



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escola Tècnica Superior d'Enginyeria Informàtica

Cerca de Patrons Melòdics a Grans Col·leccions de  
Imatges de Partitures Manuscrites Històriques

Treball Fi de Grau

Grau en Enginyeria Informàtica

AUTOR/A: Menàrguez Box, Aitana

Tutor/a: Vidal Ruiz, Enrique

Cotutor/a: Toselli, Alejandro Héctor

Director/a Experimental: VILLARREAL RUIZ, MANUEL

CURS ACADÈMIC: 2022/2023



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica  
Superior d'Enginyeria  
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica  
Universitat Politècnica de València

# **Cerca de patrons melòdics a grans col·leccions d'imatges de partitures manuscrites històriques**

**TREBALL FI DE GRAU**

Grau en Enginyeria Informàtica

*Autora:* Menàrguez Box, Aitana

*Tutor:* Vidal Ruiz, Enrique

*Co-tutor:* Toselli, Alejandro Héctor

*Director experimental:* Villarreal Ruiz, Manuel

Curs 2022-2023



# Resum

La cerca d'informació musical a grans sèries d'imatges de partitures antigues manuscrites és un problema de gran interès per persones historiadores, musicòlogues, gestores d'arxius i públic en general.

Utilitzant diverses tècniques d'indexació probabilística, s'ha desenvolupat un motor de cerca que ofereix solucions parcials a aquest problema. Mitjançant aquestes, les consultes de cerca s'expressen com seqüències de notes definides per símbols *geomètrics*. Un exemple seria «c4 &20& [L3 S2 L2 L3]», que representa una seqüència de notes el pentagrama de les quals comença en clau de do en quarta (c4) i les notes de la seqüència estan en la línia 3, l'espai 2, la línia 2 i la línia 3, respectivament.

Una versió *on-line* d'aquest buscador és públicament accessible des de <http://prhlt-carabela.prhlt.upv.es/musica>.

Aquesta forma d'expressar una consulta de cerca permet trobar informació que abans era pràcticament impossible de localitzar. Malgrat això, aquesta solució no pot considerar-se satisfactòria per al seu ús per especialistes en història de la música. Per tal que així fora, seria desitjable que les consultes pogueren expressar-se en notació musical, no *geomètrica*. En l'anterior exemple, l'objectiu seria representar-la com «la sol fa la» (o «A G F A» en notació musical anglosaxona). A més de poder expressar-la també musicalment, a través d'un teclat (simulat o real).

Encara que, en principi, la transformació requerida sembla ximple, cal tindre en compte que a les partitures històriques considerades, s'usen amb gran freqüència dos tipus de clau i no existeix una posició per omisió per a cap d'ells. D'aquesta forma, una consulta com «la sol fa la» pot aparèixer escrita de diverses formes depenent de la clau i la seua posició. Aquesta consulta en notació musical s'ha de transformar en una complexa consulta booleana AND/OR *geomètrica* en la qual la majoria dels elements són, al mateix temps, seqüències. Concretament, la consulta musical [A3 G3 F3 A3] es transformaria en «(c2 &20& [L1 S0 L0 L1 ]) || (c3 &20& [L2 S1 L1 L2 ]) || (c4 &20& [L3 S2 L2 L3 ]) || (c5 &20& [L4 S3 L3 L4 ]) || (f1 &20& [L2 S1 L1 L2 ]) || (f2 &20& [L3 S2 L2 L3 ]) || (f3 &20& [L4 S3 L3 L4 ]) || (f4 &20& [L5 S4 L4 L5 ])».

El projecte que es proposa més endavant, consisteix a donar solucions adequades als problemes anteriors i implementar tals solucions al buscador d'accés públic anteriorment citat, <http://prhlt-carabela.prhlt.upv.es/musica>.

Així mateix, abans de la implementació definitiva, el sistema desenvolupat haurà de ser avaluat empíricament, comparant les prestacions d'aquest mode de cerca amb resultats obtinguts per a un buscador basat en consultes purament *geomètriques*.

Aquest projecte es complementa amb un altre, també relacionat amb cerca de patrons melòdics en partitures manuscrites antigues, al qual l'objecte és la renderització musical (acústica) dels resultats de les consultes.

Amb la combinació d'ambdós projectes, les col·leccions de partitures antigues indexades probabilísticament podran consultar-se mitjançant notes musicals i/o un teclat (simulat o real). Els resultats (imatges de les partitures trobades) no només es podran visualitzar, sinó també escoltar en distints modes de renderització acústica.

