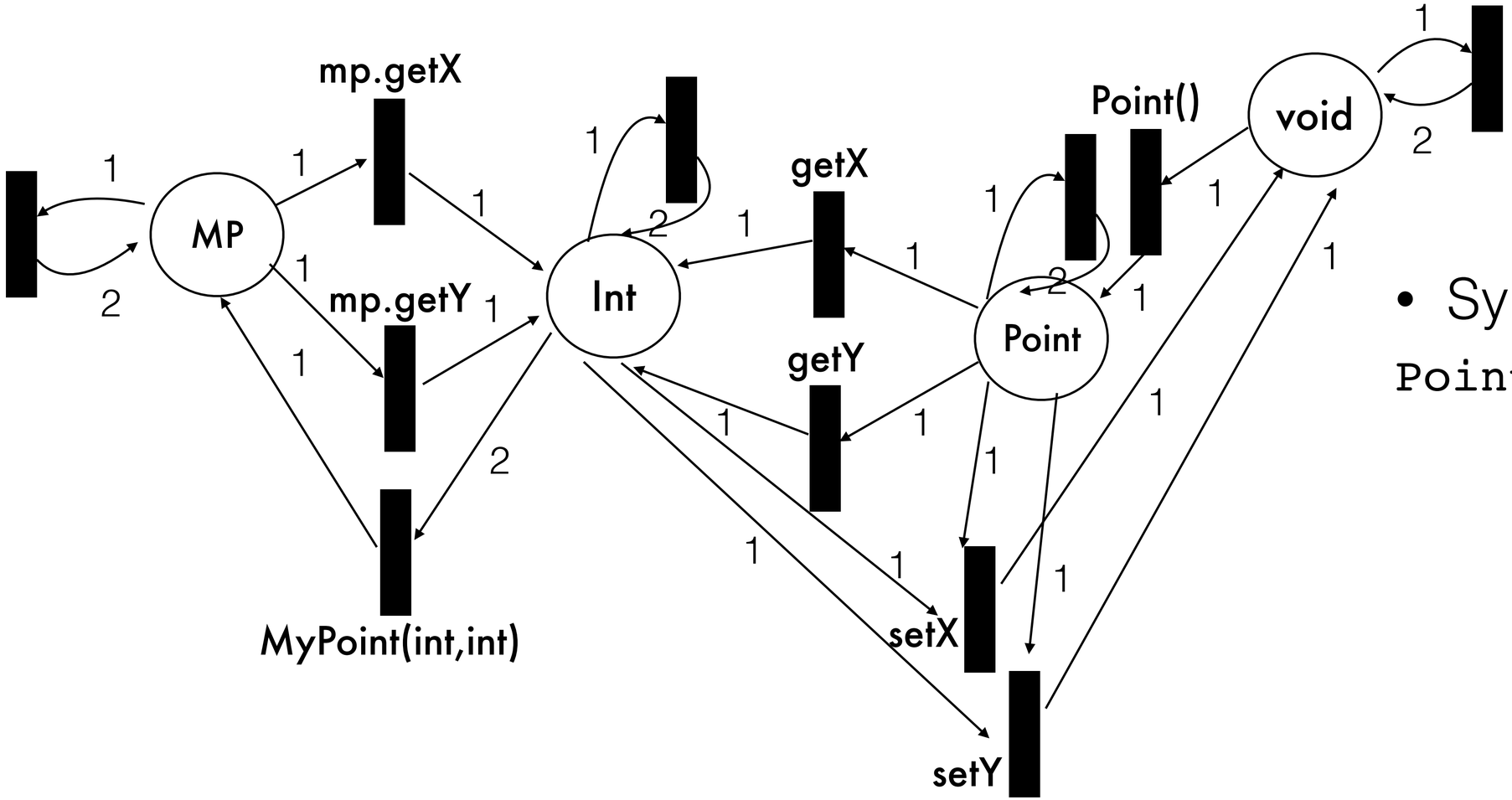

```
class Point {
  Point();
  int getX();
  int getY();
  void setX(int);
  void setY(int);
}
```

Exercise 1: Solution




```
class MyPoint {
  MyPoint(int x, int y);
  int getX();
  int getY();
}
```

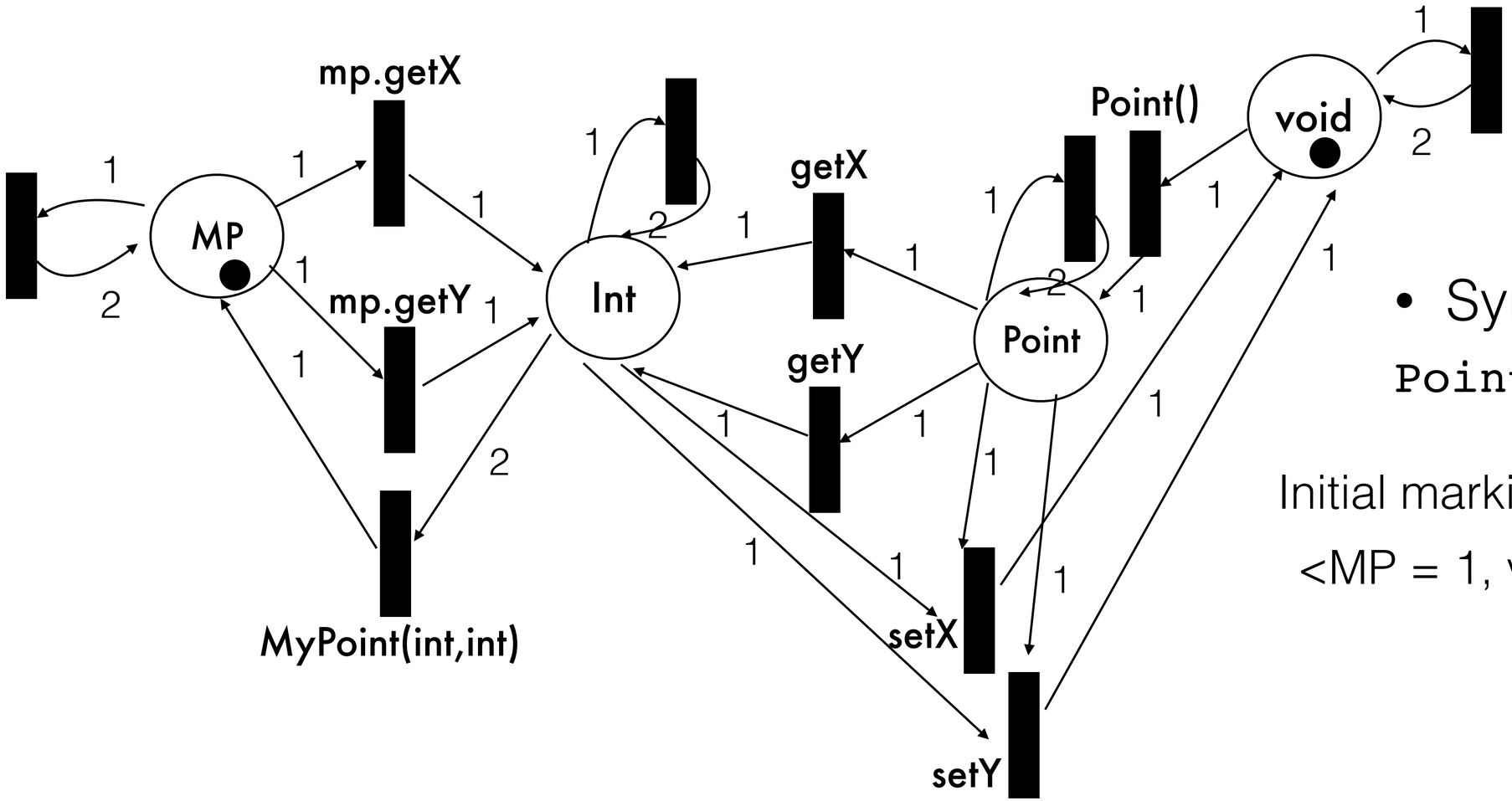
What is the initial marking?



- Synthesize this function:

```
Point convert(Mypoint pt)
```

What is the initial marking?



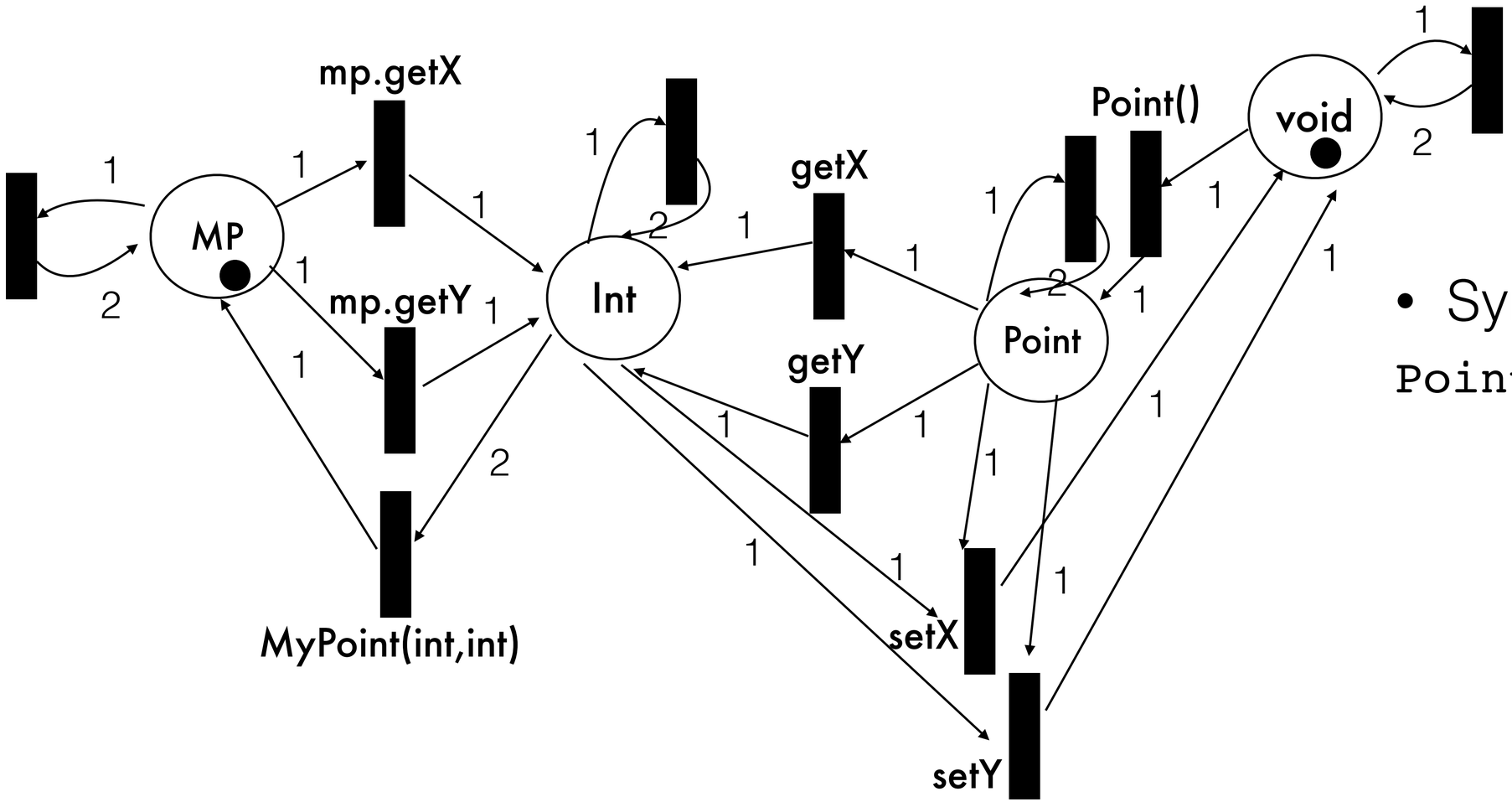
- Synthesize this function:

`Point convert(Mypoint pt)`

Initial marking:

