

Bilingual Arabic-based keyboard designed to be used with single hand

الدكتور عبد الكريم عكلة

رئيس قسم علوم هندسة البرمجيات/ كلية بغداد للعلوم الاقتصادية

Abstract

In this research we will discuss the imprecise distribution problem of the Arabic characters on keyboards. We suggest a new Arabic-based coding scheme instead of ASCII with a new bilingual keyboard suitable for it. The number of keys in the proposed keyboard is reduced to increase the speed of input characters. The proposed keyboard is designed to be used with single hand. The second hand maybe used for supporting to press the control keys only. The suggested Arabic-based coding scheme is compressed so that each character with its vowel point is stored in single byte instead of two bytes using ASCII. In other words, we reduce the capacity needed to store the Arabic text to the half of the capacity required using ASC

الخلاصة:

في هذا البحث سناقش مشكلة توزيع الحروف العربية الغير دقيق على لوحة المفاتيح المتداولة حاليا و اقترحنا ترميز جديد ثنائي اللغة يعتمد الحروف العربية كأساس بدلا من اللغة الانكليزية في جدول الترميز المعروف بالأسكي. وعلى ضوء جدول الترميز الجديد اقترحنا تصميم جديد للوحة مفاتيح ثنائية اللغة تعتمد اللغة العربية كأساس مع لغة أخرى (كالانكليزية وهي المقترحة في هذا البحث أو أية لغة أخرى يمكن للمستخدم اختيارها). تمتاز لوحة المفاتيح هذه بسرعة الإدخال لقلة الأزرار عليها بحيث تكفي أصابع اليد الواحدة لإدخال النصوص. في التمثيل المقترح يمكن خزن الحرف العربي غير المشكل في نصف بايت تقريبا (خمس ثنائيات) والحرف

المشكّل بمساحة خزنيه بحجم بايت واحد (ثمان ثنائيات)، أي بمعنى آخر تم من خلال هذا التمثيل تقليص حجم البيانات العربية المخزونة أو المنقولة إلى النصف.

الكلمات المفتاحية : لوحة مفاتيح عربية- نظام تعريب- نظام ترميز - الحروف العربية- سرعة الإدخال.

1. مقدمة

لقد تطورت طرق إدخال البيانات إلى الحاسوب كثيرا وتعددت وسائلها بالطريقة التي تكفل السرعة والدقة في الإدخال كاستخدام الفأرة والقلم الضوئي والماصح الضوئي والقارئ الضوئي وغيرها. ورغم ذلك فإن لوحة المفاتيح هي الطريقة الأكثر شيوعا في الاستخدام لقلة كلفتها وقدرتها على إدخال جميع أنواع البيانات وفي الوقت نفسه تعتبر هي الأبطأ من حيث السرعة في الإدخال لأن سرعة الإدخال تعتمد على سرعة يدي مستخدمها في الضغط على مفاتيح الحروف.

إن سرعة الإدخال من خلال لوحة المفاتيح تقاس بعدد الحروف المدخلة بالدقيقة وهي سرعة تعتمد على مدى الممارسة لعملية الإدخال وتزداد سرعة الإدخال باستخدام أصابع كلتا اليدين وحفظ مواقع الحروف على أزرار لوحة المفاتيح. وتتراوح سرعة الإدخال بين 20-50 حرفا بالدقيقة للمبتدئين ومن 50-80 حرفا بالدقيقة للمستخدمين الاعتياديين وفوق 80 حرفا بالدقيقة للمحترفين بالنسبة لحروف اللغة الانكليزية وفق ما هو متوفر في اغلب معاهد التدريب الحالية وسرعة الإدخال هذه وتكون اقل بالنسبة لحروف اللغة العربية. وعموما فإن هذه النسب مختلفة باختلاف اللغات لاختلاف عدد الحروف وطريقة توزيعها على لوحة المفاتيح.

لقد ظهرت في الأسواق العديد من البرامج المتخصصة في التدريب على زيادة سرعة الإدخال ولمختلف اللغات وهناك برامج خاصة معتمدة في بعض المعاهد والشركات المتخصصة. المهمة الرئيسية لهذه البرامج هي اعتياد المتدرب على استخدام أصابع كلتا يديه وحفظ مواقع الحروف على لوحة المفاتيح وللغة محددة فقط و لا توجد طريقة تدريب لعموم اللغات.

العوامل المؤثرة في سرعة إدخال البيانات

يستخدم مدخّل البيانات عينيه لمعرفة مواقع البيانات على لوحة المفاتيح وأصابعه للضغط على أزرارها وتدقيق المدخلات بالمتابعة على شاشة العرض وتصحيح الأخطاء إن وجدت. إن تنفيذ هذه المهمة بالسرعة الكافية يحتاج إلى قدرة خاصة يمكن امتلاكها بالتدريب الكافي ليتمكن المدخل من استخدام لوحة المفاتيح دون النظر إليها. وهذا التدريب يشمل رفع سرعة المتدرب على إدخال البيانات بدون أخطاء بحفظ مواقع الحروف على لوحة المفاتيح دون الحاجة إلى النظر إليها للبحث عن مواقع الحروف. إن هذا التدريب يحتاج إلى بعض المتطلبات

لإنجاحه منها انسيابية وشكل وأبعاد وطريقة صنع لوحة المفاتيح وحجم المفتاح وطريقة الضغط على الأزرار وغيرها من الأمور التي تتعلق بالخواص الفيزيائية للوحة المفاتيح. ولكن الأهم هي طريقة توزيع البيانات على لوحة المفاتيح بحيث تتناسب وتتناظر مع أصابع اليدين لان قوة السيطرة وسرعة الاستجابة مختلفة في أصابع اليدين كاختلاف أطوالها. فالسبابة هي الإصبع الأكثر قوة في السيطرة وسرعة الاستجابة وهي الأكثر استخداماً تليها الوسطى فالخنصر فالبنصر. أما الإبهام فيستخدم في الضغط على زر الفراغ فقط لاختلاف حجمه وشكله وموقعه في اليدين بالنسبة لبقية أصابع اليدين.

