VIRTUAL REALITY BASED END-USER ASSESSMENT TOOL for Remote Product /System Testing & Support

PRESENTATION OUTLINE

- AIM OF RESEARCH
- MOTIVATIONS
- BACKGROUND & RELATED WORK
- ANALYSIS APPROACH
- MODEL ARCHITECTURE
- EXPERIMENT & RESULTS
- CONCLUSION

MOTIVATIONS

- DESIGN
  - WOULD USERS APPRECIATE THE PRODUCT?
  - CAN THEY USE IT EASILY WITH MINIMAL EFFORT?
- MAINTENANCE & TRAINING
  - GIVING THE RIGHT TRAININGS?
  - CAN IT BE ASSEMBLED/DISASSEMBLED FOR MAINTENANCE PURPOSES?

MOTIVATIONS

- DESIGN
  - WOULD USERS APPRECIATE THE PRODUCT?

OBSERVATIONAL DATA COLLECTED FROM THE USER PRODUCT INTERACTION MAY REVEAL ANSWERS FOR THESE QUESTIONS

MAINTENANCE PURPOSES?

COLLECTION OF OBSERVATIONAL DATA

I. USE OF PHYSICAL PROTOTYPES

- EXPENSIVE
- TIME CONSUMING
- GEOGRAPHICALLY RESTRICTED APPLICATION AREA

II. USE OF VIRTUAL PROTOTYPES & AUTOMATED ANALYSIS

- INEXPENSIVE
- FAST
- MORE DETAILED DATA COLLECTION FLEXIBILITY

MOTIVATIONS

A VR BASED END USER ASSESSMENT TOOL

- IDENTIFY PROBLEMATIC ASPECTS
- REVEALS USERS' PERFORMANCE
- SHOWS PATHS FOLLOWED BY USERS & HOW THEY DIFFER FROM DESIGNER EXPECTATIONS
- REPORTS SIMILARITIES & DIFFERENCES AMONG USERS

OBJECTIVE

To design a VR based automated tool to filter, analyze and interpret behavioral data

without compromising from scale and efficiency?

APPROACH

We propose integrating existing virtual reality technology with an automated behavioral data analysis approach:

I. Construct VR based prototype testing simulations to capture human interaction in log-files

- *a prototype, digital mock-ups (DMUs)*
- *an assembly*
- *a maintenance scenario*

II. Develop an intelligent and automated data analysis technique to extract useful qualitative and quantitative information from voluminous raw data

BACKGROUND

Virtual Reality

Virtual Reality in Design & Maintenance

- 3D visualization
- Interactive simulations

Today successful applications of VR include:

- Product prototyping for testing
- Training : in medicine, military and manufacturing
- Maintenance, human factors and ergonomic studies
- Furniture industry, architecture, interior design

Maintenance

Training in Medicine

Training in Manufacturing

Architecture

RELATED WORK

Virtual Reality

Are Virtual Prototypes reliable for testing purposes?

- **The predictive power of virtual-prototypes is almost the same as physical prototypes (Dahan et al., 1998)**
- **VR features offer users the opportunity to explore virtual objects at a high level of detail that are appropriate for activity evaluation (Hurwicz, 2000)**
- **Gamberini et al. (2003) conclude that people show realistic responses to dangerous situations simulated in VR**
- **Training using computer models reported to be more effective, compared with the traditional classroom lectures (Fletcher, 1996)**
- **VR and other related technologies allow for complete recording, prevents loss of important user data compared to traditional pen-pencil based observational methods (Kempter et al., 2003)**

BACKGROUND

Automated Analysis of Data

Human-Computer Interaction

capturing user's interactions is very straightforward,
analysis is difficult
since user interface events
are very detailed and large in volume

Data Mining

knowledge-discovery in large volumes of data

Usability Engineering

- Data collection & Recoding
- Counts,metrics,statistics
- Searching sequences in data
- Comparing sequences
- Sequence extraction
- Visualization

BACKGROUND

Automated Analysis of Data

Human-Computer Interaction

capturing user's interactions is very straightforward,
analysis is difficult
since user interface events
are very detailed and la

Data Mining

knowledge-discovery in large volumes of data

What is meant by usability?

- ✓Perform main functions
- ✓Efficiency
- ✓Feelings of user
- ✓Ease of learning
- ✓Being error-proof and reliable

Usability Engineerin

- ❖Data collection & R
- ❖Counts,metrics,sta
- ❖Searching sequences in data
- ❖Comparing sequences
- ❖ Sequence extraction
- ❖ Visualization

BACKGROUND

Automated Analysis of Data

Human-Computer Interaction

capturing user's interactions is very straightforward,
analysis is difficult
since user interface events
are very detailed and la

What is meant by usability?

- ✓Perform main functions
- ✓Efficiency
- ✓Feelings of use
- ✓Ease of learnin
- ✓Being error-pro

Usability Engineeri

- ❖Data collection & R
- ❖Counts,metrics,sta
- ❖Searching sequences in data
- ❖Comparing sequences
- ❖ Sequence extraction
- ❖ Visualization

Data Mining

knowledge-discovery in large volumes of data

- ❖Statistical Methods
- ❖Association rules
- ❖Prediction techniques,
- ❖Markov processes
- ❖Sequential pattern mining techniques
- ❖Clustering
- ❖Factor analysis

RELATED WORK

Automated Analysis of Data

Human-Computer Interaction

MACSHAPA, Sanderson et al. (1994)
•spreadsheet format & pattern matching by aligning the actual logs
•visual representation of data & statistical reports

USINE, Sanderson et al. (1994)
• statistical reports
• log comparison – search for violations, cancellations etc.

Smart Agents, Hilbert and Redmiles (1998)
• modules that create a report in case of an unexpected user behavior

Maximal Repeating Patterns, Siochi and Ehrich (1990)
• detect frequent patterns in sequence of events

Data Mining

Data mining system for drop test analysis of electronic products (Zhou et al., 2001)

Data mining in software metrics (Dicks et al.,2004)

Mining Customer support call tracking database, Muehleisen (1996)

PROBLEM REQUIREMENTS

i. <u>AUTOMATED DATA COLLECTION</u>

Capturing user interaction with virtual product or system in log files

ii. <u>AUTOMATED ANALYSIS</u>

- What are the problematic points that affect performance and satisfaction?
- What are the paths users follow to accomplish certain tasks in the system?
- Do the paths followed by different users match with designer's expectations and with each other?
- What are the patterns that deviate from expected process model?
- How high is the performance of users with the system? Does it conform to expected standards?
- What are the performance characteristics of users following a certain path?

