

MEEK Project Team chat 15.5.2009

08:04:02 <Troy> Morning, everybody

15/05/2009 08:04:43 <Erkki> Good morning. Håvard cannot join us to-day, but he was ready to publish the document.

15/05/2009 08:04:57 <mike> Good morning

15/05/2009 08:05:12 <Karl> Good morning, everybody.

15/05/2009 08:05:56 <Tero> Good morning.

15/05/2009 08:05:58 <Erkki> Unfortunately I forgot to replace the languages section, but I've now sent out a note with the new text I propose.

15/05/2009 08:06:07 <Karl> As I have expressed in my mail some minutes before: Regarding the entering of characters as "precomposed" or "decomposed" (p.11 3rd/4th paragraph), in my opinion this is a detail of implementation which encoding specific and operating system specific, thus it is not necessarily appropriate to include this fact in our paper ...

15/05/2009 08:06:21 <Karl> Which characters are listed to be supported, also does not depend of the number of atomic parts used for their representation in a specific encoding. Therefore, I propose something like:

15/05/2009 08:06:40 <Karl> Letters with multiple diacritics (e.g., many of the Lithuanian accented letters) could also be entered using the above extended dead key method.

15/05/2009 08:06:49 <Karl> Another method to enter composite characters would be to enter the base letter followed by the combining diacritical marks. This method could be more suitable for those users who are not familiar with the dead key method. ...

15/05/2009 08:07:03 <Karl> or some operating systems (Microsoft, in particular), it could also be implemented easier and possibly faster than the support for multiple diacritics using the dead key method.

15/05/2009 08:08:01 <Karl> Also, in the last paragraph on p.12, simply drop teh "(precomposed)" in the fifth last line.

15/05/2009 08:09:07 <Erkki> Karl: I believe that we aren't making room for any other encoding than the 10646/Unicode.

15/05/2009 08:09:21 <Karl> read "For some operating ..." for "or some operating ..."

15/05/2009 08:11:00 <Erkki> Karl: That (precomposed) refers to what the standard states.

15/05/2009 08:11:40 <Erkki> ...since the basic mode only produces precomposed characters.

15/05/2009 08:13:53 <Karl> Erkki, while we mention 10646/Unicode as a character set, in my opinion we don't have to care about the encoding details themselves. The composed characters are valid 10646/Unicode characters anyway, whether they are precomposed or decomposed in their coded (binary) representation

15/05/2009 08:15:30 <Karl> I have no problem with retaining the "(precomposed)" in the last paragraph on p.12, as the term is used there in reference to SFS.

15/05/2009 08:15:39 <Erkki> Karl: Of course they are valid characters independently of the encoding. However, they cannot be encoded as precomposed.

08:16:08 <Erkki> ...and I believe this fact should be stated.

15/05/2009 08:16:45 <Erkki> What do the others think of this?=-

15/05/2009 08:17:49 <Karl> In my opinion, the \*encoding\* is precomposed/decomposed, but not the \*entering\*, which does not have to care about such encoding details (and we do not refer to such encoding details at other parts of our paper).

15/05/2009 08:18:29 <Troy> It depends on the context. If we are only discussing 6.1., then I am fine with dropping reference to the issue, if the text remains understandable. I guess we should edit the text, and see if we can make sense of it either way.

15/05/2009 08:19:15 <Troy> The reference might be needed to understand that such an issue must also be tackled. Afterall, we are writing a recommendation.

15/05/2009 08:20:54 <Erkki> In fact, I'd replace "and they'll thus have to be entered as ddecomposed characters" by "...be encioded as...". The text as it is now is a mistake.

15/05/2009 08:22:39 <Troy> That's correct.

15/05/2009 08:23:05 <Troy> I think my edits addressed this issue.

15/05/2009 08:23:44 <mike> This is way outside my area of expertise. However my view is that our core task is to state ways in which the keyboard should be layed out and the expected way(s) in which users are expected to type the characters on the keyboard.

15/05/2009 08:24:15 <Erkki> I'd also say "dead key method, but such characters normally would not be listed as being explicitly supported".

15/05/2009 08:24:25 <mike> If these different ways of typing the characters have implications for the encoding then we should state these - of not then forget encoding

15/05/2009 08:24:43 <mike> that shpuld be "if not"

15/05/2009 08:24:59 <mike> hat should be "should" 🙄🙄

15/05/2009 08:25:30 <mike> I give up on correcting myself 🙄

15/05/2009 08:25:33 <Erkki> Mike: If you enter the diacritics after the base character, the only feasible way would be to produce decomposed characters.

