Aoristic research in R

Correcting temporal categorizations in archaeology

Nils Müller-Scheeßel, Martin Hinz

March 2018

Institute of Pre- and Protohistoric Archaeology, Kiel University

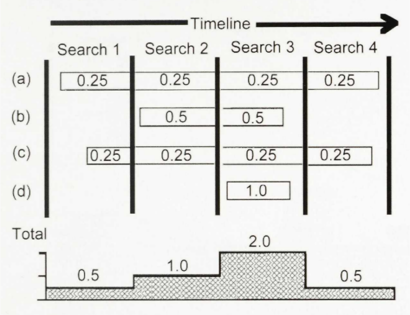
Motivation

What is aoristic?

"An aoristic analysis allows for spatially referenced objects such as geocoded crime locations to be weighted spatially according to a probability estimate. The term aoristic, one of the past tenses of the Greek verb aorist, denotes a past occurrence, with none of the limitations of other past tenses." – Ratcliffe 2002, 27

'the aorist is a class of verb forms that generally portray a situation as simple or undefined' – Wikipedia

What is aoristic?



Principle of aoristic analysis (Mischka 2004, originally Ratcliffe 2000, 671 Fig. 1)

- Production of time series from a collection of evidence with start and end date
  - in arch. situation (estimation of eg. site abundancy per time slice), these are mostly typological phases (based on typochronological artefacts)
  - in anthropological cases (estimation of eg. age distribuition), these are mostly individuals with anthropological age determination
- Widespread and intuitive tool

Aoristics in arch. applications

- background for a lot of archaeological reasoning without proper reference to the method
- earliest explicit applications in archaeology:
  - Johson 2004
  - Mischka 2004
- recently reference eg. by:
  - Crema 2012
  - Kolář et al. 2016
  - Orton et al. 2017
  - Palmisano et al. 2017

Aoristic approaches in anthropology

- „proportional method“: „It distributes the weight of single observations uniformly over the age-intervals into which they have been assigned“ (Boldsen 1988, 335f.)

i: the index of the skeleton.
j: the age in years.
$A1_i$ and $A2_i$: the lower and upper margins of the age interval into which the i'th skeleton has been age determined.

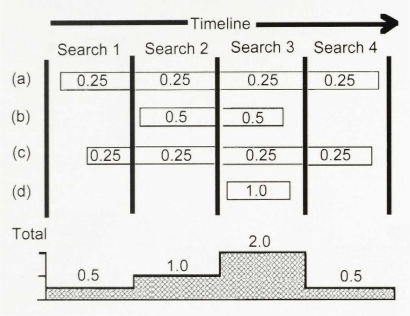
$$\alpha_{ij} = \begin{cases} 1 & \text{if } A1_i \leq j \leq A2_i \\ 0 & \text{else} \end{cases}$$

N: the number of skeletons in the analyzed sample.
Using the proportional method:

$$F(t) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=0}^{t} \{\alpha_{ij}/(A2_i - A1_i + 1)\}$$

Boldsen 1988, 336

Ordinal scales and overlapping categories: the problem



Principle of aoristic analysis (Mischka 2004)

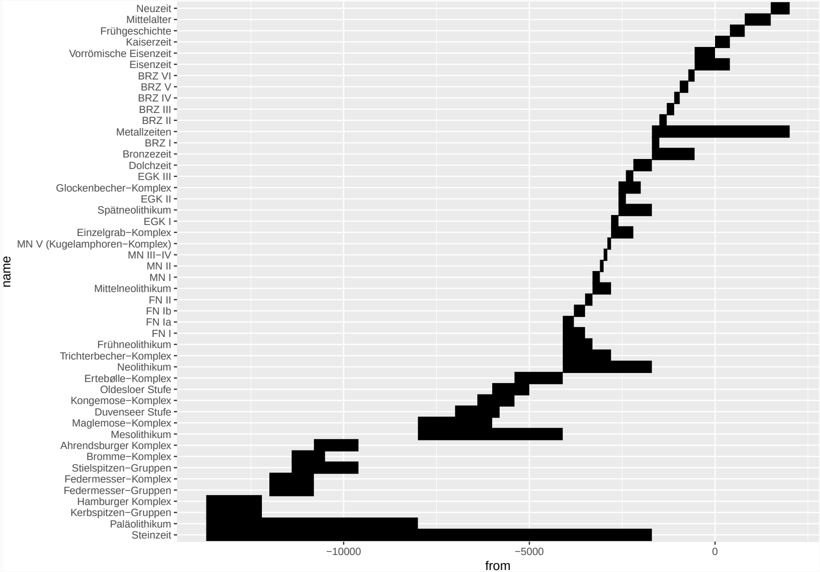
- ideal world: ordinal scales „involve discrete non-overlapping categories“ (Byers 2017, 15)
- reality, due to fragmentary nature of data:
  - archaeological phasing is most often neither exclusive nor non-overlapping
  - anthropological age categories show different grades of resolution

