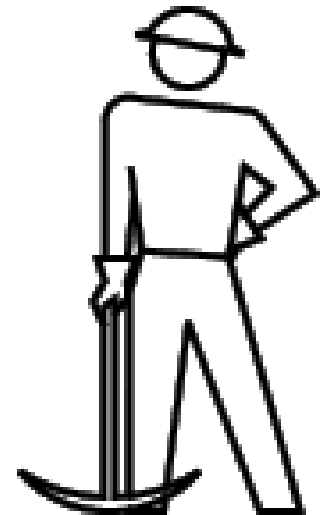
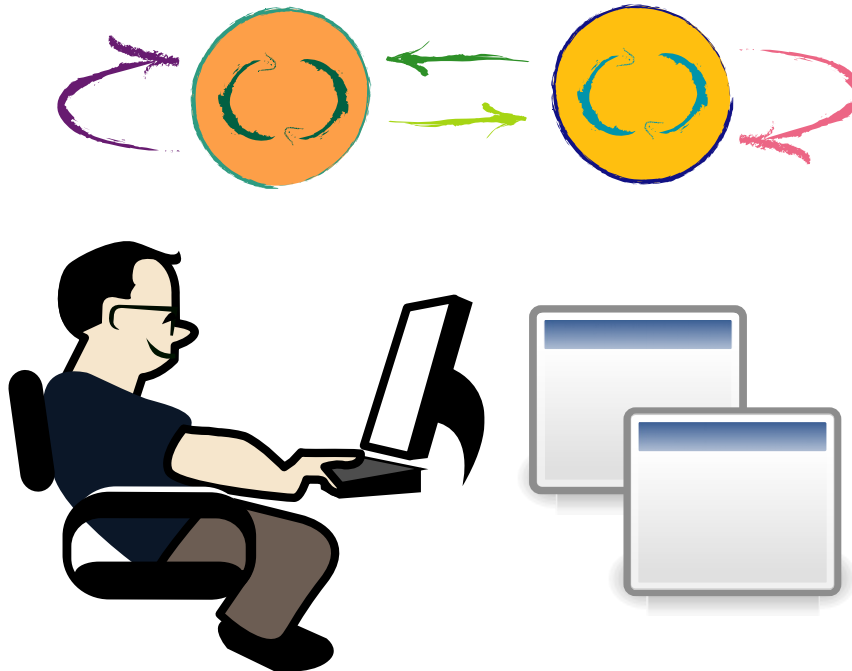


Discovery of Interaction Patterns with Graphical User Interface Usage Mining



TECHNISCHE
UNIVERSITÄT
DARMSTADT



The thesis in two parts

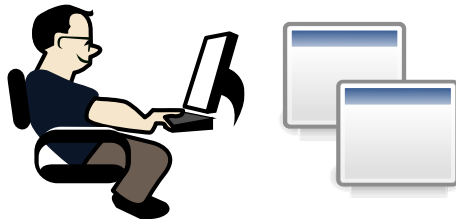


TECHNISCHE
UNIVERSITÄT
DARMSTADT

Is it possible to observe users in their day-to-day work on the Desktop
and
can this give new insights about their GUI usage behavior?

- **Graphical Software Mining**

- Observation and capturing of interactions between
 - users
 - application softwares



- **GUI Usage Mining**

- Discovery of frequent interaction sequences: Interaction patterns
 - Sequential Pattern Mining
 - Graph Mining
 - Process Mining
 - N-Gram Based Mining

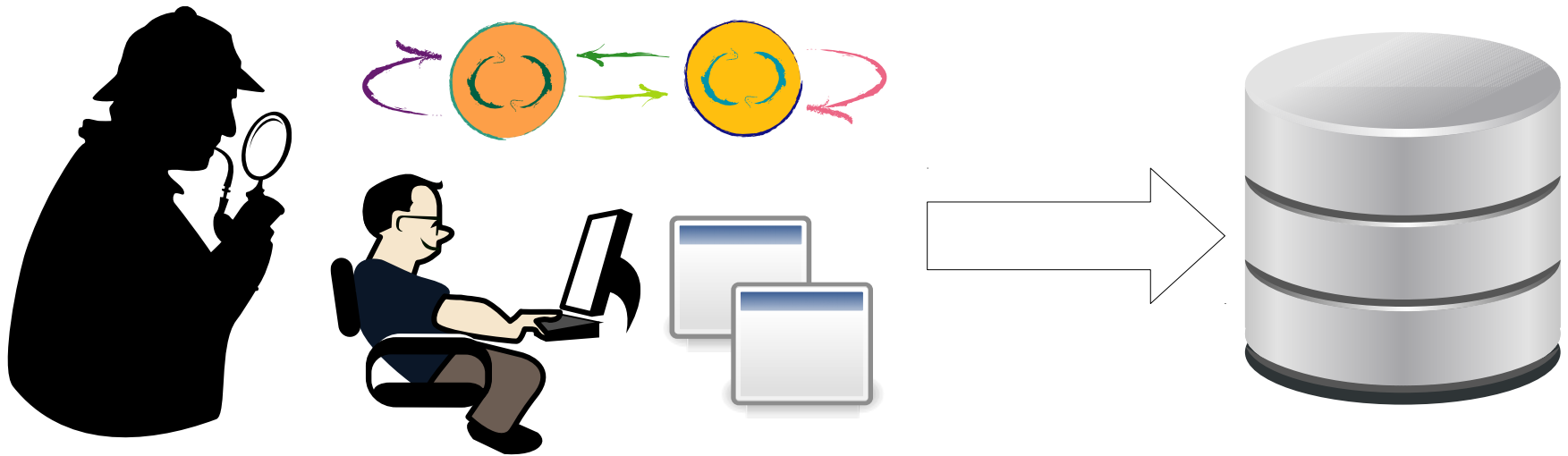


Graphical Software Mining



TECHNISCHE
UNIVERSITÄT
DARMSTADT

*... is the process of mining software **exclusively on a graphical level.***



Problem Definition

*Given a GUI environment, observe the **user** who works with some **application softwares**. While the user triggers actions, the application softwares react to them. Hence, record both side and collect the insights in a database called **interaction log**.*

Graphical Software Mining made possible by Accessibility technology

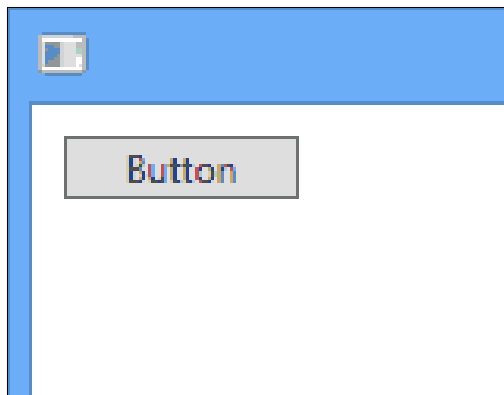


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- blind or visually impaired people
 - screen reader
 - Microsoft Active Accessibility (MSAA) [7]
 - Microsoft UI Automation (UIA) [8]



pixel based →
rendered to a pixel buffer →
throw information away



crawl

set of GUI elements with properties

- Crawl
 - Window
 - Name = "
 - Rectangle = (...)
 - Button
 - Name = 'Button'
 - Rectangle = (...)
 - Parent = Window