15/05/2009 08:26:17 <mike> This is what I thought - then these are the sort of facts that should be stated. I think we need some simple clear statements like this!

15/05/2009 08:27:21 <Karl> That depends on the operating system. It is, in principle, no problem to precompose characters also when the diacritic is input after the base letter.

15/05/2009 08:28:32 <mike> We need a solution that works for all operating systems - so if Erkki's assertion is true for just one OS then it must be stated - perhaps in a qualified way

15/05/2009 08:29:03 <Troy> The issue could also be handled in its own title. It was conceived as a solution to an MS technical problem, but it can also be a more natural way to enter text. If someone produces such a driver, it is up to users whether they wish to embrace it. It probably merits a decent size mention.

15/05/2009 08:29:52 <Karl> The problem at this time is, that in the real world we have to deal with operating systems which in fact do not support Unicode character sequences in an ideal way.

15/05/2009 08:29:53 <Troy> should be "under its own title"

15/05/2009 08:30:44 <Erkki> Karl: It would be much more difficult to create precomposed characters in a keyboard drive if the diacrotics are entered after the base character. This could fairly easily be done by a separate normalization routine, though.

15/05/2009 08:31:01 <Troy> Also, even though the solution is needed on MS Windows, there is no reason whyit couldn't be deployed on GNU/Linux and other OSes as well.

15/05/2009 08:32:00 <Erkki> Troy, I agree.

15/05/2009 08:32:52 <Troy> Since MS, X11 and OSX all have input methods for inputting text, and a full blown input method can be written to integrate normalization,

15/05/2009 08:33:42 <Troy> well, the method is probably not so problematic after all.

15/05/2009 08:36:40 <Erkki> Troy, writing a normalization routine is not overly simple, I've understood.

15/05/2009 08:37:08 <mike> So where does all this leave us - we need some fairly simple statements in the document - as we cannot assume that all (or any) of the readers will know all of the above and I hope that not all of this needs to be spelled out!

15/05/2009 08:37:21 <Erkki> What I was saying earlier that a separate normalization routine is fairly easily incorporated, not written.

15/05/2009 08:39:20 <mike> The basic outcome of our document should surely be a standard approach to keyboard layout and a standard way for users to input characters - with alternative methods for generating certain characters supported

15/05/2009 08:40:04 <Troy> For the three OSes I mentioned, this shouldn't be a problem. They all have normalization available to them. ICU has it too, so anyone can write code with normalization.

15/05/2009 08:40:09 <mike> This needs to be achievable for all platforms - the question is what needs to be said about encoding to support this?

15/05/2009 08:40:27 <Erkki> I'd propose:

15/05/2009 08:41:01 <Troy> What is your take on my proposed change? The one I sent earlier.

15/05/2009 08:41:44 <Erkki> Several of the letters with multiple diacritics (e.g., many of the Lithuanian accented letters) cannot be encoded as precomposed characters, and they'll thus have to be entered as decomposed characters. They could also be entered using the above extended dead key method, but such characters normally would not be listed as being explicitly supported.

15/05/2009 08:42:09 <mike> We should not be trying to list all technical solutions for implementing the solution on different platforms - just identifying the minimum issues that need to be clarified to support this universal solution

08:43:07 <Erkki> Sorry, I forgot to change entered to encoded.

15/05/2009 08:43:30 <Troy> "have to be entered as decomposed" should be "have to be generated as decomposed by the kbd driver

15/05/2009 08:44:33 <Erkki> .Troy, I agree (entered was a mistake, though).n

15/05/2009 08:45:03 <Troy> Also, I think the part about them being explicitly supported doesn't make sense here. The same meaning is conveyed later on, as we discussed in the last chat.

15/05/2009 08:45:45 <Erkki> Troy, could you clarify, please?

15/05/2009 08:45:50 <mike> I certainly didn't understand the bit about explicitly supported

15/05/2009 08:47:08 <Erkki> Explicitly supported means that the character and its encoding is listed in the specifications as being supported.

15/05/2009 08:48:21 <Erkki> This input mode allows for free association (which is normal with UCS/Unicode).a

15/05/2009 08:48:48 <mike> But if they can be generated why would they not be listed as explicitly supported - maybe they ought to be listed as supported only using the extended dead key method?