**Paraules clau:** reconeixement d'imatges de partitures antigues, cerca d'informació musical, indexació probabilística d'imatges (de text i música)

---

# Resumen

La cerca de información musical en grandes series de imágenes de partituras antiguas manuscritas es un problema de gran interés para historiadores, musicólogos, gestores de archivos y público en general.

Utilizando diversas técnicas de indexación probabilística, se ha desarrollado un buscador que ofrece soluciones parciales a dicho problema. Mediante éstas, las consultas de cerca se expresan como secuencias de notas definidas por símbolos *geométricos*. Un ejemplo sería  $c4 \ \&20\& \ [L3 \ S2 \ L2 \ L3]$ , que representa una secuencia de notas tales que el pentagrama empieza en clave de do en cuarta (c4) y las notas de la secuencia están en la línea 3, el espacio 2, la línea 2 y la línea 3, respectivamente.

Una versión *on-line* de este buscador es públicamente accesible desde

<http://prhlt-carabela.prhlt.upv.es/musica>.

Esta forma de expresar una consulta de cerca permite encontrar información que antes era prácticamente imposible de localizar. Sin embargo, esta solución no puede considerarse satisfactoria para su uso por especialistas en historia de la música. Para este fin, sería deseable que las consultas pudieran expresarse en notación musical, no *geométrica*. En el anterior ejemplo, el objetivo sería representar la consulta como «la sol fa la» (o «A G F A» en notación musical angolsajona). Además de poder expresarla también musicalmente, mediante un teclado (simulado o real).

Aunque, en principio, la transformación requerida parece simple, hay que tener en cuenta que en las partituras históricas consideradas se usan con gran frecuencia dos tipos de clave y no hay una posición por omisión para ninguno de ellos. Así, una consulta como «la sol fa la» puede aparecer escrita de múltiples formas dependiendo de la clave y de su posición. Esta consulta en notación musical se debe convertir en una compleja consulta booleana AND/OR *geométrica* en la que la mayoría de los elementos son, a su vez, secuencias. Concretamente, la consulta musical  $[A3 \ G3 \ F3 \ A3]$  se debería transformar en «(c2 &20& [L1 S0 L0 L1 ]) || (c3 &20& [L2 S1 L1 L2 ]) || (c4 &20& [L3 S2 L2 L3 ]) || (c5 &20& [L4 S3 L3 L4 ]) || (f1 &20& [L2 S1 L1 L2 ]) || (f2 &20& [L3 S2 L2 L3 ]) || (f3 &20& [L4 S3 L3 L4 ]) || (f4 &20& [L5 S4 L4 L5 ])».

El proyecto que aquí se propone consiste en dar soluciones adecuadas a estos problemas e implementar dichas soluciones en el buscador de acceso público anteriormente citado, <http://prhlt-carabela.prhlt.upv.es/musica>.

Asimismo, antes de la implementación definitiva, el sistema desarrollado deberá ser evaluado empíricamente, comparando las prestaciones de este modo de cerca con resultados obtenidos para un buscador basado en consultas puramente *geométricas*.

Este proyecto se complementa con otro, también relacionado con cerca de patrones melódicos en partituras manuscritas antiguas, en el que el objeto es la renderización musical (acústica) de los resultados de las consultas.

Con la combinación de ambos proyectos, las colecciones de partituras antiguas indexadas probabilísticamente podrán consultarse mediante notas musicales y/o un teclado (simulado o real). Los resultados (imágenes de las partituras encontradas) no solo se podrán visualizar, sino también oír en distintos modos de renderización acústica.

**Palabras clave:** reconocimiento de imágenes de partituras antiguas, búsqueda de información musical, indexación probabilística de imágenes (de texto y música)

---

# Abstract

Search for musical information inside large series of images of historical handwritten scores is a problem of great interest for historians, musicologists, archive managers and general public.

Using Probabilistic Indexing techniques a search engine which offers partial solutions to this problem has been developed. With this techniques search queries are expressed as sequences of "notes" defined by *geometric* symbols such as  $c4 \ \&20\& \ [L3 \ S2 \ L2 \ L3]$ , denoting a sequence of notes in which the staff starts in C clef on the fourth line (c4) and the notes of the sequence lay on line 3, space 2, line 2 and line 3, respectively.

An on-line version of this search engine is publicly accessible through <http://prhlt-carabela.prhlt.upv.es/musica>.

This way of expressing a search query allows to find information that was previously practically impossible to locate. However, this solution cannot be considered satisfactory for being used by music history specialists. To this end, it would be desirable that queries could be expressed in musical rather than *geometric* notation. In the previous example the objective would be to express the query as «A G F A» (or «la sol fa la» using solmization), apart from being able to express it musically with a (simulated or real) keyboard.

