Unicode 6.0 versus IDNA2008

Apps Area Meeting – IDNA 79
John Klensin
IDNA2008
Designed to be Unicode
Version-independent

• Special “backward-compatibility” provision in case Unicode makes an incompatible change that affects IDNA
• WG expected this provision would never, or almost-never, be used
Unicode 6.0

- Code tables now complete, book text being polished
- Three incompatible changes of properties that alter validity of characters for IDNA
- These really are corrections – the prior classifications were wrong.
- The characters are *almost certainly* not in DNS use today. But we can never prove that.
IETF Choice about IDNA

• Two characters move from DISALLOWED to PVALID
  – Almost certainly ok to just let the categories change

• One character moves from PVALID to DISALLOWED
  – Very bad idea in principle
  – In practice, this one is safe
Choice 1: Do nothing

– Character DISALLOWED in Unicode 6-conforming implementations, PVALID in Unicode 5.2-conforming ones
– Probably no registrations actually affected
– IDNA practice aligns with correct classifications
Choice 2: Modify IDNA

- Preserve old behavior, keep character valid
- Adds clutter and opens up a class of special cases we hoped to not use
- Likely to upset community whom we promised stability once they went to IDNA2008
- Misalignment with accurate classifications

\textit{draft-faltstrom-5892bis}
How to Decide?

• This is actually the only hard question

• The answer may not make much difference this time

• But the next time may be important
The Nasty Surprise

• It is now clear that any IETF protocol that
  – Needs to compare Unicode strings
  – Needs to match Unicode strings to user expectations

• Will need to deal with this type of change
  – The problem is not just theoretical

• The “Unicode Instability Considerations” section of future protocol specs?