



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

UNIDAD ACADÉMICA PROFESIONAL TIANGUISTENCO

“Optimización multicriterio de teclados basado
en ergonomía para náhuatl y maya”

TESIS

Para Obtener el Título de
Maestro en Ciencias de la Computación

Que Presenta

Ing. Luis Díaz Rosas

Tutor académico:
Dr. René Arnulfo García Hernández

Tutores adjuntos:
Dra. Yulia Nikolaevna Ledeneva
Mtro. Rafael Cruz Reyes

TIANGUISTENCO, MÉX.

Diciembre 2018



UAEM | Universidad Autónoma del Estado de México

DICTÁMEN DE AUTORIZACIÓN Y OBTENCIÓN DE GRADO DE MAESTRÍA

Tianguistenco, Méx. , a 28 de noviembre de 2018



Título del proyecto:

Optimización multicriterio de teclados basado en ergonomía para náhuatl y maya.

Asesorista:

Ing. Luis Díaz Rosas

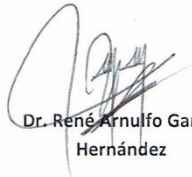
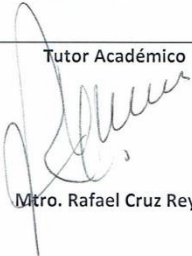

Dictamen:

No. de revisión: 6

- Rechazado
- Sujeto a modificaciones
- Aceptado, condicionado
- Aceptado

Observaciones generales:

Aceptado para la impresión
Aceptado para la defensa de grado

<p>Tutor Adjunto</p>  <p>Dr. René Arnulfo García Hernández</p>	<p>Tutor Académico</p>  <p>Mtro. Rafael Cruz Reyes</p>	<p>Tutor Adjunto</p>  <p>Dra. Yulia Nikolaevna Ledeneva</p>
---	---	--

U.A.E.M. POSGRADO CIENCIAS DE LA COMPUTACIÓN

Agradecimientos

Agradezco principalmente a dios por darme la oportunidad de culminar con mis estudios de maestría y ahora tener la oportunidad de continuar con un doctorado. También agradezco a personas que me brindaron todo su apoyo para poder lograr una meta más en mi formación personal; mi asesor el Dr. René Arnulfo García Hernández que en ningún momento dejo que mi tema o mi interés se perdiera y estuvo al pendiente de mí en cada revisión para poder culminar esta tesis, a mis padres Luis Angel Díaz Juárez e Irene Rosas Rodríguez que en todo momento pude contar con su apoyo y que alentaron a continuar estudiando.

Dedicatoria

Esta tesis se la dedico a mis padres Luis Angel Díaz Juárez e Irene Rosas Rodríguez, ya que siempre he contado con todo su apoyo para cualquier proyecto que me he propuesto, así mismo por estar siempre pendiente de mí en cada paso que doy tanto en mi vida como en mi formación personal. Les dedico esta tesis a mis padres ya que sin su apoyo no hubiese sido posible llegar a completar este objetivo de ser maestro en ciencias de la computación.

Contenido

Página

LISTA DE FIGURAS.....	I
LISTA DE TABLAS.....	III
1. INTRODUCCIÓN.....	4
1.1 Antecedentes	6
1.2 Planteamiento del problema	17
1.3 Objetivo general	18
1.3.1 Objetivos específicos.....	18
1.4 Hipótesis.....	18
1.5 Justificación.....	19
1.6 Estructura de la tesis	19
1.7 Resumen del capítulo	20
2. MARCO TEÓRICO.....	21
2.1 Algoritmos genéticos.....	21
2.1.1 Pasos de un algoritmo genético	24
2.1.2 Codificación de problemas	25
2.1.3 Evaluación de aptitud	27
2.2 Operadores genéticos.....	27
2.2.1 Operadores de selección	28
2.2.2 Operadores de cruce	29
2.2.3 Mutación.....	32
2.3 Origen de las lenguas indígenas	33
2.3.1 Descripción de alfabetos.....	36
2.4 Resumen del capítulo	36
3. ESTADO DEL ARTE	37
3.1 Optimización de teclados y sus métodos	37
3.2 Algoritmos genéticos aplicados para la optimización de distribuciones de teclados.....	49
3.3 Resumen del capítulo	53
4. MÉTODO PROPUESTO	54
4.1 Método propuesto.....	54
4.1.1 Configuración del algoritmo genético	57
4.1.2 Codificación propuesta	57
4.1.3 Función de aptitud del método propuesto.....	62

4.1.4 Determinación de frecuencia de caracteres	65
4.1.5 Operadores genéticos.....	66
4.2 Resumen del capítulo	67
5. EXPERIMENTACIÓN.....	68
5.1 Descripción de un corpus	68
5.2 Evolución del algoritmo genético y optimización de parámetros de configuración	70
5.3 Comportamiento de las 4 distribuciones de teclado optimizadas por el algoritmo genético	75
5.4 Curva de aprendizaje en nuevas distribuciones de teclado	90
5.5 Resumen del capítulo	94
6. CONCLUSIONES	95
6.1 Trabajo futuro.....	96
REFERENCIAS	98

Lista de figuras

Figura 1.1 Primer teclado presentado por Sholes en 1868 (Hempstalk, 2006).	7
Figura 1.2 Primera distribución de teclado <i>qwerty</i> presentada en 1873 (Resano, 2009).	8
Figura 1.3 Teclado en orden alfabético y teclado <i>qwerty</i> (García, 2012)	8
Figura 1.4 Diseño de la distribución del teclado estándar <i>qwerty</i> (Hempstalk, 2006).	9
Figura 1.5 Técnica de digitación para el teclado <i>qwerty</i> (Hempstalk, 2006).	10
Figura 1.6 Distribución de teclado <i>azerty</i> para el lenguaje francés presentada en 1890 (Centella, 2010).	11
Figura 1.7 Distribución de teclado <i>qwertz</i> para el lenguaje alemán (Centella, 2010).	12
Figura 1.8 Distribución de teclado <i>dvorak</i> para el lenguaje inglés presentada en 1936 (Centella, 2010).	12
Figura 1.9 Distribución de teclado <i>dvorak</i> para el idioma español (Walker, 2003).	12
Figura 1.10 Teclado <i>qwerty</i> para computadora (Cook, 2014).	15
Figura 1.11 Teclado <i>qwerty</i> (en forma ergonómica) para computadora, ejemplo 1 (Cook, 2014).	15
Figura 1.12 Teclado <i>qwerty</i> (en forma ergonómica) para computadora, ejemplo 2 (Cook, 2014).	16
Figura 2.1 Conformación de la población de un algoritmo genético.	22
Figura 2.2 Individuo y sus partes que lo caracterizan.	23
Figura 2.3 Esquema general de un algoritmo genético.	24
Figura 2.4 Pasos de un algoritmo genético.	25
Figura 2.5 Codificación binaria.	26
Figura 2.6 Codificación numérica.	27
Figura 2.7 Ejemplo de cruce en 1 punto.	30
Figura 2.8 Ejemplo de cruce en 2 puntos.	30
Figura 2.9 Ejemplo de cruce basado en PMX.	31
Figura 2.10 Ejemplo de cruce con base en orden.	31
Figura 2.11 Ejemplo de mutación por intercambio.	32
Figura 2.12 Ejemplo de mutación por inserción.	33
Figura 2.13 Ejemplo de cruce con base en orden.	33
Figura 2.14 Alfabetos de las lenguas español, náhuatl y maya.	36
Figura 3.1 Teclado optimizado por Eggers para el lenguaje inglés (Eggers, 2002).	38
Figura 3.2 Teclado optimizado por Eggers para el lenguaje alemán (Eggers, 2002).	39
Figura 3.3 Teclado optimizado por Eggers para el lenguaje francés (Eggers, 2002).	39
Figura 3.4 Teclado optimizado por Deb para el lenguaje hindi (Deb, 2003).	40
Figura 3.5 Teclado optimizado por Walker para el lenguaje inglés (Walker, 2003).	40
Figura 3.6 Teclado <i>capewell - dvorak</i> (Capewell, Michael, 2005).	41
Figura 3.7 Teclado <i>capewell - qwerty</i> (Capewell, Michael, 2005).	42
Figura 3.8 Teclado <i>capewell-adjacent</i> (Capewell, Michael, 2005).	42
Figura 3.9 Teclado <i>capewell - close</i> (Capewell, Michael, 2005).	42
Figura 3.10 Teclado <i>klauster2</i> (Capewell, Michael, 2005).	43
Figura 3.11 Teclado <i>arensito</i> (Capewell, Michael, 2005).	43
Figura 3.12 Teclado <i>qwerty</i> optimizado para la programación en el lenguaje C (Brewbaker, 2008).	44
Figura 3.13 Teclado <i>qwerty</i> optimizado para el proyecto GUTENBERG (Brewbaker, 2008).	44
Figura 3.14 Teclado <i>fitaly</i> con diseño diferente al <i>qwerty</i> (Brewbaker, 2008).	45

Figura 3.15 Teclado <i>fitaly</i> optimizado para la programación en el lenguaje C (Brewbaker, 2008).....	45
Figura 3.16 Teclado <i>fitaly</i> optimizado para el proyecto GUTENBERG (Brewbaker, 2008).	45
Figura 3.17 Teclado <i>carpalx</i> (Valle, 2011).....	46
Figura 3.18 Teclado optimizado para el lenguaje persa (Samimi, 2012).....	47
Figura 3.19 Teclado optimizado por Hosny para el lenguaje árabe (Hosny, 2015).....	47
Figura 4.1 Codificación de los individuos y sus genes.	61
Figura 4.2 Representación de los genes de un individuo sobre una distribución de teclado como una posible solución.	62
Figura 4.3 Separación de filas por columnas en el teclado.	63
Figura 4.4 Digitación de la palabra ‘TECLADO’ para calcular la distancia recorrida sobre el teclado.....	65
66	
Figura 4.5 Corpus con frecuencia de caracteres.....	66
69	
Figura 5.1 Corpus para cada optimización.....	69
Figura 5.2 Valor de la función de aptitud evaluando la distancia: experimento 1.	71
Figura 5.3 Valor de la función de aptitud evaluando la distancia: experimento 2.	72
Figura 5.4 Valor de la función de aptitud evaluando la distancia: experimento 3.	73
Figura 5.5 Valor de la función de aptitud evaluando la distancia: experimento 4.	74
Figura 5.6 Curva de aprendizaje de un teclado no optimizado: experimento 5.	93
Figura 5.7 Curva de aprendizaje de un teclado optimizado: experimento 6.	93
Figura 5.8 Comparación de una curva de aprendizaje en color azul con un teclado no optimizado (sin criterios ergonómicos) y una curva de aprendizaje en color naranja con un teclado optimizado (con criterios ergonómicos)	94

Lista de tablas

Tabla 3.1 Optimización de distribuciones de teclado implementando distintos métodos.....	48
Tabla 3.2 Optimización de distribuciones teclados implementando algoritmos genéticos.	52
Tabla 4.1 Alfabeto de la lengua náhuatl.....	58
Tabla 4.2 Alfabeto de la lengua maya.....	58
Tabla 4.3 Alfabeto de la lengua en español.....	59
Tabla 4.4 Signos y símbolos de puntuación más frecuentes de las lenguas náhuatl, maya y español.....	59
Tabla 4.5 Valor de cada gen en un individuo (decodificación).....	61
Tabla 5.1 Distancia total recorrida en metros.....	78
Tabla 5.2 Carga de trabajo por las 3 filas principales del teclado.....	78
Tabla 5.3 Carga de trabajo por cada dedo de ambas manos.....	79
Tabla 5.4 Carga de trabajo en ambas manos.....	79
Tabla 5.5 Distancia total recorrida en metros.....	81
Tabla 5.6 Carga de trabajo por las 3 filas principales del teclado.....	82
Tabla 5.7 Carga de trabajo por cada dedo de ambas manos.....	82
Tabla 5.8 Carga de trabajo en ambas manos.....	83
Tabla 5.9 Distancia total recorrida en metros.....	85
Tabla 5.10 Carga de trabajo por las 3 filas principales del teclado.....	85
Tabla 5.11 Carga de trabajo por cada dedo de ambas manos.....	86
Tabla 5.12 Carga de trabajo en ambas manos.....	86
Tabla 5.13 Distancia total recorrida en metros.....	88
Tabla 5.14 Carga de trabajo por las 3 filas principales del teclado.....	89
Tabla 5.15 Carga de trabajo por cada dedo de ambas manos.....	89
Tabla 5.16 Carga de trabajo en ambas manos.....	90
Tabla 5.17 Iteraciones con cada 5 caracteres en los lenguajes náhuatl y español.....	92



1. Introducción

México es uno de los 10 países en todo el mundo que cuenta con más lenguas originarias; ocupando el segundo lugar en Latinoamérica. De acuerdo con el instituto nacional de las lenguas indígenas (INALI), México cuenta con 69 lenguajes nacionales, 68 de estos lenguajes son lenguajes indígenas y el lenguaje español completa los 69 lenguajes. Dentro de los 68 lenguajes indígenas surgen 364 variantes lingüísticas. Los datos proporcionados por (Cultura, 2018) muestran que dentro de estas 364 variantes lingüísticas, se encuentran más de 25 millones de mexicanos que hablan alguna de estas variantes y se reconocieron como mexicanos indígenas.

Existen 11 familias lingüísticas para poder identificar el tipo de lenguaje indígena o el origen de estas, las 11 familias lingüísticas son:

1. Álgica
2. Yuto-nahua
3. Cochimí-yumana
4. Seri
5. Oto-mangue
6. Maya
7. Totonaco-tepehua
8. Tarasca
9. Mixe-zoque
10. Chontal de Oaxaca
11. Huave

El lenguaje náhuatl pertenece a la segunda familia lingüística Yuto-nahua. El lenguaje indígena náhuatl tiene 30 variantes las cuales nacen en diferentes estados de la república mexicana. Entre estas 30 variantes del lenguaje náhuatl, 3 de estas se encuentran en el estado de México y una variante más en el distrito federal.

Por otro lado, el lenguaje maya pertenece a la sexta familia lingüística y no cuenta con variantes. Sin embargo, este lenguaje se extiende cubriendo la mayor parte del estado de Campeche y Quintana Roo (Sánchez, 2010).

De acuerdo con una de las encuestas más recientes realizada en México en el año 2015, los hablantes de los lenguajes indígenas se han ido reduciendo. Los expertos apuntaron que los hablantes de los lenguajes indígenas se están reduciendo de 7.6% desde 1990 hasta 6.5% en 2015. Cada vez los hablantes de los lenguajes indígenas son menos y algunas de estos lenguajes estas en peligro de desaparición (UNIVERSAL, 2018).

El peligro de la desaparición de los lenguajes indígenas de México se debe a diferentes factores como:

- La falta de transmisión un lenguaje desde las generaciones mayores hacia la población infantil.
- Cuando un lenguaje indígena no se usa en espacios formales y se le considera exclusiva "de los indígenas", se va desterrando de sus hablantes hasta convertirse en un lenguaje con un menor porcentaje de hablantes entrando en riesgo de su desaparición.
- Los hablantes dejan de usar su lenguaje indígena por vergüenza y a causa de hechos discriminatorios.
- Cuando un lenguaje indígena deja de usarse en todos los contextos sociales comunicativos, va limitándose al uso comunal o familiar.
- Los procesos migratorios como el cambio de residencia que conlleva en muchos de los casos la necesidad de hablar un lenguaje distinto la propio. Por tanto, el lenguaje materno se deja de usar.
- La discriminación desde el interior mismo de los hablantes de los lenguajes indígenas entre su propia población.
- La preservación de los lenguajes indígenas como piezas de museo.
- La falta de información en formato digital para transmitir este conocimiento a otras partes, poblaciones o futuras generaciones.

Desde la aparición de los lenguajes indígenas, la forma de comunicación entre sus habitantes ha sido de forma verbal la mayor parte del tiempo. Poco después, con la creación de los dispositivos electrónicos los hablantes de los lenguajes indígenas han utilizado los teclados de estos medios para poder comunicarse ingresando texto. El teclado de los dispositivos electrónicos se ha utilizado para mantenerse en comunicación ya sea conversando, subiendo o consultando información en la red en algún lenguaje.

Como se mencionó anteriormente, la falta de información en formato digital para transmitir este conocimiento a otras partes, poblaciones, o futuras generaciones, es uno de los factores relevantes en la desaparición de los lenguajes indígenas, ya que las nuevas generaciones no pueden aprender a hablar y escribir algún lenguaje porque se necesita de información para que ésta pueda ser transmitida. Esto contribuye a que no se cuenta con mucha información en formato digital para que cualquier persona tenga acceso. Al no contar con la información suficiente se causa que cada lenguaje no sea transmitido a las nuevas generaciones y por consecuencia se pone en riesgo su permanencia.

En esta tesis se proponen una solución para facilitar el ingreso de información electrónica para los lenguajes maya y náhuatl. Así, cualquier persona accederá a la información recolectada para hacer uso de ella o para transmitirla a las futuras generaciones. Actualmente para almacenar y acceder a la información en formato digital, la forma más práctica es usando un dispositivo electrónico el cual cuenta con un teclado que el teclado vincula la persona con el dispositivo.

1.1 Antecedentes

¿Qué es el teclado y cómo funciona? El teclado es uno de los dispositivos que se ha convertido en un medio de comunicación indispensable, se ha utilizado para ingresar, enviar o recibir información (en formato texto) por medio de dispositivos electrónicos o mecánicos. El origen del teclado se dio en las máquinas de escribir con el objetivo de ingresar texto para poder ser interpretado y leído por cualquier persona. El primer invento de un teclado se dio a inicios de 1868 por Christopher Sholes. Este teclado tenía un diseño (distribución) similar a las teclas de un piano (figura 1.1). Este primer diseño de teclado contaba con solo dos filas en las cuales las teclas estaban ordenadas en orden alfabético (de izquierda a derecha empezando por los números seguido de las letras de la A a la Z) (Hempstalk, 2006). Este diseño contaba con muchas dificultades ya que la posición de sus teclas ocasionaba una digitación muy lenta al recorrer una gran distancia de tecla a tecla. Al momento de escribir sobre el teclado y empezar a tomar velocidad, los martillos de las teclas se quedaban atorados (García, 2012),

causando el atascamiento de dos o más teclas. Esto a su vez causaba pérdida de tiempo para el mecanógrafo al tener que regresar cada una de las teclas atoradas a su posición original para continuar con la digitación (Hempstalk, 2006).

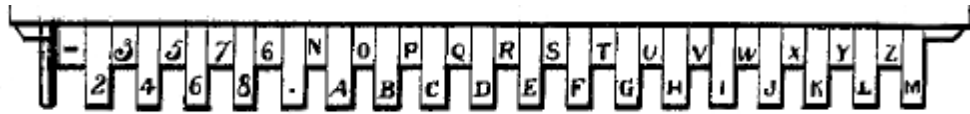


Figura 1.1 Primer teclado presentado por Sholes en 1868 (Hempstalk, 2006).

Otro de los problemas del primer teclado era que los dedos de los mecanógrafos se desplazaban demasiado por todo el teclado. Esto ocasionaba que entre los desplazamientos los dedos chocaban entre sí. Para poder solucionar este problema, se proponía distanciar las parejas o bígamas de teclas que se utilizaban con una mayor frecuencia. Sin embargo, se seguían haciendo análisis más a fondo para poder determinar una nueva ubicación de teclas. Pasaron varios años poder corregir estas dificultades de las teclas atoradas. Durante este tiempo se realizaron varios intentos tratando de crear un teclado nuevo con mejores características. La mayoría de estos intentos fueron fallidos ya que no se podía llegar a un teclado óptimo y sin dificultades de escritura; como las que presentaba el primer teclado (Fernández, 2013). El principal objetivo de crear teclados nuevos fue obtener un diseño que facilite la tarea de digitación.

En 1873 se presentó un teclado nuevo con una distribución optimizada para corregir el problema de teclas atoradas. Este teclado tenía un diseño físico totalmente distintito al del primer teclado con el objetivo de aumentar la velocidad al escribir (García, 2012). Esta distribución nueva se le denominó teclado *qwerty* (figura 1.2), nombre que provenía de las primeras 6 letras que se encontraban en la parte superior derecha de la fila superior (debajo de la fila numérica). El teclado *qwerty*, patentado por C. Sholes (Centella, 2010), aceleró el ritmo de las comunicaciones, marcó un punto importante en el desarrollo de las relaciones sociales y dio oportunidad a la mujer de poder entrar masivamente en el mundo laboral (Fernández, 2013). Las personas que por primera vez ponían sus manos sobre el teclado *qwerty*, les parecía un teclado extraño, ya que habían pasado de escribir de teclado con 2 filas en orden alfabético a un teclado con 4 filas (*qwerty*) en el que no se comprendía por qué se habían acomodado las teclas. Así, Seymour Papert describe al teclado *qwerty* como: “las

teclas de las máquina de escribir no están agrupadas de acuerdo a como más se utilizan, sino de una manera que, a primera vista parece caótica" (figura 1.3) (García, 2012). Este diseño patentado de teclado (*qwerty*) se mantuvo como un diseño estándar que con el paso del tiempo C. Sholes hizo algunas mejoras a esta distribución (figura 1.4). Así, la distribución de teclado *qwerty* pasa a ser un diseño de teclado estándar para las máquinas de escribir, las computadoras, y otros dispositivos electrónicos (Hempstalk, 2006). Cabe mencionar que la distribución de teclado *qwerty* se diseñó especialmente para el lenguaje inglés (Resano, 2009).

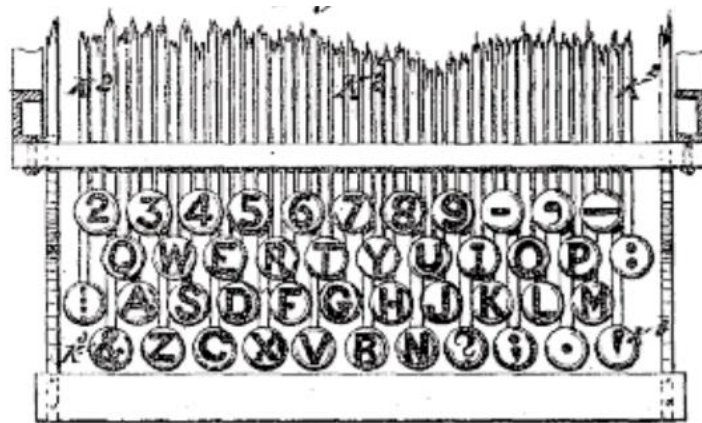


Figura 1.2 Primera distribución de teclado *qwerty* presentada en 1873 (Resano, 2009).

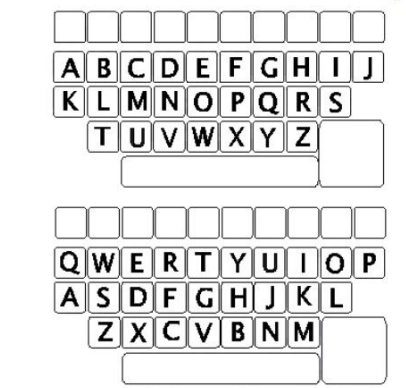


Figura 1.3 Teclado en orden alfabético y teclado *qwerty* (García, 2012)

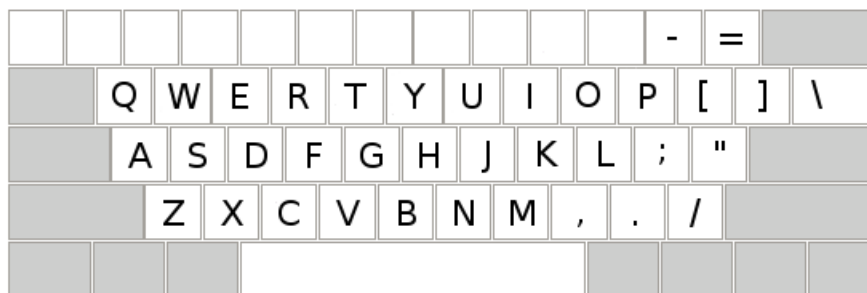


Figura 1.4 Diseño de la distribución del teclado estándar *qwerty* (Hempstalk, 2006).

En un artículo escrito por (García, 2012) se menciona que existe la posibilidad de que C. Sholes no haya contemplado el uso de los 5 dedos de ambas manos para la escritura en el teclado *qwerty*; por lo que cualquier técnica para escribir en este teclado era aceptable. Sin embargo, la mayoría de los mecanógrafos que empezaron a utilizar el teclado *qwerty* tenían diferentes formas de escribir en este teclado, lo que quiere decir que cada mecanógrafo tenía que adaptar su propia técnica de escritura. Algunos mecanógrafos utilizaban dos o tres dedos de cada mano, algunos otros escribían con un solo dedo de cada mano y en muy pocas ocasiones los mecanógrafos utilizaban los 5 dedos de ambas manos (Fernández, 2013). Para poder solucionar este problema, se consideró en cuenta la forma que tenía el teclado *qwerty*, tanto sus filas como sus columnas y así fue como se introdujo una técnica de digitación (escritura). En esta técnica de digitación se proponía utilizar los 5 dedos de ambas manos donde a cada dedo se le asignó un determinado grupo de teclas. En esta técnica de digitación se divide el teclado en 2 partes (una parte para la mano izquierda y otra parte para la mano derecha) sobre el teclado *qwerty* (figura 1.5). Así, los mecanógrafos pudieron aumentar su velocidad de digitación en la distribución de teclado *qwerty* (Hempstalk, 2006).