<MP = 1, void = 1, Int = 0, Point = 0>

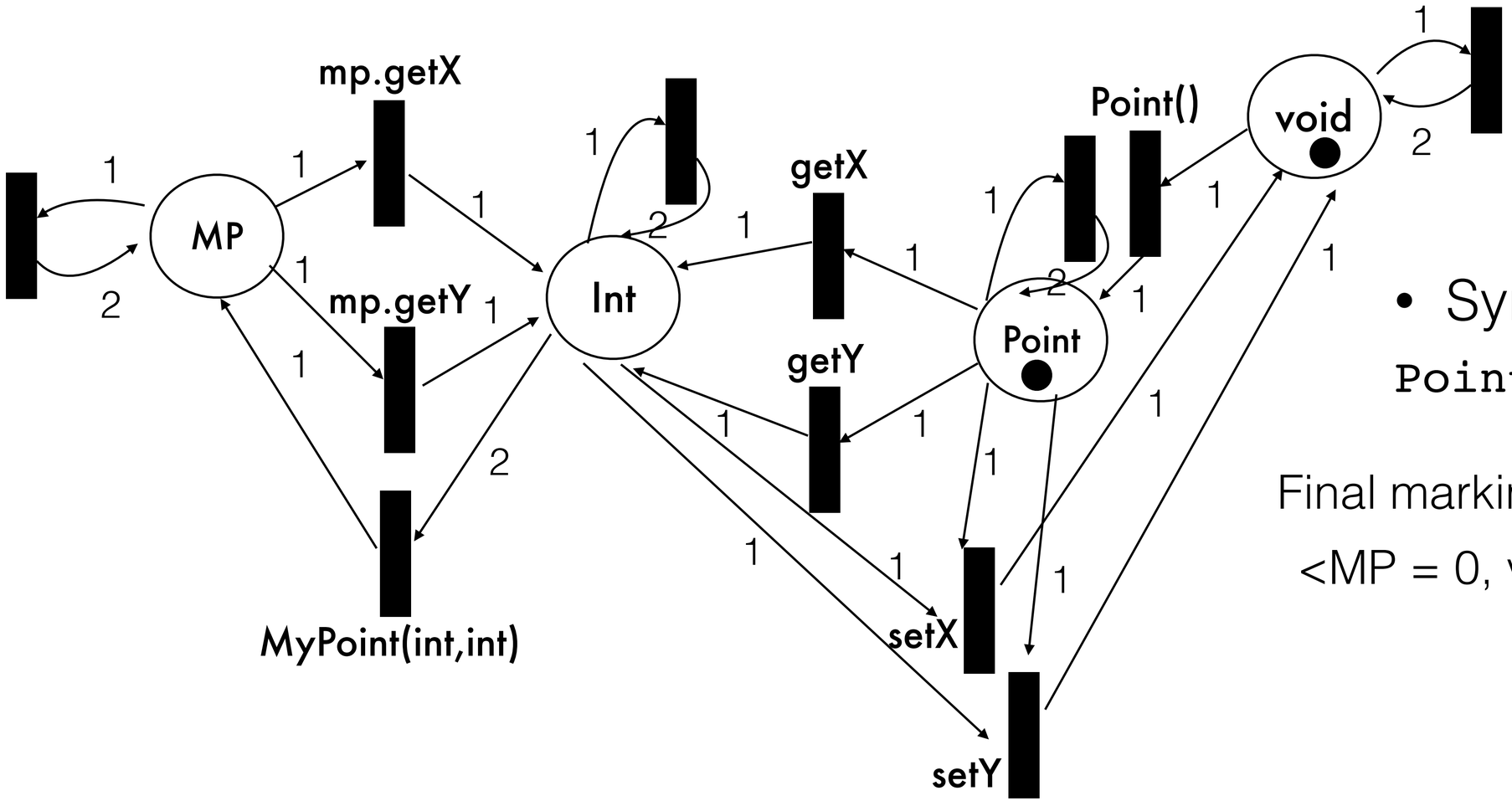
What is the final marking?



- Synthesize this function:

`Point convert(Mypoint pt)`

What is the final marking?



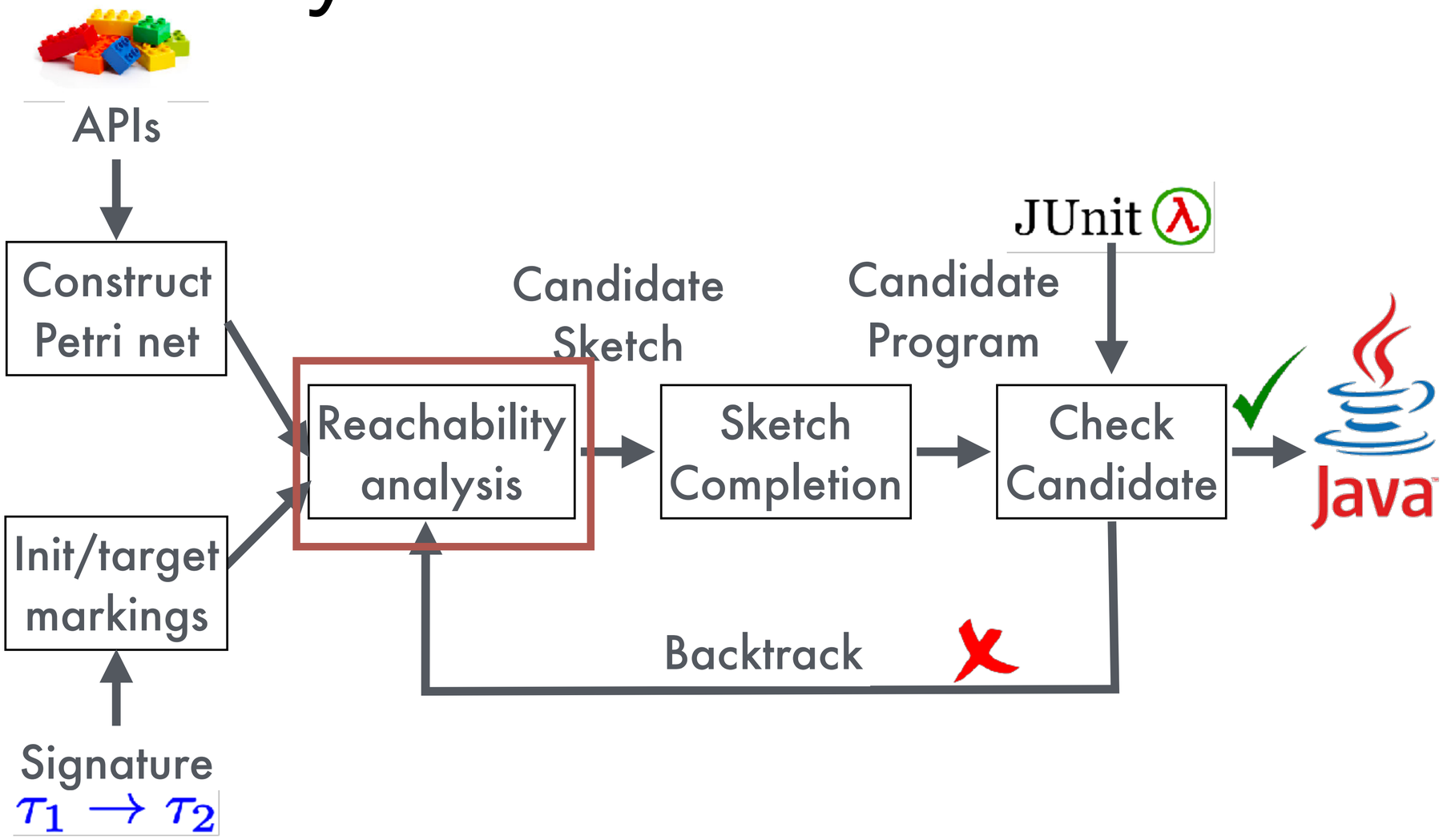
- Synthesize this function:

```
Point convert(Mypoint pt)
```

Final marking:

<MP = 0, void = *, Int = 0, Point = 1>

SyPet Architecture

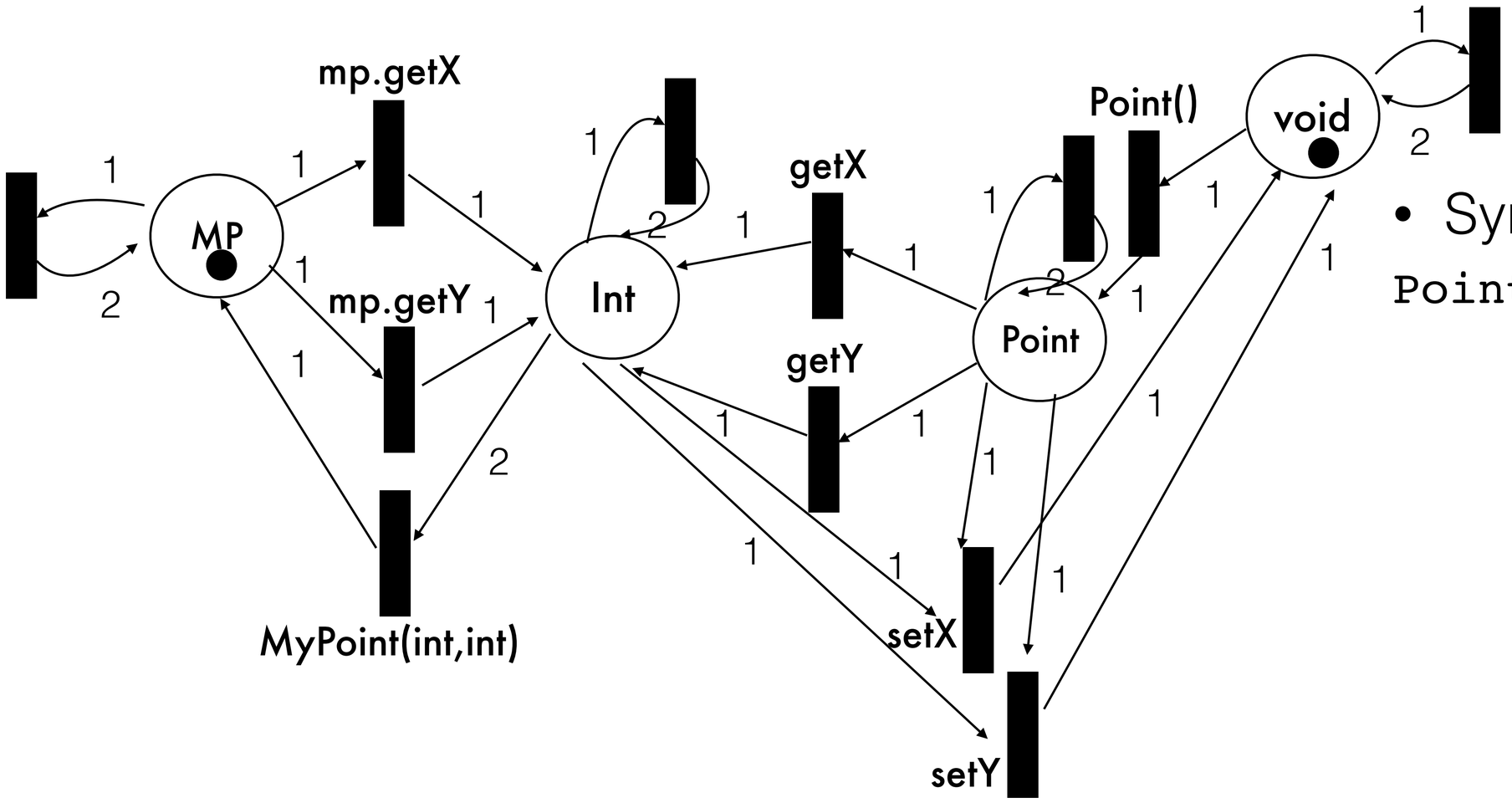


Reachability analysis

All accepting runs of Petri net correspond to method call sequences with desired type signature!

- Need to perform reachability analysis to identify accepting runs of the Petri Net
- Our solution reduces bounded reachability analysis to a SAT problem:
  - Find a reachable path of size *k*
  - Enumerate all reachable paths