Even though in principle the required transformation seems simple, it must be taken into account that inside the historical scores considered, there exist two types of clef which are used extremely frequently and it does not exist a default position for any of them. So a query like «A G F A» can appear written in multiple ways depending on the key and its position. Specifically, this query must be converted into a complex AND/OR Boolean *geometric* query in which most of the elements inside it are themselves sequences. Specifically, the musical query [A3 G3 F3 A3] is transformed into «(c2 &20& [L1 S0 L0 L1 ]) || (c3 &20& [L2 S1 L1 L2 ]) || (c4 &20& [L3 S2 L2 L3 ]) || (c5 &20& [L4 S3 L3 L4 ]) || (f1 &20& [L2 S1 L1 L2 ]) || (f2 &20& [L3 S2 L2 L3 ]) || (f3 &20& [L4 S3 L3 L4 ]) || (f4 &20& [L5 S4 L4 L5 ])

The proposed project in this document consists of providing adequate solutions to these problems and implementing these solutions in the aforementioned publicly accessible search engine, <http://prhlt-carabela.prhlt.upv.es/musica>.

Likewise, before the final implementation, the developed system must be empirically evaluated, comparing the performance of this search mode with previous results obtained for a search engine based on purely *geometric* queries.

This project is complemented by another, also related to the search for melodic patterns in old handwritten scores, in which the object is the musical (acoustic) rendering of the query results.

With the combination of both projects, it will be possible to query probabilistically indexed old collections of scores images by means of musical notes and/or a keyboard (simulated or real). And the results (images of found scores) will be not only displayed visually but also heard in different modes of acoustic rendering.

**Key words:** historical sheet music image recognition, search for musical information, probabilistic indexing of (text and music) images

---



# Índex

---

<b>Índex</b>	<b>vii</b>
<b>Índex de figures</b>	<b>ix</b>
<b>Índex de taules</b>	<b>x</b>

---

<b>1 Introducció</b>	<b>1</b>
1.1 Motivació	1
1.2 Objectius	3
1.3 Estructura de la memòria	3
<b>2 Estat de l'art</b>	<b>5</b>
2.1 Indexació probabilística: PrIx	5
2.1.1 PrIx per a símbols musicals individuals	6
2.1.2 PrIx per a seqüències de símbols musicals	6
2.2 Demostrador del motor per a seqüències musicals del PRHLT	7
2.2.1 Esquema d'implementació	7
2.2.2 Arquitectura del demostrador	7
2.2.3 Client web i consultes	8
2.2.4 Cas d'estudi: manuscrit VORAU-253	9
<b>3 Tecnologies emprades</b>	<b>11</b>
3.1 Servidor web: ús de Docker	11
3.1.1 Introducció	11
3.1.2 Funcionament	11
3.2 ferramentes de programació: llibreries i APIs	12
3.2.1 NumPy	13
3.2.2 Web MIDI API	13
3.2.3 Altres ferramentes	13
3.3 Manteniment del treball	13
3.4 Protocol MIDI	13
3.4.1 Missatges MIDI	14
<b>4 Anàlisi del problema</b>	<b>17</b>
4.1 Ús del demostrador	17
4.2 Introducció de dades musicals	17
4.3 Gestió de dades musicals	18
4.4 Tipus de consultes al buscador	18
4.5 Anàlisi del <i>ground truth</i>	20
4.6 Notes addicionals	21
<b>5 Disseny de la solució</b>	<b>23</b>
<b>6 Desenvolupament de la solució</b>	<b>25</b>
6.1 El piano a la web	25
6.1.1 Teclat	26
6.1.2 Barra de cerca	26
6.1.3 Opcions de cerca	27
6.1.4 Entrada MIDI	27