Interaction Initiation

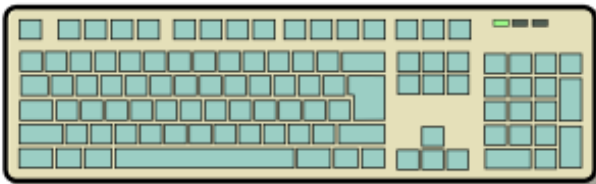


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Keyboard input device
 - no keylogger
 - only shortcuts
 - only down events

- Mouse input device
 - combination of
 - mouse entity
 - click style

$\{\text{left, middle, right, xbutton}\} \times \{\text{click, doubleclick}\}$



Program Identification



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- PID
 - volatile → persistent
- executable file
 - program hash



Program Hash

A program hash is calculated by hashing the bytes of the executable file of an application software. The used hash function is Secure Hash Algorithm (SHA) with 512 Bits.

$$f_{ph} : (b_i)_{i=1}^n \rightarrow SHA-512, b_i \in Bytes$$

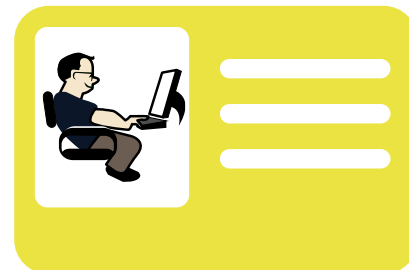
User Identification



TECHNISCHE
UNIVERSITÄT
DARMSTADT

The user ID is a hash of the following string concatenations:

- The CPU's
 - unique ID, or if not available
 - processor ID, or if not available
 - name, or if not available
 - manufacturer.
- The BIOS'
 - identification code,
 - serial number and
 - manufacturer.
- The Mainboard's
 - model,
 - name,
 - serial number and
 - manufacturer.
- The disk drive's
 - model,
 - signature,
 - total heads and
 - manufacturer.
- The video controller's
 - driver version and
 - name.
- The network adapter's
 - mac address.
- The personal computer's
 - machine name.



- Machine name
 - for debug
 - analysis in context

GUI Element of Interest

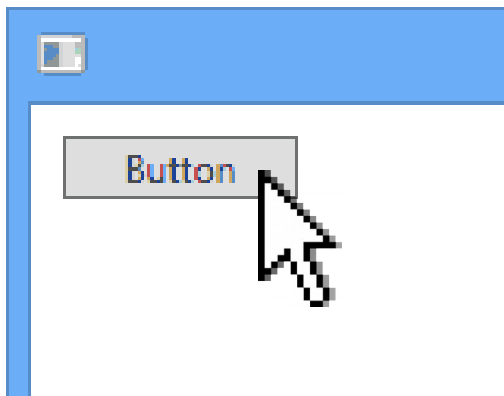


TECHNISCHE
UNIVERSITÄT
DARMSTADT

Element of Interest

The element of interest (EOI) is a GUI element the user interacts with in a certain moment.

For instance, an EOI is a clicked button, a text field the user enters text or a hovered menu item.



Name

Email

Telephone



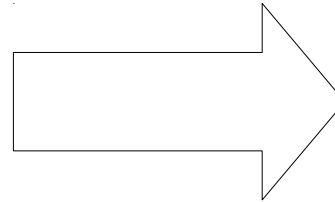
GUI Element of Interest

(1) Keyboard

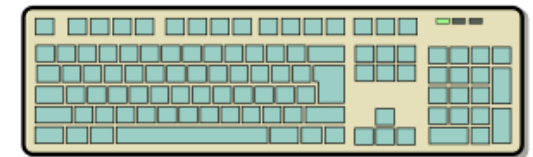


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- HasKeyboardFocus Property [6]
 - true whenever the element has the keyboard focus



Name	<input type="text" value="I"/>
Email	<input type="text"/>
Telephone	<input type="text"/>



$$\exists_{=1} e \in Crawl \quad \text{HasKeyboardFocus}(e) \equiv \text{true} \Leftrightarrow e \in \text{ElementOfInterest}$$

GUI Element of Interest

(2) Mouse



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- two methods to determine the EOI with a library call

- (2.1) *AccessibleObjectFromPoint* [4]

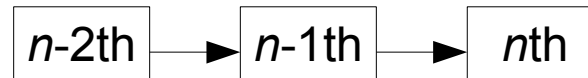
- (2.2) *ElementFromPoint* [5]

- third auxiliary method (2.3) *RankedFromPoint*

$a \equiv e \equiv r \Leftrightarrow a \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $a \equiv e \Leftrightarrow a \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $e \equiv r \Leftrightarrow e \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $a \equiv r \Leftrightarrow a \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $r \neq \text{null} \Leftrightarrow r \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $e \neq \text{null} \Leftrightarrow e \in \text{ElementOfInterest}$ else if $\text{ElementOfInterest} = \emptyset$
 $a \neq \text{null} \Leftrightarrow a \in \text{ElementOfInterest}$

- compensates errors of (2.1) and (2.2)

- $\text{Same} := n-1 \text{th Crawl} \cap n-2 \text{th Crawl}$
 $\text{Added} := n-1 \text{th Crawl} \setminus n-2 \text{th Crawl}$
 $\text{Candidates} := \text{Same} \cup \text{Added}$



$\text{PossiblyClickedElements} := \{e \mid e \in \text{Candidates} \wedge \text{intersection}(\text{rectangle}(e), \text{cursor})\}$

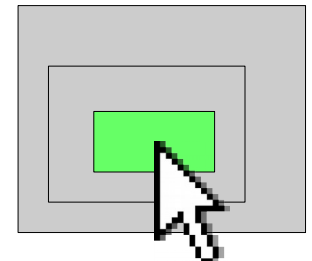
- four criteria rank

(a) rectangle area → more smaller, more likely EOI

(b) level → newer element, higher its level, more likely EOI

(c) menu item control type → menu items always on top

(d) foreground window belonging → distinguish elements from different windows



GUI Element Identification



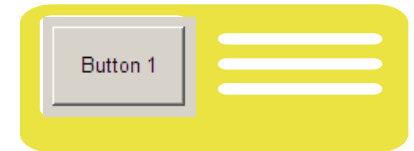
TECHNISCHE
UNIVERSITÄT
DARMSTADT

- **Between Runtime**

- properties
 - Name (string)
 - LegacyIAccessibleDescription (string)
 - AccessKey (string)
 - LegacyIAccessibleChildId (int)
 - AutomationId (string)
 - ClassName (string)
 - ControlType (int)
 - LocalizedControlType (string)
 - LegacyIAccessibleRole (int)
 - HelpText (string)
- tree structure (parent)
- child index

- **During Runtime**

- Runtime ID
- if not available → generate
- detect *dynamic* properties
 - cache hit → compared on identification property level
 - different → changed during runtime



Runtime ID: 5
Name = 'VLC'

=
≠

Runtime ID: 5
Name = 'VLC - Song01'