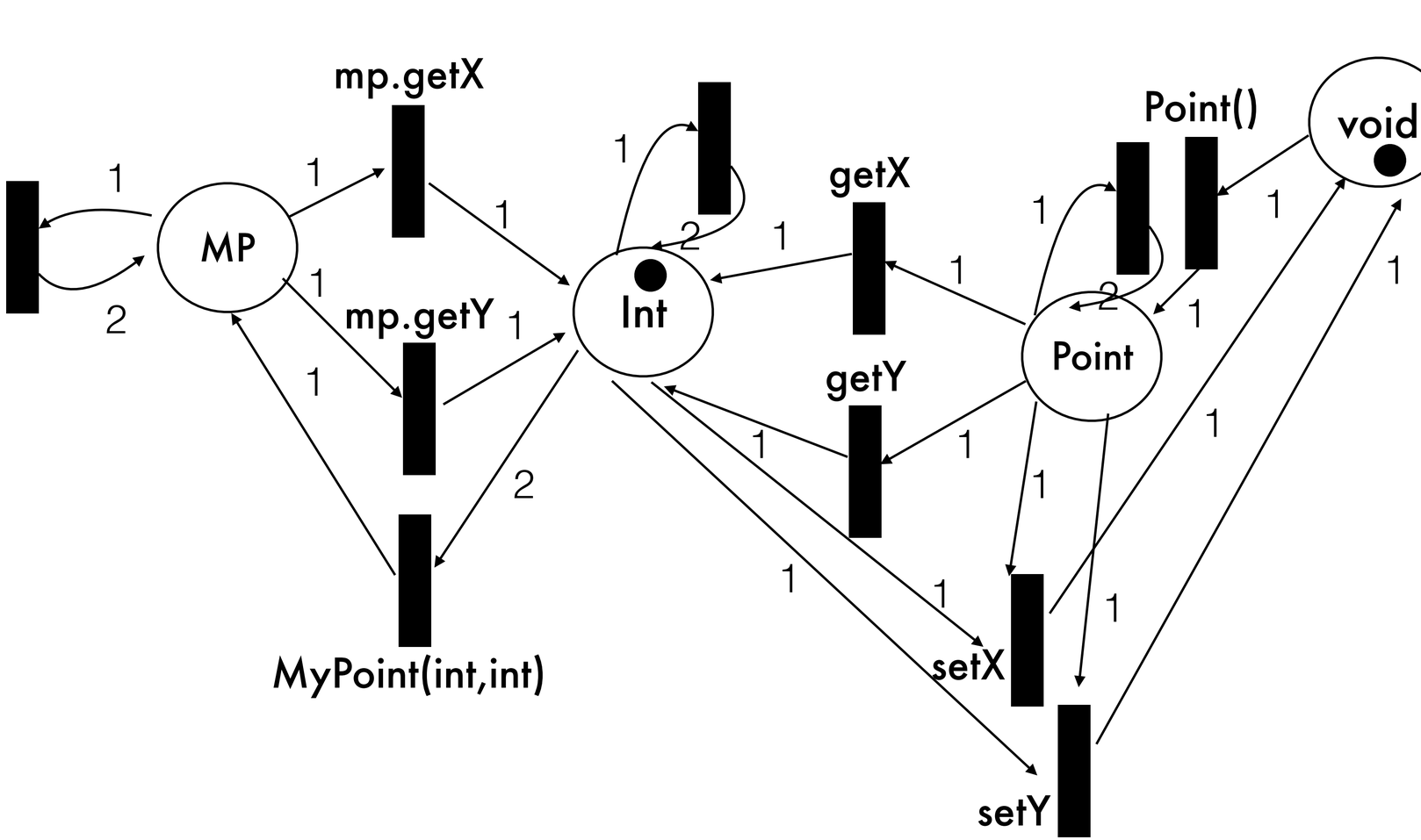
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)
```

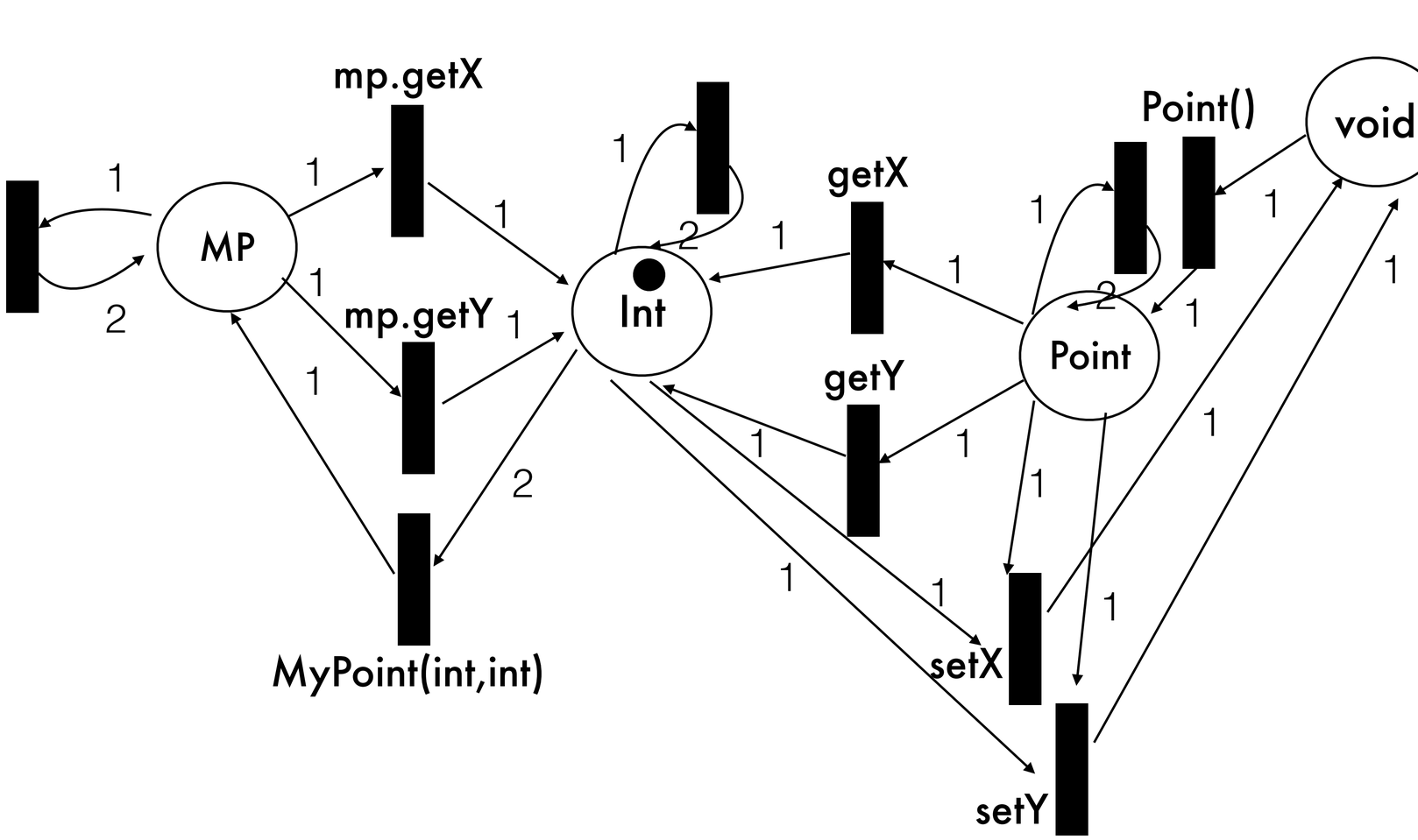
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)

1) mp.getX
```

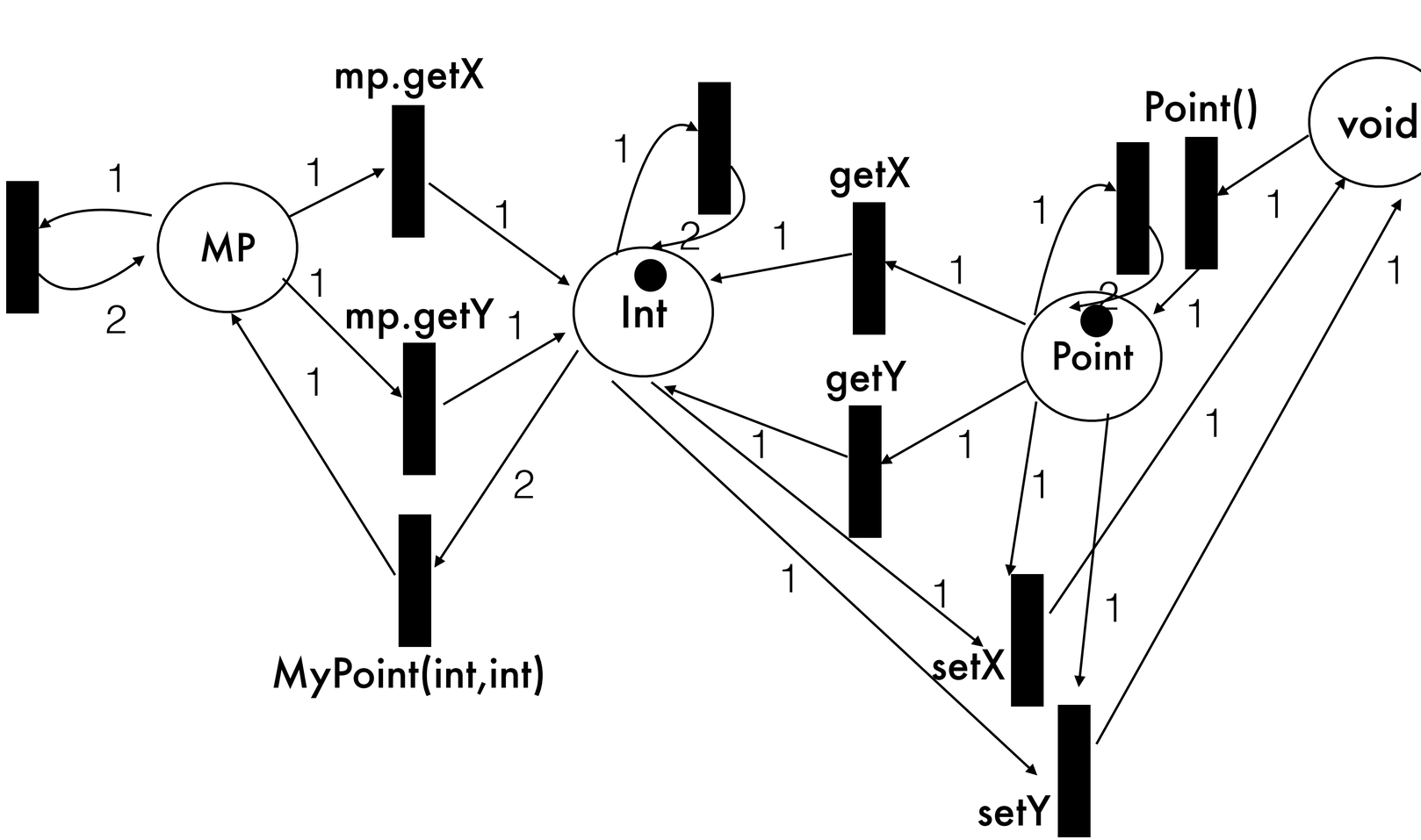
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)

1) mp.getX
2) Point()
```

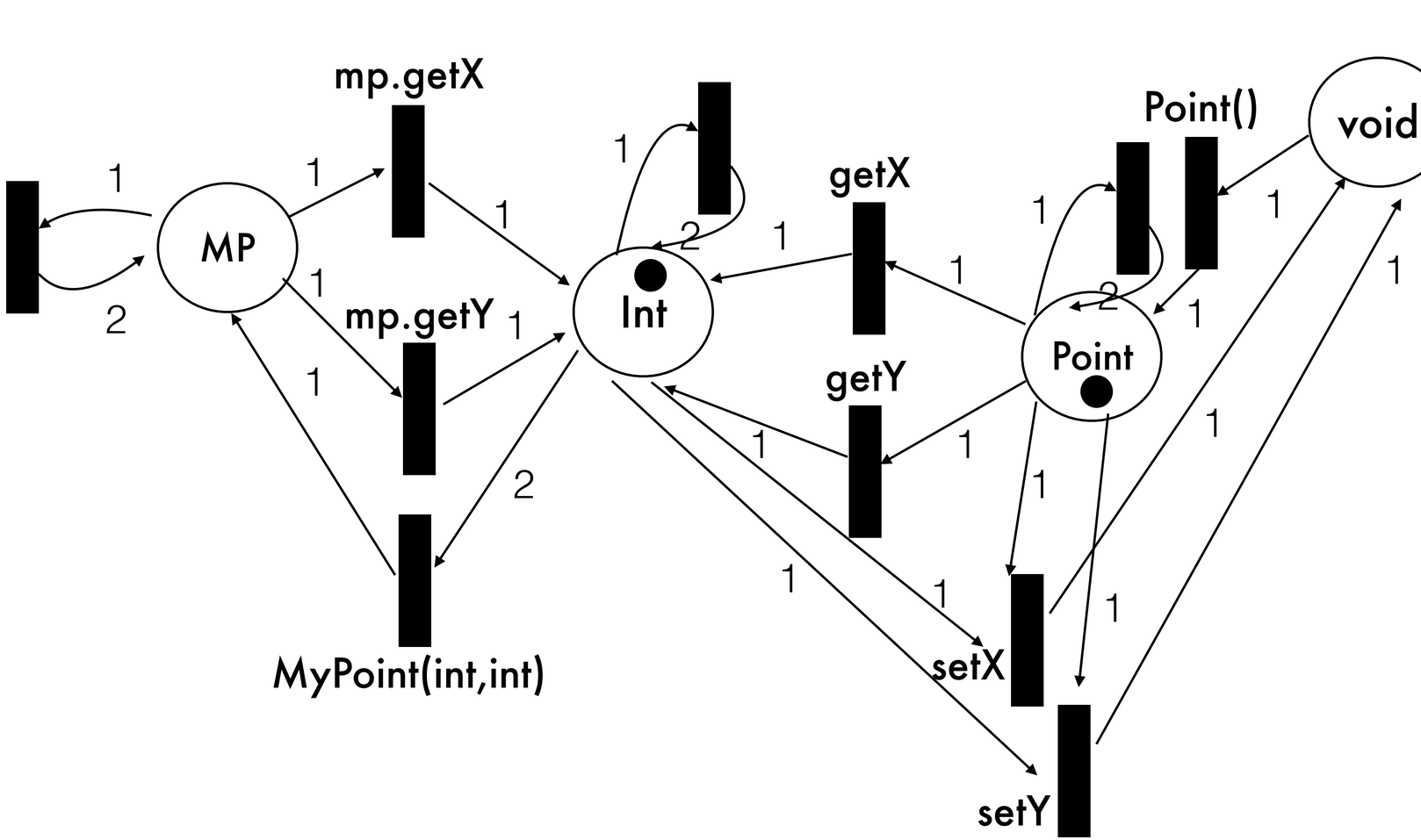
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)

1) mp.getX
2) Point()
```

