자바의 버전별역사

1) 초기 자바 ~ JDK 1.0

- 1991년, OAK 발표: GE사의 요청으로, 썬마이크로 시스템즈에서 C++의 단점을 극복하고자 만든 언어. 메모리 할당/해제의 어려움과 다중상속으로 인한 실수유발을 극복하려고 했다.

- 1996년, JDK 1.0발표 : 언어 이름을 자바라 바꾸고, Java Virtual Machine 1.0을 발표했다. Java Applet도 이때 처음 도입.

2) JDK1.1

- 1997년 : RMI, JDBC, reflection, JIT,Inner Class 개념이 포함되었다.

3) JDK 1.2

- 1998년 Java SE 1.2, ME 1.2, EE 1.2 발표

- 자바를 세가지 버전으로 나눴다. Swing이 SE에 포함, Corba IDL(이종기기간 함수호출 스펙), Collection Framework 포함

4) JDK 1.3

- 2000년도

- HotSpot(Sun에서 만든 JIT구현), JNDI(디렉토리랑 이름으로 원하는 서비스찾는것) 포함

5) JDK 1.4

- 2002년도

- JCP(Java Community Process)에 의해서 오픈소스 정책으로 자바가 관리되기 시작한 버전.

- Java 2 Security 모델의 확립(Sandbox)

- Java Web Start포함 (Java Applet이 브라우저에서 돌아가는 것과 다르게, 외부 sandbox에서 동작)

- Language: assert 도입

- API : Regular Expression, Assert keyword, Security 2 version(현재 security model), Non Blocking IO(NIO)

6) JDK 1.5

- 2004년

- 기능적으로 가장 많은 변화가 생긴버전 (Generics가 가장 대표적)

- LanguageI: Generics , annotation, auto boxing, enum,vararg ,foreach, static imports 도입

- API : java.util.concurrent API, scanner class

7) JDK 1.6

- 2006년도

- 기능에 별 차이 없슴 - 보안, 성능강화 주력.

- JVM/Swing에 있어 많은 Performance 향상(synchronization, compiler, GC,start-up time)

- G1(Garbage First) GC도입.

8) JDK 1.7

- 2011년도

- JVM : Dynamic Language support (invokedynamic - new byte operation)

- Language : Switch에서 String, try-resource, generics에서 타입추론, 숫자에서 underscore사용

- API:Concurrency 강화, NIO 강화, sort강화, crypto강화, GPU강화,

- JavaFX가 기본으로 포함

- 안정적인 ARM지원

9) JDK 1.8

- 2014

- 오라클로 인수된 후 첫번째 버전

- JDK 1.5이후 가장 큰 언어적 변화(Lambda및 함수형프로그래밍,default method)이며 러닝커브가 크다.

- JEP에 의해서 새로운 기능들이 발의되기 시작.

- Language : Lambda expression, Default Method Interface, functional programming for MapReduce style 지원, default method이용한 다중상속지원,메서드 참조

- API : Nashorn (JS엔진), new Date and Time API, stream api,Collection에 대한 함수형화 (Interface에 default가 생김으로서 가능)

- 병철처리에 접합한 구조로 진화

10) JDK 1.9

- 2016 예정

- Modular System (Jigsaw)지원예정

- Money API지원예정

- Java Shell지원예정

- 변수에 대한 타입 추론 지원예정(var,val)

- OpenCL이용한 자동화된 병렬 프로그래밍 지원예정

- value 타입 지원예정

**PART 1**

**CHAPTER1 자바 8을 눈여겨봐야 하는 이유**

**막간> Comparable, Comparator 차이**

기본형 타입이 아닌 Object 내에 특정값을 비교할 경우에는

Comparable를 implements 해서 compareTo를 구현하면 Collections.sort() , Arrays.sort() 등을 사용할 수 있음.

ex)

public class Book implements Comparable<Book>{

private int price;

public Book(int price){

this.price = price;

}

public int getPrice(){

return this.price;

}

@Override

public int compareTo(Book b) {

return this.price - b.price; //자신이 앞에 있는게 ascending order

}

}

Comparator는 compareTo 에서 구현한 기본값 외에 별도의 기준으로 구현하고 싶을 때 다음과 같이 Comparator 구현체 또는 익명 Comparator 구현체를 이용하여 나타낼 수 있다.

public static Comparator<Book> myComparator

= new Comparator<Book>() {

@Override

public int compare(Book b1, Book b2) {

return b1.getPrice() - b2.getPrice();

}

};

public static void main(String[] args) {

List<Book> myBookList = new ArrayList();

myBookList.add(new Book(300));

myBookList.add(new Book(310));

myBookList.add(new Book(100));

myBookList.add(new Book(200));

System.out.println("Original: " + myBookList);

Collections.sort(myBookList, myComparator);

for(Book book : myBookList){

System.out.println(book.getPrice());

}

}

익명 Comparator

Collections.sort(myBookList, new Comparator<Book>() {

@Override

public int compare(Book b1, Book b2) {

return b1.getPrice() - b2.getPrice();

}

});

**결론 :** Comparable과 Comparator은 상호보완 차이

더 간단하게 람다를 사용하면 다음과 같은 형태로도 가능

Collections.sort(myBookList,

(Book b1, Book b2)-> b1.getPrice() - b2.getPrice());

or

list.sort(Comparator.*comparing*(Book::getPrice)); // 오름차순  
list.sort(Comparator.*comparing*(Book::getPrice).reversed()); // 내림차순

참고 : <http://cwondev.tistory.com/15>

지금까지 대부분의 자바 프로그램은 싱글 코어만 사용되었다.

자바 8이전에는 나머지 코어를 사용하려면 스레드를 사용하는 것이 좋다

🡺 하지만 관리의 복잡성이 발생

🡺1.0에서는 스레드와 락으로 관리

🡺5에서는 스레드 풀, 병렬 실행 컬렉션(atomic)

🡺7에서는 포크/조인 프레임워크 제공

🡺8에서는 스트림 (질의 언어에서 표현식을 처리하는 것 처럼)

🡺 synchronized를 사용하지 않아도 된다.

병렬 관점이 아닌 조금 다른 관점에서 스트림 API를 보면 덕분에 두가지 기능이 생김(위험한생각)

* 메서드에 코드를 전달하는 간결 기법(메서드 레퍼런스와 람다)
* 인터페이스의 디폴트 메서드

메서드에 코드를 전달하는 자바8 기법은 함수형 프로그래밍에서 위력을 발휘한다.

**시작하기 앞서**

* 1. : 자바가 멀티코어 병렬성을 더 쉽게 이용할 수 있도록 진화하는 과정과 개념
  2. : 자바8에서 제공하는 코드를 메서드로 전달하는 기법이 어떻게 강력한 도구다 될수 있는지
  3. : 스트림 API가 어때서 강력하고 새로운 프로그래밍 도구인지
  4. : 디폴트 메서드라는 새로운 자바8의 기능을 인터페이스, 라이브러리의 간결성 유지 및

재컴파일을 줄이는데 어떻게 사용하는지

1.5 : JVM을 구성하는 자바 및 기타언어에서 함수형 프로그래밍의 영향이 어떤지

**1.1 왜 아직도 자바는 변화 하는가?**

**1.1.1 프로그래밍 언어 생태계에서 자바의 위치**

처음부터 많은 유용한 라이브러리를 포함하는 잘 설계된 객체지향 언어로 시작.

처음부터 스레드와 락을 이용한 소소한 동시성도 지원

코드를 JVM 바이트코드로 컴파일 하는 특징 때문에 자바는 인터넷 애플리케이션 프로그램의 주요 언어가 되었다.

🡺JVM 때문에 OS에 종속적이지 않다.

다양한 임베디드 컴퓨팅 분야를 성공적으로 장악하고 있다.

더 다양한 프로그래밍 도구를 제공하고, 다양한 프로그래밍 문제를 더 빠르고 정확하며 쉽게 유지보수할 수 있다는 점이 자바 8이 프로그래머에게 제공하는 개이득.

**1.1.2 스트림 처리**

스트림이란 한번에 한 개씩 만들어지는 연속적인 데이터 항목들의 모임

유닉스 C / stdin 🡺 stdout

Java System.in 🡺 System.out

조립라인은 물리적인 순서로 한 개씩 운반하지만 각각의 작업장에서는 동시에 작업 처리



Java.util.stream 패키지에 스트림 API가 추가 (조립라인 처럼 어떤 항목을 연속적으로 제공하는 기능이라고 단순히 생각하자)

🡺데이터베이스 질의처럼 고수준으로 추상화해서 일련의 스트림으로 만들어 처리

🡺스트림 파이프라인을 이용해서 입력 부분을 여러 CPU 코어에 쉽게 할당할 수 있다.

🡺스레드라는 복잡한 작업을 거치지 않아도 병렬성 획득

**1.1.3 동작 파라미터화로 메서드에 코드 전달하기**

자바8 이전에는 메서드를 다른 메서드로 전달할 방법이 없었다.

자바8에서는 가능함 (동작 파라미터화)

🡺스트림 API는 sort의 동작을 파라미터화하기 위해 compareUsingCustomerId를 넘겨주는 것처럼 어떤 동작을 파라미터화 할 수 있는 코드를 넘겨줄 수 있는 기능 제공 (아직 이해 부족)



**1.1.4 병렬성과 공유 가변 데이터**

스트림 메서드로 전달하는 코드는 다른 코드와 동시에 실행하더라도 안전하게 실행될 수 있어야 한다.

병렬에서 동시성 해결 하기

**공유 리소스 있는 경우**

🡺Synchronized / critical section (성능 저하)

🡺Atomic operation / CAS(Compare and Swap)

**공유 리소스 없는 경우**

🡺Immutable (불변)

🡺순수 함수

function add(a, b) {

return a + b;

}

**1.1.5 자바가 진화해야 하는 이유**

언어는 하드웨어나 프로그래머 기대의 변화에 부응할 수 있도록 변화 해야 한다.

🡺계속 새로운 기능을 추가하면서 인기 언어를 유지하고 있다.

**1.2 자바함수**

자바 8에서는 함수를 새로운 값의 형식으로 추가 (일급 시민, 일급 객체)

ex) scala의 일급 객체

**1) 변수나 데이터에 할당 할 수 있어야 한다**.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | object Main { | |  |  | |  | def test(): Unit ={ | |  | println("test"); | |  | } | |  |  | |  | def main(args: Array[String]): Unit = { | |  | println("Scala") | |  | **val a = test** | |  |  | |  |  | |  | } | |  | } | |

**2) 객체의 인자로 넘길 수 있어야 한다.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | object Main { | |  |  | |  | def test1(): Unit = { | |  | println("test1") | |  | } | |  |  | |  | def test2(f: => Unit): Unit = { | |  | f | |  | } | |  |  | |  | def main(args: Array[String]): Unit = { | |  | test2(test1) | |  | } | |  | } | |

**3) 객체의 리턴값으로 리턴할 수 있어야 한다.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | object TEST { | |  |  | |  | def test1(): Any => Unit = { | |  | return { x => println(x) } | |  | } | |  |  | |  | def main(args: Array[String]): Unit = { | |  | test1()("test") | |  | } | |  | } | |

**왜 함수가 필요 할까?**

🡺프로그래밍 언어의 핵심은 값을 바꾸는 것이다.

🡺자바는 이급 시민 (메서드나 클래스를 자유롭게 전달 못함 / 값으로 전달 못함)

🡺자바8에서 이급 시민 -> 일급 시민으로 바꿀 수 있는 기능 추가

**1.2.1 메서드와 람다를 일급 시민으로**

**1) 메서드 레퍼런스 (::)**

🡺메서드를 인자로 넘긴다.

|  |
| --- |
| File[] hiddenFiles = new File(“.”).listFiles(**File::isHidden**); |

**2) 람다: 익명함수(->)**

자바8에서는 메서드를 일급값으로 취급할 뿐 아니라 람다를 포함하여 함수도 값으로 취급

Utils::add1 같이 직접 구현해서 사용할 수 도 있지만

메서드가 없을 때는 (int x) -> x+1 과 같이 람다 문법을 이용하면 더 간결하게 코드를 구현

**🡺메서드화 되어 있으면 람다를 사용할 필요가 없다. (사용 할 때 선택 기준)**

**1.2.2 코드 넘겨주기**

필터 : 특정 항목을 선택해서 반환하는 동작

자바8 이전

public static List<Apple> filterGreenApples(List<Apple> inventory){  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple: inventory){  
 if ("green".equals(apple.getColor())) {  
 result.add(apple);  
 }  
 }  
 return result;  
}  
  
public static List<Apple> filterHeavyApples(List<Apple> inventory){  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple: inventory){  
 if (apple.getWeight() > 150) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

🡺비슷한 기능이지만 목적에 따라서 메서드가 여러 개 생긴다.

자바8에서는 코드를 인수로 넘겨줄 수 있기 때문에 filter 메서드를 중복으로 구현할 필요가 없다.

public static void main(String[] args) {  
 List<Apple> inventory = Arrays.*asList*(new Apple(80, "green"),  
 new Apple(155, "green"),  
 new Apple(120, "red"));

// 자바8 Predicate 방식  
 List<Apple> greeanApples = *filterApples*(inventory, **FilteringApples::*isGreeApple***); // 호출부  
  
 List<Apple> heavyApples = *filterApples*(inventory, **FilteringApples::*isHeavyApple***); // 호출부  
}  
  
public static boolean isGreeApple(Apple apple) {  
 return "green".equals(apple.getColor());  
}  
  
public static boolean isHeavyApple(Apple apple) {  
 return apple.getWeight() > 150;  
}  
  
static List<Apple> filterApples(List<Apple> inventroy, **Predicate<Apple> p**) {  
 List<Apple> result = new ArrayList<>();  
  
 for (Apple apple : inventroy) {  
 if (**p.test(apple)**) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

🡺핵심은 메서드를 전달할 수 있다는 점이다.

**프레디케이트란 무엇인가?**

수학에서는 인수로 값을 받아 true나 false로 반환하는 함수를 프레디케이트라고 한다.

Function<Apple, Boolean> 같이 구현할수도 있지만 Predicate<Apple>이 더 표준적인 방식

**1.2.3 메서드 전달에서 람다로**

Predicate를 사용하여 메서드를 값으로 전달하는 것은 유용하지만 매번 메서드 정의하는 것은 귀찮다. 이런 경우 람다를 사용

// 자바8 람다 방식  
*filterApples*(inventory, (Apple a) -> "green".equals(a.getColor()));  
*filterApples*(inventory, (Apple a) -> a.getWeight() > 150);  
*filterApples*(inventory, (Apple a) -> a.getWeight() < 80 || "brown".equals(a.getColor()));

🡺한번만 사용할 메서드는 따로 정의할 필요가 없다.

🡺하지만 람다가 몇 줄 이상 길어진다면(복잡해진다면) 익명 람다보다는 메서드를 정의하는 것이 더 명확하다.

// 다음과 같은 내장된 라이브러리를 추가하지 못한것은 병렬성 때문이다.  
Collections.filter(inventory, (Apple a) -> a.getWeight() > 150);

🡺병렬성이 없었다면 다음과 같이 라이브러리 형태로 더 간단하게 제공될수도 있었을 것이다.