6.2	La traducció	27
6.2.1	Plantejament del problema	27
6.2.2	Primer pas: traducció geomètrica	28
6.2.3	Segon pas: traducció a <i>query</i> lògica	29
6.2.4	Traducció sense tonalitat	30
6.2.5	Introducció de la traducció a la web	31
<b>7</b>	<b>Proves i avaluació</b>	<b>33</b>
7.1	Obtenció de melodies	33
7.2	Processament de les dades	34
7.2.1	Extracció de seqüències musicals	35
7.3	Obtenció de consultes	36
7.4	Tipus de detecció de les seqüències musicals	37
7.4.1	Posició lògica: LPS ( <i>Logical-position-based Spotting</i> )	37
7.4.2	Posició lògica restringida: RLPS ( <i>Restricted-logical-position-based Spotting</i> )	37
7.4.3	Posició geomètrica: GPS ( <i>Geometrical-position-based Spotting</i> )	37
7.4.4	Comparativa de les alternatives	38
7.5	Mesurant els resultats	38
7.5.1	Corba R-P	39
7.5.2	Resultats de les mètriques	39
<b>8</b>	<b>Conclusions</b>	<b>41</b>
8.1	Assoliment dels objectius	41
8.2	Sobre les proves i l'avaluació	42
8.3	Línies de treball futures	43
	<b>Bibliografia</b>	<b>45</b>
<hr/>		
	Apèndixs	
<b>A</b>	<b>Codi</b>	<b>49</b>
A.1	Codi CSS	49
A.1.1	Tecles del piano virtual	49
A.1.2	Botó de gravació	50
A.2	Codi HMTL	50
A.2.1	Diàleg del piano	50
A.2.2	Notes musicals	51
A.3	Codi JavaScript	52
A.3.1	Dades	52
A.3.2	Funcions	53
A.3.3	Codi d'inicialització	56
A.4	Codi Python	57
A.4.1	Funcions de traducció	57
<b>B</b>	<b>Sobre teoria musical</b>	<b>61</b>
B.1	Unitats bàsiques: les notes	61
B.2	Escala cromàtica: el piano	61
B.3	Representació gràfica de les notes: claus i pentagrames	62
B.4	Relacions entre les notes: els intervals i transposicions	63

# Índex de figures

---

2.1	Esquema general del motor PRHLT . . . . .	7
2.2	Esquema d'indexació del motor PRHLT . . . . .	8
2.3	Arquitectura client-servidor utilitzada pel demostrador . . . . .	8
2.4	Exemple del manuscrit VORAU-253 . . . . .	9
3.1	Arquitectura de Docker . . . . .	12
4.1	Exemple de traducció sense tonalitat . . . . .	19
4.2	Exemple de traducció amb tonalitat . . . . .	19
4.3	Anàlisi de les notes al GT . . . . .	20
4.4	Anàlisi de les claus al GT . . . . .	21
5.1	Esquema simplificat del projecte . . . . .	23
6.1	Piano dins de la web . . . . .	25
6.2	Teclat del piano web . . . . .	26
6.3	Barra de cerca del piano web . . . . .	26
7.1	Exemple amb pentagrama de notació implícita . . . . .	34
7.2	Histograma de longituds . . . . .	35
7.3	Histograma de freqüències . . . . .	36
7.4	Transformació notació implícita a notació musical . . . . .	36
7.5	Detecció GPS contra detecció LPS . . . . .	38
7.6	Gràfica R-P per a les consultes amb tonalitat . . . . .	39
7.7	Gràfica R-P per a les consultes sense tonalitat . . . . .	40
B.1	Teclat del piano: escala cromàtica . . . . .	62
B.2	Representació musical de les notes . . . . .	63
B.3	Tipus de claus musicals . . . . .	63
B.4	Exemple d'interval . . . . .	65
B.5	Exemple de transposició . . . . .	65

# Índex de taules

---

3.1	Tipus de missatges MIDI . . . . .	15
3.2	Codis MIDI de les notes . . . . .	15
4.1	Resultats de l'anàlisi del GT . . . . .	20
6.1	Equivalències entre notació musical i notació geomètrica . . . . .	28
7.1	Equivalències entre notació implícita i notació musical . . . . .	34
7.2	Resultats del càlcul d' <i>average precision</i> . . . . .	40
B.1	Classificació dels intervals musicals . . . . .	64

---

---

# CAPÍTOL 1

## Introducció

---

Dins la investigació en l'àmbit musical, una de les tasques clau a l'hora d'estudiar les partitures és l'anàlisi melòdic. Quan es tracta de partitures antigues, qui investiga aborda freqüentment documents manuscrits, molts dels quals no es troben digitalitzats en l'actualitat. Els escrits presenten dificultats de lectura i interpretació i és per això que resulta de gran utilitat l'ús de ferramentes informàtiques que permeten extraure informació (melodies, ritmes, lletres de cançons...) dels documents (partitures) de forma automàtica. Aquesta tasca es coneix com a processament i reconeixement d'imatges de documents.