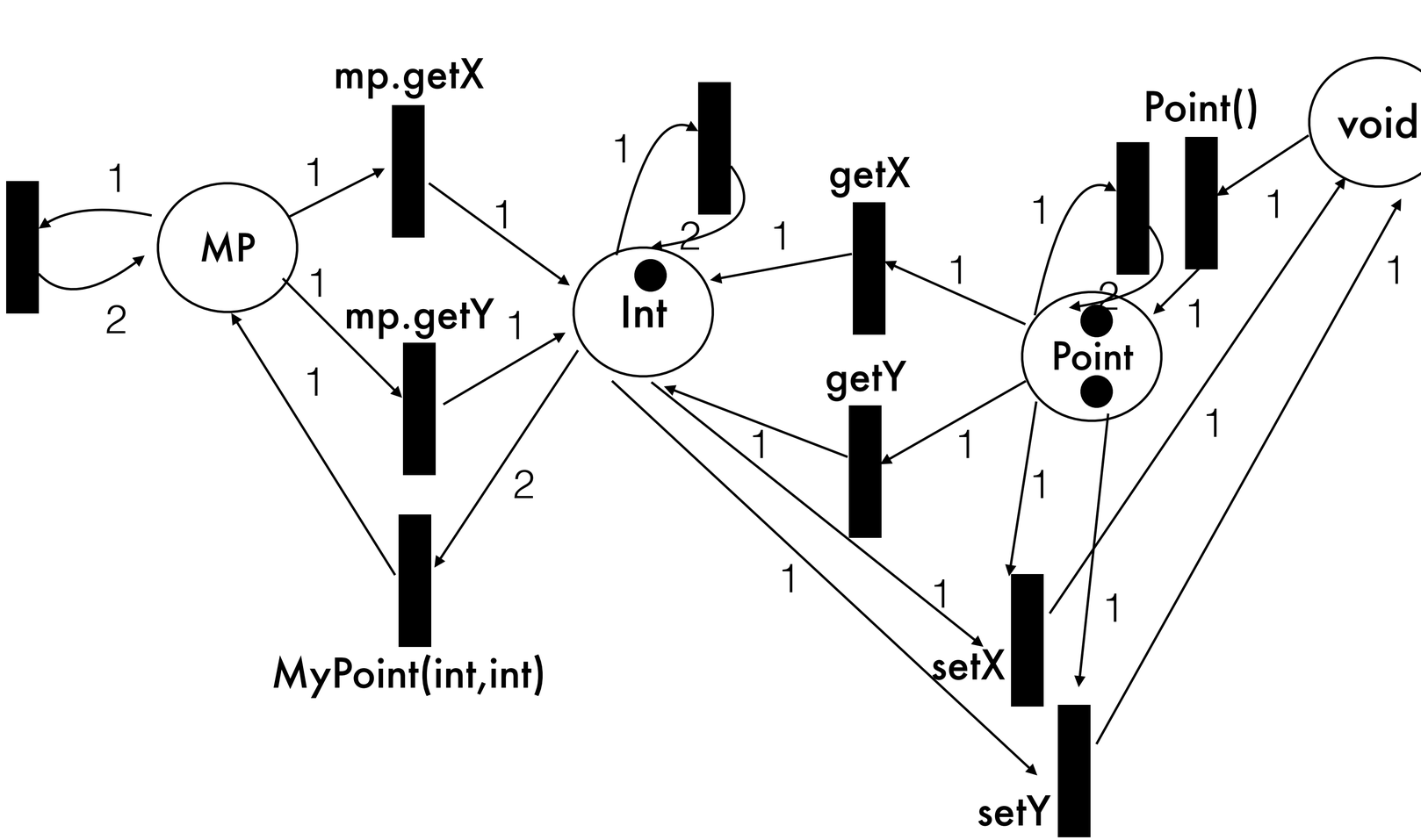
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)

1) mp.getX
2) Point()
3) Clone-Point
```

Exercise 2: Reachable paths

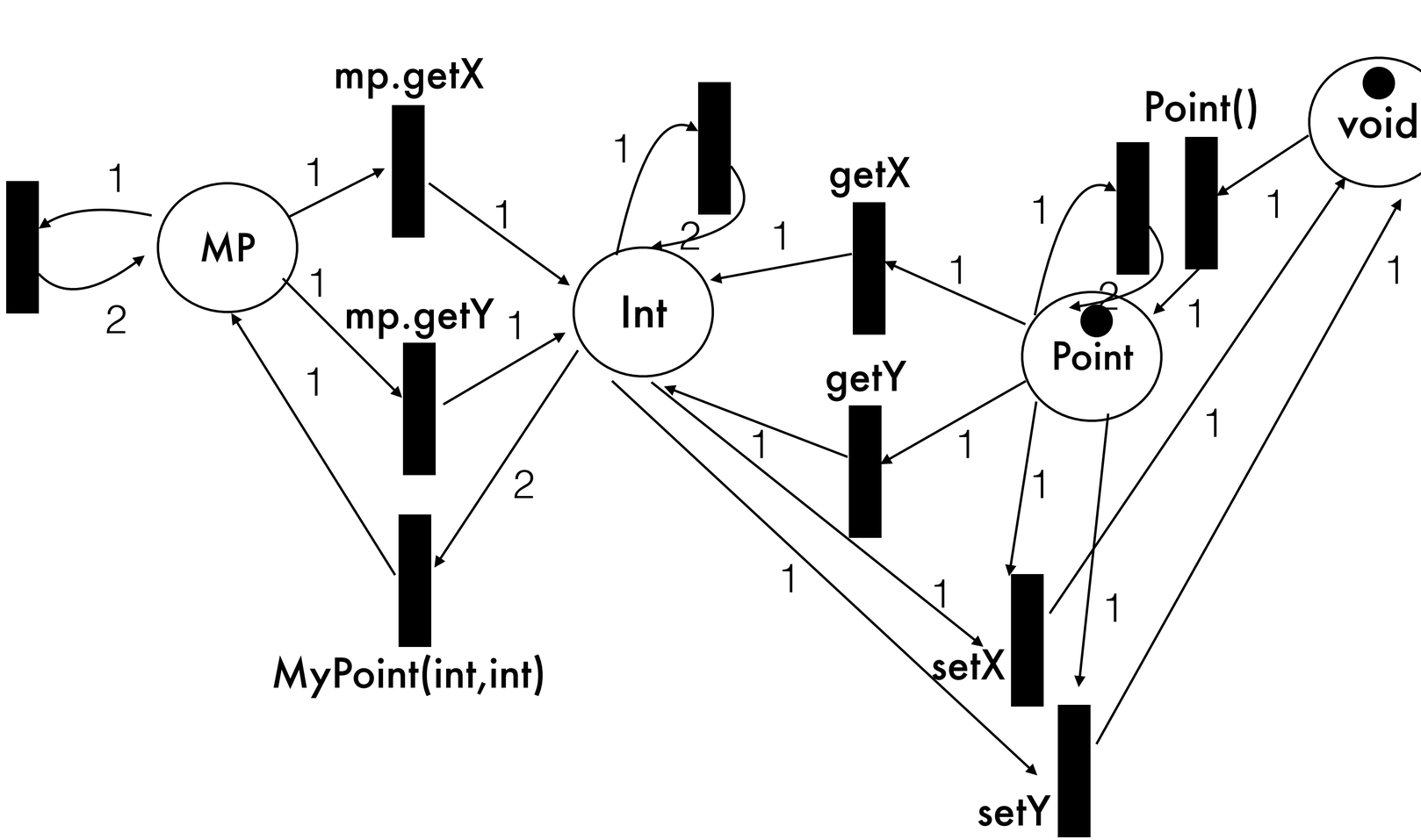


- Synthesize this function:

```
Point convert(Mypoint pt)
```

```
1) mp.getX
2) Point()
3) Clone-Point
```

Exercise 2: Reachable paths

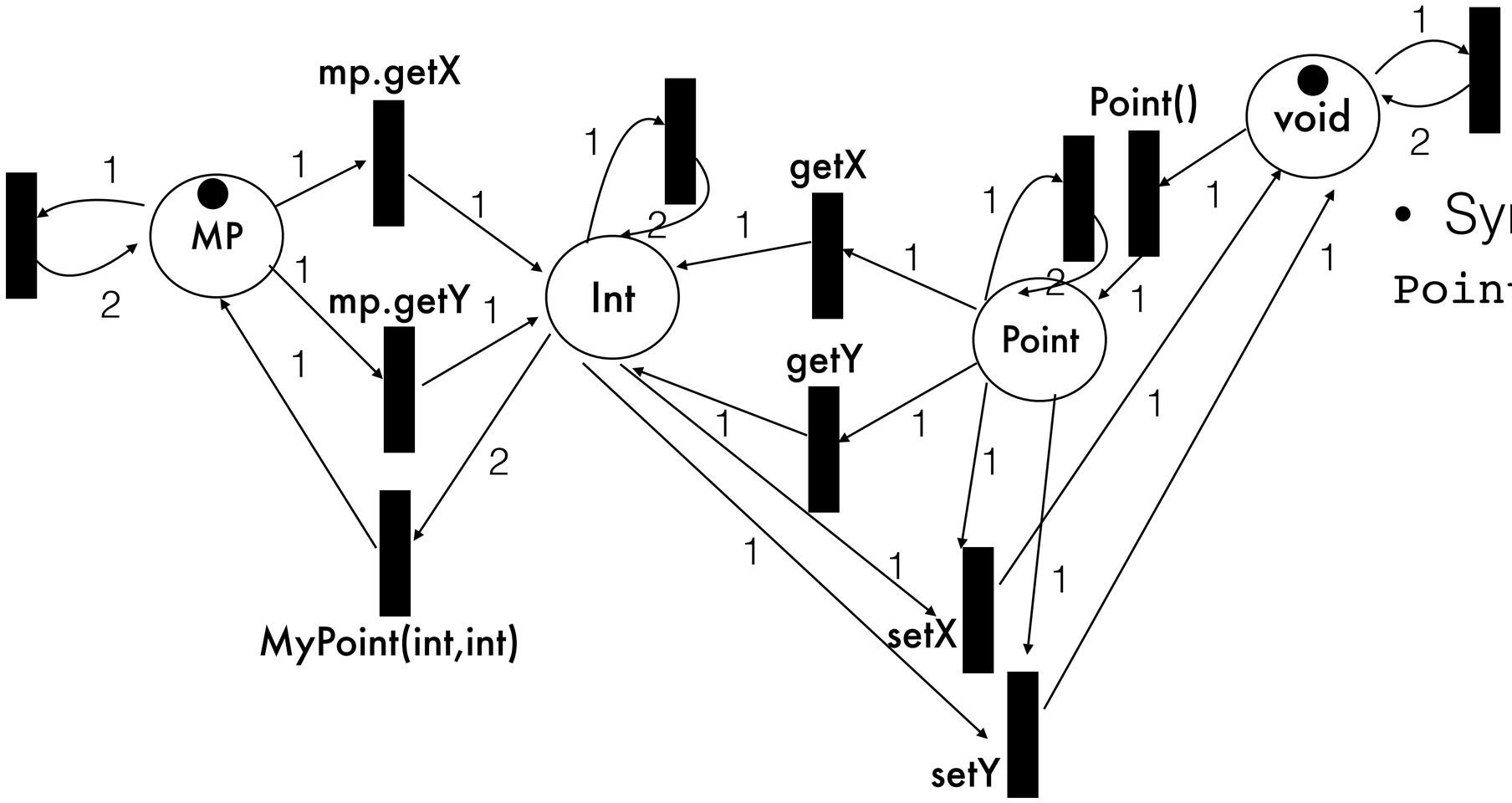


- Synthesize this function:

```
Point convert(Mypoint pt)
```

```
1) mp.getX
2) Point()
3) Clone-Point
4) setX
```

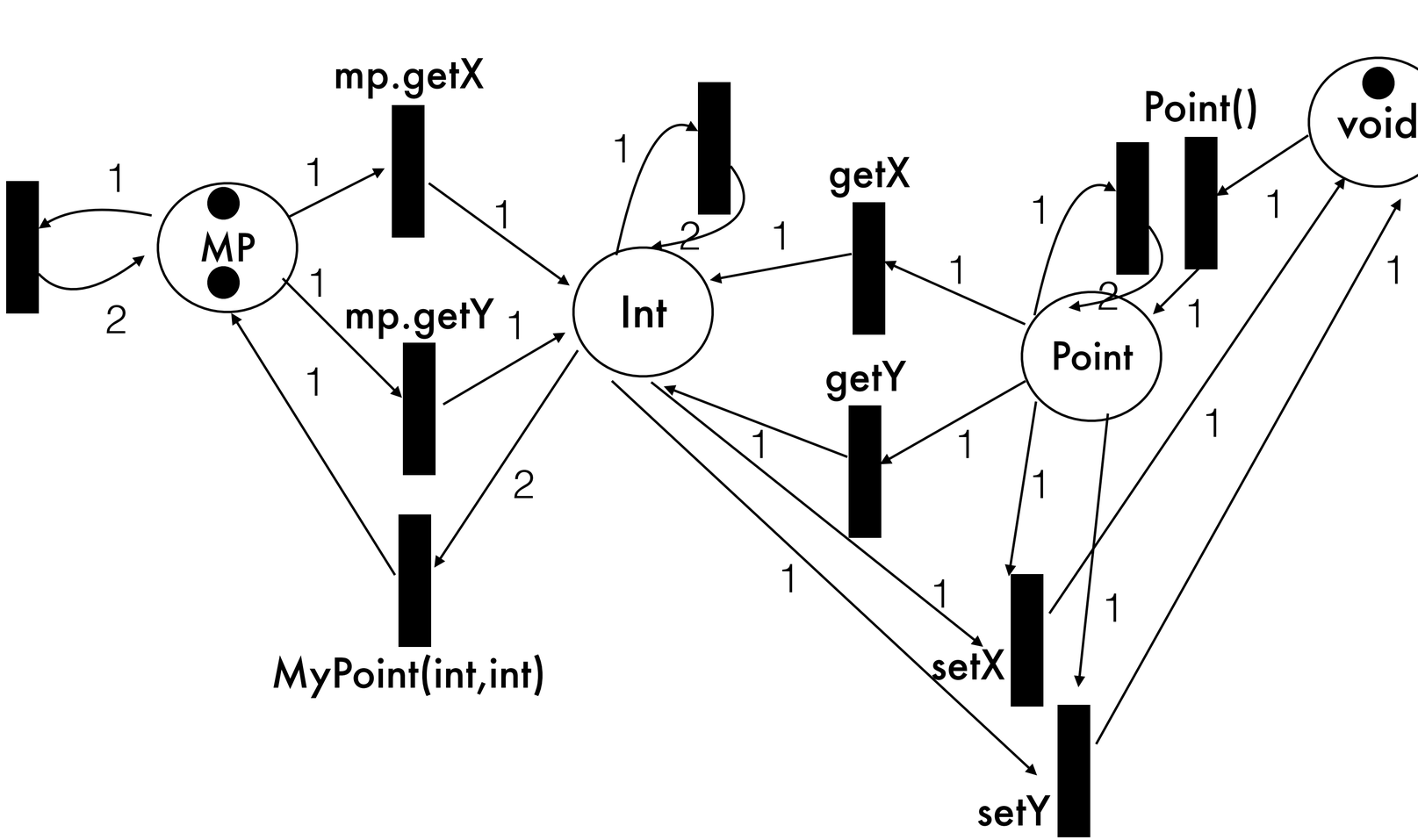
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)
```

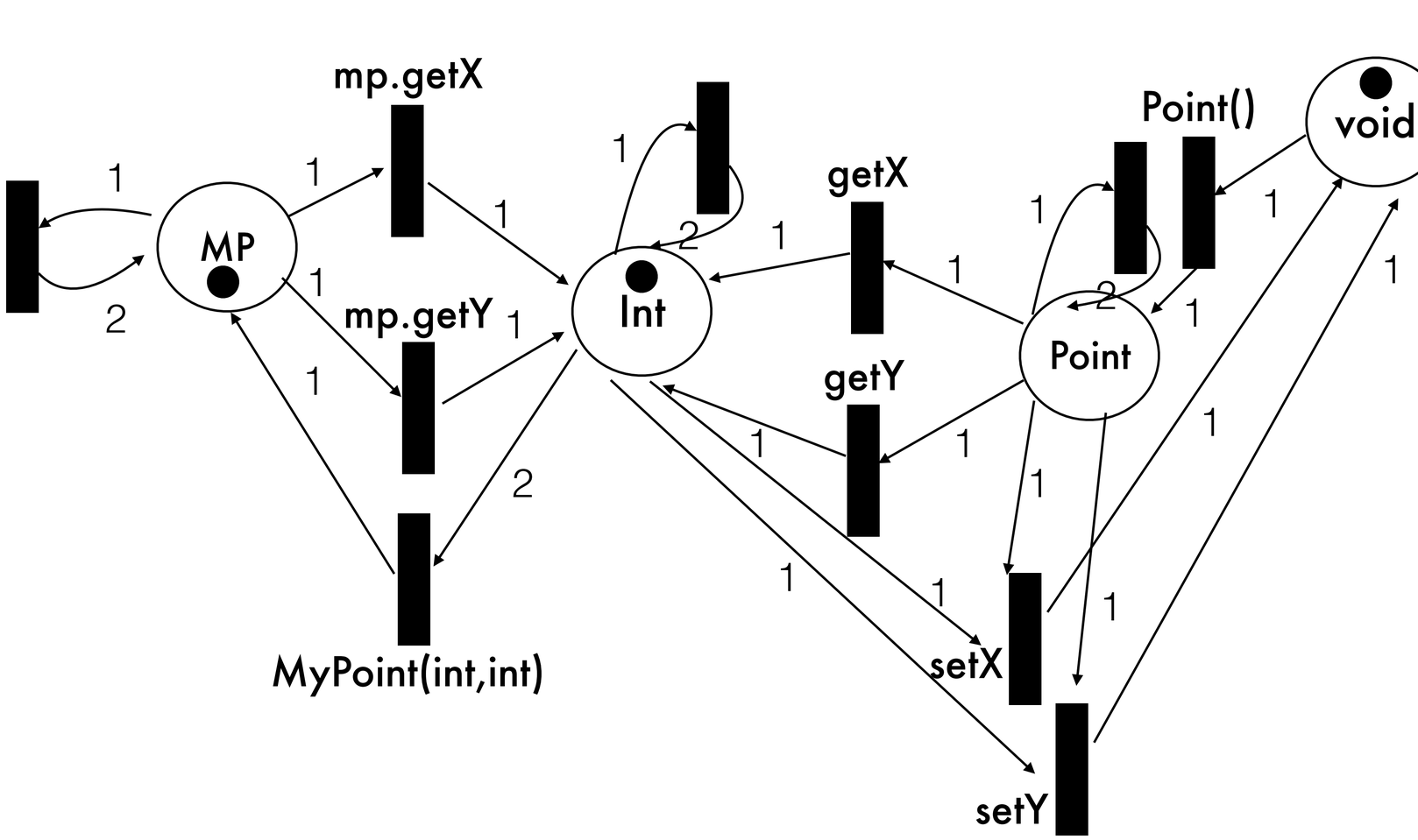
Exercise 2: Reachable paths



- Synthesize this function:

```
Point convert(Mypoint pt)
1) Clone-MP
```

Exercise 2: Reachable paths

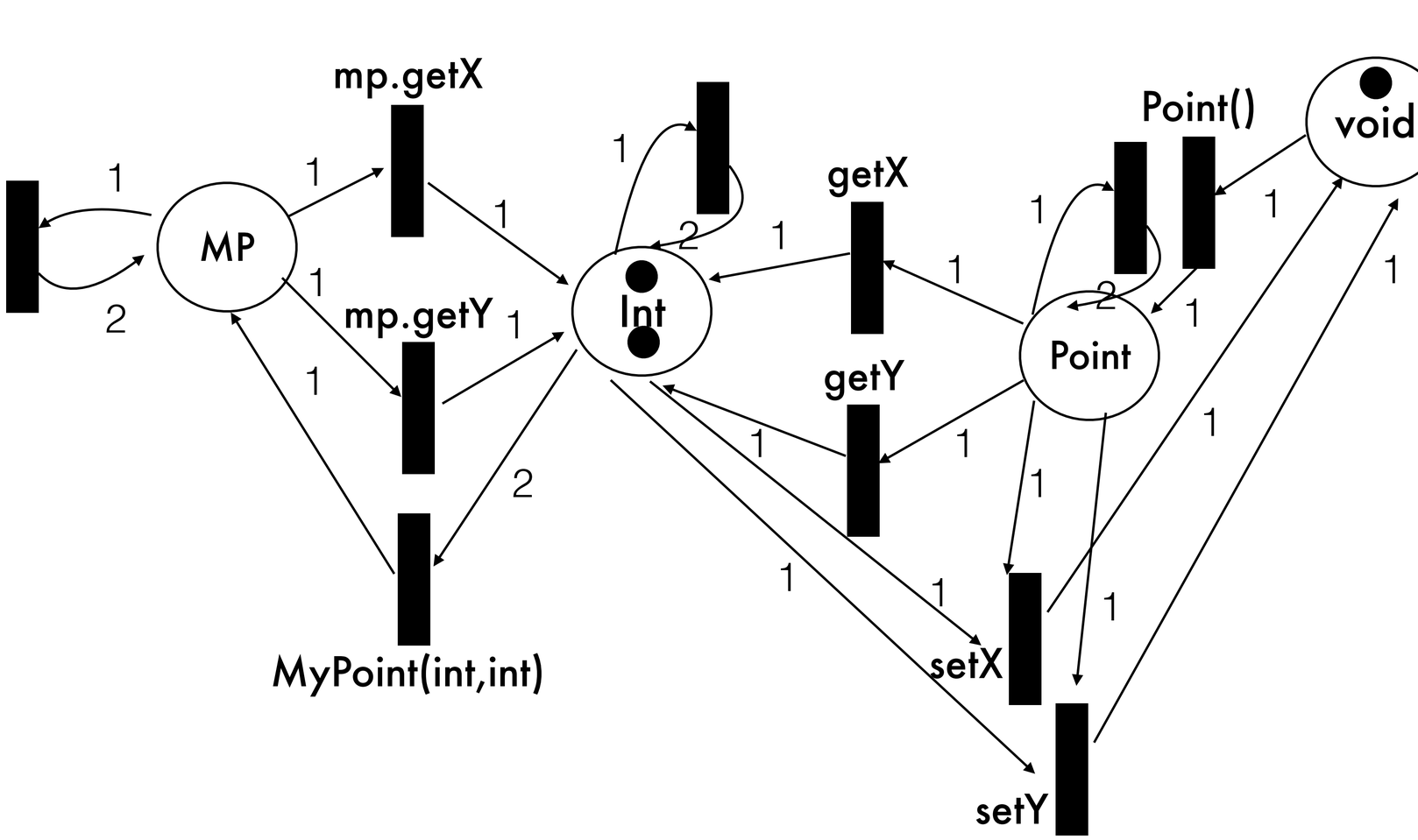


- Synthesize this function:

```
Point convert(Mypoint pt)
```

```
1) Clone-MP
2) mp.getX
```

Exercise 2: Reachable paths

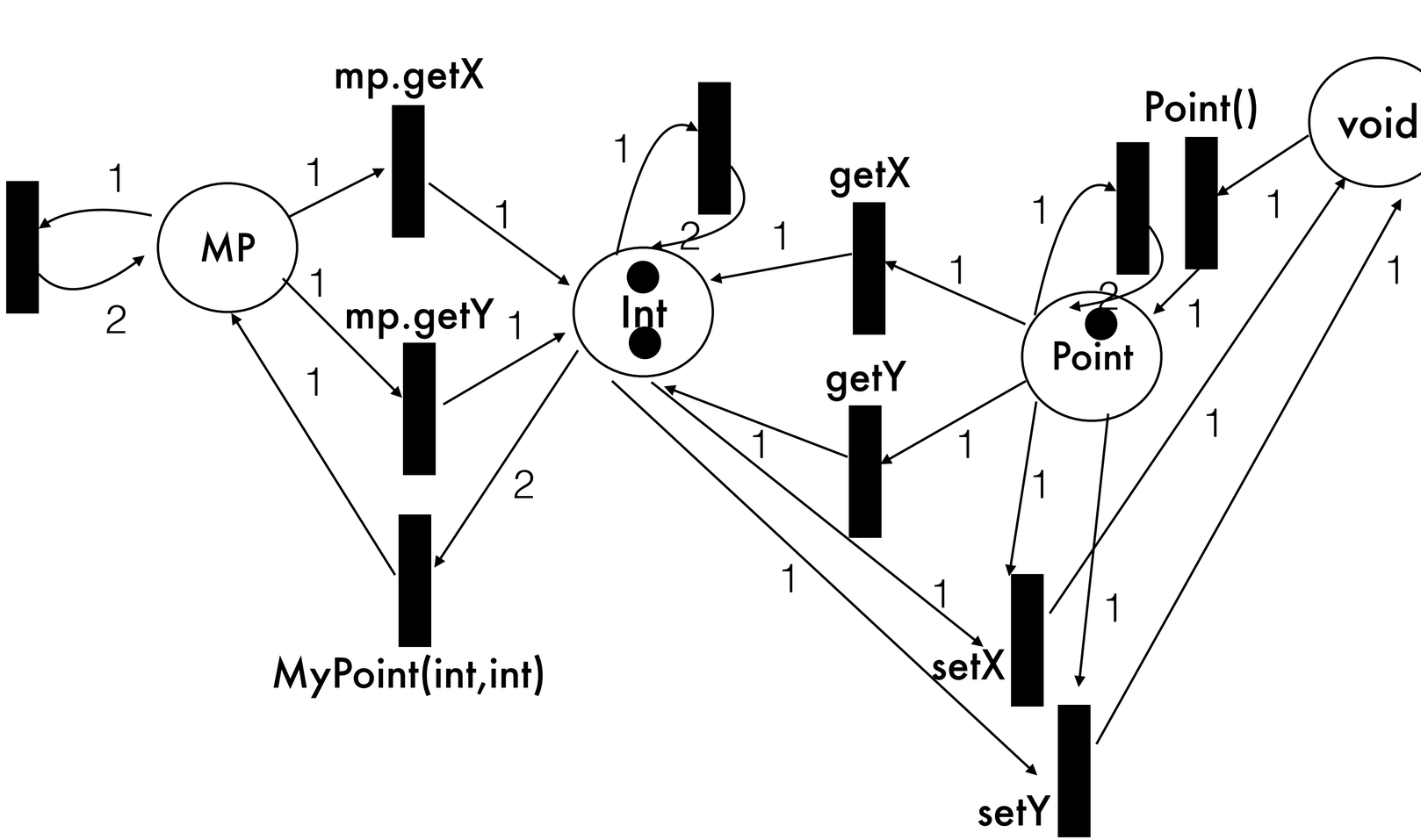


- Synthesize this function:

```
Point convert(Mypoint pt)
```

```
1) Clone-MP
2) mp.getX
3) mp.getY
```