وعلى هذا الأساس يتم توزيع الحروف الأكثر تكراراً في اللغة بشكل يناظر مواقع الأصابع الأكثر استخداماً في اليدين والحروف الأقل تكراراً في اللغة بشكل يناظر مواقع الأصابع الأقل استخداماً في اليدين. وهذا ما يجعل اللغات الأقل حروفاً أسهل وأسرع في الإدخال من اللغات الأكثر حروفاً لان زيادة الحروف تؤدي إلى زيادة الحروف التي تتكرر كثيراً وهذا يدفع إلى زيادة الأزرار التي تناظر الأصابع الأكثر استخداماً في اليدين بحيث يكون الإصبع الواحد مسؤول عن الضغط على أكثر من زر وهذا قد يؤدي إلى حصول الأخطاء في الإدخال. وعليه فمن المتوقع أن تكون عملية إدخال البيانات باللغة العربية أبطأ من إدخال البيانات باللغة الانكليزية لان اللغة العربية هي الأكثر حروفاً إذ إن الحروف الانكليزية على لوحة المفاتيح وزعت على 26 زراً و الحروف العربية على لوحة المفاتيح وزعت على 34 زراً (لا حظ الشكلين 1 و2) وهذا سبب يكفي لجعل إدخال الحروف العربية أبطأ من إدخال الحروف الانكليزية.

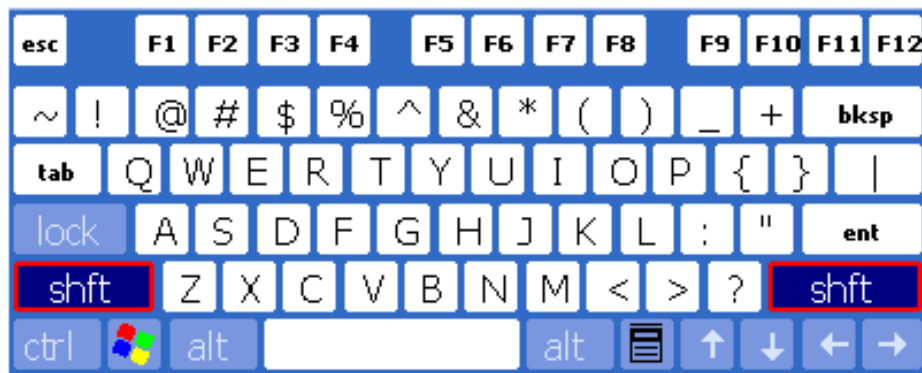
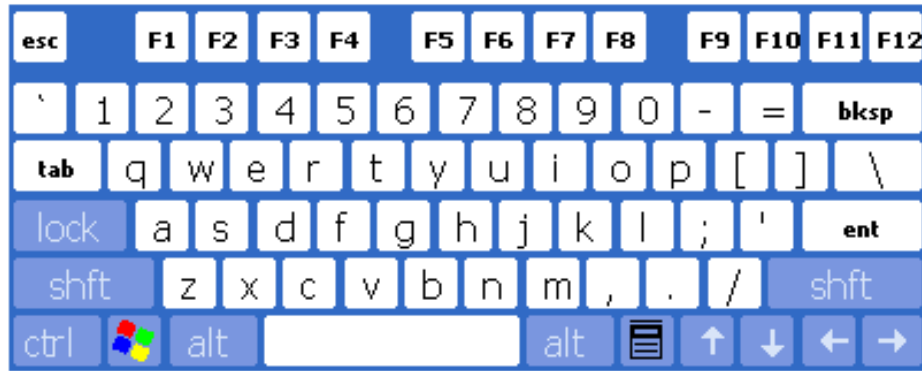
وهنا تجدر الإشارة إلى إن الحروف الانكليزية لها شكلين: حروف صغيرة وحروف كبيرة (أي إن لكل حرف صورتين)، وهذا لا يشكل زيادة في عدد الحروف لان موقع الحرف بشكله الصغير والكبير على نفس الزر و هذا الأمر مدروس بدقة بالنسبة للحروف الانكليزية عند تصميم لوحة المفاتيح ويعكسه بالنسبة لحروف اللغة العربية فان هنالك حروف لها عدة أشكال وفي أزرار مختلفة كما هو الحال بالنسبة لحرف الألف مثلاً.

3- توزيع الحروف على لوحات المفاتيح الحالية

لقد لاحظ المصممون للوحة المفاتيح عند توزيع الحروف الانكليزية على لوحة المفاتيح إن الحروف هي الأكثر تكراراً في اللغة فوضعوها بمواقع تناظر السبابة والوسطى من الأصابع في اليدين E,R,T,Y,U,I بحيث إن السبابة مسؤولة عن إدخال حرفين والوسطى مسؤولة عن إدخال حرف واحد فقط (لاحظ الشكل 3).



شكل رقم 1 : لوحة المفاتيح بالحروف العربية فقط



شكل رقم 2: لوحة المفاتيح بالحروف الانكليزية



الشكل 3. بعض أزرار لوحة المفاتيح وما يقابلها من

ولكن المصممون للوحات المفاتيح قد وزعوا الحروف هـ، ع، غ، ف، ق، ث على نفس الأزرار التي وزعوا عليها الحروف الانكليزية المذكورة أعلاه. فهل إن هذه الحروف هي الأكثر تكراراً في اللغة العربية ؟ !
 الجواب بالتأكيد لا والسؤال التالي هنا ما هو المقياس الذي استخدمه في توزيع الحروف العربية بالشكل المعروف على لوحات المفاتيح؟ وهل إن المقياس المستخدم قد اخذ بنظر الاعتبار ما تقدم في متن هذا النص؟

وعموماً مهما يكن المقياس فإن ما يؤخذ على طريقة توزيع الحروف العربية على لوحة المفاتيح إنها تؤدي إلى إرباك المدخل وإبطاء تعوده وحفظه لمواقع أزرار المفاتيح وهذا ما يجعله أبطأ في إدخال البيانات باللغة العربية من اللغة الانكليزية.

