A Simulation Based Analysis of Logico-Probabilistic Reasoning Systems

PAUL THORN AND
GERHARD SCHURZ, DÜSSELDORF

LOGICCC EUROCORES PROGRAM
OF THE ESF AND DFG.

LP Reasoning Systems

- Systems of inference with conditional that (may) characterize reasonable (but not necessarily valid) inference, assuming:

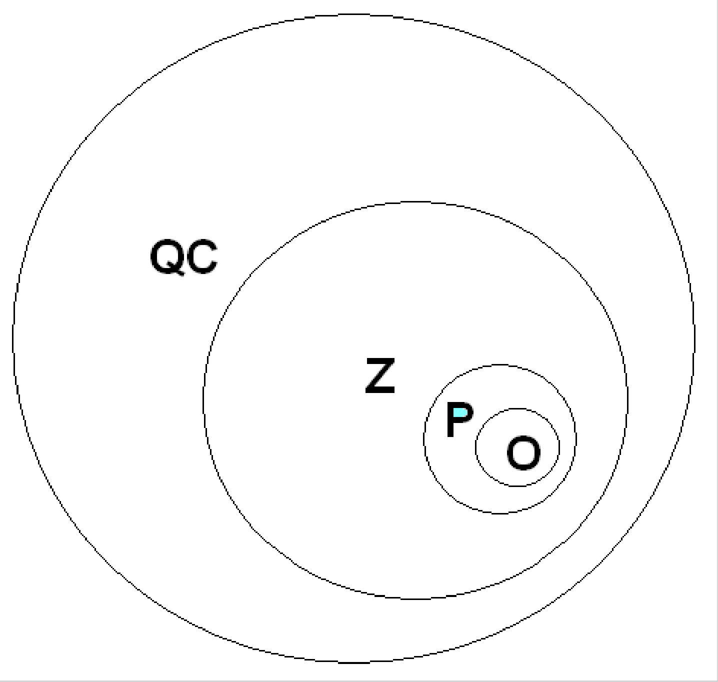
  We interpret "$C \Rightarrow D$" as expressing that $P(D|C)$ is high.

  We interpret "$C \Rightarrow_r D$" as expressing that $P(D|C) \geq r$.

Our Project

- See which of four well known LP-systems provides the best balance of **reward** versus **risk**, via computer simulations.

- $\mathbf{O} \subset \mathbf{P} \subset \mathbf{Z} \subset \mathbf{QC}$.



System **O**

REF: $\vdash_o A \Rightarrow A$

LLE: if $\vdash A \leftrightarrow B$, then $A \Rightarrow C \vdash_o B \Rightarrow C$

RW: if $\vdash B \rightarrow C$, then $A \Rightarrow B \vdash_o A \Rightarrow C$

VCM: $A \Rightarrow B \wedge C \vdash_o A \wedge B \Rightarrow C$

XOR: if $\vdash \neg(A \wedge B)$, then

$A \Rightarrow C, B \Rightarrow C \vdash_o A \wedge B \Rightarrow C$

WAND: $A \Rightarrow B, A \wedge \neg C \Rightarrow \bot \vdash_o A \Rightarrow B \wedge C$

O as an LP-System:

System **O** almost corresponds to the consequence relation **SP** (Strict Preservation):

$A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \mid\!\!\!-\!\!\!-_{\mathbf{SP}} C \Rightarrow D$

iff

for all probability functions P:

$P(D|C) \geq \min(\{ P(B_i|A_i) : 1 \leq i \leq n \})$.

Reasoning by System **O**

If one has $A \Rightarrow_{0.9} B$ and $C \Rightarrow_{0.8} D$, and

$A \Rightarrow B, C \Rightarrow D \mid\!\!-_{\mathbf{O}} E \Rightarrow F$,

then system **O** licenses the conclusion

$E \Rightarrow_{0.8} F$.

VALID

System **P**

Is characterized by the rules of system **O** along with…

AND: $A \Rightarrow B,\ A \Rightarrow C \mid\!\!-_{\mathbf{P}} A \Rightarrow B \wedge C$

System P

REF, **LLE**, **RW** (as with **O**)

AND: $A \Rightarrow B, A \Rightarrow C \mid\!\!-_{\mathbf{P}} A \Rightarrow B \wedge C$

CC: $A \Rightarrow B, A \wedge B \Rightarrow C \mid\!\!-_{\mathbf{P}} A \Rightarrow C$

CM: $A \Rightarrow B, A \Rightarrow C \mid\!\!-_{\mathbf{P}} A \wedge B \Rightarrow C$

OR: $A \Rightarrow C, B \Rightarrow C \mid\!\!-_{\mathbf{P}} A \wedge B \Rightarrow C$

P as an LP-System:

$A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \mid\!\!\!-\!\!\!-_{\mathbf{P}} C \Rightarrow D$

iff

for all probability functions P:

$U(D|C) \leq \Sigma\{\, U(B_i|A_i) : 1 \leq i \leq n\}$, where

$U(A|B) = 1 - P(A|B)$.

Reasoning by System **P**

If one has $A \Rightarrow_{0.9} B$ and $C \Rightarrow_{0.8} D$, and

$A \Rightarrow B,\ C \Rightarrow D \mid\!\!\!-_{\mathbf{P}} E \Rightarrow F$,

then system **P** licenses the conclusion

$E \Rightarrow_{0.7} F$.

VALID

System Z

- Is a strengthening of system **P** that permits contraposation and subclass inheritance (desfeasibly).

Contraposition: $A \Rightarrow B \mid\!\!\sim_Z \neg B \Rightarrow \neg A$,

but not $A \Rightarrow B, \neg B \Rightarrow A \mid\!\!\sim_Z \neg B \Rightarrow \neg A$.

Subclass Inheritance: $A \Rightarrow B \mid\!\!\sim_Z A \wedge C \Rightarrow B$,

but not $A \Rightarrow B, A \wedge C \Rightarrow \neg B \mid\!\!\sim_Z A \wedge C \Rightarrow B$.

Z Partition (System Z)

$\{ A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \}$ *tolerates* $C \Rightarrow D$

iff

$\{ A_1 \rightarrow B_1, \ldots, A_n \rightarrow B_n, C \wedge D \}$ is consistent.

The Z partition for $\Gamma = \{ A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \}$ is the set $(\Gamma_1, \ldots, \Gamma_k)$, where Γ_1 is the set of elements of Γ that are *tolerated* by Γ, and Γ_2 is the set of element of $\Gamma - \Gamma_1$ that are tolerated by $\Gamma - \Gamma_1$, etc.

Z-rank of a Conditional

$z_{\Gamma}(A \Rightarrow B) = n$ iff $A \Rightarrow B \in \Gamma_n$.