Name property is dynamic

GUI Asynchrony

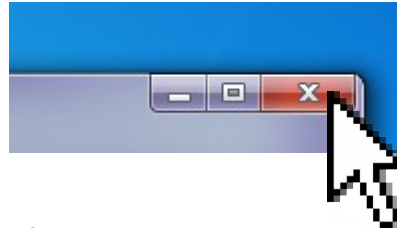


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- mouse event forwarding

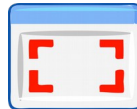


- GUI element
 - disappear after being clicked
 - other elements overlap
- intercept mouse event → determine the EOI → forward mouse event



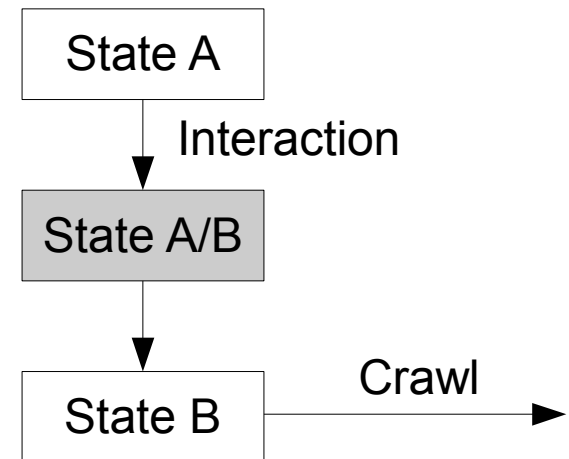
- application software waiting

- accomplished state change
 - process → wait for input idle
 - interaction state of a window (responding, wait for input)
 - hour glass cursor



- clean crawling

- meantime another state change
- detected with simultaneous
 - crawling state
 - listening for new events
- new event appears *while* crawling is not finished → invalid crawl



- censors keyboard events to password input fields
 - IsPassword property [6]
 - $\exists e \in \text{Crawl} \text{ isPassword}(e) \wedge \text{hasKeyboardFocus}(e) \Rightarrow \text{censor}(\text{keyboardEvent})$
- privacy data expressed in GUI elements
 - example → email subject



	Betreff	Von	Datum
☆	eBay-Passwort zurücksetzen	eBay	27.06.2014 17:00
☆	Bestätigung, dass Ihr eBay-Pa...	eBay	27.06.2014 17:01
☆	Allgemeine Geschäftsbeding...	Anmeldung@ebay.de	27.06.2014 20:40
☆	Herzlich willkommen bei eBay!	eBay	27.06.2014 20:40
☆	Ihr Angebot Panasonic DVD-...	eBay	27.06.2014 21:08
☆	Willkommen bei eBay: Hier is...	eBay	29.06.2014 18:02
☆	Neue Artikel, die zu Ihrer Suc...	eBay	04.07.2014 13:33
☆	Herzlichen Glückwunsch, Ihr ...	eBay	04.07.2014 21:08
☆	Ich werde die Bezahlung in H...	eBay	05.07.2014 21:34

- Crawl
 - Dataltem
 - Name = 'eBay-Passwort zurücksetzen'
 - Dataltem
 - Name = 'Bestätigung, dass Ihr eBay-Passwort geändert wurde'
 - ...

Result Interaction Log



TECHNISCHE
UNIVERSITÄT
DARMSTADT



ID	Timestamp	User	Program	Type	Arg1	Event Arg2	Arg3	Arg4	Arg5	Crawl	ValidCrawl
37701	16/10/2014 22:24:41	Part.8	Outlook.exe	KeyboardEvent	RETURN	True	False	False	53763	96CB...	True
37702	16/10/2014 22:27:52	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24686	0D5E...	True
37703	16/10/2014 22:30:57	Part.8	Outlook.exe	KeyboardEvent	VK_V	True	False	False	53769	0D5E...	True
37704	16/10/2014 22:34:04	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24687	0D5E...	True
37705	16/10/2014 22:37:17	Part.8	Outlook.exe	MouseEvent	Left	Click	null	146181	53763	50BC...	True
37706	16/10/2014 22:40:46	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196835	196681	8E91...	True
37707	16/10/2014 22:41:07	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	25106	0000...	True
37708	16/10/2014 22:41:27	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37709	16/10/2014 22:42:03	Part.8	Outlook.exe	MouseEvent	Left	Click	null	53826	null	0000...	True
37710	16/10/2014 22:42:24	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37711	16/10/2014 22:42:45	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37712	16/10/2014 22:43:06	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37713	16/10/2014 22:43:42	Part.8	Outlook.exe	MouseEvent	Left	Click	null	146181	null	0000...	True
37714	16/10/2014 22:44:03	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37715	16/10/2014 22:52:11	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24659	null	4E18...	True
37716	16/10/2014 22:58:39	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196838	24978	4E18...	True
37717	16/10/2014 23:03:49	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	3FF1...	True
37718	16/10/2014 23:09:11	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196838	24953	3FF1...	True
37719	16/10/2014 23:14:17	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	3FF1...	False
37720	16/10/2014 23:19:37	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	True
37721	16/10/2014 23:19:58	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	0000...	False
37722	16/10/2014 23:20:34	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	null	0000...	True
37723	16/10/2014 23:25:47	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	3FF1...	False
37724	16/10/2014 23:26:24	Part.8	Outlook.exe	MouseEvent	Left	DoubleClick	null	24670	24670	0000...	True
37725	16/10/2014 23:31:29	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	3FF1...	False
37726	16/10/2014 23:36:50	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	False
37727	16/10/2014 23:42:09	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	True
37728	16/10/2014 23:42:16	Part.8	TKTracker.exe	MouseEvent	Left	Click	null	null	null	0000...	True
37729	16/10/2014 23:42:37	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	53761	0000...	False
37730	16/10/2014 23:43:13	Part.8	Outlook.exe	MouseEvent	Left	DoubleClick	null	57039	null	0000...	True
37731	16/10/2014 23:49:49	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0FBA...	True
37732	16/10/2014 23:50:25	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24659	53846	0000...	False
37733	16/10/2014 23:50:46	Part.8	Outlook.exe	KeyboardEvent	VK_C	True	False	False	null	0000...	True
37734	16/10/2014 23:51:07	Part.8	Outlook.exe	KeyboardEvent	VK_V	True	False	False	null	0000...	True
37735	16/10/2014 23:57:24	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	48B8...	True

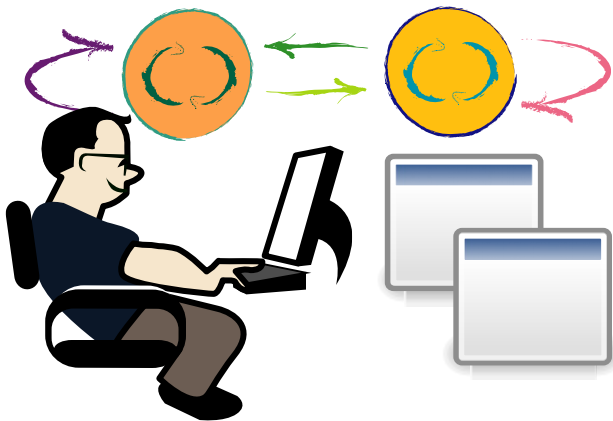
- 9 study participants
 - 12 participant-PC pairs
- 70 days
- 17759 interactions
- 160 application softwares
 - 105 distinct → version
- 247 user-program relationships
- 230963 distinct GUI elements
 - 4620 interacted with
 - 161 keyboard
 - 4459 mouse
- 4863 crawls

GUI Usage Mining

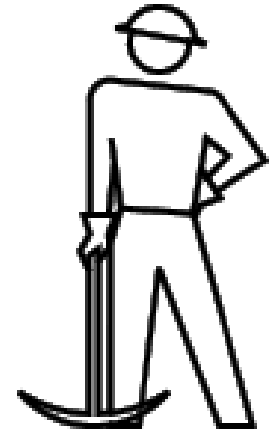


TECHNISCHE
UNIVERSITÄT
DARMSTADT

... is the automatic **discovery** of usage information from GUI interaction logs. Desktop Functionality Usage Mining focuses on the functional aspect of the desktop (e.g. GUI workflows, patterns and commands). In particular, both CLI and **GUI** give access to functionality.