4 - كيف يمكن توزيع المفاتيح على لوحة المفاتيح؟

لقد أصبح جليا مما تقدم الحاجة إلى إعادة توزيع الحروف العربية على لوحات المفاتيح وهو المدخل إلى التفكير باستنباط طريقة جديدة لتمثيل حروف اللغة العربية وإعادة توزيعها على لوحة المفاتيح بشكل يضمن الدقة والسرعة في إدخال البيانات العربية. الفكرة الرئيسية الابتدائية للبحث هي تسريع عملية إدخال البيانات باللغة العربية من خلال لوحة المفاتيح باتباع الخطوتين التاليتين:

إعادة توزيع الحروف العربية على لوحة المفاتيح بشكل يتناسب مع مواقع أصابع اليدين.
 محاولة تقليل عدد الحروف العربية على لوحة المفاتيح

الخطوة الأولى تنفذ بعد حساب تكرارات الحروف العربية في النصوص القياسية لمعرفة الحروف الأكثر تكرارا لوضعها على الأزرار التي تقابل الأصابع النشطة من اليدين. أما الخطوة الثانية فإن تنفيذها لا يتم إلا باستبدال جدول الترميز المعروف بالأسكي والذي يعتمد اللغة الانكليزية كأساس بأخر أساسه تمثيل جديد للحروف العربية بحيث نجعل كل صور الحرف الواحد على نفس الزر كما هو الحال بالنسبة لحروف اللغة الانكليزية إلا إن نتائج الدراسة دفعت باتجاه جديد وهو تصميم لوحة مفاتيح ثنائية اللغة جديدة تستخدم حروف اللغة العربية كأساس مع لغة أخرى حسب اختيار المصنّع.

لغرض تنفيذ الخطوة الاولى قمنا بحساب ترددات الحروف العربية لنصوص عديدة في مختلف المجالات (الأدبية، العلمية، الإعلامية، القرآنية) وبطول 10000 حرف لكل نص وحسبنا نسبة تردد الحروف كما في الشكل رقم 4. ولغرض تحديد الحروف قليلة التردد (الخاملة) : وهي التي ينسى موقعها على لوحة المفاتيح لقلّة تردها في النصوص) قمنا بدراسة النصوص المذكورة أعلاه لتحديد متوسط عدد الحروف في الجمل العربية ووجدنا إن معدل عدد الحروف في الجمل العربية يساوي مئة حرف تقريبا أي سطرين بالحجم العادي (14) وافترضنا إن الحروف التي تكون احتمالية تردها في الجملة المتوسطة اقل من واحد هي خاملة (أي احتمال أن لا ترد في الجملة المتوسطة) وحسب هذا الفرض صنفنا الحروف الى نشطة اي تردها اكثر أو يساوي 1% وحروف خاملة تردها اقل من 1%.

أ	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص
19.3	4.2	6	0.7	0.9	1.7	0.9	2.5	0.6	3.9	0.5	1.7	0.8	1.0

ض	ط	ظ	ع	غ	ف	ق	ك	ل	م	ن	ه	و	ي
0.3	0.6	0.3	3.0	0.6	2.7	2.2	2.5	11.2	7.2	5.9	4.4	7.3	4.7

شكل رقم 4. النسبة المئوية لتردد الحروف العربية

ومن الجدول رقم 4 يظهر إن الحروف (أ، ل، و، م، ت، ن، ي، هـ، ب، ر، ع، ف، د، ك، ق، س، ح، ص) هي حروف نشطة حسب الافتراض المذكور أعلاه وعددها ثمانية عشر حرفاً أما الحروف الخاملة فهي (خ، ج، ش، ث، ط، ذ، غ، ز، ض، ظ) وعددها عشرة حروف.

لضمان توزيع جيد للحروف العربية على لوحة المفاتيح يجب وضع الحروف كثيرة التردد بما يناظر الأصابع النشطة في اليدين وهذا ممكن، فإن السبابة في كل يد يمكن أن تكون مناظرة لعمودين من الحروف على لوحة المفاتيح (لاحظ شكل رقم واحد) مع الوسطى تناظر عمود واحد فيكون المجموع تسعة حروف لكل يد والمجموع يساوي عدد الحروف النشطة. والمتبقي هي ستة عشر حرفاً، عشرة منها الحروف الخاملة مع ثمانية حروف مشتقة أخرى مثل (ء، ئ، ؤ، ى، آ، إ، ة). علماً إن الخنصر والصغرى من أصابع اليد يمكن أن يناظرون ستة أزرار في كل يد أي بمجموع اثنا عشر حرفاً والباقي ثلاثة حروف يجب أن تعالج باستخدام الزر Shift مع احد الأزرار المناظرة للخنصر أو الصغرى من اليدين.

إن ما تقدم يمكن أن يكون حلاً مناسباً يساهم في زيادة سرعة الإدخال أي استخدام نفس مجموعة الحروف العربية الموجودة على لوحة المفاتيح الحالية مع استخدام ترتيب جديد لها مبني على أساس وضع الحروف الأكثر تكراراً في أزرار تناظر الأصابع النشطة من اليدين والحروف الأقل تكراراً في أزرار مناظرة للأصابع الخاملة من اليدين. ولكن تبقى مشكلة الحروف الخاملة التي ستكون مناظرة لإحدى الأصابع الخاملة والتي ستسبب بمرور الوقت وقلة الاستخدام مما يدفع المستخدم إلى النظر إلى لوحة المفاتيح لمعرفة مكانها وهي من المعوقات لزيادة سرعة الإدخال.

للتخلص من المشكلة المذكورة يمكن وضع حرف من الحروف الخاملة مع حرف من الحروف النشطة في نفس الزر. إن هذا الفرض دفعنا إلى ملاحظة مهمة وهي إن معظم الحروف المنقطة هي حروف خاملة (لاحظ شكل

رقم 4) ومعظم الحروف غير المنقطه هي حروف نشطة مثل الدال حرف نشط ويقابله الذال حرف خامل والراء حرف نشط يقابله الزاي حرف خامل والسين حرف نشط يقابله الشين حرف خامل وهكذا بقية الحروف، مع ملاحظة بعض الشذوذ من هذه القاعدة كما في الباء والتاء والنون والياء والفاء والقاف .

الملاحظة الجديدة أوحى لنا بوضع الحرف المنقط مع الحرف غير المنقط في نفس الزر كما في الحروف الصغيرة (Small letter) والكبيرة (Capital letter) في اللغة الانكليزية أي استخدام الزر Shift مع الحرف غير المنقط لكتابة الحرف المنقط. مع اعتبار أن الياء هو المنقط لحرف الباء (أي بإضافة نقطة لحرف الباء) وان التاء هو الحرف المنقط لحرف النون بنفس الأسلوب. مع ملاحظة إن هنالك حروف غير منقطه مثل الألف والميم والواو واللام والكاف والهاء والهمزة وان هنالك أكثر من حالة واحدة للتقطيع مثل التاء والخاء (على اعتبار إن التاء هو الحرف المنقط للنون وان الجيم هو الحرف المنقط للحاء).

