Retrospective Testing - How Good Heuristics Really Work

Andreas Marx

amarx@gega-it.de

AV-Test.org

University of Magdeburg

GEGA IT-Solutions GbR

Retrospective Testing I

- What it is: Use an old (archived) version of an anti-virus program...
- …and test it against the most current viruses (that were not known at the date of the last product update)…
- … to find out how good the heuristic and generic detection of an av program really works
- Better than using VCKs or self-written viruses!

Retrospective Testing II

- The main critical point by av researchers:
  - Such a test shows only something about the past, but nothing for the future
  - But that‘s wrong! (Why have we learned history at school?)
  - Therefore, we should learn from the past (good and bad points) for improvements in the future

Retrospective Testing III

- What can be compared?
  - Sure... detection scores for different types of malware (ITW and Zoo), but also:
  - Speed differences, database sizes (updates), number of virus signatures (what the program claims to detect), false positives, disinfection rates, scores of archived and compressed files, relations between these values etc.

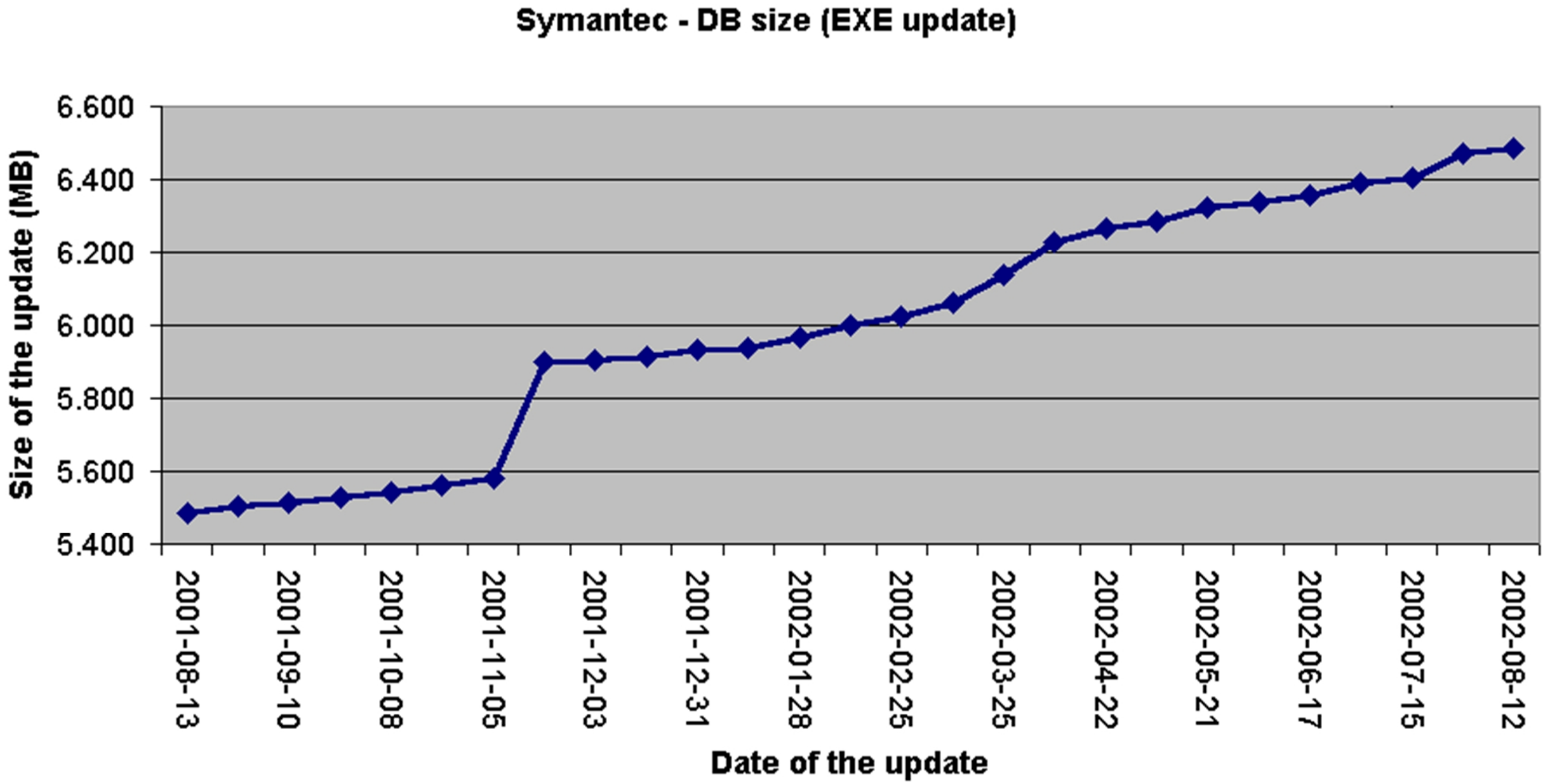
Retrospective Testing IV

- Our test methodology
  - We have compared 20 different engines (not products) for a period of more than one year now
  - We have collected all updates bi-weekly
- But I don‘t want to overflood you with all 75.000+ single entries in the XLS sheet, therefore I‘ve only picked out a few interesting issues from 15 different products

Virus Signature Database I

- Let‘s start with virus signature databases...
- The main question would be, at which ratio the databases increases per month or per year?
- What‘s the best product here with both very good detection scores and a slow increase rate?

