8.1   Materials

8.1.1   Process Log Questions

The three process logs are selected based on the availability of scientific literature discussing process insights and the technical compatibility with the context size of current Large Language Models (LLMs).

The questions are extracted in part from what is answered in case studies (Mannhardt & Blinde, 2016; Mannhardt et al., 2016; Mannhardt, 2018) and what is applicable from the questions available in the Process minin Question Forge (Zimmermann, 2024).

Road Traffic Fine Management

The process log describes the handling of road traffic fines by a local police force in Italy. It contains more than 150000 traces, each representing the life cycle of a road traffic fine. The events describe notifications, payments, and appeals (de Leoni & Mannhardt, 2015).

What does the log contain? The process log contains 150370 traces that are composed of 11 unique activities (Mannhardt et al., 2018, p. 262). All traces start with the event *Create Fine*. The majority of traces end with the event *Payment*, occurring a total of 67201 times. The process log includes 10 attributes where the *amount* field indicates the amount due to be paid for the fine and the attribute *paymentAmount* indicates the amount paid in one transaction (Mannhardt et al., 2018, p. 263). The process log covers a period of over 13 years, from January 2000 to June 2013 (Mannhardt et al., 2018, p. 262).

What is the process? At the beginning of the process, the *Create Fine* activity is executed, after which either the payment is made by the offender, or the notification is sent. In some cases, a penalty is added, or the fine is appealed. Finally, if a fine is not paid in full, the activity *Send for Credit Collection* is executed (Mannhardt et al., 2018, pp. 264, 283, 285). The Petri-net discovered by applying a standard inducive miner with a noise threshold of 0.5 to the process log (Figure 8.1) provides a process abstraction.
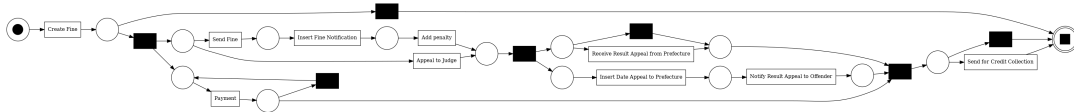

Figure 8.1: Petri Net discovered by standard inducive for
the road traffic fine management process log.

What variants are there in the process? The most common variant is *Create Fine > Send Fine > Insert Fine Notification > Add penalty > Send for Credit Collection* with 56482 occurrences which accounts for 37.56% of all traces. In 30.84% of all cases, the variant is *Create Fine > Payment*. The third most common variant with 20385 cases is *Create Fine > Send Fine* which accounts for 13.56% of the process log.

Where do process instances deviate from each other? Comparing the top two variants, the variant *Create Fine > Send Fine > Insert Fine Notification > Add penalty > Send for Credit Collection* coincides with an average fine of 73.8 and has an average trace duration of

689 days, whereas in contrast the variant *Create Fine > Payment* has an average fine of 37.4 and an average trace duration of 10.3 days.

What is the percentage of fines that are fully paid? In total, 46826, or 31.1% of all traces, were paid the full fine amount, including added penalties. 59013 or 39.2% of all traces contain a *Send for Credit Collection* event.

To what extent does the recorded process follow the regulatory requirements? Italian law constitutes a requirement that a notification for the fine has to be sent within 90 days (Mannhardt et al., 2018, p. 265). In total, 79860 traces contain a *Insert Fine Notification* of which 32267 (40.4%) are sent within 90 days after the *Create Fine* event.

What is the state of processes that ended in non-payment? Excluding all traces that contain at least one *Payment*, *Send for Credit Collection*, *Appeal to Judge* or *Send Appeal to Prefecture*, 20389 traces are in an open state. This supports the assumption that there may be issues with managing unpaid fines (Mannhardt et al., 2018, p. 267).

How often is the driver also the owner when being fined? The attribute *notificationType* indicates if the driver is the owner of the car (P) or not (C) (Mannhardt et al., 2018, p. 263). In total, 79860 traces contain a *Insert Fine Notification* event, of which 134 cases are sent to drivers who were not the car owners.

