Automatic Generation of Bayesian networks in Forensic Science

Presentation · September 2017

DOI: 10.13140/RG.2.2.17798.47689

2 authors:

Jacob De Zoete
Queen Mary, University of London
18 PUBLICATIONS **26** CITATIONS
SEE PROFILE

Norman Elliott Fenton
Queen Mary, University of London
248 PUBLICATIONS **11,924** CITATIONS
SEE PROFILE

Automatic Generation of Bayesian networks in Forensic Science

Jacob de Zoete, Norman Fenton

Queen Mary University of London

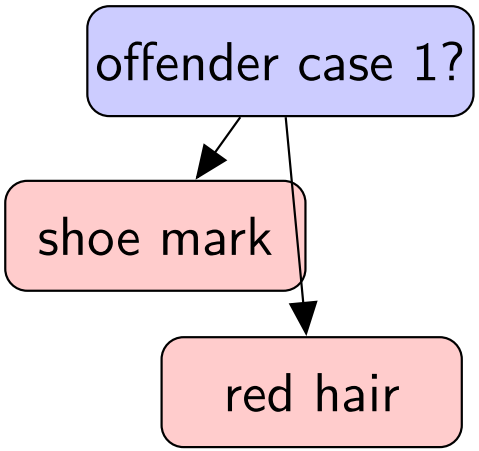
September 6, 2017

Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

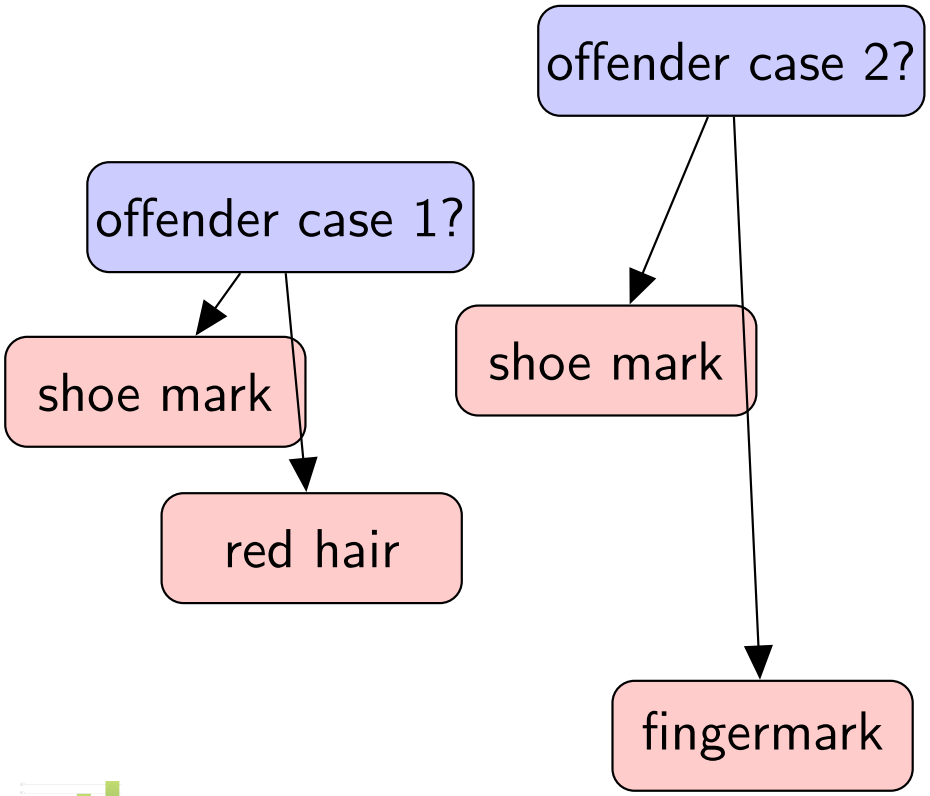
Goal

- Use the computer to generate BN structures.
- $\Rightarrow$ Example for crime linkage situations.