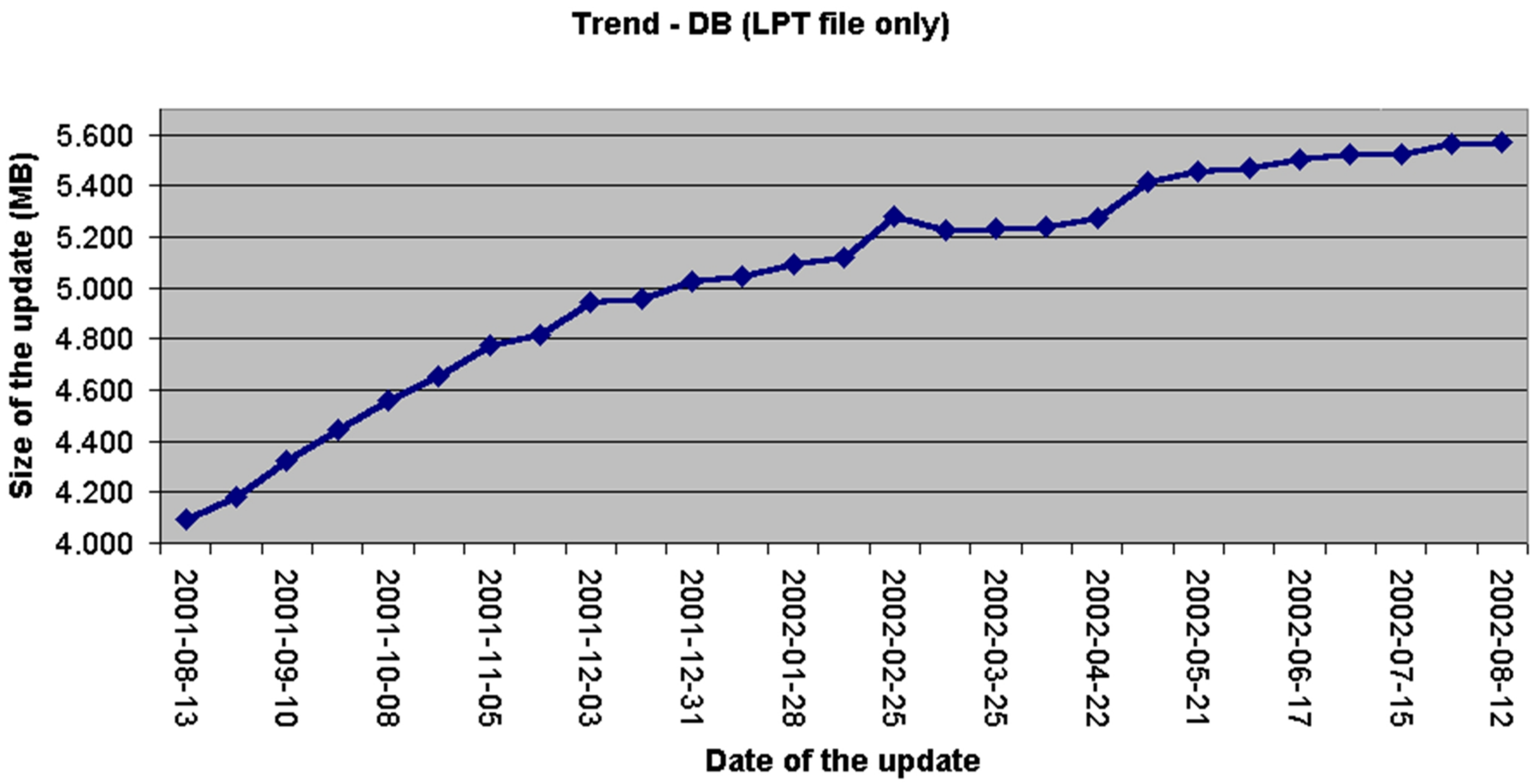
Virus Signature Database II



Virus Signature Database III

- Symantec Norton Anti-Virus (compressed EXE engine and def's installation archive)
  - Size on 2001-09-10: 5.484.077 Bytes
  - Size on 2002-09-09: 6.483.425 Bytes
  - Increase: About 1 MB last year!
    - About 83 KB a month or 18,2 % a year