ID	Timestamp	User	Program	Type	Event	Arg1	Arg2	Arg3	Arg4	Arg5	Crawl	ValidCrawl
37701	16/10/2014 22:24:41	Part.8	Outlook.exe	KeyboardEvent	RETURN	True	False	False	53763	96CB...	True	True
37702	16/10/2014 22:27:52	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24606	005E...	True	True
37703	16/10/2014 22:30:57	Part.8	Outlook.exe	KeyboardEvent	VK_V	True	False	False	53769	005E...	True	True
37704	16/10/2014 22:34:04	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24647	005E...	True	True
37705	16/10/2014 22:37:17	Part.8	Outlook.exe	MouseEvent	Left	Click	null	146181	53763	50BC...	True	True
37706	16/10/2014 22:40:46	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196835	19681	8591...	True	True
37707	16/10/2014 22:41:07	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	25106	0000...	True	True
37708	16/10/2014 22:41:27	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37709	16/10/2014 22:42:03	Part.8	Outlook.exe	MouseEvent	Left	Click	null	53826	null	0000...	True	True
37710	16/10/2014 22:42:24	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37711	16/10/2014 22:42:45	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37712	16/10/2014 22:43:06	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37713	16/10/2014 22:43:42	Part.8	Outlook.exe	MouseEvent	Left	Click	null	146181	null	0000...	True	True
37714	16/10/2014 22:44:03	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37715	16/10/2014 22:55:11	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24659	null	4E18...	True	True
37716	16/10/2014 22:58:39	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196838	24978	4E18...	True	True
37717	16/10/2014 23:03:49	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	3FF1...	True	True
37718	16/10/2014 23:09:11	Part.8	Outlook.exe	MouseEvent	Left	Click	null	196838	24953	3FF1...	True	True
37719	16/10/2014 23:14:17	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	3FF1...	False	False
37720	16/10/2014 23:19:37	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	True	True
37721	16/10/2014 23:19:58	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	24670	0000...	False	False
37722	16/10/2014 23:20:34	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	null	0000...	True	True
37723	16/10/2014 23:25:47	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	3FF1...	False	False
37724	16/10/2014 23:26:24	Part.8	Outlook.exe	MouseEvent	Left	DoubleClick	null	24670	24670	0000...	True	True
37725	16/10/2014 23:31:29	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	3FF1...	True	True
37726	16/10/2014 23:36:50	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	False	False
37727	16/10/2014 23:42:09	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24670	24670	3FF1...	True	True
37728	16/10/2014 23:42:16	Part.8	TKTracker.exe	MouseEvent	Left	Click	null	null	null	0000...	True	True
37729	16/10/2014 23:42:37	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	53761	0000...	False	False
37730	16/10/2014 23:45:13	Part.8	Outlook.exe	MouseEvent	Left	DoubleClick	null	57539	null	0000...	True	True
37731	16/10/2014 23:49:49	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	0FBA...	True	True	True
37732	16/10/2014 23:50:25	Part.8	Outlook.exe	MouseEvent	Left	Click	null	24659	53846	0000...	False	False
37733	16/10/2014 23:50:46	Part.8	Outlook.exe	KeyboardEvent	VK_C	True	False	False	null	0000...	True	True
37734	16/10/2014 23:51:07	Part.8	Outlook.exe	KeyboardEvent	VK_V	True	False	False	null	0000...	True	True
37735	16/10/2014 23:57:24	Part.8	Outlook.exe	MouseEvent	Left	Click	null	null	null	4888...	True	True

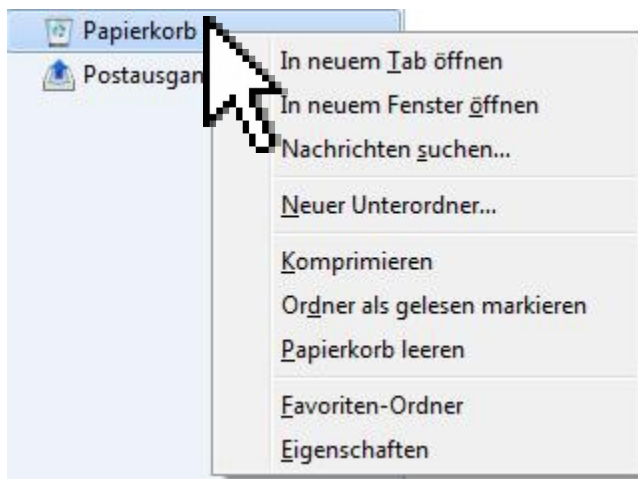


Problem Definition

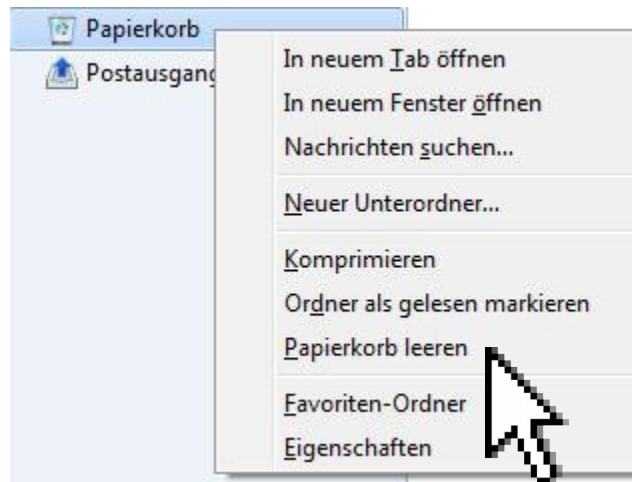
Given a stream of **interactions**, made by one user with one application software, determine frequently interesting reappearing sequences which are the product of intentional actions (**patterns**).

Pattern

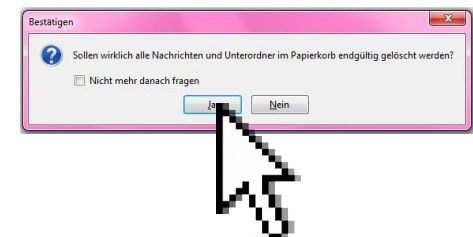
A (graphical user interface interaction) pattern is a sequence of interactions that reoccur among users to express an intentional action. Patterns are caused by the GUI design of an application software: Users are forced to perform certain sequences of interactions which represent tasks.