Exercise 2: Reachable paths

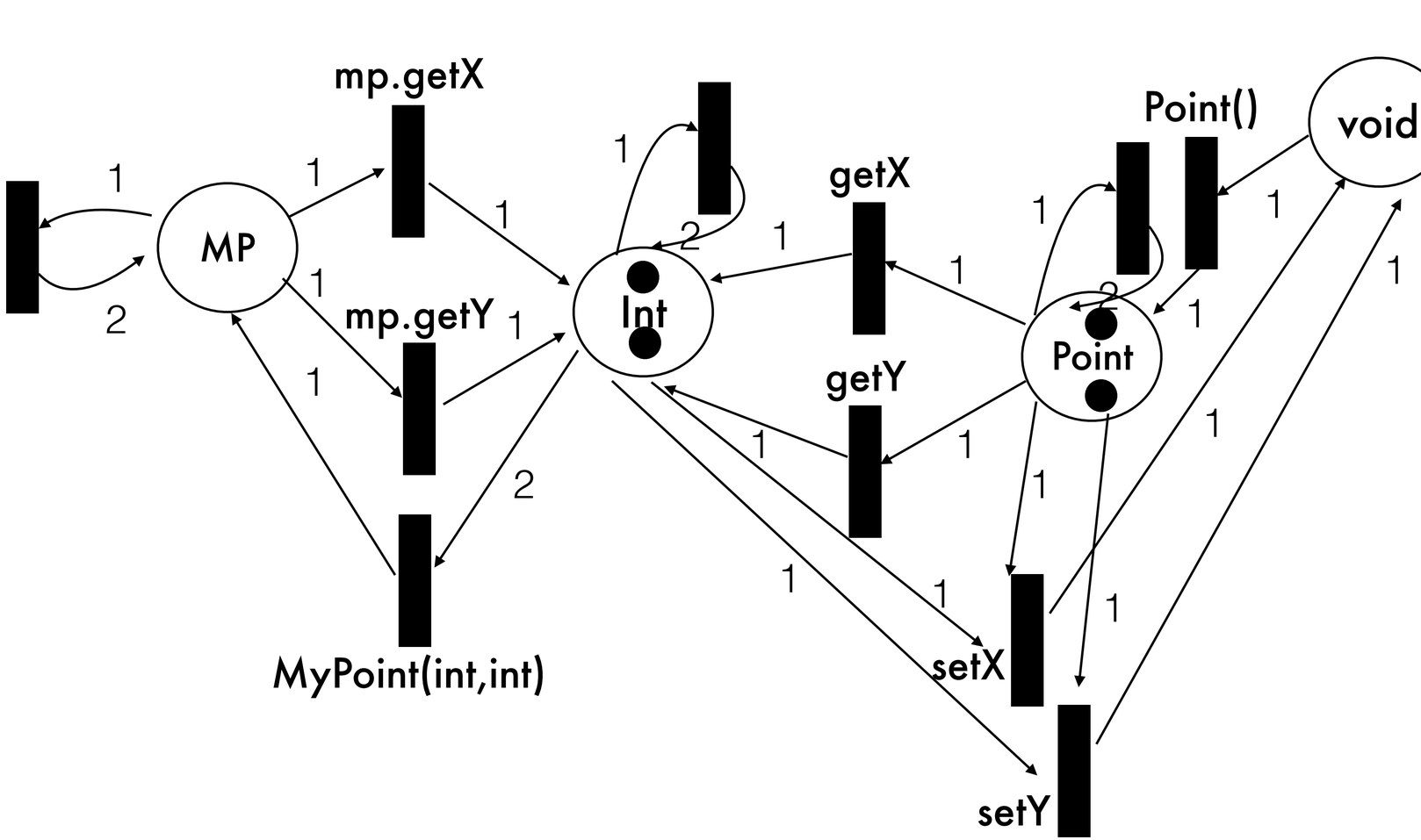


- Synthesize this function:

```
Point convert(Mypoint pt)
```

1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()