2 crimes

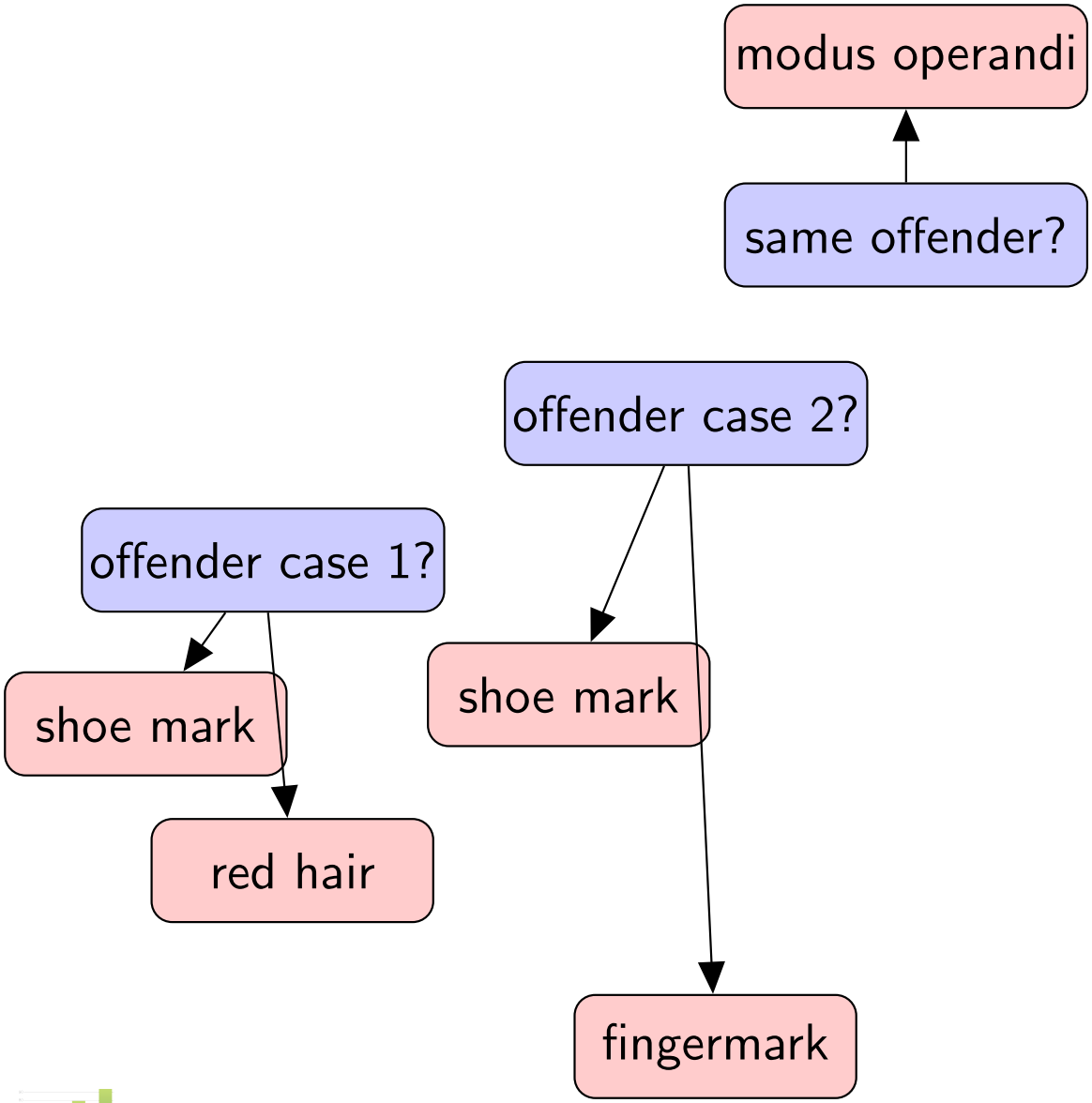


2 crimes

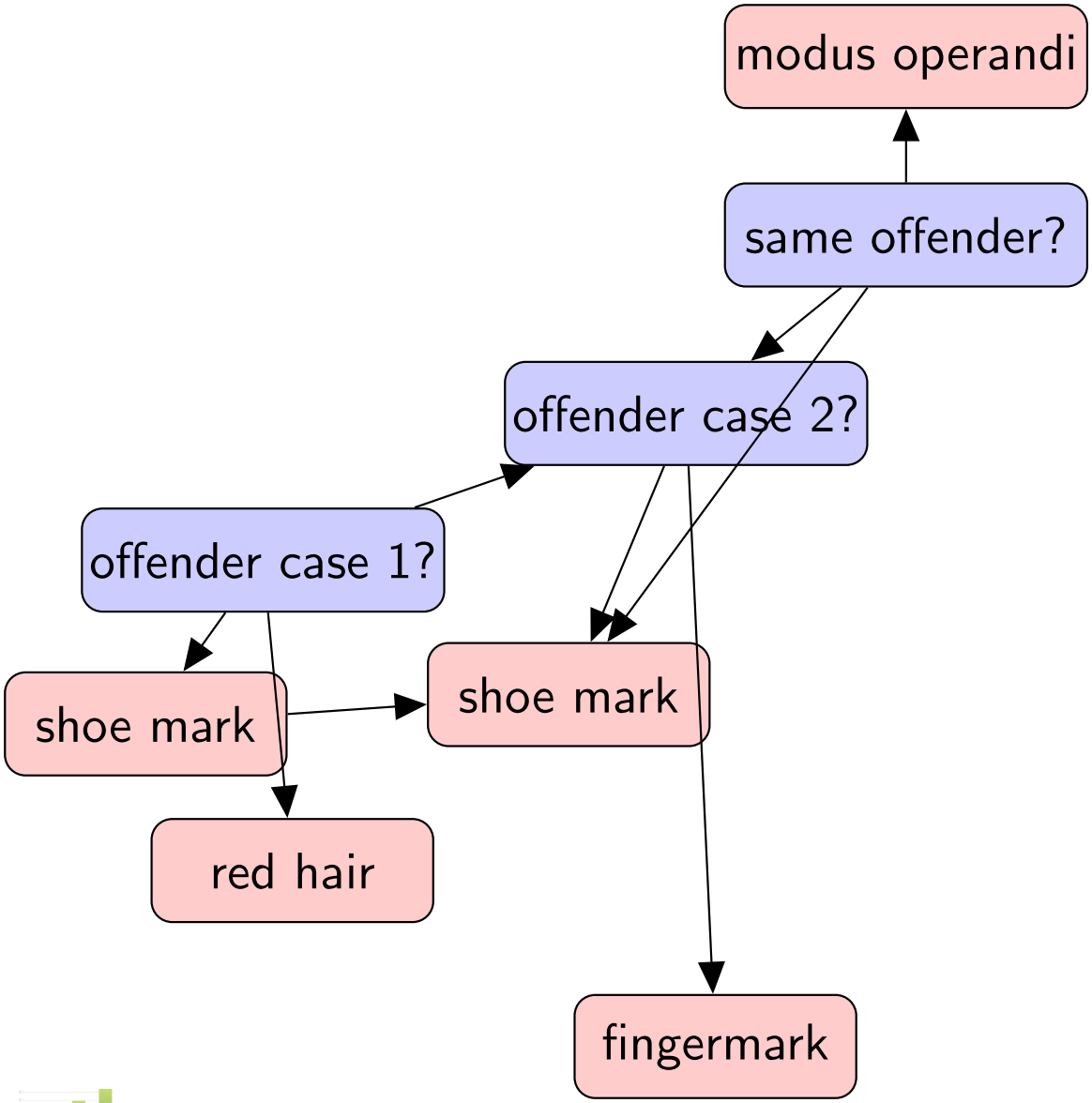


Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

2 crimes

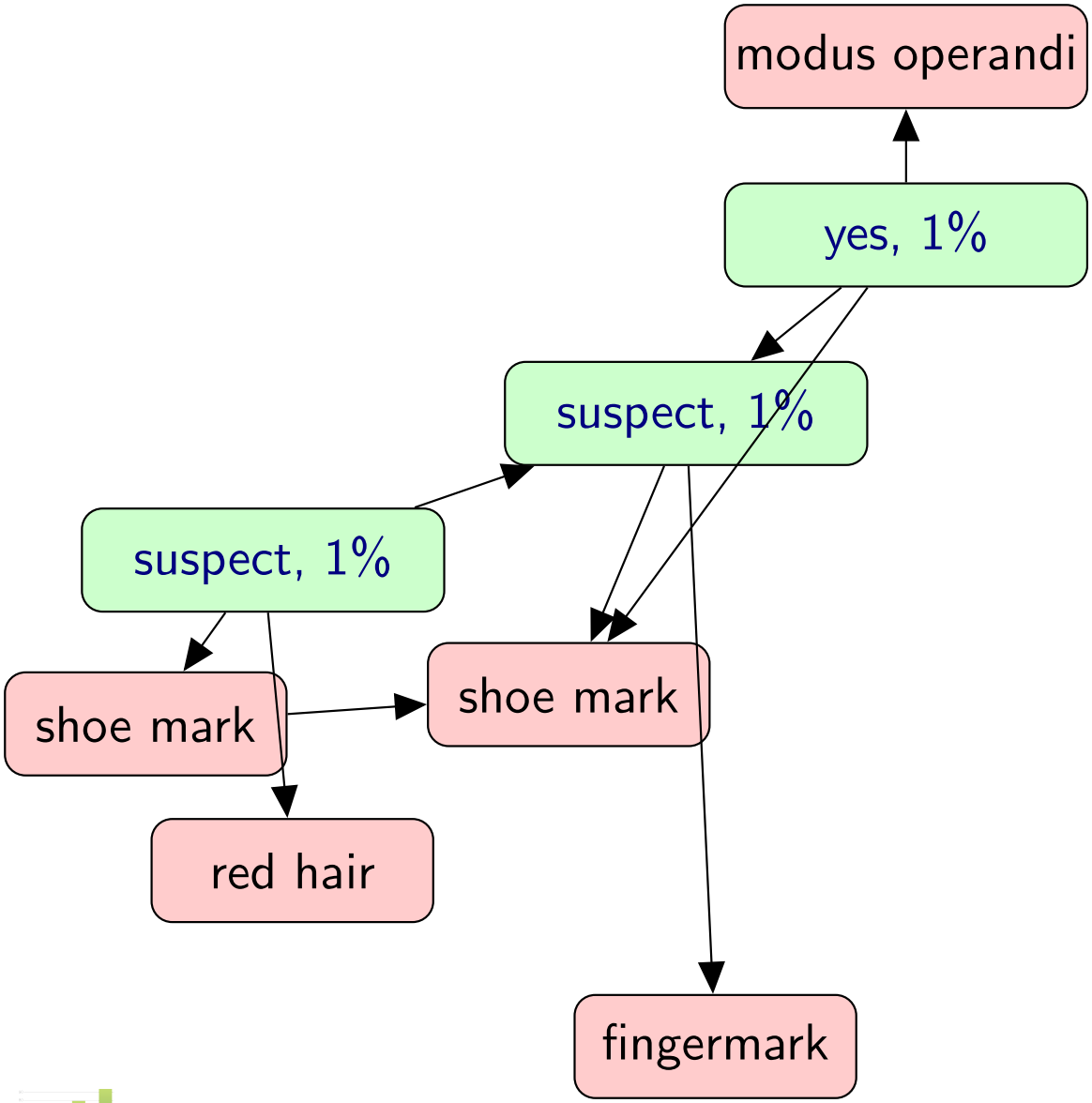

modus operandi
same offender?
offender case 2?
offender case 1?
shoe mark
shoe mark
red hair
fingermark