1. Right click on the trash icon to open a context menu



2. Left click on the menu item "empty trash"



3. Left click on the button "yes" to confirm the action

Reference Patterns



TECHNISCHE
UNIVERSITÄT
DARMSTADT

... are patterns which are annotated by experts.

Id	Time Stamp	Event	EOI	GEOI	Crawl	User	Program		show data table (1)
31512	22.09...	Mouse Left Click	Pane-137477 grid window						
31513	22.09...	Mouse Left Click	Pane-137475 panel						
31514	22.09...	Mouse Left Click	Edit-137444						
31515	22.09...	Mouse Left Click	Button-137494 Schließen Schließt das Fens...	Button-137494 S					
31516	22.09...	Mouse Left Click	Tree-51199	Tree-51199					
31517	22.09...	Mouse Left Click	Pane-137380 grid window						
31518	22.09...	Mouse Left DoubleClick	Pane-137524 stcwindow						
31519	22.09...	Keyboard Control + C	Pane-137524 stcwindow						
31520	22.09...	Mouse Left Click	Pane-137380 grid window		81B72...	AL...	pgadmin3.exe pgAdmin III ...		
31521	22.09...	Mouse Right Click	Pane-137380 grid window		7729F...	AL...	pgadmin3.exe pgAdmin III ...		
31522	22.09...	Mouse Left Click	MenuItem-51351 Daten anzeigen	MenuItem-51351 ...	769AF...	AL...	pgadmin3.exe pgAdmin III ...		
31523	22.09...	Mouse Left Click	MenuItem-51359 Die obersten (100) Zeilen ...	MenuItem-51359 ...	9DE0...	AL...	pgadmin3.exe pgAdmin III ...		
31524	22.09...	Mouse Left Click	Window-137538 Daten editieren - Kraken....	Window-137538 ...	9DE0...	AL...	pgadmin3.exe pgAdmin III ...		
31525	22.09...	Mouse Left Click	Button-137554 Schließen Schließt das Fens...	Button-137554 S...	81B72...	AL...	pgadmin3.exe pgAdmin III ...		
31526	22.09...	Mouse Left Click	TitleBar-51202 Zeigt den Fensternamen an...		81B72...	AL...	pgadmin3.exe pgAdmin III ...		
31527	22.09...	Mouse Left Click	TreeItem-124331 device_registration devic...	Tree-51199	7729F...	AL...	pgadmin3.exe pgAdmin III ...		
31528	22.09...	Mouse Right Click	TreeItem-124331 device_regis						
31529	22.09...	Mouse Left Click	Pane-137617 grid window						
31530	22.09...	Mouse Left Click	Pane-137615 panel						
31531	22.09...	Mouse Left Click	Par						
31532	22.09...	Mouse Left Click	Par						
31533	22.09...	Mouse Right Click	TreeItem-124332 device_r						
31534	22.09...	Mouse Left Click	MenuItem-51351 Daten anzeig						
31535	22.09...	Mouse Left Click	List-51220						

- reproduce made interactions
- why performed given sequences of interactions?
- give a name

(31521, 31522, 31523) [length=3]

Name SHOW DATA TABLE

Context pgAdmin III - PostgreSQL Tools 1.18.1

Problem The user wants to look at rows of a data table.

Solution (1) Mouse Right Click TreeItem-*, Mouse Left Click MenuItem-51351 Daten anzeigen, Mouse Left Click MenuItem-51360 Die letzten (100) Zeilen zeigen
(2) Mouse Right Click TreeItem-*, Mouse Left Click MenuItem-51351 Daten anzeigen, Mouse Left Click MenuItem-51360 Die oberen (100) Zeilen zeigen

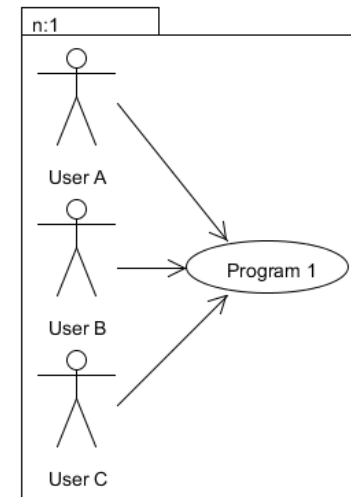
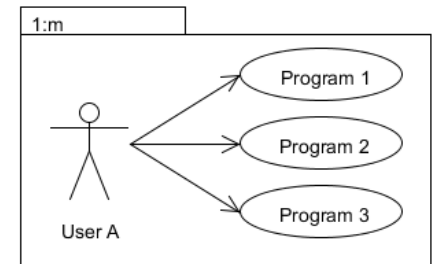
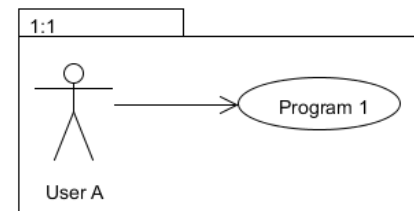
Preprocessing

1:1 relationship



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- One user interacts with one application software (1:1)
 - simplest and most suitable for discovering patterns
 - focus → **one** specific user working with **one** specific application software
 - patterns → *depend* on user & *how* application softwares used
 - every user evaluates his own discovered patterns



Preprocessing Generalized EOI



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- EOI too specific → generalize

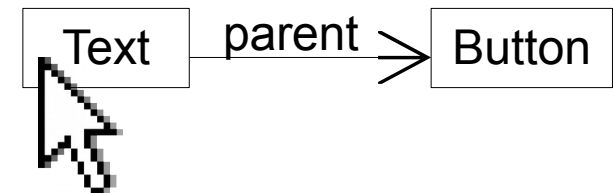


- (1) generalization not necessary

- MenuItem, Button, Tab, Tree, Table, Document, ToolBar, Group, Window, Edit

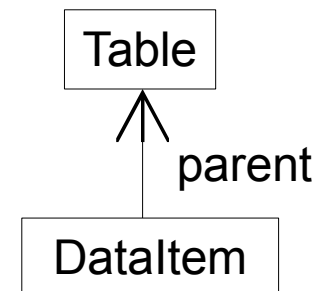
- (2) generalization can bring a benefit

- Custom, Separator, Text, Image, Hyperlink
 - more suitable element in the ancestors
 - MenuItem, Button, Tab, Tree, Table



- (3) items contained in specific containers

- TreeItem → Tree, DataItem → Table, TabItem → Tab



Preprocessing Classification



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- control type of the generalized EOI → classification
- **Structural (S)** → layout, don't visualize information nor call functions
 - Menu, Window, Group, Pane, TitleBar, List, DataGrid, Header, MenuBar, Tab, Table, ToolBar, Tree
- **Semi-Structural-Informative (SI)** → structure, but no container
 - Separator
- **Informative (I)** → visualize information, not alterable
 - Tabltee, Treeltem, Dataltem, Listltem, Text, Hyperlink, Image, Headerltem, ProgressBar, StatusBar, Thumb, ToolTip
- **Semi-Informative-Functional (IF)** → visualize information, manipulate information, alterable
 - Slider, Spinner, RadioButton, CheckBox, Calendar, Edit, Document, ScrollBar, ComboBox
- **Functional (F)** → invoke a function (or subroutine), visualize no information nor structure elements
 - Button, MenuItem (leafs)
 - special case → keyboard shortcut

Preprocessing Transaction Identification



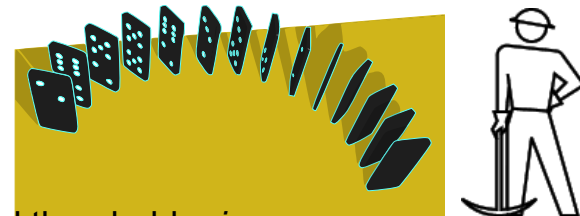
TECHNISCHE
UNIVERSITÄT
DARMSTADT

→ *meaningful clusters found in the log data.*
For creating a sequence database – a set of sequences.