Z-entailment

$\Gamma \mid\!\!\!— _{\mathbf{z}} C \Rightarrow D$

iff

$\exists n$: { $A \Rightarrow B \mid A \Rightarrow B \in \Gamma$ & $z_{\Gamma}(A \Rightarrow B) \geq n$ } *tolerates* $C \Rightarrow D$ and does not *tolerate* $C \Rightarrow \neg D$.

Z as an LP-System:

- System **Z** preserves probability (in the manner of system **P**) relative to a restricted set of probabilty functions.

Reasoning by System **Z**

If one has $A \Rightarrow_{0.9} B$ and $C \Rightarrow_{0.8} D$, and

$A \Rightarrow B, C \Rightarrow D \mid\!\!\!-\!\!-_{\mathbf{Z}} E \Rightarrow F$,

then system **Z** licenses the conclusion

$E \Rightarrow_{0.7} F$ (under certain conditions).

INVALID

System **QC**

$$A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \vdash_{\mathbf{QC}} C \Rightarrow D$$

iff

$$A_1 \supset B_1, \ldots, A_n \supset B_n \vdash C \supset D$$

QC as an LP System:

A **QC** inference, $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n \mid\!\!\!-_{QC} C \Rightarrow D$, preserves probability (in the manner of system **P**) just in case P(C) is much greater than $\Sigma\{\ U(B_i|A_i) : 1 \leq i \leq n\ \}$.

Reasoning by System QC

If one has $A \Rightarrow_{0.9} B$ and $C \Rightarrow_{0.8} D$, and

$A \Rightarrow B, C \Rightarrow D \vdash_{QC} E \Rightarrow F$,

then system **QC** licenses the conclusion

$E \Rightarrow_{0.7} F$ (under certain conditions).

INVALID

Our Project

- Evaluate which system provides the best balance of **reward** versus **risk**, by means of computer simulations.

Simulations: Basic Procedure

1. Generate a true probability distribution, and a set of *base conditionals* (premises), with associated lower probability bounds.

2. Determine which *derived conditionals* (conclusions) each system is able to derive from the given base conditionals (including lower probability bounds for these conclusions).

3. Assign scores to each system, by comparing the lower probability bounds for the derived conditionals with the true probabilities.

Language Restriction

We adopted a simple language with four binary variables: a, b, c, and d

We only considered conditionals with conjuctive antecedents and consequents (and no repetition of atoms), yielding **464** conditionals:

- 64 of the form $\pm x \wedge \pm y \wedge \pm z \Rightarrow \pm w$
- 96 of the form $\pm x \wedge \pm y \Rightarrow \pm z$
- 48 of the form $\pm x \Rightarrow \pm y$
- 96 of the form $\pm x \wedge \pm y \Rightarrow \pm z \wedge \pm w$
- 96 of the form $\pm x \Rightarrow \pm y \wedge \pm z$
- 64 of the form $\pm x \Rightarrow \pm y \wedge \pm z \wedge \pm w$

The probability distributions…

were fixed by randomly setting the following sixteen values:

$P(a)$,

$P(b|a)$, $P(b|\neg a)$,

$P(c|a \wedge b)$, $P(c|a \wedge \neg b)$, $P(c|\neg a \wedge b)$, $P(c|\neg a \wedge \neg b)$, $P(d|a \wedge b \wedge c)$, $P(d|a \wedge b \wedge \neg c)$, $P(d|a \wedge \neg b \wedge c)$, $P(d|a \wedge \neg b \wedge \neg c)$, $P(d|a \wedge b \wedge c)$, $P(d|\neg a \wedge \neg b \wedge \neg c)$, $P(d|\neg a \wedge \neg b \wedge c)$, $P(d|\neg a \wedge b \wedge c)$, and $P(d|\neg a \wedge b \wedge \neg c)$.

Scoring I

The *advantage-compared-to-guessing* score for derived conditionals:

$$\text{Score}_{AvG}(\, C \Rightarrow_r D, \mathrm{P}) = 1/3 - |r - \mathrm{P}(D|C)|.$$

Scoring II

The *subtle-price-is-right* score for derived conditionals:

$$\text{Score}_{\text{sPIR}}(\, C \Rightarrow_r D, \mathcal{P}) = r, \text{ if } r \leq \mathcal{P}(D|C),$$

$$= \mathcal{P}(D|C) - r, \text{ otherwise.}$$

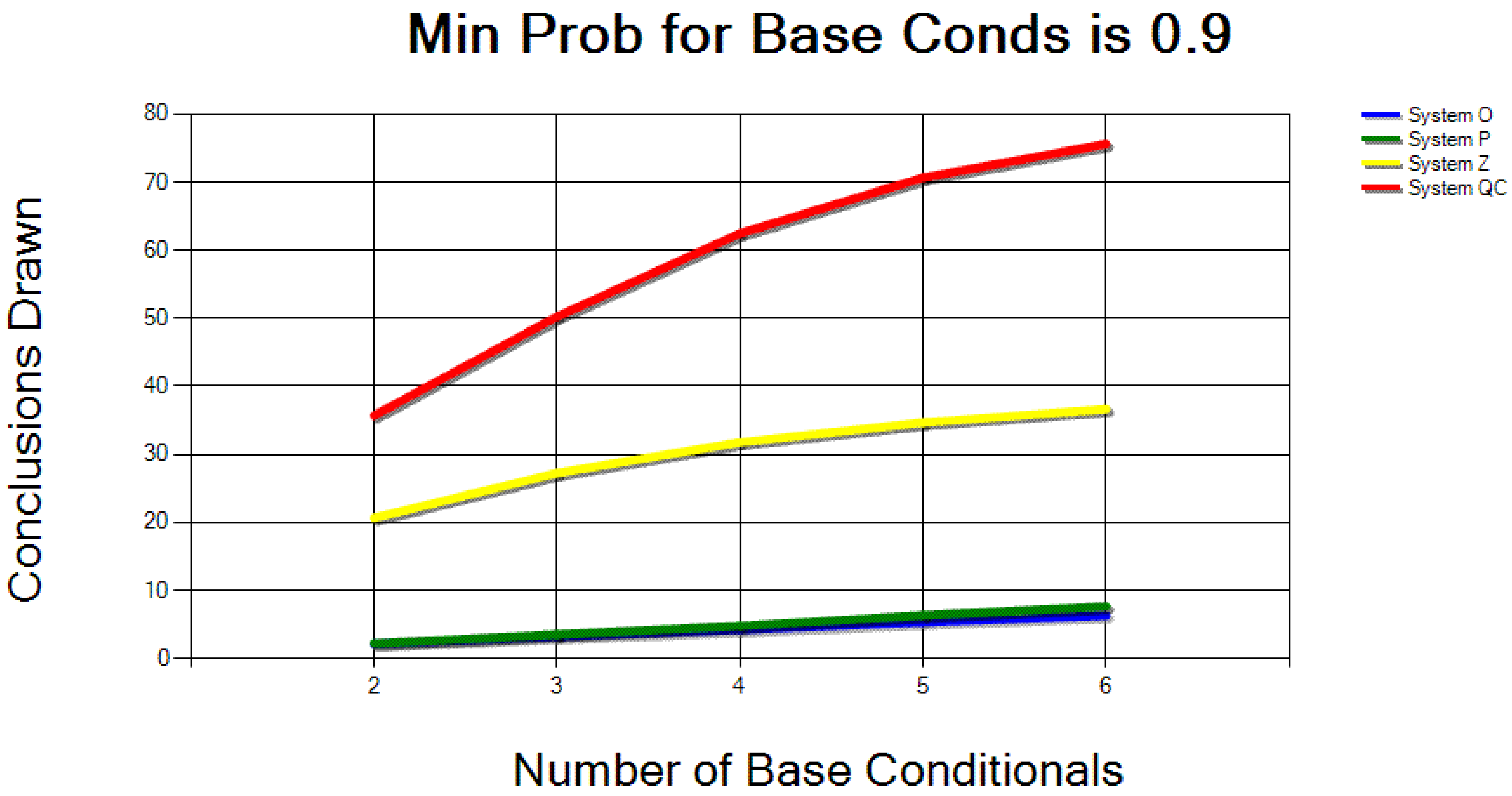
Simulations with Randomly Selected Base Conditionals