إن استخدام مفتاح Shift قلص عدد الأزرار التي تحتاجها الحروف العربية إذ إن الحروف العربية الموزعة على لوحة المفاتيح الحالية لا تستخدم المفتاح Shift إلا مع بعض أشكال حرف الألف فقط. إذن نحن تمكنا من تنفيذ الخطوة الثانية بتقليل أزرار الحروف العربية والسؤال هنا هو إلى أي مقدار يمكننا التقليل؟ وللإجابة على مثل هذا السؤال علينا التعرف أولاً على نظام التمثيل الحالي ASCII وكيفية عمل لوحة المفاتيح.

5- نظام تمثيل الحروف ASCII

ASCII هي مختصر (American Standard for Code Information Interchange)، تُلَفظ عادة أسكي، هي مجموعة رموز ونظام تمثيل مبني على الألف باء اللاتينية بالشكل الذي تستخدم به في الإنجليزية الحديثة ولغات غرب أوروبية أخرى. من أكثر الاستخدامات شيوعاً للنصوص المكتوبة بالأسكي هو استخدامها في أنظمة الحاسوب، كما تستخدم في أجهزة الاتصالات وأنظمة التحكم التي تتعامل مع نصوص.

كغيرها من أنظمة تمثيل الرموز في الحاسوب، ASCII تحدد علاقة تناظر بين قيمة رقمية لتتابع البتات (أو الثنائيات وهي الأرقام التي تكتب كسلسلة من الصفر والواحد فقط) وبين رمز أو رسم مستخدم في اللغة المكتوبة، مفسحة المجال للأجهزة الرقمية للتواصل، المعالجة والتخزين عن طريق تبادل بيانات مكونة من هذه الرموز. نظام ترميز أسكي أو امتداد متوافق معه هو المستخدم في كل الحواسيب تقريباً، وخاصة الحواسيب الشخصية ومحطات العمل.

و الأسكي (ASCII) هو نظام تمثيل من 7 بتات يستخدم سبعة أرقام ثنائية القاعدة (قيمة تتراوح بين 0 و 127) لتمثيل الحروف والرموز. وفي الوقت الذي أُقترح فيه نظام ASCII، كان العديد من الأجهزة تتعامل مع مجموعات من ثمان بتات (المعروف باسم بايت أو Octet) كأصغر وحدة معلوماتية؛ وشاع استخدام البت

الثامن parity bit لأغراض ضمان لفحص الأخطاء على خطوط الاتصال ولأهداف أخرى ترتبط بالهدف من الأجهزة المستخدمة. وكانت الأجهزة التي لم تستخدم parity bit كانت تضع القيمة 0 في البت الثامن عادة، ولكن بعض الأنظمة مثل Prime Computer والتي شغلت نظام PRIMOS كانت تضع القيمة 1 في البت الثامن.[]

يتكون جدول الآسكي من 128 مدخل، كل مدخل عبارة عن قيمة لتمثيل حرف أو رمز معين. المداخل (من 0 الى 31) هي رموز للسيطرة على لوحة المفاتيح وعلى المعلومات المرسله والمخزونة. والمداخل الباقية هي قيم لتمثيل الحروف الانكليزية المستخدمة والأرقام وبقية الرموز الأخرى المستخدمة في التطبيقات التجارية والعلمية (لاحظ شكل رقم 5).

والملاحظ في الشكل رقم 5، إن الفرق بين قيم الحروف الانكليزية الكبيرة والصغيرة هو 32 والحروف الكبيرة هي الأقل قيمة من الحروف الصغيرة. وان الحرف الصغير يمكن كتابته بضغط المفتاح Shift مع الحرف الكبير، فهذا يعني إن وظيفة المفتاح Shift هي زيادة 32 إلى قيمة الحرف المضغوط معه. والحقيقة إن المفاتيح Alt و Ctrl لها وظائف مشابهة فالأول يزيد 64 على قيمة الحرف المضغوط معه والثاني يزيد 96 على قيمة الحرف المضغوط معه. وهذه المفاتيح تستخدم لتقليل عدد الأزرار الموجودة على لوحة المفاتيح بجعل كل زر له أربعة قيم مختلفة، قيمة فيما لو ضُغَطَ لوحده وثلاث قيم أخرى فيما لو ضُغَطَ مع احد المفاتيح الثلاثة (Shift, Alt, Ctrl).

ولإضافة حروف اللغة العربية (أو أي لغة أخرى) إلى لوحة المفاتيح يستخدم البت الثامن (parity bit) لتوسيع جدول الآسكي إلى 256 مدخل ويسمى الجدول الجديد تسميات أخرى حسب الشركات المصنعة (والغالب Code Page N حيث إن N تمثل رقم محدد لكل لغة). إن بناء جدول الآسكي على أساس اللغة الانكليزية هو السبب الذي يجعل اللغة الأساسية للوحة المفاتيح هي اللغة الانكليزية ويضاف لها لغة أخرى ولهذا السبب تسمى لوحة مفاتيح ثنائية اللغة.

6- استخدام المفتاح Shift

إن الفكرة الأساسية وراء استخدام قيمة الزيادة للمفتاح Shift مساوية 32 هي إن عدد حروف الأبجدية الانكليزية هو 26 حرفاً وإن حجم التمثيل المناسب يساوي $\lceil \log_2 26 \rceil \approx 5$ بتات والتمثيل الخماسي يغطي 32 رمزاً أي يشمل جميع الحروف الانكليزية. أما اللغة العربية فإذا أضفنا الحروف (ء، ي، و، ي، آ، أ، إ، ة) إلى الأبجدية العربية يصبح العدد 36 حرفاً أي نحتاج إلى أن تكون قيمة الزيادة للمفتاح Shift تساوي 64 إذا ما حاولنا استخدام المفتاح Shift لاختصار الحروف العربية على لوحة المفاتيح كما هو الحال بالنسبة للحروف الانكليزية أي يجب أن تحتوي لوحة المفاتيح على 64 زر للحروف و الأرقام إضافة إلى أزرار السيطرة الأخرى

علما أن لوحة المفاتيح الحالية تحتوي على 45 زر للحروف والأرقام والرموز عدا رموز السيطرة. إن بناء مثل هكذا لوحة مفاتيح يؤدي زيادة عدد الأزرار وزيادة حجم لوحة المفاتيح وبالتالي إيجاد معوق آخر لزيادة سرعة الإدخال وهو عكس هدف البحث.