Exercise 2: Reachable paths

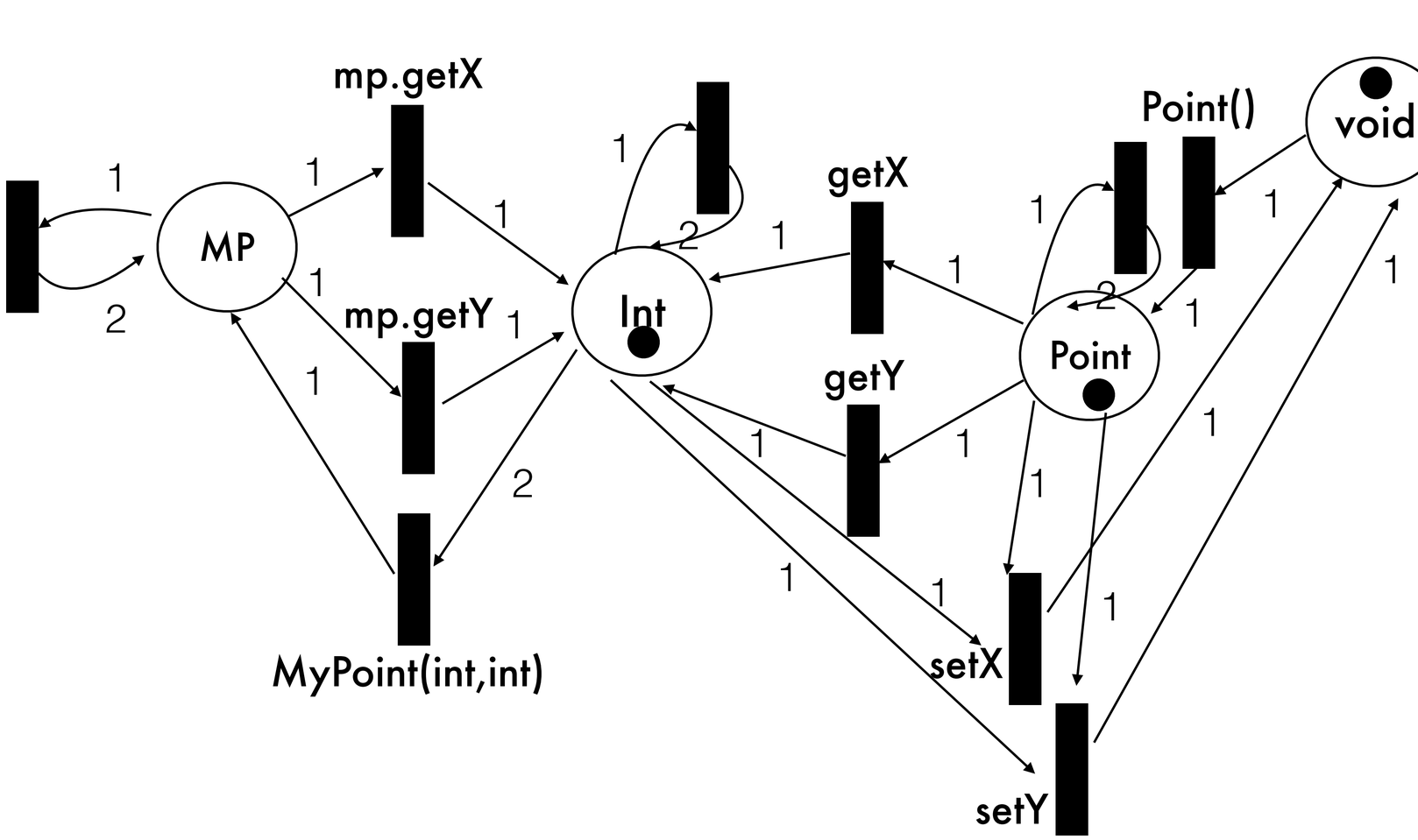


- Synthesize this function:

```
Point convert(Mypoint pt)
```

1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()
5) Clone-Point

Exercise 2: Reachable paths

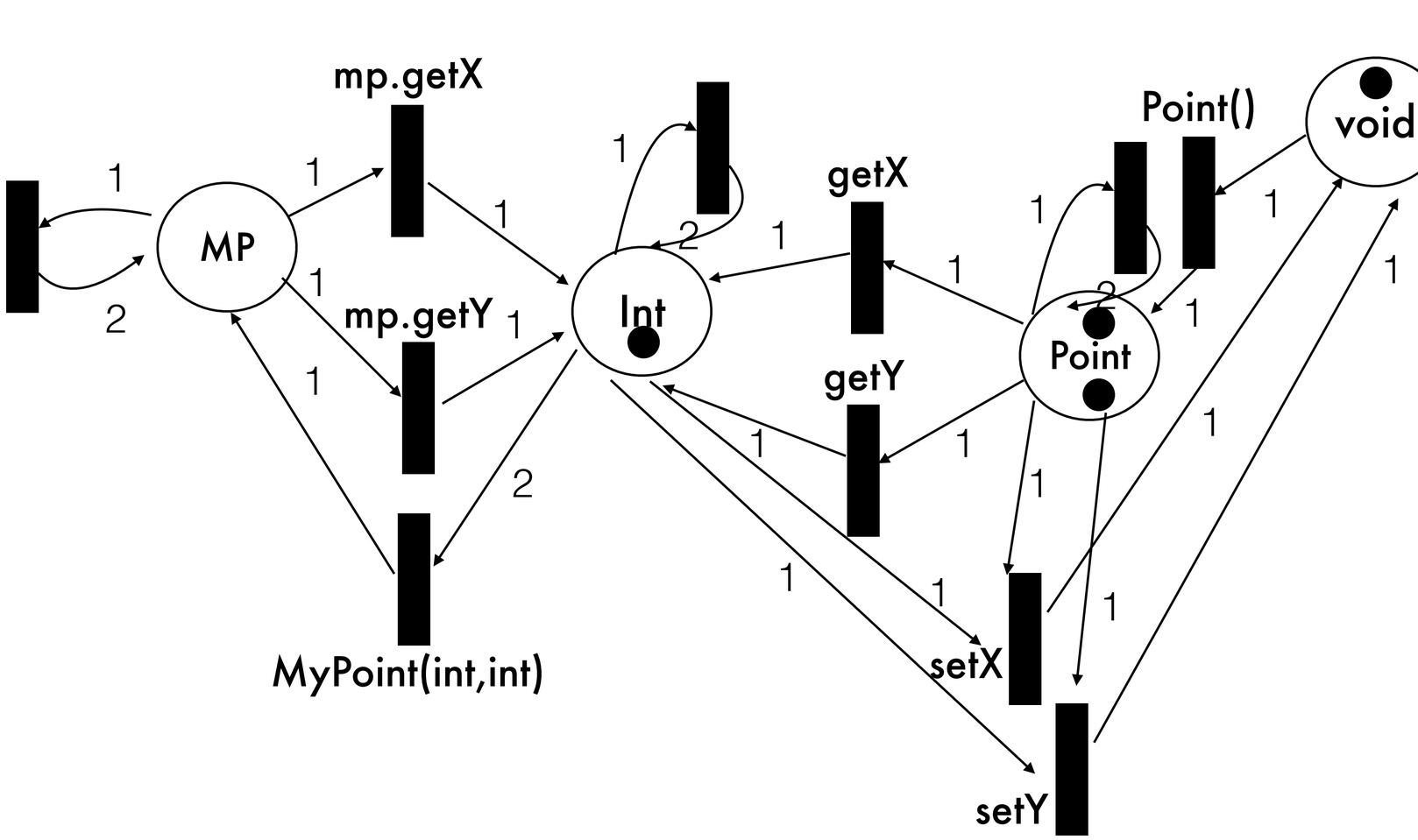


- Synthesize this function:

```
Point convert(Mypoint pt)
```

1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()
5) Clone-Point
6) setX

Exercise 2: Reachable paths

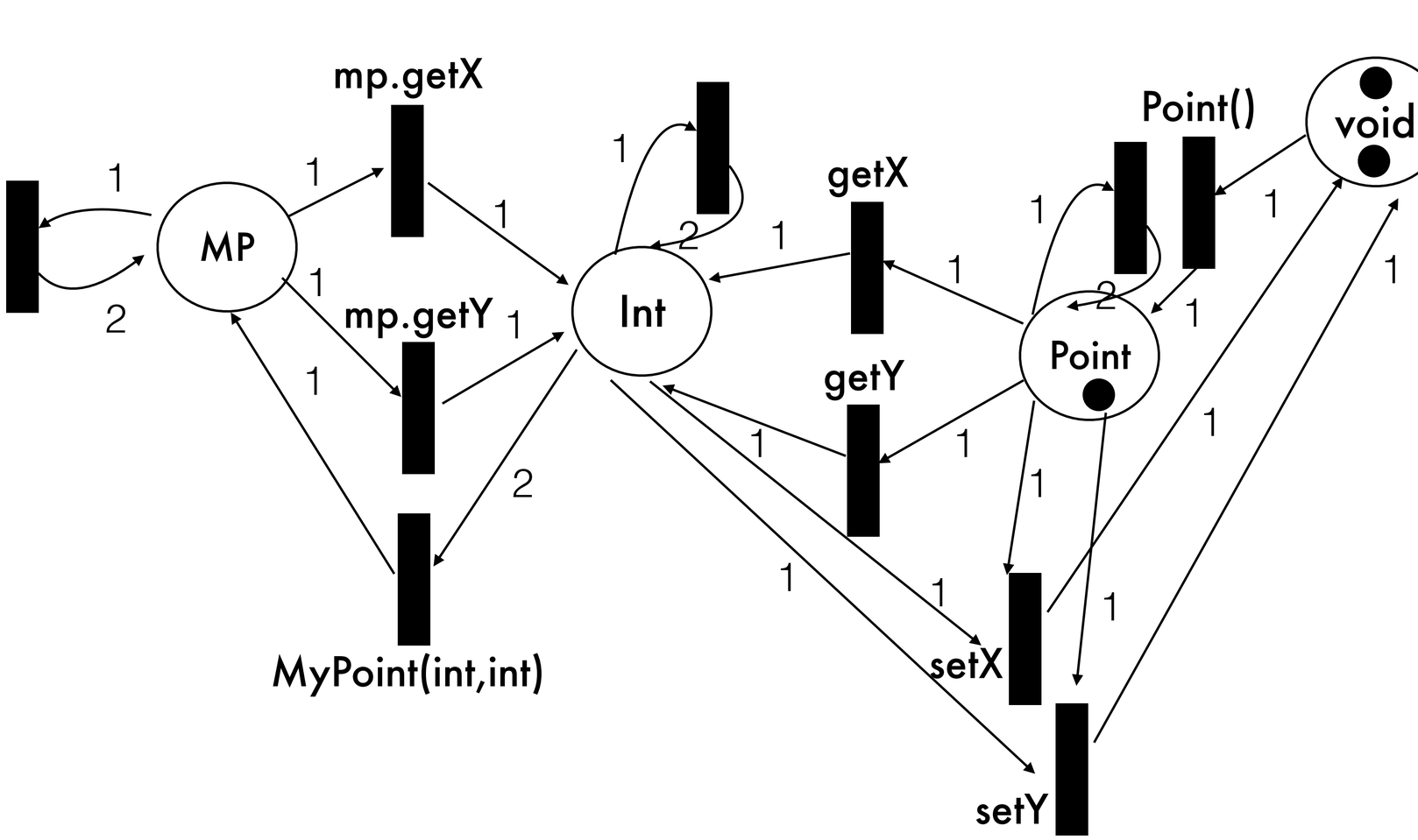


- Synthesize this function:

```
Point convert(Mypoint pt)
```

```
1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()
5) Clone-Point
6) setX
7) Clone-Point
```

Exercise 2: Reachable paths

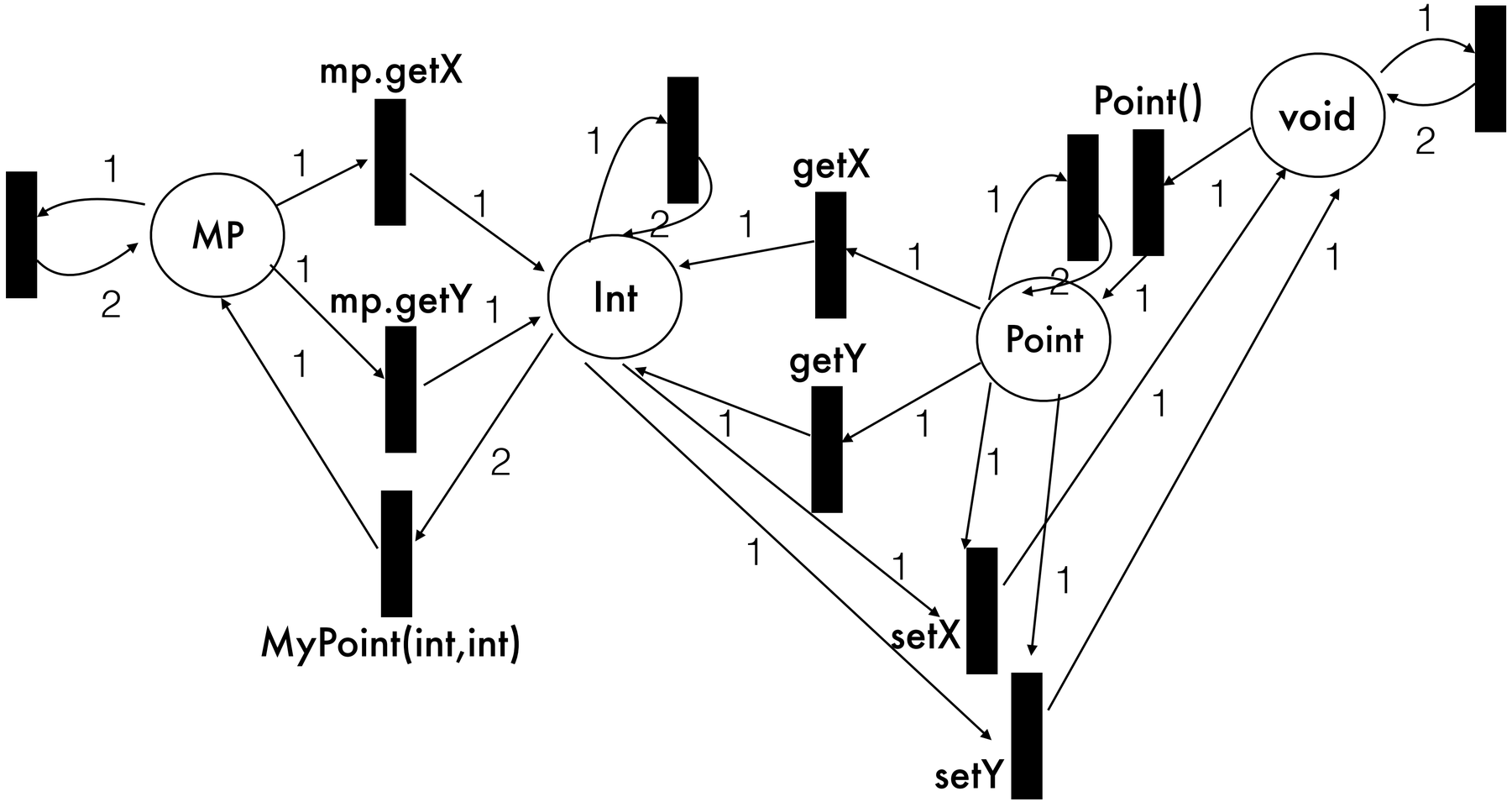


- Synthesize this function:

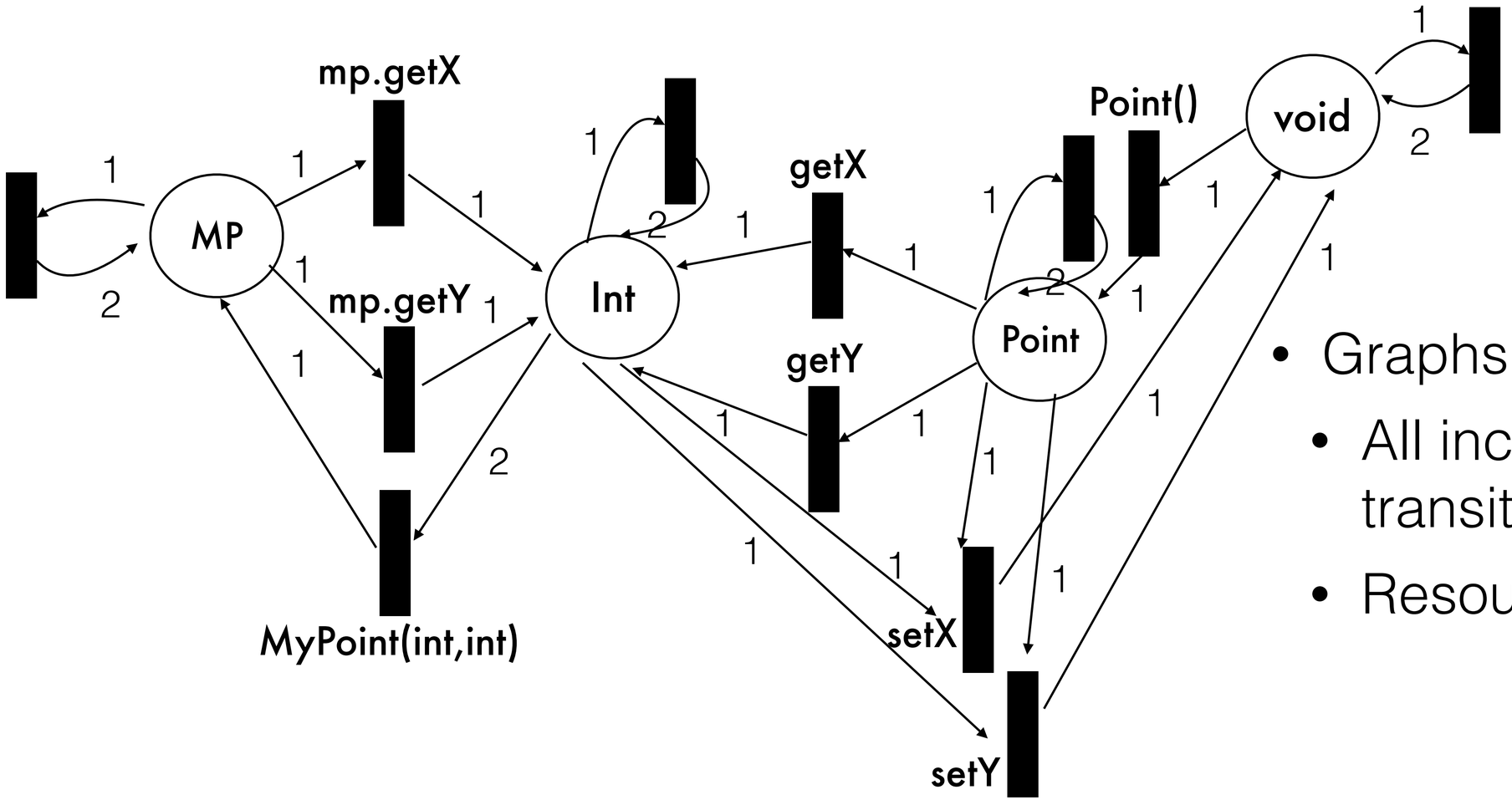
```
Point convert(Mypoint pt)
```

```
1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()
5) Clone-Point
6) setX
7) Clone-Point
8) setY
```

Why a Petri net and not a graph?

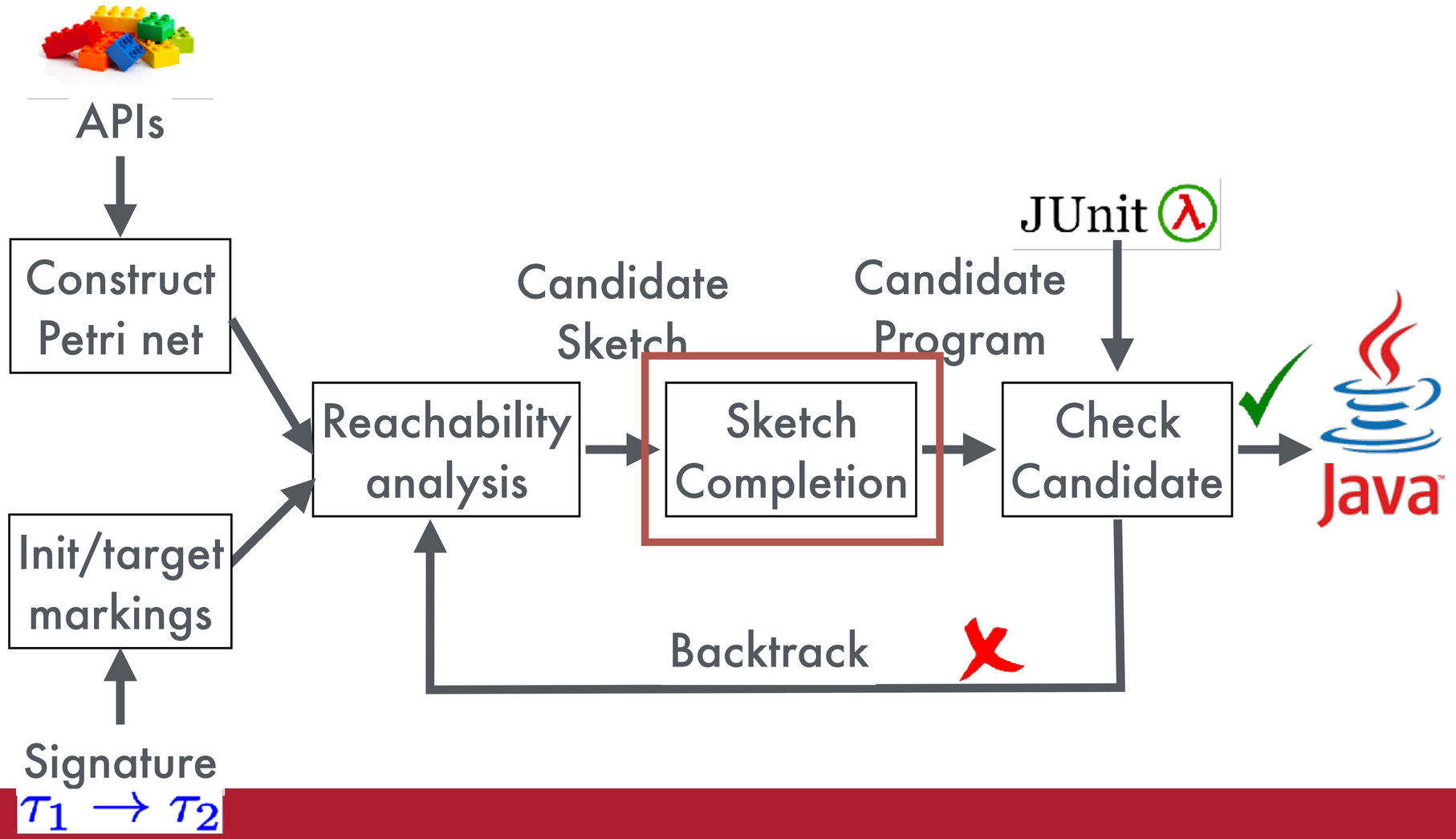


Why a Petri net and not a graph?