15/05/2009 08:49:36 <Erkki> Thus, the specific combinations that are supported (and also validated).

15/05/2009 08:49:44 <Troy> Some of those characters could be outside the languages we support.

15/05/2009 08:50:36 <mike> Eg. If you were writing a user guide for Lithuanian's using the keyboard you could have a table showing them how to generate these more obscure characters

15/05/2009 08:50:46 <Erkki> Mike, support for them in the extended mode has no impact on the implication.

15/05/2009 08:50:48 <Troy> Basically, any character with a diacritic can be entered, so we don't know about them all. Hmm, this is probably not the right way to say it, but something like that.

15/05/2009 08:51:26 <Troy> I still find the sentence out of place.

15/05/2009 08:52:10 <Erkki> Mike, a user's guide for Lithuanian would definitely list them. They are also listed with their encoding in 10646 as named sequences.

15/05/2009 08:52:54 <mike> I think the trouble with the "not listed as explicitly supported" wording is negative and leaves the question - so what needs to be done?

15/05/2009 08:53:48 <Erkki> As far as guiding the implementors: nothing needs to be done. As I said before, the user's guides are a different ball game. No

15/05/2009 08:55:08 <mike> If a formulation such as "it might be necessary to explicitly inform users how to generate each character using the extended dead key method"

15/05/2009 08:56:35 <Troy> I still think the "Several of the letters ..." sentence makes a leap in logic. It should start with a reference which makes it clear that the following paragraph no longer discusses the method of input, but the encoding of the characters instead.

15/05/2009 08:57:18 <Erkki> How about: ...but such characters would not have to be explicitly listed as being supported.

15/05/2009 08:57:47 <Troy> The section up to that point discusses the dead-key method. It should first list the other method, (the one with base letter first), and then go to explain the encoding issue.

15/05/2009 08:58:08 <Troy> "It should" refers to the flow of the text within the section.

15/05/2009 08:58:58 <Erkki> Sorry, I just realized that we should have looked at the two paragraphs, not just one.

08:59:18 <mike> I think the logic that Troy suggests sounds good to me - as I have frequently been left trying to decode the underlying implications of the current text - but unable to suggest a solution

15/05/2009 09:01:11 <Erkki> I'd now state only: They could also be entered using the above extended dead key method.

15/05/2009 09:01:23 <Troy> We have multiple issues to cover in this chapter. 1a. multiple diacritics: problem to enter. 1b. multiple diacritics: cannot be expressed precomposed. 2. dead-key method vs. "base letter first" method.

15/05/2009 09:02:10 <Troy> We start with the dead-key method. Logically, we should explain the "base letter first" method next.

15/05/2009 09:02:27 <Troy> Once we are done, explain the other issues as they relate to each of these methods.

15/05/2009 09:03:19 <Erkki> In the next We have there next: Another method to enter

15/05/2009 09:03:36 <mike> I agree that we should take one method - and explain the consequential issues - then do the same with the other method - as I said earlier, the key focus should be layout and input method issues

15/05/2009 09:06:12 <Erkki> What about, if we add the following to the end of the next paragraph?

15/05/2009 09:07:14 <Erkki> , even for those characters that could be precomposed, unless the driver would be accompanied by a normalization routine.

15/05/2009 09:08:29 <Troy> Sounds right.

15/05/2009 09:09:29 <Erkki> I'd go a couple of words back: ...precomposed encoding even for those characters...

15/05/2009 09:11:33 <Troy> Can you paste the whole sentence?

15/05/2009 09:11:42 \* Erkki quit (timeout)

15/05/2009 09:12:33 <Erkki> I'll paste the two paragraphs:

15/05/2009 09:13:05 <mike> Good - I was getting a little lost!

15/05/2009 09:13:56 <Erkki> Several of the letters with multiple diacritics (e.g., many of the Lithuanian accented letters) cannot be encoded as precomposed characters, and they'll thus have to be generated as decomposed characters by the keyboard driver. They could also be entered using the above extended dead key method.

15/05/2009 09:15:28 <Erkki> Another method to enter composite characters as decomposed characters would be to enter the base letter followed by the combining diacritical marks. This method could be more suitable for those users who are not familiar with the dead key method. It could also be easier and possibly faster to implement than the support for multiple diacritics using the dead key method. ...