2 crimes

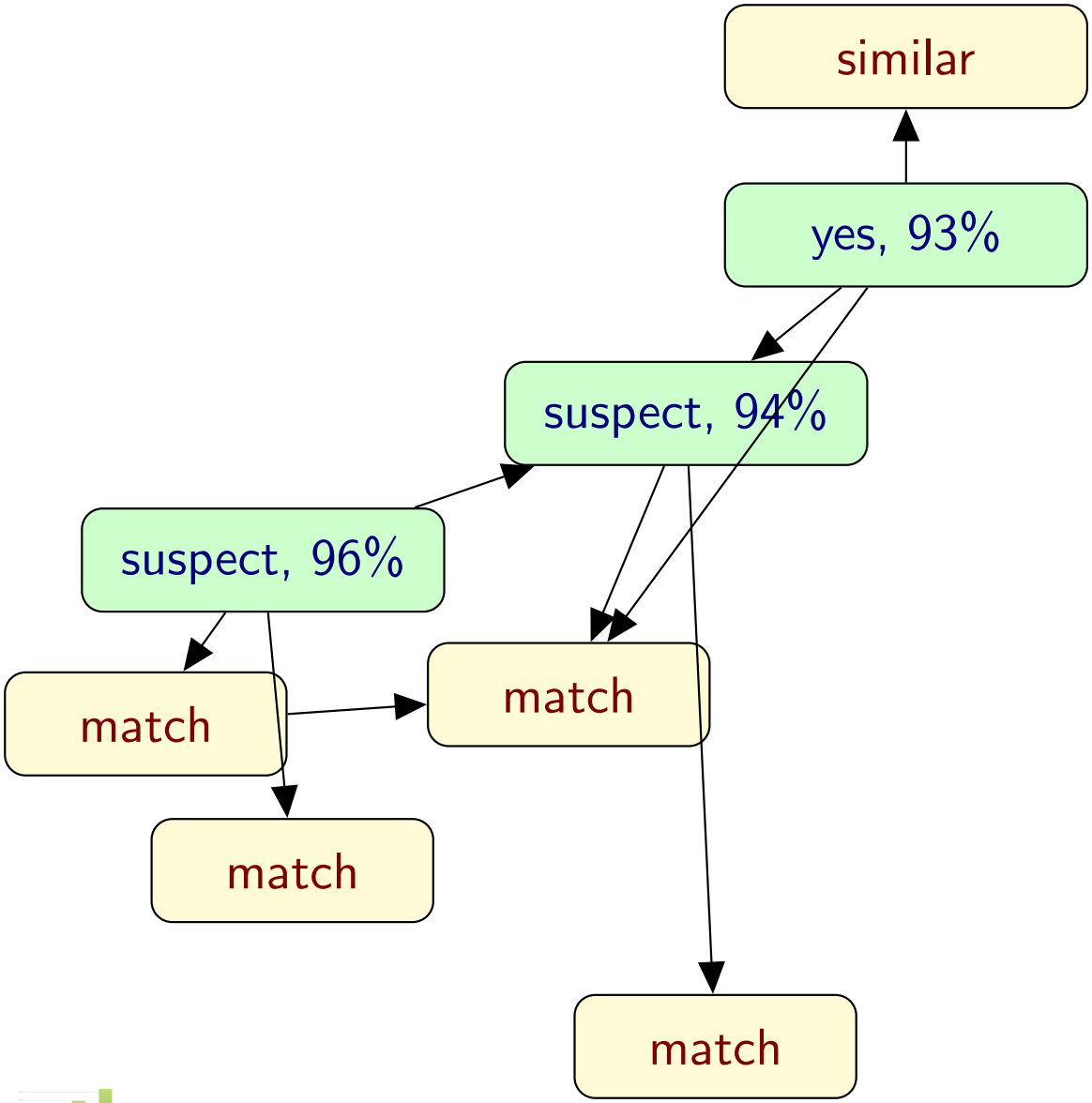


Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

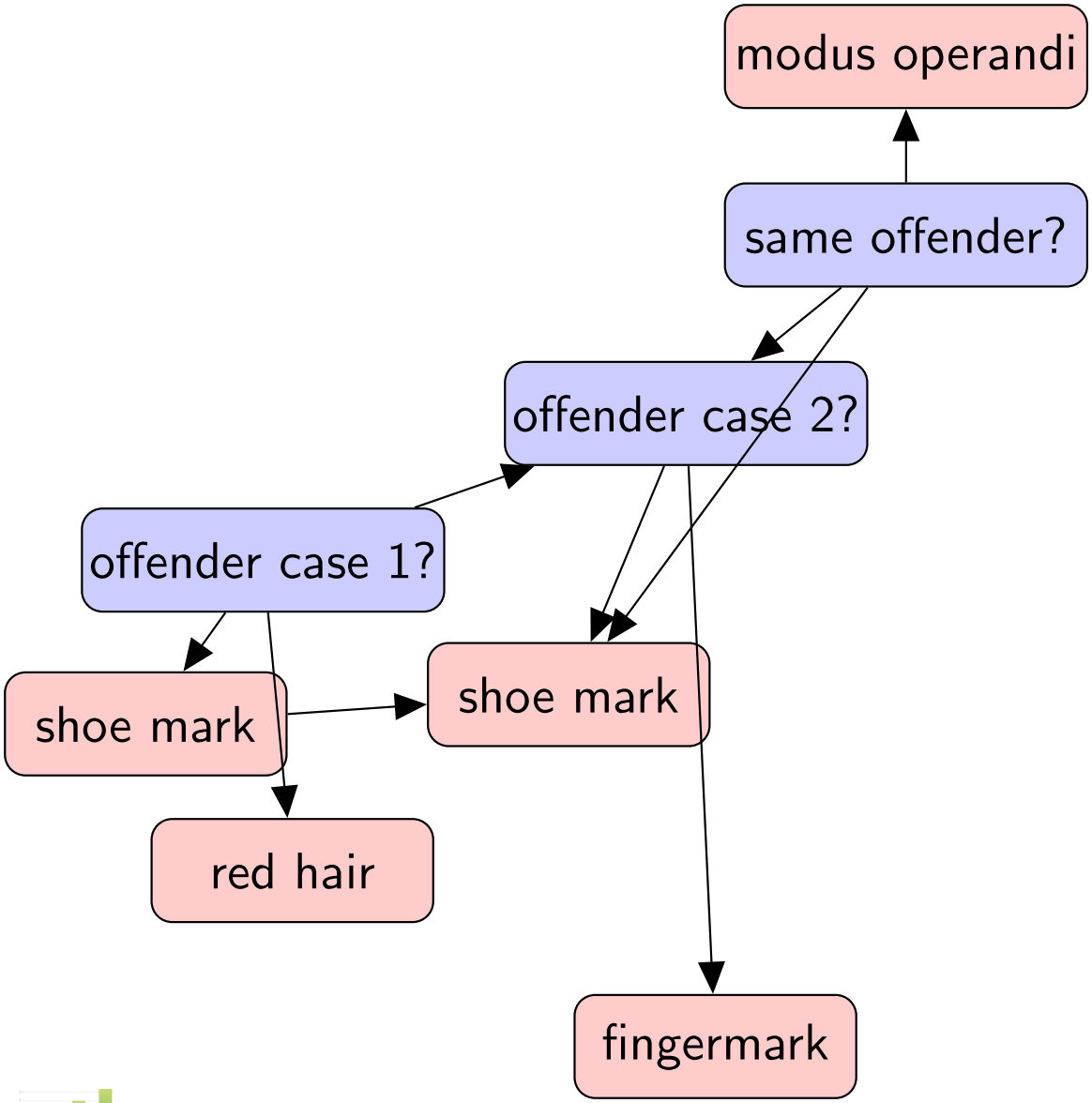
2 crimes


modus operandi
yes, 1%
suspect, 1%
suspect, 1%
shoe mark
shoe mark
red hair
fingermark