Figura 1.5 Técnica de digitación para el teclado *qwerty* (Hempstalk, 2006).

Los mecanógrafos que utilizaban el teclado *qwerty* se dieron cuenta que aún presentaba dificultades al no ser una distribución adecuada para escribir, incluso para el lenguaje inglés (Hempstalk, 2006). De esta manera, se presentaron nuevas distribuciones de teclado basadas en el diseño *qwerty*, tales como: el teclado *azerty* (figura 1.6), el cual era una modificación del mismo teclado *qwerty* cambiando la posición de sus teclas; Q por la letra A, W por la letra Z y algunos otros caracteres y signos de puntuación adecuados para el lenguaje francés.

El teclado *qwertz* (figura 1.7) es una modificación del mismo teclado *qwerty* cambiando la posición de su tecla Y por la letra Z y algunos otros caracteres y signos de puntuación adecuados para el lenguaje alemán (Centella, 2010).

En 1936, August Dvorak presenta un nuevo teclado para el lenguaje inglés llamado *dvorak* (figura 1.8). August se dio cuenta que muchas de las letras y caracteres más frecuentes en el

teclado *qwerty* no estaban en una posición cómoda y se realizaban grandes desplazamientos cuando se pasaba de una tecla a otra para la escritura en lenguaje inglés (Hempstalk, 2006). Agust afirmaba que el teclado *dvorak* tenía las características de ser más rápido en cuando a la digitación de textos en lenguaje inglés ya que su teoría fue poner las 5 vocales en la fila donde se posicionan los dedos al iniciar una digitación (fila origen). Con esto, Dvorak demostró que su teclado era un teclado más óptimo para la escritura en lenguaje inglés. Este nuevo teclado (*dvorak*) era totalmente diferente a la distribución de teclado *qwerty*. Poco después de la presentación del teclado *dvorak*, sus mecanógrafos se dieron cuenta que era más adecuado que el teclado *qwerty* por lo que se realizó una modificación del teclado *dvorak* para el lenguaje español (figura 1.9) (Walker, 2003). Sin embargo, la mayoría de sus mecanógrafos ya se habían acostumbrado a trabajar con el teclado *qwerty*, lo que hizo difícil su remplazo. Esto dio paso para que el teclado *qwerty* fuese un teclado estándar utilizándolo hasta hoy en día en diferentes dispositivos electrónicos; principalmente en el teclado de computadora (Hosny, 2015).

Desde la aparición del primer teclado, este dispositivo se ha considerado como uno de los dispositivos más importantes para la comunicación entre el humano con los dispositivos mecánicos y electrónicos, ya que permite ingresar texto que puede ser transmitido y procesado a diferentes fuentes y/ o personas (Samimi, 2012).



Figura 1.6 Distribución de teclado *azerty* para el lenguaje francés presentada en 1890 (Centella, 2010).



Figura 1.7 Distribución de teclado *qwertz* para el lenguaje alemán (Centella, 2010).



Figura 1.8 Distribución de teclado *dvorak* para el lenguaje inglés presentada en 1936 (Centella, 2010).

a o	!	"	·	\$	%	&	/	()	=	?	¿	
	1	2 @	3 #	4	5	6	7	8	9	0	'	i	
	:	;	Ñ	P	Y	F	G	C	H	L	^	*	
	.	,									~	+	
	A	O	E €	U	I	D	R	T	N	S	..	ç	ç
	>	-	Q	J	K	X	B	M	W	V	Z		
	<	_											

Figura 1.9 Distribución de teclado *dvorak* para el idioma español (Walker, 2003).

La optimización del tiempo es una de tareas más importantes para los seres humanos, cada vez se buscan diferentes alternativas para optimizar tareas con el objetivo de reducir el tiempo

de trabajo (Samimi, 2012). Un ejemplo claro es el proceso de adaptación por el cual ha pasado el teclado ya que a lo largo del tiempo ha ido evolucionando y adaptando a las necesidades del hombre. Esto refiere a que los teclados han pasado por un proceso de adaptación lingüística para optimizar el tiempo de escritura. Por ello, se han hecho modificaciones a los teclados anteriormente mencionados y se han creado nuevos teclados. De acuerdo con Samnimi, se llega a la conclusión de que la optimización de las teclas sobre un teclado puede mejorar el trabajo de los mecanógrafos en tiempo y esfuerzo. Para demostrar esto, diferentes investigaciones abordan a teclados optimizados, ya sea una modificación del teclado *qwerty* o una distribución de teclado totalmente diferente, por ejemplo:

- 1) Eggers (Eggers, 2002) implementa el método de optimización "colonia de hormigas" para optimizar tres teclados: uno para el lenguaje inglés, otro para el lenguaje francés y otro para el lenguaje alemán. Eggers retoma la técnica con los 5 dedos de ambas manos.
- 2) Deb (Deb, 2003) realiza una optimización de teclado para el lenguaje Hindi basado en un algoritmo genético tomando como referencia textos de diversas fuentes de internet.
- 3) Walker (Walker, 2003) se basa en un algoritmo genético para la optimización de teclado para el lenguaje inglés. Esta distribución de teclado se da con base en las 1,000 palabras más usadas en el lenguaje inglés.
- 4) Capewell (Capewell, Michael, 2005) usa un algoritmo genético para llegar a la optimización de diferentes teclados, tomando como base el teclado *qwerty* y el teclado *dvorak* para mejorarlos.
- 5) De manera similar, Brewbaker (Brewbaker, 2008) implementa como método de optimización un algoritmo genético para llegar a 2 distribuciones de teclado de forma optimizada. La primera distribución está enfocada a los programadores en lenguaje C y la segunda optimización está enfocada en la escritura de obras literarias en lenguaje inglés.
- 6) Samimi (Samimi, 2012) va más allá al utilizar dos métodos de optimización, ya que implementa el método de recocido simulado combinado con un algoritmo genético con el objetivo de crear una mejor distribución de teclado para el lenguaje persa; la cual se basa en las 19,092 palabras más frecuentes de dicho lenguaje.
- 7) En octubre de 2012 Bi (Bi, 2012) crea un solo teclado optimizado para los lenguajes: francés; español; alemán y chino. Bi menciona que no es necesario crear un teclado para

cada lenguaje y para ello crea un solo teclado para estos 4 lenguajes. Así mismo para poder llegar a esta optimización de teclado, Bi utiliza la distancia que existe de una tecla a otra. De este modo, Bi minimiza la mayor parte de las distancias para los 4 lenguajes implementado un modelo matemático.

- 8) Hosny (Hosny, 2015), con uno de los trabajos más recientes en el año 2015, logra optimizar una distribución de teclado para el lenguaje árabe. Hosny toma diferentes artículos de wikipedia y textos de periódicos en lenguaje árabe y así, implementa un algoritmo genético para crear su propia distribución optimizada de teclado para el lenguaje árabe.

Tal como se describe cada uno de los trabajos realizados para la optimización de distribuciones de teclado en las que se tienen diferentes objetivos, algunos de estos trabajos tratan de centrarse en un lenguaje en específico, algunos otros en uno o varios dominios de un lenguaje. Así mismo, existen teclados que fueron creados con la finalidad de ser útiles para diferentes lenguajes los cuales les llaman teclados multilenguaje. Cabe mencionar que muchos de los trabajos de optimización de teclado respetan la forma original del teclado *qwerty*, lo que quiere decir que solo se reubican las teclas de forma optimizada de acuerdo al método que se implementa o el objetivo deseado. Otras optimizaciones de teclado cambian por completo el diseño del teclado *qwerty*, lo cual se refiere a que son diseños de teclado totalmente diferente, lo que hace que sea más difícil remplazar al teclado estándar *qwerty*, ya que se tendrían que fabricar estos teclados para poder utilizarlos.

Al paso de la evolución del teclado y la aparición de nuevas distribuciones de teclado, se presentan dos formas para mejorar las dificultades del teclado, principalmente del teclado *qwerty*. La primera forma para mejorar las dificultades del teclado *qwerty* es reubicando las teclas proponiendo una nueva distribución de teclas. La segunda forma para mejorar las dificultades del teclado *qwerty* es cambiando la forma física que el teclado *qwerty* tiene. Siguiendo a la figura 1.10 se muestra un teclado *qwerty* para computadora en una forma original, respetando el diseño físico y la distribución de las teclas correspondientes. Siguiendo a la figura 1.11 se muestra un teclado *qwerty* para computadora separando al teclado en dos partes (parte para la mano izquierda y parte para la mano derecha) agregando una curvatura hacia adentro del teclado para proporcionar una mejor comodidad. Siguiendo la figura 1.12 se muestra un teclado *qwerty* para computadora con un diseño que se puede ajustar la distancia que existen entre las teclas correspondientes de la mano izquierda con las

teclas correspondientes de la mano derecha con el objetivo de que cada mecanógrafo ajuste la comodidad del teclado al escribir (Cook, 2014). Como se muestra en los diseños de las figuras 1.11 y 1.12 la distribución de teclas no se modifica, se respeta la distribución *qwerty* y solo se modifica el diseño físico del teclado.

Acorde con la definición de ergonomía dada por (Gómez, 2004) se dice que la ergonomía tiene como objetivo conseguir el mejor rendimiento: mientras el hombre piensa y acciona, el objeto con el que interactúa se va adaptando cada vez mejor a sus necesidades. El objetivo principal de la ergonomía es optimizar el trabajo a ejecutar por medio de un conjunto de criterios en un artefacto.

Por tanto, al presentar diseños de teclado que cambiaban su diseño físico (figura 1.11, figura 1.12) se podían adaptar a las necesidades de cada mecanógrafo. Estos teclados son más ergonómicos, ya que contrarrestan las dificultades que presentaba el teclado original *qwerty*.



Figura 1.10 Teclado *qwerty* para computadora (Cook, 2014).



Figura 1.11 Teclado *qwerty* (en forma ergonómica) para computadora, ejemplo 1 (Cook, 2014).



Figura 1.12 Teclado *qwerty* (en forma ergonómica) para computadora, ejemplo 2 (Cook, 2014).

Para contar con ergonomía en los teclados se puede hacer de dos formas: 1) cambiando el diseño físico, 2) re-ubicación sus teclas.

Para crear una distribución de teclado con criterios ergonómicos sin necesidad de cambiar diseño físico del teclado, se toman en cuenta diferentes criterios:

- Un conjunto de caracteres o palabras llamado "corpus"; este criterio modela la frecuencia de caracteres sobre el teclado.
- Carga de trabajo para cada dedo; este criterio balancea la carga de trabajo que tiene cada dedo dando prioridad (mayor carga) a los dedos que tienen mayor fuerza y habilidad (dedos índice y dedos medio)
- Carga de trabajo por fila; este criterio balancea la carga de trabajo que se tiene en las tres filas principales del teclado (fila superior, origen e inferior) dando prioridad a la fila donde se posicionan los dedos para iniciar una digitación y donde menor un número de movimientos realizan (fila origen).
- Carga de trabajo por mano; este criterio balancea carga de trabajo que tiene cada mano, de acuerdo con la figura 1.5 que muestra las dos partes del teclado para cada mano, se realiza un balance para tener un carga de trabajo similar en ambas manos,
- La distancia; este criterio minimiza la distancia que existe de un tecla a otra, entre menor sea la distancia entre teclas al escribir un conjunto de palabras, los movimientos de los dedos son reducidos.

Al reducir la distancia se reducen los movimientos de las manos y posiciones incomoda, al balancear la carga de los dedos se evita la fatiga para los dedos que menos fuerza y habilidad tienen, al balancear la carga de trabajo por filas se reducen los desplazamientos innecesarios hacia las filas superior e inferior, al balancear la carga de trabajo para amabas manos, se tiene un flujo continuo de escritura ya que no existe fatiga para una sola mano. En conjunto la optimización de estos criterios puede evitar lesiones futuras en ambas manos o dedos y así se da paso a una distribución de teclado con criterios ergonómicos que brindan una mayor comodidad al escribir y esto facilita el aprendizaje de digitación en comparación con teclados que no cuentan con criterios ergonómicos.

1.2 Planteamiento del problema

La falta de información de un lenguaje es la causa de que éste no pueda ser transmitido a otras personas, esto reduce el número de hablantes y en un largo plazo el lenguaje puede verse afecto estando propenso a la desaparición. Esto se debe principalmente a dos factores:

- Las personas no saben escribir el lenguaje y por tanto la información no puede ser guardada en documentos o almacenada de forma digital (en la red).
- Las personas no cuentan con los caracteres propios del lenguaje en un teclado (de un dispositivo electrónico) o que el teclado presente dificultades para escribir en el lenguaje, hace que el ingresar texto del lenguaje por medio del teclado sea una tarea difícil.

Así mismo, otro de los principales factores que aceleran la desaparición de un lenguaje, es no tener una comunicación mutua entre su propia comunidad. Este problema reside en la comunicación escrita, ya que hoy en día el medio más accesible y fácil de comunicación es mediante dispositivos electrónicos por los cuales se envía y se recibe información en forma digital utilizando un teclado para poder ingresar dicha información. Al tener un teclado que presenta dificultades para escribir en un lenguaje, las personas dejan a un lado e tipo de comunicación escrita por medio de los dispositivos electrónicos, y así este factor también aumenta el riesgo de la desaparición de un lenguaje al tener cada vez menos personas.

Anteriormente se describieron trabajos de optimización de teclados y los diferentes métodos de optimización que se han utilizado para poder llegar a la optimización de cada una de las

distribuciones de teclados. El objetivo es adaptar los teclados a uno más lenguajes en específico para que la comunicación de forma escrita y la recolección de información sobre los lenguajes sea aún mayor.

De acuerdo a los factores que afectan en la desaparición de los lenguajes, surge la pregunta: ¿Cómo facilitar distribuciones de teclado con base en criterios ergonómicos para mantener y aumentar el uso de los lenguajes náhuatl y maya?

1.3 Objetivo general

El objetivo principal de este trabajo es facilitar 4 distribuciones de teclado para la redacción de textos y que cuenten con criterios ergonómicos minimizando la distancia recorrida, balanceando la carga de trabajo en las tres filas principales del teclado, la carga para cada uno de los dedos y balanceando la carga de trabajo para cada mano para la redacción de textos en los dos primeros lenguajes indígenas más habladas de México y en lenguaje español de México.

1.3.1 Objetivos específicos

1. Buscar un conjunto de textos (corpus) para modelar la frecuencia de caracteres en cada lenguaje.
2. Proponer un método de optimización para obtener 4 distribuciones de teclado bajo criterios ergonómicos; una distribución para el náhuatl, una distribución para el maya, una distribución para náhuatl y español y una distribución para maya y español.
3. Realizar evaluaciones de los criterios ergonómicos en las 4 distribuciones de teclado.
4. Verificar el comportamiento de las distribuciones obtenidas por el método propuesto.

1.4 Hipótesis

Si se configura la posición de las teclas sobre un teclado implementando un método de optimización, entonces se podrían encontrar distribuciones de teclado que sean más óptimas para ese lenguaje bajo criterios ergonómicos (menor distancia recorrida, carga de trabajo balanceada para cada dedo, carga de trabajo balanceada en las tres filas y carga de trabajo balanceada por mano).

1.5 Justificación

Actualmente existen diferentes distribuciones de teclado, algunas de estas creadas mediante distintos métodos de optimización sobre un lenguaje, y algunas otras son solo modificaciones de distribuciones ya existentes adaptándolas para algún lenguaje. La mayoría de los trabajos que se han realizado en cuanto a la optimización de distribuciones de teclado se centran en el lenguaje inglés.

Así mismo como ya se mencionó anteriormente, no se cuenta con una distribución de teclado que aporten apoyo para los lenguajes indígenas de México. El objetivo principal de este trabajo es apoyar a estos lenguajes autóctonos por medio de distribuciones de teclado las cuales cuenten con un alfabeto específico. Así mismo se busca que la distribución de sus letras sea óptima para brindar comodidad al escribir reduciendo la fatiga y un fácil aprendizaje del uso del teclado para el usuario. De este modo, las personas hablantes de los lenguajes indígenas podrán mantener y facilitar la comunicación por medio de dispositivos electrónicos. Así mismo se podrá hacer una recopilación de toda la información posible de estos lenguajes para poder tenerla almacenada de forma digital. Logrando estos objetivos se fomentará la comunicación en dichos lenguajes, ya que las personas hablantes de algún modo podrán usar su lenguaje materno mediante la comunicación escrita.

1.6 Estructura de la tesis

En el capítulo 1 se describe las causas principales del problema (la falta de apoyo hacia los lenguajes indígenas), y las formas en que los hablantes los lenguajes indígenas se comunican cuando utilizan los medios electrónicos. Así mismo, se describen los antecedentes y evolución del teclado explicando los diferentes trabajos realizados para poder mejorar o crear un teclado con mejores criterios ergonómicos y el proceso de adaptación que ha tenido el teclado en los dispositivos electrónicos.

En el capítulo 2 se hace mención sobre la teoría del método de optimización en "algoritmos genéticos", la metodología que se sigue para poder implementar este método, la configuración y adaptación para resolver uno o varios problemas. También se describen de forma detallada cada una de sus partes (operadores genéticos) y el uso de las diferentes combinaciones de operadores genéticos de acuerdo al problema que se desea resolver.

El capítulo 3 está conformado por dos apartados principales. El primer apartado menciona los distintos métodos de optimización que se han implementado para poder optimizar diferentes distribuciones de teclado. El segundo apartado de este capítulo describe de forma detallada los trabajos donde se han implementado los algoritmos genéticos como métodos para optimizar una o varias distribuciones de teclado y las configuraciones utilizadas en estos algoritmos.

El capítulo 4 da a conocer el método propuesto para optimizar las 4 distribuciones de teclado y las partes que lo conforman. Así mismo se describe cual es la metodología que se sigue para la implementación del método propuesto.

En el capítulo 5 se muestran las pruebas, evaluaciones y experimentos realizados después de haber obtenido las 4 distribuciones de teclado por medio del algoritmo genético. Así mismo, se hace una comparación de las 4 nuevas distribuciones de teclado contra las distribuciones de teclado ya existentes para demostrar la optimización de los criterios ergonómicos de las nuevas distribuciones de teclado.

1.7 Resumen del capítulo

En este capítulo, se introdujo y se dieron los antecedentes sobre el problema principal de esta tesis que es el apoyo para mantener e incrementar el uso de los lenguajes indígenas en México. Así mismo se justifica como el teclado es uno de los dispositivos más importantes para mantener la comunicación en un lenguaje.



2. Marco teórico

En este capítulo se da una descripción detallada sobre cómo funciona uno de los métodos más utilizados para la optimización de distribuciones de teclado, como lo son los algoritmos genéticos. En este capítulo se da a conocer cómo se pueden implementar los algoritmos para la resolución de problemas, cuáles son las partes que conforman a este método y los pasos que se deben seguir para poder hacer uso de estos algoritmos. Así mismo, se describe el origen y alfabeto de los lenguajes náhuatl, maya y español.

2.1 Algoritmos genéticos

Los algoritmos genéticos se caracterizan por ser métodos adaptativos los cuales se implementan para resolver problemas de búsqueda y optimización. Estos algoritmos están basados en el principio de Darwin, en la "supervivencia del más apto" (Gestal, 2010). Fueron creados por John Holland en los años 60s con el objetivo de imitar el principio mencionado anteriormente. Para ello, Holland parte de una población inicial de individuos conformados por genes únicos y propios que los caracterizan. En la primera población se seleccionan los individuos más aptos, para posteriormente reproducirlos y finalmente mutarlos. De esta forma se obtiene la siguiente población de individuos los cuales son mucho más aptos que la población anterior (Truyol, 2015). Así mismo, dentro de los puntos más importantes que Holland describe en los algoritmos genéticos, es que hay un proceso de adaptación en el cual cada uno de los individuos interactúa y va mejorando sus aptitudes con base en un cierto ambiente.

De este modo, los individuos se van adaptando cada vez de una mejor forma y las poblaciones futuras van mejorando (Coello, 2015).

Los algoritmos genéticos son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en la teoría de la evolución de Darwin, la cual establece que los individuos más aptos de una población son los que sobreviven ya que se adaptan fácilmente a los cambios que se producen en su entorno (Coello, 2015).

Tal como se mencionó anteriormente, los algoritmos genéticos inician con soluciones creadas aleatoriamente. A este conjunto de soluciones se le llama población y está formada por individuos (figura 2.1). Cada individuo representa una solución al problema y está conformado por genes (figura 2.2) que en su conjunto lo hacen único a los demás individuos de la población. De ésta forma, cada individuo nuevo se va a adaptando al problema de mejor manera (Sebag, 2005).

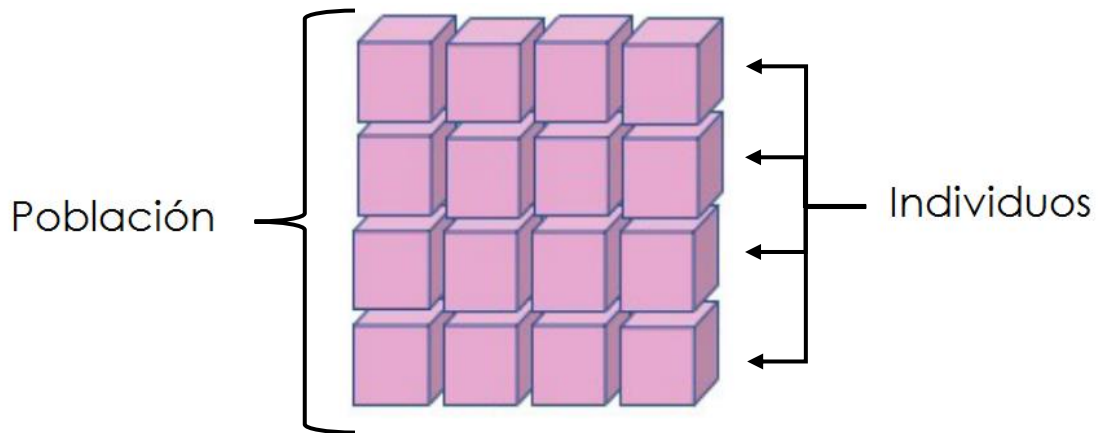


Figura 2.1 Conformación de la población de un algoritmo genético.

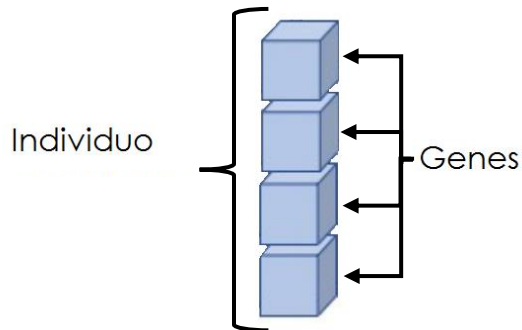


Figura 2.2 Individuo y sus partes que lo caracterizan.

Se dice que en la naturaleza todo proceso de evolución se realiza de forma natural, pero al aplicar un algoritmo genético para resolver problemas de acuerdo a dicha evolución se sigue una serie de pasos. Para ello, es necesario partir de una población inicial conformada por individuos (posibles soluciones al problema). Cada individuo es codificado de acuerdo al tipo de problema(s). Así mismo, cada individuo es evaluado con una función llamada "función de aptitud" (Coello, 2015). La población inicial puede estar establecida o en la mayoría de los casos se crea aleatoriamente. Posteriormente se siguen los siguientes pasos representados en la figura 2.3:

- Evaluar cada uno de los individuos de la población (llamados individuos padre).
- Seleccionar a los individuos más aptos de acuerdo a criterios establecidos.
- Permitir la reproducción de los individuos para empezar a crear una nueva población (llamados individuos hijos o descendientes).
- Tener una probabilidad de mutación para los genes de cada uno de los nuevos individuos (hijos).
- Organizar la nueva población (descendencia).

Estos pasos se repiten hasta una cierta condición de paro establecida, por ejemplo: alcanzar un número máximo de poblaciones, llegar a un valor deseado en la evaluación o cumplir condiciones de paro el no existir nuevos cambios en las últimas poblaciones (Truyol, 2015).

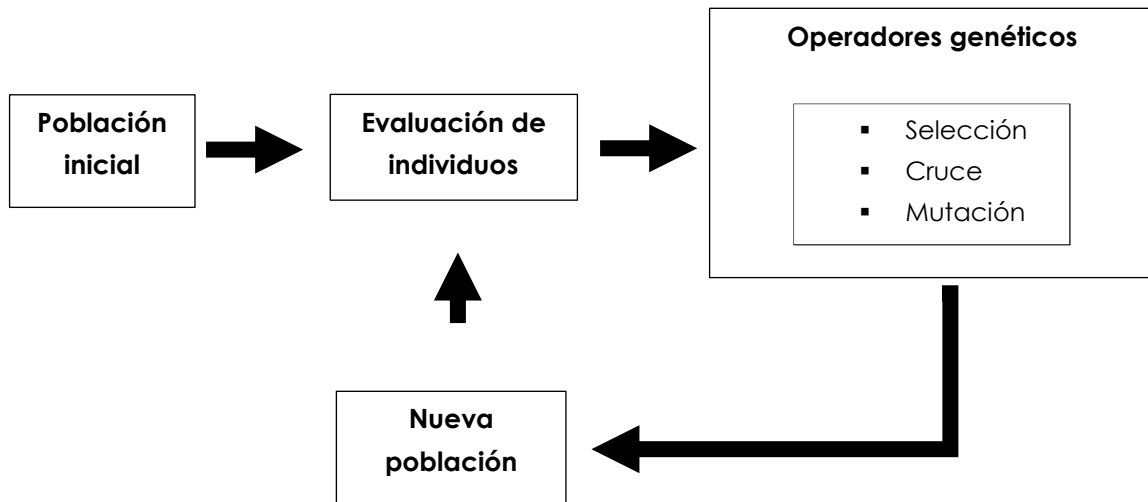


Figura 2.3 Esquema general de un algoritmo genético.