15/05/2009 09:16:05 <Erkki> ... It would not, however, produce precomposed encoding even for those characters that could be precomposed, unless the driver would be accompanied by a normalization routine.

15/05/2009 09:16:56 <Troy> "extended dead key method" refers to the SFS method, right? I hope this is clear.

15/05/2009 09:17:45 <Troy> Sounds much better.

15/05/2009 09:18:37 <Erkki> Troy, it refers to the preceding paragraph.

15/05/2009 09:18:49 <Karl> In the second part, after "Another method to enter composite characters", "as decomposed characters" seems superfluous to me. The issue is dealt with in the third part of Erkki's input anyway.

15/05/2009 09:20:21 <Erkki> Karl, I agree.

15/05/2009 09:20:35 <mike> Can we have a sentence after the description of the extended dead key method that says "This method of character entry will be referred to as the "extended dead key method" in the remainder of the document.

15/05/2009 09:21:31 <Troy> Would it be possible to enter these paragraphs in the end of the chapter? Afterall, it would be easier for the reader to first read a description of the "pragmatic solution" and then read about the problems associated with it, and the potential solutions.

09:22:00 <Erkki> Mike: Yes.,

15/05/2009 09:22:30 <Troy> Also, is "pragmatic" solution the same as "extended dead key method"? Can we write a clarification of the issue?

15/05/2009 09:23:28 <Troy> The "base letter first" approach was addressed in the SFS solution (which is the pragmatic solution). The method is essentially the extended dead key method, though.

15/05/2009 09:24:45 <Erkki> Troy, I'm lost!

15/05/2009 09:25:49 <Troy> I am not clear about the "pragmatic solution" term.

15/05/2009 09:26:14 <Erkki> The SFS standard defines all the solutions in 6.1.

15/05/2009 09:26:55 <Troy> I wonder if it becomes clear to the reader?

15/05/2009 09:27:10 <Erkki> Troy, where do you find the pragmatic solution?

15/05/2009 09:27:25 <Troy> The title of the chapter is "Pragmatic solution"

15/05/2009 09:27:33 <Erkki> It is quoted as being a pragmatic approach.

15/05/2009 09:28:04 <Troy> OK, approach. Anyway, "Pragmatic" is what stuck with me.

15/05/2009 09:30:06 <Erkki> In my understanding, this is a pragmatic approach... (based on existing layout and familiar mechanism.)

15/05/2009 09:30:42 <Troy> If everybody thinks it's clear, I have no problem with it.

15/05/2009 09:32:12 <Erkki> If we can close 6.1, what is next?

15/05/2009 09:32:24 <Erkki> We have the annexes left.

15/05/2009 09:32:25 <mike> I think as a title it is OK

15/05/2009 09:35:07 <Karl> For my part, I have no more issues.

15/05/2009 09:35:22 <Troy> What about moving the "ease of processing" paragraph after the decompose issue paragraph?

15/05/2009 09:36:49 <Troy> "ease of processing" is currently a sentence in the end of the "The underlying ..." paragraph.

15/05/2009 09:38:10 <Erkki> Troy, could it not be removed?

15/05/2009 09:38:38 <Troy> It probably could be, yes.

15/05/2009 09:39:01 <Troy> I think it is now covered in the changes we just made.

15/05/2009 09:41:35 <Erkki> OK, annexes?

15/05/2009 09:42:02 <mike> Annexes OK for me

15/05/2009 09:43:48 <Tero> Nothing new there from me, either.

15/05/2009 09:44:26 <Erkki> We can all still improve the document once it has been posted for public review.

15/05/2009 09:45:03 <Erkki> ...not frequently, though, which we all welcome, I trust.

15/05/2009 09:45:33 <Troy> I am not a stickler on the "his or her" in terms of equality. I usually prefer to use either she or he alone.

15/05/2009 09:45:59 <Troy> If you want to change it, though, I am OK with it.

15/05/2009 09:46:05 <Erkki> I'd prefer to use he or she .

15/05/2009 09:46:43 <Erkki> That change has not been implemented throughout, yet.

15/05/2009 09:47:49 <Erkki> I'm also changing they'll to they will, etc.

15/05/2009 09:49:04 <Erkki> Anything else?

15/05/2009 09:49:44 <mike> On the "she" and "he" issue, my strategy is to use the plural form wherever possible e.g. "they" as it avoids any \*\*\*ism (or clumsy attempts to avoid \*\*\*ism)

15/05/2009 09:49:45 <Erkki> Tero will be in contact with CEN for the requirements on the format at this stage.