2 crimes

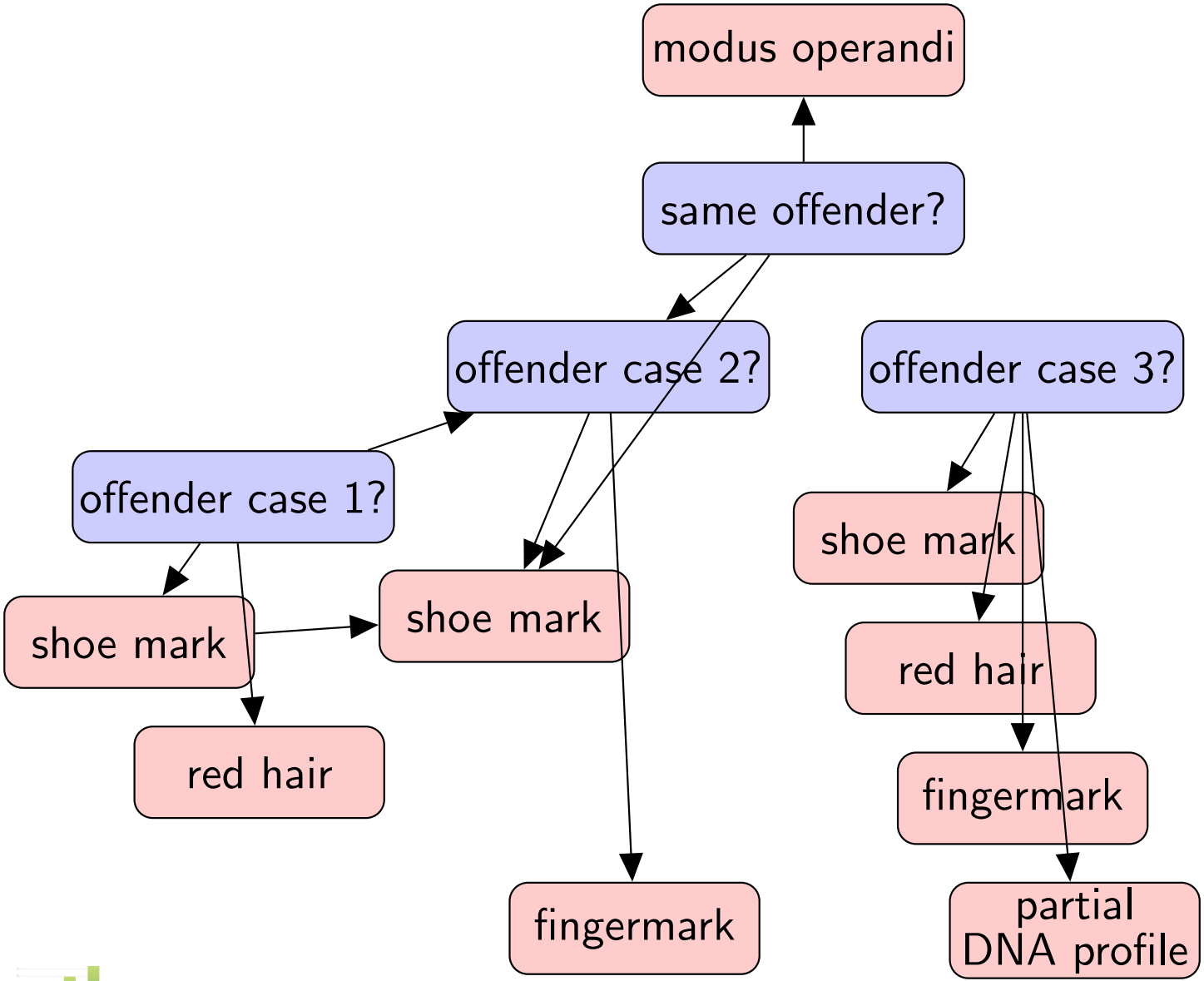

similar
yes, 93%
suspect, 94%
suspect, 96%
match
match
match
match

2 crimes

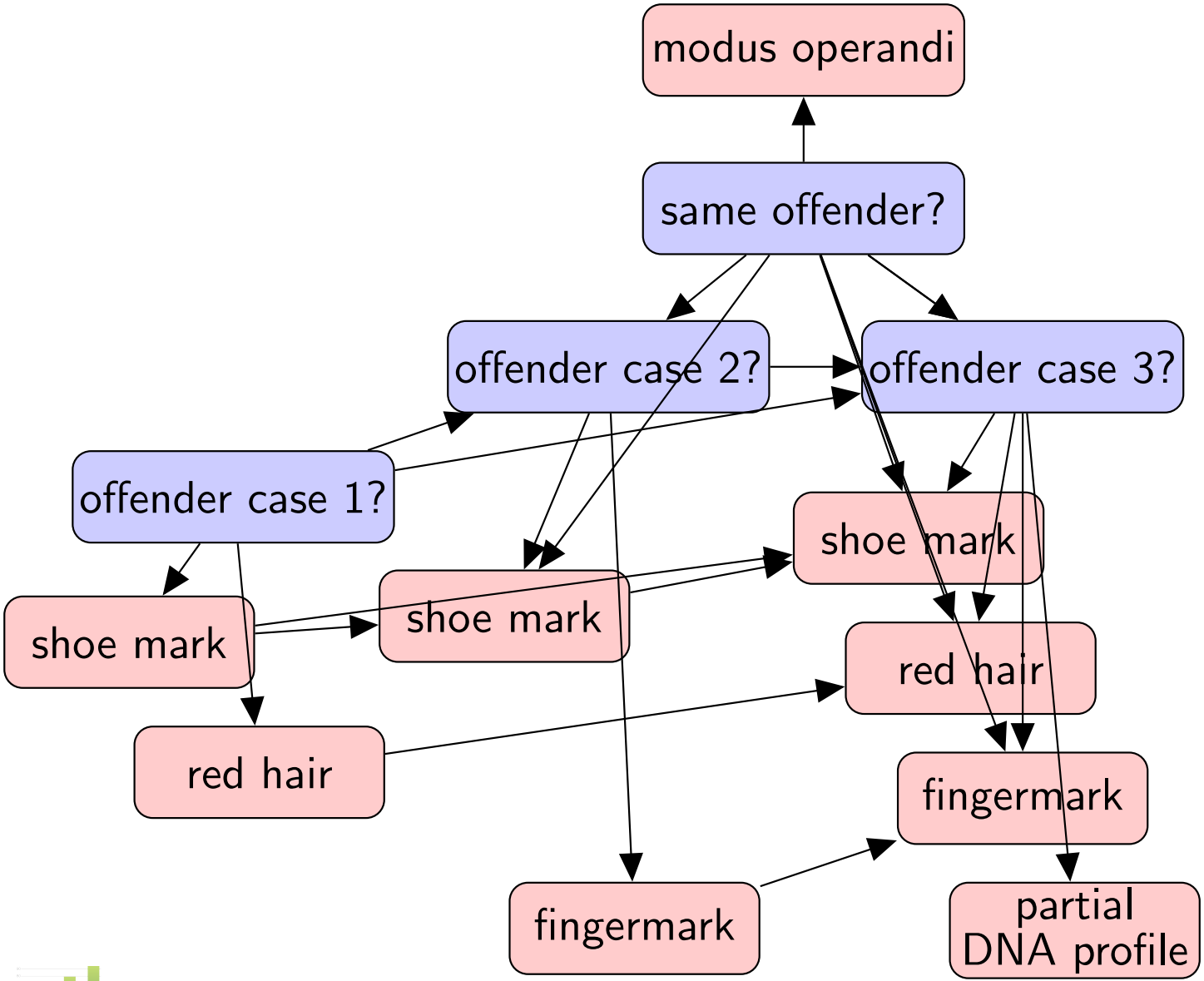


Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

3 crimes

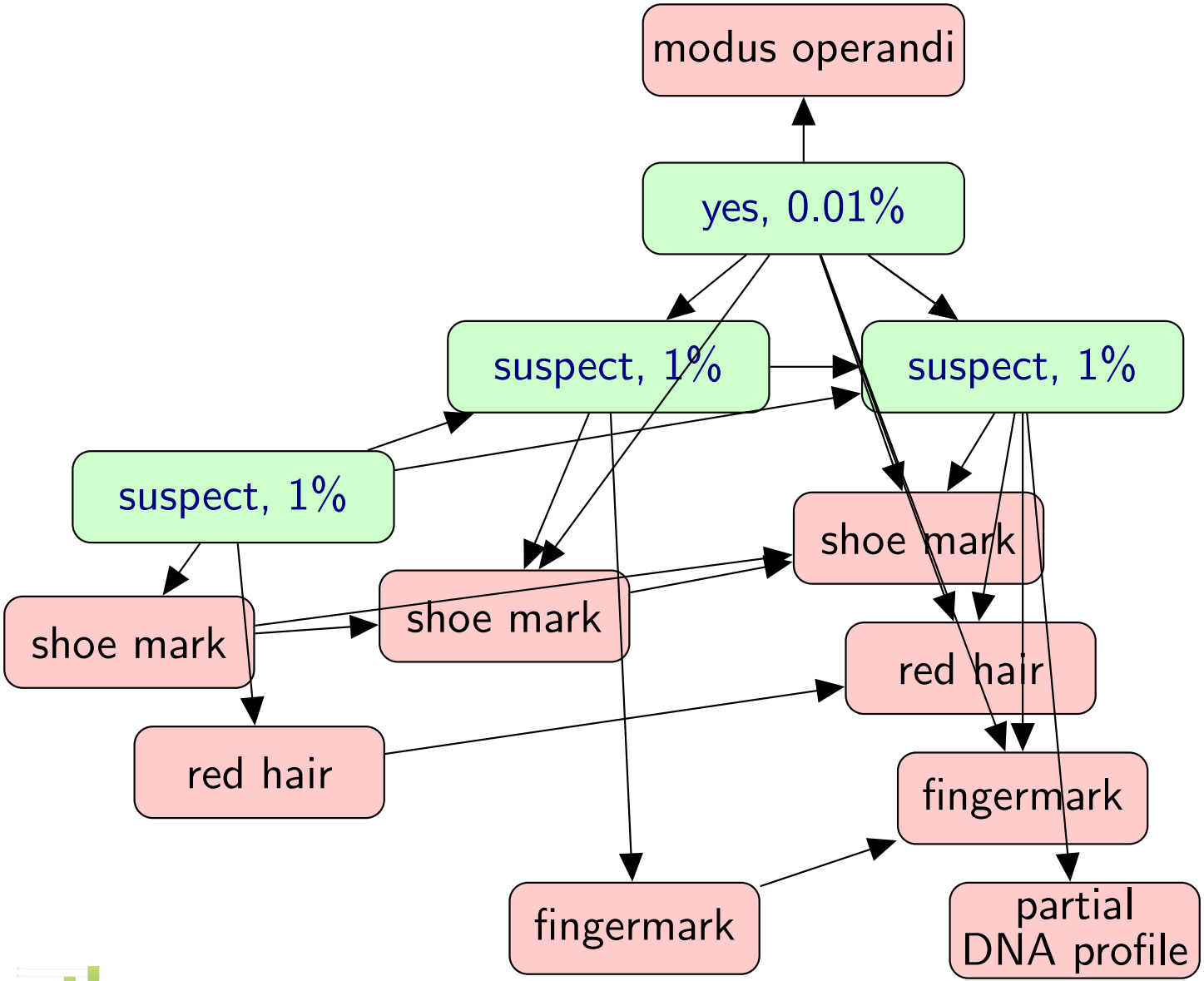

modus operandi
same offender?
offender case 2?
offender case 3?
offender case 1?
shoe mark
shoe mark
red hair
fingermark
shoe mark
red hair
fingermark
partial
DNA profile

3 crimes

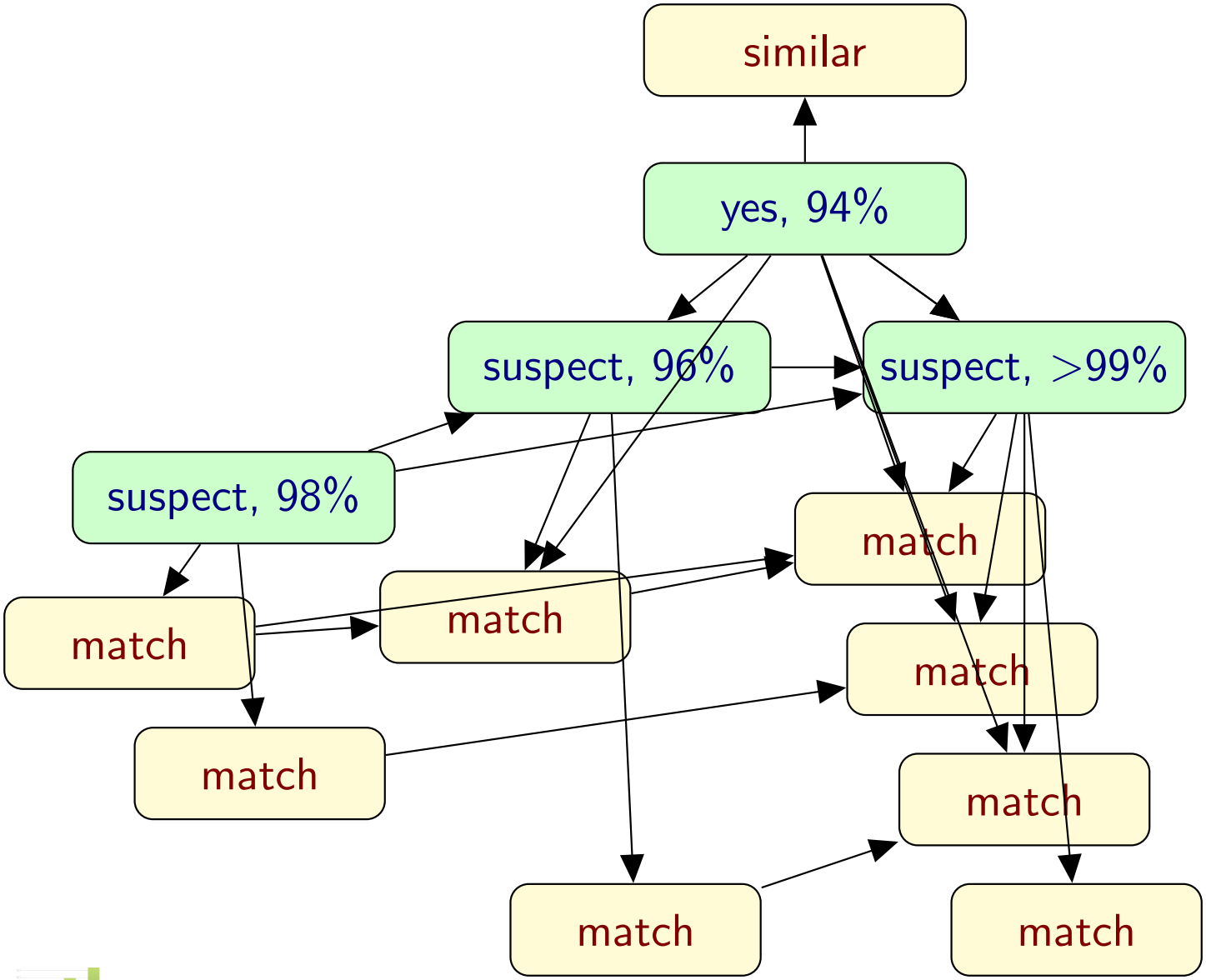


Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

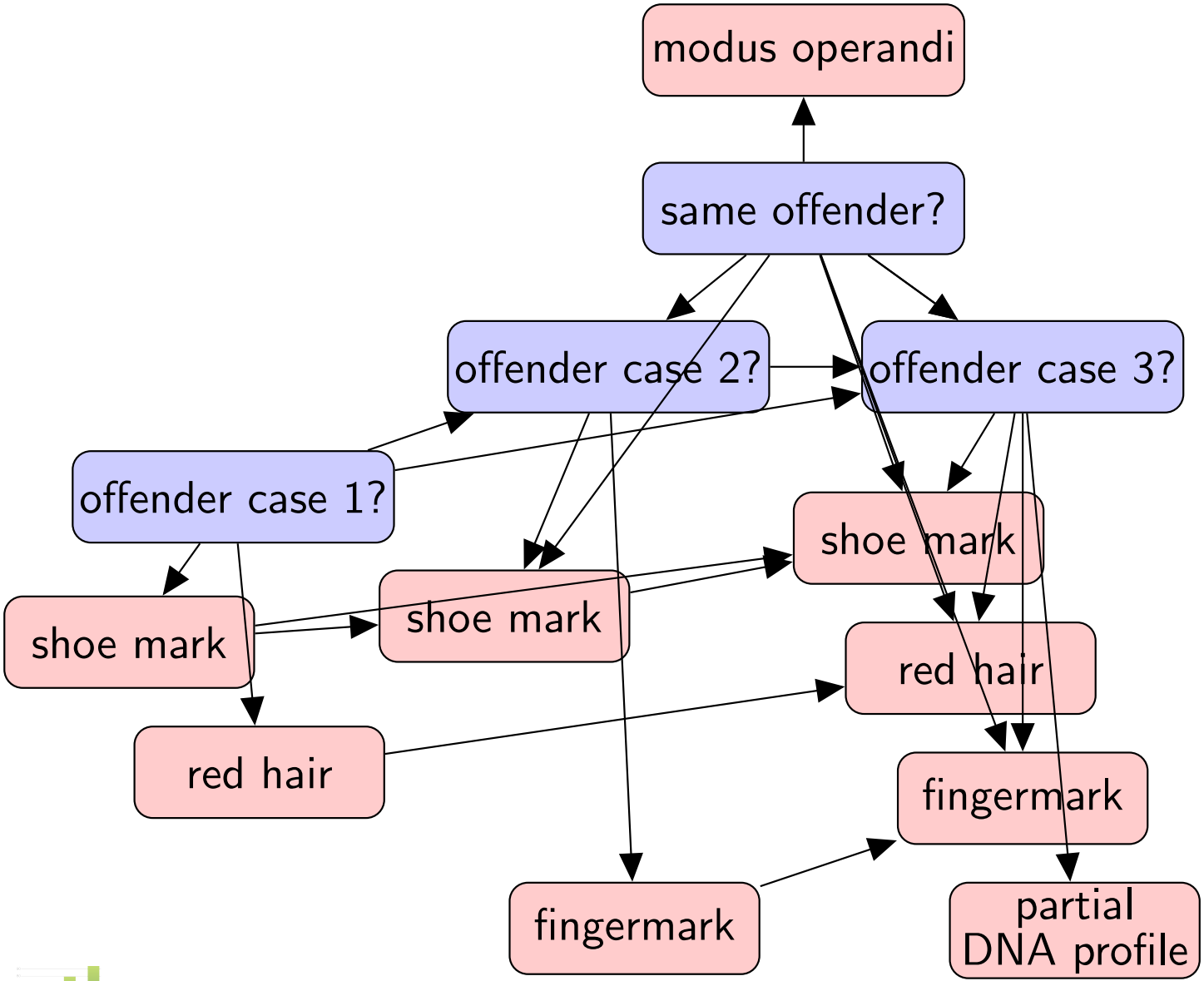
3 crimes


modus operandi
yes, 0.01%
suspect, 1%
suspect, 1%
suspect, 1%
shoe mark
shoe mark
shoe mark
red hair
red hair
fingermark
fingermark
partial
DNA profile