🡺대신 비슷한 동작을 수행하는 스트림 API(컬렉션과 비슷하지만 함수형 프로그래머에게 더 익숙한) 를 제공한다. 또한 컬렉션과 스트림 간에 변환할 수 있는 메서드 (map, reduce 등)도 제공

**1.3 스트림**

자바8 이전 컬렉션에서 필터링이 필요한 경우

Map<Currency, List<Transaction>> transactionByCurrencies = new HashMap<>();  
for(Transaction transaction : *transactions*) {  
 if(transaction.getPrice() > 1000) {  
 Currency currency = transaction.getCurrency();  
 List<Transaction> transactionForCurrency = transactionByCurrencies.get(currency);  
 if(transactionForCurrency == null) {  
 transactionForCurrency = new ArrayList<>();  
 transactionByCurrencies.put(currency, transactionForCurrency);  
 }  
 transactionForCurrency.add(transaction);  
 }  
}

🡺많은 코드를 구현해야될 뿐만 아니라 흐름 문장이 많아서 한번에 이해하기도 어렵다.

Map<Currency, List<Transaction>> transactionByCurrencies2 =  
 *transactions*.stream()  
 .filter((Transaction t) -> t.getPrice() > 1000)  
 .collect(*groupingBy*(Transaction::getCurrency)); // 통화로 그룹화함 (Map 형태)

🡺스트림 API를 사용하면 간략하게 해결할 수 있음

스트림 API를 이용하면 컬렉션 API와는 다른 방식으로 데이터를 처리할 수 있다.

컬렉션(외부반복) : 반복과정을 for-each 루프를 이용하여 각 요소를 반복하면서 처리

🡺목록이 많다면 시간이 오래 걸린다. 단일 CPU로 거대한 데이터 처리가 힘듬.

스트림(내부반복) : 라이브러리 내부에서 모든 데이터 처리

🡺멀티코어를 이용하여 빨리 처리할 수 있다.

**1.3.1 멀티스레딩은 어렵다.**

동시성 문제가 발생할 수 있다. (공유 데이터 접근)

자바8에서는 synchronized가 아닌 함수형 프로그래밍 형식의 스트림 기반 병렬성을 이용하도록 권고.



멀티스레딩일 경우 발생할 수 있는 문제점. (sum 공유자원)

자바8은 스트림 API

🡺모호함과 반복적인 코드 문제 해결

🡺멀티코어 활용 어려움 해결

기존 컬렉션 데이터는 반복패턴이 너무 많아서 라이브러리에서 이러한 반복패턴을 제공(동기)

🡺주어진 조건에 따라 데이터 필터링

🡺데이터를 추출 (특정 필드만 추출)

🡺데이터 그룹화 (Map 형태로)

🡺쉽게 병렬화 할 수 있다.

스트림 API가 parallel하게 작업할때는 내부적으로 ForkJoinPool 을 사용하는듯

<http://www.popit.kr/java8-stream%EC%9D%98-parallel-%EC%B2%98%EB%A6%AC/>

컬렉션은 어떻게 데이터를 저장하고 접근할지 중점을 둔다

스트림은 데이터에 어떤 계산을 할것인지 묘사하는 것에 중점을 둔다.

Stream API에서 병렬과 순차처리

// 순차 처리  
inventory.stream().filter((Apple a) -> a.getWeight() > 150)  
 .collect(Collectors.*toList*());  
  
// 병렬 처리  
inventory.parallelStream().filter((Apple a) -> a.getWeight() > 150)  
 .collect(Collectors.*toList*());

**1.4 디폴트 메서드**

자바8에서는 더 쉽게 변화할 수 있는 인터페이스를 만들 수 있도록 디폴트 메서드 추가

디폴트 메서드는 프로그램 구현하는데 도움을 주는 기능이 아니라 미래에 프로그램이 쉽게 변화할 수 있는 환경을 제공

🡺부모 interface를 구현한 100개의 클래스가 있는데 만약 interface에 기능이 추가된다면 100개 클래스에 모두 구현을 해야된다. 이런부분의 영향을 줄어주기 위해 디폴트 메서드 사용.

디폴트 메서드를 이용하면 기존의 코드를 건드리지 않고도 원래의 인터페이스 설계를 자유롭게 확장할 수 있다.

public interface OperateCar {  
 default public int startEngine(EncryptedKey key) {  
 // Implementation  
 }  
}  
public interface FlyCar {  
 default public int startEngine(EncryptedKey key) {  
 // Implementation  
 }  
}  
  
public class FlyingCar implements OperateCar, FlyCar {  
 public int startEngine(EncryptedKey key) {  
 FlyCar.super.startEngine(key);  
 OperateCar.super.startEngine(key);  
 }  
}

🡺다중상속을 피하기 위해 단일상속 하게 했는데 디폴트메서드로 인해 다시 다이아몬드 상속 문제가 발생하게 되었다. ( C++과 동일 )

**1.5 함수형 프로그램에서 가져온 다른 유용한 아이디어**

1) 메서드와 람다를 일급값으로 사용 (이급 -> 일급 으로 변경함으로써 람다를 사용할 수 있고

더 간략한 코딩을 할 수 있다.)

2) 가변 공유 상태가 없는 병렬 실행을 이용 (불변)

자바8에서는 NullPointer 예외를 피할 수 있게 **Optional**<T> 클래스 제공

자바의 switch 는 문자열과 기본값만 이용할 수 있는 반면에 함수형 언어는 보통 패턴 매칭을 포함한 다양한 데이터 형식을 사용 할 수 있다.

스칼라 같은 경우 패턴 매치를 통해 다양한(type, class )등의 어떤 종류의 데이터라도 매칭 가능.

object MatchTest2 extends App {  
 def matchTest(x: Any): Any = x match {  
 case 1 => "one"  
 case "two" => 2  
 case y: Int => "scala.Int"

case \_ => "no match"  
 }  
 println(matchTest("two"))  
}

**CHAPTER2 동작 파라미터화 코드 전달하기**

동작 파라미터화를 이용하면 자주 바뀌는 요구사항에 효과적으로 대응할 수 있다.

동작 파라미터화란 아직은 어떻게 실행할 것인지 결정하지 않은 코드 블록

**2.1 변화하는 요구사항에 대응하기**

요구 사항 : 녹색 사과만 필터링

// 자바8 이전의 방식  
public static List<Apple> filterGreenApples(List<Apple> inventory) {  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple : inventory) {  
 if (**"green".equals(apple.getColor())**) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

변경 사항 : 빨간 사과도 필터링

🡺앞으로 어두운 빨간색, 노란색 등 필요할 때마다 메서드를 추가할 수 없기 때문에 비슷한 코드를 구현할 경우 추상화 시켜라

**2.1.2 두번째 시도 : 색을 파라미터화**

public static List<Apple> filterApplesByColor(List<Apple> inventory, String color) {  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple : inventory) {  
 if (**apple.getColor().equals(color)**) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

🡺파라미터로 색을 받을 수 있도록 변경

호출부

// 파라미터화  
*filterApplesByColor*(inventory, "green");  
*filterApplesByColor*(inventory, "red");

변경 사항 : 150그램 이상의 무게가 나가는 사과를 무거운 사과로 구분

public static List<Apple> filterApplesByWeight(List<Apple> inventory, int weight) {  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple : inventory) {  
 if (apple.getWeight() > weight) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

🡺색처럼 무게도 받을 수 있도록 설정

**문제점 도출**

하지만 대부분 코드가 색 필터링 코드와 중복됨

**2.1.3 세번째 시도 : 가능한 모든 속성으로 필터링**

public static List<Apple> filterApples(List<Apple> inventory, String color, int weight, boolean flag) {  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple : inventory) {  
 if (**flag && apple.getColor().equals(color) ||  
 (!flag && apple.getWeight() > weight)**) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

🡺flag를 통해 구분했다.

호출부

// 플래그 추가  
*filterApples*(inventory, "green", 0, true);  
*filterApples*(inventory, "", 150, false);

**문제점 도출**

🡺요구사항이 바뀌었을 때 유연하게 대응할 수도 없다.

🡺true, false의 의미도 잘 모르겠다.

**2.2 동작 파라미터화**

좀더 유연하게 대응할 수 있는 방법이 절실하다.

수학에서는 인수로 값을 받아 true나 false로 반환하는 함수를 **프레디케이트**라고 한다.

전략 패턴 : 각 알고리즘을 캡슐화하는 알고리즘 패밀리를 정의해둔 다음에 런타임에 알고리즘을 선택하는 기법

전략 패턴은 갑작스런 알고리즘의 변화에도 유연하게 대처가능. 알고리즘을 캡슐화하기 때문이다. 따라서 언제든지 알고리즘이 변경된다면 프로세스의 큰 틀을 바꾸지 않고도 유연한 프로그래밍 가능.

템플릿 메소드과 전략 패턴 차이

-템플릿 메소드 : 보통 공통된 기능은 부모 클래스에 설정해 놓고(큰틀), 서브 클래스에서는 알고리즘 단계에서 유연하게 바뀔 수 있는 부분만 의존

-전략 패턴 : 보통 인터페이스 군으로 설정. 다형성을 이용하여 런타임 시점에 알고리즘 선택

전략패턴에서의 캡슐화는 핵심 로직을 캡슐화 하는 것이고, 템플릿 메서드 패턴의 캡슐화는 일련의 알고리즘 단계들을 캡슐화 하는 것이다.

<http://asuraiv.blogspot.kr/2014/07/template-method-pattern-1_9.html>

**2.2.1 네번째 시도 : 추상적 조건으로 필터링**

첫번재 코드에 비해 유연하고 가독성도 좋아졌고 사용하기도 쉬워졌다.

// 추상조건 필터  
public static List<Apple> abstractFilterApples(List<Apple> inventory, **ApplePredicate p**) {  
 List<Apple> result = new ArrayList<>();  
 for (Apple apple : inventory) {  
 if (p.test(apple)) {  
 result.add(apple);  
 }  
 }  
 return result;  
}

호출부

// 추상화 추가  
*abstractFilterApples*(inventory, new AppleGreenColorPredicate());  
*abstractFilterApples*(inventory, new AppleHeavyWeightPredicate());

**문제점 도출**

아직 필요한 필터 기능이 생길때마다 Appledicate 하위 class를 구현해야 되는 문제가 남아 있다. (번거롭다)

중요 : 유연한 API를 위하여 동작 파라미터화는 꼭 필요 하다.

**2.3 복잡한 과정 간소화**

새로운 동작을 전달하려면 ApplePredicate 인터페이스를 구현하는 여러 클래스를 정의해야 한다.

자바는 클래스의 선언과 인스턴스화를 동시에 수행할 수 있도록 익명 클래스라는 기법을 제공한다.

🡺하지만 익명 클래스가 모든 것을 해결하는 것은 아니다.

**2.3.1 익명 클래스**

🡺자바의 지역 클래스와 비슷한 개념이다. (블록 내부에 선언된 클래스)

🡺익명 클래스를 이용하면 클래스 선언과 인스턴스화를 동시에 할 수 있다.

**2.3.2. 다섯 번째 시도 : 익명 클래스 사용**

// 익명 클래스  
*abstractFilterApples*(inventory, new ApplePredicate() {  
 @Override  
 public boolean test(Apple apple) {  
 return "red".equals(apple.getColor());  
 }  
});

🡺익명 클래스를 이용하여 메서드의 동작을 직접 파라미터화했다.

**문제점 도출**

1) 익명 클래스는 여전히 많은 공간을 차지한다.

2) 많은 프로그래머들이 익명 클래스의 사용에 익숙하지 않다.

🡺결국 객체를 만들고 명시적으로 새로운 동작을 정의하는 메서드를 구현해야 한다는 점은 변하지 않았다.

**2.3.3 여섯 번째 시도 : 람다 표현식 사용**

자바8의 람다 표현식을 이용해서 다음처럼 간단하게 재구현할 수 있다.

// 람다  
*abstractFilterApples*(inventory, (Apple apple) -> "red".equals(apple.getColor()));

그래프로



**2.3.4 일곱 번째 시도 : 리스트 형식으로 추상화**

현재 filterApples는 Apple과 관련한 동작만 수행한다. 하지만 Apple 이외의 다양한 물건도 필터링 작동하도록 리스트 형식을 추상화 할 수 있다.

// 리스트 형식의 추상화  
public static <T> List<T> filter(List<T> list, Predicate<T> p) {  
 List<T> result = new ArrayList<>();  
 for (T e : list) {  
 if (p.test(e)) {  
 result.add(e);  
 }  
 }  
 return result;  
}

호출부

// 리스트 형식으로 추상화  
filter(inventory, (Apple apple) -> "red".equals(apple.getColor()));

자바8 아니면 불가능한 일이다.

.

**2.4 실전 예제**

**2.4.1 Comparator로 정렬하기**

// 정렬 처리  
inventory.sort(new Comparator<Apple>() {  
 @Override  
 public int compare(Apple a1, Apple a2) {  
 return a1.getWeight().compareTo(a2.getWeight());  
 }  
});

**2.4.2 Runnable로 코드 블록 실행하기**

// Runnable 사용  
Thread t = new Thread(new Runnable() {  
 @Override  
 public void run() {  
 System.*out*.println("Hello world");  
 }  
});  
  
// 람다 표현식으로 사용  
Thread t2 = new Thread(() -> System.*out*.println("Hello world"));

**2.3.4 GUI 이벤트 처리하기**

Button button = new Button("Send");  
  
button.setOnAction(new EventHandler<ActionEvent>() {  
 @Override  
 public void handle(ActionEvent event) {  
 label.setText("Sent!!");  
 }  
});  
  
// 람다 방식으로  
button.setOnAction((Action event) -> label.setText("Sent!!"));

**CHAPTER 3 람다 표현식**

일반적으로 동작 파라미터화를 이용하면 더 유연하고 재사용할 수 있는 코드를 만들 수 있다.

람다를 통해 하고 유연한 코드를 구현하는 방법을 설명한다.

**3.1 람다란 무엇인가?**

람다 표현식은 메서드로 전달할 수 있는 익명 함수를 단순화한 것.

**-익명**

보통의 메서드와 달리 일므이 없으므로 익명이다.

**-함수**

메서드처럼 특정 클래스에 종속되지 않으므로 함수라 부른다.

**-전달**

람다 표현식을 메서드 인수로 전달하거나 변수로 저장할 수 있다.

**-간결성**

익명 클래스처럼 많은 자질구레한 코드를 구현할 필요가 없다.

Comparator<Apple> byWeight = new Comparator<Apple>() {  
 @Override  
 public int compare(Apple a1, Apple a2) {  
 return a1.getWeight().compareTo(a2.getWeight());  
 }  
};

람다를 이용한 새로운 코드

Comparator<Apple> byWeight =  
 (Apple a1, Apple a2) -> a1.getWeight().compareTo(a2.getWeight());

람다는 세 부분으로 이루어진다.