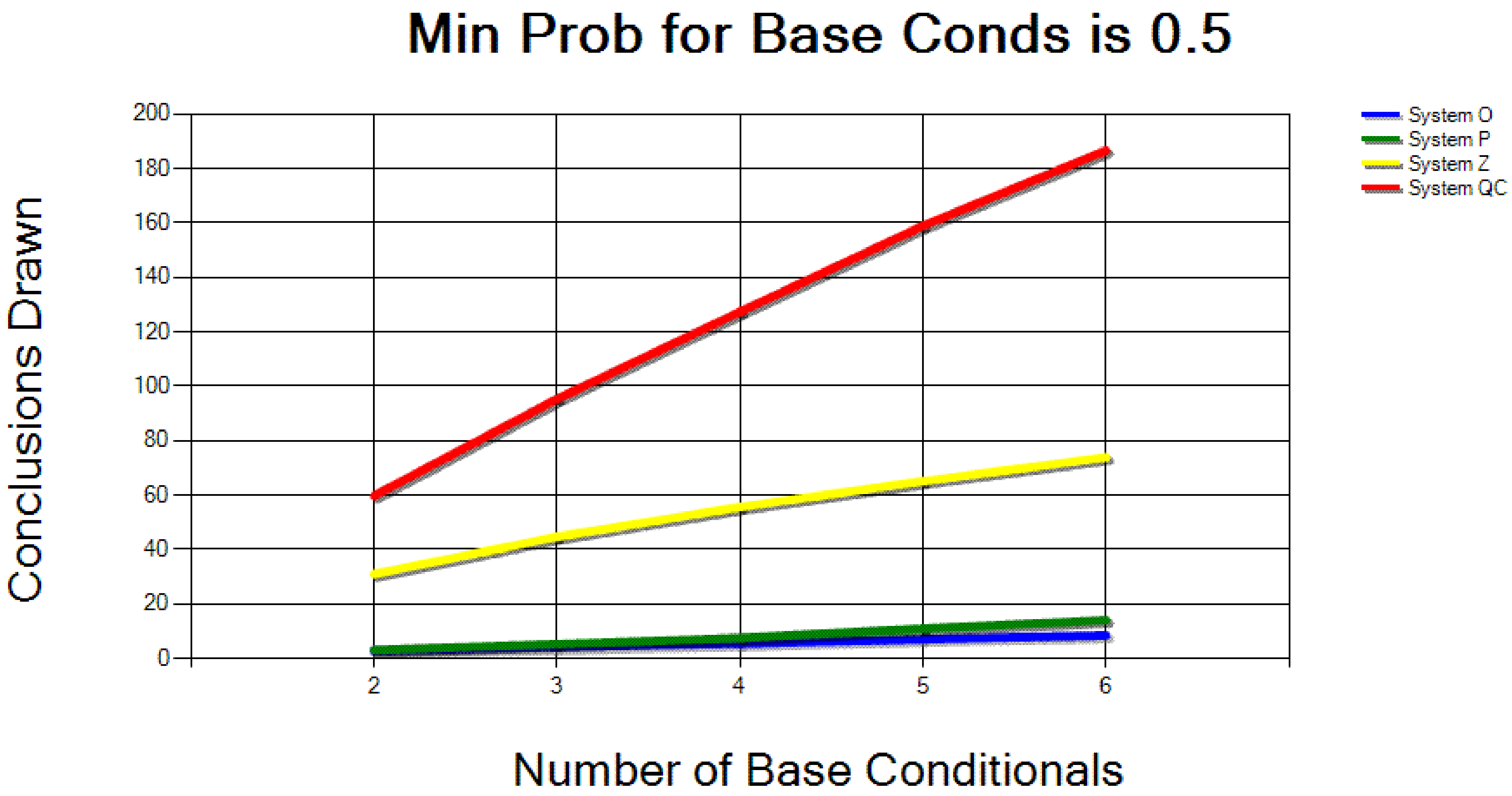
- P is determined at random, for each run.
- **n** base conditionals are selected. (**n** = 2, 3, 4, 5, or 6)
- The base conditionals are selected at random, from among those conditionals whose associated probability was at least **s**. (**s** = 0.5, 0.6, 0.7, 0.8, 0.9, or 0.9999)
- We call **s**: the minimum probability for base conditionals.
- We ran each combination of **s** and **n** one thousand times.

