Subject:	Script and Script_Extension stability policies
From:	Mark Davis
Date:	2015-04-23
Link:	http://goo.gl/luUlyw

Proposal

Ask the officers to approve the following stability policy:

> If a character's script extensions property value set contains an implicit script property value:
> (a) the set must have only one element, and
> (b) that element must be the same as the character's script property value.
>
> *(The editorial committee can, of course, make improvements in the wording.)*

Background

While doing an invariant test for the new script invariant policies, I ran across the following bad cases, that we currently do not exclude on the stability policies page:

1. scx={Common, Arabic}
2. scx={Common} and sc=Arabic

Excluding these does not represent a restriction on the expressive power of what we do, it just makes sure that the properties are consistently defined. As it turns out, we have followed this policy back to when SCX was first introduced. We just have no formal guarantee that we will continue it.

Note: we provided conditions in http://www.unicode.org/review/pri277/, in settling on Policy B. If we flesh out the conditions there to account for the singleton case, we get something like the following:

The Script_Extensions property value for a character C, aka scriptExt(C), is a non-empty, unordered set whose elements $\{E_1, \dots E_n\}$ are Script property values, and where the following conditions hold:

1. *Where scriptExt(C) has exactly one element E:*
    a. *E = script(C),* **or**
    b. $E \neq$ script(C), **and** *E is explicit,* **and** *script(C) is either Common or Inherited*
2. *Where scriptExt(C) has more than one element:*
    a. *Each E_i is explicit,* **and**
    b. *scriptExt(C) contains script(C)* **if and only if** *script(C) is explicit*

Examples:

Rule	Allows		Disallows	
	scx	**sc**	**scx**	**sc**
1a	{Common}	Common	{Common}	Arabic
	{Arabic}	Arabic	{Common}	Unknown
			{Arabic}	Latin

2a	{Latin, Arabic}	*anything*	{Common, Arabic}	*anything*
2b	{Latin, Arabic}	Common	{Latin, Arabic}	Cyrillic
	{Latin, Arabic}	Arabic		