2.1.1 Pasos de un algoritmo genético

Teniendo un problema que se desea resolver, y que a su vez ya está claramente definido se puede representar la solución con una serie de características (el individuo es la solución y la serie de características son sus genes). De esta forma, al crear nuestra población inicial con un conjunto de individuos tenemos un conjunto de soluciones al problema. Un algoritmo genético completa el ciclo de evolución, de la siguiente manera (figura 2.4) (Melanie, 1998):

Pasos sobre un algoritmo genético:

1. Comenzar con una población P generada aleatoriamente de n individuos de k genes (posibles soluciones al problema).
2. Calcular el valor de cada individuo mediante la función de aptitud $f(x)$ para cada individuo x de la generación P .
3. Repetir los siguientes pasos hasta que se hayan creado n individuos (descendientes o hijos):
 - a. Seleccionar un par de individuos (individuos padres) de la población P , siendo la probabilidad de selección, criterios establecidos para saber que individuos son más aptos. Dentro de la selección que el o los mismos individuos pueden ser seleccionado en más de una ocasión para ser padres.
 - b. Con probabilidad P_c (*probabilidad de cruce, o tasa de cruce*), cruzar el par de individuos de acuerdo al tipo de cruce elegido acorde a la codificación de los individuos para formar dos descendientes. Así como para cada tipo de codificación existen tipo de cruce, cada codificación tiene otros tipo de cruce que se pueden combinar para obtener diferentes resultados y mejorar el algoritmo genético.
 - c. Con probabilidad P_m (*probabilidad de mutación, o tasa de mutación*), mutar los nuevos individuos y colocarlos en la nueva población P' .
4. Reemplazar la población actual P con la nueva P' .
5. Volver al paso 2.

Figura 2.4 Pasos de un algoritmo genético.

2.1.2 Codificación de problemas

Cualquiera que sea la solución a un problema, en el campo de los algoritmos genéticos una solución puede ser representada mediante un conjunto de parámetros que son los genes y se codifican en forma de cadena la cual se le denomina individuo. Cada gen tiene el valor de un solo parámetro (Coello, 2015). Por lo tanto, cada individuo contiene información acerca de la solución que representa al problema que se desea resolver. Esta codificación se puede realizar de diferentes formas ya que existen diferentes tipos de codificación que pueden ser valores binarios de 1s y 0s, números enteros e incluso cadenas de palabras (Truyol, 2015). A continuación describen de manera detallada los tipos de codificación que se pueden implementar en los individuos (posibles soluciones al problema).

- **Codificación Binaria:** Este tipo de codificación es una de las más populares. Esto se debe a que los primeros algoritmos genéticos comenzaron implementando este tipo de codificación (Truyol, 2015). Esta codificación es un arreglo o cadena de bits (parámetros) con valores de 1s y 0s (figura 2.5). Se le llama subsección de individuo que codifica el valor de un solo parámetro (1 o 0), de esta forma a este conjunto de parámetros en forma de cadena se le denomina individuo o individuo de una población. (Coello, 2015).

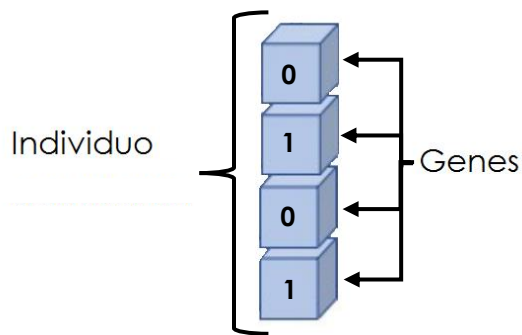


Figura 2.5 Codificación binaria.

- **Codificación numérica:** este tipo de codificación utiliza cadenas numéricas que en su mayoría se representan con una secuencia de números. Esta codificación se utiliza principalmente para el tipo de problemas en donde se desea ordenar algo y así la secuencia de número resulta muy útil (figura 2.6). Un ejemplo típico donde se implementa este tipo de codificación es el problema del agente viajero. En este caso tenemos una serie de ciudades por las que el agente debe pasar, se tiene que pasar por todas las posibles ciudades para saber la distancia más corta y así dar esa solución al problema. Ya que la variable principal es la distancia, se implementa la codificación numérica para enumerar cada una de las ciudades por las cuales pasa el viajero, así se le va dando prioridad al recorrido con menor distancia y de este modo no se repiten los traslados del agente (Truyol, 2015).

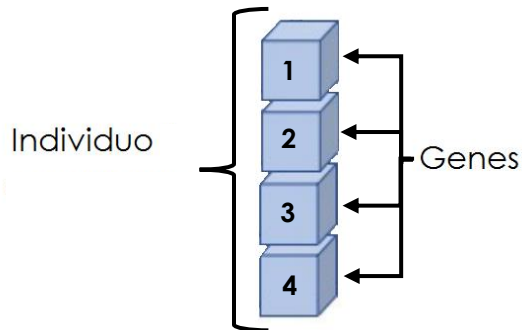


Figura 2.6 Codificación numérica.

2.1.3 Evaluación de aptitud

Una vez teniendo definida la codificación de los individuos, se debe evaluar cada individuo para saber la aptitud de cada uno. La aptitud determina qué tan buena o malo es una solución al problema, a esto se le llama función de aptitud. Este paso se lleva a cabo cada que se crea una nueva población de individuos, así se da paso a la selección de padres para seguir con el ciclo de la evolución del algoritmo genético (Gestal, 2010). El objetivo de la función de aptitud es buscar el objetivo deseado o alguna condición de paro en la evolución del algoritmo genético. La función de evaluación es parte importante en la evolución del algoritmo ya que está sincronizada con los demás operadores genéticos para que con base en un criterio obtenga la aptitud de cada individuo de cada población (Melanie, 1998).

2.2 Operadores genéticos

Dentro de un algoritmo genético se involucran 3 módulos importantes que se les conoce como operadores genéticos. (Melanie, 1998). Estos operadores son de suma importancia ya que para dar paso de una población a otra es necesario hacer uso de estos operadores. Dentro de operadores genéticos existen clases y subclases, las clases de estos operadores son:

- 1) Operadores de selección.
- 2) Operadores de cruza.
- 3) Operadores de mutación (Gestal, 2010).

2.2.1 Operadores de selección

La selección permite saber las probabilidades de cada individuo para ser cruzado con base en la teoría de Darwin (en la cual los individuos más aptos son los sobrevivientes y deben crear una descendencia) (Truyol, 2015). La función principal del operador de selección es elegir a los individuos que serán para crear una descendencia. Esto con la finalidad de imitar lo que ocurre en la naturaleza otorgando un mayor número de oportunidades a todos los individuos más aptos de toda la población (Gestal, 2010). Dentro de los operadores de selección existen varios tipos, por ejemplo:

- **Selección por torneo:** En este tipo de selección la idea principal es funcionar como un método que elige individuos haciendo comparaciones de tamaño K entre ellos y seleccionando al más fuerte. Existen dos tipos de selección por torneo: 1) torneo determinístico y 2) torneo probabilístico. La selección por torneo determinístico selecciona al azar un número de individuos (tamaño k) y entre estos individuos se selecciona al más apto para ser padre y crear descendencia. La selección por torneo probabilístico se genera un número aleatorio entre un determinado intervalo. (Gestal, 2010). Al tener los individuos ganadores se procede a generar una descendencia que reemplaza la población por los individuos menos aptos (Truyol, 2015).
- **Selección por elitismo:** Este tipo de selección se caracteriza por copiar al mejor de los individuos de la población actual en la siguiente población. Esto se debe a que al algún otro operador como la cruce o la mutación se puede perder el mejor individuo de la población. Con la selección por elitismo siempre se conserva al mejor de los individuos en la evolución del algoritmo genético (Truyol, 2015).
- **Selección por jerarquías:** Este tipo de selección es un método que no le da prioridad a la aptitud de los individuos, sino que se seleccionan con base en su rango, para ello en cada una de las poblaciones se debe ordenar de acuerdo a la aptitud que tiene cada individuo, del más apto al menos apto o viceversa. En cada población los individuos tienen un rango, los individuos con mayor rango tienen mayores posibilidades de ser padres y crear descendencia (Coello, 2015).
- **Selección "La Ruleta":** Este tipo de selección es uno de los que comúnmente se utiliza en un algoritmo genético. Es un método simple pero ineficiente ya que los individuos

son elegidos aleatoriamente, es decir, que los individuos menos aptos pueden ser seleccionados más de una vez para ser padres y crear descendencia (Coello, 2015). Este método cuenta con otras configuraciones dentro del algoritmo genético, esto se refiere a que la población puede estar ordenada jerárquicamente, de los individuos más aptos a los menos aptos, así que basta con generar un número aleatorio definiendo un intervalo como $[0..1]$ y así los más aptos tienen una mayor posibilidad de ser seleccionados para ser padres en más de una ocasión (Gestal, 2010).

2.2.2 Operadores de cruce

Es un operador que después de seleccionar y tener dos padres (individuo 1, individuo 2), indica los genes del individuo 1 y los genes del individuo 2 que pasarán a ser genes en los individuos de la nueva generación (Myers, 2005). El principal objetivo del operador de cruce es obtener una descendencia que comparta los genes de ambos padres (individuo 1 e individuo 2). La descendencia contaría probablemente con características buenas de dos individuos (padres), parte de esta nueva población debería tener una mejor adaptación que la población de padres anteriores (Gestal, 2010). A continuación se describen los diferentes tipos de cruce que se pueden implementar en un algoritmo genético. Se debe tener claro qué tipo de codificación se está utilizando, ya que dependiendo de esto se puede elegir el operador de cruce correcto.

- **Cruce con base en un 1 punto:** Teniendo una codificación binaria para los individuos de una población, en este tipo de cruce se tienen dos individuos (padre 1 y padre 2), se elige un punto de corte aleatoriamente entre los genes de ambos padres (figura 2.7). Posteriormente se copia la información genética de inicio al punto de corte de ambos padres. Esta información se inserta en los nuevos individuos (hijo 1 e hijo 2), genes de padre 1 insertados en hijo 1, genes del padre 2 insertado en el hijo 2. Posteriormente los genes restantes (del punto de corte al final del individuo) son copiados de los padres al otro hijo (genes del padre 1 insertados en el hijo 2 y genes del padre 2 insertados en el hijo 1) (Truyol, 2015).

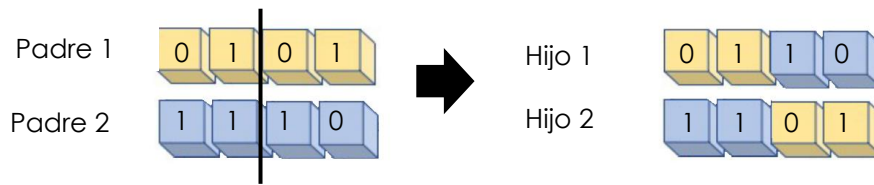


Figura 2.7 Ejemplo de cruce en 1 punto.

- Cruce con base en 2 puntos:** Teniendo una codificación binaria para los individuos de una población, en este tipo de cruce se toman los individuos candidatos (padre 1 y padre 2). Luego se eligen dos puntos de corte diferentes en un individuo padre (estos puntos serán iguales para ambos padres). Estos puntos no deben coincidir con ninguno de los extremos del individuo. De este modo, el corte se realiza de acuerdo a los dos puntos elegidos para ambos individuos (padre 1 y padre 2) (figura 2.8). Con los 3 segmentos que se tienen en ambos individuos se copia, de izquierda a derecha, el primer segmento en los hijos. Los genes del primer segmento del padre 1 insertados en el hijo 1, genes del primer segmento del padre 2 insertados en el hijo 2, genes del segundo segmento del padre 1 insertados en el hijo 2, genes del segundo segmento del padre 2 insertados en el hijo 1, genes del tercer segmento del padre 1 insertados en el hijo 1, y genes del tercer segmento del padre 2 insertados en el hijo 2 (Gestal, 2010).

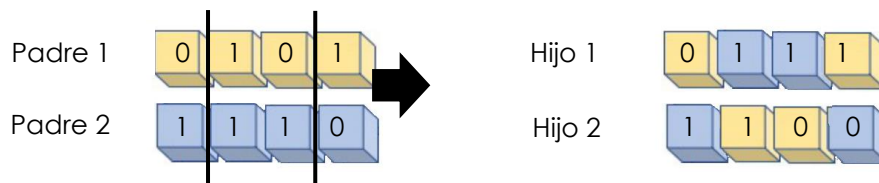


Figura 2.8 Ejemplo de cruce en 2 puntos.

- Cruce por emparejamiento parcial (PMX):** Teniendo una codificación numérica para los individuos de una población se seleccionan aleatoriamente 2 puntos de corte en los individuos padre (padre 1 y padre 2); creando 3 segmentos de genes en dichos individuos. El primer paso es tomar el segundo segmento de izquierda a derecha. Posteriormente, todos los genes que se encuentran dentro de este segmento serán intercambiados por genes que están fuera de dicho segmento sobre el mismo individuo

(padre 1 intercambia genes entre sí mismo y padre 1 intercambia genes entre sí mismo) (figura 2.9). Cabe señalar que el cambio de genes se realiza aleatoriamente. Una vez que todos los genes del segundo segmento del individuo han sido intercambiados, se copia toda la información genética del padre al hijo (genes del padre 1 insertados en el hijo 1 y genes del padre 2 insertados en el hijo 2) (Kumara Sastry, 2005).

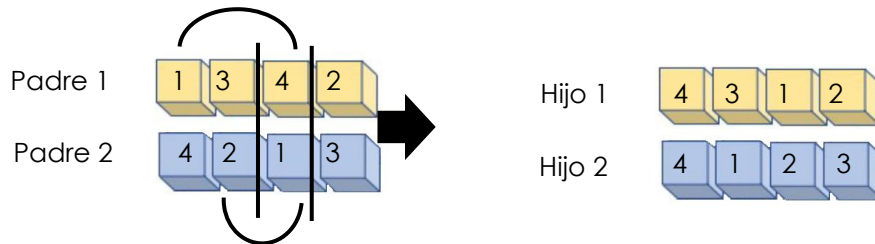


Figura 2.9 Ejemplo de cruce basado en PMX.

- Cruce con base en orden (OX):** Teniendo una codificación tipo numérica para los individuos de una población, en este tipo de cruce se seleccionan aleatoriamente 2 puntos de corte en los individuos padre (padre 1 y padre 2) creando 3 segmentos de genes en dichos individuos (Kumara Sastry, 2005). El primer paso es copiar el segundo segmento (de izquierda a derecha) de genes de los individuos padre a los individuos hijo (genes del padre 1 insertados en el hijo 1 y genes del padre 2 insertados en el hijo 2) (figura 2.10). Después se eliminan los genes del segundo segmento en los padres opuestos (tomar los genes del segundo segmento del padre 1 y eliminarlos en el padre 2 y el resto de genes del padre 2 copiarlos de izquierda a derecha en el hijo 1, así mismo tomar los genes del segundo segmento del padre 2 y eliminarlos en el padre 1 y el resto de genes del padre 1 copiarlos de izquierda a derecha en el hijo 2) (Coello, 2015).

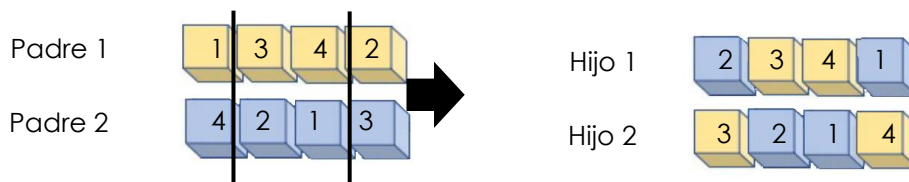


Figura 2.10 Ejemplo de cruce con base en orden.

2.2.3 Mutación

En la mutación, los individuos tienen un cambio aleatorio del valor dentro de algunos de sus genes. En algunas otras configuraciones de los algoritmos genéticos se pueden seleccionar los individuos directamente de la población, pero comúnmente la mutación es un operador que se utiliza en conjunto con el operador de cruce. Después de cruzar los individuos, se obtiene la descendencia que tendrá una cierta probabilidad de ser mutada antes de ser insertada en la nueva población (Gestal, 2010). Los operadores de mutación también dependen del tipo de codificación que los individuos tienen.

- **Mutación por intercambio:** Para una codificación numérica, la mutación por intercambio consiste en elegir 2 genes aleatoriamente y cambiarlos de posición entre ellos (figura 2.11) (Troyol, 2015). Se seleccionan los genes 2 y 4 e intercambian posiciones. Los genes restantes quedan con su valor y con su posición original (Vincent Cicirello, 2013).



Figura 2.11 Ejemplo de mutación por intercambio.

- **Mutación por inserción:** Para una codificación numérica, solo un gen del individuo es seleccionado aleatoriamente y es extraído del individuo. Posteriormente se elige un nuevo punto en el individuo (diferente al punto del gen seleccionado). A partir de la posición seleccionada los genes del individuo recorren una posición en dirección de donde se extrajo el gen seleccionado anteriormente. Ahora el gen extraído se inserta en esta nueva posición y a partir de esta posición se insertan o se recorren los demás genes hacia la posición donde fue removido el gen seleccionado. (figura 2.12) (Vincent Cicirello, 2013).

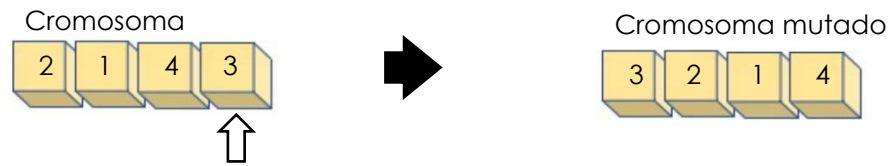


Figura 2.12 Ejemplo de mutación por inserción.

- **Mutación bit a bit:** Teniendo una codificación binaria (1s y 0s), la mutación bit a bit consiste en cambiar uno de los genes del individuo candidatos. El gen es elegido aleatoriamente y su valor simplemente se invierte (figura 2.13).

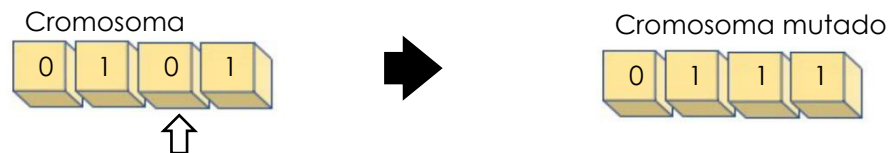


Figura 2.13 Ejemplo de cruce con base en orden.

2.3 Origen de las lenguas indígenas

México tiene con 69 lenguajes nacionales, 68 de estas lenguas son lenguas indígenas que nacen de 11 familias lingüísticas. La lengua restante para completar las 69, es el español.

Existen 11 familias lingüísticas para poder identificar el tipo de lengua indígena o el origen de estas, las 11 familias lingüísticas son:

Variantes lingüísticas de las agrupaciones de la familia álgica:

1. Kickapoo

Variantes lingüísticas de las agrupaciones de la familia yuto-nahua:

2. pápago
3. pima
4. tepehuano del norte
5. tepehuano del sur
6. tarahumara
7. guarijío
8. yaqui

- 9. mayo
- 10. cora
- 11. huichol

12. náhuatl

Variantes lingüísticas de las agrupaciones de la familia cochimí-yumana:

- 13. paipai
- 14. ku'ahl
- 15. cucapá
- 16. kumiai
- 17. kiliwa

Variantes lingüísticas de las agrupaciones de la familia seri:

- 18. seri

Variantes lingüísticas de las agrupaciones de la familia oto-mangue:

- 19. otomí
- 20. mazahua
- 21. matlatzinca
- 22. tlahuica
- 23. pame
- 24. chichimeco jonaz
- 25. chinanteco
- 26. tlapaneco
- 27. mazateco
- 28. ixcateco
- 29. chocholteco
- 30. popoloca
- 31. zapoteco
- 32. chatino
- 33. amuzgo
- 34. mixteco
- 35. cuicateco
- 36. triqui

Variantes lingüísticas de las agrupaciones de la familia maya:

- 37. huasteco
- 38. maya**
- 39. lacandón
- 40. ch'ol
- 41. chontal de Tabasco
- 42. tseltal

- 43. tsotsil
- 44. Q'anjob'al
- 45. Akateko
- 46. Jakalteko
- 47. qato'k
- 48. Chuj
- 49. tojolabal
- 50. Q'eqchi'
- 51. K'iche'
- 52. Kaqchikel
- 53. Teko
- 54. Mam
- 55. Awakateko
- 56. Ixil

Variantes lingüísticas de las agrupaciones de la familia totonaco-tepehua:

- 57. totonaco
- 58. tepehua

Variantes lingüísticas de las agrupaciones de la familia tarasca:

- 59. tarasco

Variantes lingüísticas de las agrupaciones de la familia mixe-zoque:

- 60. mixe
- 61. sayulteco
- 62. oluteco
- 63. texistepequeño
- 64. ayapaneco
- 65. popoluca de la Sierra
- 66. zoque

Variantes lingüísticas de las agrupaciones de la familia chontal de Oaxaca:

- 67. chontal de Oaxaca

Variantes lingüísticas de las agrupaciones de la familia huave:

- 68. huave

El lenguaje náhuatl se encuentra en la posición 12 y pertenece a la familia lingüística yuto-nahua.

El lenguaje maya se encuentra en la posición 38 y pertenece a la familia lingüística maya.

El lenguaje español pertenece a la familia de lenguas romances

2.3.1 Descripción de alfabetos

Los lenguajes español náhuatl y maya cuentan con alfabetos diferentes. Siguiendo la figura 2.15 se muestra que el alfabeto de la lengua en español está en forma de unigramas (un solo carácter), el alfabeto de la lengua náhuatl está en forma de unigramas y bigramas (dos caracteres), y el alfabeto de la lengua maya está en forma de unigramas, bigramas y trigramas (tres caracteres).

Alfabeto del lenguaje español																										
A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
Alfabeto del lenguaje náhuatl																										
A	E	I	O	P	T	C	QU	CU	TZ	CH	TL	Z	X	L	HU	M	N									
Alfabeto del lenguaje maya																										
A	B	CH	CH'	E	I	J	K	K'	L	M	N	O	P	P'	S	T	T'	TS	TS'	U	W	X	Y	'		

Figura 2.14 Alfabetos de las lenguas español, náhuatl y maya.

2.4 Resumen del capítulo

En este capítulo, se describió de manera detallada la información sobre los algoritmos genéticos implementados como métodos de optimización, se describe cuáles son los pasos y el ciclo de evolución por el cual pasan estos algoritmos. También se hizo mención sobre los diferentes operadores genéticos que pueden utilizarse para implementar un algoritmo genético y cuáles de estos operadores son los mejores de acuerdo al tipo de problema que se desea resolver.



3. Estado del arte

Este capítulo se centra en los trabajos que se han realizado para optimizar distribuciones de teclado. Se da a conocer la configuración de cada uno de los trabajos de optimización, qué criterios se toman en cuenta para optimizar una o más distribuciones de teclado. Así mismo, se muestra una tabla comparativa sobre las optimizaciones de teclado que están basadas específicamente en algoritmos genéticos. De igual forma se describe la configuración que se utiliza en cada algoritmo genético y los criterios utilizados para optimización de teclado.

3.1 Optimización de teclados y sus métodos

Tiempo después de que aparece el teclado *qwerty*, algunos investigadores presentan sus análisis demostrando que esta distribución de teclado que tiene dificultades para escribir incluso para el mismo lenguaje inglés (para la cual fue creada). Para tratar de resolver estos problemas con la distribución de teclado *qwerty* se realizan trabajos en los cuales presentando nuevas distribuciones de teclado o distribuciones modificadas con base en teclado *qwerty*. Algunos de estos trabajos tratan de mejorar el teclado *qwerty*, algunos otros tratan de independizarse del lenguaje creando un nuevo teclado para un lenguaje diferente y en específico. Por último otros investigadores crean diferentes teclados para un mismo lenguaje tratando de que estos teclados sean para uno o varios dominios en específico.

Los teclados optimizados surgen poco después de la presentación del teclado *dvorak* en el que se demuestra que éste es un teclado mucho más óptimo en comparación con el teclado

qwerty para a escritura de textos en inglés (Hempstalk, 2006). Sin embargo, para optimizar este teclado (*dvorak*), el creador August solo se basa en la frecuencia de bigramas (dos caracteres) y frecuencia de unigramas (un solo caracter) para llegar a la optimización deseada. Como tal August no implementa ningún método de optimización o algoritmo que tenga que ver con inteligencia artificial.

Años después, en 2002, Eggers (Eggers, 2002) presenta 3 distribuciones de teclado para diferentes lenguaje. Para realizar la optimización de estas distribuciones de teclado Eggers implementa un método de optimización llamado "colonia de hormigas". El objetivo de Eggers es optimizar un teclado para el lenguaje Alemán (Figura 3.1), un teclado para el lenguaje inglés totalmente diferente al teclado *qwerty* y *dvorak* (Figura 3.2) y un teclado para el lenguaje francés (figura 3.3). También se toma en cuenta la frecuencia de caracteres de acuerdo a un corpus de cada lenguaje, para el lenguaje alemán se utilizaron textos de la revista "Der Spiegel", para el lenguaje inglés se utilizan textos del periódico "USA Today" y para el lenguaje francés se utilizan textos del periódico "Le Monde". De este modo, con un corpus en específico para cada lenguaje se determina la frecuencia de caracteres considerando la misma técnica de digitación del teclado *qwerty* donde se usan los 5 dedos de ambas manos. Sin embargo, se puede ver que en los teclados de Eggers se tiene una forma diferente a la forma del teclado *qwerty*, lo que causa problemas al utilizar estos teclados (propuestos por Eggers) en una computadora, pero la única forma de hace uso de estos es fabricándolos especialmente con esa forma.