On to the Data

Observations (Table 1)

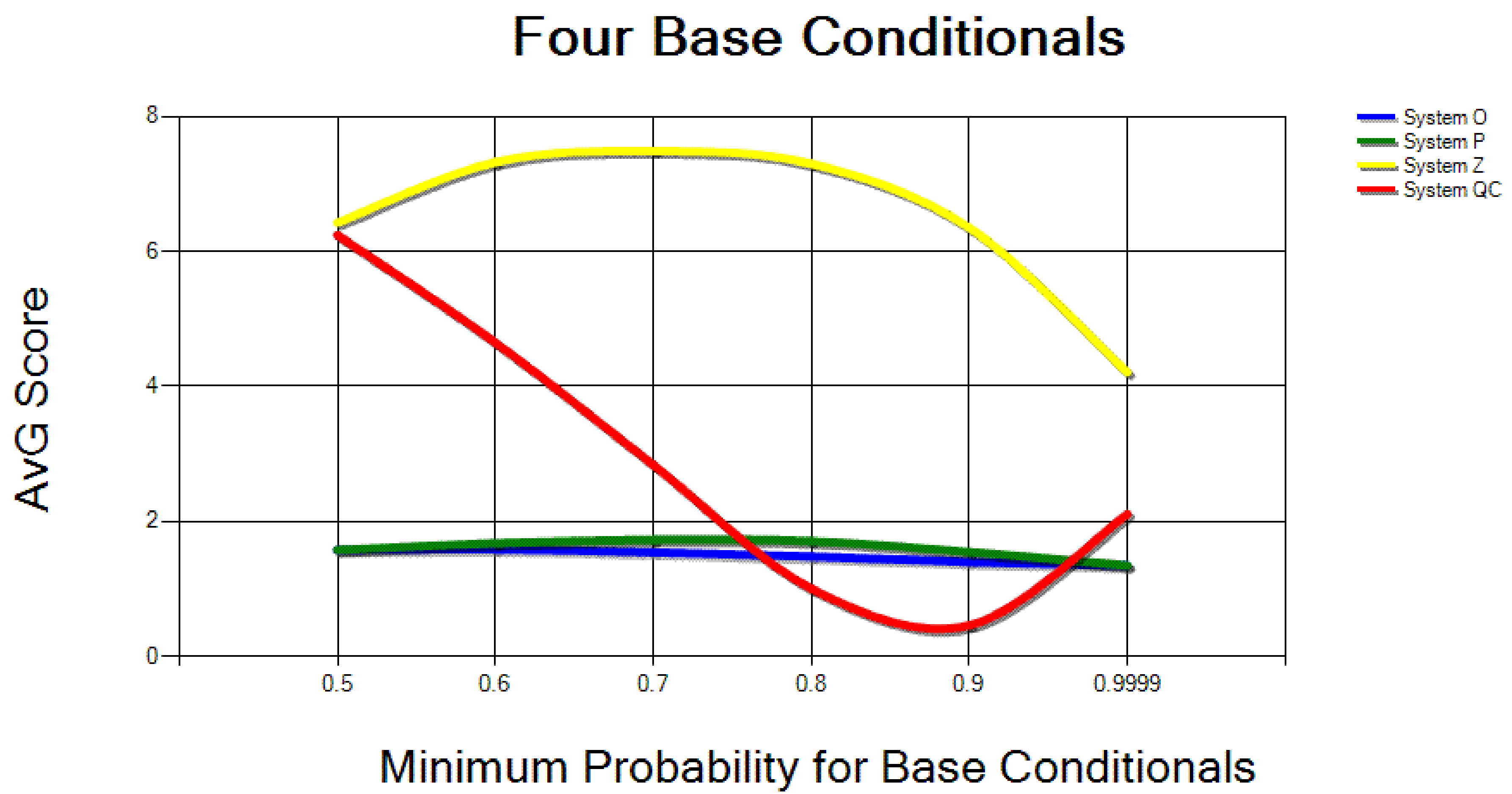
- 1. Systems **O** and **P** are very conservative, and licences very few inferences.


Min Prob for Base Conds is 0.9
Conclusions Drawn
80
70
60
50
40
30
20
10
0
2
3
4
5
6
Number of Base Conditionals
System O
System P
System Z
System QC

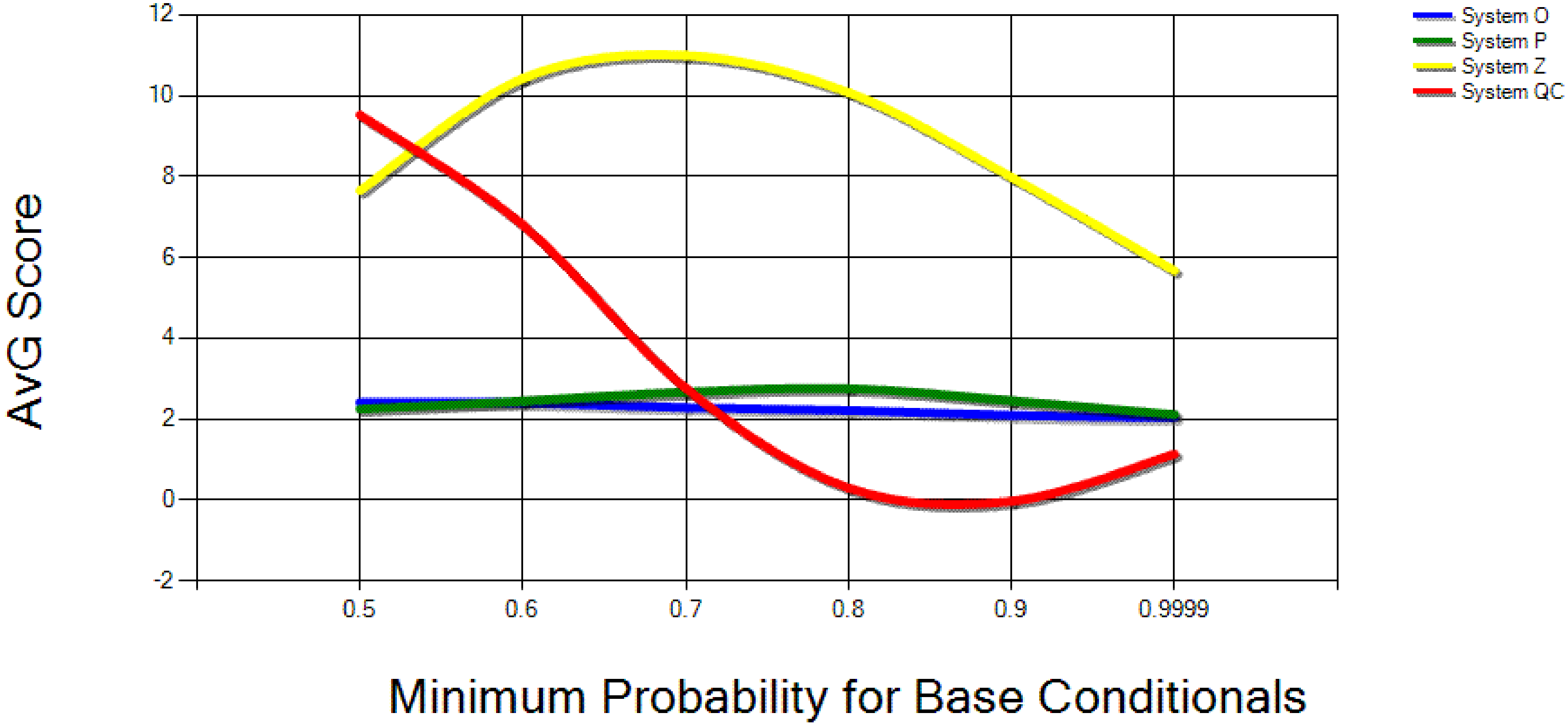

Min Prob for Base Conds is 0.5
Conclusions Drawn
200
180
160
140
120
100
80
60
40
20
0
2
3
4
5
6
Number of Base Conditionals
System O
System P
System Z
System QC

Observations (Table 2)

2. System **Z** almost always outperformed the other systems by the AvG measure (even though the AvG measure sometimes punishes non-errors).