- Graphs do not support:
  - All incoming edges to a transition are part of the path
  - Resource consumption

Accepting run as program sketch



Accepting run as program sketch

Accepting run of Petri net corresponds to program sketch because it does not specify arguments of each method call.

APIs

Construct Petri net

Init/target markings

Signature $\tau_1 \rightarrow \tau_2$

Reachability analysis

Candidate Sketch

Sketch Completion

Candidate Program

JUnit

Check Candidate

Java

Backtrack

Sketch completion

```
x = #1.getX(); y = #2.getY();
#3.setToRotation(#4, #5, #6);
a = #7.createTransformedArea(#8);
return #9;
```

- Given a path:
  - getX -> getY -> setToRotation -> createTransformedArea
- Find the arguments that should be used in each hole such that the program type checks

Exercise 3: Sketch completion

```
1) Clone-MP
2) mp.getX
3) mp.getY
4) Point()
5) Clone-Point
6) setX
7) Clone-Point
8) setY
```

```
Point convert(Mypoint pt){




}
```

- Remove the Clone transitions

Exercise 3: Sketch completion

```
2) mp.getX
3) mp.getY
4) Point()
6) setX
8) setY
```

```
Point convert(Mypoint pt){




}
```

- What is the code with holes?

Exercise 3: Sketch completion

```
2) mp.getX
3) mp.getY
4) Point()
6) setX
8) setY
```

```
Point convert(Mypoint pt){

  int x = #1.getX();
  int y = #2.getY();
  Point p = new Point();
  #3.setX(#4);
  #5.setY(#6);
  return #7;

}
```

- What is the code with holes?

Exercise 3: Sketch completion

```
2) mp.getX
3) mp.getY
4) Point()
6) setX
8) setY
```

```
Point convert(Mypoint pt){

  int x = pt.getX();
  int y = pt.getY();
  Point p = new Point();
  p.setX(#4);
  p.setY(#6);
  return p;

}
```

- What is the code with holes?

Exercise 3: Sketch completion

```
2) mp.getX
3) mp.getY
4) Point()
6) setX
8) setY
```

```
Point convert(Mypoint pt){

  int x = pt.getX();
  int y = pt.getY();
  Point p = new Point();
  p.setX(y);
  p.setY(x);
  return p;

}
```

- What is the code with holes?

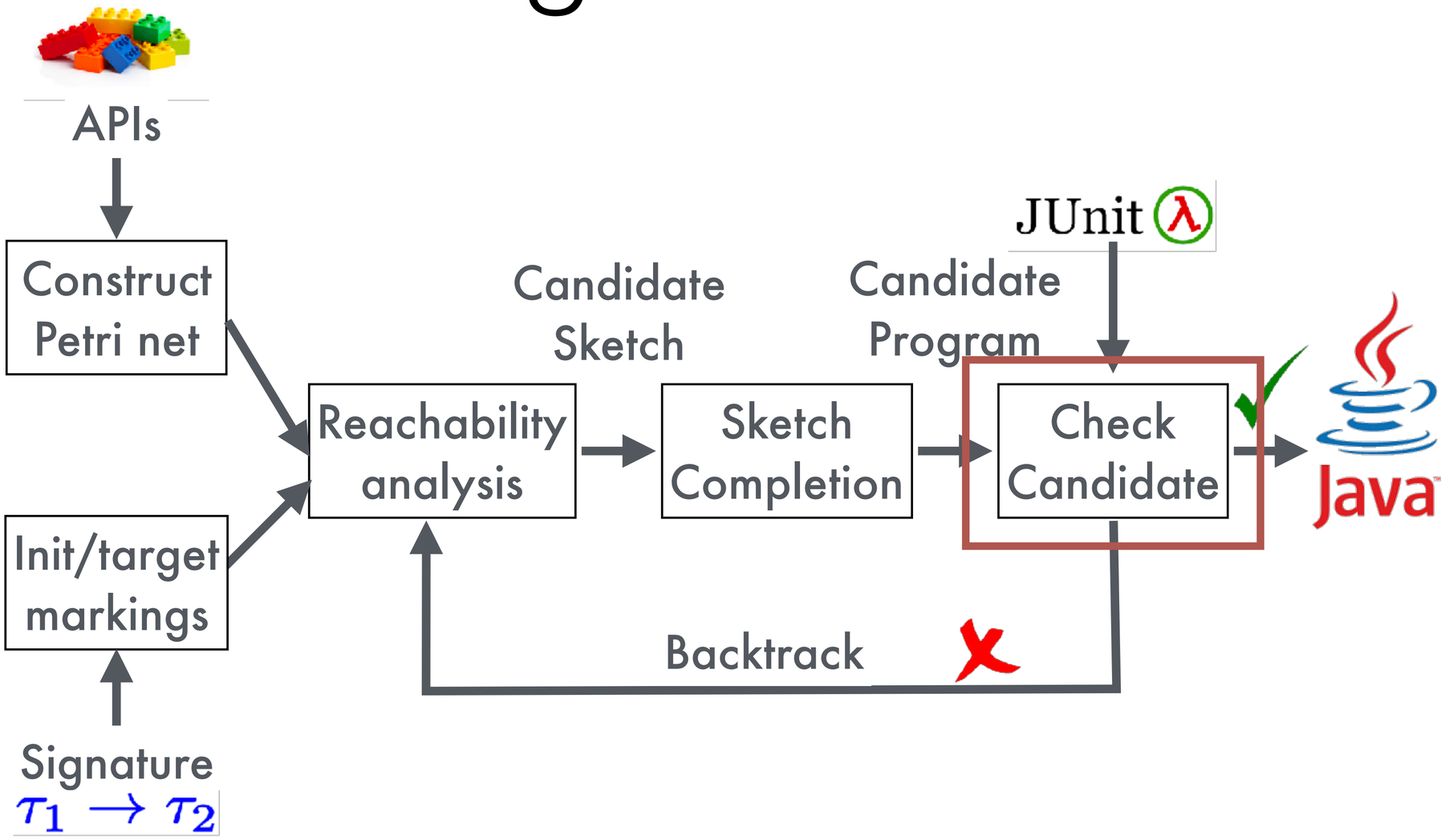
Exercise 3: Sketch completion

```
2) mp.getX
3) mp.getY
4) Point()
6) setX
8) setY
```

```
Point convert(Mypoint pt){

  int x = pt.getX();
  int y = pt.getY();
  Point p = new Point();
  p.setX(x);
  p.setY(y);
  return p;

}
```

- What is the code with holes?

Checking the candidate



Test cases

```
Point convert(Mypoint pt){

   int x = pt.getX();
   int y = pt.getY();
   Point p = new Point();
   p.setX(x);
   p.setY(y);
   return p;

}
```

- Write a test case to check the conversion

Test cases

```
Point convert(Mypoint pt){

   int x = pt.getX();
   int y = pt.getY();
   Point p = new Point();
   p.setX(x);
   p.setY(y);
   return p;

}
```

- Write a test case to check the conversion

```
bool test(){

   MyPoint mp = new MyPoint(1,2);
   Point p = convert(mp);
   return (p.getX() == 1 &&
           p.getY() == 2);

}
```

Using SyPet

- SyPet can be run in a remote server or locally
- You can send an HTTP POST request with a .json configuration file and the code will be synthesized for you!
- We also have a plug-in version for Atom (experimental)

Using SyPet

```
{
  "methodName": "distance",
  "paramNames": [
    "sypet_arg0",
    "sypet_arg1"
  ],
  "srcTypes": [
    “java.awt.geom.Point2D",
    “java.awt.geom.Point2D”
  ],
  "tgtType": "double",
  "packages": [
    "java.awt.geom"
  ],
  “testBody”:
  ”
    public static boolean test()
        throws Throwable {
    …
    }
    “
}
```

Using SyPet

```
"testBody":"
   public static boolean test() throws Throwable {
      java.awt.geom.Point2D p1 = new java.awt.geom.Point2D.Double(0,1);
      java.awt.geom.Point2D p2 = new java.awt.geom.Point2D.Double(0,4);
      return (Math.abs(distance(p1,p2)-3)==0);
   }
   "
}
```

Notes:

- You must use the full quantified name!
- Test function must throw a Throwable!

Using SyPet

How to send a request to SyPet:

```
curl -X POST -d @distance.json http://127.0.0.1:9092 --header "Content-Type:application/json
```

- Synthesized code:

```
double distance(java.awt.geom.Point2D sypet_arg0,
                java.awt.geom.Point2D sypet_arg1){

    double sypet_var1400 = sypet_arg1.distance(sypet_arg0);
    return sypet_var1400;

}
```

SyPet's strengths

SyPet's strengths

- Works for real code!
- Generic: can tackle any Java library
  - e.g. geometry, math, joda, unirest, xml, etc.
- Works well when there are many different types
- Scales to a large number of APIs

SyPet's weaknesses

SyPet's weaknesses

- Does not support conditionals
- Does not support loops
- For some applications it is hard to write test cases
- Does not scale when everything is the same type

For more information

https://utopia-group.github.io/sypet/

Outline

Code Reuse	Complex Java APIs	Data Wrangling	Learning in Synthesis
FSE'16	POPL'17	PLDI'17	PLDI'18

Program Synthesis for Data Science

Data extraction

Data cleaning

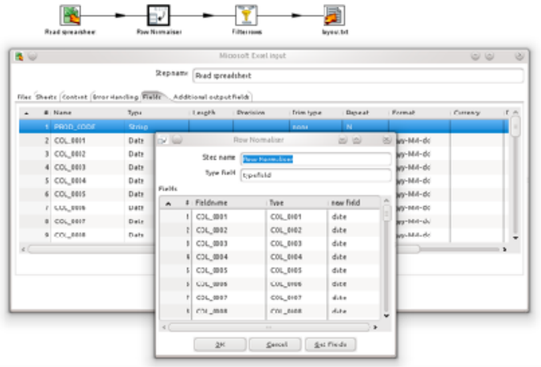


Table transformation

Program Synthesis for Data Science

Data extraction

Data cleaning

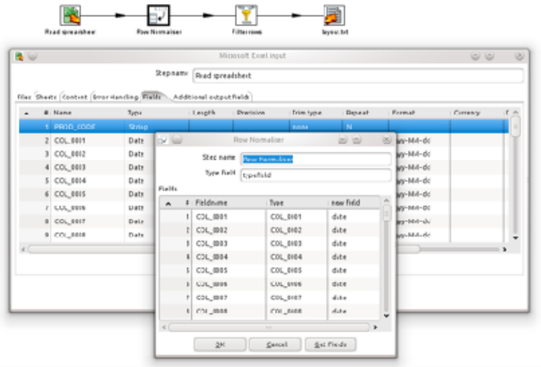


Table transformation

Data science in reality…

Data scientists spend 80% of time on tedious data wrangling tasks

Demo

- A demo of **Morpheus** is available at:

https://sat-group.github.io/ruben/files/morpheus.m4v

Applications of Program Synthesis

Data Science

Databases

Program Repair

Security

Computer-Aided Education

And many others!

Interested in Program Synthesis?

★ GHC 7.129

★ rubenm@cs.cmu.edu

Carnegie Mellon University

★ Send me an email and we can talk more!

Papers

FSE'16	Hunter: Next-Generation Code Reuse for Java
POPL'17	Component-Based Synthesis for Complex APIs
PLDI'17	Component-Based Synthesis of Table Consolidation and Transformation Tasks from Examples
PLDI'18	Program Synthesis using Conflict-Driven Learning