### 1.1 Motivació

---

Al llarg de les últimes dècades, la humanitat s'ha establert en una època de digitalització, on delega cada volta més tasques a les màquines. Quan es tracta d'activitats mecàniques simples, la incorporació exitosa de la tecnologia ha quedat assolida però no ocorre el mateix amb situacions complexes on s'ha de dotar als aparells de sentit visual. Des que es va començar la investigació en aquesta matèria, els problemes que s'han volgut abordar, han resultat ser més complexos del que semblava al començament.

Agafant com a referència la visió humana, la facilitat amb què l'ull pot interpretar el que ocorre acaba sent enganyosa. Si aquest sistema es volguera simular informàticament, suposaria l'ús d'una gran quantitat de recursos de processament a més de complexos problemes de control. És per això que la visió per computador és un dels problemes intel·lectuals més difícils d'afrontar [1].

Durant anys de recerca, no només la visió per computador ha evolucionat. El reconeixement de patrons (PR<sup>1</sup>) dins de les imatges processades ha estat també progressant. Idees bàsiques sobre aquesta matèria començaren amb l'ús de mètodes probabilístics i de la teoria de la decisió estadística [2] i el panorama canvià amb la vinguda de les xarxes neuronals artificials. A més, altres mètodes com les màquines de suport de vectors apareixeren en escena [1]. Totes aquestes tècniques donaren lloc a un nou concepte anomenat *Machine Learning*, que inclou no només la minimització de les taxes d'error sinó també la inclusió sistemàtica de probabilitat i optimització matemàtica.

Recentment, el PR s'ha vist desplaçat dels sistemes sense interacció humana a sistemes als quals els processos estan afectats pel *feedback* humà. Un exemple remarcable del reconeixement de patrons on aquest *feedback* pot ser beneficiós és la tasca de transcripció de documents manuscrits. [3]

---

<sup>1</sup>*Pattern Recognition.*

Dins del reconeixement de patrons, el reconeixement de text manuscrit (HTR<sup>2</sup>) no és un camp nou. De fet, els seus inicis es remunten als anys 60. Encara així, la tecnologia disponible en aquella època no era suficient per utilitzar-la a gran escala i és per això que les primeres investigacions se centraren en OCR<sup>3</sup> [3].

L'interès pel HTR va estar disminuint durant algun temps baix la falsa assumpció que les tecnologies computacionals modernes tornarien els documents físics inútils [3]. Malgrat això, s'ha demostrat degut a la grandíssima quantitat de manuscrits històrics que encara queden per digitalitzar que es tracta d'una tasca rellevant. L'objectiu de la digitalització d'aquests documents no només permet la seua preservació i els fa més accessibles [4]. El fet de tindre'ls en un format electrònic permet als investigadors i investigadores noves formes d'indexar i consultar aquests documents [3], la qual cosa pot obrir noves línies de recerca.

Dintre dels diferents tipus de documents manuscrits, es troben les partitures. Grans col·leccions d'aquest tipus d'arxius són preservades en catedrals i abadies i, sense transcripcions fiables, resulten inaccessibles [5]. Es tracta d'un cas especial de transcripcions, ja que localitzar símbols musicals aïllats no resulta de gran utilitat (la majoria d'ells apareixen en totes les pàgines). En canvi, si l'objectiu de les cerques són patrons melòdics, es pot obtindre major informació, fent així accessible al públic la música plasmada a les partitures [5, 6].

A més, molts dels problemes derivats de la gran quantitat de dades que es tracten es tornen menys crítics si, en compte de perseguir realitzar transcripcions exactes, l'objectiu se centra en determinar la probabilitat que un element es trobe en una regió de la imatge [4]. Eines com la indexació probabilística (PrIx<sup>4</sup>[4]) o el KWS<sup>5</sup> [7], les quals es tractaran en els propers capítols, han sorgit amb aquesta motivació.

Una eina creada a partir de les tècniques esmentades anteriorment és el motor de cerca d'indexació probabilística per a textos manuscrits creat pel PRHLT *Research Center* [8]. En concret, el demostrador creat per a la cerca de seqüències de símbols musicals en imatges és accessible a través de <http://prhlt-carabela.prhlt.upv.es/musica>. El propòsit d'aquest treball és l'estudi i implementació de noves funcionalitats aplicables al demostrador, deixant d'aquesta forma una porta oberta per a la creació d'una nova eina que es pugui beneficiar de les conclusions que s'hi extreuen.

La línia principal de modificació ve donada pel fet que les cerques al buscador es basen en consultes en **notació geomètrica** que, com es s'exposarà més endavant, no resulta un mètode accessible des del punt de vista de la teoria musical.

Baix aquestes condicions, es podria adaptar la cerca per a un ús amb una notació específica musical. És a dir, s'hauria de canviar la forma en que es fan les consultes emprant no només una notació diferent, sinó també dispositius com teclats (piano) virtuals o MIDI (vore 3.4) per introduir les dades.