3 crimes

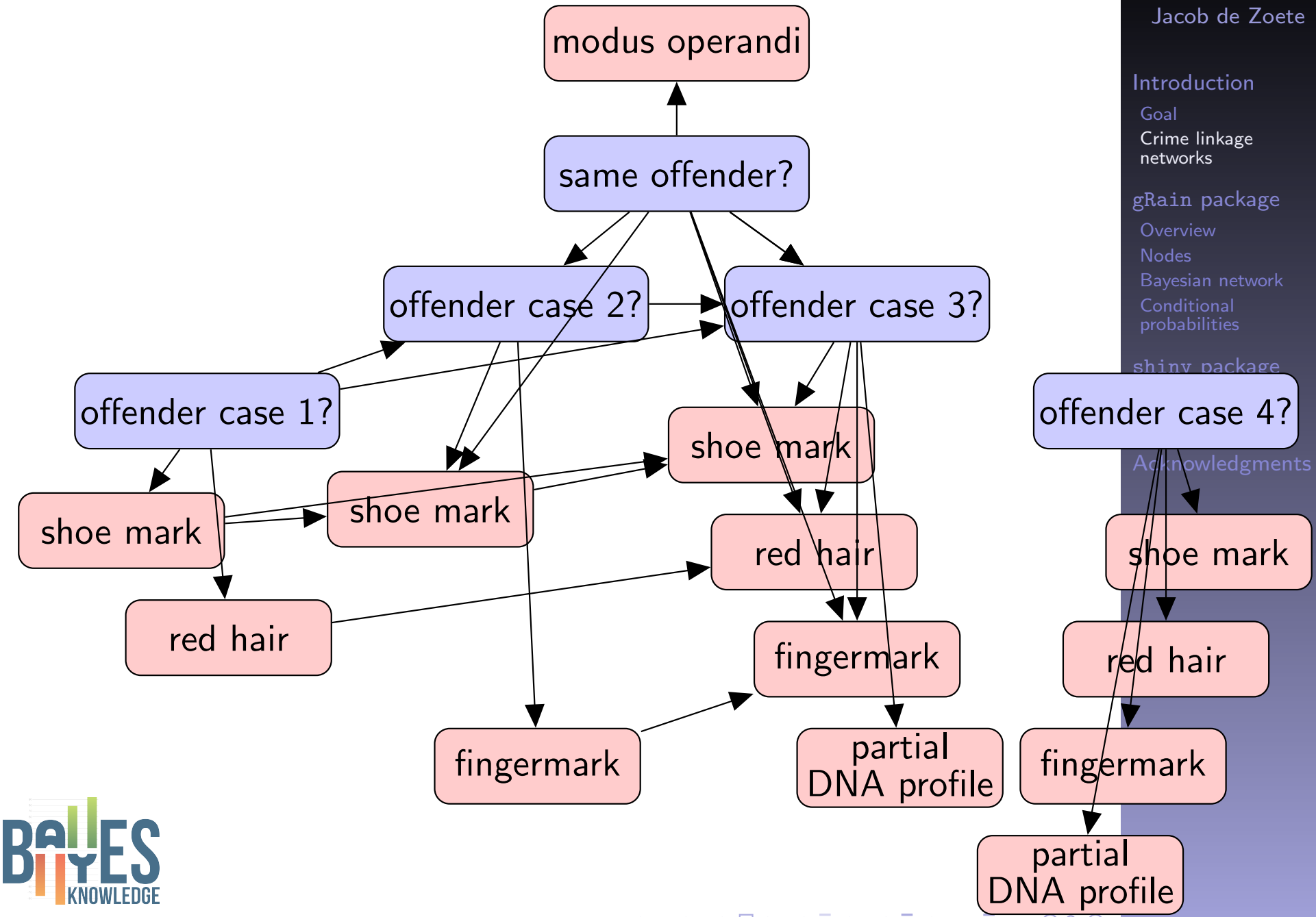

similar
yes, 94%
suspect, 96%
suspect, >99%
suspect, 98%
match
match
match
match
match
match
match
match

