Data Exploration

MM

2/8/2017

Adding the time cuts and adding columns for patterns of 2 and 3 combined events

library (dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library (ggplot2)  
encounter <- read.csv("encounter.csv", stringsAsFactors = FALSE)  
  
OP2\_pattern<-matrix(NA,30367,1)  
encounter= data.frame(encounter,OP2\_pattern)  
OP3\_pattern<-matrix(NA,30367,1)  
encounter= data.frame(encounter,OP3\_pattern)  
for (i in 1:30366) {  
   
 if (encounter[i,3]==1 & encounter[i+1,3]==2){  
 encounter[i+1,4]='A'  
 }  
 else if (encounter[i,3]==1 & encounter[i+1,3]==3){  
 encounter[i+1,4]='B'  
 }  
 else if (encounter[i,3]==2 & encounter[i+1,3]==3){  
 encounter[i+1,4]='C'  
 }  
 else if (encounter[i,3]==3 & encounter[i+1,3]==2){  
 encounter[i+1,4]='D'  
 }  
   
}  
  
for (i in 1:30365) {  
   
 if (encounter[i,3]==1 & encounter[i+1,3]==2 & encounter[i+2,3]==3){  
 encounter[i+1,5]='P1'  
 }  
 else if (encounter[i,3]==1 & encounter[i+1,3]==3 & encounter[i+2,3]==2){  
 encounter[i+1,5]='P2'  
 }  
   
}  
encounter <- mutate(encounter,   
 enc\_id = as.factor(enc\_id),  
 time\_to\_event = as.factor(time\_to\_event),  
 event\_type\_name = ifelse(as.double(event\_type) == 1, " accessing a problem", ifelse(as.double(event\_type) == 2, " adding a note to a problem", ifelse(as.double(event\_type) == 3, " adding a task for a problem ","other operation"))),  
 time\_cat = ifelse(as.double(time\_to\_event) < 100, "100s",ifelse(as.double(time\_to\_event) < 200, "200s",ifelse(as.double(time\_to\_event) < 300, "300s", ifelse(as.double(time\_to\_event) < 400, "400s",ifelse(as.double(time\_to\_event) < 500, "500s", ifelse(as.double(time\_to\_event) < 600, "600s", ifelse(as.double(time\_to\_event) < 700, "700s",ifelse(as.double(time\_to\_event) < 800, "800s", ifelse(as.double(time\_to\_event) < 900, "900s",ifelse(as.double(time\_to\_event) < 1000, "1000s",ifelse(as.double(time\_to\_event) < 1100, "1100s",ifelse(as.double(time\_to\_event) < 1200, "1200s",ifelse(as.double(time\_to\_event) < 1300, "1300s",ifelse(as.double(time\_to\_event) < 1400, "1400s","after\_1400s")))))))))))))),  
 time\_cat = factor(time\_cat, levels = c("100s", "200s", "300s","400s", "500s", "600s", "700s","800s", "900s","1000s","1100s","1200s","1300s","1400s","after\_1400s")),  
 OP2\_pattern\_names = ifelse(as.character(OP2\_pattern) == 'A', " Adding note after accessing a problem", ifelse(as.character(OP2\_pattern) == 'B', " adding task after accessing a problem", ifelse(as.character(OP2\_pattern) == 'C' , " adding a task after adding a note ",ifelse(as.character(OP2\_pattern) == 'D', " adding a note after adding a task ","other patterns")))),  
 OP3\_pattern\_names = ifelse(as.character(OP3\_pattern) == 'P1', " Access-> add note -> add task", ifelse(as.character(OP3\_pattern) == 'P2', " Access-> add task -> add note", "other patterns")))

## Warning in ifelse(as.double(c("1", "2", "3", "3", "1", "1", "2", "1",  
## "3", : NAs introduced by coercion

## Warning in ifelse(as.double(c("1", "2", "3", "3", "1", "1", "2", "1",  
## "3", : NAs introduced by coercion

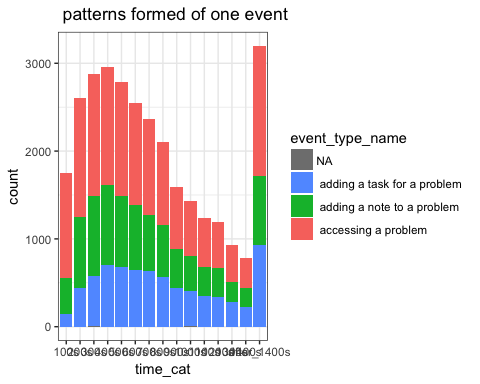
## Warning in ifelse(as.double(c("1", "2", "3", "3", "1", "1", "2", "1",  
## "3", : NAs introduced by coercion

summary(encounter)