TO ANSWER THESE QUESTIONS WE PROPOSE A FOUR-PHASE APPROACH…

ANALYSIS APPROACH

PHASES

1. **Record user interaction with the system as set of events**

2. **Determine metrics that are indicators of performance in the system**

   Ex: Task Completion Times , Cancelled Tasks, Repeated Tasks…

   Calculating these values for each user

   Using statistical techniques and/or other measurements

3. **Cluster users with respect to the metrics**

4. **Extract followed paths in each cluster**

   Compare these paths with the designer data and other user clusters

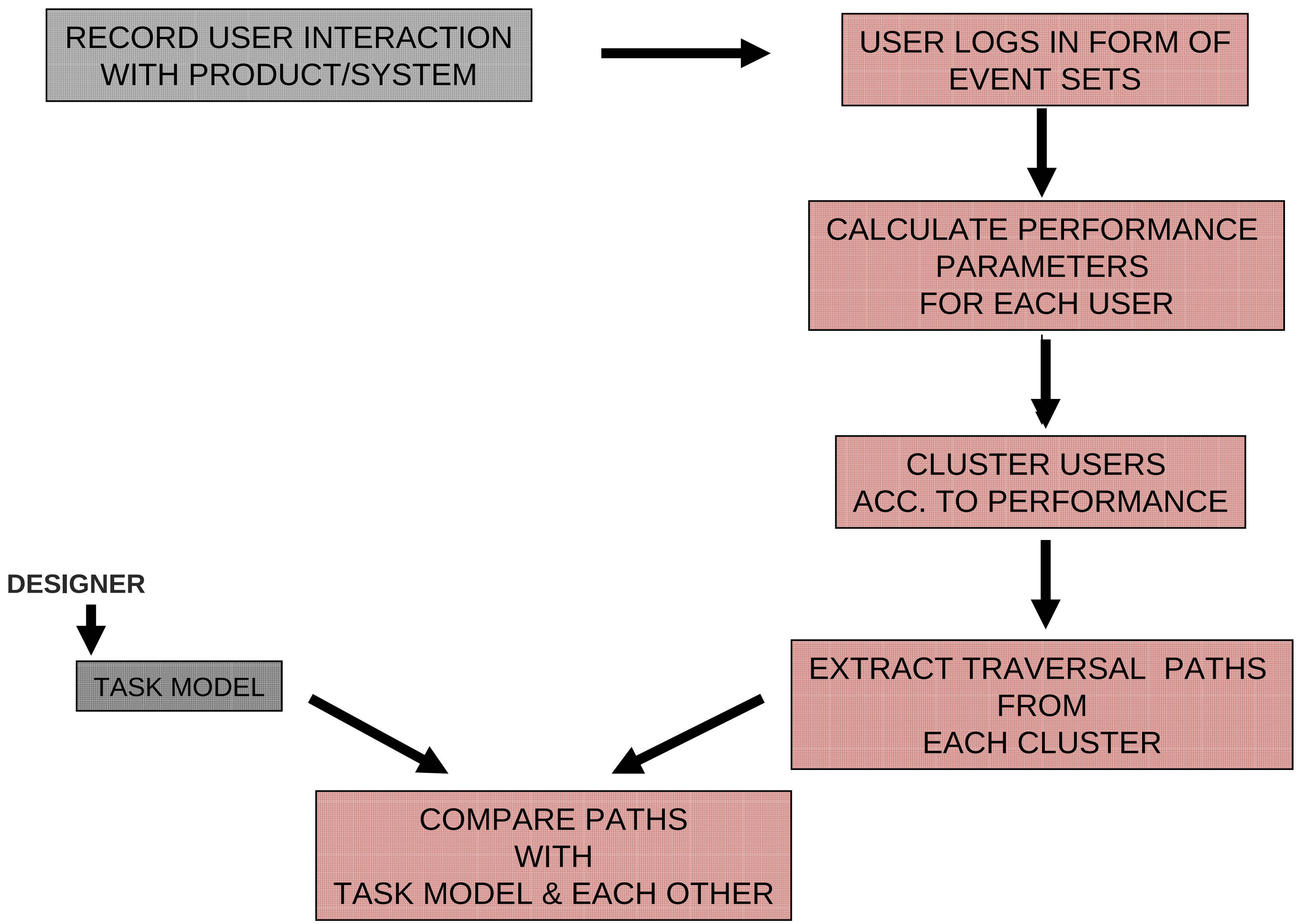
   Expert Logs
   Process Model

OUTCOME

Designed VR based Behavioral Data Analysis Tool

- Reveals problems in the design
- Suggests more efficient design & training options
- Compare alternatives in terms of usability

MODEL ARCHITECTURE



ANALYSIS METHODOLOGY

PHASE I

Capturing User Interaction With The System

For each session, record:

I. All user input and system output as a set of events

II. Necessary measurements as set of measured attributes in log files

$$R_i = \langle (i), (e_1, e_2 \ldots\ldots\ldots e_n), (a_1, a_2 \ldots\ldots\ldots a_m) \rangle$$

ANALYSIS METHODOLOGY

OUTPUT OF PHASE 1

Log files of subjects

Log file for User$_i$: Ri

$$R_i = \langle (i), (e_1, e_2 \ldots\ldots\ldots e_n), (a_1, a_2 \ldots\ldots\ldots a_m) \rangle$$

where

$$E_i = (e_1, e_2 \ldots\ldots\ldots e_n)$$

Event Stream for User i

$$A_i = (a_1, a_2 \ldots\ldots\ldots a_m)$$

Set of measured values for the session of User i

$$R_1 = \langle (1), (e_1, e_2 \ldots\ldots e_n), (a_1, a_2 \ldots\ldots a_m) \rangle$$

$$R_2 = \langle (2), (e_1, e_2 \ldots\ldots e_n), (a_1, a_2 \ldots\ldots a_m) \rangle$$

$$\vdots$$

$$R_n = \langle (n), (e_1, e_2 \ldots\ldots e_n), (a_1, a_2 \ldots\ldots a_m) \rangle$$

ANALYSIS METHODOLOGY

PHASE II* : *Performance Parameters

Determine and calculate the usability related numerical performance parameters from the session log file

❖***For any system/product designers collaboratively determine potential indicators that will apply best to their purposes***

Examples from HCI literature:

•task completion time
•repetitions
• cancellations and undo's
•the frequency and patterns in use of help
• time spent in a window

❖***Selected Indicators are calculated for each user***

OUTPUT OF PHASE II

USERS (U_i)	INDICATORS			
	I_1	I_2	$\cdots$	I_m
U_1	i_{11}	i_{12}	$\cdots$	i_{1m}
U_2	i_{21}	i_{22}	$\cdots$	i_{2m}
.	.			
U_n	i_{n1}	i_{n2}	$\cdots$	i_{nm}

ANALYSIS METHODOLOGY

***PHASE III* : Clustering Subjects with respect to the results**

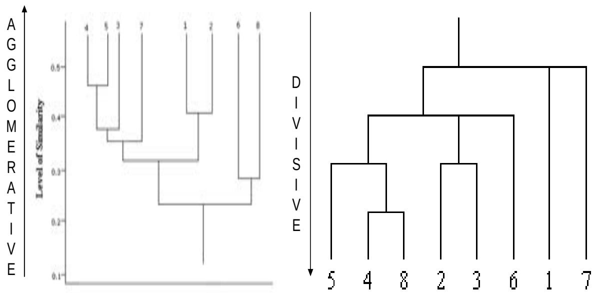
Objective: To find clusters that will represent the sample space in the least costly way

Clustering in Data Mining

Partitional Clustering
Partition the data set into *k clusters*

- *K-Means*
- *K-Medoids*

Hierarchical Clustering
Produce a nested sequence set of partitions



ANALYSIS METHODOLOGY

PHASE III **: *Clustering Subjects with respect to the results***

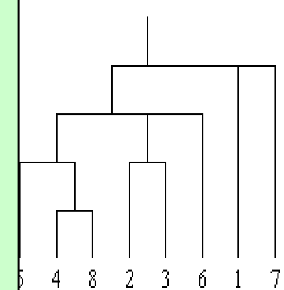
Objective: To find clusters that will represent the sample space in the least costly way

Clustering in Data Mining

Partitio
Partitio
k clust

•*K-Mea*
•*K-Med*

stering
d sequence



Choice depends on the properties of the system in question:

❖the number and kind of included attributes
❖number of subjects
❖correlation between dimensions
❖noise in parameters

ANALYSIS METHODOLOGY

PHASE IV : Approach I

Extracting recurring paths in each cluster and comparing these paths with the designer data and/or among user clusters.

Apriori Algorithms (Agrawal and Srikant, 1995) : Mining traversal patterns in a cluster & choosing the frequent patterns

Log files in Cluster k

$$R_i = \langle (i), (e_1, e_2 \ldots\ldots\ldots e_n) \rangle$$

$$R_{i+1} = \langle (i+1), (e_1, e_2 \ldots\ldots\ldots e_n) \rangle$$

.
.
.
.
.
.
.
.
.

$$R_{i+1} = \langle (i+1), (e_1, e_2 \ldots\ldots\ldots e_n) \rangle$$

Patterns for Cluster k

Path 1: $e_{11} \rightarrow e_{12} \ldots\ldots\ldots\ldots \rightarrow e_{1n}$

Path 2: $e_{21} \rightarrow e_{22} \ldots \rightarrow e_{2k}$

.
.
.
.
.
.

Path m: $e_{m1} \rightarrow e_{m2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \rightarrow e_{mv}$

ANALYSIS METHODOLOGY

PHASE IV : Approach II

Building a process model for each cluster from execution logs

Probabilistic Methods

Event Streams in Cluster k

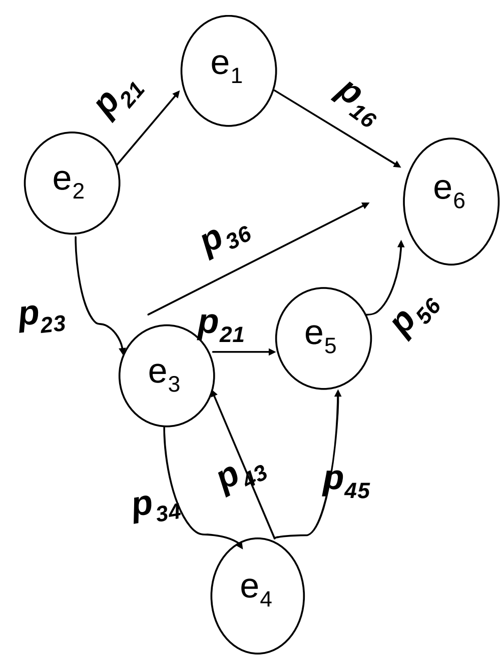
$$R_i = \langle (i), (e_1, e_2 ....e_n) \rangle$$

$$R_{i+1} = \langle (i+1), (e_1, e_2 ....e_n) \rangle$$

.
.
.
.
.
.
.