3 crimes

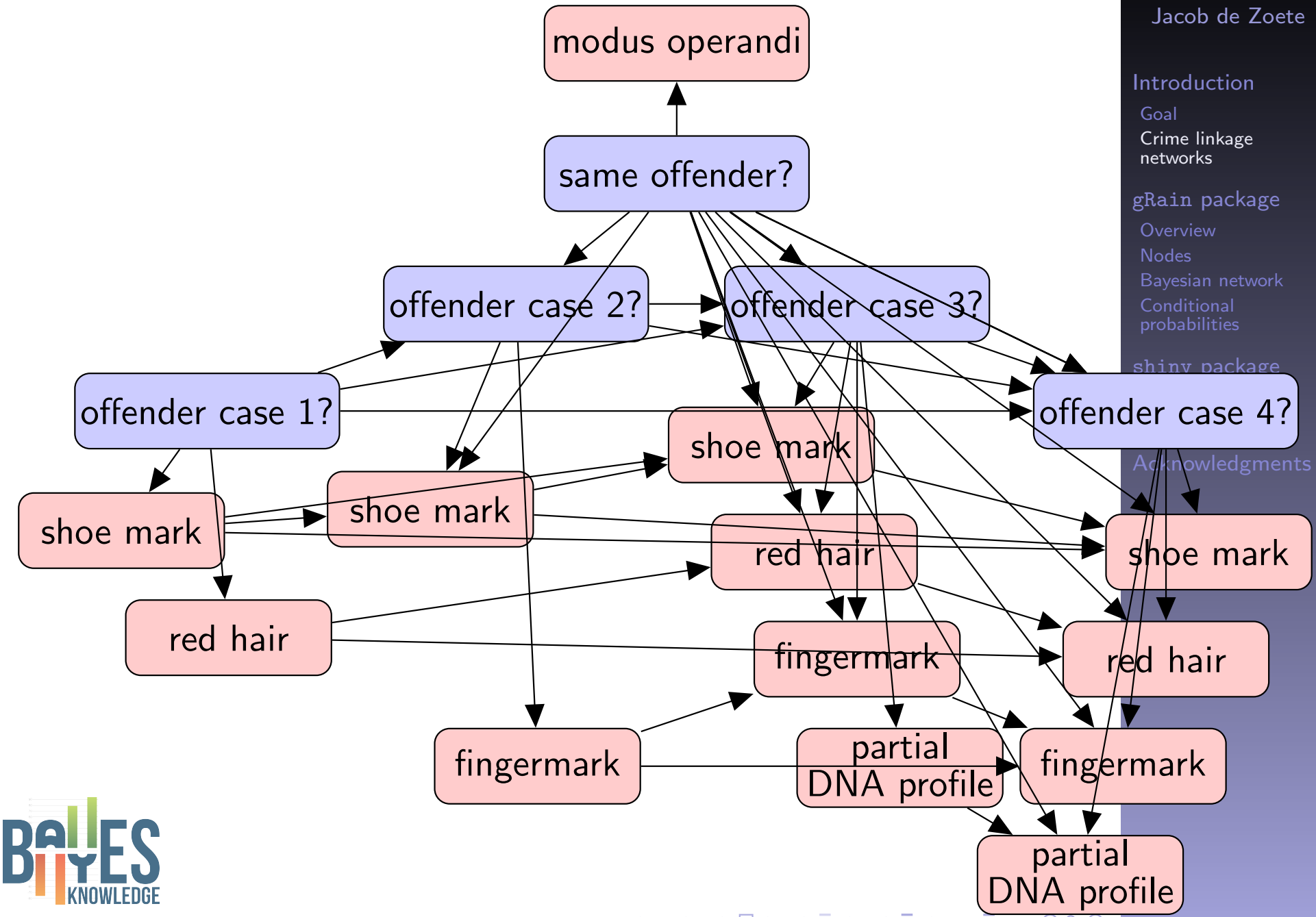


Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

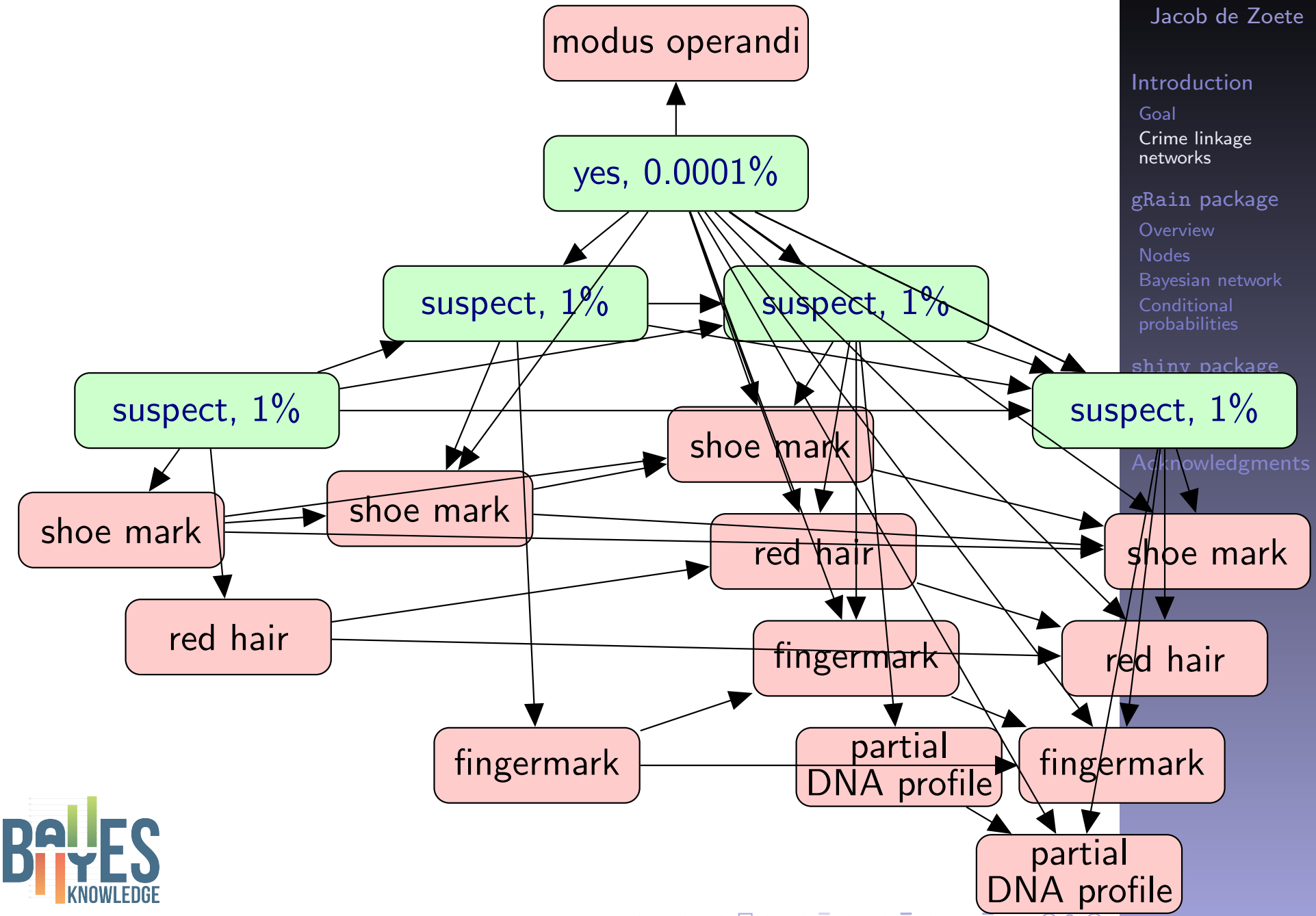
4 crimes


modus operandi
same offender?
offender case 2?
offender case 3?
offender case 1?
offender case 4?
shoe mark
shoe mark
shoe mark
shoe mark
red hair
red hair
fingermark
fingermark
partial DNA profile
fingermark
partial DNA profile

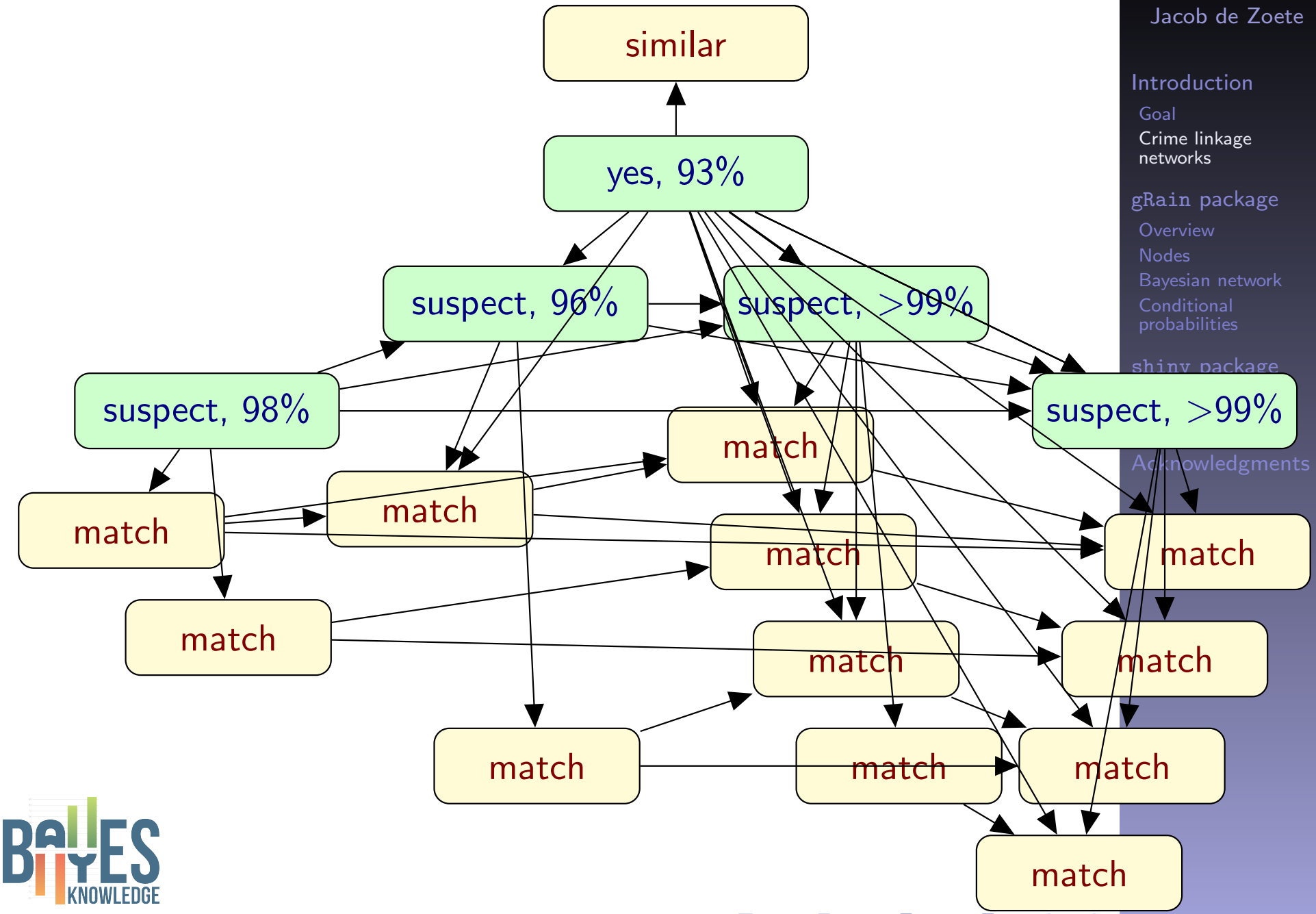
4 crimes


modus operandi
same offender?
offender case 2?
offender case 3?
offender case 1?
offender case 4?
shoe mark
shoe mark
shoe mark
red hair
shoe mark
red hair
fingermark
red hair
fingermark
partial DNA profile
fingermark
partial DNA profile

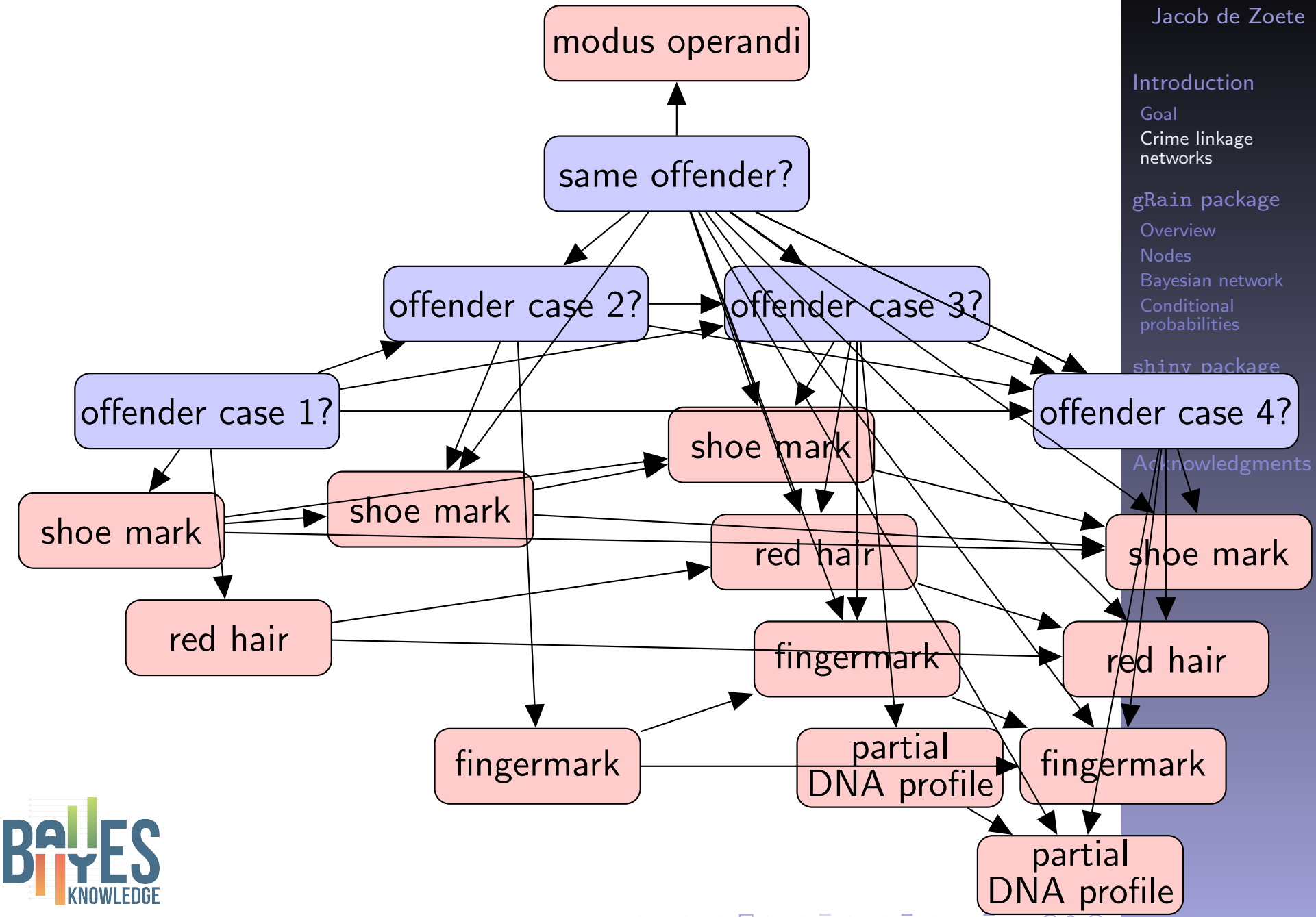
4 crimes


modus operandi
yes, 0.0001%
suspect, 1%
suspect, 1%
suspect, 1%
suspect, 1%
shoe mark
shoe mark
shoe mark
red hair
shoe mark
red hair
fingermark
red hair
fingermark
partial DNA profile
fingermark
partial DNA profile