## enc\_id time\_to\_event event\_type OP2\_pattern   
## 2088 : 83 322 : 42 Length:30367 Length:30367   
## 112 : 58 410 : 42 Class :character Class :character   
## 131 : 48 378 : 41 Mode :character Mode :character   
## 193 : 46 376 : 40   
## 562 : 44 486 : 40   
## 50 : 43 215 : 39   
## (Other):30045 (Other):30123   
## OP3\_pattern event\_type\_name time\_cat   
## Length:30367 Length:30367 after\_1400s: 3198   
## Class :character Class :character 400s : 2962   
## Mode :character Mode :character 300s : 2883   
## 500s : 2787   
## 200s : 2609   
## 600s : 2545   
## (Other) :13383   
## OP2\_pattern\_names OP3\_pattern\_names   
## Length:30367 Length:30367   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
##

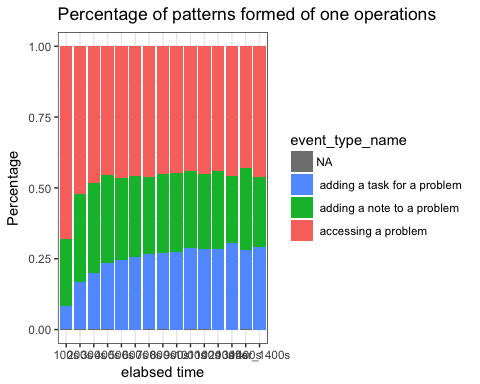
Patterns of one event:

ggplot(encounter, aes(x = time\_cat, fill = event\_type\_name)) + geom\_bar(position = "stack") +guides(fill = guide\_legend(reverse = TRUE)) + theme\_bw()+ ggtitle(" patterns formed of one event")



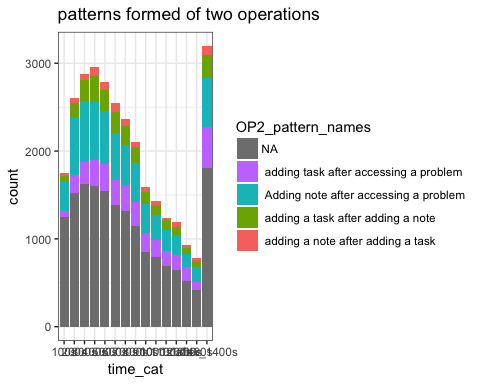
Pourcentages of patterns of one event:

ggplot(encounter, aes(x = time\_cat, fill = event\_type\_name)) +  
 geom\_bar(position = "fill") +  
 guides(fill = guide\_legend(reverse = TRUE)) +  
 theme\_bw() +  
 labs(x="elabsed time",y = "Percentage")+  
 ggtitle("Percentage of patterns formed of one operations")



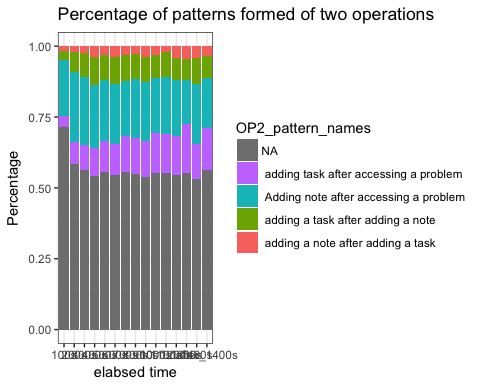
Patterns of two event:

ggplot(encounter, aes(x = time\_cat, fill = OP2\_pattern\_names)) + geom\_bar(position = "stack") +guides(fill = guide\_legend(reverse = TRUE)) + theme\_bw()+ggtitle("patterns formed of two operations")



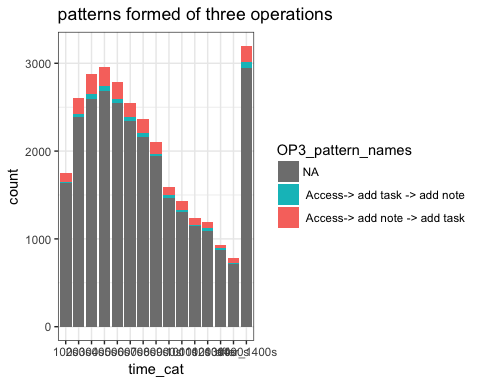
Pourcentages of patterns of two events:

ggplot(encounter, aes(x = time\_cat, fill = OP2\_pattern\_names)) +  
 geom\_bar(position = "fill") +  
 guides(fill = guide\_legend(reverse = TRUE)) +  
 theme\_bw() +  
 labs(x="elabsed time",y = "Percentage")+  
 ggtitle("Percentage of patterns formed of two operations")



Patterns of three event:

ggplot(encounter, aes(x = time\_cat, fill = OP3\_pattern\_names)) + geom\_bar(position = "stack") +guides(fill = guide\_legend(reverse = TRUE)) + theme\_bw()+ggtitle("patterns formed of three operations")



Pourcentages of patterns of three events:

ggplot(encounter, aes(x = time\_cat, fill = OP3\_pattern\_names)) +  
 geom\_bar(position = "fill") +  
 guides(fill = guide\_legend(reverse = TRUE)) +  
 theme\_bw() +  
 labs(x="elabsed time",y = "Percentage")+  
 ggtitle("Percentage of patterns formed of three operations")