Sepsis

The process log describes the trajectories of 1050 sepsis patients who were admitted to the emergency room. The events it contains describe the patients' registration, triage, laboratory tests, and discharge.

What does the log contain? The process log contains 1050 traces that are composed of 16 unique activities (Mannhardt et al., 2018, p. 292). 995 traces start with the event *ER Registration*. 393 traces end with the event *Release A* and 291 of the traces end with the event *Return ER*. The process log includes 29 attributes where the *Age* field indicates the patient's age and the attribute *InfectionSuspected* indicates whether an infection is suspected (Mannhardt et al., 2018, pp. 292, 293). The process log covers a period of almost 2 years, from November 2013 to June 2015.

What is the process? At the beginning of the process (995 times out of 1050), the *ER Registration* activity is executed, after which the *ER Treiage* and *ER Sepsis Triage* are recorded. In some cases, measurements are conducted before the registration is completed. Finally, one of the release events is executed, after which, in some cases, a *Return ER* can occur. (Mannhardt et al., 2018, p. 295). The Petri-net discovered by applying a standard inducive miner with a noise threshold of 0.5 to the process log (Figure 8.2) provides a process abstraction.

What variants are there in the process? The most common variant is *ER Registration > ER Triage > ER Sepsis Triage* with 35 occurrences which accounts for 2.86% of all traces. Excluding measurement activities, the variants *ER Registration > ER Triage > ER Sepsis Triage > Release A* and *ER Registration > ER Triage > ER Sepsis Triage > Release A > Return ER* emerge as the most frequent, having 388 and 267 occurrences, respectively which combined account for 62.4% of the process log.

Where do process instances deviate from each other? 800 traces contain an *Admission NC* event, 110 traces contain an *Admission IC* event, and in 140 cases, no admission
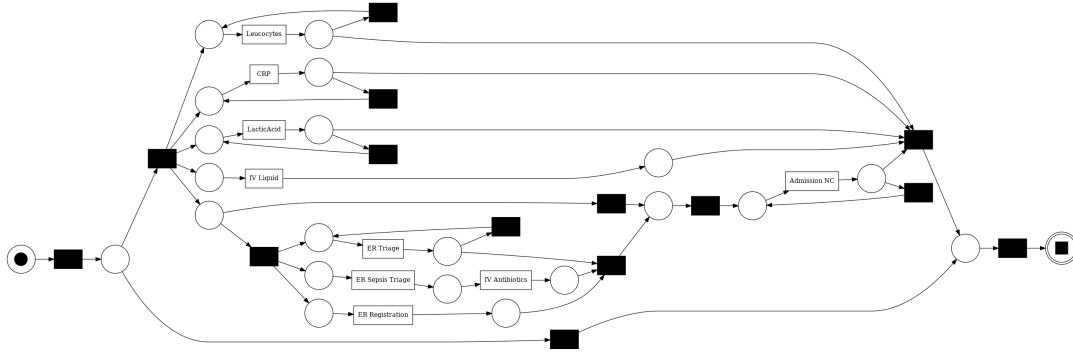
Figure 8.2: Petri Net discovered by standard inducive for
the sepsis process log.

event is recorded. In total, 10 traces the *Admission IC* event without the occurrence
of an *Admission NC* event. Overall, 294 of all traces contain an *Return ER* event
(Mannhardt & Blinde, 2016, p. 78).

What can be said about the ward admission? 97 traces (9.2%) describe first an *Admission
IC* and afterwards an *Admission NC* event. In 39 traces (3.7%) the *Admission NC* event
comes first and the *Admission IC* event second. On average the patients age is 66.75
years for *Admission IC > Admission NC* cases and 65.26 years for *Admission NC >
Admission IC* cases.

Has the delay in antibiotic administration an effect if patients return to the ER? In total,
it contains 823 traces of an *ER Sepsis Triage* as well as an *IV Antibiotics* event, of
which 342 cases conform to the one-hour rule. For all traces that indicate an ER
readmission, the mean time for antibiotics administration is 1.84 hours and a median
time of 1.75 hours, where as traces without an ER readmission indicate an average
time of 1.86 hours and a median time of 1.48 hours.