$$R_m \langle (m), (e_1, e_2 ....e_n) \rangle$$

Process Model for Cluster k



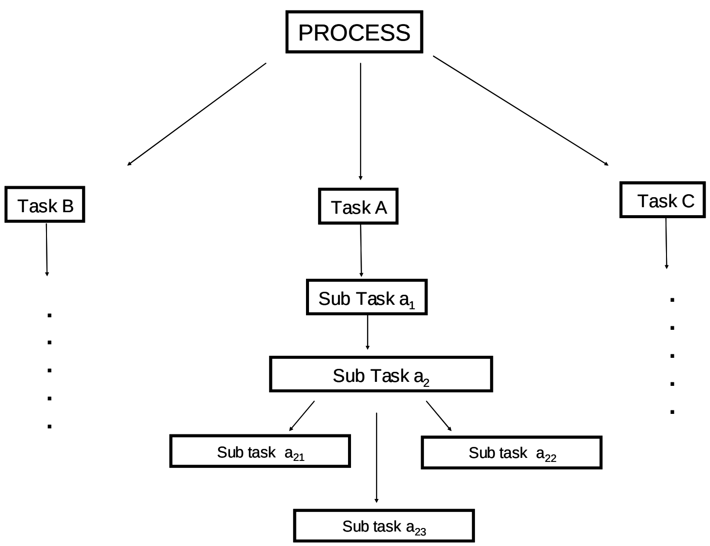
Process Model for Cluster k

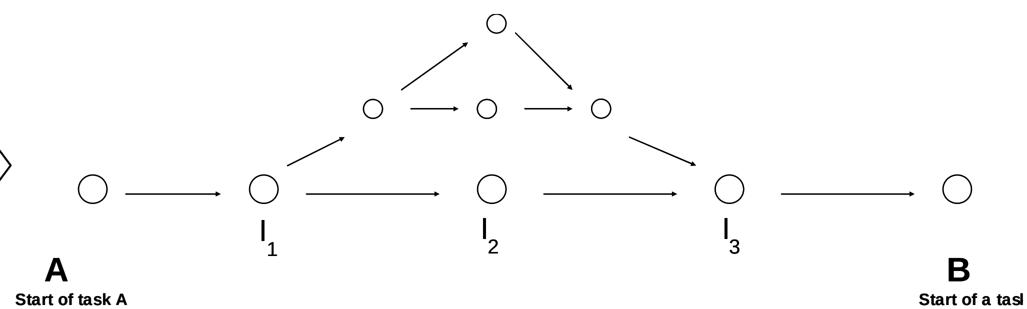
ANALYSIS METHODOLOGY

PHASE IV : Approach III
Sequence Comparison

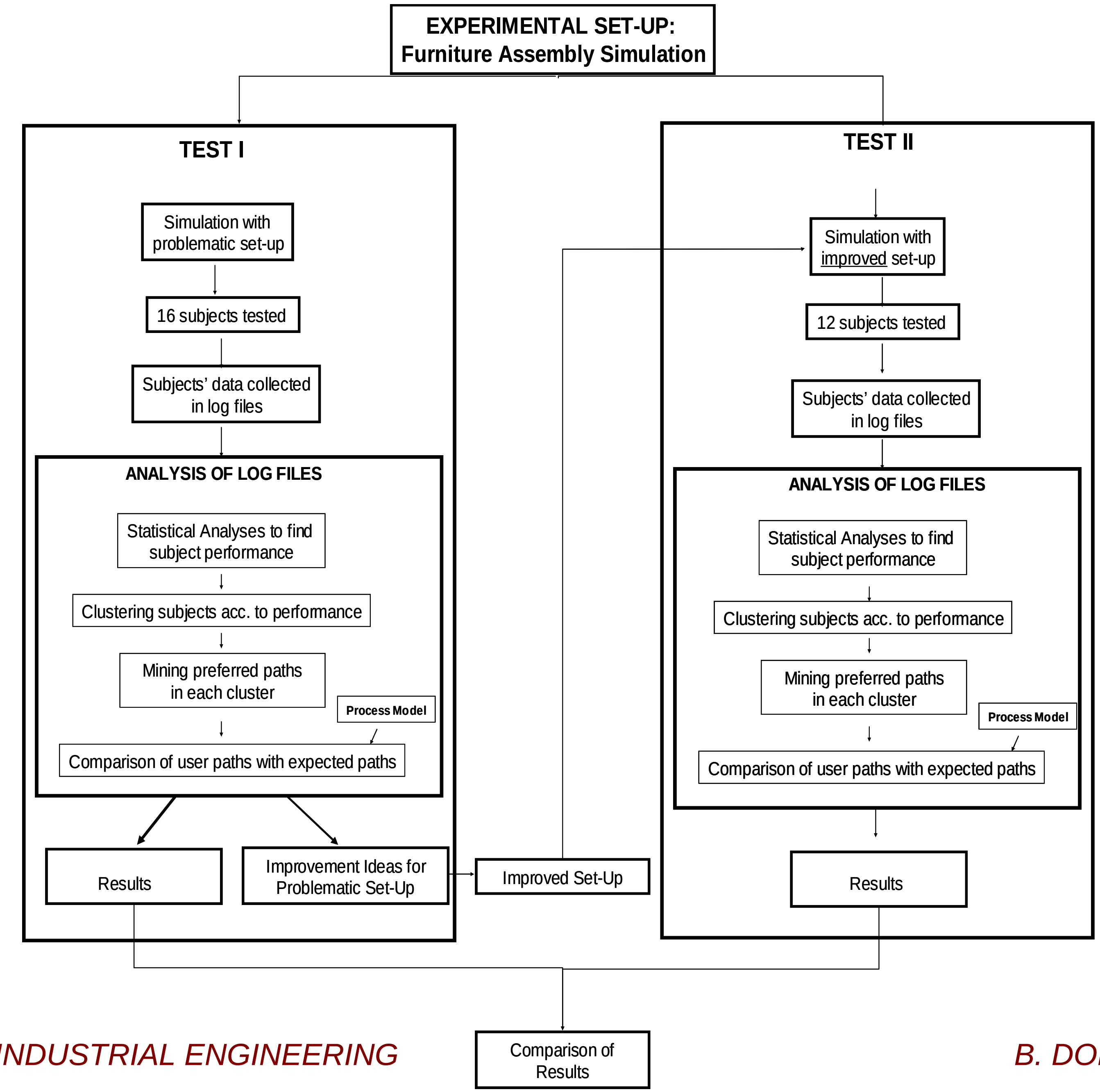
Objective: *Detect paths that diverge from target sequences*
Detect points at which divergence starts and ends

PROCESS MODEL → EXPECTED PATHS → PATH COMPARISON





EXPERIMENT


EXPERIMENTAL SET-UP:
Furniture Assembly Simulation
TEST I
Simulation with problematic set-up
16 subjects tested
Subjects' data collected in log files
ANALYSIS OF LOG FILES
Statistical Analyses to find subject performance
Clustering subjects acc. to performance
Mining preferred paths in each cluster
Process Model
Comparison of user paths with expected paths
Results
Improvement Ideas for Problematic Set-Up
Improved Set-Up
TEST II
Simulation with improved set-up
12 subjects tested
Subjects' data collected in log files
ANALYSIS OF LOG FILES
Statistical Analyses to find subject performance
Clustering subjects acc. to performance
Mining preferred paths in each cluster
Process Model
Comparison of user paths with expected paths
Results
Comparison of Results

EXPERIMENT

TEST ENVIRONMENT

Simulation of assembly process of newly purchased home furniture

- **28 subjects**
  -27 Concordia University students,1 university graduate
  -No financial compensation
- **In Concordia University graduate laboratories**
- **Time spent is 15 minutes**

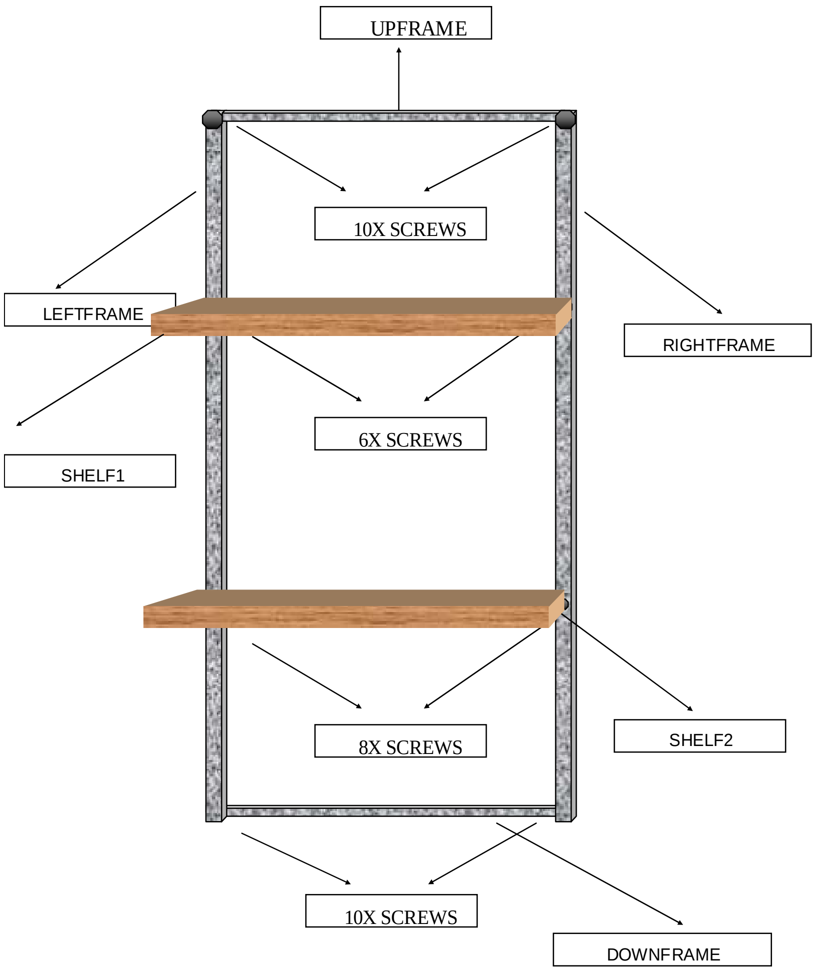
ITEM TO BE ASSEMBLED

Wall mounted shelving unit that consists of :

- **Side frames**
- **Up & down frames**
- **Shelves**
- **Screws of different types**

EXPERIMENT

ASSEMBLY INSTRUCTIONS FOR THE WALL MOUNTED SHELVING UNIT



1. Assemble left frame.

2. Pick a 10x screw and place to the top hole of the right frame.

3. Pick the second 10x screw and place to the top hole of left frame.

4.Pick a frame and assemble to the top holes on the left and right frames.