4 crimes


similar
yes, 93%
suspect, 96%
suspect, >99%
suspect, 98%
suspect, >99%
match
match
match
match
match
match
match
match
match
match
match

4 crimes


modus operandi
same offender?
offender case 2?
offender case 3?
offender case 1?
offender case 4?
shoe mark
shoe mark
shoe mark
red hair
shoe mark
red hair
fingermark
red hair
fingermark
partial DNA profile
fingermark
partial DNA profile

4 crimes


modus operandi
NEVER SEND A HUMAN
TO DO A MACHINE'S JOB
fingermark
DNA profile
fingermark
partial
DNA profile

gRain package

- R is an open source programming language for statistical computing.
- The capabilities of R are extended through user-created packages.
- gRain is an R package for probability propagation in Bayesian networks.

gRain package

```
of1 <-cptable(~offender_1,                                    1
              levels  = c("suspect","unknown"),               2
              values  = c(0.1,0.9))                           3
```

Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,  → node name                    1
              levels  = c("suspect","unknown"),            2
              values  = c(0.1,0.9))                        3
```

Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,                                    1
              levels  = c("suspect","unknown"),  → states     2
              values  = c(0.1,0.9))                           3
```

Generating BNs

Jacob de Zoete

Introduction
Goal
Crime linkage networks

gRain package
Overview
Nodes
Bayesian network
Conditional probabilities

shiny package
Overview
Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,                                    1
              levels  = c("suspect","unknown"),               2
              values  = c(0.1,0.9)) → probability table       3
```

Generating BNs

Jacob de Zoete

Introduction
Goal
Crime linkage networks

gRain package
Overview
Nodes
Bayesian network
Conditional probabilities

shiny package
Overview
Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,
              levels  = c("suspect","unknown"),
              values  = c(0.1,0.9))
```

```
e11<-cptable(~evidence_11 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  0.01, 0.99))
```

gRain package

```
of1 <-cptable(~offender_1,                                    1
              levels  = c("suspect","unknown"),               2
              values  = c(0.1,0.9))                           3

e11<-cptable(~evidence_11 + offender_1,  → node parents       4
              levels  = c("match","no match"),                5
              values  = c(1,0,  0.01, 0.99))                  6
```

Generating BNs
Jacob de Zoete
Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,
              levels  = c("suspect","unknown"),
              values  = c(0.1,0.9))

e11<-cptable(~evidence_11 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  0.01, 0.99))

e12<-cptable(~evidence_12 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  0.02, 0.98))
```


Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,
              levels  = c("suspect","unknown"),
              values  = c(0.1,0.9))

e11<-cptable(~evidence_11 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  0.01, 0.99))

e12<-cptable(~evidence_12 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  0.02, 0.98))

plist <-compileCPT(list(of1,e11,e12))
pn    <-grain(plist)
```

Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,                                    1
              levels  = c("suspect","unknown"),               2
              values  = c(0.1,0.9))                           3

e11<-cptable(~evidence_11 + offender_1,                       4
              levels  = c("match","no match"),                5
              values  = c(1,0,  0.01, 0.99))                  6

e12<-cptable(~evidence_12 + offender_1,                       7
              levels  = c("match","no match"),                8
              values  = c(1,0,  0.02, 0.98))                  9

plist <-compileCPT(list(of1,e11,e12))   → compile             10
pn    <-grain(plist)                                          11
```

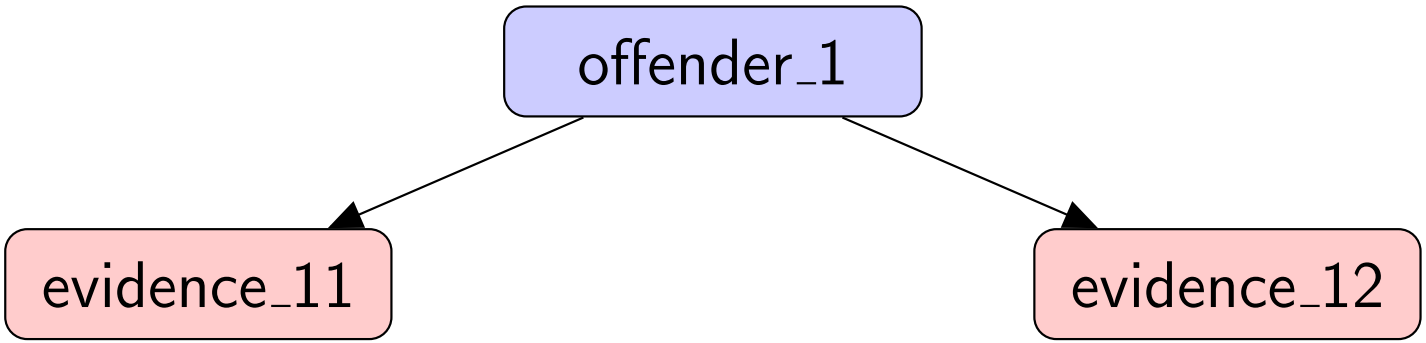
Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,                                  1
              levels  = c("suspect","unknown"),             2
              values  = c(0.1,0.9))                         3

e11<-cptable(~evidence_11 + offender_1,                     4
              levels  = c("match","no match"),              5
              values  = c(1,0,  0.01, 0.99))                6

e12<-cptable(~evidence_12 + offender_1,                     7
              levels  = c("match","no match"),              8
              values  = c(1,0,  0.02, 0.98))                9

plist <-compileCPT(list(of1,e11,e12))                       10
pn    <-grain(plist)    → construct BN                      11
```

Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain network



gRain package - second offender

```
sof <-cptable(~same_offender,                    10
              levels  = c("yes","no"),           11
              values  = c(0.1,0.9))              12
```

Generating BNs

Jacob de Zoete

Introduction
Goal
Crime linkage networks

gRain package
Overview
Nodes
Bayesian network
Conditional probabilities

shiny package
Overview
Application

Acknowledgments

gRain package - second offender

```
sof <-cptable(~same_offender,                                    10
              levels  = c("yes","no"),                           11
              values  = c(0.1,0.9))                              12

of2 <-cptable(~offender_2 + offender_1 + same_offender,          13
              levels  = c("suspect","unknown"),                  14
              values  = c(1,0,  0,1,  0.1,0.9, 0.1,0.9))         15
```

gRain package - second offender

```
sof <-cptable(~same_offender,                                          10
              levels  = c("yes","no"),                                 11
              values  = c(0.1,0.9))                                    12

of2 <-cptable(~offender_2 + offender_1 + same_offender,                13
              levels  = c("suspect","unknown"),                        14
              values  = c(1,0,  0,1,  0.1,0.9, 0.1,0.9))               15

e21<-cptable(~evidence_21 + offender_2,                                16
              levels  = c("match","no match"),                         17
              values  = c(1,0,  0.01, 0.99))                           18
```

gRain package - second offender

```
sof <-cptable(~same_offender,                                    10
              levels  = c("yes","no"),                           11
              values  = c(0.1,0.9))                              12

of2 <-cptable(~offender_2 + offender_1 + same_offender,          13
              levels  = c("suspect","unknown"),                  14
              values  = c(1,0,  0,1,  0.1,0.9, 0.1,0.9))         15

e21<-cptable(~evidence_21 + offender_2,                          16
              levels  = c("match","no match"),                   17
              values  = c(1,0,  0.01, 0.99))                     18
```

gRain package - second offender

```
parents_e21<-c("offender_2")
```

gRain package - second offender

```
parents_e21<-c("offender_2")

if ( exists ("e11"){
  parents_e21<-paste(parents_e21, "+ evidence_11 + same_offender")
}
```

gRain package - second offender

```
parents_e21<-c("offender_2")

if (exists ("e11")){    → check whether variable exists
  parents_e21<-paste(parents_e21, "+ evidence_11 + same_offender")
}
```

Generating BNs
Jacob de Zoete
Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package - second offender

```
parents_e21<-c("offender_2")                                          1

if ( exists ("e11"){                                                  2
  parents_e21<-paste(parents_e21, "+ evidence_11 + same_offender")    3
}                                                                     4
```

↓

"offender_2"

gRain package - second offender

```
parents_e21<-c("offender_2")                                          1

if ( exists ("e11"){                                                  2
  parents_e21<-paste(parents_e21, "+ evidence_11 + same_offender")    3
}                                                                     4
```

↓

"offender_2 + evidence_11 + same_offender"

Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

gRain package - second offender

```
sof <-cptable(~same_offender,                                          10
              levels  = c("yes","no"),                                 11
              values  = c(0.1,0.9))                                    12

of2 <-cptable(~offender_2 + offender_1 + same_offender,                13
              levels  = c("suspect","unknown"),                        14
              values  = c(1,0,  0,1,  0.1,0.9, 0.1,0.9))               15

e21<-cptable(~evidence_21 + parents_e21,                               16
              levels  = c("match","no match"),                         17
              values  = c(1,0,      1,     0,  0,1,     0,    1,       18
                          1,0,  0.01, 0.99, 1,0, 0.01, 0.99))          19

plist <-compileCPT(list(of1,e11,e12,sof,of2,e21))                      20
pn    <-grain(plist)                                                   21
```