Are patients administered antibiotics within one hour? Antibiotics are, on average, ad-
ministered within 1.7 hours after the *ER Sepsis Triage* event. Only in 41.5% of the
cases is the process conforming to the one-hour rule (Mannhardt & Blinde, 2016,
p. 76; Mannhardt et al., 2018, pp. 296, 308).

Are patients administered lactic acid measurements within three hours? The lactic acid
measurements are on average, performed within 1.7 hours. In total 99.3% of all
traces are conformant with this rule (Mannhardt & Blinde, 2016, p. 76; Mannhardt
et al., 2018, p. 296)

Hospital Billing

The process log contains 100000 billing-related cases from a hospital in the Nether-
lands. The events describe a billing package's process, including creation, change,
billing, and storno.

What does the log contain? The process log contains 100000 traces that are composed
of 17 unique activities (Mannhardt et al., 2018, p. 327). All traces start with the event
NEW. 63498 traces end with the event *BILLED* and 8215 of the traces end with the
event *DELETE*. The process log includes 18 attributes where the *caseType* the type of
the billing package and the attribute *diagnosis* indicates the diagnosis being treated

(Mannhardt et al., 2018, p. 327). The process log covers a period of 3 years, from December 2012 to January 2016.

What is the process? At the beginning of the process, the *NEW* activity is executed. In 22373 (22.3%), *NEW* is the only event recorded. In some cases the diagnosis is changed after the billing package is created. Moreover, some traces show that packages can be reopened, canceled or rejected after the *RELEASE* event (Mannhardt et al., 2018, p. 328). The Petri-net discovered by applying a standard inducive miner with a noise threshold of 0.5 to the process log (Figure 8.3) provides a process abstraction.
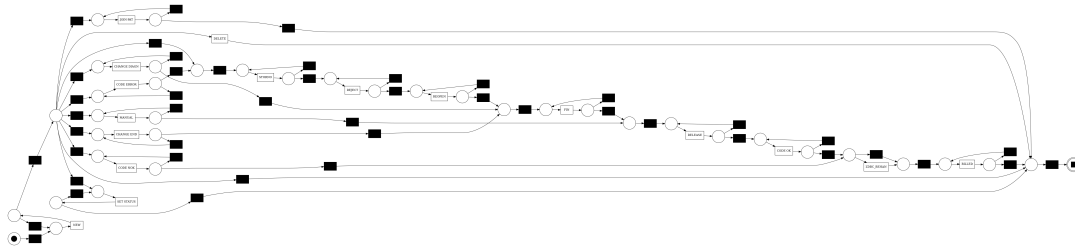
Figure 8.3: Petri Net discovered by standard inducive for
the hospital billing process log.

What variants are there in the process? The most common variant is *NEW > CHANGE DIAGN > FIN > RELEASE > CODE OK > BILLED* with 33673 occurrences which accounts for 33.7% of all traces. The variant *NEW, NEW > FIN NEW > DELETE* combined account for 30694 traces (30.7%) of the process log.

Where do process instances deviate from each other? 8221 traces (8.2%) contain a *DELETE* event of which 331 have been also billed. 2548 traces (2.5%) contain a *STORNO* event of which 2547 have been also billed. In 42% of the process instances (42314 traces), a *CHANGE DIAGN* event is present.

How many cases are never billed? In total, 35146 (35.1%) of all traces do not contain a *BILLED* activity, of which 32905 traces do also not contain a *RELEASE* activity. The process log contains 692 traces that have been released but not billed.

What cases are deleted? 4813 traces are created and directly after deleted. The median duration is 9.1 days with a average duration of 51.1 days. Comparing the deletion rate based on the attribute *speciality*, traces with the value E (1299 deleted traces out of 7628: 17%) and with the value M (1060 deleted traces out of 9153: 11%) are more likely to be deleted as traces with the value L (880 deleted traces out of 9477: 9.3%) or with the value D (764 deleted traces out of 7654: 9.9%).

How often are cases manually edited? There are, in total, 243 traces that contain a *MANUAL* activity.