إن المعوق أعلاه يدفعنا في هذا البحث إلى إيجاد طريقة لتمثيل الحروف العربية وتقليص عددها على لوحة المفاتيح باستخدام المفتاح Shift بحيث تكون قيمة الزيادة 16 وليست 32 لأن القيمة 32 مستخدمة حالياً ولا تؤدي إلى تقليل عدد الأزرار على لوحة المفاتيح.

أن تكون قيمة الزيادة في المفتاح Shift تساوي 16 يعني أن يغطي التمثيل الرباعي كل حروف الأبجدية العربية ويكلام آخر اختصار الأبجدية العربية إلى 15 حرف يضاف إليه الفراغ كحرف أساسي. معنى الاختصار هنا هو أن يكون كل زر يمثل حرفين لا يختلفان كثيراً بالشكل بحيث تسهل عملية تذكر مواقعها على لوحة المفاتيح وبشكل مماثل للحروف الانكليزية.

وعلى هذا الأساس فالمطلوب عرض 15 حرفاً أساسياً على 15 زر من أزرار لوحة المفاتيح بحيث تتمكن من تذكر مواقع لحروف مشابهة على نفس الأزرار باستخدام المفتاح Shift. أي إن هذا التمثيل سيغطي ثلاثون حرفاً وحروف اللغة العربية أكثر من ذلك.

7- نظام التمثيل الجديد

لقد تمكنا في بحث سابق [1] من وضع تمثيل للحروف العربية يستخدم فيه التمثيل الخماسي كما هو موضح في الشكل رقم 6.

الحرف الأصلي	المشتق بإضافة نقطة	الحرف الأصلي	المشتق بإضافة نقطة	الحرف الأصلي	المشتق بإضافة نقطة
الألف	ء	ر	ز	ف	ق
ب	ي	س	ش	ل	ك
ن	ت	ص	ض	م	ث
ح	ج	ط	ظ	ه	ة
□	ذ	ع	غ	و	خ

شكل رقم 6. الحروف العربية الأصلية والمنقطة

1-7 تمثيل الحروف

في الشكل رقم 6, اعتبرنا إن هو الحرف المنقوت لحرف الألف وهي تسمية مجردة للتماثل مع بقية الحروف وكذلك الحال بالسبب للكاف مع اللام والتاء المربوطة مع الهاء أما في حالة التاء والحاء فإنها حالات إضافية ثلاثة للنون والحاء ولا محل لها في جدولنا وفي نفس الوقت فإن حروف الميم والواو حروف لا توجد مشتقة من التنقيط منها لذلك اعتبرنا التاء هو المشتق من تنقيط الميم والحاء هو المشتق من تنقيط الواو ولذلك ضرورتان الأولى: إن حرفا الميم والواو هما من الحروف العالية التردد في النصوص العربية (لاحظ الشكل رقم 4) وإن حرفا التاء والحاء هما من الحروف القليلة التردد في النصوص العربية ولابد لنا من وضع عالي التردد مع قليل التردد ليسهل تذكر موقع الأخير كما أن وجود مثل هاتان الحالتان لا تشكل معضلة بالنسبة للفكرة الأساسية للبحث لأن التاء والحاء هما من الحروف القليلة جدا في النصوص, أما الضرورة الثانية : فإنه لا توجد طريقة أخرى لترتيب الجدول بحيث نستطيع أن نضع خمسة عشر حرفا في صورة جدول لكل مدخل فيه وجهان بدون مثل هذا التجاوز بسيط. علما إن المدخل السادس عشر لهذا الجدول هو رمز الفراغ وله مدخل بوجه واحد (لاحظ الشكل رقم 7).

بهذا نكون قد تمكنا من وضع تمثيل متسلسل فيه الفرق بين الحرف الأصلي والحرف المشتق بالتنقيط قيمة تساوي 16 لجعل قيمة الزيادة للمفتاح Shift تساوي 16 وإن الفراغ له تمثيلان كما هو الحال في لوحة المفاتيح الحالية (أي إن للفراغ قيمتان الفرق بينهما قيمة الزيادة للمفتاح Shift والقيمتان هما 32 و 64). في الشكل رقم 7 نجد إن قيمة حرف الألف هي 1 والهمزة 17 وحرف الباء 2 والياء 18 وهكذا. أما الفراغ فله قيمتان هما الصفر و 16.

أما الحروف الغير واردة في الجدول مثل (ء, و, ي, آ, أ, إ, ا) فسيرد شرح تمثيلها بعد شرح التشكيل في الفقرات التالية.

شكل رقم 5. جدول الآسكي

Dec	Hex	Char	Name	Dec	Hex	Char	Name
32	20		SPACE	79	4F	O	LATIN CAPITAL LETTER O
33	21	!	EXCLAMATION MARK	80	50	P	LATIN CAPITAL LETTER P
34	22	"	QUOTATION MARK	81	51	Q	LATIN CAPITAL LETTER Q
35	23	#	NUMBER SIGN	82	52	R	LATIN CAPITAL LETTER R
36	24	\$	DOLLAR SIGN	83	53	S	LATIN CAPITAL LETTER S
37	25	%	PERCENT SIGN	84	54	T	LATIN CAPITAL LETTER T
38	26	&	AMPERSAND	85	55	U	LATIN CAPITAL LETTER U
39	27	'	APOSTROPHE	86	56	V	LATIN CAPITAL LETTER V
40	28	(LEFT PARENTHESIS	87	57	W	LATIN CAPITAL LETTER W