	\	1	2	3	4	5		0		6	7	8	9	0	.	/		
	:	}	L	D	B	'		1	Ret	;	<	"	(?	~			
	Y	K	W	C	Q	J	lSh	2	rSh	*	-	&		!	+			
	R	G	N	S	T	P	lSp	3	rSp	A	O	E	I	H	U			
	r	g	n	s	t	p	lSp	4	rAGr	{	>	#	^	%)			
	X	Z	=	M	V	F	lAGr	5			,							
	X	Z	=	m	v	f					,							

Figura 3.1 Teclado optimizado por Eggers para el lenguaje inglés (Eggers, 2002).

	\	1	2 ₂	3 ₃	4	5		0		6	7	8	9	0	.	/		
	Ö ö	}	Ü ü	O o	Y y	:		1	Ret	* +	~ ~	M m	W w	B b	=	:		
	!	> <		"	&	(Ä ä	lSh	2	rSh	F f	V v	{	K k	§	#		
	G g	H h	A a	E e	I i	U u	lSp	3	rSp	C c	T t	S s	D d	N n	R r			
	%]	[)	_	?	lAGr	4	rAGr	Q q	J j	L l	X x	Z z	P p			
					.			5			,							

Figura 3.2 Teclado optimizado por Eggers para el lenguaje alemán (Eggers, 2002).

	\	1	2 ₂	3 ₃	4	5		0		6	7	8	9	0	.	/		
	+ w	G g	J j	V v	Ç ç			1	Ret	È è	(À à		*	{	<		
	Z z	B b	F f	M m	L l	Q q	lSh	2	rSh	?	-	%	É é] &	È è			
	P p	N n	S s	R r	T t	D d	lSp	3	rSp	O o	A a	E e	I i	U u	.			
	Y y	K k	X x	C c	H h	U ü	lAGr	4	rAGr	µ #	:	!>	~	=	^	!		
					.			5			,							

Figura 3.3 Teclado optimizado por Eggers para el lenguaje francés (Eggers, 2002).

Un año después en 2003, Deb (Deb, 2003) presenta un teclado optimizado para el lenguaje Hindi, en este trabajo Deb se basa en un algoritmo genético para obtener un teclado optimizado basado en la frecuencia de caracteres y la distancia que existe entre teclas. Para determinar la frecuencia de caracteres en el lenguaje hindi se utilizan diferentes textos tomados de diversas fuentes de internet. Deb de igual forma contempla una técnica de digitación utilizando los 5 dedos de ambas manos, y de acuerdo con la figura 3.4, se cambia la forma física del teclado ya que la disposición de sus teclas es diferentes a la distribución del teclado *qwerty*, esto también causa problemas al utilizar este teclado en una computadora. A diferencia del teclado *qwerty*, Deb agrupa de forma más específica cada columna de teclas correspondientes a cada dedo, y en vez de tener una barra espaciadora en la parte inferior, ésta se encuentra en la parte media del teclado (figura 3.4).

LEFT HAND								RIGHT HAND								
	Little		Ring	Mid	Index		Thumb	Thumb	Index		Mid	Ring	Little			
0	ऑ	-	ी	-	।	ए	ओ	-	-	-	-	-	न	-		
	ऑ	-	-	-	-	-	-	-	-	-	-	-	ज	ड		
1	-	ठ	रा	INV	-	ई	ये	-	-	-	ळ	ण	ळ	ड		
	-	ु	ी	ै	ी	उ	इ	-	-	ड	द	श	प	ष		
2	-	-	ख	ग	र	ं	ओ	-	-	श्र	भ	ल	ं	ॉ		
	-	-	य	ै	ा	ि	्	-	SPC	त	न	र	क	व		
3	-	-	घ	ं	ै	यं	ऊ	-	-	ै	रे	झ	ह	फ		
	-	SHF	आ	अ	ी	ू	्र	-	-	थ	स	ल	ज	ट		
														ऑ		
														च		
														RET		
														SHF		
	7	6	5	4	3	2	1	0	0	1	2	3	4	5	6	7

Figura 3.4 Teclado optimizado por Deb para el lenguaje hindi (Deb, 2003).

En ese mismo año (2003) Walker (Walker, 2003) realiza un trabajo implementando un algoritmo genético para optimizar un teclado para el lenguaje inglés totalmente diferente a la distribución de teclado *qwerty*. Este teclado respeta la forma del teclado *qwerty* y por tanto, se adoptó la técnica de digitación para los 5 dedos de ambas manos. Por otra parte, para llegar a la optimización de este teclado, Walker se basa en la distancia entre teclas y frecuencia de caracteres. Para obtener esta frecuencia Walker toma las 1000 palabras más frecuentes del lenguaje inglés y con base en esta lista de palabras obtiene la optimización de su teclado (figura 3.5)

		Q	*	H	J	*	N	L	K	M	D	*	*				
		I	U	A	O	E	P	W	C	R	B	*					
		X	Y	*	*	S	T	G	F	Z	V						

Figura 3.5 Teclado optimizado por Walker para el lenguaje inglés (Walker, 2003).

En el año 2005, Capewell (Capewell, Michael, 2005) realiza un trabajo en que logra optimizar 6 distribuciones de teclado respetando el diseño original del teclado *qwerty* y así mismo retoma la misma técnica de digitación contemplando los 5 dedos de ambas manos. El objetivo principal en los teclados de Capewell es reducir el movimiento de los dedos basándose en la distancia que existen entre las teclas. Para poder llegar al objetivo deseado Capewell utiliza teclados existentes como *qwerty* e implementa un algoritmo genético como método de optimización para modificar estas distribuciones de teclado, logrando así optimizar dichas distribuciones. La figura 3.6 muestra una distribución optimizada del teclado *dvorak* para el lenguaje inglés, después de realizar cambios mediante el algoritmo genético Capewell agrega su apellido a este teclado modificado, por tanto le llama *capewell - dvorak*. La figura 3.7 muestra una distribución optimizada del teclado *qwerty* para el lenguaje inglés, después de realizar cambios mediante el algoritmo genético Capewell agrega su apellido a este teclado modificado, por tanto le llama *capewell - qwerty*. La figura 3.8 muestra una distribución optimizada del teclado *adjacent* para el lenguaje inglés, después de realizar cambios mediante el algoritmo genético Capewell agrega su apellido a este teclado modificado, por tanto le llama *capewell - adjacent*. La figura 3.9 muestra una distribución optimizada del teclado *close* para el lenguaje inglés, después de realizar cambios mediante el algoritmo genético Capewell agregó su apellido a este teclado modificado, por tanto le llama *capewell - close*. Finalmente Capewell presenta 2 teclados más. La figura 3.10 muestra una distribución optimizada del teclado *klauster*, después de realizar cambios mediante el algoritmo genético Capewell lo presenta simplemente como *klauster2*. La figura 3.11 muestra una distribución optimizada del teclado *arensito*, después de realizar cambios mediante el algoritmo genético Capewell lo presenta simplemente con el nombre original del teclado *arensito*.

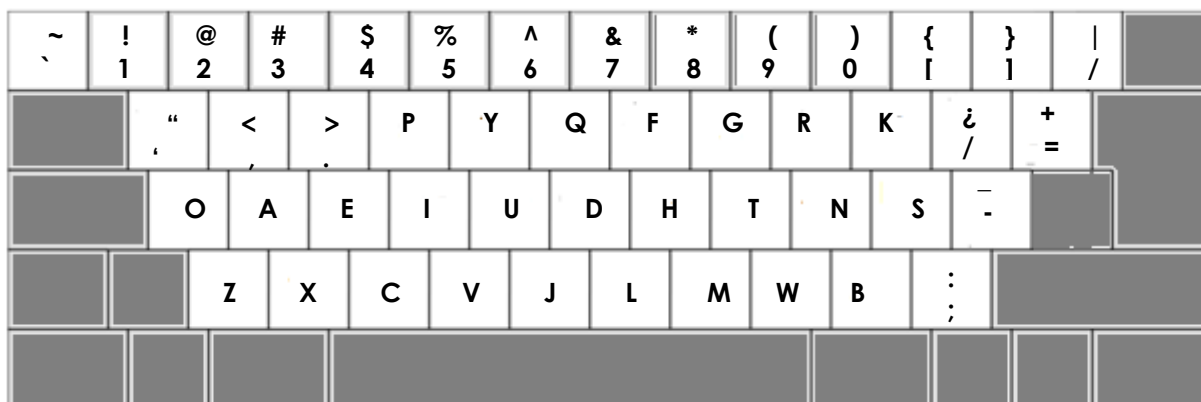


Figura 3.6 Teclado *capewell - dvorak* (Capewell, Michael, 2005).



Figura 3.7 Teclado *capewell - qwerty* (Capewell, Michael, 2005).

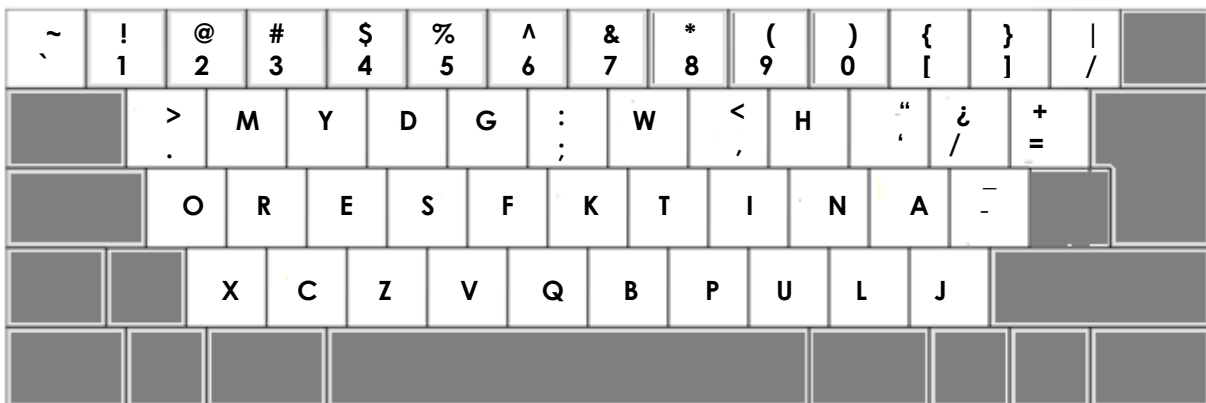


Figura 3.8 Teclado *capewell-adjacent* (Capewell, Michael, 2005).

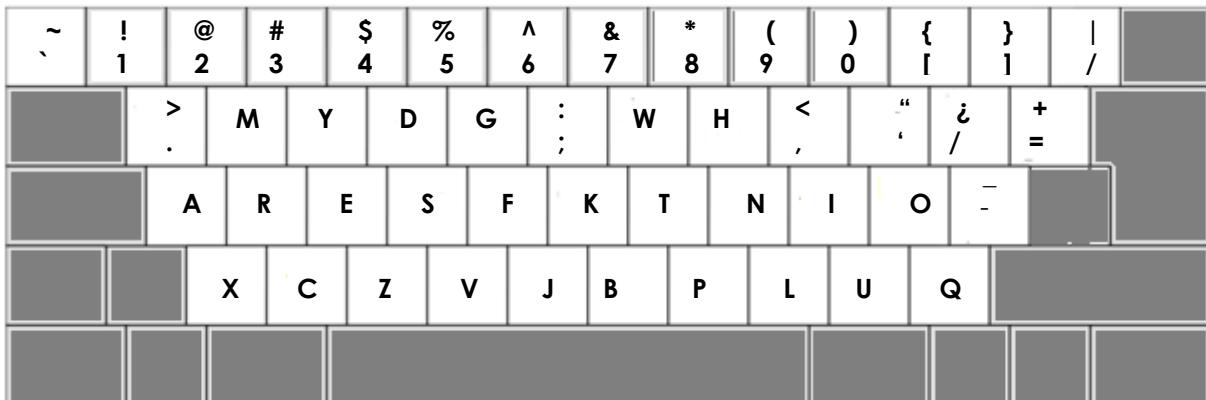


Figura 3.9 Teclado *capewell - close* (Capewell, Michael, 2005).

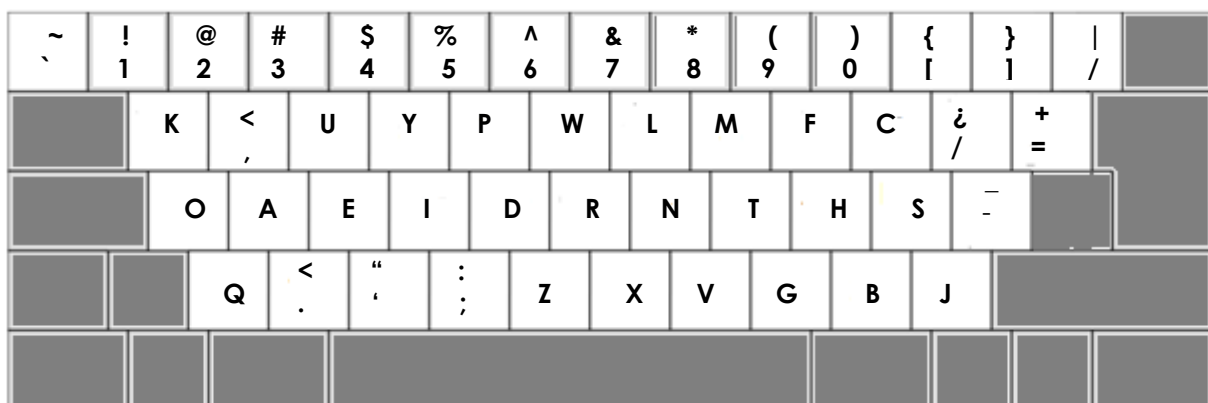


Figura 3.10 Teclado *klauster2* (Capewell, Michael, 2005).

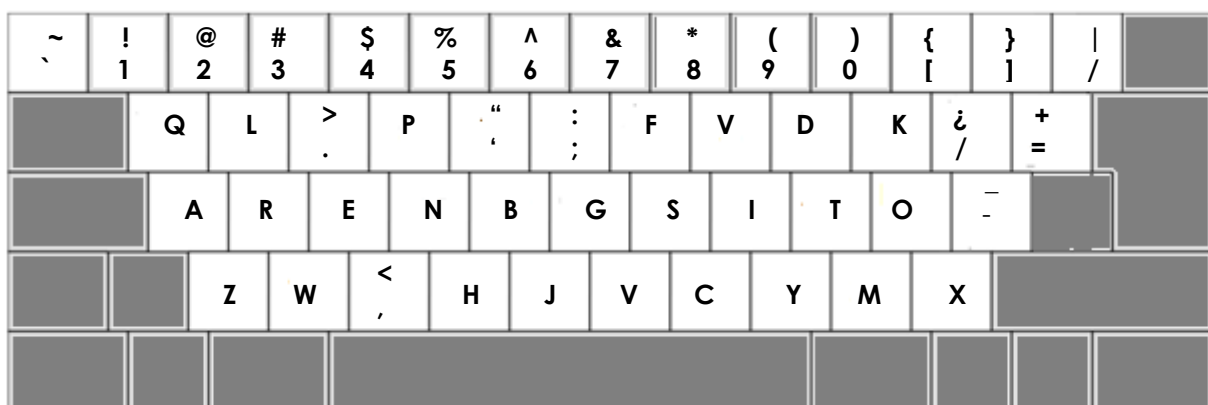


Figura 3.11 Teclado *arensito* (Capewell, Michael, 2005).

En 2008, el investigador Brewbaker (Brewbaker, 2008) programa un algoritmo genético para optimizar 4 distribuciones de teclado para el lenguaje inglés. Dos de estas distribuciones respetan la forma que tiene el teclado *qwerty* y así mismo retoma la técnica de digitación de los 5 dedos de ambas manos. Las primeras 2 distribuciones de teclado, a pesar de que son creadas para el lenguaje inglés tienen un enfoque diferente ya que cada una es optimizada para dos dominios diferentes del lenguaje inglés. La primera distribución de teclado se optimiza especialmente para todos los programadores que utilizaban el lenguaje C (figura 3.12). La segunda distribución de teclado se optimiza para el proyecto GUTENBERG (proyecto para la creación de libros electrónicos) (figura 3.13). Para las otras dos distribuciones de teclado Brewbaker se basa en la forma del teclado *fitaly* (Figura 3.14) para optimizar nuevamente dos distribuciones de teclado con el mismo enfoque que las anteriores. Una distribución de teclado

se optimiza especialmente para todos los programadores que utilizan el lenguaje C (figura 3.15), ya otra distribución de teclado se optimiza para el proyecto GUTENBERG (figura 3.16).

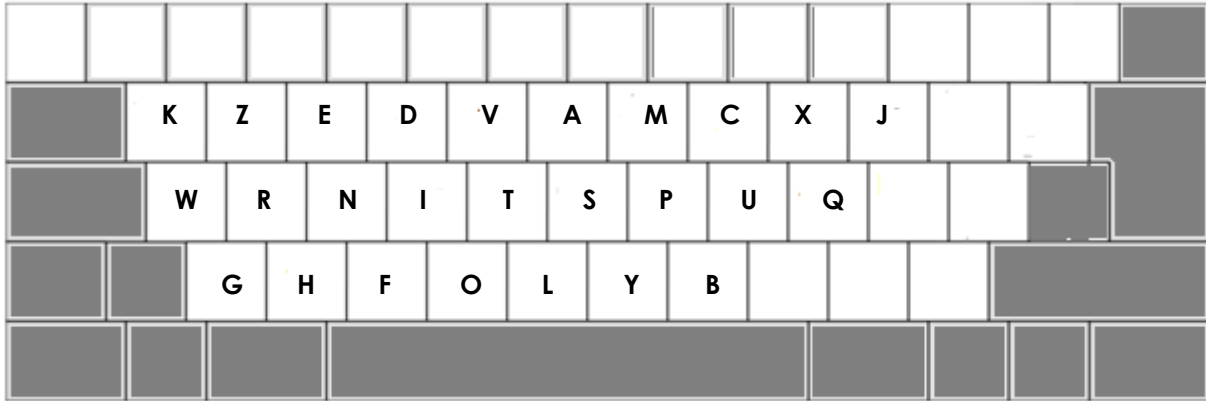


Figura 3.12 Teclado *qwerty* optimizado para la programación en el lenguaje C (Brewbaker, 2008).



Figura 3.13 Teclado *qwerty* optimizado para el proyecto GUTENBERG (Brewbaker, 2008).

<i>z</i>	<i>v</i>	<i>c</i>	<i>h</i>	<i>w</i>	<i>k</i>
<i>f</i>	<i>i</i>	<i>t</i>	<i>a</i>	<i>l</i>	<i>y</i>
		<i>n</i>	<i>e</i>		
<i>g</i>	<i>d</i>	<i>o</i>	<i>r</i>	<i>s</i>	<i>b</i>
<i>q</i>	<i>i</i>	<i>u</i>	<i>m</i>	<i>p</i>	<i>x</i>

Figura 3.14 Teclado *fitaly* con diseño diferente al *qwerty* (Brewbaker, 2008).

<i>k</i>	<i>v</i>	<i>s</i>	<i>g</i>	<i>h</i>	<i>w</i>
<i>x</i>	<i>c</i>	<i>t</i>	<i>n</i>	<i>i</i>	<i>f</i>
		<i>a</i>	<i>e</i>		
<i>q</i>	<i>u</i>	<i>l</i>	<i>d</i>	<i>r</i>	<i>z</i>
<i>j</i>	<i>b</i>	<i>m</i>	<i>o</i>	<i>p</i>	<i>y</i>

Figura 3.15 Teclado *fitaly* optimizado para la programación en el lenguaje C (Brewbaker, 2008).

<i>q</i>	<i>u</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>c</i>
<i>j</i>	<i>p</i>	<i>o</i>	<i>s</i>	<i>t</i>	<i>x</i>
		<i>r</i>	<i>a</i>		
<i>y</i>	<i>m</i>	<i>e</i>	<i>l</i>	<i>i</i>	<i>w</i>
<i>b</i>	<i>z</i>	<i>v</i>	<i>d</i>	<i>n</i>	<i>k</i>

Figura 3.16 Teclado *fitaly* optimizado para el proyecto GUTENBERG (Brewbaker, 2008).

Después de hacer varias evaluaciones al teclado *qwerty*, en el año 2011, Valle (Valle, 2011) presenta una distribución de teclado con base en el teclado *qwerty*. Valle analiza y optimiza diferentes distribuciones de teclado en las que se va mejorando las dificultades del teclado *qwerty* con el objetivo de reducir el esfuerzo de escritura. Principalmente, Valle tiene como objetivo llegar a una distribución de teclado optimizada para escribir especialmente en inglés, francés y español; pero después de realizar varias pruebas también se demuestra que es una distribución de teclado óptima para lenguajes de programación como C y Python. Valle utiliza recocido simulado como método de optimización y respetando el diseño original del teclado *qwerty*. Esta distribución nueva de teclado se presenta con el nombre de teclado *carpalx*, en la figura 3.17 se observa que es un teclado que se puede utilizar en una computadora solo cambiando la posición de las teclas mediante un software, y no hay la necesidad de fabricar este teclado.

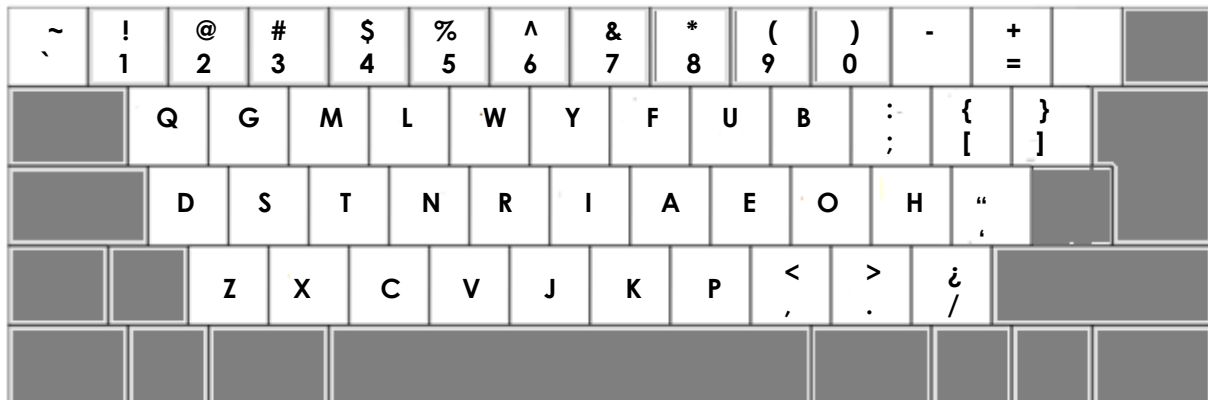


Figura 3.17 Teclado *carpalx* (Valle, 2011).

En el trabajo de investigación realizado por Samimi (Samimi, 2012) se implementa una combinación de métodos de optimización: recocido simulado y un algoritmo genético para crear una distribución para el lenguaje persa respetando la forma del teclado *qwerty*, y por tanto, se utiliza la misma técnica de digitación de los 5 dedos de ambas manos. Para lograr esta distribución de teclado optimizada, Samimi se basa en la frecuencia de caracteres creando una lista con las 19,092 palabras más frecuentes del lenguaje persa. En la figura 3.18 se muestra el teclado optimizado, y es una distribución que cuenta físicamente con las mismas características del teclado *qwerty*, la ventaja de retoma la forma original del teclado *qwerty* es que para hacer uso de este teclado (propuesto por Samimi) en una computadora simplemente se debe re-etiquetar o re-ordenar las teclas sin necesidad de fabricar un teclado especial.



Figura 3.18 Teclado optimizado para el lenguaje persa (Samimi, 2012).

En el año 2015, Hosny optimiza su propia distribución de teclado implementando un algoritmo genético para crear a una distribución de teclado para el lenguaje árabe (figura 3.19). Al igual que algunos de los trabajos de optimización de teclado mencionados anteriormente, Hosny también se basa en la frecuencia de caracteres creando un corpus con artículos de wikipedia y textos de periódico en lenguaje árabe. Lo que hace diferente a la optimización de Hosny es que solo considera un dedo para la escritura sobre su distribución de teclado ya que es un teclado que se propone ser utilizado en dispositivos táctiles donde en su mayoría de veces se utiliza un solo dedo para escribir o se hace uso de un lápiz óptico.

ذ	إ	أ	هـ	ن	م	ع	د	ق	ث	ض
ى	ص	ح	س	ل	ا	ي	ب	ة	ط	غ
ظ	خ	ش	ك	ت	و	ر	ف	ج	ز	ئ

Figura 3.19 Teclado optimizado por Hosny para el lenguaje árabe (Hosny, 2015).