**1) 파라미터 리스트**

Comparator의 compare 메서드의 파라미터

**2) 화살표**

화살표(->)는 람다의 파라미터 리스트와 바디를 구분한다.

**3) 람다의 바디**

두 사과의 무게를 비교한다. 람다의 반환값에 해당하는 표현식이다.

1) 2) 3)

**(Apple a1, Apple a2)** -> a1.getWeight().compareTo(a2.getWeight());

자바8에서 지원하는 다섯가지 람다 표현식

1) String 형식의 파라미터 하나를 가지며 **int**를 반환한다. (return은 함축)

(String s) -> s.length()

2) Apple 형식의 파라미터 하나를 가지며 **boolean**을 반환 한다.

(Apple a) -> a.getWeight() > 150

3) int 형식의 파라미터 두개를 가지며 리턴값이 없다. 여러 행의 문장을 포함할 수 있다.

(int x, int y) -> {  
 System.*out*.println(“Result:”);  
 System.*out*.println(x+y);  
}

4) 파라미터가 없으며 int를 반환

() -> 42

5) Apple 형식의 파라미터를 두개 가지며 int를 반환한다.

(Apple a1, Apple a2) -> a1.getWeight().compareTo(a2.getWeight())

**매우 중요!!**

람다의 기본 문법

1) **(parameters) -> expression**

ex)

() -> “Iron Man”

(String s) -> “Iron Man”

**OR(또는)**

2) (**parameters) -> { statements; }**

ex)

() -> {}

() -> { return “Iron Man”; }

(String s) -> { return “Iron Man”; }

(Integer i) -> { return “Alan” + I }

**3.2 어디에, 어떻게 람다를 사용할까?**

람다는 함수형 인페이스(Functional Interface) 라는 문맥에서 람다 표현식을 사용할 수 있다.

**3.2.1 함수형 인터페이스**

앞에서 본 Predicate<T>가 함수형 인터페이스이다. 오직 **하나의 추상 메서드**만 지정하기 때문이다.

@FunctionalInterface  
public interface Predicate<T> {  
boolean test(T t);

}

함수형 인터페이스는 정확히 하나의 추상 메서드를 지정하는 인터페이스이다.

🡺많은 디폴트 메서드가 있더라도 추상 메서드가 오직 하나면 함수형 인터페이스이다.

함수형 인터페이스로 뭘 할 수 있을까?

🡺람다 표현식으로 함수형 인터페이스의 추상 메서드 구현을 직접 전달할 수 있으므로 전체 표현식을 **함수형 인터페이스의 인스턴스로 취급 (내부에서 추상 메서드를 구현)**

🡺함수형 인터페이스는 람다 뿐만 아니라 **익명 내부 클래스로도 같은 기능 구현 가능**

public static void main(String[] args) {  
  
 Runnable r1 = () -> System.*out*.println("Hello World 1");  
  
 Runnable r2 = new Runnable() { // 내부 클래스  
 @Override  
 public void run() {  
 System.*out*.println("Hello World 2");  
 }  
 };  
  
 *process*(r1);  
 *process*(r2);  
 *process*(() -> System.*out*.println("Hello World 3"));  
  
}  
  
public static void process(Runnable r) {  
 r.run();  
}

**3.2.2 함수 디스크립터**

함수형 인터페이스의 추상 메서드 **시그너처**는 람다 표현식의 **시그너처** 를 가르킨다.

🡺람다 표현식의 시그너처를 서술하는 메서드를 **함수 디스크립터** 라고 부른다.

Ex) Runnable 인터페이스의 유일한 추상 메서드 run은 인수와 반환값이 없으므로 인수와 반환값이 없는 시그너처로 생각할 수 있다.

Remind!) 람다 표현식

🡺변수에 할당할 수 있다. (일급 시민)

🡺함수형 인터페이스를 인수로 받는 메서드로 전달할 수 있다.

🡺함수형 인터페이스의 추상 메서드와 같은 시그너처를 갖는다.

**왜 함수형 인터페이스를 인수로 받는 메서드에만 람다 표현식을 사용할 수 있을까? (나도 궁금)**

🡺자바 프로그래머가 하나의 추상 메서드를 갖는 인터페이스에 익숙하다

🡺언어를 더 복잡하게 만들지 않는 현재의 방법 선택

**@FunctionalInterface란?**

함수형 인터페이스를 가리키는 어노테이션이다. 추상 메서드가 하나만 선언.

**3.3 람다 활용 : 실행 어라운드 패턴**

아래와 같이 설정과 정리 과정에 둘러 싸여 있는 형태를 **실행 어라운드 패턴**(execute around pattern)

|  |
| --- |
| 초기화/준비 코드 |
| 작업 A |
| 정리/마무리 코드 |

try-with-resources (자동 자원 반환)

public static String processFile() throws IOException {  
 try (BufferedReader br = new BufferedReader(new FileReader("data.txt"))) {  
 return br.readLine();  
 }  
}

**3.3.1 1단계:동작 파라미터화를 기억하라**

위 코드는 파일에서 한번에 한줄만 읽을 수 있다. 두줄을 읽거나 자주 사용되는 단어를 반환 하려면?

🡺processFile을 동작 파라미터화 시킨다.

String result = *processFile*((BufferedReader br) ->  
 br.readLine() + br.readLine());

🡺BufferedReader를 인수로 받아서 String을 반환 하는 람다가 필요하다.

**3.3.2 2단계:함수형 인터페이스를 이용해서 동작 전달**

BufferedReader -> String과 IOException을 던질 수 있는 시그너처와 일치하는 함수형 인터페이스를 만들어야 한다.

@FunctionalInterface  
public interface BufferedReaderProcessor {  
 String process(BufferedReader b) throws IOException; // 이게 함수 디스크립터 인감?  
}

정의한 인터페이스를 processFile 메서드의 인스로 전달할 수 있다. (일급 객체)

public static String processFile(**BufferedReaderProcessor p**) throws IOException {  
  
}

**3.3.3 3단계:동작 실행!**

이제 BufferedReaderProcessor에 정의된 process 메서드의 시그너처(BufferedReader -> String)와 일치하는 람다를 전달할 수 있다.

public static String processFile(BufferedReaderProcessor p) throws IOException {  
 try (BufferedReader br = new BufferedReader(new FileReader("data.txt"))) {  
 return p.process(br);  
 }  
}

**3.3.4 4단계:람다 전달**

람다를 이용해서 다양한 동작을 메서드로 전달할 수 있다.

String oneLine = *processFile*((BufferedReader br) -> br.readLine());  
String towLines = *processFile*((BufferedReader br) -> br.readLine() + br.readLine());

**3.4 함수형 인터페이스 사용**

함수형 인터페이스는 오직 하나의 추상 메서드를 지정한다. 함수형 인터페이스의 추상 메서드는 람다 표현식의 시그너처를 묘사한다.

함수형 인터페이스의 추상 메서드 시그너처를 함수 **디스크립터**라고 한다.

내가 이해한거는

@FunctionalInterface 내에 추상 메서드

String process(BufferedReader b) 의 형태와 (반환 타입 String, 인자 타입 BufferedReader)

과 람다 호출하는 형태의 타입과 같아야 한다.

(BufferedReader br) -> br.readLine()) (반환 타입 String, 인자 타입 BufferedReader)

**3.4.1 Predicate**

따로 정의할 필요 없이 바로 사용할 수 있다는 점이 특징이다.

@FunctionalInterface  
public interface Predicate<T> {  
boolean test(T t);

}

**3.4.2 Consumer**

제네릭 형식 T를 받아서 void를 반환하는 추상메서드를 정의한다. T 형식의 객체를 인수로 받아서 어떤 동작을 수행하고 싶을 때 Consumer 인터페이스를 사용.

@FunctionalInterface  
public interface Consumer<T> {  
void accept(T t);

}

기본형

public static void main(String[] args) {  
 *forEach*(  
 Arrays.*asList*(1,2,3,4,5),  
 (Integer i) -> System.*out*.println(i)  
 );  
}  
  
public static<T> void forEach(List<T> list, Consumer<T> c) {  
 for(T i : list) {  
 c.accept(i);  
 }

실제 사용 예 TDA)

//선언부

public void etAppUpdateResponse(com.tmoncorp.tda.module.version.domain.Version lastAppActiveVersion, Consumer<Version> addProperty) {

**addProperty.accept(versionInfo);**

}

//호출부

appUpdateResponse.setAppUpdateResponse(

lastAppActiveVersion,

(versionInfo) ->

{

        if (paramVersion.getIsActive() == false) {

               versionInfo.setEssential(true);

              }

}

);

**3.4.3 Function**

제네릭 형식 T를 인수로 받아서 제네릭 형식 R 객체를 반환하는 apply 추상메서드 정의.

입력을 출력으로 매핑하는 람다를 정의할 때 Function 인터페이스를 활용할 수 있다.

(예를들어 사과의 무게 정보를 추출하거나 문자열을 길이와 매핑)

@FunctionalInterface  
public interface Function<T, R> {R apply(T t);

}

**기본형 특화**

자바의 모든 형식은 참조형 + 기본형 이다. 하지만 제네릭 파라미터에는 참조형만 사용할 수 있다.

🡺C#은 제약이 없고, 스칼라는 참조형만 있다.

기본형 🡺 참조형 (박싱)

참조형 🡺 기본형 (언박싱)

박싱과 언박싱이 자동으로 이루어지는 것을 오토박싱

🡺이런 과정들은 메모리 더 소모. 래퍼클래스는 힙에 저장되고 기본형을 가져올대도 메모리 탐색과정 필요

자바8에서는 기본형을 입출력으로 사용하는 상황에서 오토박싱 동작을 피할 수 있도록 **특별한 버전의 함수형 인터페이스를 제공**

내 생각 : 제네릭을 잘 사용하면 좀더 활용범위가 넓다.

|  |  |
| --- | --- |
| **함수형 인터페이스** | **함수 디스크립터** |
| Predicate<T> | T -> boolean |
| Consumer<T> | T -> void |
| Function<T, R> | T -> R |
| Supplier<T> | () -> T |
| UnarayOperator<T> | T -> T |
| BinaryOperator<T> | (T, T) -> T |
| BiPredicate<L, R> | (L, R) -> boolean |
| BiConsumer<T, U> | (T, U) -> void |
| BiFunction<T, U, R> | (T, U) -> R |

**Supplier () -> T**

Supplier<Date> dateSupplier= SupplierFunctionExample::getSystemDate;  
Date systemDate = dateSupplier.get();  
System.*out*.println("systemDate->" + systemDate);

public static Date getSystemDate() {  
 return new Date();  
}

🡺 결과 : Wed Dec 16 19:18:15 IST 2015

**UnarayOperator T -> T**

UnaryOperator<Integer> operator = t -> t \* 2;  
  
System.*out*.println(operator.apply(5));  
System.*out*.println(operator.apply(10));  
System.*out*.println(operator.apply(15));

🡺 10, 20, 30

**BinaryOperator (T, T) -> T**

BinaryOperator<Integer> adder = (n1, n2) -> n1 + n2;  
System.*out*.println(adder.apply(3, 4));

🡺 7

**BiPredicate (L, R) -> boolean**

BiPredicate<Integer, Integer> bi = (x, y) -> x > y;  
System.*out*.println(bi.test(2, 3));

🡺 false

**BiConsumer (T, U) -> void**

BiConsumer<String, String> biConsumer = (x, y) -> {  
 System.*out*.println(x);  
 System.*out*.println(y);  
};  
biConsumer.accept("java2s.com", " tutorials");

java2s.com

tutorials

**BiFunction (T, U) -> R**

BiFunction<String, String,String> bi = (x, y) -> {  
 return x + y;  
};  
  
System.*out*.println(bi.apply("java2s.com", " tutorial"));

Java2s.com tutorial

**예외, 람다, 함수형 인터페이스의 관계**

함수형 인터페이스는 확인된 예외를 던지는 동작을 허용하지 않는다.

🡺try/ catch를 사용하거나 예외를 선언하는 함수형 인터페이스를 직접 정의

1)인터페이스에 예외 직접 선언

@FunctionalInterface  
public interface BufferedReaderProcessor {  
 String process(BufferedReader b) throws IOException;  
}

2)try/catch 사용

Function<BufferedReader, String> f =  
 (BufferedReader b) -> {  
 try {  
 return b.readLine();  
 } catch (IOException e) {  
 throw new RuntimeException(e);  
 }  
 };

**3.5 형식 검사, 형식 추론, 제약**

람다 표현식 자체에는 람다가 어떤 함수형 인터페이스를 구현하는지의 정보가 포함되어 있지 않다. 따라서 람다 표현식을 더 제대로 이해하려면 람다의 실제 형식을 파악해야 한다.

**3.5.1 형식 검사**

람다가 사용되는 콘텍스트(context)를 이용해서 람다의 형식(type)을 추론할 수 있다.

어떤 콘텍스트(람다가 전달될 메서드 파라미터나 람다가 할당되는 변수 등) 에서 기대되는 람다 표현식의 형식을 **대상 형식(**target type) 이라고 부른다.

List<Apple>heavierthan105g = filter(inventory, (Apple a) -> a.getWeight() > 150);

1. filter 메서드의 선언을 확인

2. filter 메서드는 두 번째 파라미터로 Predicate<Apple> 형식을 기대한다.

3. Predicate<Apple>은 test라는 한 개의 추상 메서드를 정의하는 함수형 인터페이스

4. test 메서드는 Apple을 받아 boolean을 반환하는 함수 디스크립터를 묘사한다.

5. filter 메서드로 전달된 인수는 이와 같은 요구사항을 만족해야 한다.

**3.4.2 같은 람다, 다른 함수형 인터페이스**

대상 형식이라는 특징 때문에 같은 람다 표현식이더라도 호환되는 추상 메서드를 가진 다른 함수형 인터페이스로 사용될 수 있다.

@FunctionalInterface  
public interface Callable<V> {V call() throws Exception;  
}

public interface PrivilegedAction<T> {T run();  
}

**호출부**

public static void main(String[] args) {  
 Callable<Integer> c = () -> 42;  
 PrivilegedAction<Integer> p = () -> 42;  
}

🡺둘다 인수를 받지 않고 제네릭 형식 T를 반환하는 함수를 정의한다. 둘다 유효한 코드

**람다**

콘텍스트를 이용해서 람다의 형식을 추론 filter(inventory, **(Apple a) -> a.getWeight() > 150**);

**다이아몬드 연산자 (<>)**

다이아몬드 연산자로 콘텍스트에 따른 제네릭 형식을 추론 List<String listOfStrings = new **ArrayList<>();**

**특별한 void 호환 규칙**

람다의 바디에 일반 표현식이 있으면 void를 반환하는 함수 디스크립터와 호환이 된다.

물론 파라미터 리스트도 호환되어야 한다.

Predicate<String> p = s -> list.add(s); // boolean  
Consumer<String> b = s -> list.add(s); // void

🡺list.add는 boolean을 반환하지만 void랑도 호환이 된다.

**3.5.3 형식 추론**

자바 컴파일러는 람다 표현식이 사용된 콘텍스트(대상 형식)를 이용해서 람다 표현식과 관련된 함수형 인터페이스를 추론한다.