5. Pick a 10x screw and place to the bottom hole on the right frame.

6. Pick the last 10x screw and place to the bottom hole.

7. Pick the second frame and assemble to the down holes on the left and right frames.

8. Pick a 6x screw and place on the upper middle hole of right frame.

9. Pick a 6x screw and place on the upper middle hole of left frame.

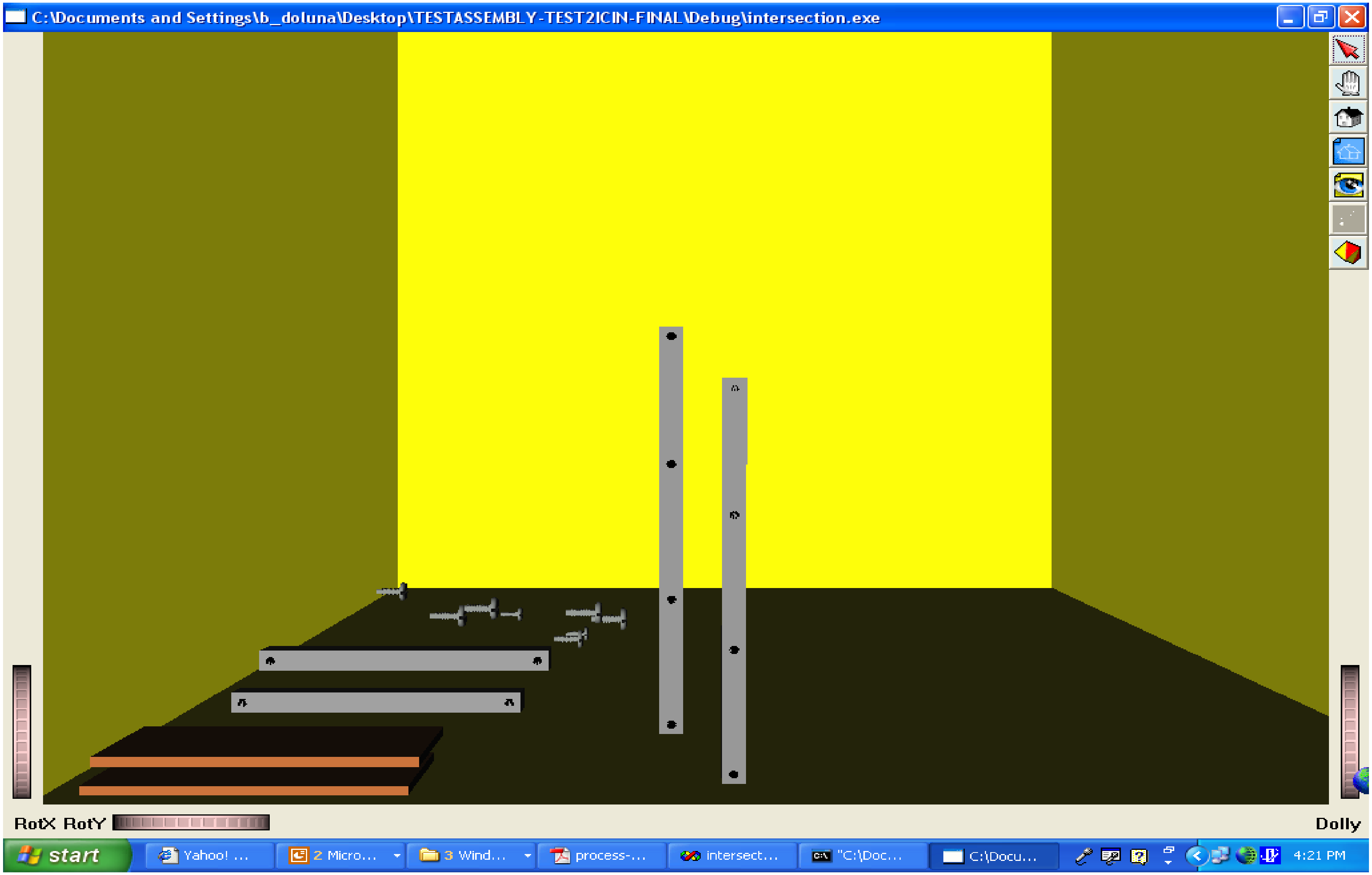
10. Pick a shelf and assemble on the upper holes.

11. Pick an 8x screw and place on the lower hole of right frame.

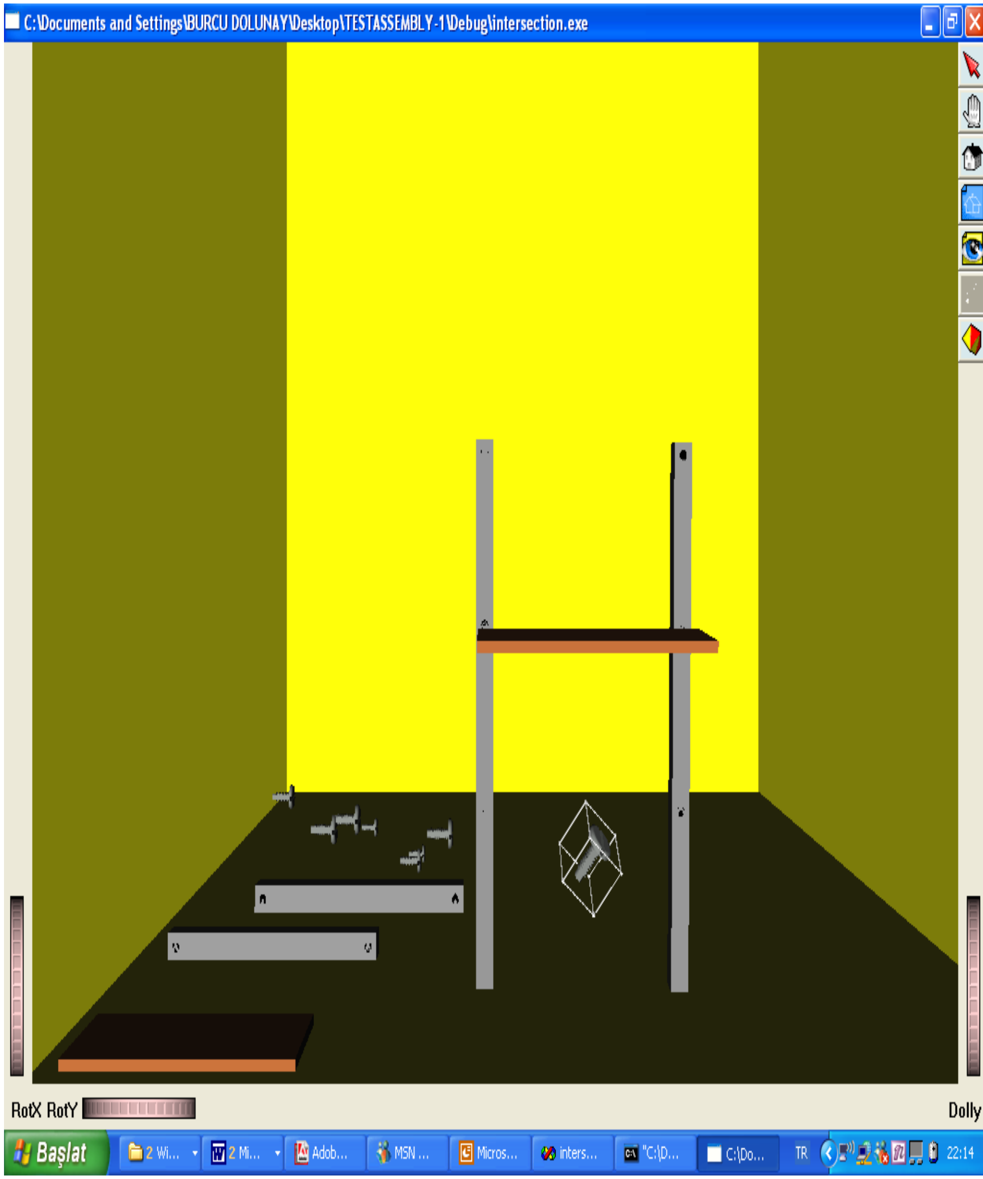
12. Pick an 8x screw and place on the lower hole of left frame.