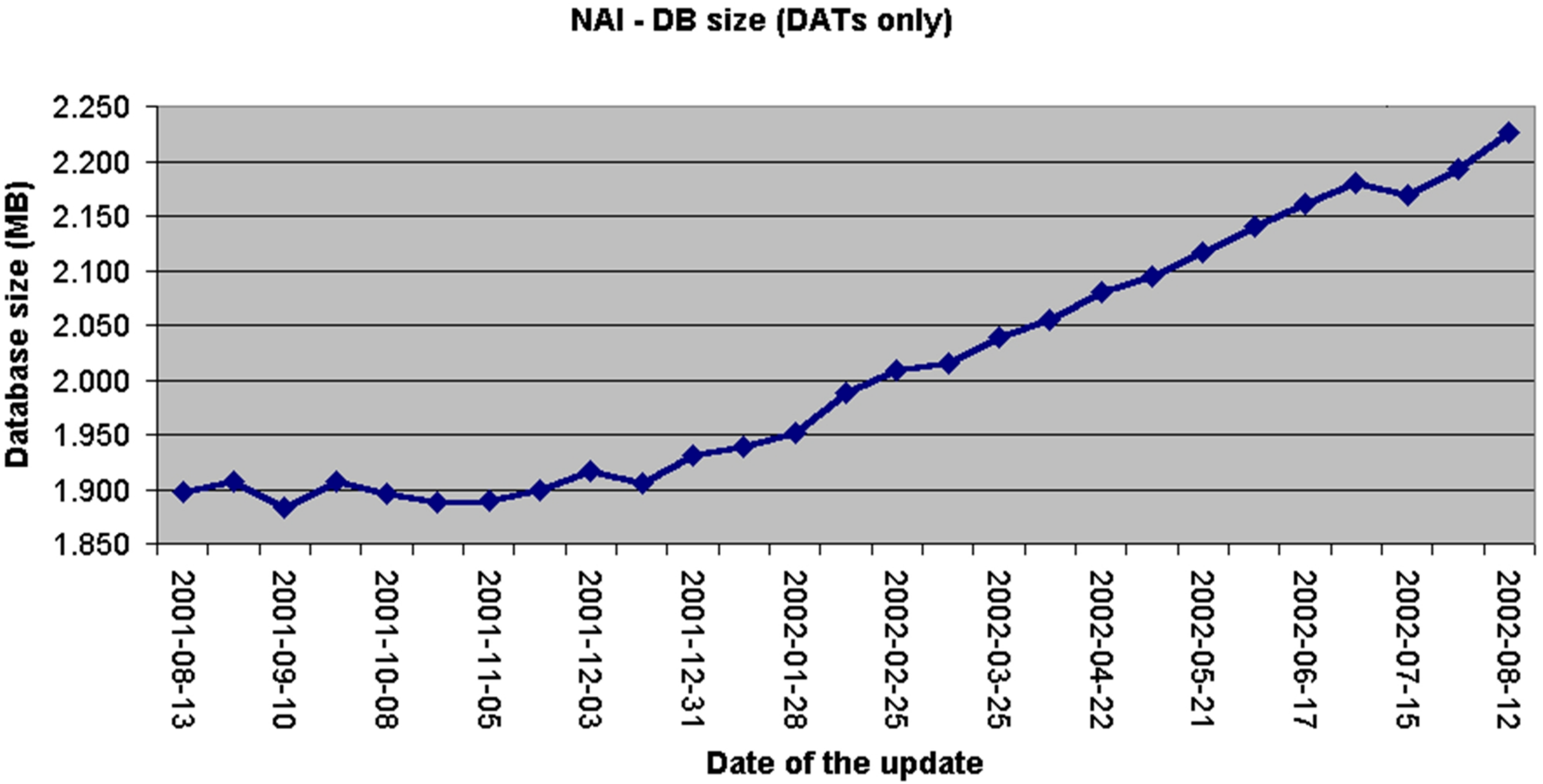
Virus Signature Database IV



Virus Signature Database V

- Trend Micro (uncompressed LPT virus definition file only)
  - Size on 2001-09-10: 4.093.616 Bytes
  - Size on 2002-09-09: 5.574.396 Bytes
  - Increase: About 1,5 MB last year!
    - About 123 KB a month or 26,5 % a year
    - ZIP-compressed, the file was growing by 683 KB

Virus Signature Database VI



Virus Signature Database VII

- NAI/McAfee (DAT files, uncompressed)
  - Size on 2001-09-10: 1.898.159 Bytes
  - Size on 2002-09-09: 2.226.803 Bytes
  - Increase: About 329 KB last year!
    - About 27 KB a month or 14,8 % a year
    - For a period of more than 4 months, the DAT size was decreasing rather than increasing... due to a major clean-up of all virus definition (less exact detection)