Resulting problem:

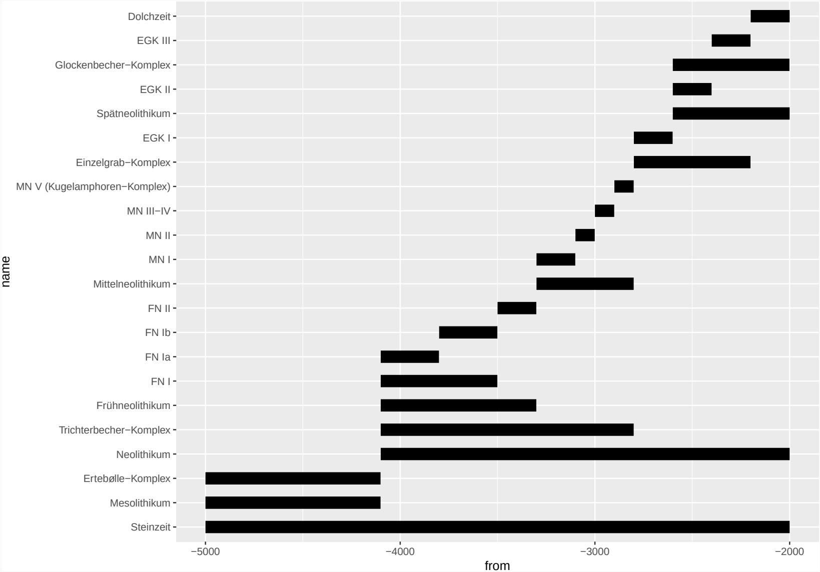
- values where categories overlap tend to get overemphasized