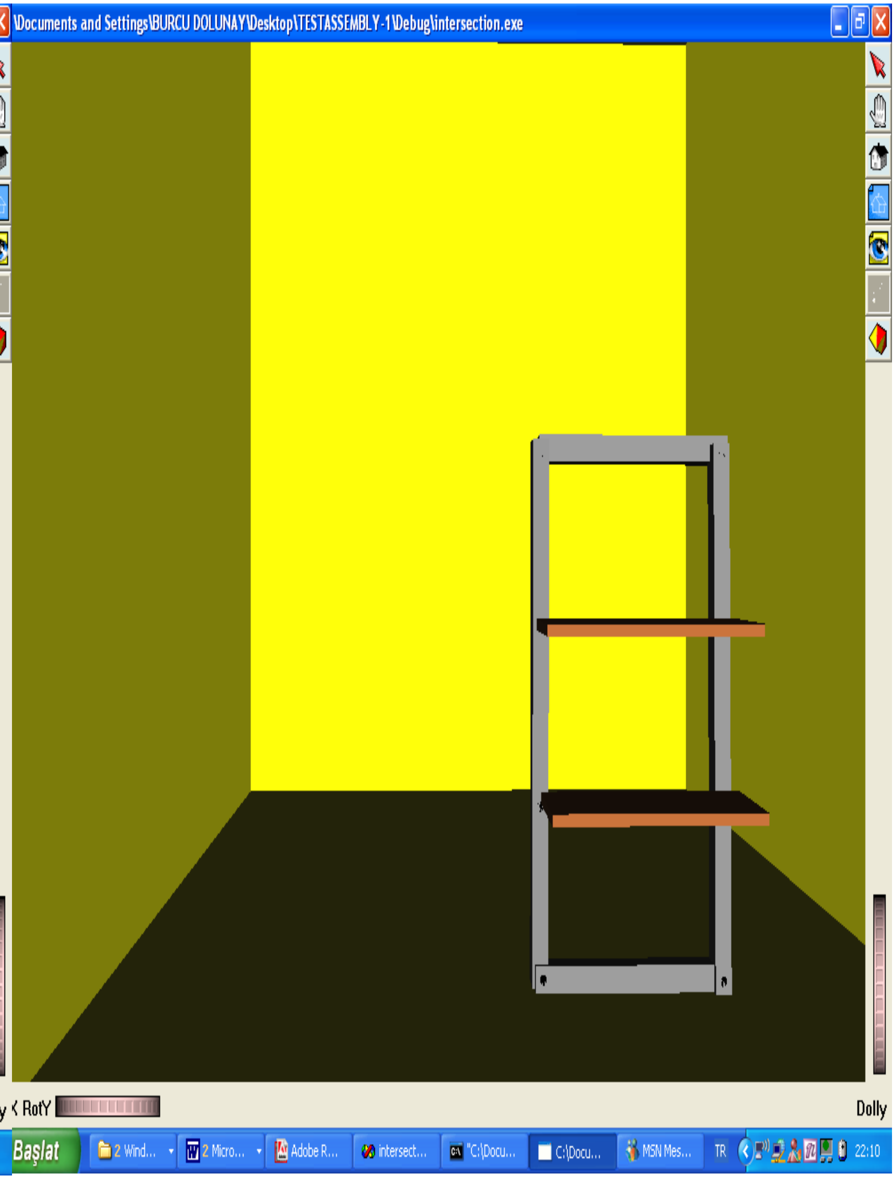
13. Pick the other shelf and assemble on the lower holes.

TEST ENVIRONMENT



TEST ENVIRONMENT





DURING ***COMPLETED***

DATA ANALYSIS

INDICATORS TO BE COLLECTED

- **Task Completion Times & Repetitions & Cancellations**

CLUSTERING

- **Correlation Analysis : Pearson Correlation Coefficient**

Repetitions & Cancellations : 0.127501763
Time & Repetitions : 0.262578038
Time & Cancellations : 0.210673528

No Evident Linear Relationship

Clustering on 3-dimensions : 3 Clusters (CL1: 11, CL2:3, CL3:2)
Data Standardization (Kaufman et al., 1990)

- **K-means Clustering Algorithm**

  - ✓ **Simple, effective**
  - ✗ **Outlier sensitivity**
  - ✗ **One data point in one cluster**

DATA ANALYSIS

Approach I - Extracting recurring paths in each cluster and comparing these paths with the designer data and/or among user clusters.

Maximal Forward Sequences Algorithm (Chen et al.,1994)

- **Find *maximal forward sequences* for each user**
  Eliminate repetitions, cancellations
  Can not be a subsequence of any other path

- **Collect the logs of a cluster in the same file**
- **Scan the file for repeating patterns - Similarity coefficient**

Approach II - Sequence Comparison

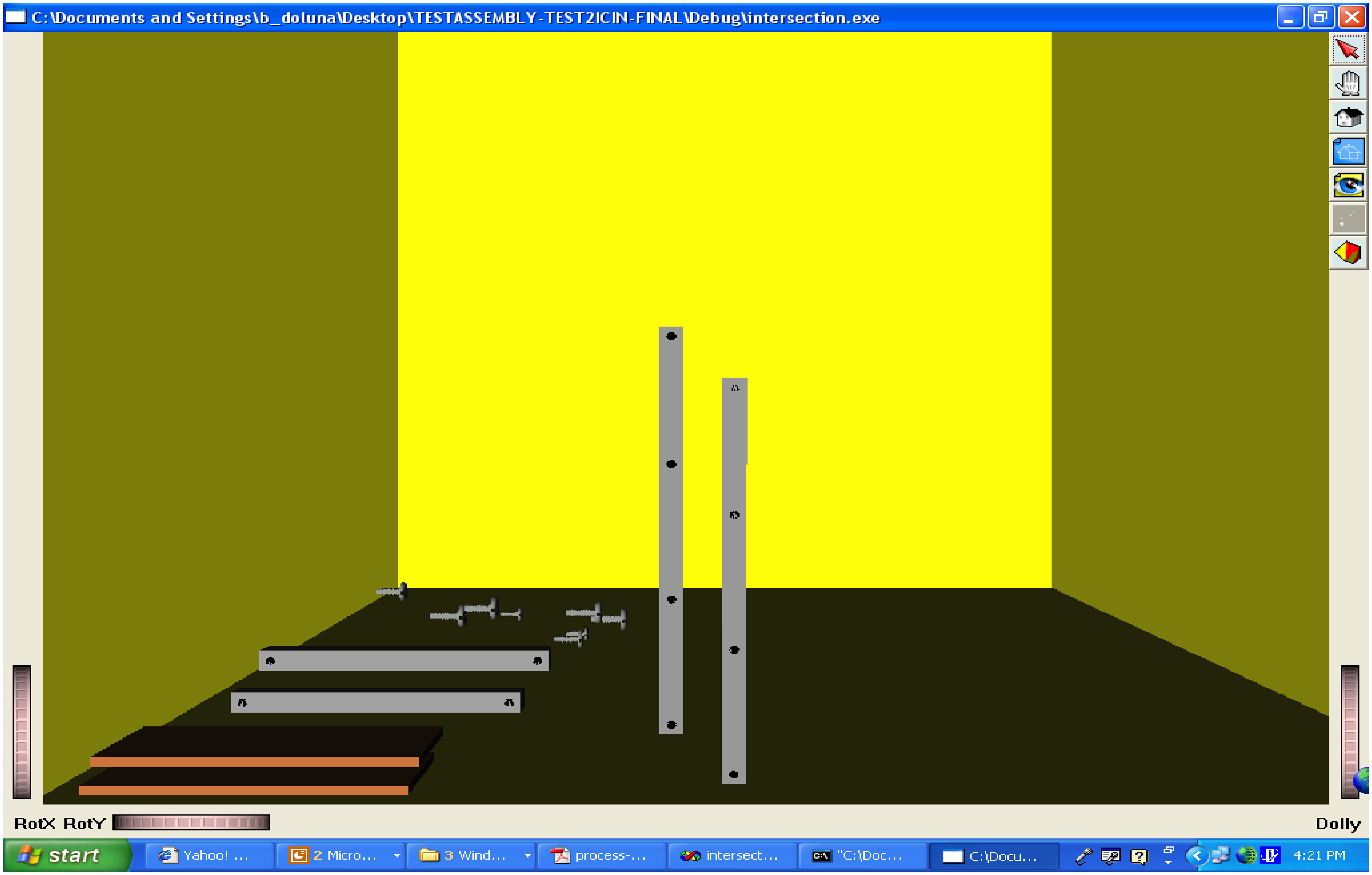
- Obtain sub-paths from process model

- Compare with subject paths

EXPERIMENT I

- **SIMULATION WITH PROBLEMATIC SCREWS**
- **16 SUBJECTS**
- **COMMON PRACTICE IN FURNITURE STORES**
- **FACED IN DAILY LIFE BY NOVICE PEOPLE**