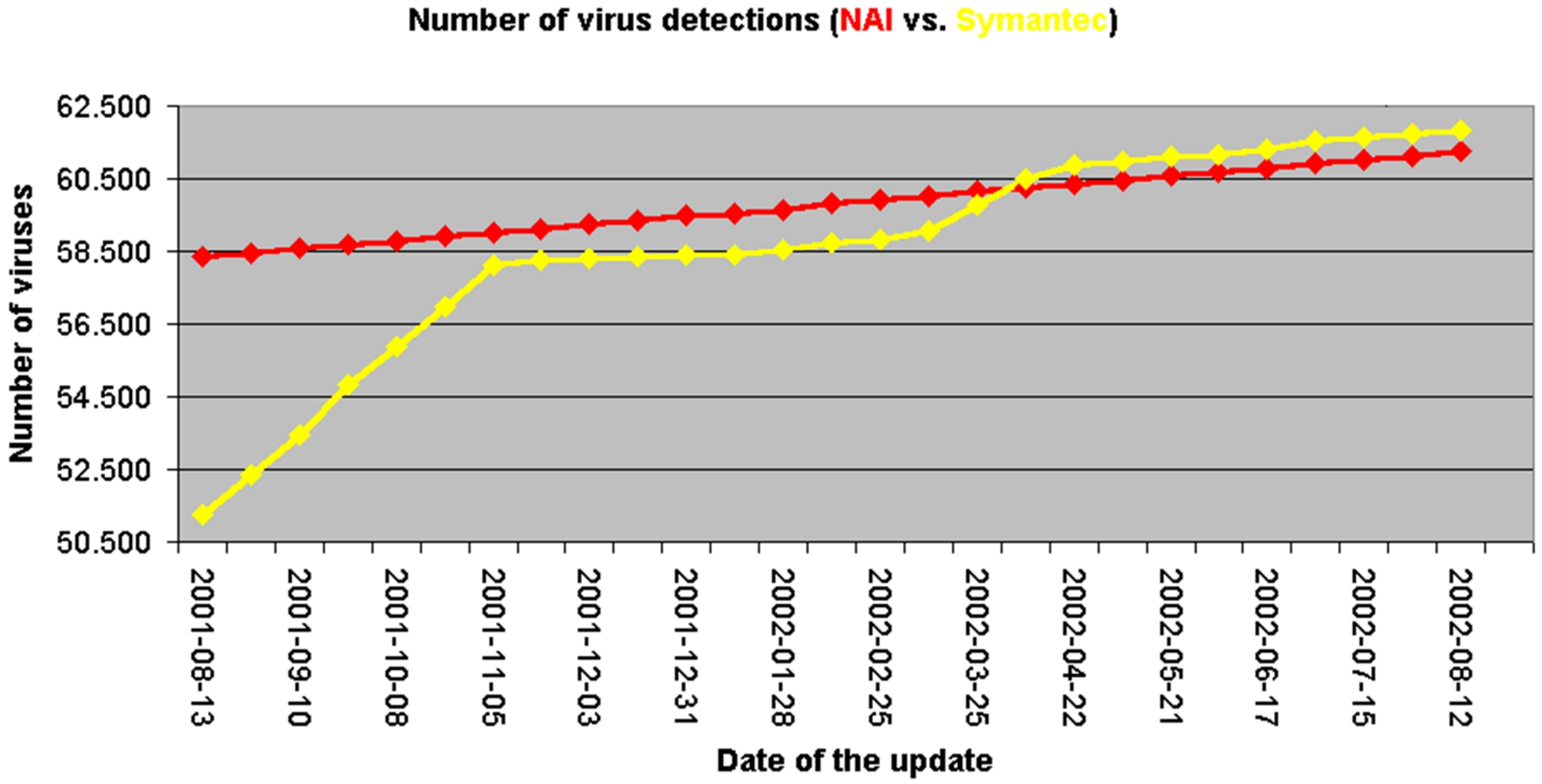
Virus Signature Database VIII

- Norman Virus Control (Main scan DLL, cmd-line scanner and full virus database)
  - Size on 2001-09-10: 1.259.267 Bytes
  - Size on 2002-09-09: 1.374.790 Bytes
  - Increase: Only 115.523 Bytes last year!
    - About 9,6 KB a month or 8,5 % a year
    - Nearly the same detection rate as all the other scanners! And with version 5.40 it will be < 1 MB

Number of Virus Detections I

- The number of viruses a program claims to detect is often PR-driven - the current range in our mid-September 2002 testset shows numbers between 27.000 and 73.000 “detectable viruses”
- An interesting point is actually, how Symantec got a much higher number than McAfee now (see the following slide)

Number of Virus Detections II



Speed differences

- Actually, most anti-virus programs are still as fast as one year ago, therefore, the new virus detection has not decreased the speed
- But there are a few update peaks, where the speed was slowing down a lot, but returned with the next update (likely due to adding detection of complex polymorphic viruses)