Four Base Conditionals
AvG Score
Minimum Probability for Base Conditionals
8
6
4
2
0
0.5
0.6
0.7
0.8
0.9
0.9999
System O
System P
System Z
System QC

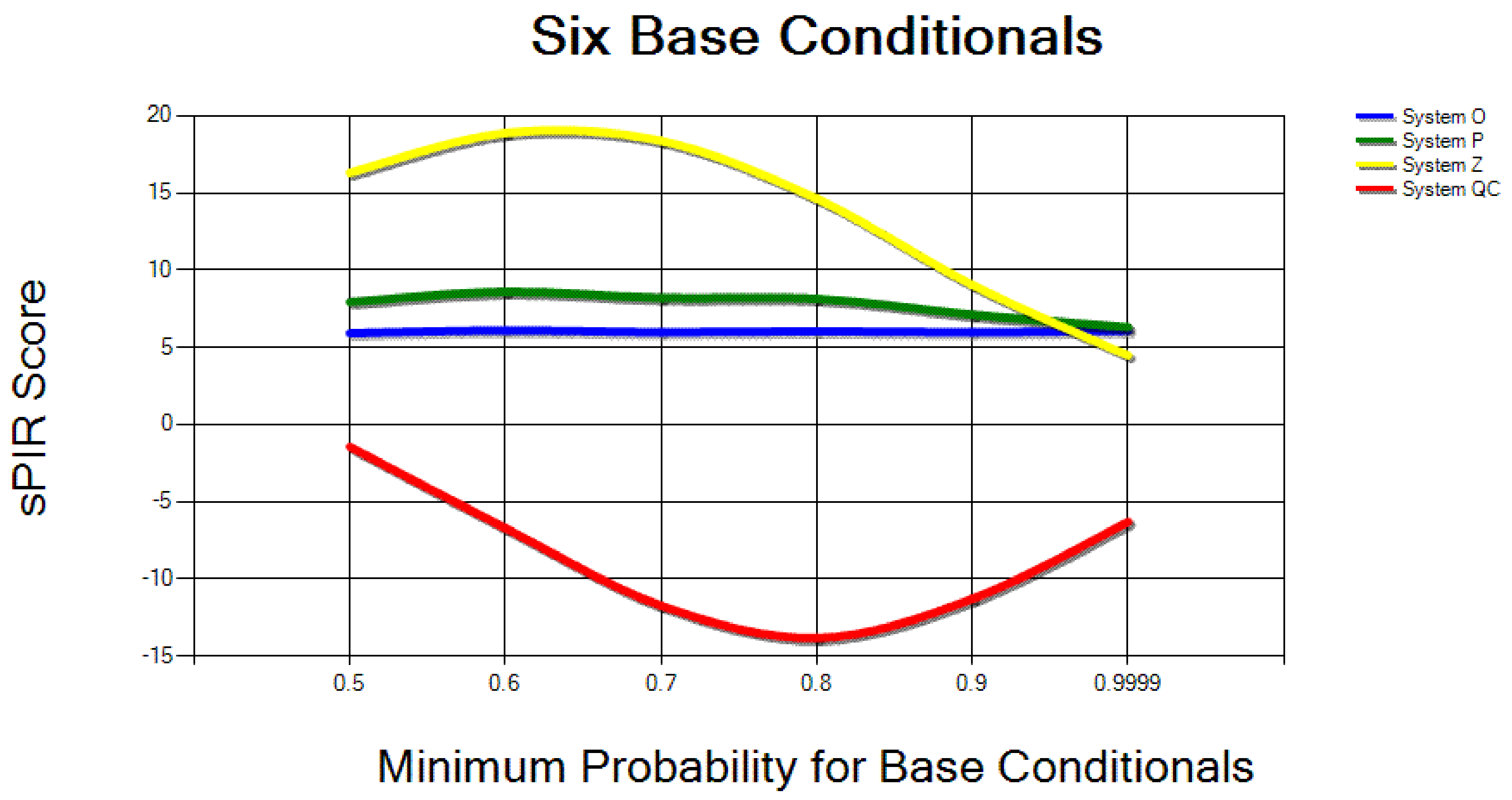
Six Base Conditionals



Observations (Table 2)

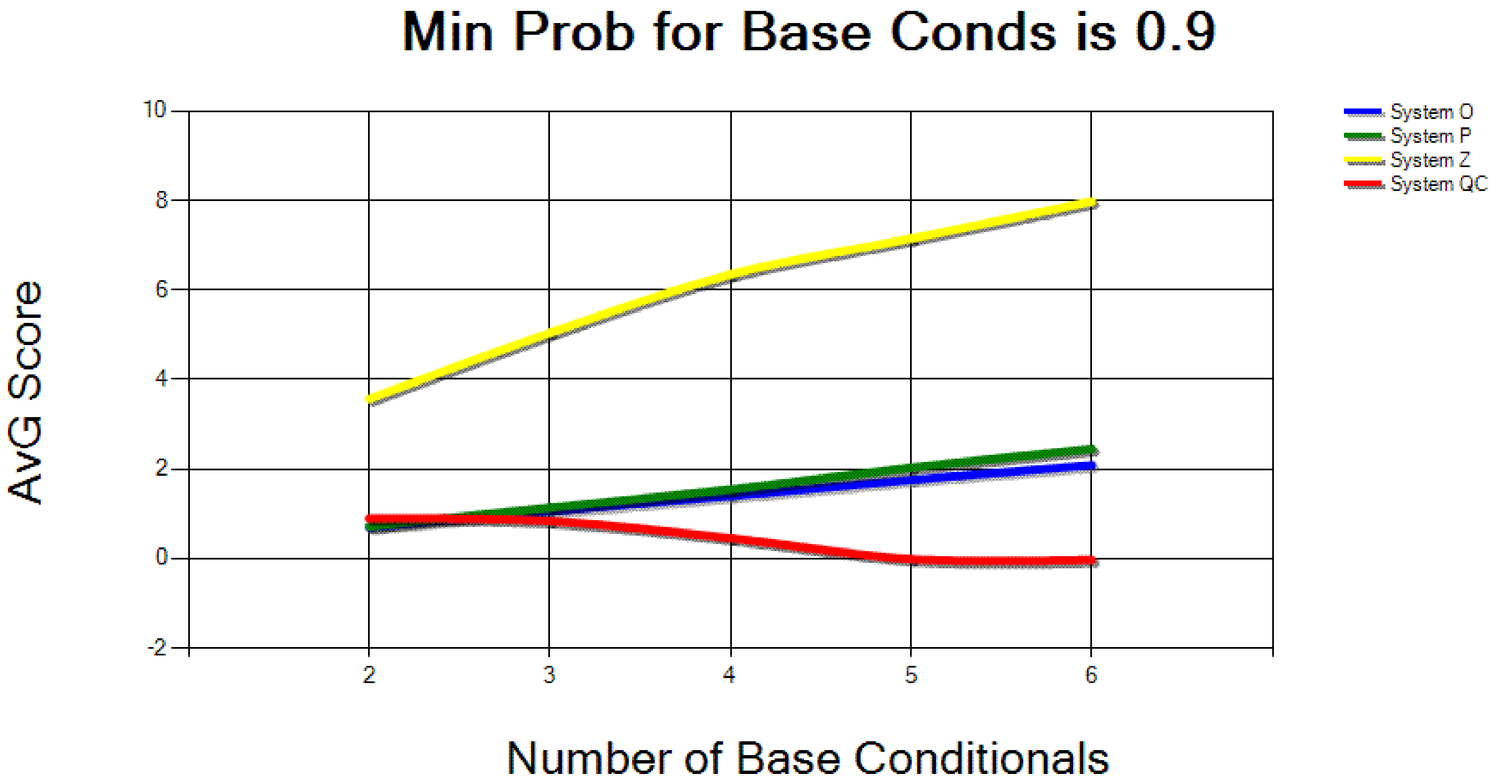
3. System **Z** generally achieved positive sPIR scores, and outscored all of the systems by this measure (save in the case where s = 0.9999).