For what cases is the diagnosis changed after the billing package is created? In 42314 traces, there is a *CHANGE DIAGN* activity present. Traces with the *caseType B* are more likely to have the diagnosis changed (Mannhardt, 2018, p. 330). The process log contains 40568 out of 42501 traces of *caseType B* that also have the activity *CHANGE DIAGN* present. For traces that display *caseType A*, 1489 out of 32681 traces have the diagnosis change.

8.1.2 Process Logs Data

Table 8.1: Road traffic fine management data sample.
Column name abreventions: nT: notificationType, tPA: totalPaymentAmount, vC: vehicleClass

amount	org:resource	dismissal	concept:name	vC	tPA	lifecycle:transition	time:timestamp	article	points	case:concept:name	expense	nT	lastSent	paymentAmount	matricola
35.0	561	NIL	Create Fine	A	0.0	complete	2006-07-24 00:00:00+00:00	157.0	0.0	A1					
			Send Fine			complete	2006-12-05 00:00:00+00:00			A1	11.0				
35.0	561	NIL	Create Fine	A	0.0	complete	2006-08-02 00:00:00+00:00	157.0	0.0	A100					
			Send Fine			complete	2006-12-12 00:00:00+00:00			A100	11.0				
			Insert Fine Notification			complete	2007-01-15 00:00:00+00:00			A100		P	P		
71.5			Add penalty			complete	2007-03-16 00:00:00+00:00			A100					
			Send for Credit Collection			complete	2009-03-30 00:00:00+00:00			A100					
36.0	561	NIL	Create Fine	A	0.0	complete	2007-03-09 00:00:00+00:00	157.0	0.0	A10000					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10000	13.0				
			Insert Fine Notification			complete	2007-08-02 00:00:00+00:00			A10000		P	P		
74.0			Add penalty			complete	2007-10-01 00:00:00+00:00			A10000					
			Payment		87.0	complete	2008-09-09 00:00:00+00:00			A10000				87.0	
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-19 00:00:00+00:00	157.0	0.0	A10001					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10001	13.0				
			Insert Fine Notification			complete	2007-07-25 00:00:00+00:00			A10001		P	P		
			Insert Date Appeal to Prefecture			complete	2007-08-02 00:00:00+00:00			A10001					
74.0			Add penalty			complete	2007-09-23 00:00:00+00:00			A10001					
		#	Send Appeal to Prefecture			complete	2007-09-24 00:00:00+00:00			A10001					
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	157.0	0.0	A10004					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10004	13.0				
			Insert Fine Notification			complete	2007-07-24 00:00:00+00:00			A10004		P	P		
74.0			Add penalty			complete	2007-09-22 00:00:00+00:00			A10004					
			Send for Credit Collection			complete	2009-03-30 00:00:00+00:00			A10004					
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	157.0	0.0	A10005					
			Payment		36.0	complete	2007-03-21 00:00:00+00:00			A10005				36.0	
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	157.0	0.0	A10007					
			Payment		36.0	complete	2007-03-23 00:00:00+00:00			A10007				36.0	
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	157.0	0.0	A10008					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10008	13.0				
			Insert Fine Notification			complete	2007-08-02 00:00:00+00:00			A10008		P	P		
74.0			Add penalty			complete	2007-10-01 00:00:00+00:00			A10008					
			Send for Credit Collection			complete	2009-03-30 00:00:00+00:00			A10008					
22.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	7.0	0.0	A10009					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10009	13.0				
			Insert Fine Notification			complete	2007-07-23 00:00:00+00:00			A10009		P	P		
44.0			Add penalty			complete	2007-09-21 00:00:00+00:00			A10009					
			Payment		35.0	complete	2007-10-01 00:00:00+00:00			A10009				35.0	
			Payment		57.0	complete	2007-10-31 00:00:00+00:00			A10009				22.0	
21.0	550	NIL	Create Fine	A	0.0	complete	2006-08-02 00:00:00+00:00	7.0	0.0	A1001					
			Send Fine			complete	2006-12-12 00:00:00+00:00			A1001	11.0				
			Insert Fine Notification			complete	2006-12-31 00:00:00+00:00			A1001		P	P		
42.5			Add penalty			complete	2007-03-01 00:00:00+00:00			A1001					
			Send for Credit Collection			complete	2009-03-30 00:00:00+00:00			A1001					
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-20 00:00:00+00:00	157.0	0.0	A10010					
			Send Fine			complete	2007-07-17 00:00:00+00:00			A10010	13.0				
			Insert Fine Notification			complete	2007-07-23 00:00:00+00:00			A10010		P	P		
74.0			Add penalty			complete	2007-09-21 00:00:00+00:00			A10010					
			Send for Credit Collection			complete	2009-03-30 00:00:00+00:00			A10010					
36.0	537	NIL	Create Fine	A	0.0	complete	2007-03-21 00:00:00+00:00	157.0	0.0	A10011					
			Send Fine			complete	2007-09-29 00:00:00+00:00			A10011	18.6				