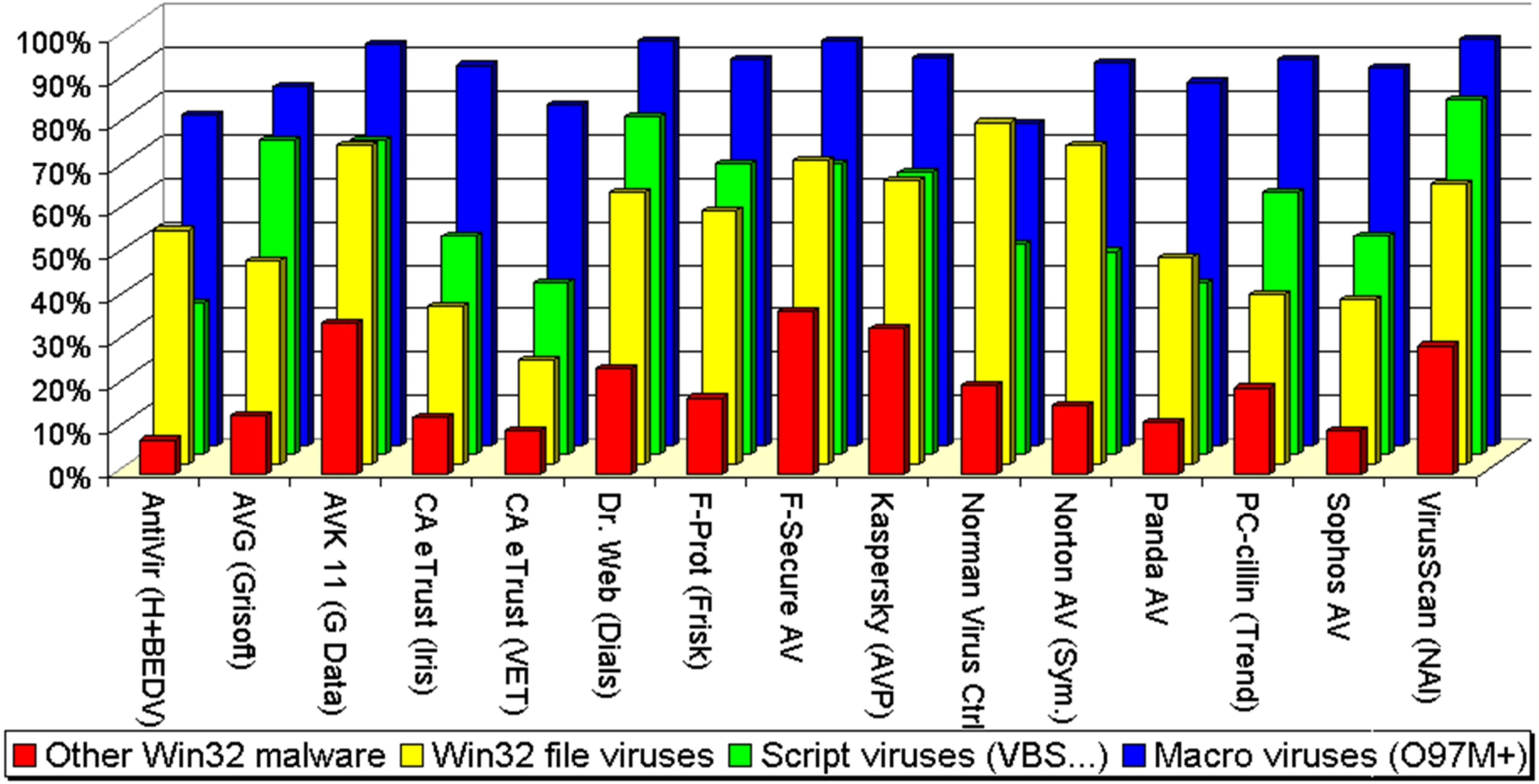
Archived and Compressed Files

- A few new archive formats were added to a small number of programs, but we did not saw dramatic changes at all
- One program (NAI) had an increasing score on compressed files in a few signature updates without any engine changes (Reason: detection routines now looks more on “uncompressable“ malware parts)

