

IPI Compare Tool

HBI-5a-T: Eindopdracht Java I voor Bio-Informatica

Casus

Een afdeling bij het UMCG heeft een groot aantal bestanden met daarin opgeslagen proteïne informatie. Deze bestanden verschillen echter soms op een enkel punt. Er zijn verschillende tools in omloop om tekstbestanden met elkaar te vergelijken. Wat ze echter zoeken is een tool om specifiek deze IPI bestanden met elkaar te vergelijken. De nieuw te ontwikkelen tool dient namelijk ook een indicatie af te geven van de ernst van de afwijking.

Voorbeeld bestand: een tab delimited file met 4 kolommen

IPI:IPI00000816.1	IPI00000816	ISOFORM 1 OF 14-3-3 PROTEIN EPSILON.	255
IPI:IPI00003421.3	IPI00003421	T-BOX BRAIN PROTEIN 1.	682
IPI:IPI00004358.4	IPI00004358	GLYCOGEN PHOSPHORYLASE, BRAIN FORM.	843
IPI:IPI00005467.3	IPI00005467	BRAIN-SPECIFIC SERINE PROTEASE 4.	317
IPI:IPI00005614.6	IPI00005614	ISOFORM LONG OF SPECTRIN BETA CHAIN, BRAIN 1.	2364
IPI:IPI00007812.1	IPI00007812	V-TYPE PROTON ATPASE SUBUNIT B, BRAIN ISOFORM.	511

Opdracht IPI Compare Tool(100 punten)

Ontwikkel een tool die in staat is om bestanden met elkaar te vergelijken. Als de bestanden exact overeen komen geeft het programma de melding dat het identieke bestanden betreft. Indien er afwijkingen zijn aangetroffen geeft de tool aan in welke regel en op welke positie de afwijking is aangetroffen.

Functionaliteit

- 1) De IPI Compare Tool controleert de opbouw van het bestand. Het bestand dient te bestaan uit 4 kolommen:
 - a. IPI: met een code
 - b. Code zonder toevoeging
 - c. Beschrijving van de functionaliteit
 - d. Grootte van het eiwit
- 2) Indien het bestand niet voldoet aan het onder 1 gestelde format werpt het programma een custom exception op `NotValidIPIFormat`.
- 3) Het programma vangt de custom exception af en toont een pop-up waarin wordt aangegeven dat het gelezen bestand corrupt is.
- 4) Het programma doorloopt regel voor regel beide bestanden gelijktijdig. Indien er een inconsistentie wordt geconstateerd wordt aangegeven op welk regel deze zich bevindt en in welke kolom deze staat. Let op: hiervoor zijn geen arrays en andere complexe datastructuren nodig. Het vereist alleen het gelijktijdig doorlopen van beide bestanden.
- 5) Ontwikkel tevens een visualisatie van het verloop van de analyse. Een voortgangsbalk die aangeeft waar er een fout is geconstateerd. Zie Figure 4. Met kleuren geef je aan welk type verschil is aangetroffen. Bijvoorbeeld geel is een verschil in kolom 4 (grootte). Blauw is een verschil in kolom 2 (code). Dit kun je aangeven op de relatieve positie.

Voorbeeldschermen

De bijgaande voorbeeldschermen geven een indruk van de gewenste functionaliteit. De uiteindelijke opmaak mag afwijken mits de functionaliteit overeenkomt.