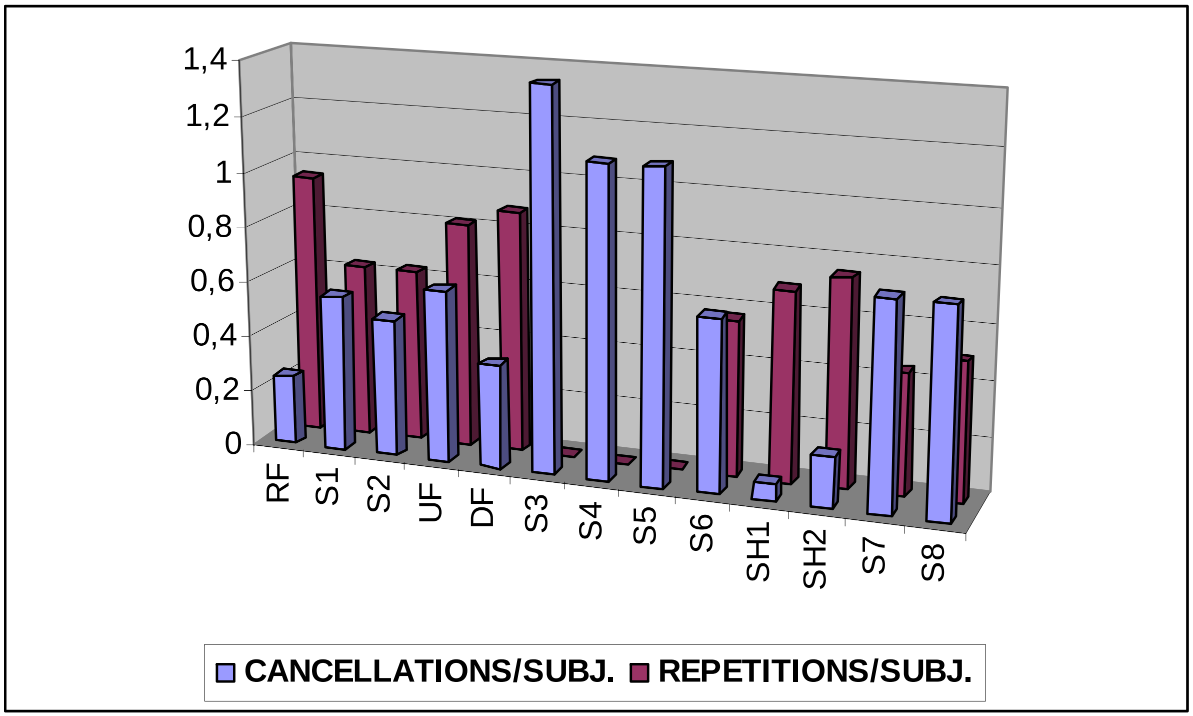
EXPERIMENT I



EXPERIMENT RESULTS

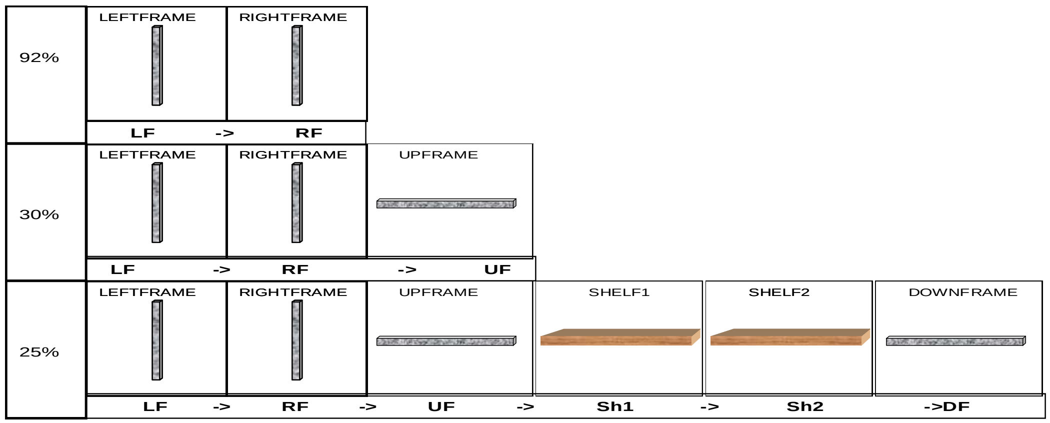
TEST I : Statistical Results

•Mean task completion time : 9.522 minutes,
•Mean number of repetitions : 3.9
•Mean number of cancellations : 4.67

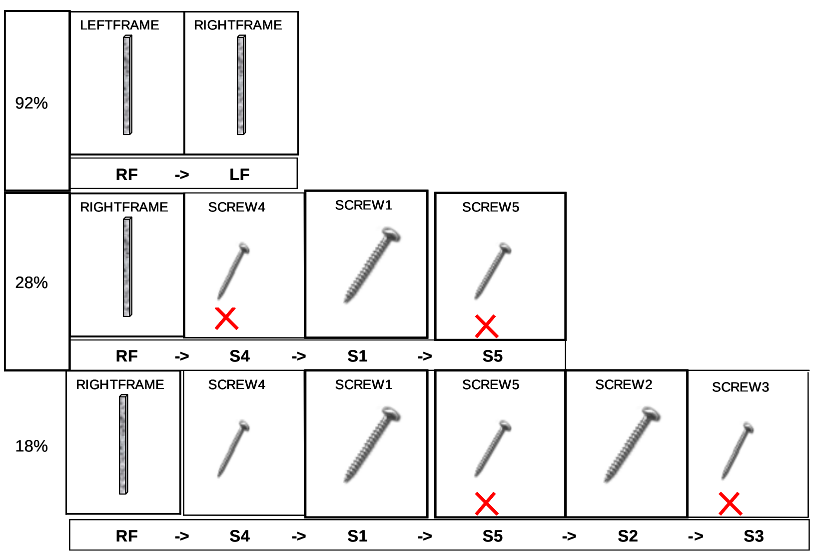


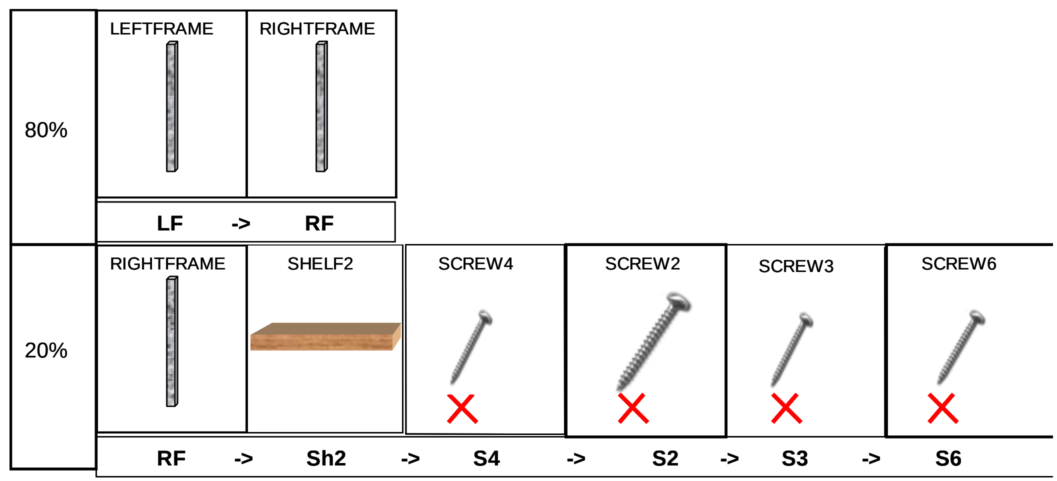
EXPERIMENT RESULTS

TEST I : Path Analysis



Paths from Cluster I





Paths from Cluster III

Paths from Cluster II

EXPERIMENT RESULTS

TEST I : Path Analysis – Comparison with Expected

❖ ***Number of violating paths is 57***

❖ ***Mean deviation: 3.56 per user***

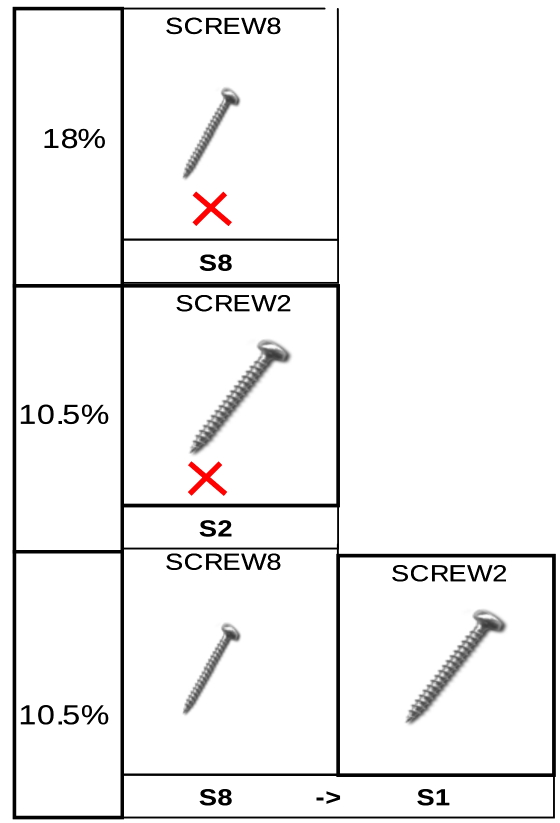
❖ ***Mean path length is 3.14***

- shortest pattern consists of one event,
- longest pattern consists of 11 events