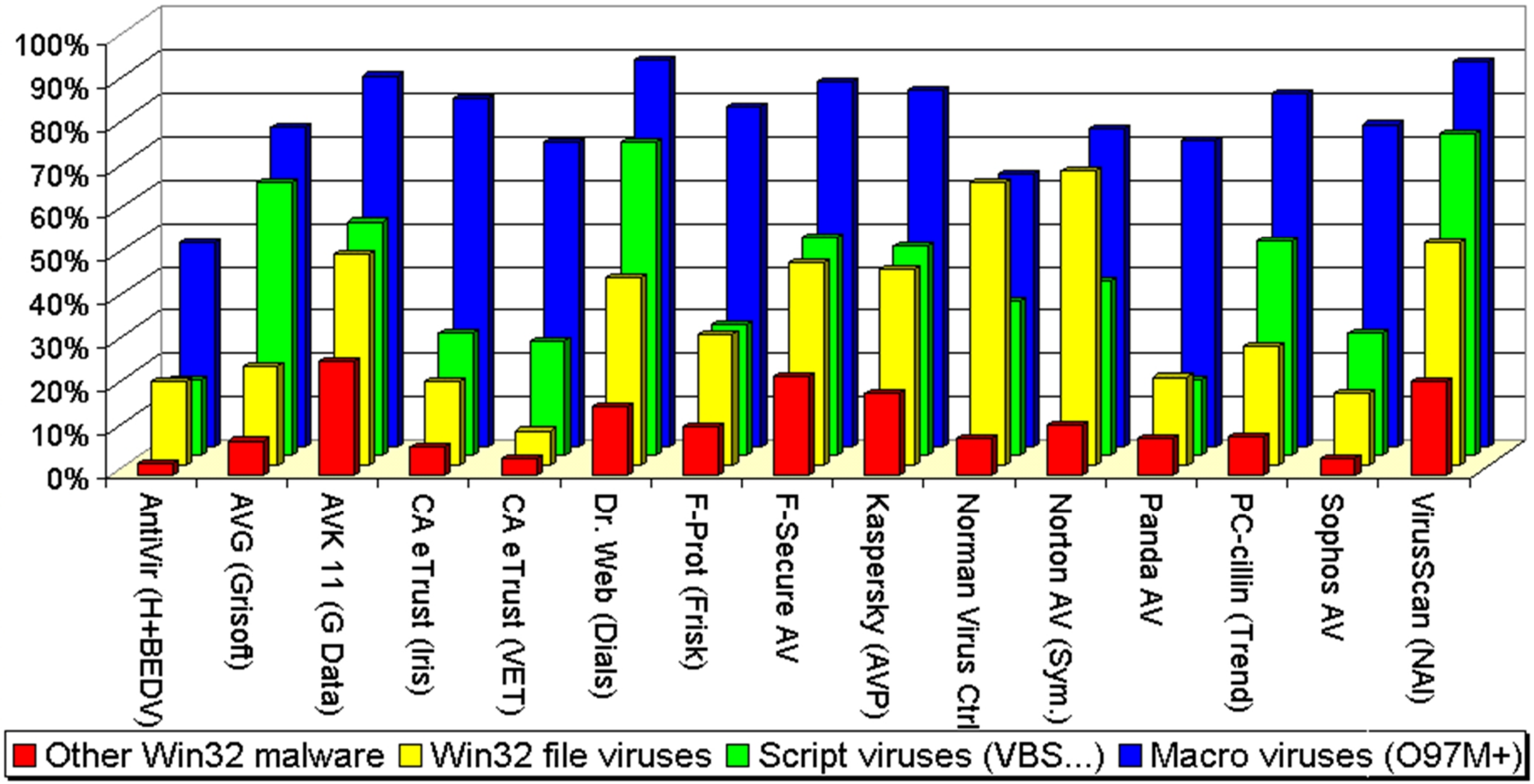
Detection Scores I

- OK, now to the most interesting part…
- Actually, we have quite a lot of data... I‘ve just picked out one test (out of 27 performed) which has also been used for an av test in the German c‘t magazine (AV-Test-ID 2002-05)
- Three and six month old scanners were used for a test performed mid-April 2002

Detection Scores II (3 months)



Detection Scores III (6 months)



Detection Scores IV

- Summary for three months old scanners I
  - Quite good detection of macro viruses
    - At least 74%, best detection was 94% with an average of 86,5%
  - Still good script virus detection rates
    - Worst program detected only about 35%, but the best one found 81,5%, average was 58%

Detection Scores V

- Summary for three month old scanners II
  - Relatively poor detection of Win32 file viruses
    - 24% for the worst program, but a very good rate for the best program (78,5%), average was 55,5%
  - Extremely bad detection of other Win32 malware like trojans and backdoors
    - The best program detected 37%, but the worst only 7,5%, the average result was 20%

Detection Scores VI

- Summary of six month old scanners
  - Detection rates dropped significantly for a very high number of tested av programs
  - But there are still a few ones with a very good detection of both macro and script viruses
  - However, nearly all performed quite poor on Win32 viruses and especially on other Win32 malware (developers need to do something here)

Summary I

- Databases of all scanners are increasing fast, we need to stop this or we see 10 MB virus definition files at the end of next year!
  - Developers need to “compress” all virus signatures better by replacing old virus patters with more generic ones - esp. for DOS viruses
- Numbers like „detectable viruses“ does not show anything

Summary II

- Heuristic and generic detection for macro viruses and script malware is very good and for Win32 viruses is OK from what we can expect
- There are still improvements needed for other Win32 malware in all programs!

Retrospective Testing

- Are there any…

QUESTIONS?