D'altra banda, moltes voltes no només resulten rellevants les consultes musicals respecte de les notes en si, sinó respecte d'altres característiques. Així, per exemple, no importa el nom de les notes al resultat, sinó els intervals que les separen. Per això, es planteja també l'adició d'un nou mode de cerca referent a la clau de la melodia<sup>6</sup>.

---

<sup>2</sup>*Handwritten Text Recognition.*

<sup>3</sup>*Optical Character Recognition.* Tracta el reconeixement individual de caràcters aïllats, normalment utilitzant tècniques de reconeixement de patrons simples [3].

<sup>4</sup>*Probabilistic Indexing.*

<sup>5</sup>*Keyword Spotting.*

<sup>6</sup>Una explicació més àmplia sobre aquest i altres conceptes musicals es pot consultar a l'apèndix B.

Aquestes propostes de canvi seran provades i avaluades tant a nivell d'interacció d'usuari com a nivell de rendiment del sistema; cal respondre si noves formes de realitzar consultes aporten qualcuna milloria en l'actuació del buscador. Totes aquestes proposicions es fonamenten, per tant, dins d'un marc d'investigació de cara a futures implementacions.

D'altra banda, els canvis no només poden ser aplicats en relació a l'entrada de dades, sinó també a l'eixida i la visualització. Aquest és un problema que abordarà Carmen Bueno Hurtado en el seu TFG *Renderización acústica de patrones melódicos localizados en grandes colecciones de imágenes de partituras manuscritas históricas* i per tant, en aquest projecte només es tractaran els continguts relacionats amb l'entrada de dades.

## 1.2 Objectius

---

A partir de la premissa anterior, la introducció de noves funcionalitats dins de la interfície perquè siga musicalment accessible, es poden derivar els següents objectius a escometre:

1. Adició de noves entrades en format musical, que són:
  - (a) Per teclat d'ordinador
  - (b) Per teclat virtual web
  - (c) Per teclat MIDI
2. Creació d'un algoritme de traducció de notació geomètrica a notació musical
3. Adició de nous tipus de cerques relatives a la melodia d'entrada
  - (a) Mantenint el to original
  - (b) Sense tindre en compte el to
4. Aplicació dels canvis dins de la interfície web
5. Avaluació dels nous tipus de cerca introduïts

## 1.3 Estructura de la memòria

---

Els objectius esmentats anteriorment es tracten al llarg dels 7 capítols vinents. On, en primer lloc, s'explica el mar teòric dins del qual s'engloba la ferramenta a canviar, a més d'una exposició del seu funcionament i disseny intern. Més endavant, es troba la documentació mínima requerida per entendre els fonaments de cada tecnologia emprada per assolir objectius proposats.

A continuació, es realitza un anàlisi per parts del problema a abordar, tenint en compte l'estat inicial, les dades i l'ús que es farà de la ferramenta. A través d'aquest anàlisi s'arriba en el següent capítol a un esquema bàsic del sistema que es realitzarà. Aquest esquema es desenvoluparà més endavant dividint-se en dos blocs principals, un algorítmic i altre sobre la web, que després s'uniran i es tractaran d'integrar.

Més tard, es troben les proves i avaluacions realitzades per tal de determinar la qualitat de les modificacions implantades, juntament amb els detalls teòrics per entendre la validesa d'aquestes i les conclusions extretes. Com a cloenda, es troben conclusions en les quals es farà una revisió comparativa de l'estat de partida i l'estat final, valorant què s'ha aconseguit tenint en compte els objectius inicials.

Al final del document, queden tant la bibliografia amb totes les fonts emprades i dos apèndixs: un amb tot el codi desenvolupat al llarg del projecte i un altre sobre teoria musical bàsica per poder ubicar millor al lector o lectora en el desenrotllament del treball.



---

---

## CAPÍTOL 2

# Estat de l'art

---

Abans d'endinsar-se en l'explicació del treball realitzat al projecte, és necessari fer una revisió de l'estat inicial previ al seu desenvolupament. A continuació, es resumeix el marc teòric en què es basa el demostrador dins del qual s'inclouran les modificacions a més del seu disseny, estructura i funcionament previ a la seua modificació.

### 2.1 Indexació probabilística: PrIx

---

Com s'ha mencionat anteriorment, el problema ja no es tracta d'aconseguir una traducció precisa paraula a paraula en el document, sinó determinar com de probable és que una paraula donada estiga o no escrita dins d'alguna regió indexable de la imatge, ja siga una línia, un paràgraf o, en el cas de les partitures, un **pentagrama**.