❖ ***Long paths were not frequent***

❖ ***%12.5 started after the "assembly of down frame***

❖ ***14% turned back to normal with "picking screw1" and "picking screw2"***

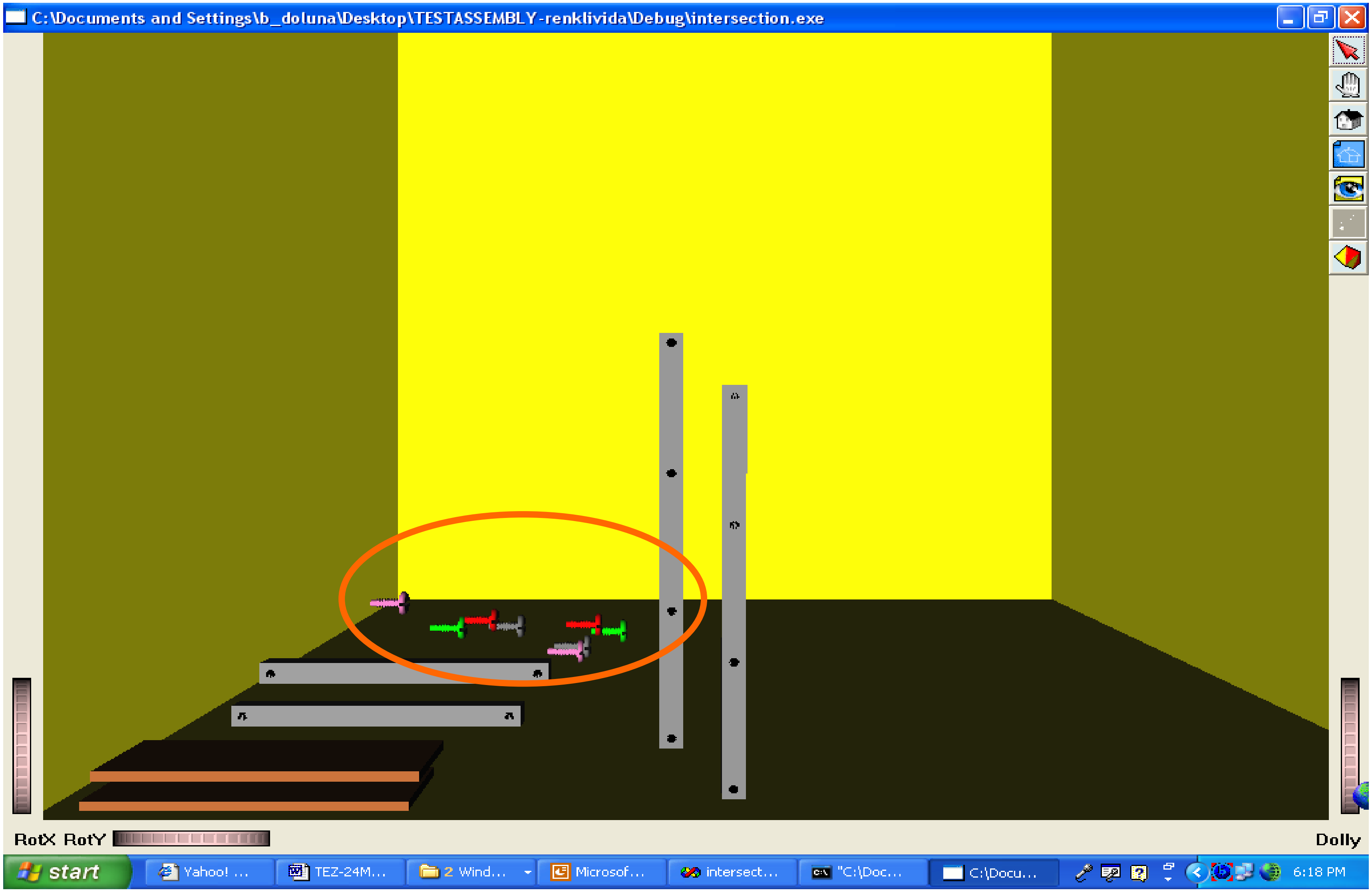


Violating sub paths

EXPERIMENT II

- SIMULATION WITH <u>DIFFERENT</u> COLOR SCREWS
- 12 SUBJECTS

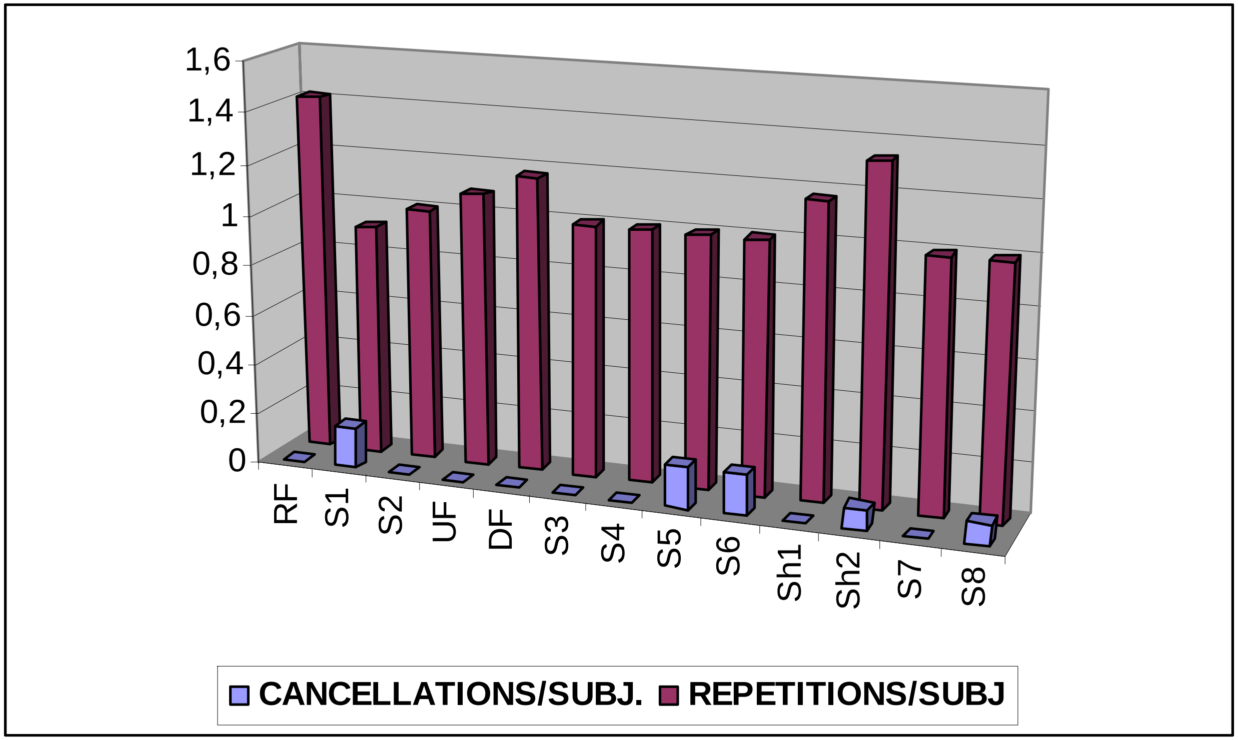
TEST ENVIRONMENT



EXPERIMENT RESULTS

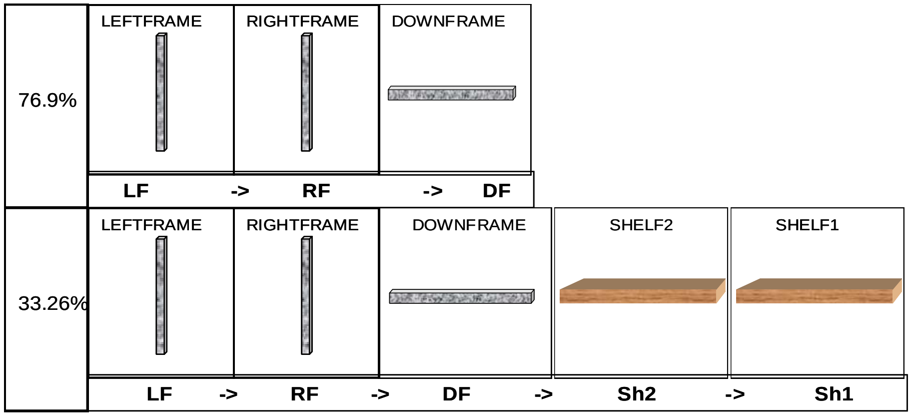
TEST II : Statistical Results

•Mean task completion time : 2.397 minutes,
•Mean number of repetitions : 2.25
•Mean number of cancellations : 0.667