대상 형식 => 함수 디스크립터를 알 수 있음 => 람다의 시그너처도 추론할 수 있음

*filter*(inventory, (Apple apple) -> "red".equals(apple.getColor()));

🡺  
*filter*(inventory, apple -> "red".equals(apple.getColor()));

🡺컴파일러는 표현식의 파라미터 형식에 접근할 수 있으므로 람다 문법에서 이를 생략할 수 있다.

🡺어떤게 좋은지 정해진 규칙은 없음. 개발자가 가독성 좋은걸로 선택

**3.5.4 지역 변수 사용**

지금까지 람다 표 현식은 인수를 자신의 바디 안에서만 사용했다.

하지만 람다 표현식에서 익명 함수가 하는 것처럼 **자유 변수 (외부 정의된 변수)**를 활용할 수 있다.

🡺람다 캡처링이라고 부른다.

int portNumber = 1337;  
Runnable r = () -> System.*out*.println(portNumber);

인스턴스 변수와 정적 변수를 자유롭게 사용할 수 있다. 그러려면 **지역 변수는 명시적으로 final로 선언**되어 있어야 하거나 **실질적으로 final로 선언된 변수와 똑같이 사용**되어야 한다.

**지역 변수의 제약**

인스턴스 변수는 힙에 저장되는 반면 지역 변수는 스택에 위치한다.

스레드에서 실행 시, 지역 변수 접근하려고 할 수 있기 때문에(지역 변수는 이미 해제되었을 수도 있다.) 그래서 지역 변수의 복사본을 제공한다. 따라서 복사본의 값이 변경되지 않아야 하므로 제약이 생긴 것이다. 또한 값이 변경되면 병렬 처리 때 방해요소가 된다.

**클로저**

람다와 클로저의 공통점은 메서드의 인수로 전달될 수 있으며 자신의 외부 영역의 변수에 접근할 수 있다.

차이점은 람다는 정의된 메서드의 지역 변수의 값은 바꿀 수 없다. (지역 변수 값은 final)

**3.6 메서드 레퍼런스**

메서드 레퍼런스는 특정 람다 표현식을 축약한 것이라고 생각하면 된다.

inventory.sort((a1, a2) -> a1.getWeight().compareTo(a2.getWeight()));  
🡺  
inventory.sort(Comparator.*comparing*(Apple::getWeight));

**3.6.1 요약**

명시적으로 메서드명을 참조함으로써 가독성을 높일 수 있다. (::) 형태로 사용.

Ex)

|  |  |
| --- | --- |
| 람다 | 메서드 레퍼런스 단축표현 |
| (Apple a) -> a.getWeight() | Apple::getWeight |
| () -> Thread.currentThread().dumpStak() | Thread.currentThread()::dumpStack |
| (str, i) -> str.substring(i) | String::substring |
| (String s) -> System.out.println(s) | System.out::println |

🡺메서드 레퍼런스는 새로운 기능이 아니라 하나의 메소드를 참조하는 람다를 편리하게 표현할 수 있는 문법

**메서드 레퍼런스를 만드는 방법 (3가지)**

**1) 정적 메서드 래퍼런스**

🡺예를 들면 Integer의 parseInt 메서드는 Integer::parseInt로 표현할 수 있다.

**2) 다양한 형식의 인스턴스 메서드 레퍼런스**

🡺예를 들어 String의 length 메서드는 String::length로 표현할 수 있다.

**3) 기존 객체의 인스턴스 메서드 레퍼런스**

🡺Transaction expensiveTransaction 이 있으면 expensiveTransaction::getValue 로 표현 가능

List<String> str = Arrays.*asList*("a", "b", "A", "B");  
  
str.sort((s1, s2) -> s1.compareToIgnoreCase(s2)); // 람다 표현식  
str.sort(String::compareToIgnoreCase); // 메서드 레퍼런스

🡺컴파일러는 람다 표현식의 형식을 검사하던 방식과 비슷한 과정으로 메서드 레퍼런스가 주어진 함수형 인터페이스와 호환 하는지 확인한다.

**3.6.2 생성자 레퍼런스**

ClassName::new 처럼 클래스명과 new 키워드를 이용해서 기존 생성자의 레퍼런스를 만들 수 있다. (정적 메서드 레퍼런스를 만드는것과 비슷)

**Supplier () -> T**

Supplier<Apple> c1 = () -> new Apple();  
Apple a1 = c1.get();  
🡺 생성자 레퍼런스 방식 (인자가 없는)  
Supplier<Apple> c2 = Apple::new;  
Apple a2 = c2.get();

**Function T -> R**

Function<Integer, Apple> f1 = (weight) -> new Apple(weight);  
Apple b1 = f1.apply(110);  
🡺 생성자 레퍼런스 방식 (인자가 있는)  
Function<Integer, Apple> f2 = Apple::new;  
Apple b2 = f2.apply(110);

Apple(int weight) 를 호출

// 호출부

List<Integer> weights = Arrays.*asList*(7, 3, 4, 10);  
List<Apple> apples = *map*(weights, Apple::new);  
  
  
public static List<Apple> map(List<Integer> list, Function<Integer, Apple> f) {  
 List<Apple> result = new ArrayList<>();  
 for(Integer e : list) {  
 result.add(f.apply(e));  
 }  
 return result;  
}

**BiFunction (T, U) -> R**

// BiFuntion  
BiFunction<String, Integer, Apple> g1 =  
 (color, weight) -> new Apple(color, weight);  
Apple c1 = g1.apply("green", 110);  
  
BiFunction<String, Integer, Apple> g2 = Apple::new;  
Apple c2 = g2.apply("green", 110);

인스턴스화 하지 않고도 생성자에 접근할 수 있게 다양하게 응용할 수 있다.

static Map<String, Function<Integer, Fruit>> *map* = new HashMap<>();  
  
static {  
 *map*.put("apple", Apple::new);  
 *map*.put("orange", Orange::new);  
}  
  
public static Fruit giveMeFruit(String fruit, Integer weight) {  
 return *map*.get(fruit.toLowerCase())  
 .apply(weight);  
}

3개 이상 생성자일 경우 직접 만들어서 사용해야 한다.

**3.7 람다, 메서드 레퍼런스 활용하기**

동작 파라미터화, 익명 클래스, 람다 표현식, 메서드 레퍼런스등 정리

**3.7.1 1단계:코드 전달**

Comparator 객체를 인수로 받아 비교

public static void main(String... args) {  
 List<Apple> inventory = new ArrayList<>();  
 inventory.addAll(Arrays.*asList*(new Apple(80, "green"), new Apple(155, "green"), new Apple(120, "red")));  
  
 inventory.sort(new AppleComparator());  
}  
  
static class AppleComparator implements Comparator<Apple> {  
 public int compare(Apple a1, Apple a2) {  
 return a1.getWeight().compareTo(a2.getWeight());  
 }  
}

**3.7.2 2단계:익명 클래스 사용**

위 코드는 Comparator 코드를 구현해야 되기 때문에 익명 클래스를 이용하는 것이 좋다.

// 익명 클래스  
inventory.sort(new Comparator<Apple>() {  
 @Override  
 public int compare(Apple a1, Apple a2) {  
 return a1.getWeight().compareTo(a2.getWeight());  
 }  
});

**3.7.3 3단계:람다 표현식 사용**

람다 표현식을 이용하여 경량화된 문법을 이용해서 코드 전달. 함수형 인터페이스를 사용하는 곳 어디서나 람다 표현식 사용 가능.

추상 메서드의 시그너처(함수 디스크립터)는 람다 표현식의 시그너처를 정의 한다.

Comparator는 (T, T) -> int다

// 람다  
inventory.sort((a1, a2) -> a1.getWeight().compareTo(a2.getWeight()));

형식 추론

컴파일러에서 콘텍스트를 활용해서 형식 추론

(Apple a1, Apple a2) 🡺 (a1, a2) 로 줄일 수 있다.

Comparing 메서드를 사용하여 더 간소화

// Comparator의 comparing을 이용하여 더 간소화  
inventory.sort(Comparator.*comparing*((a) -> a.getWeight()));

**3.7.4 4단계:메서드 레퍼런스 사용**

// 메서드 레퍼런스  
inventory.sort(Comparator.*comparing*(Apple::getWeight));

코드만 짧아진 것이 아니라 의미도 명확해졌다.

**3.8 람다 표현식을 조합할 수 있는 유용한 메서드**

간단한 여러 개의 람다 표현식을 조합해서 복잡한 람다 표현식을 만들 수 있다.

디폴트 메서드 사용

**3.8.1 Comparator 조합**

역정렬

// 역정렬  
inventory.sort(Comparator.*comparing*(Apple::getWeight).reversed());

**Comparator 연결**

무게가 같은 두 사과가 존재한다면 어떻게 해야 할까? 이럴 땐 비교 결과를 더 다듬을 수 있는 두번째 Comparator를 만들 수 있다.

// 두번째 정렬  
inventory.sort(Comparator.*comparing*(Apple::getWeight).reversed().thenComparing(Apple::getColor));

🡺thenComparing 메서드로 두번째 비교자를 만들 수 있다.

**3.8.2 Predicate 조합**

Predicate는 복잡한 프레디케이트를 만들 수 있도록 negate, and, or 세가지 메서드를 제공한다.

// negate 빨간 사과가 아니다.  
Predicate<Sorting.Apple> notRedApple = redApple.negate();  
  
// and 빨간 사과이면서 무게가 150이상이다.  
Predicate<Sorting.Apple> redAndHeavyApple = redApple.and(a -> a.getWeight() > 150);  
  
// or 빨간 사과이면서 무게가 150이상 또는 그냥 녹색 사과  
Predicate<Sorting.Apple> redAndHeavyAppleOrGree = redApple.and(a -> a.getWeight() > 150)  
 .or(a -> "green".equals(a.getColor()));

**3.8.3 Function 조합**

andThen, compose 두가지 디폴트 메서드를 제공한다.

**andThen** 메서드는 주어진 함수를 먼저 적용한 결과를 다른 함수의 입력으로 전달하는 함수를 반환

Function<Integer, Integer> f = x -> x + 1;  
Function<Integer, Integer> g = x -> x \* 2;  
Function<Integer, Integer> h = f.andThen(g);  
int result = h.apply(1); // (1 + 1) \* 2 = 4

**Compose** 메서드는 인수로 주어진 함수를 먼저 실행한 다음에 그 결과를 외부 함수의 인수로 제공한다.

Function<Integer, Integer> f = x -> x + 1;  
Function<Integer, Integer> g = x -> x \* 2;  
Function<Integer, Integer> h = f.compose(g);  
int result = h.apply(1); // 1 + (1 \* 2) = 3

🡺andThen과 순서 반대

**3.9 비슷한 수학적 개념**

**3.9.1 적분**

**3.9.2 자바8 람다로 연결**

**3.10 요약**

-람다 표현식은 익명 함수의 일종이다. 이름은 없지만, 파라미터 리스트, 바디 반환 형식을 가지며 예외를 던질 수 있다.

-람다 표현식으로 간결한 코드를 구현할 수 있다.

-함수형 인터페이스는 하나의 추상 메서드만을 정의하는 인터페이스다.

-함수형 인터페이스를 기대하는 곳에서만 람다 표현식을 사용할 수 있다.

-람다 표현식을 이용해서 함수형 인터페이스의 추상 메서드를 즉석으로 제공할 수 있으며 람다 표현식 전체가 함수형 인터페이스의 인스턴스로 취급된다.

-java.util.function패키지는 Predicate<T>, Function<T, R>, Supplier<T>, Consumer<T>, BinaryOperator<T> 등을 포함해서 자주 사용하는 다양한 함수형 인터페이스를 제공한다.

-자바8은 Predicate<T>와 Function<T, R> 같은 제네릭 함수형 인터페이스와 관련한 박싱 동작을 피할 수 있도록 IntPredicate, IntToLongFunction 등과 같은 기본형 특화 인터페이스도 제공한다.

-실행 어라운드 패턴(자원할당, 자원 정리 등 코드 중간에 실행해야 하는 메서드에 꼭 필요한 코드)을 람다와 활용하면 유연성과 재사용성을 추가로 얻을 수 있다.

-람다 표현식의 기대 형식을 대상 형식이라고 한다.

-메서드 레퍼런스를 이용하면 기존의 메서드 구현을 재사용하고 직접 전달할 수 있다.

-Comparator, Predicate, Function 같은 함수형 인터페이스는 람다 표현식을 조합할 수 있는 다양한 디폴트 메서드를 제공한다.

**3장의 용어정리**

**대상형식**

람다 표현식으로 구현하고자 하는 함수형 인터페이스.

Function<String> f = str -> "hello"+str; 에서 **Funtion<String>**이 대상형식이 됨

**메서드 시그너처 and 함수 디스크립터**

메서드 시그너처를 람다 표현식 형태로 표현한 것.

**boolean test(T t)** 가 메서드 시그너처라면 함수 디스크립터로 표현할시 T -> boolean 과 같은 형태로 표현할 수 있음

**PART 2. 함수형 데이터 처리**

**CHATER 4. 스트림 소개**

-질의어 / ( select \* from table where color = “red” 🡺 반복자 필요 없다. );

-멀티코어 / (멀티코어 아키텍쳐 병렬 처리<fork join> 🡺 복잡하고 어렵다. )

컬렉션을 많이 사용하지만 완벽한 컬렉션 관련 연산을 지원하려면 한참 멀었다. (점점 발전..)