41	29)	RIGHT PARENTHESIS	88	58	X	LATIN CAPITAL LETTER X
42	2A	*	ASTERISK	89	59	Y	LATIN CAPITAL LETTER Y
43	2B	+	PLUS SIGN	90	5A	Z	LATIN CAPITAL LETTER Z
44	2C	,	COMMA	91	5B	[LEFT SQUARE BRACKET
45	2D	-	HYPHEN-MINUS	92	5C	\	REVERSE SOLIDUS
46	2E	.	FULL STOP	93	5D]	RIGHT SQUARE BRACKET
47	2F	/	SOLIDUS	94	5E	^	CIRCUMFLEX ACCENT
48	30	0	DIGIT ZERO	95	5F	_	LOW LINE
49	31	1	DIGIT ONE	96	60	`	GRAVE ACCENT
50	32	2	DIGIT TWO	97	61	a	LATIN SMALL LETTER A
51	33	3	DIGIT THREE	98	62	b	LATIN SMALL LETTER B
52	34	4	DIGIT FOUR	99	63	c	LATIN SMALL LETTER C
53	35	5	DIGIT FIVE	100	64	d	LATIN SMALL LETTER D
54	36	6	DIGIT SIX	101	65	e	LATIN SMALL LETTER E
55	37	7	DIGIT SEVEN	102	66	f	LATIN SMALL LETTER F
56	38	8	DIGIT EIGHT	103	67	g	LATIN SMALL LETTER G
57	39	9	DIGIT NINE	104	68	h	LATIN SMALL LETTER H
58	3A	:	COLON	105	69	i	LATIN SMALL LETTER I
59	3B	;	SEMICOLON	106	6A	j	LATIN SMALL LETTER J
60	3C	<	LESS-THAN SIGN	107	6B	k	LATIN SMALL LETTER K
61	3D	=	EQUALS SIGN	108	6C	l	LATIN SMALL LETTER L
62	3E	>	GREATER-THAN SIGN	109	6D	m	LATIN SMALL LETTER M
63	3F	?	QUESTION MARK	110	6E	n	LATIN SMALL LETTER N
64	40	@	COMMERCIAL AT	111	6F	o	LATIN SMALL LETTER O
65	41	A	LATIN CAPITAL LETTER A	112	70	p	LATIN SMALL LETTER P
66	42	B	LATIN CAPITAL LETTER B	113	71	q	LATIN SMALL LETTER Q
67	43	C	LATIN CAPITAL LETTER C	114	72	r	LATIN SMALL LETTER R
68	44	D	LATIN CAPITAL LETTER D	115	73	s	LATIN SMALL LETTER S
69	45	E	LATIN CAPITAL LETTER E	116	74	t	LATIN SMALL LETTER T
70	46	F	LATIN CAPITAL LETTER F	117	75	u	LATIN SMALL LETTER U
71	47	G	LATIN CAPITAL LETTER G	118	76	v	LATIN SMALL LETTER V
72	48	H	LATIN CAPITAL LETTER H	119	77	w	LATIN SMALL LETTER W
73	49	I	LATIN CAPITAL LETTER I	120	78	x	LATIN SMALL LETTER X
74	4A	J	LATIN CAPITAL LETTER J	121	79	y	LATIN SMALL LETTER Y
75	4B	K	LATIN CAPITAL LETTER K	122	7A	z	LATIN SMALL LETTER Z
76	4C	L	LATIN CAPITAL LETTER L	123	7B	{	LEFT CURLY BRACKET
77	4D	M	LATIN CAPITAL LETTER M	124	7C		VERTICAL LINE
78	4E	N	LATIN CAPITAL LETTER N	125	7D	}	RIGHT CURLY BRACKET
				126	7E	~	TILDE

2-7 تمثيل التشكيل في اللغة العربية

تختلف اللغة العربية عن اللغة الانكليزية واغلب اللغات الحية بخاصية التشكيل وهي عملية رسم إحدى العلامات

السبع التالية :

السكون ().

الكسرة ().

الفتحة ().

الضمة ().

تتوين الكسر ().

تتوين الفتح ().

تتوين الضم ().

وظيفتها تغيير النطق الصوتي للحرف وتسمى حركات. وقد عولج التشكيل في لوحة المفاتيح الحالية على إنه حرف قائم بذاته له تمثيل ثماني كبقية الحروف لكنه يرسم فوق الحرف المحرك بدون حاجة إلى حيز كافٍ لحرف وهذا طبعاً زيادة في الإدخال فالمستخدم للوحة المفاتيح إذا أراد إدخال أربعة حروف مشكلة فإنه يستخدم لوحة المفاتيح بالضغط على ثمانية أزرار متوالية ويستخدم المفتاح Shift مع احد أزرار الحروف لإدخال الحركات ومع ذلك فإنه لم يدخل غير أربعة حروف وهذا ما يجعل سرعة إدخال الحروف الانكليزية (أو أي لغة أخرى خالية من التشكيل) هي ضعف سرعة إدخال النصوص العربية المشكلة. وهذا السبب وراء انتشار النصوص غير المشكلة في جميع النشريات والمؤلفات عدا المختصة منها كالأدبية والدينية. والحقيقة إن عدم استخدام النصوص المشكلة هو احد الأسباب الرئيسية وراء عدم توفر معالجات تطبيقية لفهم اللغة العربية حاسوبياً وهو سبب عدم وجود ترجمة آلية دقيقة للغة العربية.

B ₃	B ₂	B ₁	B ₀	عمود سطر	B ₄	
					0	1
0	0	0	0	0	فراغ	فراغ
0	0	0	1	1	أ	ء
0	0	1	0	2	ب	پ
0	0	1	1	3	ن	ت
0	1	0	0	4	ح	ج
0	1	0	1	5	د	ذ
0	1	1	0	6	ر	ز
0	1	1	1	7	س	ش
1	0	0	0	8	ص	ض
1	0	0	1	9	ط	ظ
1	0	1	0	10	ع	غ
1	0	1	1	11	ف	ق
1	1	0	0	12	ل	ك
1	1	0	1	13	م	ث
1	1	1	0	14	هـ	ة
1	1	1	1	15	و	خ

شكل رقم 7. التمثيل الخماسي لحروف اللغة العربية.

نرى في هذا البحث إن التشكيل أو الحركات هي خصائص متعلقة بالحرف العربي وليس لها معنى بدونه ويجب أن تكون جزء منه وليست كيانات حرفية قائمة بحد ذاتها. بمعنى آخر لا توجد فتحة لوحدها بل يوجد حرف مفتوح ولا ضمة لوحدها بل يوجد حرف مضموم وهكذا بالنسبة لبقية الحركات. وهذا يتطلب أن نوفر تمثيل لكل الحروف المشكلة ضمن نظام التمثيل الجديد. أما إمكانية إدخالها من خلال لوحة المفاتيح فهي بتوفير مفتاح خاص لكل حركة يعمل بنفس أسلوب المفتاح Shift في الحروف المنقطة. ونرى أن تكون للفتحة والضمة والكسرة مفاتيح تزيد قيمة الحرف بقيم هي 32 و 64 و 96 على التوالي. أي لو فرضنا إن قيمة تمثيل الحرف ب في التمثيل الجديد تساوي 31 مثلا فإن قيم تمثيلات الباء المفتوحة يساوي 63 والباء المضمومة يساوي 95 والباء

المكسورة يساوي 127 وهكذا بالنسبة لبقية الحروف, أما الحركة نفسها فيمكن كتابتها لوحدها بضغط المفتاح الخاص بها مع الفراغ.