Case study 1: Archaeological context

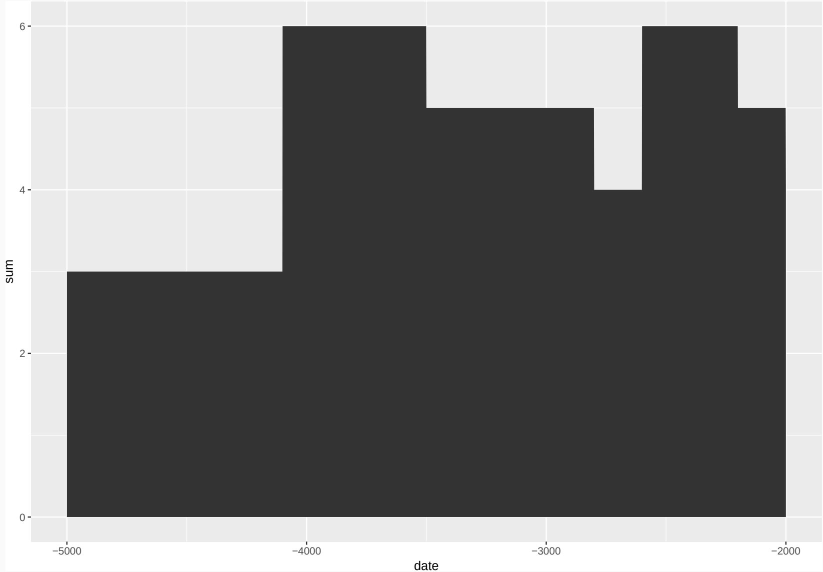
The archaeological situation



The archaeological situation, a bit more reduced



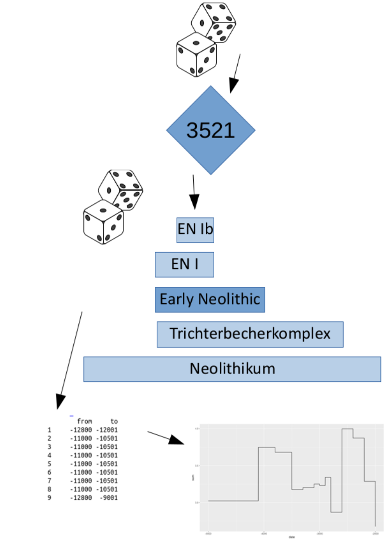
Number of overlapping periods



Simulating the typological dating process

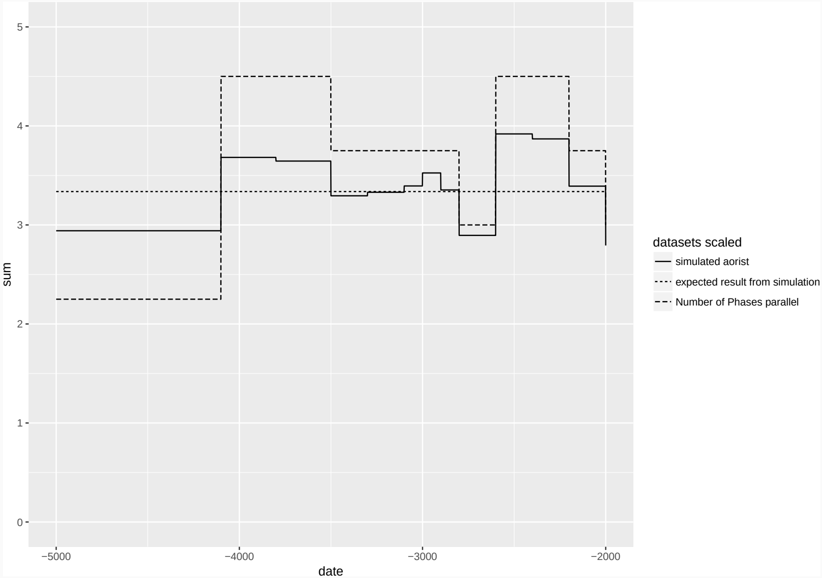
Simulation

- Equal distribution of data by randomly selecting a date within the time interval
- Preservation of objects is independent from their typochronological diagnostical value → random selection of possible archaeological dating
- Resulting dates analysed aoristically

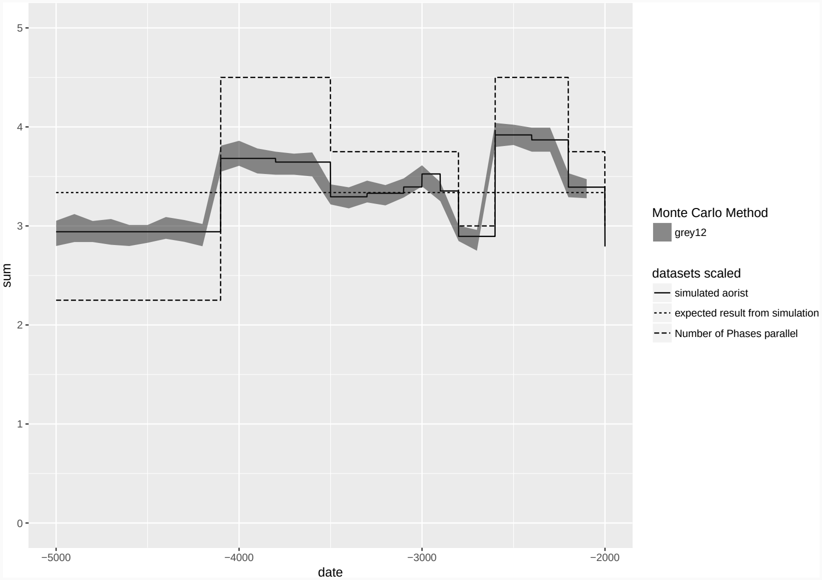


Monte Carlo simulation archaeological periods

Result of a naive aoristic analysis



Result of monte carlo method (Crema 2012 etc., using archSeries package)