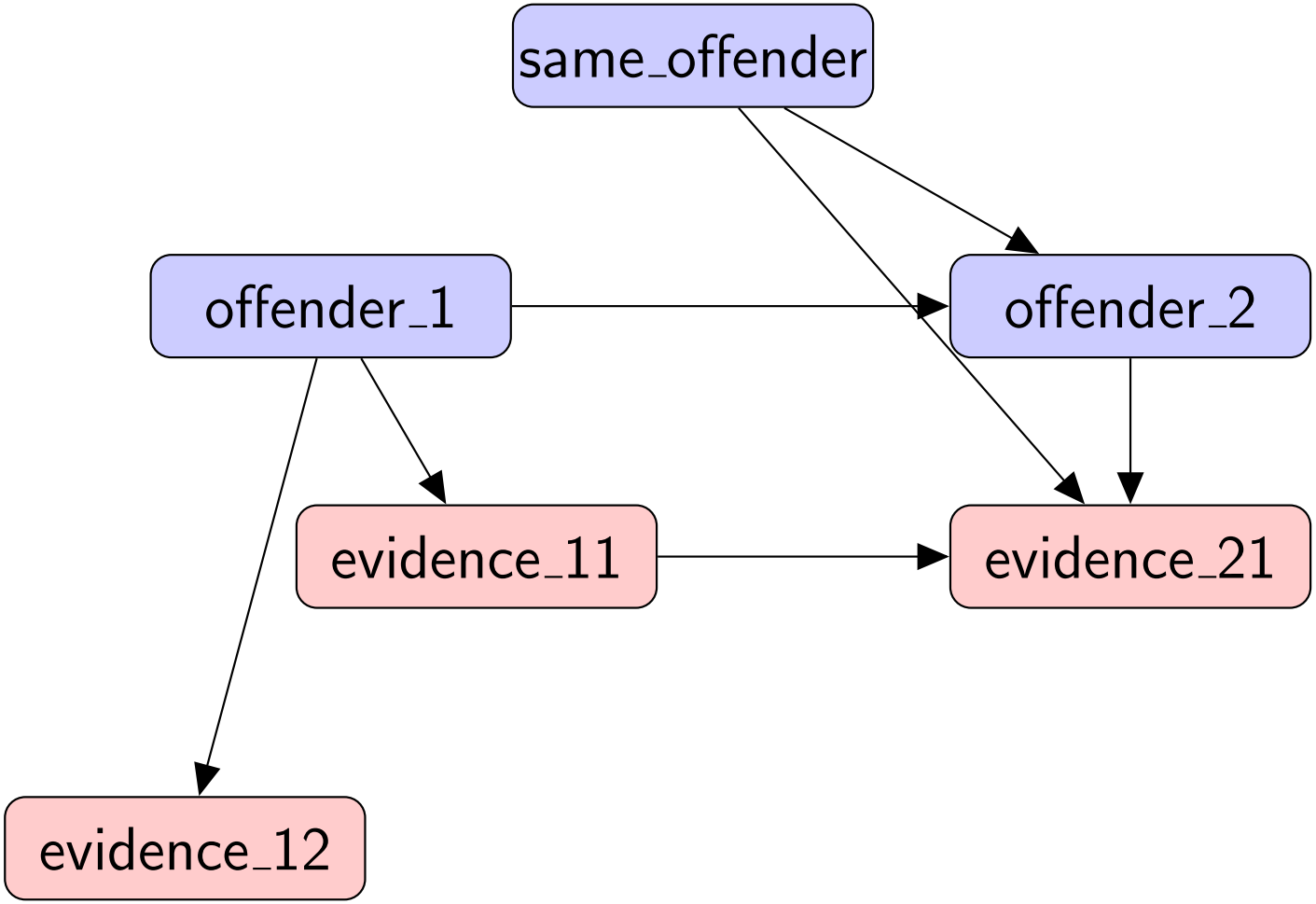
Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package - second offender

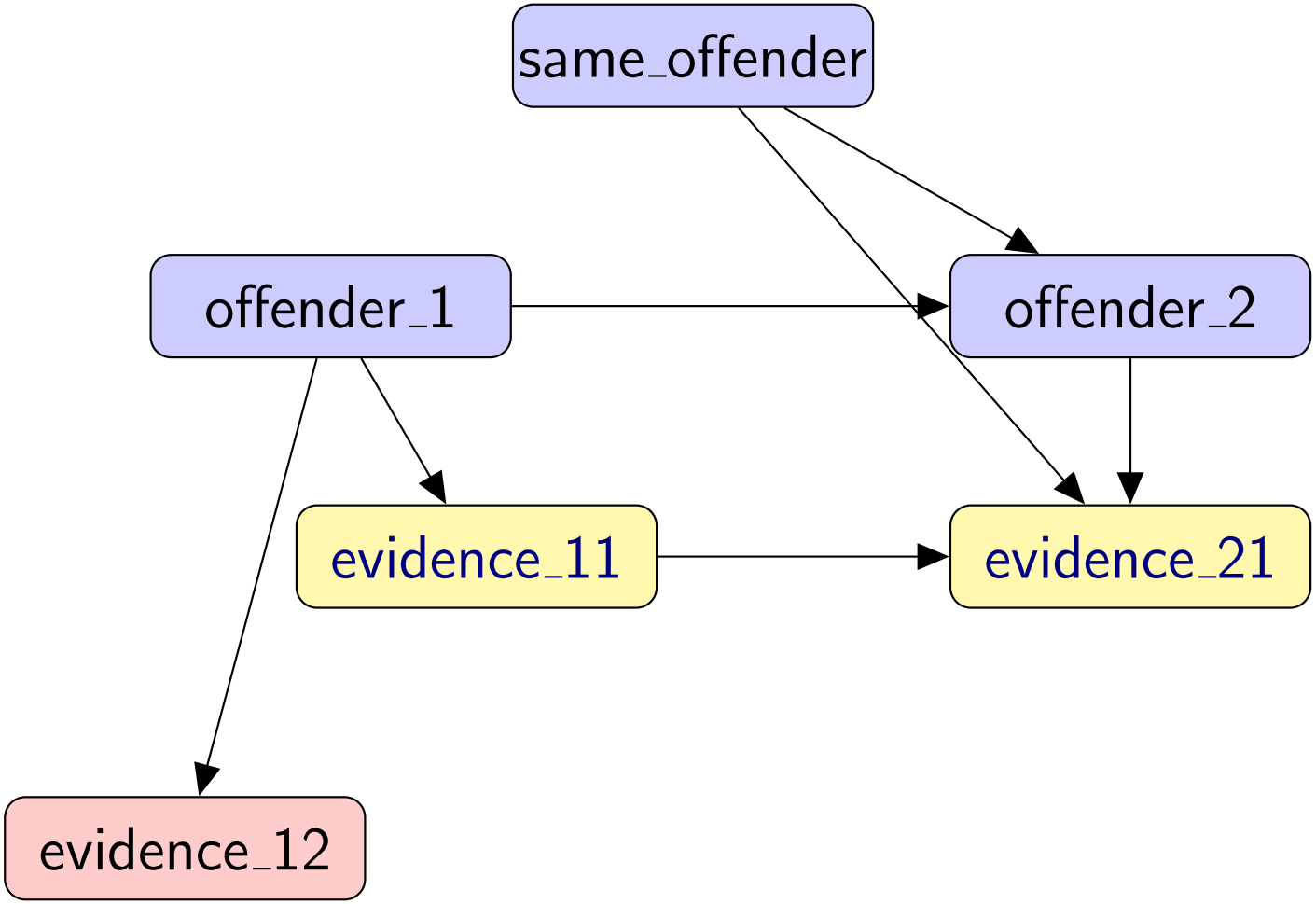
```
sof <-cptable(~same_offender,                                          10
              levels  = c("yes","no"),                                 11
              values  = c(0.1,0.9))                                    12

of2 <-cptable(~offender_2 + offender_1 + same_offender,                13
              levels  = c("suspect","unknown"),                        14
              values  = c(1,0,  0,1,  0.1,0.9, 0.1,0.9))               15

e21<-cptable(~evidence_21 + parents_e21,                               16
              levels  = c("match","no match"),                         17
              values  = c(1,0,      1,     0,  0,1,     0,    1,       18
                          1,0,  0.01, 0.99, 1,0, 0.01, 0.99))          19

plist <-compileCPT(list(of1,e11,e12,sof,of2,e21))                      20
pn    <-grain(plist)                                                   21
```

Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain network



gRain network



Conditional probabilities

evidence_11		
offender_1	*suspect*	*unknown*
match	1	0.01
no match	0	0.99

evidence_21					
offender_2	*suspect*	*unknown*			
same_off.	*	*yes*		*no*	
evidence_11	*	*match*	*no match*	*match*	*no match*
match	1	1	0	0.01	0.01
no match	0	0	1	0.99	0.99

Conditional probabilities

evidence_11		
offender_1	*suspect*	*unknown*
match	1	0.01
no match	0	0.99

evidence_21					
offender_2	*suspect*	*unknown*			
same_off.	*	*yes*		*no*	
evidence_11	*	*match*	*no match*	*match*	*no match*
match	1	1	0	0.01	0.01
no match	0	0	1	0.99	0.99

Conditional probabilities

evidence_11		
offender_1	*suspect*	*unknown*
match	1	p_evid_1
no match	0	1−p_evid_1

evidence_21					
offender_2	*suspect*	*unknown*			
same_off.	*	*yes*		*no*	
evidence_11	*	*match*	*no match*	*match*	*no match*
match	1	1	0	p_evid_1	p_evid_1
no match	0	0	1	1−p_evid_1	1−p_evid_1

Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,
              levels  = c("suspect","unknown"),
              values  = c(1/N, 1-1/N))

e11<-cptable(~evidence_11 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  p_evid_1, 1-p_evid_1))

e12<-cptable(~evidence_12 + offender_1,
              levels  = c("match","no match"),
              values  = c(1,0,  p_evid_2, 1-p_evid_2))
```

Introduction
Goal
Crime linkage networks
gRain package
Overview
Nodes
Bayesian network
Conditional probabilities
shiny package
Overview
Application
Acknowledgments

gRain package

```
of1 <-cptable(~offender_1,                                  1
              levels  = c("suspect","unknown"),             2
              values  = c(1/N, 1-1/N))                      3

e11<-cptable(~evidence_11 + offender_1,                     4
              levels  = c("match","no match"),              5
              values  = c(1,0,  p_evid_1, 1-p_evid_1))      6

e12<-cptable(~evidence_12 + offender_1,                     7
              levels  = c("match","no match"),              8
              values  = c(1,0,  p_evid_2, 1-p_evid_2))      9
```


Generating BNs

Jacob de Zoete

Introduction
- Goal
- Crime linkage networks

gRain package
- Overview
- Nodes
- Bayesian network
- Conditional probabilities

shiny package
- Overview
- Application

Acknowledgments

shiny package

- shiny is an R package for building interactive web applications.
- Using your own R code, it is relatively easy to construct a GUI.
- Furthermore, this application can be accessed online, which removes the necessity of having R installed.

shiny application

http://bayes-knowledge.com/

Acknowledgments

This work is supported by the ERC project BAYES_KNOWLEDGE (**ERC-2013-AdG339182**)

European Research Council

Established by the European Commission

Generating BNs for crime linkage scenarios

References

[de Zoete et al., 2015] Modelling crime linkage with Bayesian networks. *Science & Justice*, 55(3): pp 209 – 217.

[de Zoete et al., 2017] Evaluating evidence in linked crimes with multiple offenders. *Science & Justice*, 57(3): pp 228 – 238.

[Shiny, 2016] shiny: Web Application Framework for R. `R package version 0.13.2`.

[Højsgaard, 2012] Graphical independence networks with the gRain package for R. *Journal of Statistical Software*, 46(10): pp 1 – 26.