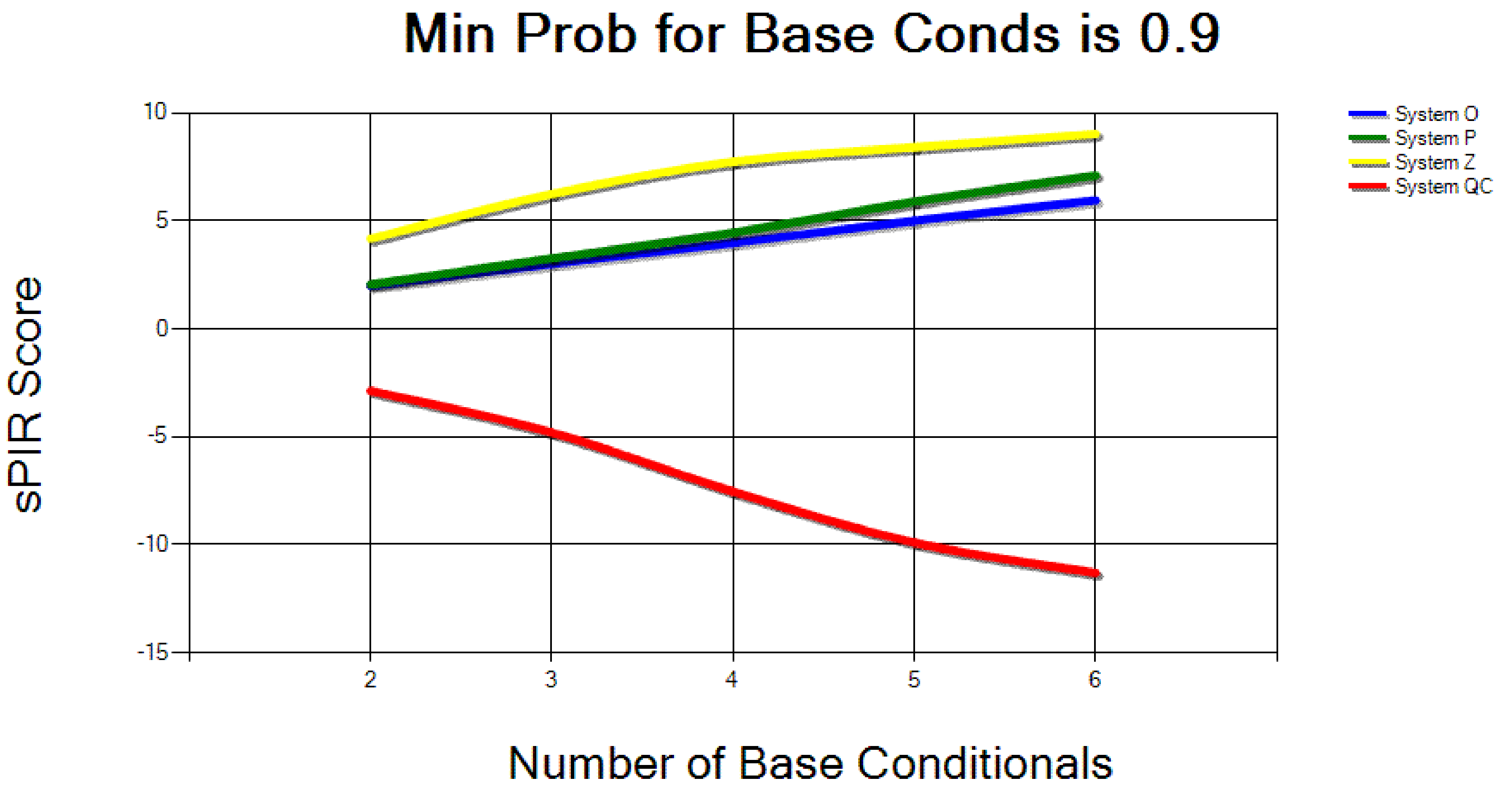
4. **QC** usually obtained negative sPIR scores, due to frequent errors.


Six Base Conditionals
20
15
10
5
0
-5
-10
-15
0.5
0.6
0.7
0.8
0.9
0.9999
System O
System P
System Z
System QC
sPIR Score
Minimum Probability for Base Conditionals

Observations (Table 3)

5. Where s is fixed, increasing the number of base conditionals tends to increase the scores (for all measures) for systems **O**, **P**, and **Z** (but not for **QC**).