Solution to the problem

Solution

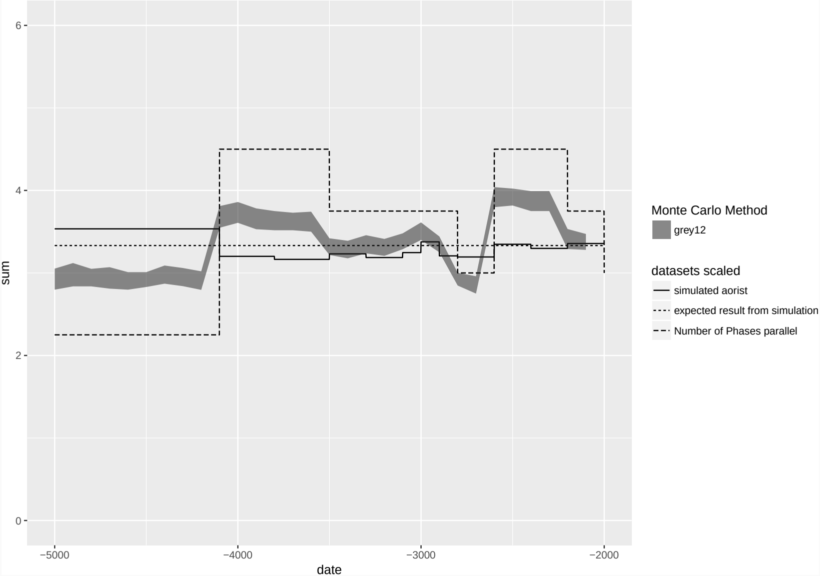
- Correct the aoristic weight by the number of parallel phases:
- Aoristic weight for each interval is 1/n parallel phases
- Normalize to row sum = 1

	3499-3400	3399-3300	3299-3200
Phase 1	1	1	1
Phase 2	0	1	0
Phase 3	1	1	0

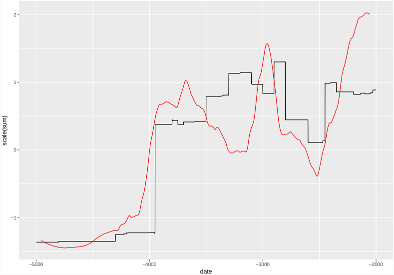
	3499-3400	3399-3300	3299-3200
Phase 1	0.5	0.3333333	1
Phase 2	0.0	0.3333333	0
Phase 3	0.5	0.3333333	0

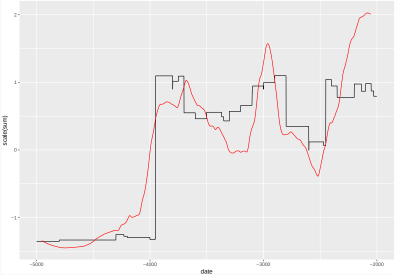
	Phase 1	Phase 2	Phase 3
Site 1	0.2727273	0.1818182	0.5454545
Site 2	0.0000000	1.0000000	0.0000000
Site 3	0.6000000	0.4000000	0.0000000

Result of a corrected aoristic analysis



Real World: Neolithic at Jutland





Real World Example: Aoristic analysis of the Neolithic od Northern Jutland. Left uncorrected, right corrected. Black aoristic result, red sum calibration, smoothed by 100 year moving window.

Case study 2: Anthropological context

A simulated „real-world“ anthropological case study

```r
magdalenenberg
```

```
##        a    Dx
## 1    0-4  3.79
## 2    5-9  4.62
## 3  10-14  4.54
## 4  15-19  4.21
## 5  20-24 14.99
## 6  25-29 20.61
## 7  30-34 17.20
## 8  35-39 14.39
## 9  40-44  6.68
## 10 45-49  4.04
## 11 50-54  5.49
## 12 55-59  5.72
## 13  60-x  4.72
```

```r
sc
```

```
##              lx     ls
## 1  0.96585586 0.8499
## 2  0.92423423 0.7691
## 3  0.88333333 0.7502
## 4  0.84540541 0.7362
## 5  0.71036036 0.7130
## 6  0.52468468 0.6826
## 7  0.36972973 0.6525
## 8  0.24009009 0.6223
## 9  0.17990991 0.5898
## 10 0.14351351 0.5534
## 11 0.09405405 0.5106
## 12 0.04252252 0.4590
## 13 0.02126126 0.3965
```