Aquest tipus de processament té relació amb KWS, on l'objectiu és determinar localitzacions en una imatge de text les quals tinguen una alta probabilitat de contindre instàncies de les paraules consultades (sense una transcripció explícita del document) [7]. Aquest és també el propòsit de la indexació probabilística (PrIx). La diferència és que aquesta tècnica se centra, més que en paraules clau específiques, en les localitzacions més probables de totes les paraules candidates per ser paraules clau; al mateix temps que es determinen, també s'indexen amb les seues corresponents probabilitats [4].

La taxonomia tradicional en KWS distingeix dues formulacions diferents: *Query-by-Example* (QbE), on la consulta (o *query*) s'especifica a partir d'imatges d'exemples o *Query-by-String* (QbS), on al sistema li és donada una representació simbòlica en forma de cadena del contingut que es vol buscar [9].

Dins l'àmbit musical, existeix investigació prèvia en detecció de símbols musicals (MSS<sup>1</sup>) amb QbE [10]. Pel contrari, el marc de treball desenvolupat a [9], explicat més endavant, se centra en la detecció de símbols musicals a partir de QbS (QbS MSS a partir d'ara), donada la semblança amb la cerca musical basada en contingut tradicional.

La pregunta fonamental que intenta contestar la detecció de símbols musicals és si donades una consulta (una seqüència de símbols musicals, una melodia) i una imatge d'una partitura (els pentagrames), la consulta es troba escrita en la imatge donada. A continuació, s'expressa de forma condensada el procés d'indexació probabilística desenvolupat a [9] que se segueix en el demostrador.

---

<sup>1</sup>*Music-Symbol Spotting*



### 2.1.1. PrIx per a símbols musicals individuals

Des d'una visió probabilística, es defineix una variable  $R$  que indica si una imatge d'una partitura  $x$  és rellevant per al símbol musical  $v$ . Una imatge es considera rellevant si  $v$  es troba escrita en ella. La distribució de probabilitat  $P(R|Q, X)$  pot utilitzar-se per resoldre la qüestió [9]. El seu càlcul és:

$$P(R | v, x) \approx \max_{b \in \mathcal{B}} P(v | x, b) \quad (2.1)$$

On  $\mathcal{B}$  és un conjunt de caixes d'inclusió ( $\mathcal{BB}^2$ ). Com en  $P(v | x, b)$   $b$  ve donada, es pot dir que aquesta probabilitat és la probabilitat a posteriori associada a algun classificador de paraules aïllades (és a dir retallades de la imatge  $x$  mitjançant  $b$ ).

L'equació 2.1 s'usa generalment a nivell de línia d'imatge; és a dir, s'assumeix que  $x$  és la imatge retallada d'una línia dins de la imatge que la conté. En aquest cas, els  $\mathcal{BB}$  es defineixen simplement com a intervals de posició horitzontal. En l'actualitat la detecció automàtica de línies de text és un problema que es considera raonablement resolt per a imatges de text en les quals les línies segueixen un patró suficientment regular i homogeni [11, 12].

Per tal d'aconseguir una precisió d'indexació alta, es requereix que  $P(v | x, b)$  vinga donada per un classificador que tinga en compte el context lingüístic de  $v$ .

Actualment, la forma d'aconseguir els millors resultats és mitjançant l'ús de reconeixadors holístics d'imatges de text. Aquests són capaços d'obtenir la seqüència de paraules més probable continguda en una línia dada d'una imatge de text. En [13] i [14] es descriu com calcular  $\max_{b \in \mathcal{B}} P(v | x, b)$  (i amb això  $P(R | v, x)$ ) mitjançant grafs de paraula (*Word-Graphs*), obtinguts amb un reconeixedor d'aquest tipus.

### 2.1.2. PrIx per a seqüències de símbols musicals

Normalment en àmbit d'investigació musical no es volen recuperar només note aïllades, sinó que interessa recuperar seqüències de notes o melodies senceres.

Una forma d'oferir aquesta possibilitat és estendre la teoria descrita per detectar notes individuals i dur-la per gestionar les consultes de seqüències com a consultes booleanes (combinacions de diferents símbols musicals) amb tres operadors bàsics: AND, OR i NOT; denotats com  $\wedge$ ,  $\vee$  i  $\neg$  respectivament.