Table 8.2: Sepsis data sample.
Diagnostic* and SIRS* columns are omitted

InfectionSuspected	org:group	DisfuncOrg	Hypotensie	Infusion	concept:name	Age	time:timestamp	Oligurie	Diagnose	Hypoxie	case:concept:name	Leucocytes	CRP	LacticAcid
True	A	True	True	True	ER Registration	85.0	2014-10-22 11:15:41+00:00	False	A	False	A			
	B				Leucocytes		2014-10-22 11:27:00+00:00				A	9.6		
	B				CRP		2014-10-22 11:27:00+00:00				A		21.0	
	B				LacticAcid		2014-10-22 11:27:00+00:00				A			2.2
	C				ER Triage		2014-10-22 11:33:37+00:00				A			
	A				ER Sepsis Triage		2014-10-22 11:34:00+00:00				A			
	A				IV Liquid		2014-10-22 14:03:47+00:00				A			
	A				IV Antibiotics		2014-10-22 14:03:47+00:00				A			
	D				Admission NC		2014-10-22 14:13:19+00:00				A			
	B				CRP		2014-10-24 09:00:00+00:00				A		109.0	
	B				Leucocytes		2014-10-24 09:00:00+00:00				A	8.7		
	B				Leucocytes		2014-10-26 08:00:00+00:00				A	9.6		
	B				CRP		2014-10-26 08:00:00+00:00				A		47.0	
	B				Leucocytes		2014-10-28 08:00:00+00:00				A	10.7		
	B				CRP		2014-10-28 08:00:00+00:00				A		15.0	
	B				CRP		2014-10-30 08:00:00+00:00				A		9.0	
	B				Leucocytes		2014-10-30 08:00:00+00:00				A	13.0		
	B				Leucocytes		2014-10-31 08:00:00+00:00				A	11.3		
	B				CRP		2014-10-31 08:00:00+00:00				A		9.0	
	B				CRP		2014-11-02 08:00:00+00:00				A		6.0	
	B				Leucocytes		2014-11-02 08:00:00+00:00				A	10.9		
	E				Release A		2014-11-02 15:15:00+00:00				A			
True	A	False	False	True	ER Registration	45.0	2014-12-21 11:04:24+00:00	False	B	False	B			
	C				ER Triage		2014-12-21 11:17:19+00:00				B			
	B				CRP		2014-12-21 11:36:00+00:00				B		240.0	
	B				LacticAcid		2014-12-21 11:36:00+00:00				B			0.8
	B				Leucocytes		2014-12-21 11:36:00+00:00				B	13.8		
	A				ER Sepsis Triage		2014-12-21 12:15:45+00:00				B			
	A				IV Liquid		2014-12-21 12:33:48+00:00				B			
	A				IV Antibiotics		2014-12-21 12:33:55+00:00				B			
	F				Admission NC		2014-12-21 13:17:08+00:00				B			
	B				CRP		2014-12-24 08:00:00+00:00				B		226.0	
	B				CRP		2014-12-26 14:00:00+00:00				B		75.0	
	E				Release A		2014-12-26 18:00:00+00:00				B			