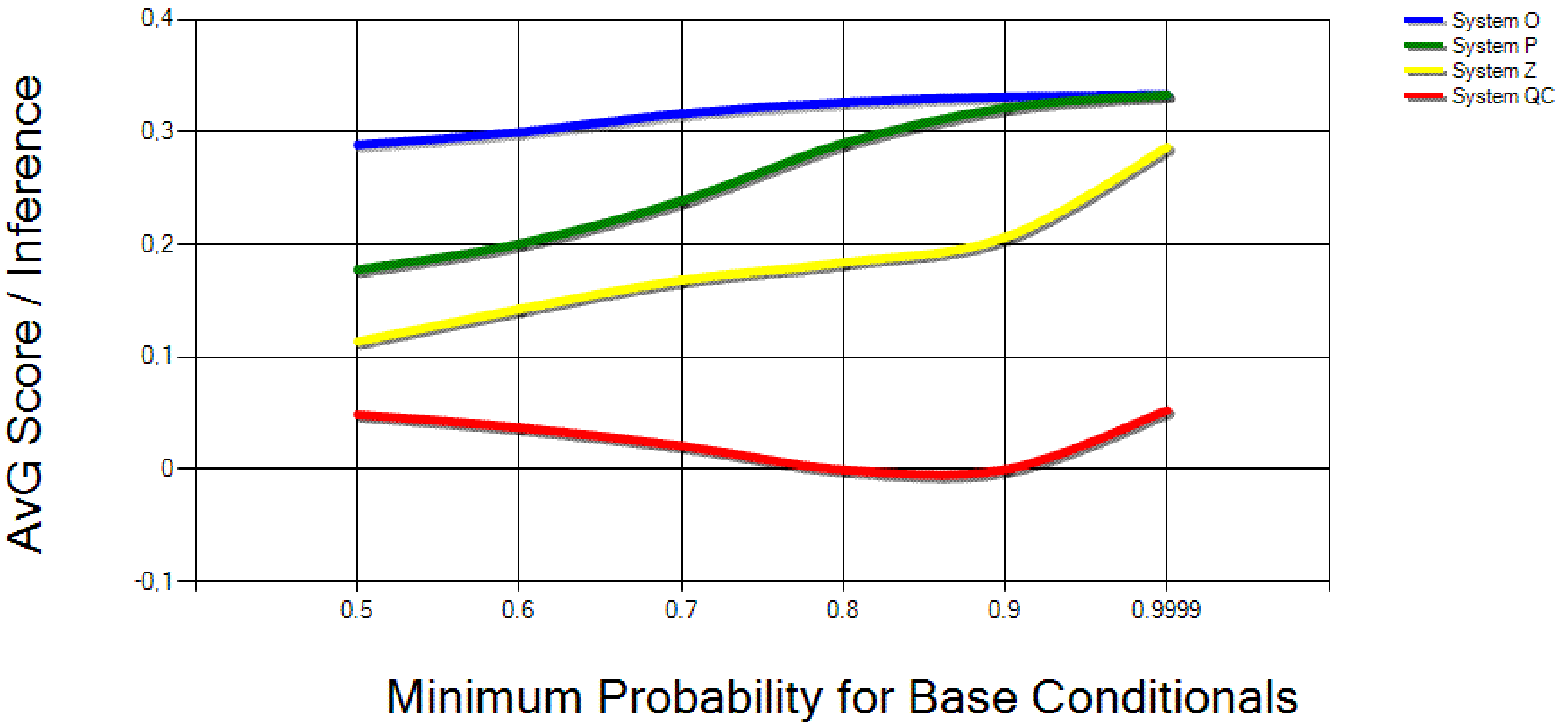

Min Prob for Base Conds is 0.9
AvG Score
Number of Base Conditionals
10
8
6
4
2
0
-2
2
3
4
5
6
System O
System P
System Z
System QC


Min Prob for Base Conds is 0.9
System O
System P
System Z
System QC
10
5
0
-5
-10
-15
2
3
4
5
6
sPIR Score
Number of Base Conditionals

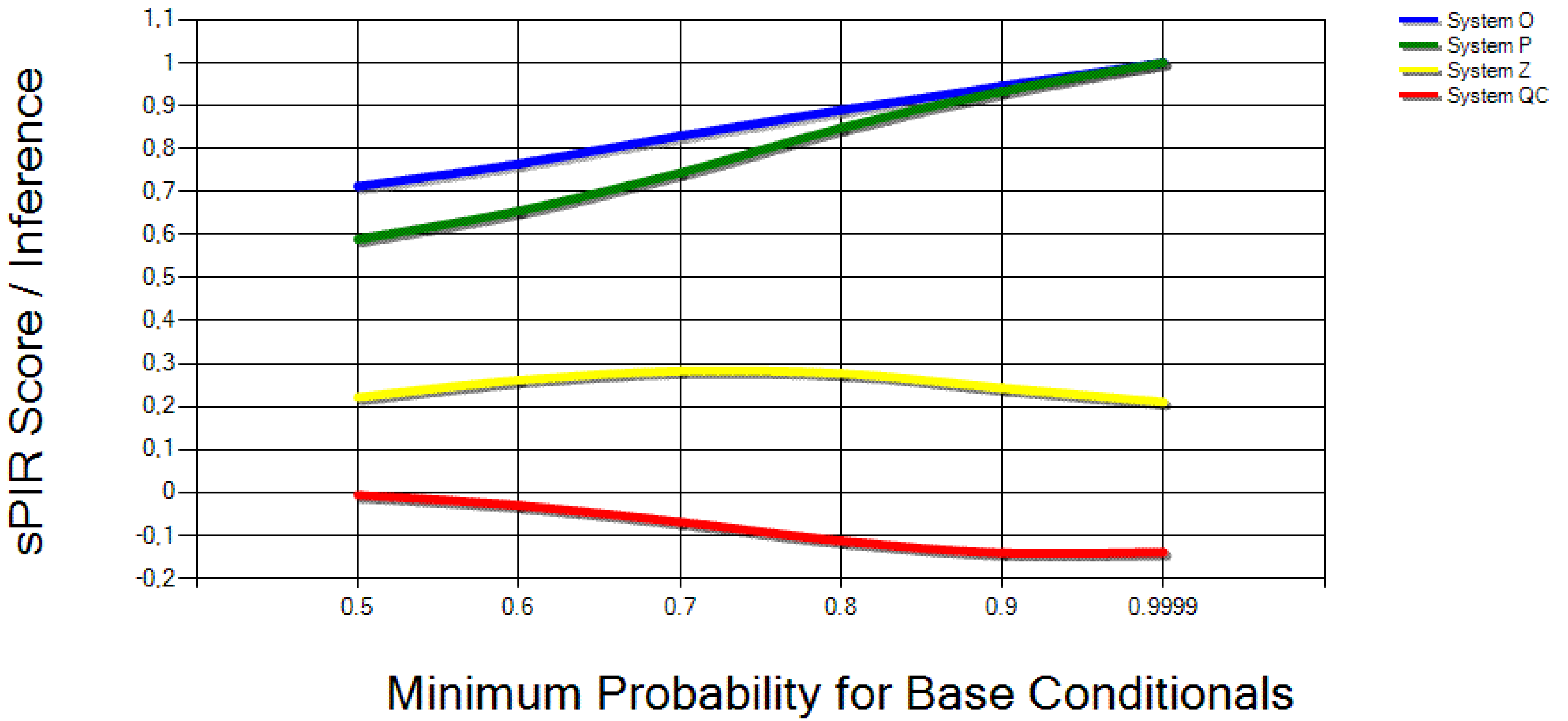
Observations (Table 4)