Los trabajos descritos anteriormente son trabajos realizados para optimizar distribuciones de teclado, creando nuevas distribuciones de teclado que cuenten con una forma totalmente diferente a la forma del teclado estándar *qwerty* o considerando la distribución del teclado *qwerty* para hacer mejoras en una o varias distribuciones nuevas de teclado. Para lograr cumplir estos objetivos se implementan diferentes métodos de optimización como: algoritmos genéticos; colonia de hormigas y recocido simulado. Dentro de estas implementaciones se

tienen diferentes configuraciones para cada método, ya que los investigadores realizan las optimizaciones de los teclados con diferentes objetivos, algunos de estos con la finalidad de mejorar la distribución de teclado *qwerty*, crear un nuevo teclado para uno o más lenguajes, y así mismo algunos están enfocados a teclados de computadora y otros a teclados para dispositivos móviles. En la siguiente tabla 3.1 se describe de manera general los trabajos realizados para la optimización de distribuciones o creación de teclados, el método que se implementa para optimizar uno o varios teclados, el lenguaje o los lenguajes para los cuales son creados y así mismo el uso de los dedos de las manos para poder hacer uso de estos teclados.

Tabla 3.1 Optimización de distribuciones de teclado implementando distintos métodos.

TÍTULO DE LA INVESTIGACIÓN	MÉTODO DE OPTIMIZACIÓN	LENGUAJE	TÉCNICA DE DIGITACIÓN
<i>An Ant Colony Optimization Algorithm for the Optimization of the Keyboard Arrangement Problem</i> (Eggers, 2002)	Colonia de hormigas	Francés, inglés y alemán	Uso de los 5 dedos de ambas manos
<i>Evolving a More Optimal Keyboard</i> (Walker, 2003)	Algoritmo Genético	Inglés	Uso de los 5 dedos de ambas manos
<i>Design of an Optimal Hindi Keyboard for Convenient and Efficient Use</i> (Deb, 2003)	Algoritmo Genético	Hindi	Uso de los 5 dedos de ambas manos
<i>Alternative keyboard layouts</i> (Capewell, Michael, 2005)	Algoritmo Genético	Inglés	Uso de los 5 dedos de ambas manos
<i>Optimizing stylus keyboard layouts with a genetic algorithm: customization and internationalization</i> (Brewbaker, 2008)	Algoritmo Genético	Inglés	Uso de los 5 dedos de ambas manos
<i>Scientific hands on comfort</i> (Valle, 2011)	Recocido Simulado	Inglés, español,	Uso de los 5 dedos de ambas manos

		Francés y lenguajes de programación (C y Python)	
<i>Optimization of Farsi Letter Arrangement on Keyboard by Simulated Annealing and Genetic Algorithms (Samimi, 2012)</i>	Recocido Simulado y Algoritmo Genético	Persa	Uso de los 5 dedos de ambas manos
<i>A Genetic Algorithm Approach for Optimizing a Single-Finger Arabic Keyboard Layout (Hosny, 2015)</i>	Algoritmo Genético	Árabe	Uso de un solo dedo o un puntero

3.2 Algoritmos genéticos aplicados para la optimización de distribuciones de teclados

A lo largo de este capítulo se mencionan los diferentes métodos (colonia de hormigas, recocido simulado y algoritmos genéricos) que se implementan para crear o modificar distribuciones de teclado con el objetivo de llegar a una o varias distribuciones de teclado mejorando las anteriores. La mayoría de los investigadores buscan superar las dificultades del teclado *qwerty* ya que es un teclado que se ha convertido en un estándar como el teclado de computadora que hoy en día se utiliza. Cada método de optimización tiene sus propias configuraciones y variables, lo que los hace diferentes entre ellos principalmente es el lenguaje para el cual se optimiza, el corpus de entrada para determinar la frecuencia de caracteres y la técnica de digitación (número de dedos que se utilizan al digitar).

La tabla 3.2 tiene específicamente una lista de las optimizaciones de teclado en las que se implementan algoritmos genéticos y se describe de manera detallada la configuración de cada algoritmo genético. Esta configuración consiste en la combinación de operadores genéticos (selección, cruce y mutación) que se utilizaron en cada algoritmo, así mismo los criterios (distancia entre teclas, frecuencia de caracteres, número de los dedos para digitar,

balance en ambas manos, etc.) que se toman en cuenta para llegar a la distribución de teclado deseada.

La tabla 3.2 muestra el uso de los algoritmos genéticos con distintas configuraciones de tipos de codificaciones, operadores de selección, cruza y mutación; así mismo los criterios de evaluación utilizados en la función de aptitud.

En uno de los primeros trabajos de optimización de teclado presentado por Walker (Walker, 2003), se configura con los operadores: 1) operador de selección por muestreo estocástico y 2) operador de mutación por intercambio. Walker utiliza este mismo operador como operador de cruza ya que no considera importante tener un operador de cruza. Walker menciona que dentro de la mutación pueden presentarse cambios importantes dentro de la población de individuos y son cambios suficientes para que esta población pueda evolucionar. Para la función de evaluación o función de aptitud se evalúan diferentes criterios como: 1) la carga que tenía cada dedo al digitar un conjunto de caracteres, esto para evitar que los caracteres más frecuentes queden para un solo dedo y que este tengan sobre carga; 2) uso de la misma mano, un criterio que va a proporcionar un balance en ambas manos, lo que quiere decir que los caracteres más frecuentes no van a estar de un solo lado del teclado sino que deberían estar distribuidos con un balance aproximado del 50% del lado izquierdo y el otro 50% del lado derecho; para evitar sobre carga en alguna de las manos de los mecanógrafos; 3) así mismo un criterio para saber qué tanta carga tiene el dedo meñique ya que es uno de los dedos con menos fuerza y menos habilidad de movimiento. La codificación para los individuos de las poblaciones no se menciona. Walker afirma que en su distribución de teclado evaluada con los criterios anteriores se puede escribir con mayor velocidad en comparación con la distribución *qwerty*.

En el siguiente trabajo de optimización de distribuciones de teclado presentado por Deb (Deb, 2003), utiliza una codificación numérica para los individuos. Los operadores genéticos que usa en su algoritmo son: 1) operador de selección (2 operadores) selección por ruleta y selección por torneo, Deb utiliza estos dos operadores de selección ya que después de hacer varias pruebas se menciona que tiene mejores resultados al utilizar 2 operadores en la selección; 2) operador de cruza con base orden (OX); 3) operador de mutación por intercambio. Para la función de aptitud se evalúan diferentes criterios como: 1) alternancia entre ambas manos

para evitar que alguna de las dos manos tenga mayor carga y esto ocasione fatiga al mecanógrafo; 2) el uso de un mismo dedo consecutivamente, esto para determinar de manera particular la carga sobre cada uno de los dedos, darle mayor prioridad a los dedos índice (dedos con mayor fuerza y habilidad) y así realizar un balance y distribuir toda la carga a lo largo y ancho de todo el teclado. Con estos criterios Deb disminuye el esfuerzo al escribir y reducir la carga de trabajo para los dedos meñique, anular y medio, dándole una mayor carga a los dedos índice.

En el trabajo de optimización presentado por Brewbaker (Brewbaker, 2008), se implementa un algoritmo genético y la configuración utilizada es: una codificación numérica para los individuos. Los operadores genéticos que usa el algoritmo son: 1) operador de selección por ruleta; 2) operador de cruza en dos puntos; 3) operador de mutación por intercambio. Para la función de evaluación o función de aptitud se evalúan diferentes criterios como: 1) tiempo en cada pulsación (tiempo promedio que tarda un mecanógrafo en digitar un tecla y pasar a la siguiente), 2) distancia de una tecla a otra, con el objetivo de reducir el movimiento de los dedos mediante el desplazamiento, 3) tamaño de teclas, este criterio está ligado al criterio anterior ya que depende del tamaño del teclado es la distancia entre cada tecla. Así mismo Brewbaker implementa su algoritmo genético para optimizar una distribución de teclado diferente a la que él propone. Una vez definido el tamaño del alfabeto se puede optimizar cualquier distribución de teclado para cualquier lenguaje y cualquier dominio, los criterios en la función de evaluación son los mismos pero se pueden realizar diferentes optimizaciones simplemente con tener el corpus del lenguaje deseado a optimizar.

En uno de los últimos trabajos de optimización de teclado realizado por Hosny (Hosny, 2015), también usa un algoritmo genético y la configuración es una codificación distinta a las mencionadas anteriormente ya que en este caso se codifica el individuo de forma binaria. Los operadores genéticos que usa en su algoritmo son: 1) operador de selección por ruleta; 2) operadores de cruza dos puntos de corte (TPX) y cruce uniforme en un punto. Hosny usa estos dos operadores de cruza para comparar los resultados teniendo uno y dos puntos de corte; 3) operadores de mutación: por intercambio, por inserción y por recocido simulado, utilizando estos 3 operadores de cruza para hacer diferentes configuraciones a su algoritmo genético y en cada una de las combinaciones posibles con los diferentes operadores de cruza y mutación va mejorando los resultados de optimización. Para la función de función de aptitud se evalúan

diferentes criterios como: 1) las frecuencia de pares de letras (bigramas) y sus distancias en teclado de tecla a tecla; 2) la frecuencia de un solo caracter y en que fila se encuentra; 3) y así mismo saber en qué parte del teclado se tiene una mayor carga de trabajo. Sabiendo los valores de estos criterios. Hosny menciona que su objetivo es mejorar la velocidad y precisión con la que se digita un conjunto de caracteres utilizando un solo dedo o un lápiz óptico en los dispositivos móviles.

Tabla 3.2 Optimización de distribuciones teclados implementando algoritmos genéticos.

TÍTULO DE LA OPTIMIZACIÓN	OPERADORES GENÉTICOS			FUNCIÓN DE APTITUD
	SELECCIÓN	CRUZA	MUTACIÓN	
<i>Evolving a More Optimal Keyboard</i> (Walker, 2003)	Muestreo estocástico universal	Ningún operador de cruza	intercambio	Uso de una sola misma mano (sobre carga). Balance en ambas manos. Uso de una tecla en la fila origen. Uso del dedo meñique en la fila inferior o superior. Uso del dedo índice. Uso del dedo medio. Uso del dedo anular. Uso del dedo meñique.
<i>Design of an Optimal Hindi Keyboard for Convenient and Efficient Use</i> (Deb, 2003)	Ruleta Torneo	Cruce en base orden (OX)	intercambio	Alternancia en manos. Uso de un mismo dedo de forma consecutiva (carga de trabajo).
<i>Optimizing stylus keyboard layouts with a genetic algorithm: customization and internationalization</i> (Brewbaker, 2008)	Ruleta	Cruce en dos puntos de corte) (TPX)	Intercambio	El valor promedio de tiempo por pulsación de tecla. Probabilidad de pasar de una tecla a otra (teclas adyacentes). Distancias entre teclas.

				Tamaños de teclas (dependiendo de tamaño del teclado). Tamaño del alfabeto (número de caracteres).
<i>A Genetic Algorithm Approach for Optimizing a Single-Finger Arabic Keyboard Layout</i> (Hosny, 2015)	Ruleta	Cruce en dos puntos de corte (TPX). Cruce en un punto.	Intercambio Inserción Recocido simulado	Las frecuencias de pares de letras (bigramas) y la distancia de estas en el teclado (tecla a tecla). La frecuencia de una sola letra (unigrama) y la fila a la que está asignada. Parte del teclado donde se tiene mayor carga de trabajo.

3.3 Resumen del capítulo

En este capítulo, se hace mención sobre los diferentes trabajos realizados para la optimización de nuevas distribuciones de teclado, se describe de manera detallada el método de optimización utilizado en cada trabajo y los criterios que se tomaron en cuenta para poder optimizar una o varias distribuciones de teclado.



4. Método propuesto

Recordando el problema principal de esta tesis, el cual consiste en facilitar distribuciones de teclado con criterios ergonómicos. En este capítulo se describe de manera detallada el método que se propone implementar para realizar las distribuciones de teclado para cada lengua y a su vez que cuenten con criterios ergonómicos. También se describe cual es la configuración que se utiliza en el método propuesto y la metodología que se sigue para la optimizaciones de las distribuciones de teclado.

4.1 Método propuesto

En el capítulo anterior se habla de los trabajos que se han realizado para optimizar distribuciones de teclado, se hace mención de los diferentes métodos de optimización que se implementan. También se describe de manera detallada una lista de los trabajos que implementan algoritmos genéticos como método de optimización.

En esta tesis se propone implementar un algoritmo genético como método para optimizar los criterios de distribuciones de teclado y obtener 4 distribuciones con criterios ergonómicos: 1) una distribución de teclado para la lengua indígena náhuatl; 2) una distribución de teclado para la lengua indígena maya; 3) una distribución de teclado para las lenguas español – náhuatl; 4) y una distribución de teclado para las lenguas español - maya. Para llegar a las 4 distribuciones de teclado deseadas, se realizan los siguientes pasos:

El primer paso consiste en elegir el o los lenguajes para optimizar, por ejemplo: lenguaje maya; se usa el corpus (conjunto de palabras) de este lenguaje.

El segundo paso consiste en hacer uso del algoritmo genético siguiendo los pasos:

1. Definir la codificación de los individuos de las poblaciones. Para implementar el algoritmo genético se hizo uso de una codificación numérica con valores del número 1 al 37, por tanto cada individuo tiene 37 genes que comprenden a los caracteres del teclado. Por medio de esta codificación se puede representar cada uno de los individuos como posible solución al problema (cada individuo es una combinación de caracteres que representa una distribución de teclado)
2. Crear una población aleatoria (el tamaño de la población es definido anteriormente). Cada individuo es creado con genes aleatorios con valores en 1 y 37 hasta llenar el tamaño de la población deseada. De esta forma cada uno de los individuos es una distribución de teclado aleatoria.
3. Evaluar cada uno de los individuos de la población aleatoria (creada anteriormente). En este paso la función de aptitud evalúa los siguientes criterios: 1) la distancia total recorrida del corpus de cada lenguaje que se desea optimizar; 2) la carga de trabajo que tiene cada uno de los dedos de ambas manos al digitar el corpus de cada lenguaje; 3) la carga de trabajo que tiene cada una de las 3 filas principales del teclado; fila superior, fila origen y fila inferior; 4) la carga de trabajo que tiene cada una de las manos por separado; mano izquierda y mano derecha. Para determinar la evaluación de cada uno de los criterios mencionados, se toma como referencia el corpus del lenguaje seleccionado (en este ejemplo el corpus en lenguaje maya), y con base en este corpus se evalúa cada uno de los criterios en cada uno de los individuos.
4. Se hace uso de los operadores genéticos; operador de selección, operador cruza (se programaron dos operadores de selección: cruza con base en orden y cruza por emparejamiento parcial) y operador de mutación. En la selección por torneo se define previamente el tamaño del torneo, una vez seleccionados a los individuos (selección de los individuos más aptos de la población), se realiza el torneo y los ganadores (individuos más aptos del torneo) son candidatos a ser padres. Posteriormente se define el operador de cruza que se utiliza en el algoritmo

genético. Una vez definido el operador de cruce se toman los individuos candidatos a ser padres y se realiza la combinación de genes entre ellos para crear la nueva descendencia. Posteriormente se aplica una probabilidad de mutación a los individuos que serán mutados y se les aplica una segunda probabilidad de mutación para cada uno de sus genes. Después de ser o no mutados, los individuos son insertados en la nueva población. Este paso 4 se repite hasta crear una población del mismo tamaño a la anterior. La condición para detener el algoritmo genético es definida anteriormente y puede ser hasta llegar a un valor de aptitud deseada o detener la evolución del algoritmo genético en un número máximo de poblaciones creadas.

Después de hacer uso del algoritmo genético, el tercer paso consiste en verificar el comportamiento de la distribución de teclado obtenida por el algoritmo genético. En este caso la distribución es para el lenguaje maya. Para verificar el comportamiento de la distribución obtenida por el algoritmo genético se compara esta distribución contra distribuciones de teclado existentes como: *qwerty*, *dvorak*, etc. Después se verifica que los criterios ergonómicos (distancia recorrida, carga de trabajo por cada dedo, carga de trabajo en las tres filas principales del teclado, carga de trabajo por mano) sean más óptimos en comparación con las distribuciones existentes. Con esto se demuestra que tan óptima es la distribución de teclado obtenida por el algoritmo genético en cuanto a criterios ergonómicos.

Para optimizar una distribución de teclado para escribir en lenguaje maya y español, se elige el o los lenguajes para optimizar (tal como se describe en el paso uno), por ejemplo para optimizar una distribución de teclado para los lenguajes maya y español, se usa el corpus en lenguaje maya y el corpus en el lenguaje español uniendo estos dos corpus para formar uno solo. Posteriormente se procede de nuevo al paso dos donde se usa el algoritmo genético y se termina con el paso 3 evaluando los criterios ergonómicos de la nueva distribución obtenida por el algoritmo genético. De este modo, haciendo los cambios de corpus en los diferentes lenguajes, se optimizan las 4 distribuciones de teclado propuestas.

4.1.1 Configuración del algoritmo genético

Para la implementación del algoritmo genético como método de optimización se tiene la siguiente configuración que consta en el tipo de codificación, los operadores genéticos (operadores de selección, operadores de cruce, operadores de mutación), criterios en la función de evaluación, corpus de entrada de acuerdo a cada lengua a optimizar y tamaño del alfabeto.

4.1.2 Codificación propuesta

La codificación que se implementa en cada uno de los individuos es una codificación tipo numérica, los valores van de 1 a 37, cada individuo tiene 37 genes. El tamaño 37 fue definido de acuerdo al siguiente análisis: en la tabla 4.1 se muestra el alfabeto de la lengua náhuatl de tamaño 18, en este alfabeto se tienen unigramas (A, E, I, O, P, T, C, Z, X, L, Y, M, N) y bigramas (QU, CU, TZ, CH, TL, HU). En la tabla 4.2 se muestra el alfabeto de la lengua maya de tamaño 25, en este alfabeto se tienen unigramas (A, B, E, I, J, K, L, M, N, O, P, S, T, U, W, X, Y), bigramas (CH, K', P', T', TS) y trigramas (CH', TS'). En la tabla 4.3 se muestra el alfabeto de la lengua español de tamaño 27 en el cual solo se tienen unigramas.

Después de obtener los alfabetos correspondientes de cada lengua se obtienen los signos de puntuación y símbolos más frecuentes (20 símbolos agrupados en bigramas) de acuerdo a la lengua náhuatl, maya y español (tabla 4.4), el conjunto de estos signos y símbolos es de tamaño 10.

Tabla 4.1 Alfabeto de la lengua náhuatl.

1 = A	7 = C	13 = Z
2 = E	8 = QU	14 = X
3 = I	9 = CU	15 = L
4 = O	10 = TZ	16 = HU
5 = P	11 = CH	17 = M
6 = T	12 = TL	18 = N

Tabla 4.2 Alfabeto de la lengua maya.

1 = A	10 = L	19 = TS
2 = B	11 = M	20 = TS'
3 = CH	12 = N	21 = U
4 = CH'	13 = O	22 = W
5 = E	14 = P	23 = X
6 = I	15 = P'	24 = Y
7 = J	16 = S	25 = ' (apostrophe)
8 = K	17 = T	
9 = K'	18 = T'	

Tabla 4.3 Alfabeto de la lengua en español.

1 = A	10 = J	19 = R
2 = B	11 = K	20 = S
3 = C	12 = L	21 = T
4 = D	13 = M	22 = U
5 = E	14 = N	23 = V
6 = F	15 = Ñ	24 = W
7 = G	16 = O	25 = X
8 = H	17 = P	26 = Y
9 = I	18 = Q	27 = Z

Tabla 4.4 Signos y símbolos de puntuación más frecuentes de las lenguas náhuatl, maya y español.

1 = - (5 =] }	9 = . :
2 = ')	6 = < >	10 = / *
3 = + #	7 = & \	
4 = [{	8 = , ;	

Para el alfabeto de las lenguas maya y náhuatl se tienen bigramas y trigramas a diferencia del lenguaje español donde solo se tiene bigramas. Los bigramas y trigramas de las lenguas náhuatl y maya se consideran como una sola tecla o un solo carácter sobre un teclado. Por ejemplo, en el lenguaje náhuatl el bigrama TZ es una sola tecla en el teclado o se considera como un solo carácter en su alfabeto, en la lengua maya el trigrama 'CH' es una sola tecla en el teclado o se considera como un solo carácter en su alfabeto. Los 3 lenguajes tienen diferentes tamaños de alfabeto. Para poder pasar a una optimización de teclado en la que se pueda escribir no solo en alguna de las dos lenguas indígenas más habladas de México (náhuatl y maya) sino también el lenguaje español, se unifica el tamaño de los 3 alfabetos, ya que todos los bigramas y trigramas que se tienen en los alfabetos de la lengua náhuatl y de la lengua maya son una combinación de unigramas los cuales aparecen en el alfabeto del lenguaje español. Por tanto, si se tiene el alfabeto del lenguaje español de tamaño 27 (tabla 4.3) se pueden obtener todas las combinaciones para poder escribir en la lengua náhuatl y en la lengua maya. A este conjunto de caracteres se le agregan los signos y símbolos mencionados en la tabla 4.4, y ahora se tiene un conjunto de tamaño 37, cada número del 1 al 37 tiene un valor asignado y así es como hace una decodificación en un individuo (tabla 4.5).

La figura 4.1 representa un individuo codificando la configuración acorde con la tabla 4.5, donde cada número del 1 al 37 es un gen y de acuerdo a la tabla se tiene un valor asignado para cada gen, por ejemplo "1=A, 2=B, 3=C". Así para todos los genes del individuo.

Tabla 4.5 Valor de cada gen en un individuo (decodificación).

1 = A	14 = N	27 = Z
2 = B	15 = Ñ	28 = - (
3 = C	16 = O	29 = =)
4 = D	17 = P	30 = + #
5 = E	18 = Q	31 = [{
6 = F	19 = R	32 =] }
7 = G	20 = S	33 = < >
8 = H	21 = T	34 = & \
9 = I	22 = U	35 = , ;
10 = J	23 = V	36 = . :
11 = K	24 = W	37 = / *
12 = L	25 = X	
13 = M	26 = Y	

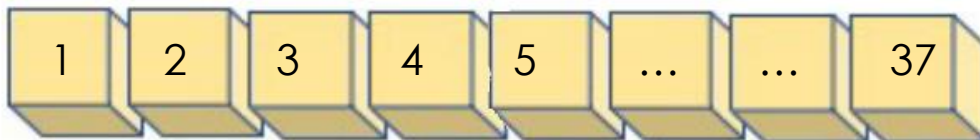


Figura 4.1 Codificación de los individuos y sus genes.

La figura 4.2 representa un individuo como una distribución de teclado (una posible solución al problema). De esta forma, cada individuo tiene un orden diferente en sus genes y por tanto es una nueva distribución de teclado (una solución diferente). Tal como se observa en la figura 4.2, se modifican solo las teclas que tienen un número asignado. Esta configuración se usa para obtener las 4 distribuciones de teclado.

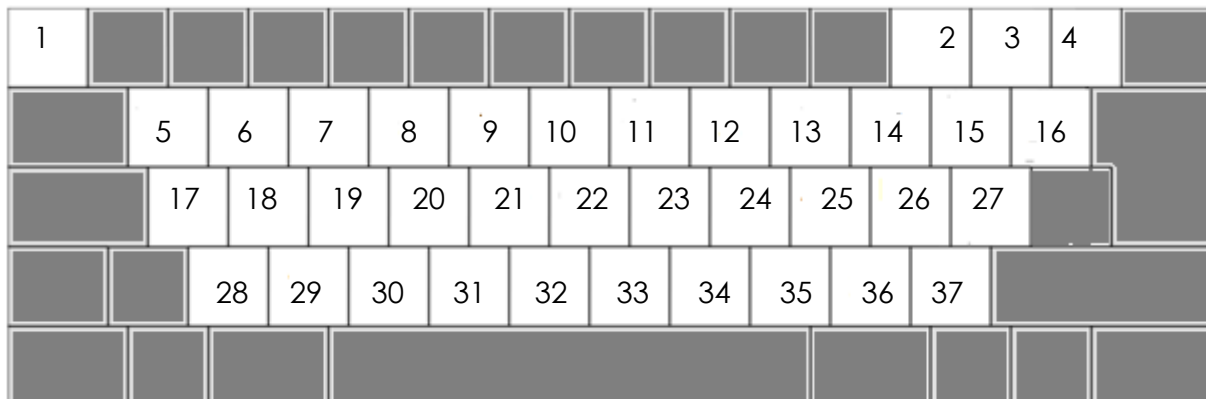


Figura 4.2 Representación de los genes de un individuo sobre una distribución de teclado como una posible solución.

4.1.3 Función de aptitud del método propuesto

Dentro de la función de evaluación se toman en cuenta diferentes criterios para evaluar a cada uno de los individuos de la población, los criterios que son:

- El uso de los 5 dedos de ambas manos para respetar la forma original del teclado *qwerty*.
- La distancia que se genera en un conjunto de teclas al realizar diferentes movimientos con los dedos de ambas manos.
- La carga de trabajo que se tiene por las 3 filas principales del teclado, siguiendo la figura 4.3 se indica la superior (1), fila origen (2) y fila inferior (3). La función de aptitud evalúa que porcentaje de carga de trabajo tienen cada una de las filas.
- La carga de trabajo que se tiene por las 8 columnas del teclado, siguiendo la figura 4.3 se indican las columnas de acuerdo a la técnica de digitación del teclado *qwerty* mencionada en el capítulo 1 (figura 1.5). La columna (6) es un conjunto de teclas que

le corresponde al dedo meñique de la mano izquierda, la columna (7) es un conjunto de teclas que le corresponde al dedo anular de la mano izquierda, la columna (8) es un conjunto de teclas que le corresponde al dedo medio de la mano izquierda, la columna (9) es un conjunto de teclas que le corresponde al dedo índice de la mano izquierda, la columna (10) es un conjunto de teclas que le corresponde al dedo índice de la mano derecha, la columna (7) es un conjunto de teclas que le corresponde al dedo medio de la mano derecha, la columna (8) es un conjunto de teclas que le corresponde al dedo anular de la mano derecha, la columna (9) es un conjunto de teclas que le corresponde al dedo meñique de la mano derecha. La función de aptitud evalúa que porcentaje de carga de trabajo tienen cada una de las columnas y de esta forma se sabe la carga de trabajo que tiene cada uno de los dedos de ambas manos.