- **(TI1)** Functional Interaction split
 - indicate the end of a task resp. pattern
- **(TI2)** Reference length
 - either a navigation or a content → less time on navigation, more time on content
 - cut-off time → guess of the percentage of navigation
- **(TI3)** Maximal forward reference
 - backward reference → occurred earlier in history
 - two clues → crawl or the generalized EOI
- **(TI4)** Time window
 - in specified time interval

Sequential Pattern Mining (S1)

- interactions occur in sequence → suitable for the problem
- sequential patterns
 - **frequent** sequential patterns → minimal frequency above a specified threshold *minsup*
 - **closed** sequential patterns → not included in another pattern having the same support
 - **maximal** sequential patterns → closed pattern not included in another closed pattern
- maximality eliminates very similar patterns
- sequence database → item is generalized EOI ID
- Vertical mining of Maximal Sequential Patterns (VMSP) [1]



```
49483 -1 64819 -1 -2
49444 -1 49483 -1 64819 -1 -2
49482 -1 82419 -1 -2
49482 -1 82419 -1 -2
66206 -1 49482 -1 -2
49485 -1 129693 -1 49330 -1 -2
```

```
===== Algorithm VMSP - STATISTICS =====
Total time ~ 267 ms
Frequent sequences count : 8
Max memory (mb) : 4.31258
minsup 1
Intersection count 5416
=====
```

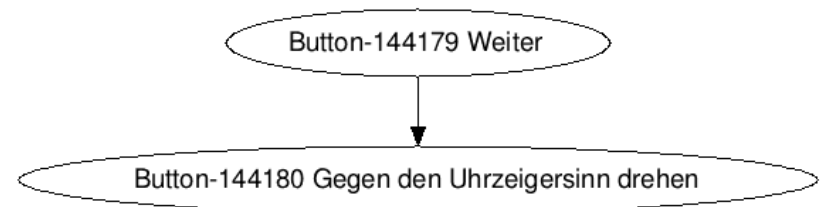
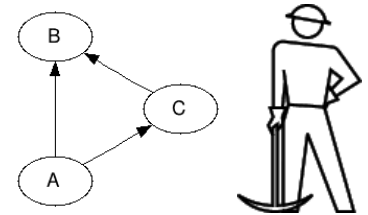
```
82419 -1 SUP: 2
66206 -1 SUP: 3
49485 -1 SUP: 2
49330 -1 SUP: 2
49482 -1 82419 -1 SUP: 2
49483 -1 64819 -1 SUP: 2
```

Graph Mining (S2)



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- discover frequent subgraphs
 - allow modeling branches and loops
- transformation to create a graph database
 - element graph
 - vertices → generalized EOI IDs
 - edge → consecutive interaction
- graph-based Substructure pattern mining (gSpan) [2]

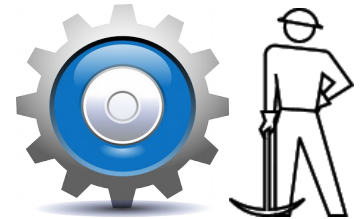


Process Mining (S3)



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- extracts process models from a given event log
 - spaghetti-like process models → abstractions
 - abstractions of processes → patterns
- repeats: *"similar regions (sequence of activities) common within a trace and/or across a set of traces in an event log signifies some set of common functionality accessed by the process"* [3]
 - repeat alphabet → set of activities occurring in the repeat
 - abstraction → maximal elements of partial ordering of repeat alphabets
- interaction log → event log



N-Gram Based Mining (S4)

- problems
 - transaction identification difficult
 - short sequences which reappear in a slightly different way
 - adjustable $n \rightarrow$ uncertainty of the pattern length
- n-gram \rightarrow sequence with n contiguous items from a given sequence
- skip-grams \rightarrow k items can be skipped in between
- functional skip-grams \rightarrow at least one functional interaction
 - Equality
 - functional \rightarrow compared by ID of the (generalized) EOI
 - non-functional \rightarrow compared by classification
 - classification equal \rightarrow compared by control type



0	1
413 (S) Mit der Maus Rechts Klick Tabelle-2637	418 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
465 (S) Mit der Maus Rechts Klick Tabelle-2637	466 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
471 (S) Mit der Maus Rechts Klick Tabelle-2637	472 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
478 (S) Mit der Maus Rechts Klick Tabelle-2637	479 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
480 (S) Mit der Maus Rechts Klick Tabelle-2637	481 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
489 (S) Mit der Maus Rechts Klick Tabelle-2637	490 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
5866 (S) Mit der Maus Rechts Klick Tabelle-2637	5879 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
5919 (S) Mit der Maus Rechts Klick Tabelle-2637	5921 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
6098 (S) Mit der Maus Rechts Klick Tabelle-2637	6102 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
8054 (S) Mit der Maus Rechts Klick Tabelle-2637	8057 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9044 (S) Mit der Maus Rechts Klick Tabelle-2637	9045 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9050 (S) Mit der Maus Rechts Klick Tabelle-2637	9051 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9431 (S) Mit der Maus Rechts Klick Tabelle-2637	9432 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9436 (S) Mit der Maus Rechts Klick Tabelle-2637	9437 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9494 (S) Mit der Maus Rechts Klick Tabelle-2637	9495 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9679 (S) Mit der Maus Rechts Klick Tabelle-2637	9680 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9706 (S) Mit der Maus Rechts Klick Tabelle-2637	9707 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9856 (S) Mit der Maus Rechts Klick Tabelle-2637	9857 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
9892 (S) Mit der Maus Rechts Klick Tabelle-2637	9893 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10049 (S) Mit der Maus Rechts Klick Tabelle-2637	10050 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10088 (S) Mit der Maus Rechts Klick Tabelle-2637	10089 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10430 (S) Mit der Maus Rechts Klick Tabelle-2637	10431 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10432 (S) Mit der Maus Rechts Klick Tabelle-2637	10433 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10456 (S) Mit der Maus Rechts Klick Tabelle-2637	10457 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10632 (S) Mit der Maus Rechts Klick Tabelle-2637	10634 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10648 (S) Mit der Maus Rechts Klick Tabelle-2637	10649 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10656 (S) Mit der Maus Links Klick Tabelle-2637	10658 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
10670 (S) Mit der Maus Rechts Klick Tabelle-2637	10671 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
11317 (S) Mit der Maus Rechts Klick Tabelle-2637	11321 (F) Mit der Maus Links Klick Menüitem-4249 Löschen
11423 (S) Mit der Maus Rechts Klick Tabelle-2637	11442 (F) Mit der Maus Links Klick Menüitem-4249 Löschen

Evaluation Setup



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- 25 promising user-program pairs → most distinct functional interactions
- pattern score $\in \{-3, -2, -1, 0, 1, 2, 3\}$
 - "This recurring interaction pattern describes a task accomplishment"
 - $-3 \rightarrow$ "doesn't apply", $3 \rightarrow$ "applies"
- optional name/description → "How would you name the task accomplishment?
Why did you perform these interactions?"
- special selection policy: clueless → score 0

Dieses immer wiederkehrende Interaktionsmuster beschreibt eine Aufgabenbewältigung?