🡺ex) 교집합, 합집합 차집합

**Java**

// 합집합

public <T> List<T> union(List<T> list1, List<T> list2) {  
 Set<T> set = new HashSet<T>();  
  
 set.addAll(list1);  
 set.addAll(list2);  
  
 return new ArrayList<T>(set);  
}  
// 교집합  
public <T> List<T> intersection(List<T> list1, List<T> list2) {  
 List<T> list = new ArrayList<T>();  
  
 for (T t : list1) {  
 if (list2.contains(t)) {  
 list.add(t);  
 }  
 }  
  
 return list;  
}

**Java (apache CollectionUtils)**

// 차집합

Collection<String> except = CoolectionUtils.usbtract(list, list2);

// 교집합

Collection<String> intersection = CoolectionUtils.union(list, list2);

// 합집합

Collection<String> union = CoolectionUtils.intersecation(list, list2);

**C# Linq**

int[] first = { 2, 4, 6, 8, 10 };  
int[] second = { 3, 6, 9, 12 };  
  
// 차집합

var except = first.Except(second); // 2, 4, 8, 10  
  
// 교집합  
var intersection = first.Intersect(second); // 6  
  
// 합집합  
var union = first.Union(second); // 12, 10, 9, 8, 6, 4, 3, 2

**4.1 스트림이란 무엇인가?**

-질의어 가능

-병렬 처리 쉽게 가능

**자바7 이하**

// 자바7 스타일  
ArrayList<Dish> menu = new ArrayList<>();  
menu.add(new Dish("스테이크", 500));  
menu.add(new Dish("라면", 450));  
menu.add(new Dish("김치찌개", 300));

// 필터

List<Dish> lowCaloricDishes = new ArrayList<>();  
for (Dish d : menu) {  
 if (d.getColories() < 400) {  
 lowCaloricDishes.add(d);  
 }  
}

// 정렬  
Collections.*sort*(lowCaloricDishes, new Comparator<Dish>() {  
 @Override  
 public int compare(Dish d1, Dish d2) {  
 return Integer.*compare*(d1.getColories(), d2.getColories());  
 }  
});  
  
// 특정 형태로 변환.  
List<String> lowCaloricDishesName = new ArrayList<>();  
for (Dish d : lowCaloricDishes) {  
 lowCaloricDishesName.add(d.getName());  
}

**🡺연산을 여러 번 해야되기 때문에 코드가 길어진다.**

**자바8 (Stream API는 고수준 빌딩 블록)**

// 자바8 스타일  
List<String> lowCaloricDishesName2 = menu.parallelStream() // 병렬  
 .filter(d -> d.getColories() < 400) // 필터  
 .sorted(Comparator.*comparing*(Dish::getColories)) // 정렬  
 .map(Dish::getName) // 특정 형태로 변환  
 .collect(Collectors.*toList*()); // 출력 형태

****

🡺선언형으로 코드 구현

🡺빌딩 블록 연산을 연결해서 복잡한 데이터 처리 파이프라인 생성

**🡺주의 할 것이 빌딩 블록의 순서에 따라서 성능 영향을 미친다.**

**고수준 빌딩 블록**

🡺특정 스레딩 모델에 제한되지 않고 자유롭게 어떤 상황에서든 사용할 수 있다. (병렬성)

**구아바, 아파치, 람다j 왜 생겼을까?**

🡺자바가 지원을 안해서..

**선언형 이란?**

해법을 정의하기보다는 문제를 설명하는 고급언어이다

**그럼 고급언어는?**

사람이 이해하기 쉽게 작성된 프로그래밍 언어

**스트림 API 특징**

**1) 선언형**

🡺이해하기 쉽게 작성되었다. (가독성이 좋아짐)

**2) 조립할 수 있음**

🡺유연성 확장성이 좋아짐

**3)병렬화**

🡺상태가 불변하기 때문에 병렬화가 쉬워지고 그로 인해 성능이 좋아짐

**왜 컬렉션 대신 스트림을 사용할까?**

🡺결국 컬렉션으로는 불가능한 작업을 처리하기 위해

**4.2 스트림 시작하기**

**그래서 스트림이 뭐임?**



**자료의 입출력을 도와 주는 중간매개체**

**데이터 처리 연산** 을 지원하도록 **소스**에서 추출된 **연속된 요소** (잘 안와닿음)

**데이터 처리 연산**

순차적 또는 병렬로 filter, map, reduce, find match, sort (고수준)등을 이용해서 데이터를 연산

(컬렉션의 주제는 데이터고 스트림의 주제는 계산이다)

**소스**

컬렉션, 배열, I/O 등으로부터 데이터를 얻어오는 것이고 (순서 유지)

Files.lines("test.txt")  
Arrays.*stream*(array)

**연속된 요소**

연속된 값 집합의 인터페이스

List<String> stringList = Arrays.*asList*("test", "test2");  
stringList // 소스  
 .stream() // 소스에서 추출된 연속된 요소  
 .filter(x -> x.contains("test")) // 데이터 처리 연산 (파이프 라이닝 구성)  
 .collect(Collectors.*toList*());

🡺스트림 연산

**그 외** / **파이프 라이닝**

대부분의 스트림 연산은 스트림 연산끼리 연결해서 커다란 파이프라인을 구성

(lazy, short circuit)



**그 외2 / 내부 반복**

for loop (외부 반복) 반복 반대.. 자세한건 뒤에서

**결론 스트림 API는 파이프라인을 더 최적화할 수 있는 유연성 제공**

**4.3 스트림과 컬렉션**

컬렉션과 새로운 스트림 모두 **연속된** 요소 형식의 값을 저장하는 자료구조의 인터페이스 제공

🡺**연속된** : 순차적으로 **값**에 접근 (iterator)

**map은 Collection 아님** / 연속된 요소 없기 때문에 entrySet, entryKey 등으로 변환후 작업

new HashMap<>().entrySet().stream()

**컬렉션 :** 추가, 삭제 가능 하며, 메모리에 다 가지고 있는다.

(DVD)

**스트림 :** 추가, 삭제 불가 하며, 요청할 때만 필요한 요소를 계산하는 고정된 자료구조.

(인터넷 스트리밍)

문제) 무한의 수가 저장되어 있는 데이터셋에서 10개를 꺼내서 합을 구하시오

**Collcetion**

List<Integer> list = new ArrayList<>();  
  
for(int i = 0; i < 100000000000000000; i ++) {  
 list.add(i); // 메모리에 적재를 해야 되기 때문에 불가  
}  
  
int sum = 0;  
for(int i = 0; i < 10; i ++) {  
 sum += list.get(i);  
}

**Stream**

int totalSum = IntStream.*range*(0, 100000000000000000)  
 .limit(10)  
 .sum();

**4.3.1 딱 한번만 탐색할 수 있다.**

ex) Stream Deep copy…

탐색된 스트림의 요소는 소비. 다시 탐색하려면 초기 데이터 소스에서 새로운 스트림 생성

List<String> title = Arrays.*asList*("Java8", "In", "Action");  
Stream<String> s = title.stream();  
s.forEach(System.*out*::println);  
s.forEach(System.*out*::println);

**접근시, 해당 에러 나옴**

Exception in thread "main" java.lang.IllegalStateException: stream has already been operated upon or closed

**4.3.2 외부 반복과 내부 반복**

컬렉션 : 사용자가 직접 요소를 반복 해야 된다 - **외부 반복** (for-each)

스트림 : 반복을 알아서 처리하고 결과값을 어딘가에 저장해 준다 – **내부 반복**

// for-each (루프 외부 반복)  
List<String> names = new ArrayList<>();  
for(Dish d: menu) {  
 names.add(d.getName());  
}  
  
// iterator (내부적으로 숨겨졌던 반복자를 사용한 외부 반복)  
List<String> names2 = new ArrayList<>();  
Iterator<Dish> iterator = menu.iterator();  
while(iterator.hasNext()) {  
 Dish d = iterator.next();  
 names2.add(d.getName());  
}  
  
// stream (내부반복)  
List<String> names3 = menu.stream().map(Dish::getName).collect(Collectors.*toList*());

🡺내부 반복을 사용하면 투명하게 병렬로 처리하거나 최적화된 다양한 순서로 처리

🡺아무래도 외부 반복은 노출되어 있기 때문에 상태 변화 가능

🡺filter, map 등 순서에 따라 성능이 달라짐

**4.4 스트림 연산**

List<String> name = menu.stream() // 스트림 생성  
 .filter(d -> d.getCalories() > 300) // 중간 연산 (파  
 .map(Dish::getName) // 중간 연산 이  
 .limit(3) // 중간 연산 프 라인 형성)  
 .collect(Collectors.*toList*()); // 최종 연산 (파이프 라인 실행 후 해제)

**4.4.1 중간 연산**

최종 연산 하기 전에 아무 연산도 수행하지 않는다. (lazy)

최종 연산에서 중간연산을 합쳐서 한꺼번에 처리 한다.

// 중간 연산  
List<String> name =  
 menu.stream()  
 .filter(d -> {  
 System.*out*.println("filtering : " + d.getName());  
 return d.getCalories() > 500;  
 })  
 .map(d ->{  
 System.*out*.println("mapping : " + d);  
 return d.getName();  
 })  
 .limit(3)  
 .collect(Collectors.*toList*());

🡺쇼트셔킷 기법 때문에 처음 3개만 선택됨

ex)쇼트셔킷 논리연산자 and, or

🡺filter, map은 서로 다른 연산이지만 한과정으로 병합 ( 루프 퓨전 – 수직적 처리 )

**4.4.2 최종연산**

최종 연산은 보통 List(collect), Integer(count), void(foreach)등 결과가 반환

**4.4.3 스트림 이용하기**

List<String> name = menu.stream() // 1) 스트림 생성  
 .filter(d -> d.getCalories() > 300) // 2) 중간 연산   
 .map(Dish::getName) // 중간 연산   
 .limit(3) // 중간 연산  
 .collect(Collectors.*toList*()); // 3) 최종 연산

빌더 패턴과 비슷

PersonInfo result = personInfoBuilder

.setName("MISTAKE")

.setAge(20)

.setFavoriteAnimal("cat")

.setFavoriteColor("black")

.setFavoriteNumber(7)

// 마지막에 .build() 메소드를 호출해서 최종적인 결과물을 만들어 반환

.build();

**4.5 요약**

-내부 반복지원 (투명성)

-스트림 생성, 중간연산, 최종연산

-중간 연산자는 결과 생성 못함 🡺 최종연산에서 종합

-스트림의 요소는 요청할 때만 계산

**CHATER 5. 스트림 활용**

앞에서 filter를 이용하여 내부 반복을 외부 반복으로 바꾸는 방법을 보았다.

**// 외부 반복**  
List<Dish> vegetarianDishes = new ArrayList<>();  
for (Dish d : menu) {  
 if (d.isVegetarian()) {  
 vegetarianDishes.add(d);  
 }  
}  
  
**// 내부 반복**  
List<Dish> vegetarianDishes2 = menu.stream()  
 .filter(Dish::isVegetarian)  
 .collect(Collectors.*toList*());

**5.1 필터링과 슬라이싱(나눈다)**

**5.1.1 프레디케이트로 필터링**

Stream<T> filter(**Predicate**<? super T> predicate);  
  
@FunctionalInterface  
public interface Predicate<T> {  
 boolean test(T t);  
}

프레디케이트는 불린을 반환 하는 함수이기 때문에 Dish::isVegetarian 을 함수 디스크립션으로 사용 가능

**5.1.2 고유 요소 필터링**

menu.stream()  
 .filter(Dish::isVegetarian)  
 **.distinct()** .forEach(System.*out*::println);

distinct를 이용하여 고유 요소 반환 가능 (hashCode, equals가 같아야 한다)

@Override  
public boolean equals(Object o) {  
 if (this.name.equals(((Dish) o).name))  
 return true;  
 return false;  
}  
  
@Override  
public int hashCode() { return name.hashCode(); }

**5.1.3 스트림 축소**

// filtering  
menu.stream()  
 .filter(d -> d.getCalories() > 300)  
 **.limit(3)** .collect(Collectors.*toList*());

n개의 요소 반환

Set 같이 정렬되지 않은 스트림에 사용되면 limit도 정렬되지 않으 상태로 반환 (숏셔킷 이점)

**5.1.4 요소 건너뛰기**

menu.stream()  
 .filter(d -> d.getCalories() > 300)  
 **.skip(2)** .collect(Collectors.*toList*());

n개의 요소 제외

skip과 limit는 상호 보완적이다.

**5.2 매핑**

객체에서 특정 데이터를 선택하는 작업 (ex SQL 테이블에서 특정 열만 선택)

**5.2.1 스트림의 각 요소에 함수 적용하기**

함수를 적용한 결과가 새로운 요소로 매핑된다 (고치는게 아니라 새로운 버전을 만든다.)

// map  
menu.stream()

.map(Dish::getName)

.collect(Collectors.*toList*());

특정 컬럼의 길이를 알고 싶을때는 map 메서드를 연결해서 구할 수 있다.

// 요리 명의 길이  
menu.stream()

.**map(Dish::getName)**

**.map(String::length)** .collect(Collectors.*toList*());

**5.2.2 스트림 평면화**

최종 목표는 [“Hello”, “World”] 🡺 [“H”, “e”, “l”, “o”, “W”, “r”, “d”]

🡺 map 형태의 반환이 Stream<String>로 되어야 한다.

**참고) Split의 String[]을 반환 한다.**

public String[] split(String regex) {  
 return split(regex, 0);  
}

**첫번째 시도**

strList.stream()  
 .**map(x -> x.split(""))** // stream<String[]> ["H", "e", "l", "l" "o"]["W", "o", "r", "l", "d"] .distinct()  
 .collect(Collectors.*toList*());

map 메소드가 반환한 스트림 형식은 Stream<String[]> 이다.

🡺distinct는 equals, hashCode 비교기 때문에 중복 제거를 할 수 없다.

**두번째 시도**

strList.stream()  
 .map(x -> x.split(""))  
 **.map(Arrays::*stream*)** .distinct()  
 .collect(Collectors.*toList*());

**Arrays::stream** : 문자열을 받아서 스트림을 만드는 메소드

배열 스트림 대신, 문자열 스트림이 필요하다.

map 메소드가 반환한 스트림 형식은 Stream<Stream>이다.

🡺결국 distinct 실패

List<Tmon> list = Arrays.*asList*(new Tmon("심준보", 32), new Tmon("조우현", 32), new Tmon("조우현", 32), new Tmon("송두리", 32));  
  
// map 으로 반환  
List<String> names = list.stream()  
 .map(Tmon::getName) // Stream<String> => 심준보, 조우현, 조우현, 송두리  
 .collect(Collectors.*toList*());  
  
// stream 형태로 반환  
String[] array = names.toArray(new String[names.size()]);  
Stream<String> streamOfName = Arrays.*stream*(array); // Stream<String> => 심준보, 조우현, 조우현, 송두리  
  
// distinct  
List<String> str = streamOfName  
 .distinct().collect(Collectors.*toList*());  
  
for (String s : str) {  
 System.*out*.println(s); // Stream<String> => 심준보, 조우현, 송두리  
}  
  
// split  
List<String> strList = Arrays.*asList*("Hello", "World");  
List<String[]> strArrayList = strList.stream()  
 .map((word -> word.split(""))) // Stream<String[]> => ["H","e","l","l","o"], ["W","o","r","l","d"]  
 .collect(Collectors.*toList*());  
  
// Arrays::stream  
List<Stream<String>> streamList = strList.stream()  
 .map(word -> word.split("")) // Stream<String[]> => ["H","e","l","l","o"], ["W","o","r","l","d"]  
 .map(wordArray -> Arrays.*stream*(wordArray)) // Stream<Stream<String>> => ?  
 .distinct()  
 .collect(Collectors.*toList*()); // List<Stream<String>>  
  
// flatMap  
List<String> streamList2 = strList.stream()  
 .map(word -> word.split("")) // Stream<String[]> => ["H","e","l","l","o"], ["W","o","r","l","d"]  
 .flatMap(wordArray -> Arrays.*stream*(wordArray)) // Stream<Stream<String>> => Stream<String>  
 .distinct()  
 .collect(Collectors.*toList*());  
  
  
List<String> a1 = new ArrayList<>();  
a1.add("test1");  
a1.add("test2");  
  
List<String> a2 = new ArrayList<>();  
a2.add("test2");  
a2.add("test3");  
  
// list 콘텐츠  
a1.addAll(a2);

**flatMap 사용**

strList.stream()  
 .map(x -> x.split(""))  
 .flatMap(Arrays::*stream*)  
 .distinct()  
 .collect(Collectors.*toList*());

flatMap은 스트림이 아닌 **스트림의 콘텐츠로 매핑한다**.