- La carga de trabajo para ambas manos, de acuerdo con la técnica de digitación, el conjunto de teclas de la parte (4) le corresponde a la mano izquierda y el conjunto de teclas de la parte (5) le corresponde a la mano derecha. La función de aptitud evalúa que porcentaje de carga de trabajo tienen cada una de estas dos partes del teclado, para saber qué mano tiene una mayor carga de trabajo que la otra y así poder realizar un respectivo balance (aproximadamente un 50% de carga de trabajo del lado izquierdo y un 50% de carga de trabajo del lado derecho).

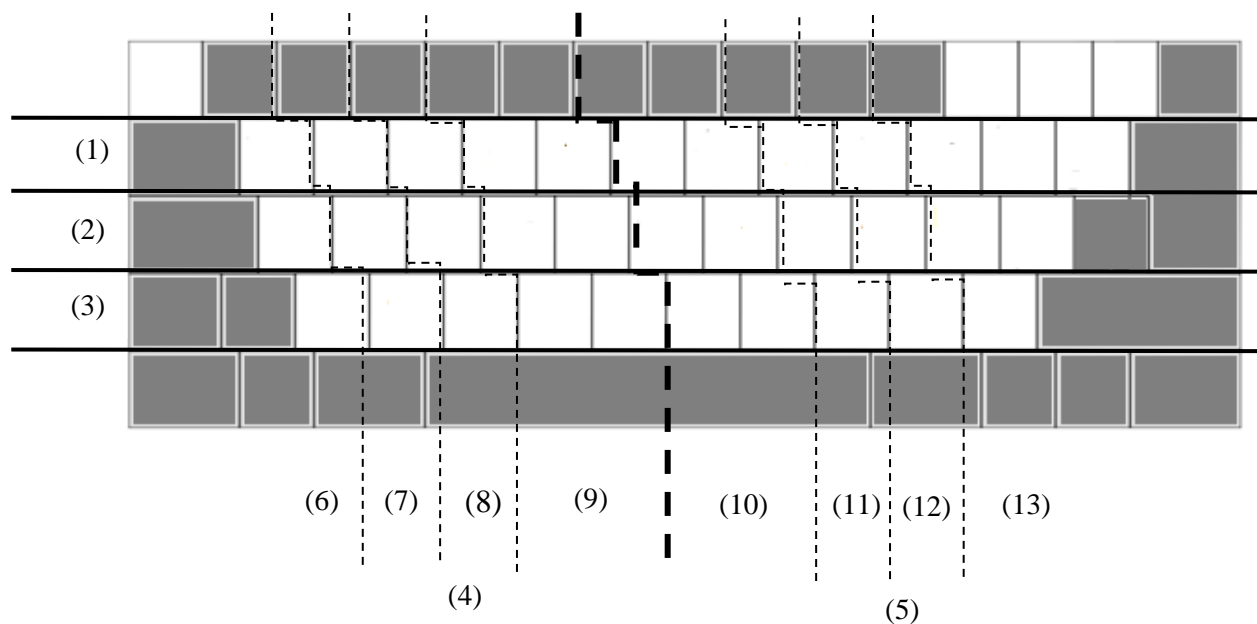


Figura 4.3 Separación de filas por columnas en el teclado.

Para calcular la distancia total recorrida al escribir un conjunto de palabras se implementa una herramienta calculado la distancia entre teclas. El estándar de una tecla a otra es de 0.018 metros y en algunas teclas esta distancia es mayor. La herramienta proporciona la distancia total recorrida al digitar un conjunto de palabras. La figura 4.4 muestra un ejemplo al escribir la palabra "TECLADO". En la figura 4.4 se observa la sumatoria de la distancia generada con cada movimiento de los dedos, y de acuerdo la técnica de digitación del teclado *qwerty* y el orden de los movimientos realizados por cada dedo. Del mismo modo, la herramienta indica la parte del teclado que se está escribiendo con mayor frecuencia, tanto por filas como por columnas. Siguiendo la figura 4.4, para escribir la palabra "TECLADO" los dedos se posicionan en las teclas de la fila origen (marcadas con una línea color negro), ésta es la posición inicial de los dedos para empezar una digitación, ahora el dedo índice de la mano izquierda se encuentra en la letra F y se desplaza a la letra T en la cual se tiene una distancia de 0.021 metros. Este mismo dedo debe regresar a su posición inicial (letra F) y en su regreso genera la misma distancia de 0.021 metros, por tanto ahora la distancia recorrida es de 0.042 metros (sumatoria del desplazamiento y el regreso), el segundo movimiento se produce en el dedo medio de la mano izquierda desde la letra D (posición inicial) a la letra E en la que se tiene una distancia de 0.018 metros, sumada a la distancia anterior se tiene ahora 0.060 metros. Este mismo dedo debe regresar a su posición inicial (letra D) y vuelve a generar una distancia de 0.018 metros. Ahora la distancia total es de 0.078 metros, el tercer movimiento se realiza con este mismo dedo (dedo medio mano izquierda) desde su posición inicial (letra D) a la letra C generando una distancia de 0.020 metros. Por tanto, en el regreso sumaría 0.040 metros y agregado a la distancia que se lleva da un total de 0.118 metros. El cuarto movimiento se da con el dedo anular de la mano izquierda pulsando la letra L (hasta ahora se ha digitado "TECL") y como es una posición inicial de los dedos, esta no genera desplazamiento alguno y por tanto, su distancia al digitar la letra L es de 0 metros y se mantiene la distancia de 0.118 metros. El quinto movimiento para poder escribir la letra A de igual forma se genera en una posición inicial de los dedos, en el dedo meñique de la mano izquierda con una distancia de 0 metros continuando con la misma distancia de 0.118 metros. Para escribir el sexto carácter (D) de la palabra "TECLADO" este se da en una posición inicial con el dedo medio de la mano izquierda generando una distancia de 0 metros y manteniendo como distancia total 0.118 metros. El último movimiento se da en el dedo anular de la mano derecha desde su posición inicial (letra L) a la letra O generando un distancia de 0.018 metros y el dedo debe regresar a su posición inicial generando otra distancia de 0.018 metros. Estas dos distancias se suman a la

distancia recorrida anterior y se tiene un total de 0.154 metros (esta es la distancia total recorrida al escribir la palabra "TECLADO" sobre la distribución qwerty).



Figura 4.4 Digitación de la palabra ‘TECLADO’ para calcular la distancia recorrida sobre el teclado.

4.1.4 Determinación de frecuencia de caracteres

La frecuencia de caracteres es una de las variables más importantes en la optimización de las distribuciones de teclado y juega un papel muy importante dentro de la función de evaluación ya que con esto se sabe qué tan óptima es una distribución de teclado para uno o varios lenguajes.

Para que la función de aptitud indique si una distribución de teclado es óptima para uno o varios lenguajes, se debe tener como referencia un corpus de dichos lenguajes para el o los cuales se desea optimizar la distribución. Para esta tesis el objetivo es optimizar 4 distribuciones de teclado que son: 1) distribución de teclado para la lengua indígena náhuatl; 2) distribución de teclado para la lengua indígena maya, 3) distribución de teclado para las lenguas español – náhuatl; 4) distribución de teclado para las lenguas español – maya. Como tal solo intervienen 3 lenguajes en esta optimización que son; náhuatl, maya y español. Por tanto se tienen 3 corpus diferentes (figura 4.5) para determinar la frecuencia de caracteres en cada lenguaje, un corpus para determinar la frecuencia de caracteres en la lengua náhuatl, un corpus para determinar la frecuencia de caracteres en la lengua maya, un corpus para determinar la frecuencia de caracteres en el español. Cada uno de estos corpus contiene se crea con las palabras más frecuentes de cada lenguaje.

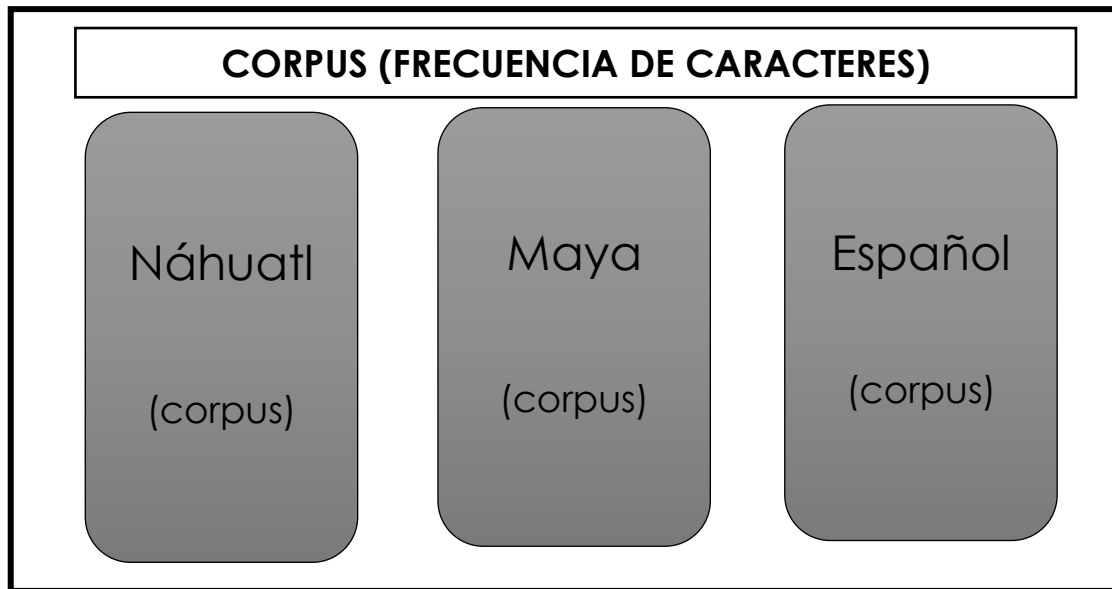


Figura 4.5 Corpus con frecuencia de caracteres.

4.1.5 Operadores genéticos

El algoritmo genético del método propuesto se compone con operadores de selección, cruce (dos operadores de cruce para obtener y comparar los diferentes resultados) y mutación de la siguiente manera:

- **Operador de selección:** SELECCIÓN POR TORNEO (descrito en el capítulo 2), es un operador genético de selección en el algoritmo genético, ya que con base en la teoría, es este operador se descarta a los individuos con menor aptitud y se seleccionan a los más fuertes.
- **Operador de cruce Nº 1:** CRUZA CON BASE EN ORDEN OX (descrito en el capítulo 2), es un operador genético de cruce para intercambiar genes entre dos individuos y así crear la descendencia o nueva población que cada vez se adapta más como solución al problema y va evolucionando teniendo cada vez mejores aptitudes en cada uno de los genes de cada individuo.
- **Operador de cruce Nº 2:** CRUZA POR EMPAREJAMIENTO PARCIAL (descrito en el capítulo 2) es un segundo operador genético de cruce para poder intercambiar genes entre

dos individuos y así crear la descendencia o nueva población que cada vez se adapta más como solución al problema. Los resultados son diferentes al utilizar un operador de cruza diferente ya que se siguen otros pasos.

- **Operador de mutación:** MUTACIÓN POR INTERCAMBIO (descrito en el capítulo 2), es un operador genético que realiza su función después del operador de cruza, el principal objetivo es aplicar una probabilidad de mutación a cada uno de los individuos y si alguno de estos es candidato a ser mutado, sus genes son modificados de acuerdo a los pasos que se siguen en la mutación por intercambio y esto puede representar una mejora en la población actual.
-

4.2 Resumen del capítulo

En este capítulo, se da a conocer el método propuesto el cual es un algoritmo genético, se hace mención sobre la configuración que se utiliza y cada uno de los operadores genéticos. Así mismo la conformación de los corpus que se utilizan para llegar a las optimización de teclado y los criterios que se evalúan para determinar si las distribuciones de teclado son distribuciones optimizadas y con criterios ergonómicos.



5. Experimentación

Con base en la hipótesis planteada en esta tesis que consiste en demostrar que si se configuran las posiciones de las teclas sobre un teclado implementando un método de optimización (algoritmo genético), se podrían encontrar distribuciones de teclado más óptimas en cuanto a criterios ergonómicos en comparación con la distribución *qwerty*.

En este capítulo se muestran los experimentos realizados durante la implementación del método propuesto (algoritmo genético), se muestra el comportamiento de las distribuciones optimizadas por el algoritmo genético y se hace una comparación de estas distribuciones contra distribuciones de teclado ya existentes (*qwerty*, *dvorak*, etc). También se muestra la curva de aprendizaje promedio para poder aprender a digitar sobre una distribución de teclado optimizada con criterios económicos y la curva de aprendizaje para una distribución de teclado no optimizada y sin criterios ergonómicos.

5.1 Descripción de un corpus

Para cada una de las 4 optimizaciones de las distribuciones de teclado se crean los corpus correspondientes para que al hacer uso de estos se determine la aptitud de cada uno de los individuos.

Para la creación de los 3 corpus principales, primero se obtiene un conjunto de palabras con tamaño de 16,000 caracteres. Este conjunto de palabras se obtiene para cada lenguaje (español, náhuatl y maya), y dentro de este conjunto se tienen las palabras más frecuentes de cada lenguaje.

Tal como se menciona anteriormente, solo intervienen 3 lenguajes en la optimización de las 4 distribuciones de teclado, se cuenta con el lenguaje náhuatl, maya y el español. Por tanto se utilizan 4 corpus diferentes para cada optimización, los corpus se crearon y se configuraron de la siguiente manera (figura 5.1):

- Corpus en lenguaje náhuatl conformado por un conjunto de palabras en lenguaje náhuatl (16,000 caracteres), se utiliza para realizar la primera optimización de teclado.
- Corpus en lenguaje maya conformado por un conjunto de palabras en lenguaje maya (16,000 caracteres), se utiliza para realizar la segunda optimización de teclado.
- Corpus en los lenguajes español - náhuatl conformado por un conjunto de palabras en lenguaje español y náhuatl (32,000 caracteres), se utiliza para realizar la tercera optimización de teclado,
- Corpus en los lenguajes español - maya conformado por un conjunto de palabras en lenguaje español y maya (32,000 caracteres), se utiliza para realizar la cuarta optimización de teclado,

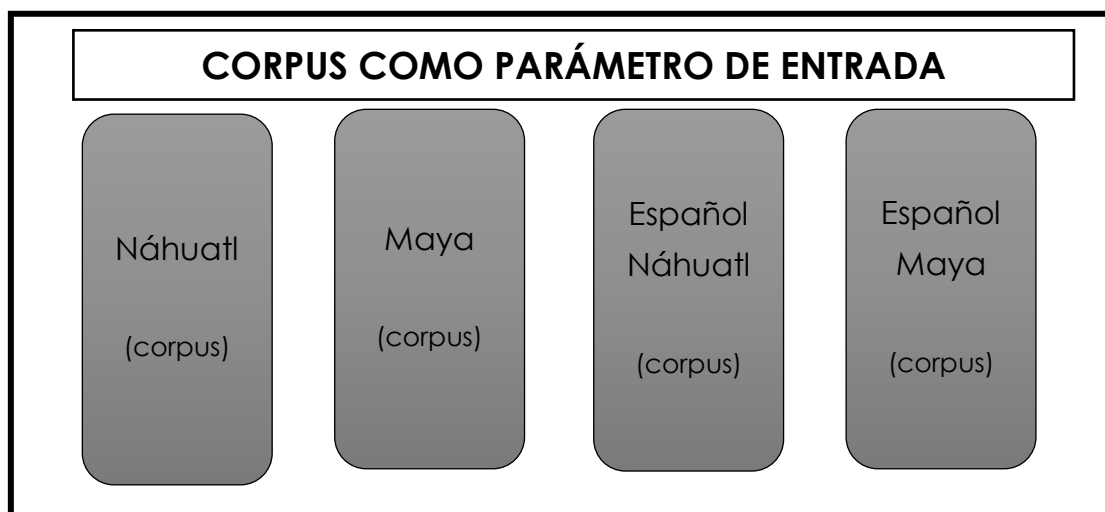


Figura 5.1 Corpus para cada optimización.

5.2 Evolución del algoritmo genético y optimización de parámetros de configuración

Una vez teniendo la codificación se define el tamaño de la población (número de individuos), valor de la función de aptitud que se desea alcanzar, tamaño del torneo para el operador de cruza, número de poblaciones máximas que puede alcanzar el algoritmo genético para detener el ciclo, la probabilidad de mutación para cada individuo y la probabilidad de mutación para los genes de cada individuo.

Posteriormente se realizan diferentes pruebas y diferentes combinaciones de parámetros para llegar a la optimización deseada (para cada una de las 4 distribuciones de teclado) en el menor tiempo.

La primera optimización con el algoritmo genético se realiza con los siguientes parámetros:

- **Corpus de entrada:** corpus de la lengua náhuatl.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 400 poblaciones
- **Operador de cruza:** cruza con base en orden
- **Probabilidad de mutación por individuo:** 1.5% de probabilidad
- **Probabilidad de mutación por genes:** 5% de probabilidad

La figura 5.2 muestra el comportamiento que tiene el algoritmo genético evaluando la distancia en metros. En el eje X se tiene el número de poblaciones máximas, que son 500 poblaciones, y en el eje Y el valor de la aptitud del mejor individuo, su valor de aptitud del mejor individuo de la población inicial empieza en 4,120 metros que es un valor alto en comparación con el valor aproximado que se desea alcanzar. Esto se debe a que la primera población se crea con genes de forma aleatoria. Para poder aproximarse al valor deseado, el algoritmo genético evoluciona en cada población minimizando cada vez más el valor de aptitud, tal como se muestra en la figura 5.2. Esta configuración optimiza el valor de la aptitud hasta 1,397 metros (el individuo más apto de las 500 poblaciones).

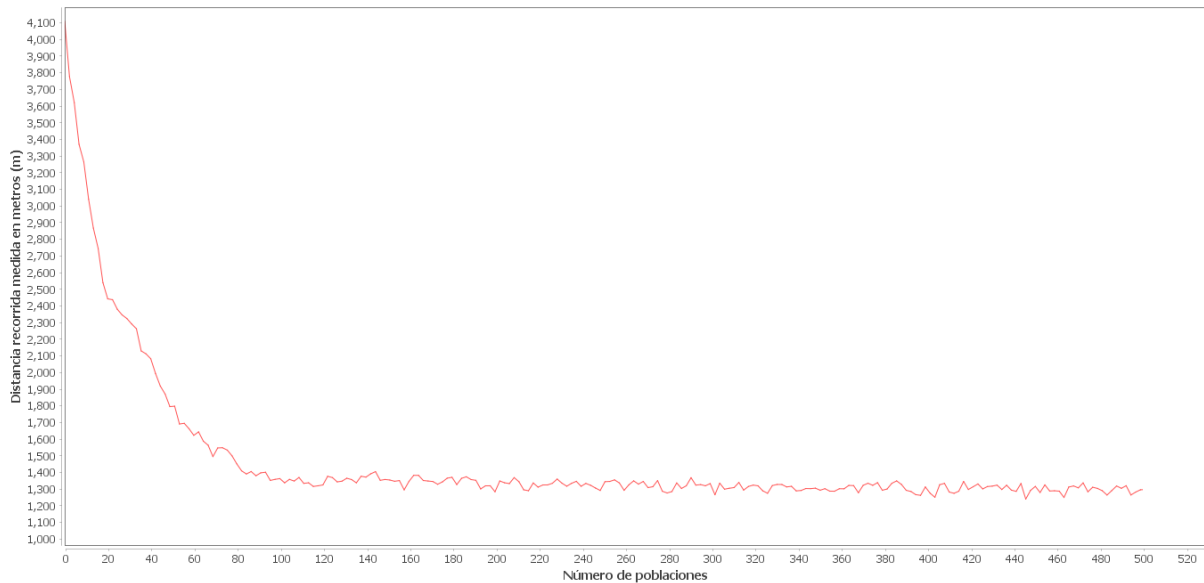


Figura 5.2 Valor de la función de aptitud evaluando la distancia: experimento 1.

La segunda prueba de optimización con el algoritmo genético se realiza con los siguientes parámetros:

- **Corpus de entrada:** corpus de la lengua náhuatl.
- **Tamaño de torneo (operador de selección):** 4 individuos
- **Número máximo de poblaciones:** 500 poblaciones
- **Operador de cruza:** cruza con base en orden
- **Probabilidad de mutación por individuo:** 2.5% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

La figura 5.3 muestra el comportamiento que tiene el algoritmo genético evaluando la aptitud en metros. En el eje X se tiene el número de poblaciones máximas, que son 500 poblaciones, y en el eje Y el valor de la distancia del mejor individuo, su valor de aptitud del mejor individuo de la población inicial empieza en 4,020 metros que es un valor alto en comparación con el valor aproximado que se desea alcanzar. Esto se debe a que la primera población se crea con genes de forma aleatoria. Para poder aproximarse al valor deseado, el algoritmo genético evoluciona en cada población minimizando cada vez más el valor de aptitud tal como se muestra en la figura 5.3. Esta configuración optimiza el valor de la aptitud hasta 1,697 metros (el individuo más apto de las 500 poblaciones).

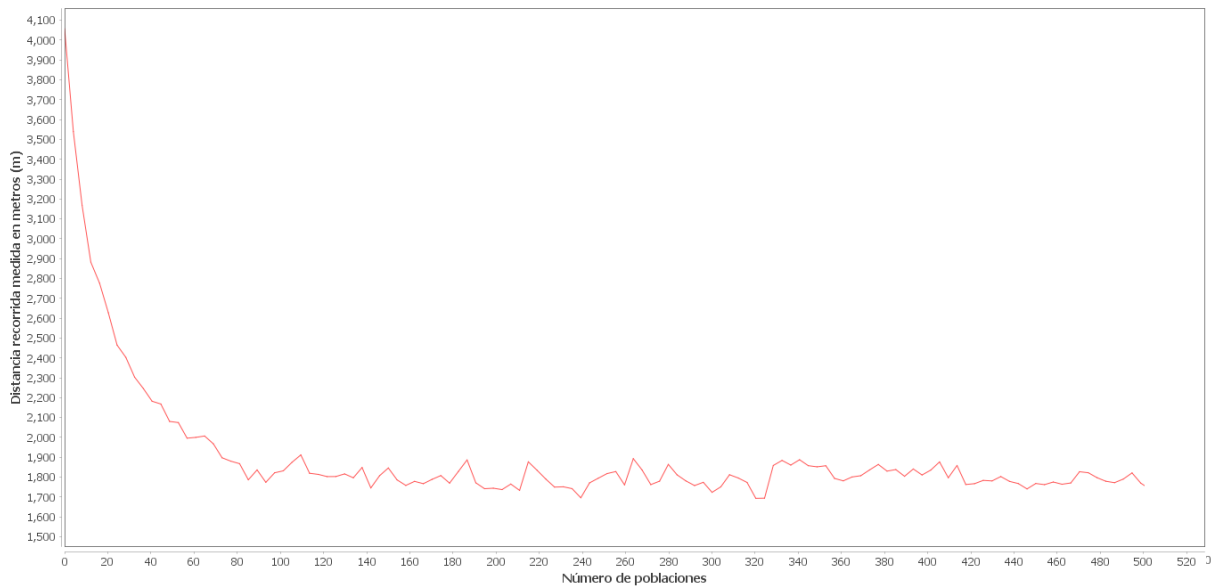


Figura 5.3 Valor de la función de aptitud evaluando la distancia: experimento 2.

La tercera prueba de optimización con el algoritmo genético se realiza con los siguientes parámetros:

- **Corpus de entrada:** corpus de la lengua náhuatl.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 500 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 1.5% de probabilidad
- **Probabilidad de mutación por genes:** 5% de probabilidad

La figura 5.4 muestra el comportamiento que tiene el algoritmo genético evaluando la aptitud en metros. En el eje X se tiene el número de poblaciones máximas, que son 500 poblaciones, y en el eje Y el valor de la distancia del mejor individuo, su valor de aptitud del mejor individuo de la población inicial empieza en 4,420 metros que es un valor alto en comparación con el valor aproximado que se desea alcanzar. Esto se debe a que la primera población se crea con genes de forma aleatoria. Para poder aproximarse al valor deseado, el algoritmo genético evoluciona en cada población minimizando cada vez más el valor de aptitud tal como se muestra en la figura 5.4. Esta configuración optimiza el valor de la aptitud hasta 1,357 metros (el individuo más apto de las 500 poblaciones).