Trifft nicht zu ☐ ☐ ☐ ☒ ☐ ☐ ☐ Trifft zu

Wie würdest du die Aufgabenbewältigung nennen?
Warum hast du diese Interaktionen ausgeführt?

Evaluation Result

Pattern Quantity



TECHNISCHE
UNIVERSITÄT
DARMSTADT

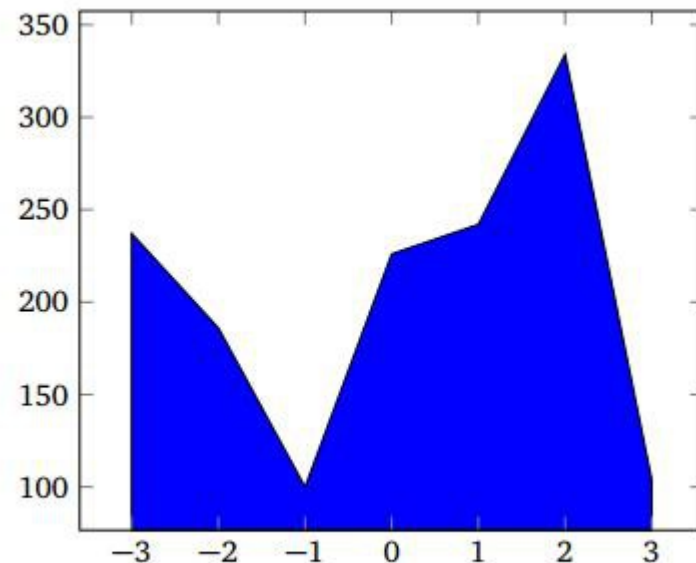
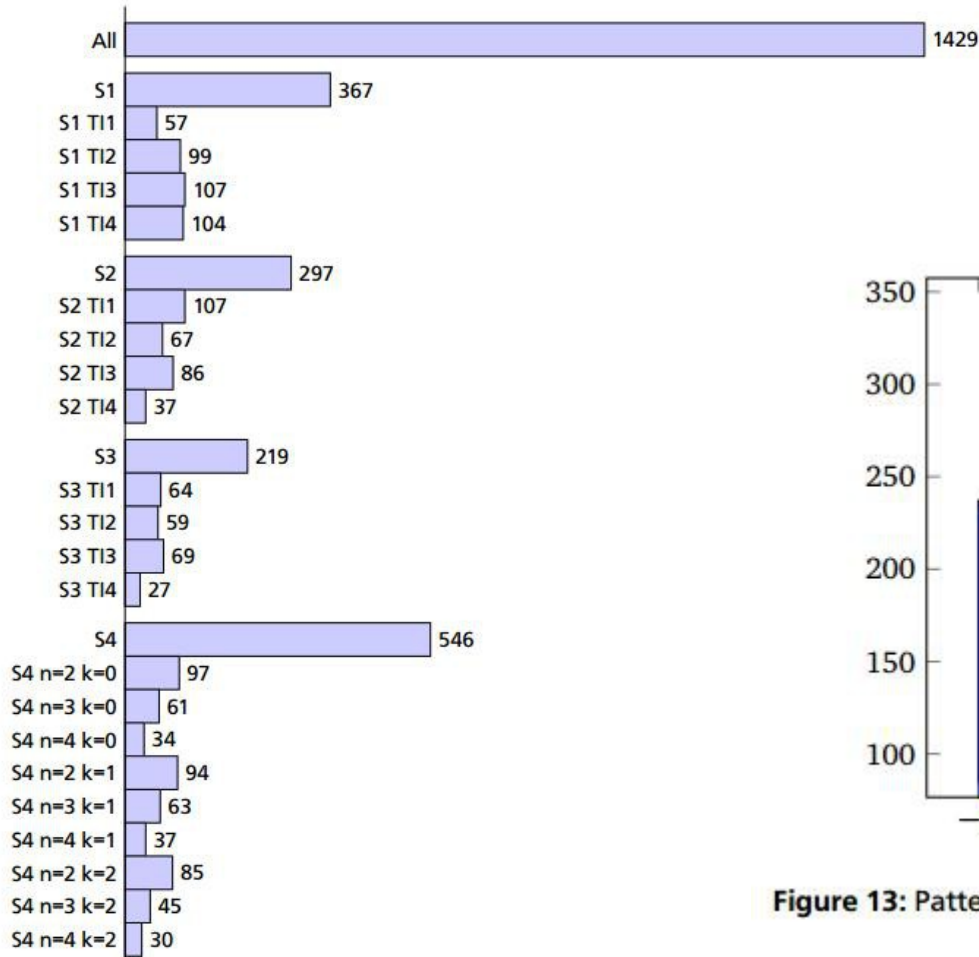