أما التتوين والسكون فتمت معالجتهما بإضافة مفتاح خاص أيضاً إذا ضُغَط مع الحرف كان الحرف ساكن وإذا ضُغَط مع مفتاح حركة وحرف كان تتوين الحركة فوق الحرف. والمعالجة الرمزية لتمثيل الحرف الساكن والمنون تكون بإضافة 128. وعلى نفس مثالنا السابق فإن الباء الساكن (ب) تكون قيمة تمثيله في النظام الجديد $128+31=159$, و الباء المنون بالفتح (ب) فقيمة تمثيله تساوي $128+63=191$, والباء المنون بالضم (ب) فقيمة تمثيله تساوي $128+95=223$, والباء المنون بالكسر (ب) فقيمة تمثيله تساوي $128+127=255$ والشكل رقم 8 يوضح مثال لحرف الصاد.

الملاحظة	القيمة العشرية	التمثيل الثماني	الحرف
حرف الصلّا بدون حركة	16	00010000	ص
32+16	48	00110000	ص
64+16	80	01010000	ص
96+16	112	01110000	ص
128+16	144	10010000	ص
128+48	176	10110000	ص
128+80	208	11010000	ص
128+112	240	11110000	ص

شكل رقم 8. تمثيل حرف الصاد وقيمه مع الحركات

ومن الجدير بالذكر إن الحرف في التمثيل الجديد له ثمانية صور فمثلا هنالك ثمانية صور لحرف الباء هي: ب, ب, ب, ب, ب, ب, ب, ب, والصورة الأولى (صورة عدم وجود حركة) هي حالة جواز أكثر من حالة حركة كما في القرآن الكريم أو حالة الكتابة الغير مشكّلة.

هنالك بعض الحروف الموجودة في لوحة المفاتيح الحالية وغير موجودة في لوحة المفاتيح الجديدة المقترحة وهي (و, ئ, ي, آ, أ, إ, ا) وهي موجودة في ولكن لست بشكل ظاهر على إحدى المفاتيح وإنما من ضغط حرفي الهمزة والألف مع إحدى الحركات وحسب القواعد الإملائية العربية عدا الألف المقصورة فتكتب بالضغظ على حرف الألف بدون الحركة مرتين والألف الممدودة وهي ألف ساكنة يليها ألف مفتوحة علماً إن الألف الموصولة (ا) فهي ألف ساكنة.

3-7 المفتاح الجديد "ENG"

إن جداول الترميز عادة يجب أن تحتوي على حروف سيطرة على الطباعة و لوحة المفاتيح وعلى تراسل البيانات وتسمى الحروف غير المطبوعة (non-printable characters) وعددها 32 حرفاً وهي في جدول الآسكي تمثّل بالقيم من (0-31)، وعليه فأن عدد الحروف التي يجب أن يتضمنها جدول الترميز يساوي عدد الحروف المطبوعة مضافاً لها عدد الحروف الغير مطبوعة (أي 240+32) ويساوي 272 وهذا العدد من الحروف يفوق العدد المحدد للتمثيل الثماني (ثمان بتات) ونحن في هذا البحث لا نرغب باستخدام تمثيل آخر لأنه لا يتناسب مع البناء الجهازي للحاسوب ولا نريد زيادة الحجم المخصص لخزن الحرف الواحد لذا لابد من إيجاد طريقة لحل هذا الإشكال بما يضمن لنا استخدام التمثيل الثماني للحروف.

الحل المناسب لهذه المشكلة هو بإضافة مفتاح على لوحة المفاتيح يتيح لنا الانتقال بين جدولين من التمثيلات ليكن الجدول الأول هو جدول الآسكي والثاني سيكون جدول تمثيل الحروف العربية المقترح بحيث تكون مدخل قيمة المفتاح غير مستخدم في الجدولين وهذه القيمة هي 255. سيحمل المفتاح الجديد تسمية "ENG" وطريقة عمله تشبه طريقة عمل المفتاح "Caps Lock" أي عند الضغط عليه نستخدم الترميز المعروف بالآسكي ولا ينتقل إلى الترميز الجديد إلا بالضغط عليه مرة أخرى لإطفائه.

الفائدة الأساسية للمفتاح الجديد "ENG" هي الانتقال بين التمثيل الجديد و الآسكي بإضافة حرف قيمته 255 في متن النص للدلالة على الانتقال، وهناك فائدة لا تقل أهمية عن سابقتها وهي المطابقة "Compatibility" مع لوحات المفاتيح الموجودة حالياً.

5-7 توزيع الحروف العربية على اللوحة المقترحة

ولأن لكي نوزع الحروف العربية على لوحة المفاتيح المقترحة نحتاج لدراسة تكرارات الحروف العربية ليكون كل حرف في المفتاح المقابل للإصبع الذي يتناسب مع كمية تكراره وحسب ما تم توضيحه في الفقرات السابقة. والشكل رقم 9 يمثل تكرارات الحروف بعد دمج الحروف المنقطه مع الحروف الغير منقطه.

أ	ب	ن	ح	د	ر	س	ص	ط	ع	ف	ل	م	ه	و
٤	ي	ت	ج	ذ	ز	ش	ض	ظ	غ	ق	ك	ث	ة	خ
19.3	8.9	9.5	2.6	3.1	4.4	2.5	1.3	0.9	3.6	4.9	13.7	7.9	7.4	8.9

شكل رقم 9. تكرارات المفاتيح للوحة المقترحة

الجدول الجديد يحتوي لحد الآن على 241 مدخل (240 حرف + حرف الانتقال 255) والمتبقي من الجدول هو 14 مدخل تمثل الأرقام العشرة مع حرف الكشيده (التطويلة) وثلاثة مداخل تترك لاستخدامات حروف السيطرة.