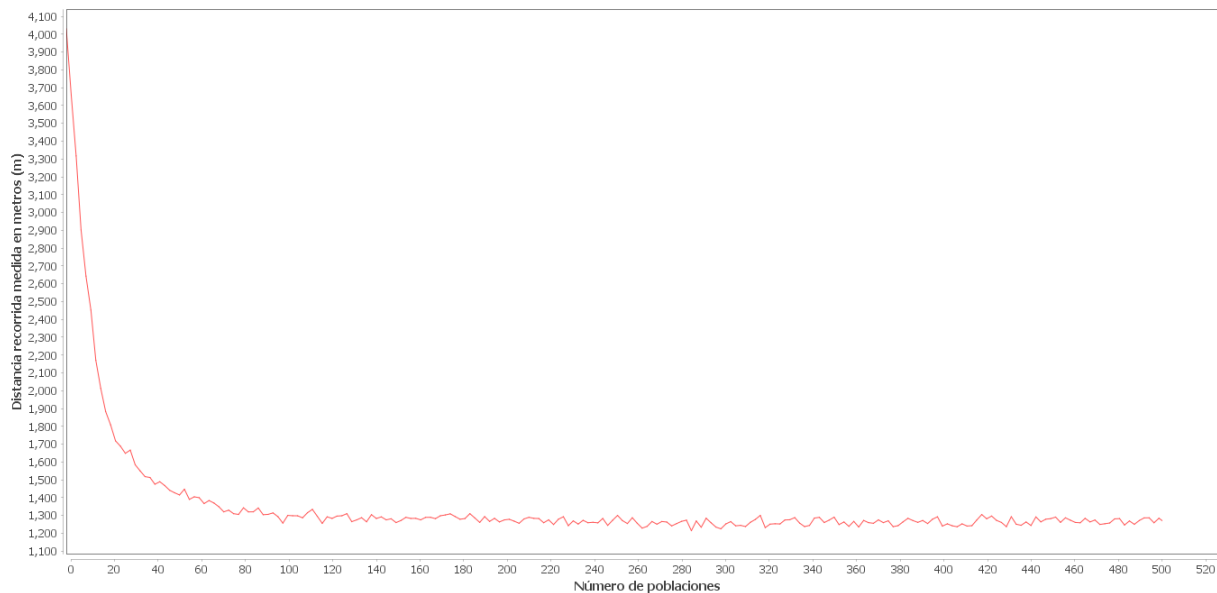


Figura 5.4 Valor de la función de aptitud evaluando la distancia: experimento 3.

La cuarta prueba de optimización con el algoritmo genético se realiza con los siguientes parámetros:

- **Corpus de entrada:** corpus de la lengua náhuatl.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 500 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 3% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

En la figura 5.5 se muestra el comportamiento que tiene el algoritmo genético evaluando la aptitud en metros. En el eje X se tiene el número de poblaciones máximas que son 500 poblaciones, y en el eje Y el valor de la distancia del mejor individuo, su valor de aptitud del mejor individuo de la población inicial empieza en 4,100 metros que es un valor alto en comparación con el valor aproximado que se desea alcanza. Esto se debe a que la primera población se crea con genes de forma aleatoria. Para poder aproximarse al valor deseado, el algoritmo genético evoluciona en cada población minimizando cada vez más el valor de aptitud tal como se muestra en la figura 5.5. Esta configuración optimiza el valor de la aptitud hasta 1,327 metros (el individuo más apto de las 500 poblaciones).



Figura 5.5 Valor de la función de aptitud evaluando la distancia: experimento 4.

Después de realizar diferentes combinaciones de parámetros de entrada para el algoritmo genético, las pruebas para optimizar las demás distribuciones de teclado se realizan con la siguiente configuración:

La optimización de una distribución de teclado para escribir en la lengua náhuatl se realiza con los siguientes parámetros de entrada:

- **Corpus de entrada:** corpus de la lengua náhuatl.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 500 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 3% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

La optimización de una distribución de teclado para escribir en la lengua maya se realiza con los siguientes parámetros de entrada:

- **Corpus de entrada:** corpus de la lengua maya.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 500 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 3% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

La optimización de una distribución de teclado para escribir en las lenguas español y náhuatl se realiza con los siguientes parámetros de entrada:

- **Corpus de entrada:** corpus del lenguaje náhuatl y español.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 700 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 3% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

La optimización de una distribución de teclado para escribir en las lenguas español y maya se realiza con los siguientes parámetros de entrada:

- **Corpus de entrada:** corpus del lenguaje maya y español.
- **Tamaño de torneo (operador de selección):** 2 individuos
- **Número máximo de poblaciones:** 700 poblaciones
- **Operador de cruza:** emparejamiento parcial
- **Probabilidad de mutación por individuo:** 3% de probabilidad
- **Probabilidad de mutación por genes:** 7% de probabilidad

5.3 Comportamiento de las 4 distribuciones de teclado optimizadas por el algoritmo genético

Después de obtener las 4 distribuciones de teclado optimizadas por el algoritmo genético, se realizan diferentes pruebas para mostrar cada uno de los criterios optimizados en las 4 distribuciones. Estas pruebas son comparativas con otras distribuciones de teclado existentes. En especial se comparan las nuevas 4 distribuciones de teclado (obtenidas por el algoritmo genético) contra el teclados ya existentes que cuentan con las mismas características de forma física pero principalmente se compara contra el teclado *qwerty*, ya que es el teclado

estándar más utilizado desde su aparición en las máquinas de escribir y ahora en los teclados de computadora.

Para observar el comportamiento de la distribución de teclado para la lengua náhuatl obtenida por el algoritmo genético, se realiza una comparación contra otras distribuciones de teclado existentes como; *qwerty*, *dvorak*, *arensito*, *capewell – qwerty*, *capewell - dvorak*. Para hacer esta comparación se usa el corpus de la lengua náhuatl (utilizado en el algoritmo) y se compararon los siguientes criterios:

- En la tabla 5.1 se muestra la distancia de las diferentes distribuciones de teclado y la distribución de teclado para la lengua náhuatl, la distancia de la distribución de teclado para la lengua náhuatl en comparación con las distancias de las otras distribuciones de la tabla 5.1 es la menor distancia (1327 metros). Con este criterio, entre menor sea la distancia, se reducen los movimientos innecesarios para todos los dedos de ambas manos.
- En la tabla 5.2 se muestra la carga de trabajo en las 3 filas principales del teclado (fila superior, fila origen y fila inferior). Anteriormente se había mencionado la importancia de la fila origen, es la fila más importante del teclado ya que en esta fila se posicionan inicialmente los dedos de ambas manos para iniciar una digitación y es donde la distancia es 0 metros al digitar las teclas que se encuentran en una posición inicial. Como se puede observar en la tabla 5.2, la distribución de teclado para la lengua náhuatl tiene la mayor carga de trabajo (79.2%) en la fila origen. Este criterio, al tener la mayor carga de trabajo en la fila origen (posición inicial de los dedos), reduce el número de desplazamientos hacia arriba (desplazarse a la parte superior del teclado) y hacia abajo (desplazarse a la parte inferior del teclado) para todos los dedos de ambas manos.
- En la tabla 5.3 se muestra la carga de trabajo de cada uno de los dedos de ambas manos. Como se puede observar en la tabla 5.3, la distribución de teclado para la lengua náhuatl tiene un balance de forma jerárquica dando prioridad en primer lugar a los dedos índices (con la mayor carga de trabajo), en segundo lugar los dedos medios, en tercer lugar a los dedos anulares, y en cuarto lugar a los dedos meñiques (con la menor carga de trabajo). Este criterio tiene como objetivo proporcionar la carga de trabajo de acuerdo a la habilidad y fuerza de cada dedo. Esto evita que

alguno de los dedos se sobre cargue y provoque una futura lesión. Tal como se observa la distribución de teclado para la lengua náhuatl tiene, en los dedos índices, la mayor carga de trabajo; ya que son los dedos más fuertes y con mayor habilidad. Los dedos meñiques tienen la menor carga de trabajo por ser los dedos con menor fuerza y habilidad de movimiento.

- La tabla 5.4 muestra el balance que tienen ambas manos e indica la mano con más carga de trabajo. El teclado *qwerty* tiene una carga de trabajo de 34% en la mano izquierda y una carga de trabajo de 63% en la mano derecha; la diferencia entre estos porcentajes es de 29 %, lo que quiere decir que la mano derecha está sobrecargada (al tener el mayor porcentaje de carga de trabajo) con un 29% más, que la mano izquierda. Este criterio (balance de ambas manos) tiene como objetivo balancear la carga de trabajo en ambas partes del teclado para ambas manos, al aproximarse a los valores 50% de carga de trabajo en el lado izquierdo y 50% de carga de trabajo en el lado derecho se da paso a un flujo continuo de digitación, se evitan futuras lesiones en cualquiera de las muñecas de ambas manos y se evita la fatiga causada por sobre carga en alguna de las dos manos. Con esta tabla 5.4 se demuestra que la distribución de teclado para la lengua náhuatl está balanceado de forma optimizada en comparación con las demás distribuciones de la tabla ya que la carga en ambas manos tienen una diferencia mínima indicando que la mano derecha está un 2% más sobrecargada que la mano izquierda.

Tabla 5.1 Distancia total recorrida en metros.

Distancia recorrida (corpus en lenguaje náhuatl)	
Teclado	Distancia
Teclado propuesto (náhuatl)	1327 m
<i>qwerty</i>	3091 m
<i>dvorak</i>	2087 m
<i>arensito</i>	1901 m
<i>capewell - qwerty</i>	1958 m
<i>capewell - dvorak</i>	2087 m

Tabla 5.2 Carga de trabajo por las 3 filas principales del teclado.

Carga de trabajo por fila (corpus en lenguaje náhuatl)			
Teclado	Filas superior	Fila origen	Fila inferior
Teclado propuesto (náhuatl)	16.7 %	79.2%	3.8%
<i>qwerty</i>	44.2 %	36.5 %	19.1 %
<i>dvorak</i>	16.3 %	68.1 %	15.4 %
<i>arensito</i>	22.2 %	61.5 %	16.1 %
<i>capewell - qwerty</i>	20.5 %	61.6 %	10.2 %
<i>capewell - dvorak</i>	15 %	68.1 %	16.6 %

Tabla 5.3 Carga de trabajo por cada dedo de ambas manos.

Carga de trabajo sobre los dedos de ambas manos (corpus en lenguaje náhuatl)								
Teclado	Dedos mano izquierda				Dedos mano derecha			
	Meñique	Anular	Medio	Índice	Índice	Medio	Anular	Meñique
Teclado propuesto (náhuatl)	5 %	13 %	14 %	15 %	16 %	12 %	13 %	8 %
<i>qwerty</i>	15 %	4 %	7 %	8 %	23 %	23 %	13 %	4 %
<i>dvorak</i>	14 %	8 %	8 %	30 %	7 %	10 %	8 %	11 %
<i>arensito</i>	15 %	6 %	7 %	12 %	7 %	21 %	11 %	16 %
<i>capewell - qwerty</i>	15 %	4 %	7 %	8 %	23 %	17 %	13 %	10 %
<i>capewell - dvorak</i>	7%	17 %	8 %	23 %	14 %	8 %	8 %	11 %

Tabla 5.4 Carga de trabajo en ambas manos.

Carga de trabajo por mano (corpus en lenguaje náhuatl)			
Teclado	Mano izquierda	Mano derecha	Diferencia de carga en cada mano
Teclado propuesto (náhuatl)	47 %	49 %	2 %
<i>qwerty</i>	34 %	63 %	29 %
<i>dvorak</i>	60 %	36 %	24 %
<i>arensito</i>	40 %	55 %	15 %
<i>capewell - qwerty</i>	34 %	63 %	29 %
<i>capewell - dvorak</i>	55 %	41 %	14 %

Para observar el comportamiento de la distribución de teclado para la lengua maya obtenida por el algoritmo genético, se realiza una comparación contra otras distribuciones de teclado existentes como; *qwerty*, *dvorak*, *arensito*, *capewell – qwerty*, *capewell - dvorak*. Para hacer esta comparación, se usa el corpus de la lengua maya (utilizado en el algoritmo) y se compararon los siguientes criterios:

- En la tabla 5.5 se muestra la distancia de las diferentes distribuciones de teclado y la distribución de teclado para la lengua maya, la distancia de la distribución de teclado para la lengua maya en comparación con las distancias de las otras distribuciones de la tabla 5.5 es la menor distancia (1351 metros). Con este criterio, entre menor sea la distancia, se reducen los movimientos innecesarios para todos los dedos de ambas manos.
- En la tabla 5.6 se muestra la carga de trabajo en las 3 filas principales del teclado (fila superior, fila origen y fila inferior). Como se puede observar en la tabla 5.6, la distribución de teclado para la lengua maya tiene la mayor carga de trabajo 69.2% en la fila origen. Este criterio, al tener la mayor carga de trabajo en la fila origen (posición inicial de los dedos), se reduce el número de desplazamientos hacia arriba (desplazarse a la parte superior del teclado) y hacia abajo (desplazarse a la parte inferior del teclado) para todos los dedos de ambas manos.
- En la tabla 5.7 se muestra la carga de trabajo de cada uno de los dedos de ambas manos. Como se puede observar en la tabla 5.7 la distribución de teclado para la lengua maya tiene un balance de forma jerárquica dando prioridad en primer lugar a los dedos índices (con la mayor carga de trabajo), en segundo lugar los dedos medios, en tercer lugar a los dedos anulares, y en cuarto lugar a los dedos meñiques (con la menor carga de trabajo). Este criterio tiene como objetivo proporcionar la carga de trabajo a cada dedo de acuerdo a la habilidad y fuerza que tiene cada uno de estos para evitar que alguno de los dedos este sobre cargado y provocar una futura lesión. Tal como se observa la distribución de teclado para la lengua maya tiene en los dedos índices tienen la mayor carga de trabajo ya que son los dedos más fuertes y con mayor habilidad, los dedos meñiques tienen la menor carga de trabajo por ser los dedos con menor fuerza y habilidad de movimiento.
- La tabla 5.8 muestra el balance de carga de trabajo que tienen ambas manos. Así mismo se indica cuál de las dos manos tiene sobre carga de trabajo. El teclado *qwerty*

tiene una carga de trabajo de 36% en la mano izquierda y una carga de trabajo de 60% en la mano derecha, la diferencia entre estos porcentajes es de 24%, esto quiere decir que la mano derecha esta sobre cargada (al tener el mayor porcentaje de carga de trabajo) un 24% más que la mano izquierda. Este criterio (balance de ambas manos) tiene como objetivo balancear la carga de trabajo en ambas partes del teclado para ambas manos. Al aproximarse a los valores 50% de carga de trabajo en el lado izquierdo y 50% de carga de trabajo en el lado derecho se da paso a un flujo continuo de digitación, se evitan futuras lesiones en cualquiera de las muñecas de ambas manos y se evita la fatiga causada por sobre carga en alguna de las dos manos. La tabla 5.8 muestra que la distribución de teclado para la lengua maya está balanceado de forma optimizada en comparación con las demás distribuciones ya que la carga en ambas manos tienen una diferencia mínima indicando que la mano derecha está un 3% más sobrecargada que la mano izquierda.

Tabla 5.5 Distancia total recorrida en metros.

Distancia recorrida (corpus en lenguaje maya)	
Teclado	Distancia
Teclado propuesto (maya)	1351 m
<i>qwerty</i>	2257 m
<i>dvorak</i>	1631 m
<i>arensito</i>	1807 m
<i>capewell – qwerty</i>	1796 m
<i>capewell – dvorak</i>	1723 m

Tabla 5.6 Carga de trabajo por las 3 filas principales del teclado.

Carga de trabajo por fila (corpus en lenguaje maya)			
Teclado	Filas superior	Fila origen	Fila inferior
Teclado propuesto (maya)	22.3 %	69.2 %	8.1 %
<i>qwerty</i>	40.1 %	41.5 %	17.6 %
<i>dvorak</i>	16.7 %	63.8 %	19.2 %
<i>arensito</i>	26.4 %	57.1 %	16.2 %
<i>capewell - qwerty</i>	22.6 %	60.1 %	16.5 %
<i>capewell - dvorak</i>	15.1 %	63.8 %	20.8 %

Tabla 5.7 Carga de trabajo por cada dedo de ambas manos.

Carga de trabajo sobre los dedos de ambas manos (corpus en lenguaje maya)								
Teclado	Dedos mano izquierda				Dedos mano derecha			
	Meñique	Anular	Medio	Índice	Índice	Medio	Anular	Meñique
Teclado propuesto (maya)	5 %	13 %	13 %	16 %	16 %	16 %	11 %	7 %
<i>qwerty</i>	16 %	4 %	10 %	6 %	25 %	15 %	14 %	5 %
<i>dvorak</i>	17 %	7 %	13 %	25 %	9 %	9 %	5 %	11 %
<i>arensito</i>	16 %	7 %	10 %	13 %	8 %	17 %	8 %	17 %
<i>capewell - qwerty</i>	16 %	4 %	10 %	6 %	26 %	8 %	14 %	12 %
<i>capewell - dvorak</i>	7 %	18 %	11 %	18 %	14 %	7 %	9 %	12 %

Tabla 5.8 Carga de trabajo en ambas manos.

Carga de trabajo por mano (corpus en lenguaje maya)			
Teclado	Mano izquierda	Mano derecha	Diferencia de carga en cada mano
Teclado propuesto (maya)	47 %	50 %	3 %
<i>qwerty</i>	36 %	60 %	24 %
<i>dvorak</i>	62 %	34 %	28 %
<i>arensito</i>	46 %	50 %	4 %
<i>capewell - qwerty</i>	36 %	60 %	24 %
<i>capewell -dvorak</i>	54 %	42 %	12 %

Para observar el comportamiento de la distribución de teclado para los lenguajes náhuatl - español obtenida por el algoritmo genético, se realiza una comparación contra otras distribuciones teclado existentes como; *qwerty*, *dvorak*, *arensito*, *capewell – qwerty*, *capewell - dvorak*. Para hacer esta comparación, se usa el corpus de los lenguajes náhuatl y español comparados con los siguientes criterios:

- En la tabla 5.9 se muestra la distancia de las diferentes distribuciones de teclado y la distribución de teclado para los lenguajes náhuatl - español, la distancia de la distribución de teclado para los lenguajes náhuatl - español en comparación con las distancias de las otras distribuciones de la tabla 5.9 que es la menor distancia (3,009 metros). En este criterio mientras menor sea la distancia, más se reducen los movimientos innecesarios para todos los dedos de ambas manos.
- En la tabla 5.10 se muestra la carga de trabajo en las 3 filas principales del teclado (fila superior, fila origen y fila inferior). Como se puede observar en la tabla 5.10 la distribución de teclado para los lenguajes náhuatl - español tiene la mayor carga de trabajo (70.8%) en la fila origen. Para este criterio, al tener la mayor carga de trabajo en la fila origen (posición inicial de los dedos) reduce el número de desplazamientos

hacia arriba (desplazarse a la parte superior del teclado) y hacia abajo (desplazarse a la parte inferior del teclado) para todos los dedos de ambas manos.

- En la tabla 5.11 se muestra la carga de trabajo de cada uno de los dedos de ambas manos. Como se puede observar en la tabla 5.11, la distribución de teclado para los lenguajes náhuatl - español tiene un balance de forma jerárquica dando prioridad en primer lugar a los dedos índices (con la mayor carga de trabajo), en segundo lugar los dedos medios, en tercer lugar a los dedos anulares, y en cuarto lugar a los dedos meñiques (con la menor carga de trabajo). Este criterio tiene como objetivo proporcionar la carga de trabajo a cada dedo de acuerdo a la habilidad y fuerza que tiene cada uno de estos para evitar que alguno de los dedos este sobre cargado y provocar una futura lesión. Tal como se observa en la tabla 5.11, la distribución de teclado para los lenguajes náhuatl – español tiene en los dedos índices tienen la mayor carga de trabajo ya que son los dedos más fuertes y con mayor habilidad, los dedos meñiques tienen la menor carga de trabajo por ser los dedos con menor fuerza y habilidad de movimiento.
- La tabla 5.12 muestra el balance de carga de trabajo que tienen ambas manos. Así mismo indica la mano que tiene más carga de trabajo. El teclado *qwerty* tiene una carga de trabajo de 43% en la mano izquierda y una carga de trabajo de 53 % en la mano derecha, la diferencia entre estos porcentajes es de 7%. Esto quiere decir que la mano derecha esta sobre cargada (al tener el mayor porcentaje de carga de trabajo) un 7% más que la mano izquierda. Este criterio (balance de ambas manos) tiene como objetivo balancear la carga de trabajo en ambas partes del teclado para ambas manos, al aproximarse a los valores 50% de carga de trabajo en el lado izquierdo y 50% de carga de trabajo en el lado derecho se da paso a un flujo continuo de digitación, se evitan futuras lesiones en cualquiera de las muñecas de ambas manos y se evita la fatiga causada por sobre carga en alguna de las dos manos. Con esta tabla 5.12 se demuestra que la distribución de teclado para los lenguajes náhuatl - español está balanceado de forma optimizada en comparación con las demás distribuciones de la tabla ya que la carga en ambas manos tienen una diferencia mínima indicando que la mano izquierda está un 4% más sobrecargada que la mano derecha.

Tabla 5.9 Distancia total recorrida en metros.

Distancia recorrida (corpus en lenguajes náhuatl y español)	
Teclado	Distancia
Teclado propuesto (náhuatl – español)	3009 m
<i>qwerty</i>	5630 m
<i>dvorak</i>	3583 m
<i>arensito</i>	3191 m
<i>capewell - qwerty</i>	3484 m
<i>capewell - dvorak</i>	3279 m

Tabla 5.10 Carga de trabajo por las 3 filas principales del teclado.

Carga de trabajo por fila (corpus en lenguajes náhuatl y español)			
Teclado	Filas superior	Fila origen	Fila inferior
Teclado propuesto (náhuatl – español)	15.7 %	70.8 %	13.6 %
<i>qwerty</i>	45.6 %	34.4 %	19.5 %
<i>dvorak</i>	18.8 %	68.3 %	12.3 %
<i>arensito</i>	20.6 %	64 %	14.8 %
<i>capewell - qwerty</i>	20.7 %	61.1 %	17.6 %
<i>capewell - dvorak</i>	15.2 %	68.3 %	15.9 %

Tabla 5.11 Carga de trabajo por cada dedo de ambas manos.

Carga de trabajo sobre los dedos de ambas manos (corpus en lenguajes náhuatl y español)								
Teclado	Dedos mano izquierda				Dedos mano derecha			
	Meñique	Anular	Medio	Índice	Índice	Medio	Anular	Meñique
Teclado propuesto (náhuatl – español)	6 %	11 %	13 %	19 %	15 %	14 %	10 %	6 %
<i>qwerty</i>	14 %	6 %	13 %	10 %	20 %	16 %	14 %	3 %
<i>dvorak</i>	13 %	10 %	10 %	22 %	9 %	9 %	10 %	12 %
<i>arensito</i>	14 %	8 %	11 %	12 %	10 %	16 %	12 %	13 %
<i>capewell - qwerty</i>	14 %	6 %	13 %	10%	20 %	13 %	14 %	7 %
<i>capewell - dvorak</i>	8 %	15 %	12 %	19 %	15 %	6 %	10 %	10 %

Tabla 5.12 Carga de trabajo en ambas manos.

Carga de trabajo por mano (corpus en lenguajes náhuatl y español)			
Teclado	Mano izquierda	Mano derecha	Diferencia de carga en cada mano
Teclado propuesto (náhuatl – español)	49 %	45 %	4 %
<i>qwerty</i>	43 %	53 %	7 %
<i>dvorak</i>	55 %	40 %	15 %
<i>arensito</i>	45 %	51 %	6 %
<i>capewell - qwerty</i>	43 %	54 %	11 %
<i>capewell -dvorak</i>	54 %	41 %	13 %

Para observar el comportamiento de la distribución de teclado para los lenguajes maya - español obtenida por el algoritmo genético, se realiza una comparación contra otras distribuciones de teclado existentes como; *qwerty*, *dvorak*, *arensito*, *capewell – qwerty*, *capewell – dvorak*. Para hacer esta comparación, se usa el corpus de los lenguajes maya y español comparados con los siguientes criterios:

- En la tabla 5.13 se muestra la distancia de las diferentes distribuciones de teclado y la distribución de teclado para los lenguajes maya - español, la distancia de la distribución de teclado para los lenguajes maya - español en comparación con las distancias de las otras distribuciones de la tabla 5.13 es la menor distancia (3,001 metros). Con este criterio, entre menor sea la distancia, se reducen los movimientos innecesarios para todos los dedos de ambas manos.
- En la tabla 5.14 se muestra la carga de trabajo en las 3 filas principales del teclado (fila superior, fila origen y fila inferior). Como se puede observar en la tabla (5.14) la distribución de teclado para los lenguajes maya - español tiene la mayor carga de trabajo (67.6 %) en la fila origen. Este criterio, al tener la mayor carga de trabajo en la fila origen (posición inicial de los dedos) reduce el número de desplazamientos hacia arriba (desplazarse a la parte superior del teclado) y hacia abajo (desplazarse a la parte inferior del teclado) para todos los dedos de ambas manos.
- En la tabla 5.15 se muestra la carga de trabajo de cada uno de los dedos de ambas manos. Como se puede observar en la tabla (5.15) la distribución de teclado para los lenguajes maya - español tiene un balance de forma jerárquica dando prioridad en primer lugar a los dedos índices (con la mayor carga de trabajo), en segundo lugar los dedos medios, en tercer lugar a los dedos anulares, y en cuarto lugar a los dedos meñiques (con la menor carga de trabajo). Este criterio tiene como objetivo proporcionar la carga de trabajo a cada dedo de acuerdo a la habilidad y fuerza que tiene cada uno de estos para evitar que alguno de los dedos este sobre cargado y provocar una futura lesión. Tal como se observa la distribución de teclado para los lenguajes maya – español tiene en los dedos índices tienen la mayor carga de trabajo ya que son los dedos más fuertes y con mayor habilidad, los dedos meñiques tienen la menor carga de trabajo por ser los dedos con menor fuerza y habilidad de movimiento.
- En la tabla 5.16 se muestra el balance de carga de trabajo que tienen ambas manos. Así mismo se indica la mano que tiene sobre carga de trabajo. El teclado *qwerty* tiene

una carga de trabajo de 44% en la mano izquierda y una carga de trabajo de 52% en la mano derecha, la diferencia entre estos porcentajes es de 8%, esto quiere decir que la mano derecha esta sobre cargada (al tener el mayor porcentaje de carga de trabajo) un 8% más que la mano izquierda. Para este criterio (balance de ambas manos) tiene como objetivo balancear la carga de trabajo en ambas partes del teclado para ambas manos, al aproximarse a los valores 50% de carga de trabajo en el lado izquierdo y 50% de carga de trabajo en el lado derecho se da paso a un flujo continuo de digitación, se evitan futuras lesiones en cualquiera de las muñecas de ambas manos y se evita la fatiga causada por sobre carga en alguna de las dos manos. La tabla 5.16 muestra que la distribución de teclado para los lenguajes maya - español está balanceado de forma optimizada en comparación con las demás distribuciones de la tabla ya que la carga en ambas manos tienen una diferencia mínima indicando que la mano derecha está un 4% más sobrecargada que la mano izquierda.