15/05/2009 09:50:14 <Tero> I'll also start making preparations for the meeting.

15/05/2009 09:50:26 <Tero> ... the date of which we should now decide.

15/05/2009 09:50:33 <mike> I see this chat software has a filtering system - I suspect that you realise what the 3 asterisks were - crazy!

15/05/2009 09:50:39 <Erkki> I'll look through the document in the light of Mike's statement.

15/05/2009 09:51:05 <Troy> haha

15/05/2009 09:51:23 <Troy> \*\*\*tant is a verboted word too, then.

15/05/2009 09:51:25 <Tero> "They" works, that's true, but isn't it considered a bit colloquial?

15/05/2009 09:51:28 <Troy> verboten

15/05/2009 09:51:56 <mike> Not really

09:52:39 <Troy> They form is only a problem because it sometimes forces one to use "unapproved" grammar. It does often make text flow better.

15/05/2009 09:53:41 <mike> True - but usually it works fine

15/05/2009 09:54:00 <Troy> Thre are still some places in Annex 3 where the '\' is not changed to '+'.

15/05/2009 09:54:01 <Tero> We really should've written the whole document in Finnish. That would have solved all \*\*\* problems.

15/05/2009 09:54:21 <Troy> Yep, or Sami.

15/05/2009 09:54:44 <Tero> Or Hungarian. Or Korean etc.

15/05/2009 09:54:50 <Troy> We should also conduct these chats in arabic, in order to avoid sensoring.

15/05/2009 09:54:54 <Erkki> Well, we still need to decide the date for the public meeting.

15/05/2009 09:54:59 <Tero> Indeed.

15/05/2009 09:55:08 <Tero> The first week of September has been suggested.

15/05/2009 09:55:12 <Tero> Any objections?

15/05/2009 09:55:34 <Tero> I have also checked that the CEN facilities are available at that time.

15/05/2009 09:56:19 <mike> Is this the week beginning 31/08 - if so fine - but not the next week

15/05/2009 09:56:48 <mike> Or the following one

15/05/2009 09:56:49 <Erkki> Mike, yes it is.

15/05/2009 09:57:02 <mike> Perfect

15/05/2009 09:57:07 <Karl> For me, any day is OK in that week.

15/05/2009 09:57:11 <Erkki> So, which weekday?

15/05/2009 09:57:23 <Karl> Friday 4th?

15/05/2009 09:58:08 <Erkki> People want to leave early particularly on Fridays!

15/05/2009 09:58:17 <mike> Friday would be my least preferred day - as I fly off to Nice for 2 weeks on the Sunday - but I can do it if that is best for everyone else

15/05/2009 09:58:22 <Tero> Maybe Thursday would be better?

15/05/2009 09:58:38 <mike> Good for me

15/05/2009 09:58:52 <Karl> OK for me.

15/05/2009 09:59:05 <Troy> Works for me.

15/05/2009 09:59:17 <Erkki> Fine, then.

15/05/2009 09:59:30 <Tero> Then the next question is, how many participants are we expecting?

15/05/2009 09:59:38 <Tero> I need to give this information to Mina.

15/05/2009 10:00:31 <Troy> Maybe ask via email?

15/05/2009 10:00:32 <Erkki> If you state that it will be less than 20, I believe you are safe in your estimate.

15/05/2009 10:00:41 <Tero> Sounds good.

10:01:26 <Tero> I'll relay this information.

15/05/2009 10:02:09 <Erkki> OK, on my part I thank you all for your great patience and valuable contributions.

15/05/2009 10:03:00 <Troy> Thanks to you, Erkki.

15/05/2009 10:03:21 <mike> Thanks also from me - I think we have all worked well together

15/05/2009 10:04:07 <Troy> Yes, I agree. This has worked out nicely.

15/05/2009 10:05:28 <Karl> Thanks from me, also.

15/05/2009 10:05:40 <Tero> My first experience as a workshop secretary has also been very pleasant.

15/05/2009 10:06:21 <Erkki> Tero, little do you know yet!

15/05/2009 10:06:35 <Tero> Probably so.

15/05/2009 10:06:47 <Erkki> ...since you'll have to perform the final edits by the rules.

15/05/2009 10:07:09 <Tero> But now thank you all and see you in September!