6. For systems **O**, **P**, and **Z**, the score earned per inference *tends* to increase for higher values of s.

Five Base Conditionals



Five Base Conditionals



Conclusions

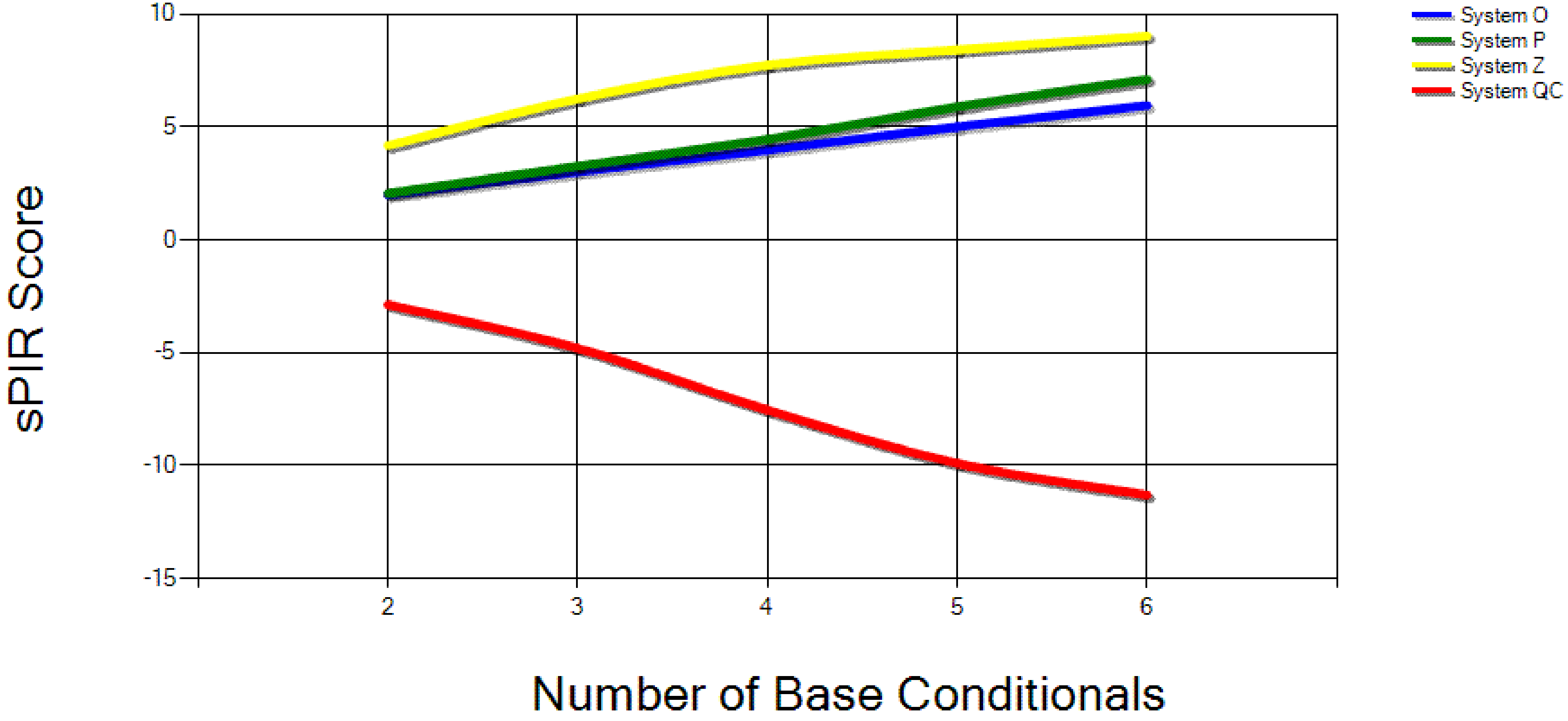
1. System **P** offers the same security as system **O**, and licences more inferences.

2. System **Z** licences far more inferences than system **P**, and generally infers lower probability bounds that are close to the true probability values (which is revealed by AvG scoring).

Conclusions

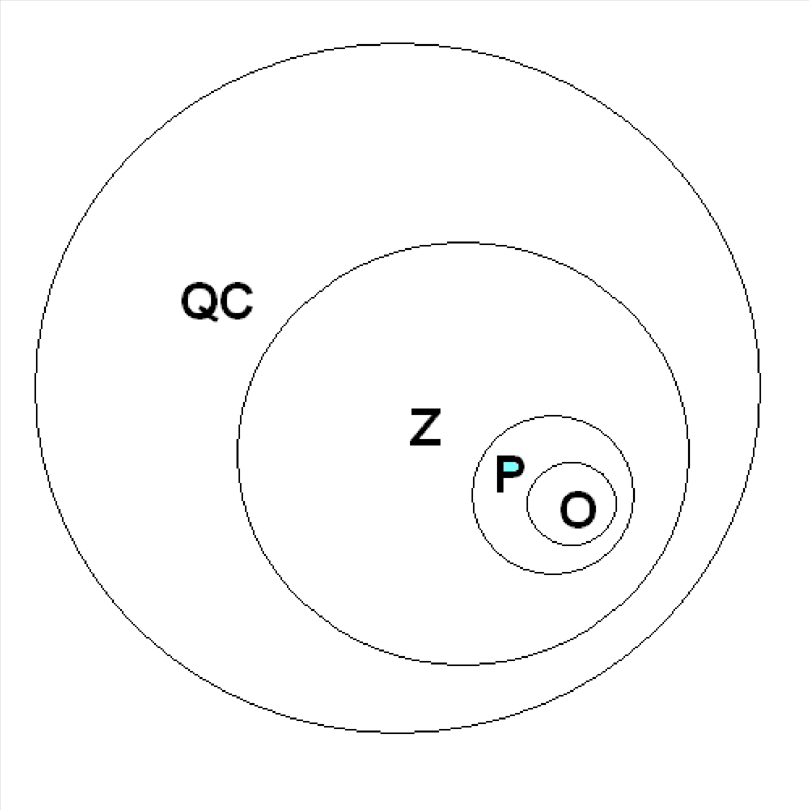
3. **QC** inferences are too risky, and very few of the inferences sanctioned by **QC**, and not by **Z**, should be made. (Consider the **QC–Z** inferences.)

Min Prob for Base Conds is 0.9



Conclusions

4. Of the four systems, system **Z** offers the best balance of safety versus inferential power.



The End.