Figure 13: Pattern score chart which shows the quantity of scores

Evaluation Result

Pattern Score



TECHNISCHE
UNIVERSITÄT
DARMSTADT

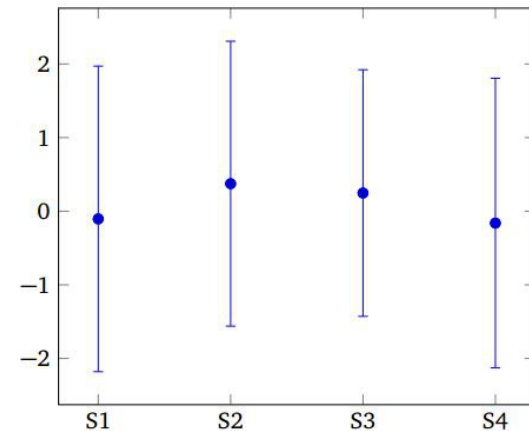
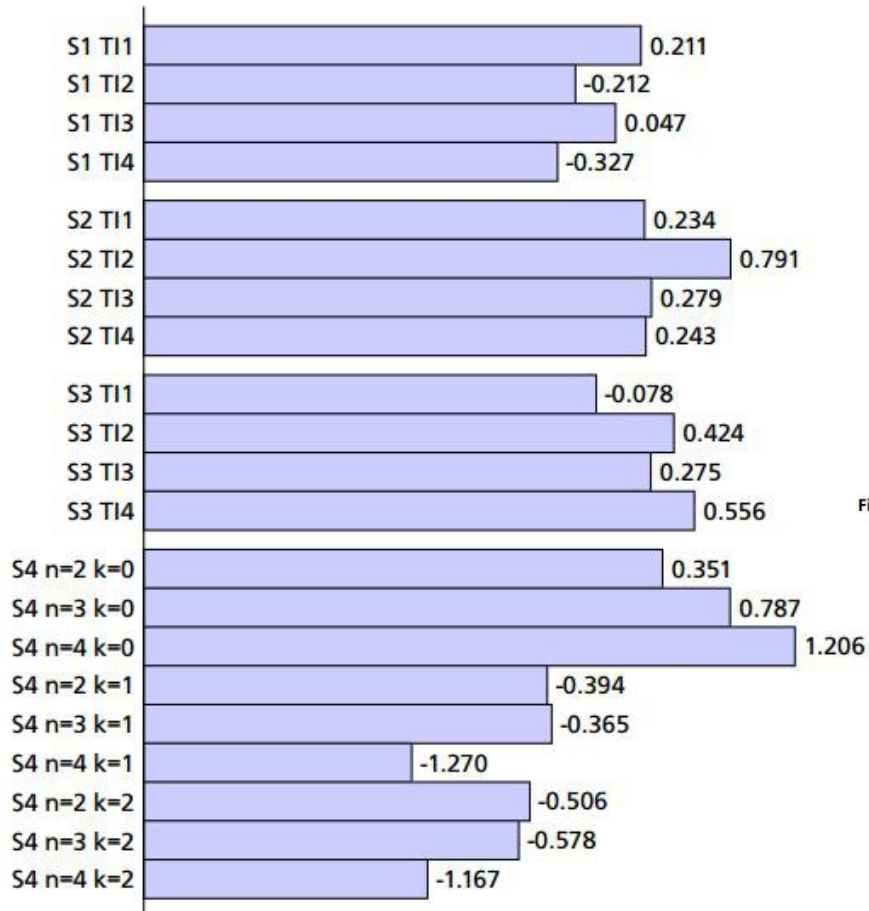
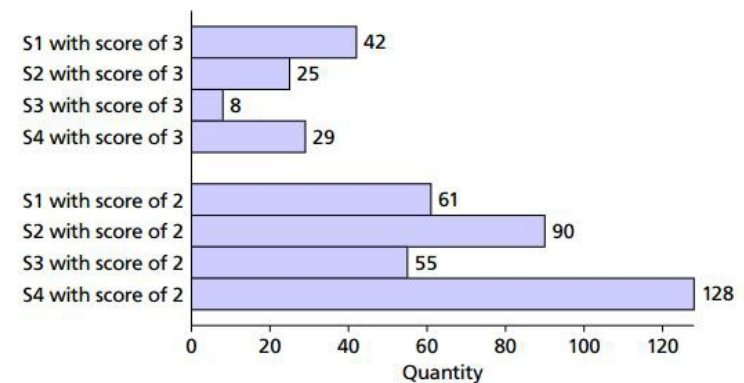


Figure 18: Comparison of the four strategies with an error chart ignoring the transaction identification approaches and n/k values



Evaluation Result

k-skip-n-gram analysis



TECHNISCHE
UNIVERSITÄT
DARMSTADT

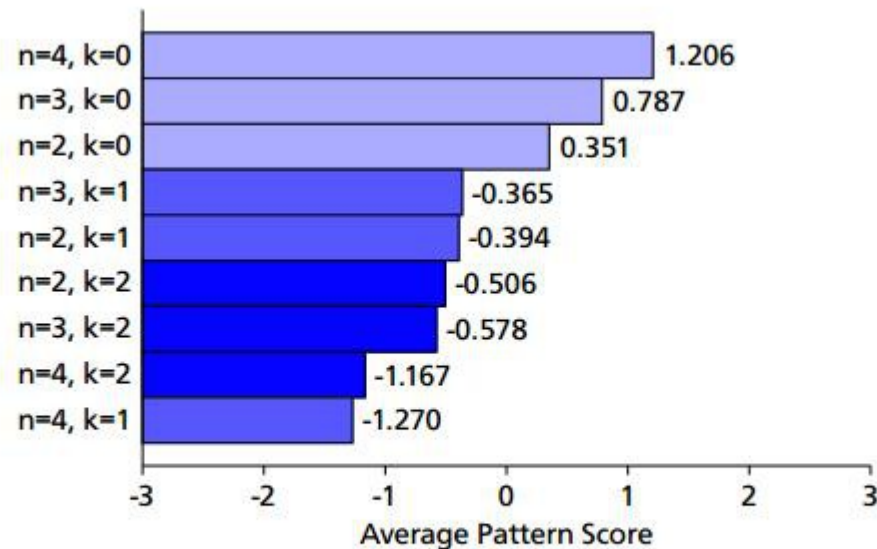


Figure 14: Ordered average pattern score depending on n and k parameter of the k -skip- n -gram strategy (S4)

Conclusion

- Patterns good abstraction
 - too complex → sequential & time ordered
 - simpler associations
 - reference patterns good starting point
 - only 11 reference patterns of 3 annotators
 - 1:1 relationship
 - n:1 relationships more interesting
 - generalized EOI
 - matching errors → alienate generalized interaction log
 - transaction identification
 - discovering meaningful segments is actually the same as discovering patterns
 - transactions malformed → discovery algorithm fails
 - too few data points are used in the evaluation
 - no significance could be argued
 - no strategy returned only acceptable patterns
- **bottom line**
 - patterns exists & 104 very acceptable could be discovered
 - best strategy unclear → insufficient data points
 - however, first steps made

Thank you ...



TECHNISCHE
UNIVERSITÄT
DARMSTADT

... for paying attention



Questions?

References



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- [1] Philippe Fournier-Viger et al. “VMSP: Efficient Vertical Mining of Maximal Sequential Patterns”. In: Advances in Artificial Intelligence. Vol. 8436. 2014, pp. 83–94.
- [2] Xifeng Yan and Jiawei Han. “gSpan: Graph-Based Substructure Pattern Mining”. In: 2002 IEEE International Conference on Data Mining, 2002. Proceedings. (2002).
- [3] R.P. Jagadeesh Chandra Bose and W.M.P. van der Aalst. “Abstractions in process mining: A taxonomy of patterns”. In: Business Process Management (2009).
- [4] Microsoft Developer Network. AccessibleObjectFromPoint function. URL: <http://msdn.microsoft.com/en-us/library/windows/desktop/dd317977%28v=vs.85%29.aspx> (visited on 01/01/2015).
- [5] Microsoft Developer Network. IUIAutomation::ElementFromPoint method. URL: <http://msdn.microsoft.com/enus/library/windows/desktop/ee671538%28v=vs.85%29.aspx> (visited on 01/01/2015).
- [6] Microsoft Developer Network. Automation Element Property Identifiers. URL: <http://msdn.microsoft.com/enus/library/windows/desktop/ee684017%28v=vs.85%29.aspx> (visited on 01/01/2015).
- [7] Microsoft Developer Network. Accessibility. URL: <http://msdn.microsoft.com/en-us/library/ms753388%28v=vs.110%29.aspx> (visited on 01/01/2015).
- [8] Microsoft Developer Network. UI Automation and Microsoft Active Accessibility. URL: <http://msdn.microsoft.com/en-us/library/ms788733%28v=vs.110%29.aspx> (visited on 01/01/2015).

Image References



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- <http://www.pdclipart.org/>
- <https://openclipart.org/>