즉 flatMap은 하나의 평면화된 스트림을 반환한다.

**5.3 검색과 매칭**

**5.3.1 프레디케이트가 적어도 한 요소와 일치하는지 확인**

anyMatch

if(menu.stream().anyMatch(Dish::isVegetarian)) {  
 System.*out*.println("the menu is (some what) vegetarian friendly!!");  
}

**5.3.2 프레디케이트가 모든 요소와 일치하는지 검사**

allMatch

boolean isHealthy = menu.stream()  
 .allMatch(d -> d.getCalories() < 100);

noneMatch (allMatch 반대)

boolean isHealthy2 = menu.stream()  
 .noneMatch(d -> d.getCalories() >= 1000);

쇼트서킷 : &&, ||

**5.3.3 요소 검색**

**Optional 이란?**

Optional은 값의 존재나 부재 여부를 표현하는 컨테이너 클래스.

🡺컨테이너란 ? 같은 타입의 여러 객체를 저장하는 일종의 집합

🡺null 관련 버그를 피하기 위해

|  |  |
| --- | --- |
| isPresent | Optional이 값이 있으면 true, 없으면 false |
| ifPresent | @FunctionalInterface public interface Consumer<T> { void accept(T t);  }  값이 있으면 주어진 코드 블록 실행 |
| T get() | 존재하면 값을 반환. 없으면 NoSuchElementException발생 |
| T orElse | 값이 있으면 값을 반환, 없으면 기본값을 반환 |

**5.3.4 첫 번째 요소 찾기**

리스트나 정렬된 연속 데이터로 생성된 스트림은 순서가 정해져 있다.

**findFirst vs findAny**

findfirst : 첫번째 요소 (순서 상관 있음)

findAny: 첫번째 요소 (병렬성에서 처럼, 순서를 찾기 어려울 경우 순서상관 없이 첫번째 요소)

**5.4 리듀싱**

**모든 스트림 요소** 를 처리해서 값으로 도출 (폴드 라고도 부른다)

**5.4.1 요소의 합**

List<Integer> numbers = Arrays.*asList*(1, 2, 3, 4, 5);  
**// Java7 이전 방식**

int sum = 0; // 초기값  
for (int x : numbers) {  
 sum += x; // 모든 요소 연산(+)  
}

**// reduce 이용 방식 (람다)**  
int sum2 = numbers.stream()  
 .reduce(0, (a, b) -> a + b); // 초기값, 모든 연산 (+)

int multiply = numbers.stream()  
 .reduce(0, (a, b) -> a \* b);

**// reduce 이용 방식 (메서드 레퍼런스)**

int sum3 = numbers.stream()  
 .reduce(0, Integer::*sum*); // 초기값, 모든 연산 (+)

🡺이점 reduce 방식은 내부 반복으로 병렬이 쉽다.

reduce 파라미터

T reduce(T identity, BinaryOperator<T> accumulator);

BinaryOperator는 BiFunction을 상속받은 FunctionalInterface

@FunctionalInterface  
public interface BiFunction<T, U, R> {  
R apply(T t, U u);

}

@FunctionalInterface  
public interface BinaryOperator<T> extends BiFunction<T,T,T> {  
public static <T> BinaryOperator<T> minBy(Comparator<? super T> comparator) {  
 Objects.*requireNonNull*(comparator);  
 return (a, b) -> comparator.compare(a, b) <= 0 ? a : b;  
 }  
public static <T> BinaryOperator<T> maxBy(Comparator<? super T> comparator) {  
 Objects.*requireNonNull*(comparator);  
 return (a, b) -> comparator.compare(a, b) >= 0 ? a : b;  
 }  
}

**자세한 계산 방식은 P. 162 그림 참조**

**응용 문제 number의 값은?**

List<Integer> numbers = Arrays.*asList*(1, 2, 3, 4, 5);  
int number = numbers.stream()  
 .reduce(0, Integer::*min*);

**초깃값 없음**

Optional<Integer> sum = numbers.stream()  
 .reduce((a, b) -> (a + b));

초깃값이 없으면 Optional 사용, 있으면 해당 Type 선언 (당연한건가…)

**5.4.2 최댓값과 최솟값**

**퀴즈 5-3**

// 맵 리듀스 패턴  
int count = menu.stream()  
 .map(d -> 1)

.distinct() // 전체를 1로 매핑했기 때문에 distinct 하면 1됨  
 .reduce(0, (a, b) -> a + b);  
  
  
long count = menu.stream().distinct().count(); // equals, hashcode 를 구현했기 때문에 중복 빼고 계산

**번외>** JAVA8의 Interface

public interface Calculator {  
 public int plus(int pre, int post);  
  
 public int multi(int pre, int post);  
  
 default int execPlus(int pre, int post) {  
 return pre + post;  
 }  
  
 **static int execMulti(int pre, int post) {  
 return pre \* post;  
 }**  
}

사용할 때

**Calculator.execMulti(10, 10);**

**스트림 연산 : 상태 없음과 있음**

**내부 상태 없다.**

map, filter 등은 상태 필요 없다.

🡺단순한 결과 출력 (누적이나 이력이 필요 없다)

**내부 상태 있다. (bound)**

reduce, sum, max 같은 연산은 결과를 누적할 내부 상태가 필요하다.

🡺int나 double과 같은 내부 상태 사용 (크기 한정)

**내부 상태 있다 (unbound)**

sorted, distinct는 과거의 이력을 알고 있어야 한다.

🡺 연산을 수행하는 데 필요한 저장소 크기는 정해져 있지 않다.

따라서 데이터 스트림의 크기가 크거나 무한이라면 문제가 생길 수 있다.

**응용 문제result의 값은?**

참고> int형 범위 **(-2147483648 ~ 2147483647)**

List<Integer> list = Arrays.*asList*(2147483647, 1);  
long result = list.stream().reduce(0, Integer::*sum*);

답 : -2147483648 (오버플로우 발생)

**🡺문제 자주 생길 것 같음 (특히 int 형)**

**5.5 실전 연습**

**내가 푼답**

// 2011년에 일어난 모든 트랜잭션을 찾아 값을 오름차순으로 정리하시오.  
List<Integer> values = transactions.stream()  
 .filter(year -> year.getYear() == 2011)  
 .map(Transcatino::getValue)  
 .sorted()  
 .collect(Collectors.*toList*());  
  
// 거래자가 근무하는 모든 도시를 중복 없이 나열하시오.  
List<String> country = transactions.stream()  
 .map(city -> city.getTrader().getCity())  
 .distinct()  
 .collect(Collectors.*toList*());  
  
// 케임브리지에서 근무하는 모든 거래자를 찾아서 이름순으로 정렬하시오.  
List<String> names = transactions.stream()  
 .filter(city -> city.getTrader().getCity() == "Combridge")  
 .map(name -> name.getTrader().getName())  
 .distinct()  
 .sorted()  
 .collect(Collectors.*toList*());  
  
// 모든 거래자의 이름을 알파벳순으로 정렬해서 반환하시오.  
List<String> alphabets = transactions.stream()  
 .map(name -> name.getTrader().getName().split(""))  
 .flatMap(Arrays::*stream*)  
 .distinct()  
 .sorted()  
 .collect(Collectors.*toList*());  
  
// 밀라노(Milano) 에 거래자가 있는가?  
boolean isMilano = transactions.stream()  
 .map(city -> city.getTrader().getCity())  
 .anyMatch(city -> city.equals("Milan"));  
  
// 케임브리지에 거주하는 거래자의 모든 트랜잭션값을 출력하시오.  
List<Integer> values2 = transactions.stream()  
 .filter(city -> city.getTrader().getCity() == "Combridge")  
 .map(Transcatino::getValue)  
 .collect(Collectors.*toList*());  
  
// 전체 트랜잭션 중 최댓값은 얼마인가?  
int max = transactions.stream()  
 .map(Transcatino::getValue)  
 .reduce(0, Integer::*max*);  
  
// 전체 트랜잭션 중 최소값은 얼마인가?  
Optional min = transactions.stream()  
 .map(Transcatino::getValue)  
 .reduce(Integer::*min*);

**5.6 숫자형 스트림**

// 숫자형 스트림  
int calories = menu.stream()  
 .map(Dish::getCalories)  
 .reduce(0, Integer::*sum*);

스트림 요소 형식은 Stream<Integer>이지만 sum 메서드가 없는 이유는 Stream<Dish> 형식의 요소는 sum을 수행할 수 없기 때문에.

🡺스트림 API 숫자 스트림을 효율적으로 처리하도록 **기본 특화 스트림** 을 제공

**5.6.1 기본형 특화 스트림**

IntStream, DoubleStream, LongStream 제공

**숫자 스트림으로 매핑**

int calories = menu.stream()  
 .mapToInt(Dish::getCalories) // IntStream 반환  
 .sum();

IntStream != Strema<Integer>

**객체 스트림으로 복원하기**

IntStream intStream = menu.stream().mapToInt(Dish::getCalories);  
Stream<Integer> stream = intStream.boxed();

**기본값:OptionalInt**

OptionalInt, OptionalDouble, OptionalLong

OptionalInt maxCalories = menu.stream()  
 .mapToInt(Dish::getCalories)  
 .max();  
  
int max = maxCalories.orElse(1);

min,max는 초기값에 따라서 잘못된 결과가 도출될 수 있음. orElse 로 명시 설정 가능

Ex) reduce(0, Integer::min);

**5.6.2 숫자 범위**

IntStream evenNumbers = IntStream.range(1, 10).filter(i -> i % 2 == 0);  
System.out.println(evenNumbers.count()); **// 결과 4**  
IntStream evenNumbers2 = IntStream.rangeClosed(1, 10).filter(i -> i % 2 == 0);  
System.out.println(evenNumbers2.count()); **// 결과 5**

Range와 rangeClosed 차이점은 마지막 숫자 포함 유/무 (range는 미포함, rangeClosed 포함)

**5.6.3 숫자 스트림 활용타고라스 수**

패스

**5.7 스트림 만들기**

다양한 방식으로 스트림 만들기

**5.7.1 값으로 스트림 만들기**

Stream<String> stream = **Stream.of**("Java 8 ", "Lambdas ", "In ", "Action");  
stream.map(String::toUpperCase).forEach(System.out::println);

**5.7.2 배열로 스트림 만들기**

int[] numbers = {2, 3, 5, 7, 11, 13};  
int sum = **Arrays.stream**(numbers).sum();

**5.7.3 파일로 스트림 만들기**

long uniqueWords = 0;  
try (Stream<String> lines = **Files.lines**(Paths.get("data.txt"), Charset.defaultCharset())) {  
 uniqueWords = lines.flatMap(line -> Arrays.stream(line.split(" ")))  
 .distinct()  
 .count();  
  
} catch (IOException e) {  
  
}

**5.7.4 함수로 무한 스트림 만들기**

**Stream.iterate, Stream.generate**를 이용하여 **무한 스트림을** 만들 수 있다..

// iterate

**Stream.iterate**(0, n -> n + 2)  
 .limit(10)  
 .forEach(System.out::println);

@FunctionalInterface  
public interface UnaryOperator<T> extends Function<T, T> {

}

@FunctionalInterface  
public interface Function<T, R> {R apply(T t);

}

// generate  
**Stream.generate**(Math::random)  
 .limit(10)  
 .forEach(System.out::println);

@FunctionalInterface  
public interface Supplier<T> {  
 T get();  
}

**내부 상태 있다 (unbound)**

🡺 연산을 수행하는 데 필요한 저장소 크기는 정해져 있지 않다.

따라서 데이터 스트림의 크기가 크거나 무한이라면 문제가 생길 수 있다.

병렬 코드에서 공급자에 상태가 있으면 안전하지 않다.

iterate는 불변이였는데 generate는 가변으로 처리할 수 있기 때문에 병렬처리 하면 올바른 결과를 못 얻을 수 있다.

🡺iterate, generate는 unbound 이기 때문에 크기 제한 안하면 최종 연산 수행 시 아무 결과도 계산되지 않는다. (무한 반복됨 위험!!)

**CHATER 6. 스트림으로 데이터 수집**

-컬렉션(Collection)

-컬렉터(Collector)

-컬렉트(collect)

**6.1 컬렉터란 무엇인가?**

List<Dish> vegetarianDishes2 = menu.stream()  
 .filter(Dish::isVegetarian)  
 .collect(Collectors.toList());

Collector인터페이스는 스트림의 요소를 어떤 식으로 도출할지 지정한다.

Map<Currency, List<Traction>> transactionsByCurrencies =  
 transactions.stream().collect(groupingBy(Transaction::getCurrency));

GroupingBy를 이용하여 Map형태로도 만들 수 있다.

**6.1.1 고급 리듀싱 기능을 수행하는 컬렉터**

-훌륭하게 설계된 함수형 API는 높은 수준의 조합성과 재사용성을 꼽을 수 있다.

-collect는 내부적으로 리듀싱 연산이 일어 난다.

🡺최종 결과를 저장하는 자료구조에 값을 누적한다.

🡺Collector 인터페서드를 어떻게 구현하느냐에 따라 어떤 리듀싱 연산을 수행할지 결정된다.

ex) Collecotrs.toList()

**6.1.2 미리 정의된 컬렉터**

Collectors 에서 제공하는 기능 크게 3가지

1. 스트림 요소를 하나의 값으로 리듀스하고 요약 (count, sum, min, max등)
2. 요소 그룹화
3. 요소 분할

**6.2 리듀싱과 요약**

long howManyDishes = menu.stream().collect(Collectors.counting()); // Collectors  
  
long howManyDishes2 = menu.stream().count(); // Stream

컬렉터로 스트림의 모든 항목을 하나의 결과로 합칠 수 있다. (내부적으로 reducing 사용)

**6.2.1 스트림값에서 최댓값과 최솟값 검색**

Comparator<Dish> dishCaloriesComparator =  
 Comparator.comparingInt(Dish::getCalories);  
  
// 최댓값  
Optional<Dish> mostCalorieDish =  
 menu.stream()  
 .collect(Collectors.maxBy(dishCaloriesComparator));

Comparator를 구현해서 Collectors.maxBy로 전달

**6.2.2 요약 연산**

스트림에 있는 객체의 숫자 필드의 **합계나 평균** 등도 리듀싱 기능이 자주 사용된다. (요약 연산)

// 합계  
int totalCalories = menu.stream().collect(Collectors.summingInt(Dish::getCalories));  
// 평균  
double avgCalories = menu.stream().collect(Collectors.averagingInt(Dish::getCalories));  
// IntSummaryStatistics (count, sum, min, average, max 다포함)  
IntSummaryStatistics menuStatistics = menu.stream().collect(Collectors.summarizingInt(Dish::getCalories));

**6.2.3 문자열 연결**

// 문자열 연결  
String shortMenu = menu.stream().map(Dish::getName).collect(Collectors.joining(", "));

**6.2.4 범용 리듀싱 요약 연산**

int totalCalories = menu.stream().collect(Collectors  
 .reducing(0, Dish::getCalories, (i, j) -> i + j));

-Static 팩토리 메서드로도 정의할 수 있다.