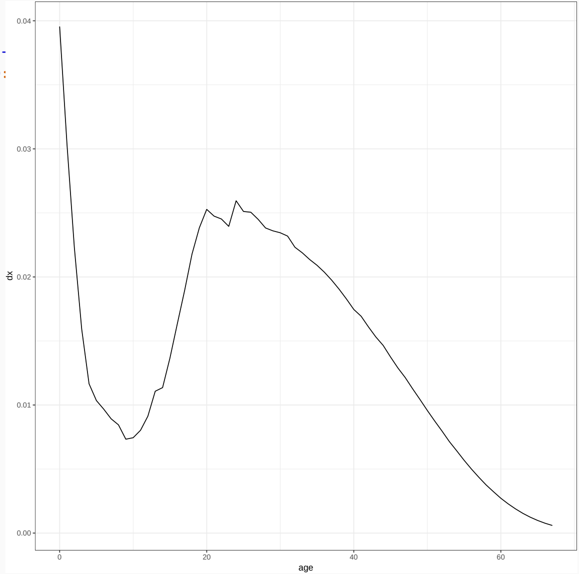
```r
sc <- mutate(sc, yx = 0.5*log( (1-lx)/lx),
             ys = 0.5*log( (1-ls)/ls))
lf <- lm(yx ~ ys, data=sc)

# Data available at
# http://data.princeton.edu/eco572/datasets/brassrlm1.dta

rlm1 <- read.dta("brassrlm1.dta")
yfit1 <- predict(lf, newdata=rlm1)
yearly_lx <- cbind(1:69, 1/(1 + exp(2 * yfit1[1:69])))
yearly_lx[1:17,]
```

```
##    [,1]      [,2]
## 1     1 0.9796006
## 2     2 0.9400476
## 3     3 0.9097007
## 4     4 0.8874003
## 5     5 0.8714727
## 6     6 0.8598166
## 7     7 0.8494700
## 8     8 0.8397969
## 9     9 0.8308738
## 10   10 0.8224225
## 11   11 0.8150926
## 12   12 0.8076455
## 13   13 0.7996179
## 14   14 0.7904883
## 15   15 0.7794131
## 16   16 0.7680557
## 17   17 0.7543888
```

simulating the Magdalenenberg Population

```
dx <- diff(yearly_lx[,2]) * -
magdasim <- data.frame(age=0:
                    dx=dx)

magdasim[1:17,]
```

```
##    age          dx
## 2    0 0.039552917
## 3    1 0.030346919
## 4    2 0.022300374
## 5    3 0.015927641
## 6    4 0.011656132
## 7    5 0.010346601
## 8    6 0.009673067
## 9    7 0.008923062
## 10   8 0.008451354
## 11   9 0.007329895
## 12  10 0.007447123
## 13  11 0.008027555
## 14  12 0.009129596
## 15  13 0.011075182
## 16  14 0.011357435
## 17  15 0.013666906
## 18  16 0.016332366
```



Real world Age Cohorts

A typical „real-world“ example of anthropological age categories (adapted from Moghaddam et al. 2016 on the cemetery of Münsingen-Rain)

```
##    from to
## 1     3  4
## 2     7  8
## 3     9 10
## 4     9 11
## 5    10 10
## 6    12 13
## 7    12 14
## 8    13 14
## 9    15 19
## 10   18 27
## 11   18 29
## 12   20 25
## 13   20 29
## 14   20 49
## 15   25 34
## 16   25 39
## 17   25 44
## 18   30 39
## 19   30 44
## 20   30 49
## 21   34 49
## 22   35 44
## 23   35 49
## 24   40 49
## 25   40 54
## 26   40 59
## 27   45 59
## 28   50 99
## 29   60 99
## 30    0  0
## 31    0  6
## 32    1  2
## 33    3  6
## 34    1  3
```