والشكل رقم 10 يمثل التوزيع المقترح للحروف العربية على لوحة المفاتيح الجديدة.



شكل رقم 10
توزيع الحروف العربية على لوحة المفاتيح الجديدة

5-7 توزيع الحروف اللاتينية

لوحة المفاتيح المقترحة تحتوي على 15 مفتاحاً للحروف وهي مساوية لعدد الحروف العربية الغير منقطه والموضحة في الفقرات السابقة ولكي تكون لوحة المفاتيح ثنائية اللغة يجب أن تتسع لإضافة حروف لغة أخرى غير العربية و اللغة الانكليزية هي اللغة المقترحة في بحثنا.

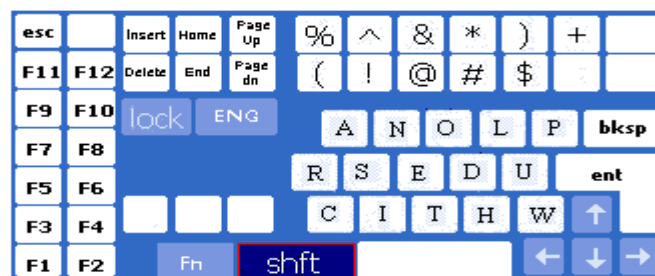
والمشكلة هنا تكمن في كون عدد حروف اللغة الانكليزية يساوي 26 حرفا وان عدد مفاتيح اللوحة المقترحة هو 15 مفتاحاً أي أن هنالك 11 حرفاً ليس لها مجال على لوحة المفاتيح. والحل المقترح هو بتوزيع أعلى خمسة عشر حرفاً تكراراً من الحروف الانكليزية على مفاتيح اللوحة الجديدة وتوزيع الحروف المتبقية بصورة

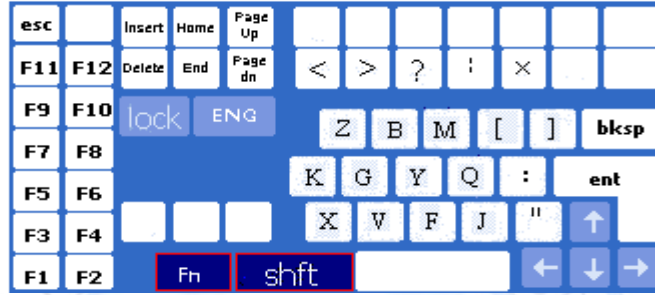
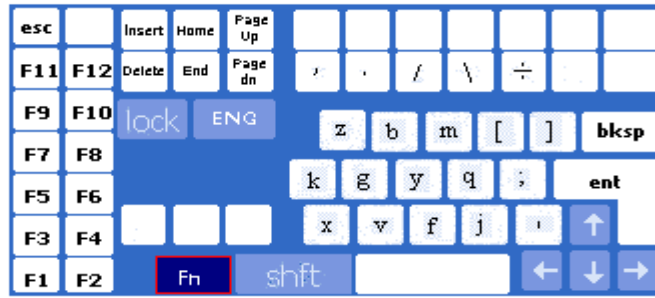
مشتركة مع الحروف الموزعة سابقاً ويستخدم مفتاح التتوين (أو FN كما هو موجود في لوحات المفاتيح الحالية (للانتقال إليها علماً إن مفتاح "Shift" يستخدم للانتقال إلى الحروف الانكليزية الكبيرة. والشكل رقم 11 يمثل نسبة تكرارات الحروف الانكليزية في النصوص القياسية.

Z	Y	X	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B	A
0.1	2.3	0.1	2.3	1.0	3.5	8.0	7.3	7.0	0.02	2.4	8.1	7.1	1.7	4.1	0.9	0.1	7.1	2.7	1.6	1.7	11.4	2.9	4.5	1.4	6.3

شكل رقم 11 نسب تكرارات الحروف الانكليزية في النصوص القياسية

أما الشكل رقم 12 فانه يمثل مقترحنا في توزيع الحروف الانكليزية على لوحة المفاتيح الجديدة. ويحتوي الشكل على أربعة نسخ من لوحة المفاتيح النسختين الأولى والثانية الحروف التي تظهر على لوحة المفاتيح عند الضغط على المفتاح ENG والأولى بدون استخدام المفتاح Shift والثانية بالضغط على المفتاح Shift مع الحرف. أما النسختين الثالثة والرابعة فهي الحروف التي يمكن استخدامها على لوحة المفاتيح بعد الضغط على المفتاح FN. وفي الشكل ترى النسخة الرابعة تظهر باستخدام المفتاح Shift بالإضافة للمفتاح FN. ومن الجدير بالذكر إن تمثيل البيانات اللاتينية يتم باستخدام نظام الترميز القديم وهو الآسكي. وفي حالة المعلومات المخزونة فان الفاصلة بين المعلومات الممثلة بالترميز الجديد والمعلومات الممثلة بالترميز القديم هي البيت ذو القيمة 255 كما اشرنا سابقاً في قيمة المفتاح ENG.





شكل رقم 12
توزيع الحروف الانكليزية على لوحة المفاتيح الجديدة

الاستنتاجات

بيننا في هذا البحث إن من صنع الحاسبة ووجد التمثيل الرياضي للحروف كان قد أوجده لكي يلائم حروف اللغة العربية وليست الحروف اللاتينية. فالتمثيل الثماني للمعلومات يبدو انه قد تم إعداده ليتناسب مع تمثيل الحروف العربية (خمسة بتات لمثيل الحروف وثلاثة بتات لتمثيل الحركات والتفتيط). لاحظ الجدول رقم 7. تطبيق البحث بحاجة إلى تبنى من قبل إحدى الشركات الكبرى في مجال التطبيقات الحاسوبية إذ لا يمكن تنفيذ و بناء لوحة المفاتيح المقترحة بدون اعتماد لطريقة تمثيل الحروف العربية في التطبيقات الحاسوبية. يمكن اعتماد نفس لوحة المفاتيح المقترحة في لغات أخرى تستخدم شكل الحرف العربي مثل اللغة الكردية والفارسية وغيرها .

المصادر

1. عبد الكريم عكله, "نظام تعريب وتقليص جديد" مجلة كلية بغداد للعلوم الاقتصادية, العدد الثاني والعشرون, 2009.