Tabla 5.13 Distancia total recorrida en metros.

Distancia recorrida (corpus en lenguajes maya y español)	
Teclado	Distancia
Teclado propuesto (maya - español)	3001 m
qwerty	4796 m
dvorak	3127 m
arensito	3196 m
capewell - qwerty	3322 m
capewell - dvorak	3195 m

Tabla 5.14 Carga de trabajo por las 3 filas principales del teclado.

Carga de trabajo por fila (corpus en lenguajes náhuatl y español)			
Teclado	Filas superior	Fila origen	Fila inferior
Teclado propuesto (maya – español)	17.7 %	67.6 %	13.8 %
<i>qwerty</i>	43.6 %	36.7 %	18.7 %
<i>dvorak</i>	19.4 %	66.2 %	13.9 %
<i>arensito</i>	22.6 %	62 %	14.8 %
<i>capewell - qwerty</i>	21.8 %	60.2 %	17 %
<i>capewell - dvorak</i>	15.3 %	66.2 %	17.9 %

Tabla 5.15 Carga de trabajo por cada dedo de ambas manos.

Carga de trabajo sobre los dedos de ambas manos (corpus en lenguajes maya y español)								
Teclado	Dedos mano izquierda				Dedos mano derecha			
	Meñique	Anular	Medio	Índice	Índice	Medio	Anular	Meñique
Teclado propuesto (maya – español)	4 %	11 %	15 %	16 %	17 %	15 %	11%	7 %
<i>qwerty</i>	14 %	6 %	15 %	9 %	22 %	12 %	14 %	4 %
<i>dvorak</i>	14 %	10 %	13 %	19 %	10 %	8 %	9 %	12 %
<i>arensito</i>	14 %	7 %	13 %	12 %	11 %	14 %	10 %	16 %
<i>capewell - qwerty</i>	14 %	6 %	15 %	9 %	22 %	8 %	14 %	8 %
<i>capewell - dvorak</i>	8 %	16 %	14 %	16 %	15 %	6 %	11 %	10 %

Tabla 5.16 Carga de trabajo en ambas manos.

Carga de trabajo por mano (corpus en lenguajes maya y español)			
Teclado	Mano izquierda	Mano derecha	Diferencia de carga en cada mano
Teclado propuesto (maya – español)	46 %	50 %	4 %
<i>qwerty</i>	44 %	52 %	8 %
<i>dvorak</i>	56 %	39 %	17 %
<i>arensito</i>	46 %	51 %	5 %
<i>capewell - qwerty</i>	44 %	52 %	8 %
<i>capewell -dvorak</i>	54 %	42 %	12 %

5.4 Curva de aprendizaje en nuevas distribuciones de teclado

Después de revisar los trabajos relacionados con la optimización de distribuciones de teclado, ninguno de estos trabajos hace mención sobre la curva de aprendizaje para sus distribuciones de teclado propuestas. Esto hace que otros investigadores o lectores no le tomen un mayor interés a estos trabajos ya que no se da una explicación de qué tan viable podría ser el aprender a digitar sobre una nueva distribución de teclado, qué tiempo se tardarían en aprender la nueva ubicación de teclas propuesta qué tan fácil sería o si habría alguna dificultad en aprender una nueva distribución de teclado completamente diferente a la distribución de teclado estándar *qwerty*. Para poder abordar este problema, en esta tesis se realizan pruebas mostrando una curva de aprendizaje promedio. En estas pruebas demuestra que es más fácil y rápido aprender un teclado optimizado que cuenta con criterios ergonómicos ya que la habilidad de los dedos y posiciones de las manos se adaptan más rápido y de mejor forma a los criterios ergonómicos, entre más optimizados sean los criterios ergonómicos la comodidad al digitar es mayor y la curva de aprendizaje se vuelve más rápida.

Para poder llevar acabo la demostración de una curva de aprendizaje con una nueva distribución de teclado, se realiza de la siguiente manera:

1. Se eligen dos distribuciones de teclado en las cuales nunca se ha escrito, el objetivo de tomar dos distribuciones de teclados diferentes, tomar una distribución de teclado no optimizada (una de las distribuciones de teclado de las cuales son creadas aleatoriamente en la población inicial del algoritmo genético) y sin criterios económicos, y la otra distribución de teclado es la distribución optimizada con criterios ergonómicos obtenida por el algoritmo genético para las lenguas náhuatl y español.
2. Posteriormente se implementa una herramienta para poder medir el tiempo (en mili segundos) de digitación. En esta herramienta determina el tiempo que se genera al escribir un determinado número de caracteres.
 - o De acuerdo con el artículo de (Ballesteros, 1999) donde se menciona que la memoria a corto plazo en promedio puede retener 7 ± 2 elementos (esto quiere decir que como mínimo se pueden recordar 5 y como máximo 9 elementos). En esta prueba se toma como referencia el mínimo de elementos (5 caracteres).
3. Teniendo definido el número de elementos (caracteres) para poder medir el tiempo, se crea un corpus con 5,000 caracteres; 2,500 caracteres en lenguaje náhuatl y 2,500 caracteres en lenguaje español ya que se aprende digitar en dos distribuciones de teclado diferentes pero que se enfocan en estos dos lenguajes: náhuatl y español.
4. Teniendo el corpus con los 5,000 caracteres se divide en cada 5 caracteres para tener 1,000 iteraciones (500 iteraciones en lenguaje náhuatl y 500 iteraciones en lenguaje español) tal como se muestra en la tabla 5.17, en la lista de iteraciones para el lenguaje náhuatl, de la iteración 1 a la iteración 10 se tienen las palabras "ICA IPAN IN TONALTIN AMOCA HUELOQUINEQUIA QUIHUICAZ IN IPIL" separadas por cada 5 caracteres (omitiendo los espacios en blanco). en la lista de iteraciones para el lenguaje español, de la iteración 1 a la iteración 10 se tienen las palabras "HOLA COMO ESTAS SOLO PASABA A SALUDAR Y SABER QUE ESTAS BIEN" separadas por cada 5 caracteres (omitiendo los espacios en blanco).
5. La digitación inicia cuando se pulsa la tecla "ENTER" del teclado, con esto se indica cada cambio de iteración, de este modo la herramienta guarda el lapso de tiempo en milisegundos cada que se escribe una nueva iteración (5 caracteres). La digitación termina cuando se completan las 1,000 iteraciones (5,000 caracteres digitados en el teclado) y se grafican los valores obtenidos en tiempo para cada iteración.

Tabla 5.17 Iteraciones con cada 5 caracteres en los lenguajes náhuatl y español.

ITERACIONES EN NÁHUATL		ITERACIONES EN ESPAÑOL	
1	ICAIP	1	HOLAC
2	NINTO	2	OMOES
3	NALTI	3	TASSO
4	NAMOC	4	LOPAS
5	AHULE	5	ABAAS
6	LOQUI	6	SALUD
7	NEQUI	7	ARYSA
8	AQUIH	8	BERQU
9	UICAZ	9	EESTA
10	INPIL	10	SBIEN
.		.	
.		.	
.		.	
2500	CEMIL	2500	SALUD

La figura 5.6 muestra la gráfica de los lapsos de tiempo en milisegundos generados al escribir 5,000 caracteres (1,000 iteraciones) en una distribución de teclado no optimizada (distribución creada aleatoriamente por el algoritmo genético en la población inicial). Esta gráfica representada en la figura 5.6 muestra la curva de aprendizaje para una distribución de teclado no optimizada y sin criterios ergonómicos. En el eje X se tiene el número de iteraciones realizadas (1,000 iteraciones), y en el eje Y se tiene el tiempo en milisegundos.

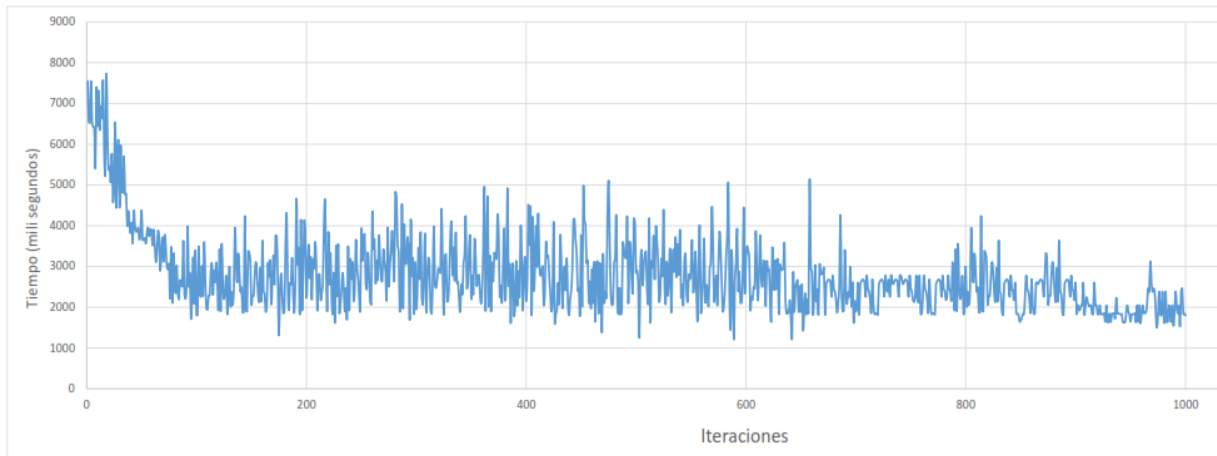


Figura 5.6 Curva de aprendizaje de un teclado no optimizado: experimento 5.

La figura 5.7 muestra la gráfica de los lapsos de tiempo en milisegundos generados al escribir 5,000 caracteres (1,000 iteraciones) en una distribución de teclado optimizada obtenida por el algoritmo genético. Esta gráfica representada en la figura 5.7 muestra la curva de aprendizaje para una distribución de teclado optimizada y con criterios ergonómicos.

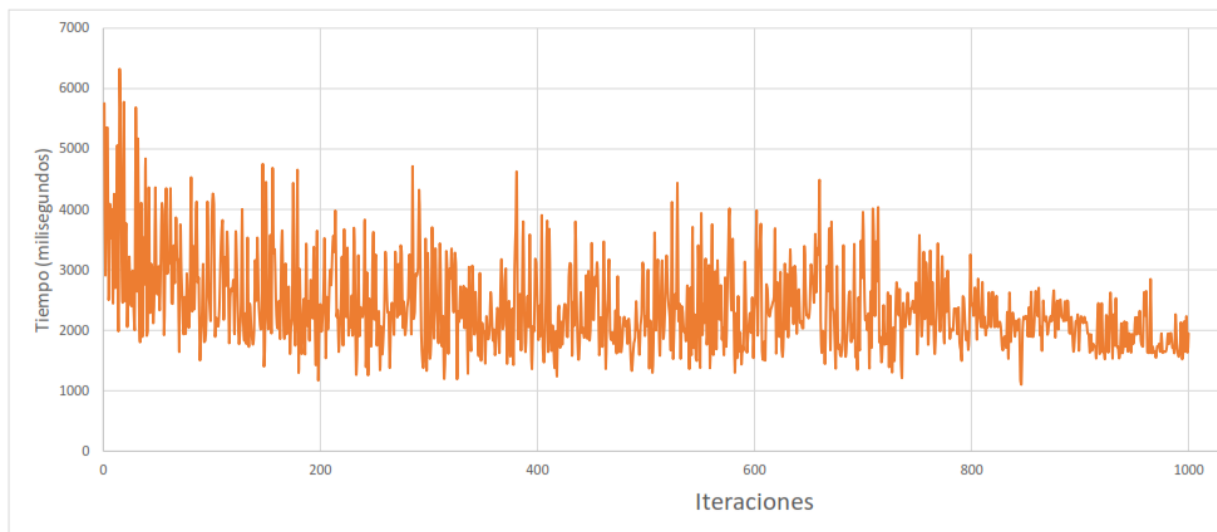


Figura 5.7 Curva de aprendizaje de un teclado optimizado: experimento 6.

En la figura 5.8, se observa de color azul la curva de aprendizaje con una distribución de teclado no optimizada y sin criterios ergonómicos, de color marrón se representa la curva de aprendizaje con una distribución de teclado optimizada (obtenida por el algoritmo genético) y con criterios ergonómicos. Tal como se puede observar, la curva de aprendizaje en una distribución de teclado no optimizada, presenta dificultades y lleva un poco más de tiempo aprender en comparación con una distribución de teclado optimizada que desde las primeras iteraciones de puede notar el ahorro de tiempo en aprender.

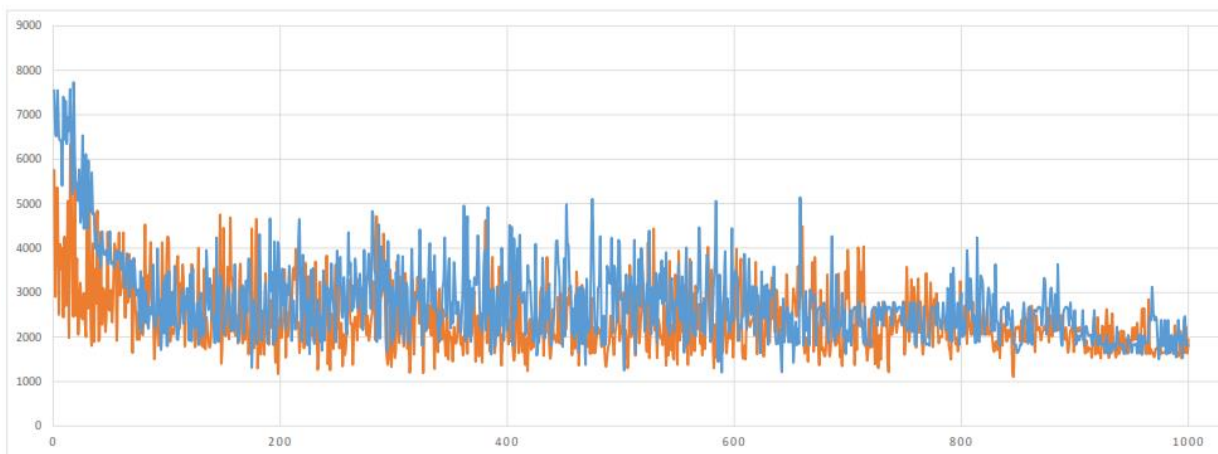


Figura 5.8 Comparación de una curva de aprendizaje en color azul con un teclado no optimizado (sin criterios ergonómicos) y una curva de aprendizaje en color naranja con un teclado optimizado (con criterios ergonómicos) .

5.5 Resumen del capítulo

En este capítulo, se mostraron todos los resultados y experimentos realizados para poder evaluar a cada una de las distribuciones de teclado y hacer una comparación con otras distribuciones para demostrar que tan óptimas son las distribuciones de teclado obtenidas por el algoritmo genético. Así mismo se dio a conocer la curva de aprendizaje promedio para poder aprender las nuevas distribuciones de teclado.



6. Conclusiones

Las personas hablantes de las lenguas indígenas náhuatl y maya, al utilizar un dispositivo electrónico en su mayoría de veces utilizan el teclado *qwerty* para poder ingresar texto de su propia lengua indígena náhuatl o maya. Por medio de las 4 distribuciones de teclado se brinda el apoyo a las lenguas indígenas náhuatl y maya. El objetivo es que las personas hablantes de estas lenguas cuenten con un teclado ergonómico y fácil de aprender para que se puedan mantener en comunicación, para que puedan transmitir información en estas lenguas a través de medios electrónicos y que sea de ayuda para otros hablantes de dichas lenguas (náhuatl y maya).

Las 4 distribuciones de teclado para los lenguajes (1) náhuatl, (2) maya, (3) náhuatl – español, (4) maya – español, cuentan con criterios ergonómicos como; minimización de la distancia, carga de trabajo en las tres filas principales del teclado (fila superior, fila origen y fila inferior) dando prioridad a la fila origen del teclado (fila donde se posicionan inicialmente los dedos y no se realizan desplazamientos con los dedos de ambas manos), carga balanceada de forma jerárquica (carga asignada de acuerdo a la habilidad natural de movimiento y fuerza de cada dedo) en todos los dedos de ambas manos, carga balanceada en ambas manos con el objetivo de aproximar los valores de 50 % de carga de trabajo en la mano izquierda y 50 % de carga de trabajo en la mano derecha. Cada uno de estos criterios dan paso a la distribución de teclado de forma ergonómica y optimizada para cada lenguaje, el objetivo es contar con un teclado que se multicriterio y a su vez multilinguaje que ayude a minimizar el tiempo, reduzca la fatiga de digitación, reduzca los errores al digitar ocasionados por

posiciones incómodas, que se tenga un flujo continuo de la digitación al tener un arreglo de teclas optimizado, que ayude a prevenir lesiones ocasionadas por posiciones incómodas de las manos o sobrecarga de trabajo en alguna de estas, que sea una distribución de teclado fácil de aprender y en poco tiempo.

Tras realizar diferentes pruebas y haciendo comparaciones de las distribuciones de teclado optimizadas por el algoritmo genético con otras distribuciones de teclado existentes, se demostró que cada una de 4 distribuciones de teclado optimizada por el algoritmo genético cuenta con criterios ergonómicos de forma optimizada especialmente comparadas con el teclado estándar *qwerty*.

Para poder hacer uso de alguna de estas distribuciones en una computadora, se puede hacer de diferentes formas sin necesidad de fabricar los teclados propuestos, una de las formas por la cual se puede hacer uso de una de las 4 distribuciones de teclado optimizadas por el algoritmo genético es, re-etiquetar cada una de las teclas de acuerdo a la distribución que se desea utilizar y por medio de un software cambiar la entrada digital que tienen el teclado, de este modo se puede hacer uso de alguna de las 4 distribuciones de teclado optimizadas por el algoritmo genético. La segunda forma es, desarmar un teclado y re-ordenar las teclas de acuerdo a la distribución de teclado que se desea utilizar, y así mismo la salida digital de la tecla se realiza mediante un software.

6.1 Trabajo futuro

El trabajo que se pretende realizar para dar continuidad a esta tesis, es agregar el sonido de cada letra en el teclado de acuerdo a cada lenguaje. El objetivo de agregar el sonido en un teclado, es mejorar la curva de aprendizaje, es decir, si una distribución de teclado que cuenta con criterios ergonómico es más fácil de aprender en comparación de una distribución sin criterios ergonómicos, entonces al escuchar el sonidos de cada tecla se puede reforzar el aprendizaje y esto da paso a un aprendizaje mucho más rápido.

Una de las ventajas principales al agregar el sonido de las letras en las teclas del teclado, es reforzar el aprendizaje de los lenguajes náhuatl, maya y español, ya que se memoriza el sonido

Conclusiones

de cada una de las letras de cada lenguaje. Así mismo sirve como método de aprendizaje para personas que no saben escribir alguno de estos lenguajes.

Dentro del trabajo futuro también se pretenden crear lecciones para aprender alguno de los tres lenguajes: náhuatl, maya y español; por medio de memorización del sonido de las letras agregadas en el teclado y al mismo tiempo se aprende a escribir sobre las nuevas distribuciones de teclado con criterios ergonómicos.

Referencias

- Ballesteros, S. (1999). Memoria humana: investigación y teoría. *Psicothema*, 11, 705-723. Retrieved from redalyc: <http://www.redalyc.org/pdf/727/72711401.pdf>
- Bi, X. (2012). Multilingual Touchscreen Keyboard Design and Optimization. *Human-Computer Interaction*, 34.
- Bodenhofer, U. (2003). Genetic Algorithms: Theory and Applications. *Fuzzy Logic Laboratorium Linz-Hagenberg*, 17-65.
- Brewbaker, C. R. (2008). Optimizing stylus keyboard layouts with a genetic algorithm: customization and internationalization. *Dept. of Computer Science Iowa State University*, 1-8.
- Capewell, M. (2005). *Keyboard Comparison Applet*. Retrieved from Alternative Keyboard Layouts: <http://www.michaelcapewell.com/projects/keyboard/index.htm>
- Capewell, Michael. (2005). *Alternative Keyboard Layouts*. Retrieved from Michael Capewell.com: http://www.michaelcapewell.com/projects/keyboard/layout_capewell.htm
- Centella, R. N. (2010). La evolución de una especie tecnológica. *Museo nacional de ciencia y tecnología*, 130.
- Coello, C. A. (2015). Introducción a la computación evolutiva. *CINVESTAV-IPN, Departamento de Ingeniería Eléctrica, Sección de Computación. México, DF.*, 7-14.
- Cook, A. M. (2014). *Assistive technologies*. Canada: Elsevier Health Sciences.
- Cultura, S. d. (2018, Febrero 21). Gob. Retrieved from Gob: <https://www.gob.mx/cultura/articulos/lenguas-indigenas?idiom=es>
- Deb, K. (2003). Design of an Optimal Hindi Keyboard for Convenient and Efficient Use. *Indian Institute of Technology, Kanpur*.
- Eggers, J. (2002). An Ant Colony Optimization Algorithm for the Optimization of the Keyboard Arrangement Problem. *European Journal of Operational Research*, pp. 672-686.
- Fernández, I. (2013). LA MÁQUINA DE ESCRIBIR. 2º curso del Programa para , 10.
- García, J. L. (2012). Tecnología: la historia del teclado convencional de las máquinas de escribir. *Maestría en gestión de proyectos educativos*, 1-7.
- Gen, M. (2007). *genetic algorithms & engineering design*. Canada: John Wiley & Sons, Inc.
- Gestal, M. (2010). *Introducción a los Algoritmos Genéticos y la Programación Genética*. Coruña: Universidade da Coruña.
- Gómez, J. A. (2004). *Principios de ergonomía*. Colombia: U. Jorge Tadeo Lozano.

-
- Hempstalk, K. (2006). The Great Keyboard Debate: QWERTY versus DVORAK. *University of Waikato*.
- Hosny, M. I. (2015). A Genetic Algorithm Approach for Optimizing a Single-Finger Arabic Keyboard Layout. In M. I. Hosny, *Studies in Computational Intelligence* (pp. 261-277). Saudi Arabia: Springer International Publishing Switzerland.
- Kumara Sastry, D. G. (2005). Genetic Algorithms. In D. G. Kumara Sastry, *Search Methodologies* (pp. 97-125). USA: Springer US.
- Melanie, M. (1998). *An introduction to genetic algorithms*. England: A Bradford Book.
- Myers, R. (2005). Genetic algorithms for structural editing. *departament of computer science*, 160-168.
- Resano, R. P. (2009). Smith Premier No. 10, utilidad y belleza en el Museo Cerralbo. *Pieza del mes*.
- Samimi, N. (2012). Optimization of Farsi Letter Arrangement on Keyboard by Simulated Annealing and Genetic Algorithms. *Department of Computer, Omidiyeh Branch, Islamic Azad University*.
- Sánchez, J. L. (2010). Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. *Catálogo de las Lenguas Indígenas Nacionales*, 67-301.
- Sebag, M. (2005). An Induction-based Control for Genetic Algorithms. In M. Sebag, *Machine Learning: ECML-95* (pp. 352-355). Europa: Springer.
- Sivanandam, S. (2008). *Introduction to Genetic Algorithms*. India: Springer.
- Truyol, A. P. (2015). Algoritmos genéticos. *Universidad Carlos III*. Retrieved from <http://www.it.uc3m.es/jvillena/irc/practicas/06-07/05.pdf>
- UNIVERSAL, E. (2018, 10 2). *EL UNIVERSAL*. Retrieved from *EL UNIVERSAL*: <http://www.eluniversal.com.mx/cultura/expertos-advierten-riesgo-de-desaparicion-de-lenguas-indigenas-en-mexico>
- Valle, J. (2011). *Scientific hands of comfort*. Retrieved from *CARPALX*: <http://mkweb.bcgsc.ca/carpalx>
- Vincent Cicirello, R. C. (2013). Profiling the Distance Characteristics of Mutation Operators for Permutation-Based Genetic Algorithms. *Proceedings of the Twenty-Sixth International Florida Artificial Intelligence Research Society Conference* (pp. 46-51). Florida: Association for the Advancement of Artificial.
- Walker, C. P. (2003). Evolving a More Optimal Keyboard. *Course project: Introduction to evolutionary computation, Missouri University of Science & Technology*, 5-8.