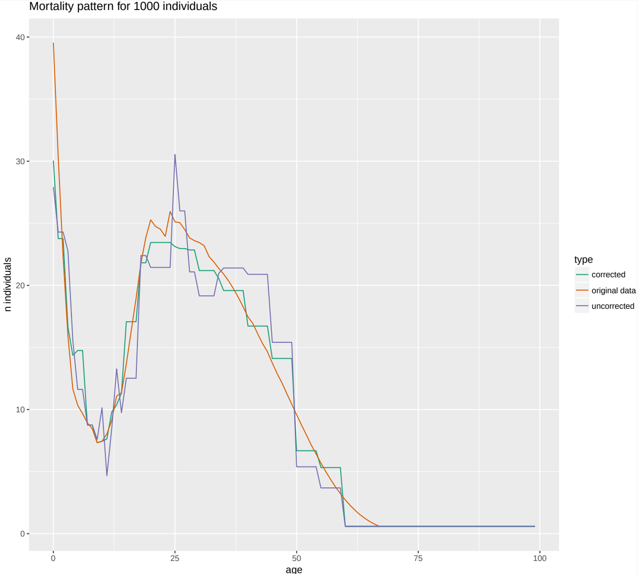
The simulation

```
# simulate 10000 individuals based on their dx

sim_pop_age <- sample(magdasim$age,                             # from the ages
                      size = 10000,                             # take 10000 individuals
                      replace = T,                              # multible times is ok
                      prob = magdasim$dx)                       # with probability according to their dx

sim_result <- sapply(sim_pop_age,                               # for all individuals
                     function(x){
                       periods_reduced %>%                      # from the periodes
                         subset(from <= x & to >= x) %>%        # take those within which the age falls
                         sample_n(size = 1, replace = T) %>% # take on of those randomly
                         unlist                                 # reformat
                     }
)

interval_collector <- sim_result %>% t() %>% as.data.frame() # reformat the result

head(interval_collector)
```

```
##   from to
## 1   35 49
## 2   12 14
## 3   18 27
## 4   35 49
## 5   20 49
## 6   25 34
```

The simulation

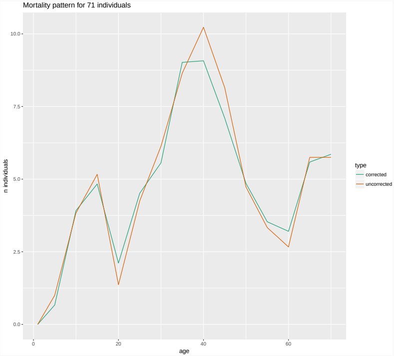
```
##
## Attaching package: 'aoristAAR'
## The following object is masked from 'package:archSeries':
##
##     aorist
```




```
# simulate 10000 individuals b

sim_pop_age <- sample(magdasin
                      size = 1
                      replace
                      prob = m

sim_result <- sapply(sim_pop_a
                     function(
                       periods
                         subse
                         sampl
                         unlis
                     }
)

interval_collector <- sim_resu
```

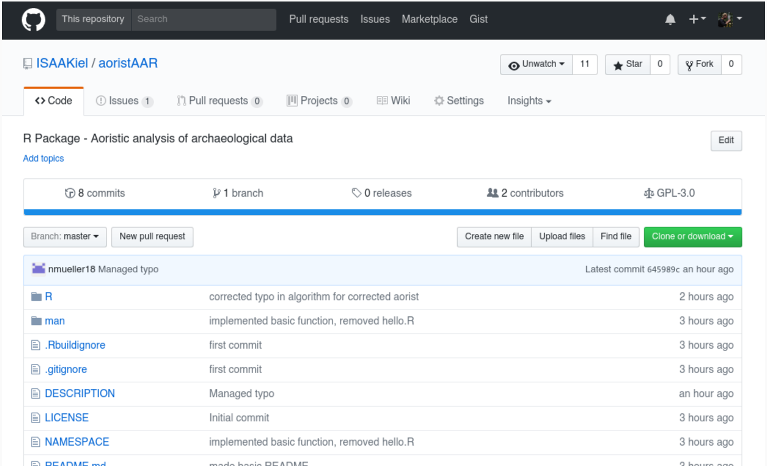
A real „real-world“ example

the cemetery of Münsingen-Rain (processing with the R-package „mortAAR“, data after Moghaddam et al. 2016)



Conclusion

- Aoristic analysis is still a valuable tool, might need a bit calibration
- Monte Carlo method ala Crema 2012 etc. is unfortunately not sufficient for removing bias of overlapping phases
- R package for doing (corrected!) aoristic analysis is available at the Github Repository of ISAAK: https://github.com/ISAAKiel/aoristAAR



Thank you

Nils Müller-Scheeßel

nils.mueller-scheessel@ufg.uni-kiel.de

Martin Hinz

martin.hinz@ufg.uni-kiel.de

The code of this presentation is available at:

https://github.com/MartinHinz/aoristAAR_pres