Table 8.3: Hospital billing data sample.
act*, flag* and lifecycle:transition columns are omitted

isCancelled	diagnosis	time:timestamp	caseType	speciality	org:resource	concept:name	blocked	isClosed	state	case:concept:name	closeCode	msgCount	version	msgType	msgCode
False	A	2012-12-16 19:33:10+00:00	A	A	ResA	NEW	False	True	In progress	A					
		2013-12-15 19:00:37+00:00				FIN			Closed	A	A				
		2013-12-16 03:53:38+00:00				RELEASE			Released	A					
		2013-12-17 12:56:29+00:00				CODE OK				A		0.0	A		
		2013-12-19 03:44:31+00:00			ResB	BILLED			Billed	A					
False	B	2012-12-16 19:33:50+00:00	A	A	ResA	NEW	False	False	In progress	B					
True		2013-10-19 12:37:05+00:00			ResC	DELETE			In progress	B					
False	C	2013-01-13 21:04:24+00:00	A	B	ResA	NEW	False	True	In progress	C					
		2013-04-17 19:59:43+00:00				FIN			Closed	C	B				
		2013-04-18 02:30:35+00:00				RELEASE			Released	C					
		2013-04-21 01:55:12+00:00				CODE OK				C		0.0	B		
		2013-05-01 14:41:32+00:00			ResD	REOPEN			In progress	C					
	D	2013-05-01 14:41:54+00:00				CHANGE DIAGN			In progress	C					
		2013-05-02 20:33:39+00:00			ResA	FIN			Closed	C					
		2013-05-03 04:03:17+00:00				RELEASE			Released	C					
		2013-05-03 05:21:24+00:00				CODE OK				C					
		2013-05-23 07:32:15+00:00			ResB	BILLED			Billed	C					
False	E	2013-01-16 21:05:37+00:00	A	B	ResA	NEW	False	True	In progress	D					
		2014-01-15 20:15:25+00:00				FIN			Closed	D	A				
		2014-01-16 03:50:37+00:00				RELEASE			Released	D					
		2014-01-22 03:54:27+00:00				CODE OK				D		0.0	A		
		2014-02-09 21:56:28+00:00			ResB	BILLED			Billed	D					
False		2013-01-18 09:07:53+00:00	B	C	ResE	NEW	False	True	In progress	E					
	F	2013-01-18 09:08:04+00:00				CHANGE DIAGN			In progress	E					
		2013-04-21 20:50:27+00:00			ResA	FIN			Closed	E	C				
		2013-04-22 02:41:46+00:00				RELEASE			Released	E					
		2013-04-24 02:35:19+00:00				CODE OK				E		0.0	B		
		2013-05-21 09:50:35+00:00			ResB	BILLED			Billed	E					
False		2013-01-10 08:41:37+00:00	B	D	ResF	NEW	False	True	In progress	F					
	G	2013-01-26 08:31:57+00:00			ResG	CHANGE DIAGN			In progress	F					
		2013-03-12 20:12:03+00:00			ResA	FIN			Closed	F	B				
		2013-03-12 23:34:59+00:00				RELEASE			Released	F					
		2013-03-15 03:27:41+00:00				CODE OK				F		0.0	C		
		2013-08-04 22:46:17+00:00			ResB	BILLED			Billed	F					
False		2012-12-18 09:07:50+00:00	B	A	ResC	NEW	False	True	In progress	G					
	H	2012-12-18 09:08:40+00:00				CHANGE DIAGN			In progress	G					
		2013-02-09 19:24:43+00:00			ResA	FIN			Closed	G	B				
		2013-02-10 01:16:23+00:00				RELEASE			Released	G					
		2013-02-13 01:45:46+00:00				CODE OK				G		0.0	C		
		2013-07-12 22:16:45+00:00			ResB	BILLED			Billed	G					
False		2013-01-22 09:35:15+00:00	B	D	ResH	NEW	False	True	In progress	H					
	I	2013-01-22 09:35:34+00:00				CHANGE DIAGN			In progress	H					
		2013-04-23 20:00:40+00:00			ResA	FIN			Closed	H	C				
		2013-04-24 00:24:39+00:00				RELEASE			Released	H					
		2013-04-28 01:44:04+00:00				CODE OK				H		0.0	B		