-특화 컬렉터를(maxBy, Sum 등) 사용한 이유는 프로그래밍적 편의성

**reducing(0, Dish::getCalories, (i, j) -> i + j));**

1. 초기값
2. 변환 함수
3. BinaryOperator (T, T) -> T

**collect와 reduce 차이**

둘다 기능은 비슷하지만 (스트림 요소들을 다른 방식으로 결합)

-collect는 도출하려는 결과를 누적하는 컨테이너를 바꾸도록 설계된 메서드.

-reduce는 두 값을 하나로 도출하는 불변형 연산.

🡺 reduce를 잘못 구현하면 병렬수행할 수가 없다.

🡺 병렬성을 확보하려면 collect 메서드로 리듀싱 연산을 구현하는게 좋다

**컬렉션 프레임워크 유연성 : 같은 연산도 다양한 방식으로 수행할 수 있다.**

int totalCalories = menu.stream().map(Dish::getCalories).reduce(Integer::sum).get();



🡺orElse, orElseGet 등을 사용하는 것이 좋다.

**자신의 상황에 맞는 최적의 해법 선택**

// CASE1

int totalCalories = menu.stream().collect(Collectors  
 .reducing(0, Dish::getCalories, (i, j) -> i + j));

// CASE2

int totalCalories = menu.stream().collect(Collectors

.reducing(0, Dish::getCalories, Integer::sum));

// CASE3

int totalCalories = menu.stream().mapToInt(Dish::getCalories).sum();

CASE 1 : reducing 메서드 사용 (람다)

CASE 2 : reducing 메서드 사용 (메서드 레퍼런스)

CASE 3 : 가독성이 가장 좋고, mapToInt를 이용하여 자동 언박싱 과정을 피함으로 성능까지 좋음

**6.3 그룹화**

Map<Dish.Type, List<Dish>> dishesByType =  
 menu.stream().collect(Collectors.groupingBy(Dish::getType));

🡺 {FISH=[prawns, salmon], MEAT=[pork, beef, chicken], OTHER=[french fries, rice, season fruit, pizza]}

🡺기준으로 스트림이 그룹화되므로 이를 **분류 함수** 라고 한다.

Map<CaloricLevel, List<Dish>> dishesByCaloriceLevel = menu.stream().collect(  
 Collectors.groupingBy(dish -> {  
 if (dish.getCalories() <= 400) {  
 return CaloricLevel.DIET;  
 } else if (dish.getCalories() <= 700) {  
 return CaloricLevel.NORMAL;  
 } else {  
 return CaloricLevel.FAT;  
 }  
 }));

람다 표현식으로도 로직을 구현할 수 있다.

**6.3.1 다수준 그룹화**

Map<Dish.Type, Map<CaloricLevel, List<Dish>>> dishesByTypeCaloricLevel =  
 menu.stream().collect(  
 Collectors.groupingBy(Dish::getType,  
 Collectors.groupingBy(dish -> {  
 if (dish.getCalories() <= 400) {  
 return CaloricLevel.DIET;  
 } else if (dish.getCalories() <= 700) {  
 return CaloricLevel.NORMAL;  
 } else {  
 return CaloricLevel.FAT;  
 }  
 })));

groupingBy 메서드 내에 두번째 기준을 정의하는 내부 groupingBy를 전달해서 두 수준으로 스트림의 항목을 그룹화 할 수 있다. (n 수준 그룹화)

|  |  |  |  |
| --- | --- | --- | --- |
| 칼로리 종류 | FISH | MEAT | OTHER |
| DIET | prawns | chicken | Pizza  Fries |
| NORMAL | salmon | beef | fruit  rice |
| FAT |  | pork |  |

**6.3.2 서브그룹으로 데이터 수집**

Map<Dish.Type, Long> typesCount = menu.stream()  
 .collect(Collectors.groupingBy(Dish::getType, Collectors.counting()));

🡺{FISH=2, MEAT=3, OTHER=5}

요리 종류 중 가장 칼로리 높은 음식

Map<Dish.Type, Optional<Dish>> mostCaloricByType =  
 menu.stream()  
 .collect(Collectors.groupingBy(Dish::getType,

Collectors.maxBy(Comparator.comparingInt(Dish::getCalories))));

🡺{FISH=Optional[prawns], MEAT=Optional[pork], OTHER=Optional[pizza]}

groupingBy(f) 는 groupingBy(f, toList()) 의 축약

public static <T, K> Collector<T, ?, Map<K, List<T>>>  
groupingBy(Function<? super T, ? extends K> classifier) {  
 return *groupingBy*(classifier, *toList*());  
}

**컬렉터 결과를 다른 형식에 적용하기**

Map<Dish.Type, Dish> mostCaloricByType2 =  
 menu.stream()  
 .collect(groupingBy(Dish::getType, // 분류 함수   
 collectingAndThen(  
 maxBy(comparingInt(Dish::getCalories)), // 감싸인 컬렉터

Optional::get))); // 변환 함수

-collectingAndThen은 적용할 컬렉터와 변환 함수를 인수로 받아 다른 컬렉터를 반환한다.

-반환되는 컬렉터는 기존 컬렉터의 래퍼 역할.

-collect의 마지막 과정에서 변환 함수로 자신이 반환하는 값을 매핑.

-리듀싱 컬렉터는 절대 Optional.empty()를 반환하지 않는다.

**groupingBy와 함께 사용하는 다른 컬렉터 예제**

Map<Dish.Type, Integer> totalCaloriesByType =  
 menu.stream().collect(groupingBy(Dish::getType, summingInt(Dish::getCalories) ));

같은 그룹으로 분류된 모든 요소에 리듀싱 작업을 수행할 때는 팩토리 메서드 groupingBy 두 번째 인수로 전달한 컬렉터를 사용한다.

Map<Dish.Type, Set<CaloricLevel>> caloricLevelsByType =  
 menu.stream().collect(  
 groupingBy(Dish::getType,  
 mapping(dish -> {  
 if (dish.getCalories() <= 400) {  
 return CaloricLevel.DIET;  
 } else if (dish.getCalories() <= 700) {  
 return CaloricLevel.NORMAL;  
 } else {  
 return CaloricLevel.FAT;  
 }  
  
 }, toSet()))  
 );

Mapping 메서드도 groupingBy와 자주 사용된다.

toSet을 이용하면 리스트가 아닌 집합으로 스트림의 요소가 누적된다.

*mapping*(dish -> {  
 if (dish.getCalories() <= 400) {  
 return CaloricLevel.*DIET*;  
 } else if (dish.getCalories() <= 700) {  
 return CaloricLevel.*NORMAL*;  
 } else {  
 return CaloricLevel.*FAT*;  
 }  
  
}, *toCollection*(HashSet::new))));

toCollection을 이용하면 원하는 방식를 제거 가능

**6.4 분할**

분할은 분할 함수라 불리는 프레디케이트를 분류 함수로 사용하는 특수한 그룹화기능이다.

List<Dish> vegetarianDishes =  
 menu.stream().collect(partitioningBy(Dish::isVegetarian)).get(true);  
  
List<Dish> vegetarianDishes2 =  
 menu.stream().filter(Dish::isVegetarian).collect(Collectors.toList());

collect에서 partitioningBy(분할 함수)를 사용 하거나 filter(프레디케이트)로 필터하거나 같은 결과를 얻을 수 있다.

**6.4.1 분할의 장점**

분할 함수가 반환하는 참, 거짓 두 가지 요소의 스트림 리스트를 모두 유지한다는 것이 분할의 장점이다.

Map<Boolean, Map<Dish.Type, List<Dish>>> vegetarianDishesByType =  
 menu.stream().collect(  
 partitioningBy(Dish::isVegetarian,  
 groupingBy(Dish::getType)));

분할은 groupingBy 대신 partitioningBy를 사용한다.

**6.4.2 숫자를 소수와 비소수로 분할하기**

**static factory method**

일반 생성자

public Blog();  
public Blog( String powerBloggerID );

스태틱 팩토리 메서드 (생성자와 달리 자기 나름의 이름을 가질 수 있다.)

public static Blog **newBlogInstance**();  
public static Blog **newPowerBlogInstance**( String powerBloggerID );

**6.5 Collector 인터페이스**

public interface Collector<T, A, R> {  
Supplier<A> supplier();BiConsumer<A, T> accumulator();BinaryOperator<A> combiner();  
Function<A, R> finisher();  
Set<Characteristics> characteristics();

}

-T는 수집될 스트림 항목의 제네릭 형식

-A는 누적자, 즉 중간 결과를 누적하는 객체의 형식

-R은 수집 연산 결과 객체의 형식

**6.5.1 Collecotr 인터페이스의 메서드 살펴 보기**

네 개의 메서드는 collect 메서드에서 실행하는 함수를 반환

다섯 번째 메서드 characteristics는 collect 메서드가 어떤 최적화를 이용해서 리듀싱 연산을 수행할 것인가 힌트 특성 집합

**Supplier<A> supplier 메서드 : 새로운 결과 컨테이너 만들기**

public Supplier<List,T>> supplier() {  
 return () -> new ArrayList<T>()  
}  
  
public Supplier<List<T>> supplier() {  
 return ArrayList::new;  
}

supplier는 수집 과정에서 빈 누적자 인스턴스를 만드는 파라미터가 없는 함수.

ex)만약 ToList가 비어 있을 경우 비어있는 스트림의 수집 과정의 결과가 될 수 있다.

**BiConsumer<A,T> accumulator 메서드 : 결과 컨테이너에 요소 추가 하기**

public BiConsumer<List<T>, T> accumulator() {  
 return (list, item) -> list.add(item);  
}  
  
public BiConsumer<List<T>, T> accumulator() {  
 return List::add;  
}

스트림에서 n번째 요소를 탐색할 때 두 인수, 즉 누적자(n-1개 수지상ㅌ)와 n 번째 요소를 함수에 적용

**Function<A, R> finisher 메서드 : 최종 변환값을 결과 컨테이너로 적용하기**

public Function<List<T>, List<T>> finisher() {  
 return Function.*identity*();  
}

스트림 탐색을 끝내고 누적자 객체를 최종 결과로 변환하면서 누적 과정을 끝낼 때 호출할 함수를 반환

**BinaryOperator<A> combiner 메서드 : 두 결과 컨테이너 병합 (병렬과 연관이 깊다)**

public BinaryOperator<List<T>> combiner() {  
 return (list1, list2) -> {  
 list1.add(list2);  
 return list1;  
 }  
}

스트림의 서로 다른 서브파트를 병렬로 처리할 때 누적자가 이 결과를 어떻게 처리할지 정의한다.

**Set<Characteristics> characteristics 메서드**

컬렉터의 연산을 정의하는 Characteristics 형식의 불변 집합을 반환한다.

Characteristics는 스트림을 병렬로 리듀스할 것인지 병렬로 리듀스 한다면 어떤 최적화를 선택해야 할지 힌트를 제공한다.

Characteristics는 다음 세 항목을 포함하는 열거형이다.

**1)UNORDERED**

리듀싱 결과는 스트림 요소의 방문 순서나 누적 순서에 영향을 받지 않는다.

**2)CONCURRENT**

다중 스레드에서 accumulator 함수를 동시에 호출할 수 있으며 이 컬렉터는 스트림의 병렬 리듀싱을 수행할 수 있다.

UNORDERED를 함께 설정하지 않았다면 정렬되지 않은(set) 상황에서만 병렬 리듀싱 수행가능

**3)IDENTITY\_FINISH**

finisher 메서드가 반환하는 함수는 단순히 identity를 적용할 뿐이므로 이를 생략할 수 있다.

따라서 리듀싱 과정의 최종 결과로 누적자 객체를 바로 사용할 수 있다.

**6.5.2 응용하기**

public class ToListCollector<T> implements Collector<T, List<T>, List<T>> {  
  
 @Override  
 public Supplier<List<T>> supplier() {  
 return ArrayList::new;  
 }  
  
 @Override  
 public BiConsumer<List<T>, T> accumulator() {  
 return List::add;  
 }  
  
 @Override  
 public Function<List<T>, List<T>> finisher() {  
 return Function.*identity*();  
 }  
  
 @Override  
 public BinaryOperator<List<T>> combiner() {  
 return (list1, list2) -> {  
 list1.addAll(list2);  
 return list1;  
 };  
 }  
  
  
 @Override  
 public Set<Characteristics> characteristics() {  
 return Collections.*unmodifiableSet*(EnumSet.*of*(  
 *IDENTITY\_FINISH*, *CONCURRENT* ));  
 }  
}

List<Dish> dishes = menu.stream().collect(new ToListCollector<Dish>());

toList는 스태틱 팩토리 메서드이지만 ToListCollector는 new로 인스턴스화 한다는 점이 다르다.

**6.6 커스텀 컬렉터를 구현해서 성능 개선하기**

생략

**6.7 요약**

-collect는 스트림의 요소를 요약 결과로 누적하는 다양한 방법(컬렉터)을 인수로 갖는 최종 연산

-스트림 요소를 하나의 값으로 리듀스하고 요약하는 컬렉터뿐 아니라 최솟값, 최댓값, 평균값을 계산하는 컬렉터 등이 미리 정의되어 있다.

-미리 정의된 컬렉터인 groupingBy로 스트림의 요소를 **그룹화** 할 수 있다.

-partitioningBy로 스트림의 요소를 **분할**할 수 있다.

-컬렉터는 다수준의 그룹화, 분할, 리듀싱 연산에 적합하게 설계되어 있다.

-Collector 인터페이스에 정의된 메서드를 구현해서 커스텀 컬렉터 개발할 수 있음.

**CHATER7. 병렬 데이터 처리와 성능**

자바7은 더 쉽게 병렬화를 수행하면서 에러를 최소화할 수 있도록 **포크/조인 프레임워크** 라는 기능을 제공한다.

**7.1 병렬 스트림**

parallelStream을 호출하면 **병렬 스트림** 이 생성된다.

병렬 스트림이란 각각의 스레드에서 처리할 수 있도록 스트림 요소를 여러 청크로 분할한 스트림이다.

**전통적 자바방식**

public static long iterativeSum(long n) {  
 long result = 0;  
 for (long i = 1L; i <= n; i++) {  
 result += i;  
 }  
 return result;  
}

**스트림 연산 방식**

public static long sequentialSum(long n) {  
 return Stream.*iterate*(1L, i -> i + 1)  
 .limit(n)  
 .reduce(0L, Long::*sum*);  
}

n (limit)가 커진다면 병렬 처리하는 것이 좋다.

동기화는? 몇 개의 스레드를? 결과는 어떻게 더할까?

🡺병렬 스트림을 이용하면 걱정 끝

**7.1.1 순차 스트림을 병렬 스트림으로 변환하기**

public static long parallerSum(long n) {  
 return Stream.*iterate*(1L, i -> i + 1)  
 .limit(n)  
 .parallel() // 병렬 스트림으로 변환  
 .reduce(0L, Long::*sum*);  
}



<병렬 리듀싱 연산>

여러 청크로 분할되어 병렬로 수행 후, 마지막으로 리듀싱 연산으로 생성된 부분결과를 다시 리듀싱 연산으로 합쳐서 전체 스트림의 리듀싱 결과를 도출한다.