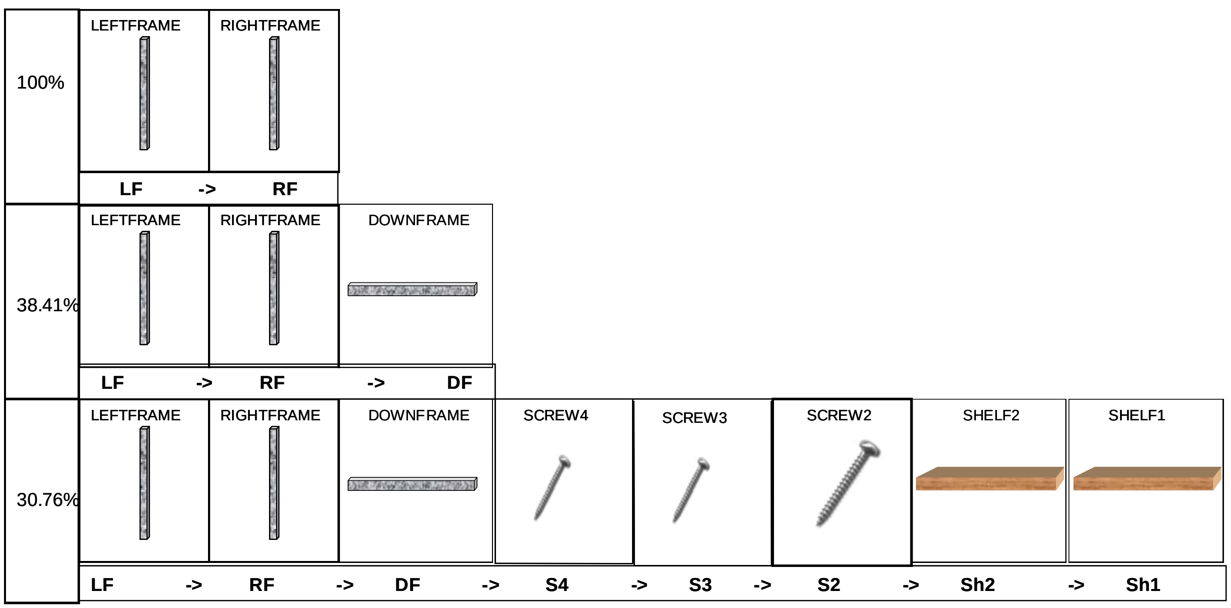


EXPERIMENT RESULTS

TEST II : Path Analysis



Paths from Cluster I



Paths from Cluster III

EXPERIMENT RESULTS

TEST II : Path Analysis – Comparison with Expected

❖***Number of violating paths was is 7***

❖***Mean deviation: 0.58 tasks per user***

❖***Mean path length is 1.14***

- shortest pattern consists of one event,
- longest pattern consists of two events

❖***Only one 2-event path***

COMPARISON OF RESULTS

- ***Mean assembly completion time decreased by 28.57% ($p<0.05$)***
- ***Mean number of cancellations decreased by 43.2% ($p<0.025$)***
- ***Decrease in the number of repetitions is not significant***
- ***Total number of deviating paths***

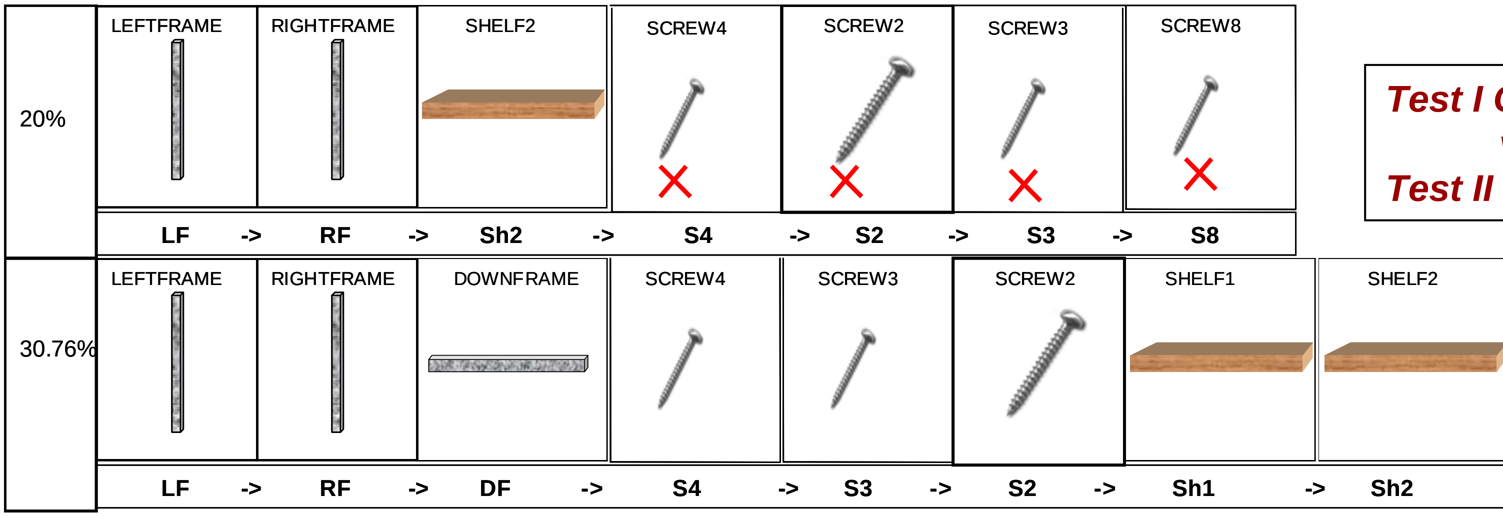
57 in the first test vs. 7 in the second test

- ***3.56deviations per subject*** ***0.58 deviations per subject***

- ***The mean length of deviation: 3.14 in the first test vs. 1.14 in the second test***

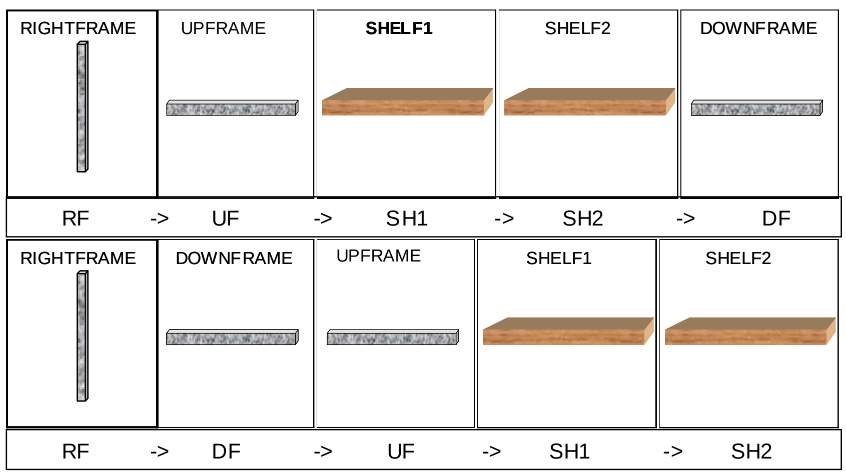
COMPARISON OF RESULTS

Path Analysis



Test I Cluster III
versus
Test II Cluster III

In both groups



POTENTIAL APPLICATIONS OF THIS WORK

IN THE AREA OF A NEW PRODUCT DEVELOPMENT

- **Provides valuable insight about users' experience with product, enables more efficient design changes**
- **Enables testing products in early stages of product life**

FOR EXISTING SYSTEMS/PRODUCTS : AFTER SALES SUPPORT

Training:

- Seeing user patterns means better insight about the weak points
- Increased efficiency in the training programs
- Increased user satisfaction

Maintenance:

For complex systems on a wide geographical region:

- Ability to trace the processes followed by different branches for maintenance of systems,
- Detect deviations from the expected maintenance process
- Ability to provide remote support without having to send people to the actual place

CONLUSION & FUTURE WORK

The methodology is applicable whenever there is a need to understand if the product is functioning as expected on user's side

ADVANTAGES:

- ***Automated data collection and analysis on large sample spaces***
- ***No location and time constraints***

FUTURE WORK:

- ***Collecting all tools and proposed data analysis methodology in a user friendly interface***
- ***Enabling different choices in the interface at each phase***
- ***Integration of a visualization aid for data presentation after analysis***
- ***Analysis of concurrent activities, activities that happen at the same time***

THANK YOU…

QUESTIONS ?