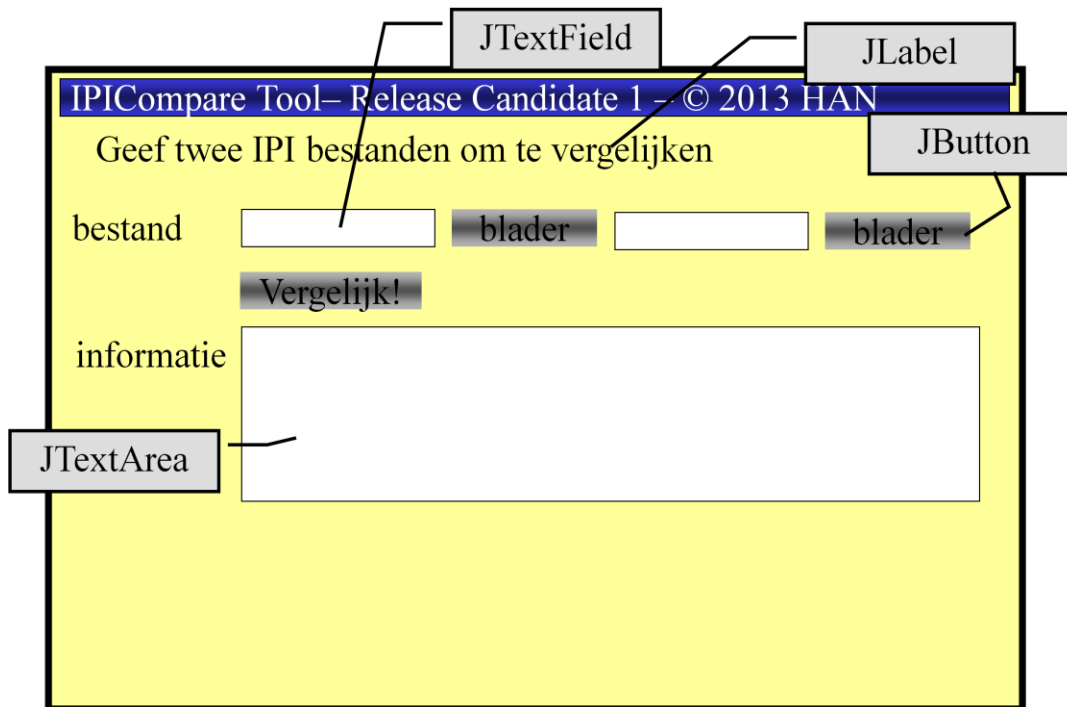


Figure 1: Weergave van grafische elementen (eigen werk auteur)

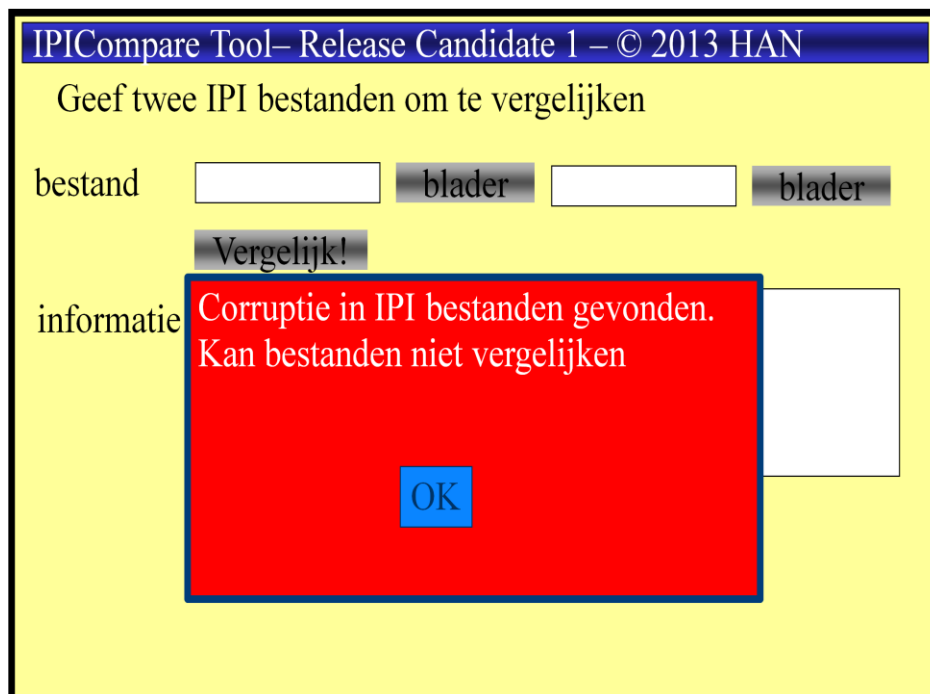


Figure 2: Pop up indien IPI bestanden niet voldoen aan eisen van een IPI bestand. (eigen werk auteur)

IPICompare Tool– Release Candidate 1 – © 2013 HAN

Geef twee IPI bestanden om te vergelijken

bestand

informatie

Bestanden komen niet overeen op regel 71:
Verskil in kolom 4, grootte

Figure 3: Melding indien bestanden verschillen op een regel (eigen werk auteur)

IPICompare Tool– Release Candidate 1 – © 2013 HAN

Geef twee IPI bestanden om te vergelijken

bestand

informatie

Bestanden komen niet overeen op regel 71:
Verskil in kolom 4, grootte
Bestanden komen niet overeen op regel 92:
Verskil in kolom 2, code

Figure 4: Grafische weergave van verloop analyse. De gekleurde lijnen geven een indicatie van de positie en het geconstateerde verschil. (eigen werk auteur)

Beoordelingsformulier Thematoets HBI-cou5a

Naam:

Competentie 2.2 (1, 2, 4, 5, 6, 7, 8, 11)		Beoordeling door	
Onderdeel	Max Score	Student	Docent
1. GUI functionaliteit			
a. Juiste opbouw GUI	10		
b. Juiste weergave van gevraagde GUI elementen	20		
2. Exception Handling			
a. Afvangen standaard Exceptions	10		
b. Custom Exception Handling	20		
3. Applicatie logica			
a. Efficiënte en correcte verwerking van bestand	10		
b. Juiste berekening	10		
c. Code is generiek geschreven	10		
d. Code voldoet aan de Java conventies	10		
4. Documentatie en commentaar			
a. Geen goed gebruik van commentaar	-10		
b. Geen documentatie met Javadoc	-10		
Totaal	Max: 100		

Inleveren

Geschreven code en een eigen beoordeling van je geschreven code aan de hand van bovenstaande beoordelingscriteria. Vul daartoe de kolom student in. N.B. Het eindoordeel van de docent is hetgeen geadministreerd wordt.

Het cijfer is het aantal behaalde punten delen door het maximaal aantal te behalen punten maal 10.

Everything in this document was made by the author ©2013 Martijn van der Bruggen
HAN University of Applied Sciences