.filter()  
.sequential()  
.map()  
**.parallel()**  
.reduce();

Parallel과 sequential 두 메서드 중 최종적으로 호출된 메서드가 전체 파이프라인에 영향을 미친다.

**병렬 스트림에서 사용하는 스레드 풀 설정**

기본적으로 ForkJoinPoll은 프로세서 수,

즉 Runtime.getRuntime().availableProcessors()가 반환하는 값에 상응하는 스레드를 갖는다.

// CPU 코어의 수만큼 최대 스레드를 사용하는 스레드풀을 생성

Runtime**.**getRuntime**().**availableProcessors**()**

🡺기본값 권장

**7.1.2 스트림 성능 측정**

// 반복 Iterative sum done in: 7 msecs (생각보다 빠름)  
public static long iterativeSum(long n) {  
 long result = 0;  
 for (long i = 1L; i <= n; i++) {  
 result += i;  
 }  
 return result;  
}  
  
// 순차 Sequential sum done in: 118 msecs  
public static long sequentialSum(long n) {  
 return Stream.*iterate*(1L, i -> i + 1)  
 .limit(n)  
 .reduce(0L, Long::*sum*);  
}  
  
// 병렬 Parallel sum done in: 399 msecs (실망...)  
public static long parallerSum(long n) {  
 return Stream.*iterate*(1L, i -> i + 1)  
 .limit(n)  
 .parallel() // 병렬 스트림으로 변환  
 .reduce(0L, Long::*sum*);  
}

for 루프는 저수준으로 동작하며 기본값을 박싱하거나 언박싱할 필요가 없으므로 수행 속도가 빠르다.

**왜 병렬이 느릴까?**

-iterate가 박싱된 객체를 생성하므로 이를 다시 언박싱하는 과정이 필요했다.

-iterate는 병렬로 실행될 수 있도록 독립적인 청크로 분할하기가 어렵다.

Iterate는 본질적으로 순차적이기 때문에 청크로 분할하기가 어렵다.

🡺리듀싱 연산이 수행되지 않는다. (전체 숫자를 모른다.)

이처럼 병렬 프로그래밍은 까다롭다.

🡺parallel 메서드 호출 시, 내부적으로 어떤 일이 일어나는지 꼭 이해해야 한다. (테스트도 해보고)

**더 특화된 메서드 사용**

LongStream.rangeClosed

-LongStream.rangeClosed 는 기본형 long을 사용하기 때문에 박싱 언박싱 오버헤드 없음.

-LongStream.rangeClosed는 쉽게 청크로 분할할 수 있는 숫자 범위를 생산한다.

ex) 1-5, 6-10, 11-15, 16-20 범위로 숫자 분할

// 순차 Sequential Ranged sum done in 8 msecs  
public static long rangedSum(long n) {  
 return LongStream.*rangeClosed*(1, n)  
 .reduce(0L, Long::*sum*);  
}  
  
// 병렬 Parallel Ranged sum done in 3 msecs (드디어 for보다 더 빠름)  
public static long parallelRangedSum(long n) {  
 return LongStream.*rangeClosed*(1, n)  
 .parallel()  
 .reduce(0L, Long::*sum*);  
}

올바른 자료구조를 선택해야 병렬실행도 최적의 성능을 발휘할 수 있다는 사실을 확인할 수 있다.

하지만 병렬화는 완전 공짜는 아니다.

-각 서브스트림을 서로 다른 스레드의 리듀싱 연산으로 할당하고

-이들 결과를 하나의 값으로 합쳐야 한다

🡺멀티코어 간의 데이터 이동은 생각보다 비싸다.

🡺상황에 따라서 병렬화를 이용할 수 있거나 없을 수도 있다.

🡺스트림을 병렬화해서 코드 실행 속도를 빠르게 하고 싶으면 항상 병렬화를 올바르게 사용

**7.1.3 병렬 스트림의 올바른 사용법**

// 순차 Sequential SideEffect sum 8 msecs / Result: 50000005000000 (10번다 동일)  
public static long sideEffectSum(long n) {  
 Accumulator accumulator = new Accumulator();  
 LongStream.*rangeClosed*(1, n).forEach(accumulator::add);  
 return accumulator.total;  
}  
  
// 병렬 Parallel SideEffect sum 3 msecs / Result: 24895627783262 (10번다 다름)  
public static long sideEffectParallelSum(long n) {  
 Accumulator accumulator = new Accumulator();  
 LongStream.*rangeClosed*(1, n).parallel().forEach(accumulator::add);  
 return accumulator.total;  
}

class Accumulator {  
 public long total = 0;  
  
 public void add(long value) {  
 total += value;  
 }  
}

total += value를 실행하면 이런 문제가 생긴다.

total += value는 아토믹 연산이 아니다.

공유된 가변 상태를 피해야 한다.

**7.1.4 병렬 스트림 효과적으로 사용하기**

고정된 양을 기준으로 병렬 스트림을 사용하는 것은 적절하지 않음

🡺상황이 달라지면 (기기 등) 기준이 제 역할을 못함

-확신이 서지 않는다면 직접 측정하라. (적절한 베치마크로 직접 성능 측정)

-박싱을 주의 하라.

🡺자동 박싱과 언박싱은 성능을 크게 저하시키는 요소

🡺박싱을 피하도록 기본 특화 스트림(IntStream, LongStream, DoubleStream)을 제공

-순차 스트림보다 병렬 스트림에서 성능이 떨어지는 연산이 있다.

🡺요소의 순서에 의존하는(limit, findFirst)연산을 병렬 스트림에서 수행하려면 비싼 비용을 치러야 한다. (순서를 유지 해야 되기 때문)

🡺findAny는 요소의 순서와 상관없이 연산하므로 findFirst 보다 성능이 좋다.

🡺정렬된 스트림에 unordered를 호출하면 비정렬된 스트림을 얻을 수 있다.

-스트림에서 수행하는 전체 파이프라인 연산 비용을 고려하라.

🡺N(요소수)\*Q(비용) = N\*Q(처리비용). Q가 높아진다는 것은 병렬 스트림으로 개선할 수 있다는 가능성을 의미

-소량의 데이터에서는 병렬 스트림이 도움 되지 않는다.

🡺병렬화 과정에서 생기는 부가 비용을 상쇄할 수 있을 만큼 이득이 없다.

-스트림을 구성하는 자료구조가 적절한지 확인하라

🡺예를들어 ArrayList를 LinkedList 보다 효율적으로 분할할 수 있다.

🡺range 팩토리 메서드로 만든 기본형 스트림도 쉽게 분해할 수 있다.

🡺결국 IntStream, LongStream, DoubleStream (range를 쓰려면) 사용하는게 좋다.

-스트림의 특성과 파이프라인의 중간 연산이 스트림의 ㄹ 어떻게 바꾸는지에 따라 분해 과정의 성능이 달라질 수 있다.

🡺필터 연산이 있으면 스트림의 길이를 예측할 수 없어서 효과적으로 병렬처리 못함

-최종 연산의 병합 과정의 비용을 살펴 보라

🡺병합 과정의 비용이 비싸다면 병렬 스트림으로 얻은 성능의 이익이 합치는 과정에서 상쇄될 수 있다.

|  |  |
| --- | --- |
| 소스 | 분해성 |
| ArrayList | 훌륭함 |
| LinkedList | 나쁨 |
| IntStream.range | 훌륭함 |
| Stream.iterate | 나쁨 (무한스트림) |
| HashSet | 좋음 |
| TreeSet | 좋음 |

**7.2. 포크/조인 프레임워크**

포크/조인 프레임워크는 병렬화할 수 있는 작업을 재귀적으로 작은 작업으로 분활한 다음에 서브태스크 각각의 결과를 합쳐서 전체 결과를 만든다 (분할 / 정복)

7.2.1 RecursiveTask 활용

스레드 풀을 이용하려면 RecursiveTask<R>의 서브클래스를 만들어야 한다.

🡺R은 결과 형식 또는 결과가 없을 때는 RecursiveAction 형식

RecursiveTask를 정의하려면 추상 메서드 compute를 구현

protected abstract R compute();

compute구현은 다음과 같은 형식을 유지

if(태스크가 충분히 작거나 더 이상 분할할 수 없으면) {  
 순차적으로 태스크 계산  
}else {  
 태스크를 두 서브태스크로 분할  
 태스크가 다시 서브태스크로 분할되도록 이 메서드를 재귀적으로 호출함  
 모든 서브태스크의 연산이 완료될 때까지 기다림  
 각 서브태스크의 결과를 합침  
}

package chater7;  
  
*/\*\*  
 \* Created by simjunbo on 2018-02-26.  
 \*/*public class ForkJoinSumCalculator extends java.util.concurrent.RecursiveTask<Long> {  
  
 private final long[] numbers;  
 private final int start;  
 private final int end;  
 public static final long *THRESHOLD* = 10\_000;  
  
 public ForkJoinSumCalculator(long[] numbers) {  
 this(numbers, 0, numbers.length);  
 }  
  
 private ForkJoinSumCalculator(long[] numbers, int start, int end) {  
 this.numbers = numbers;  
 this.start = start;  
 this.end = end;  
 }  
  
 @Override  
 protected Long compute() {  
 int length = end - start;  
 if (length <= *THRESHOLD*) {  
 return computeSequentially();  
 }  
 ForkJoinSumCalculator leftTask =  
 new ForkJoinSumCalculator(numbers, start, start + length / 2);  
  
 leftTask.fork(); //비동기로  
  
 ForkJoinSumCalculator rightTask =  
 new ForkJoinSumCalculator(numbers, start + length / 2, end);  
  
 Long rightResult = rightTask.compute(); // 오른쪽 동기로  
 Long leftResult = leftTask.join(); // 왼쪽 취합  
  
 return leftResult + rightResult;  
 }  
  
 private long computeSequentially() {  
 long sum = 0;  
 for (int i = start; i < end; i++) {  
 sum += numbers[i];  
 }  
 return sum;  
 }  
}

다음 코드 처럼 ForkJoinSumCalculator의 생성자로 원하는 수의 배열로 넘겨줄 수 있다.

public static long forkJoinSum(long n) {  
 long[] numbers = LongStream.*rangeClosed*(1, n).toArray();  
 ForkJoinTask<Long> task = new ForkJoinSumCalculator(numbers);  
 return new ForkJoinPool().invoke(task);  
}

**7.2.2 포크/조인 프레임워크 제대로 사용 하는 방법**

-join 메서드를 태스크에 호출하면 태스크가 생산하는 결과가 준비될 때까지 호출자를 블록시킨다.

🡺따라서 두 서브태스크가 모두 시작된 다음에 join을 호출해야 한다. 그러면 다른 태스크가 끝나길 기다리는 일이 발생하며 원래 순차 보다 느리고 복잡해 진다.

-RecursiveTask 내에 ForkJoinPool의 invoke 메서드를 사용하지 말아야 한다.

대신 compute나 fork 메서드를 직접 호출할 수 있다. 순차 코드에서 병렬 계산을 시작할 때만

Invoke를 사용한다.

-서브태스크에 fork 메서드를 호출해서 ForkJoinPool의 일정을 조절할 수 있다.

🡺왼쪽, 오른쪽 모두 fork를 호출하는것보다 한쪽 작업에는 compute를 것이 효율적이다.

🡺그러면 두 태스크에서 한 태스크에는 같은 스레드를 재사용할 수 있으므로 불필요한 태스크를 할당하는 오버헤드를 피할 수 있다.

-포크/조인 프레임워크를 이용하여 병렬 계산은 디버깅하기 어렵다.

-병렬 스트림에서 살펴본 것처럼 멀티코어에 포크/조인 프레임워크를 사용하는 것이

순차처리보다 무조건 빠를거라는 생각은 버려야 한다.

**7.2.3 작업 훔치기**

실제로는 코어 개수와 관계없이 적절한 크기로 분할된 많은 태스크를 포킹하는 것이 바람직하다.

훨씬 복잡한 시나리오가 사용되는 현실에는 각각의 서브태스크의 작업 완료 시간이 달라 질 수 있다.

🡺포크/조인 프레임워크에서는 **작업 훔치기** 라는 기법으로 이 문제를 해결한다.



**7.3 Spliterator**

Spliterator는 분할할 수 있는 반복자라는 의미이다.

Iterator(반복자) 랑 비슷하지만 Spliterator는 병렬작업에 특화되어 있다.

Spliterator가 어떻게 동작하는지 이해하면 병렬 스트림 동작과 관련한 통찰력을 얻을 수 있다.

public interface Spliterator<T> {  
 boolean tryAdvance(Consumer<? super T> action);  
 Spliterator<T> trySplit();  
 long estimateSize();  
 int characteristics();  
}

-tryAdvance 메서드는 Spliterator의 요소를 하나씩 순차적으로 수비하면서 탐색해야 할 요소가 남아 있으면 반환한다. (Iterator와 같다)

-trySplit 메서드는 Spliterator의 일부 요소(자신이 반환 요소)를 분할해서 두 번째 Spliterator를 생성하는 메서드다.

-estimateSize 메서드로 탐색해야 할 요소 수 정보를 제공할 수 있다.

-characteristics 메서드는 Spliterator 자체의 특성 집합을 포함하는 int를 반환한다.

**7.3.1 분할 과정**

final String SENTENCE =  
 " Nel mezzo del cammin di nostra vita " +  
 "mi ritrovai in una selva oscura" +  
 " ch la dritta via era smarrita ";  
System.*out*.println("Found " + countWordsIteratively(SENTENCE) + " words");

public static int countWordsIteratively(String s) {  
 int counter = 0;  
 boolean lastSpace = true;  
 for (char c : s.toCharArray()) {  
 if (Character.*isWhitespace*(c)) {  
 lastSpace = true;  
 } else {  
 if (lastSpace) {  
 counter++;  
 }  
 lastSpace = false;  
 }  
 }  
 return counter;  
}

🡺Found 19 words

**함수형 단어 개수 계산 메서드 재구현하기**

**7.4 요약**

-내부 반복을 이용하면 명시적으로 다른 스레드를 사용하지 않고도 스트림을 병렬로 처리 가능

-병렬처리가 항상 빠른건 아니다. 성능을 직접 측정해봐야 한다.

-데이터가 아주 많거나 각 요소를 처리하는데 오랜 시간이 걸릴 때 성능을 높일 수 있다.

-가능하면 기본형 특화 스트림을 사용하는 등 올바른 자료구조 선택하느냐에 따라 성능적으로

영향을 미칠 수 있다.

-포크/조인 프레임워크에서는 병렬화할 수 있는 태스크를 작은 태스크로 분할한 다음에 분할된

태스크를 각각의 스레드로 실행하며 서브태스크 각각의 결과를 합쳐서 최종 결과를 생성한다.

-Spliterator는 탐색하려는 데이터를 포함하는 스트림을 어떻게 병렬화할 것인지 정의한다.

보충

Java 병렬에서 하드웨어 적 사용 (CAS)

http://mygumi.tistory.com/112