Una aproximació suficientment bona per computar de forma eficient les probabilitats associades a combinacions arbitràriament complexes d'operadors booleans por trobar-se a [15]. Gràcies a aquestes, es pot computar de forma senzilla una aproximació de la rellevància de qualsevol combinació booleana de *queries* de símbols individuals.

Cal destacar que la formulació actual no permet una cerca de seqüències apropiada. L'operació booleana AND és una condició necessària però no suficient, ja que és necessari afegir una restricció respecte de la posició geomètrica dels símbols. [9]

De tota forma, el marc de treball considerat fa ús dels grafs de paraula, que també proveïen informació sobre la localització dels símbols en la secció de la partitura.

La tasca de la implementació de la detecció per a seqüències de símbols musicals serà matèria de futurs treballs.

<sup>2</sup>*Bounding Boxes*

## 2.2 Demostrador del motor per a seqüències musicals del PRHLT

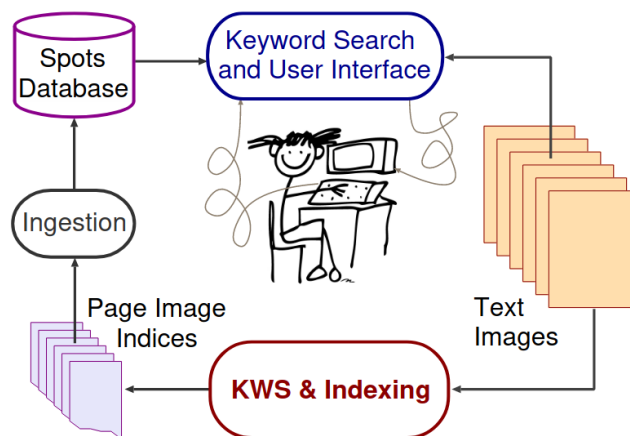
El demostrador web esmentat anteriorment, accessible a través de <http://prhlt-carabela.prhlt.upv.es/musica/>, permet utilitzar el motor de cerca i indexació de símbols i seqüències musicals sobre partitures manuscrites antigues (creat pel PRHLT) de forma visual a través d'una interfície d'usuari. A continuació, es mostren els seus elements principals, així com el funcionament i l'estructura que el suporta.

### 2.2.1. Esquema d'implementació

L'esquema general d'implementació es pot veure a la figura 2.1.

La ingesta de dades (part esquerra) s'ha fet de forma *off line*, creant una base de dades amb els resultats de les deteccions; es tracta d'un procés computacionalment barat.

En la part superior central es troba la interfície web d'usuari, que serà on es realitzen les modificacions al llarg d'aquest projecte. Des d'aquesta es podran generar consultes que s'analitzaran de forma *on line* (part inferior) per trobar la informació demanada i presentar les imatges recuperades. Per a aquesta tasca es necessiten temps de resposta curts. [16]



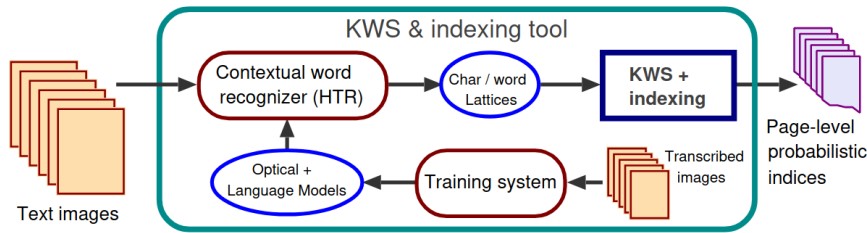
**Figura 2.1:** Esquema general d'implementació de la indexació probabilística i cerca d'imatges de text del motor del PRHLT. Font: [16]

L'esquema de la ferramenta d'indexació es pot veure a 2.2. Es basa en tecnologies KWS, utilitzant reconeixadors contextuals de paraules que requereix models entrenats per les imatges transcrites. Tant el reconeixedor contextual com el sistema d'entrenament solen ser peces de *software* separades, que no s'inclouen en la ferramenta. En general, KWS i la indexació poden ser molt exigents quan a capacitat de còmput. [16]

### 2.2.2. Arquitectura del demostrador

El demostrador segueix una arquitectura client-servidor formada per quatre components principals: client web (*Web client*), servidor web (*Web server*), servidor de dades (*Data server*) i servidor PrIx (*PrIx server*); com en la figura 2.3.

L'objectiu principal del *Web server* és oferir pàgines HTML de la interfície i preparar les respostes a les sol·licituds de l'usuari. Per a això, s'ha utilitzat un servidor HTTP (Apache) amb suport PHP. Aquest es connecta als altres servidors quan la informació sol·licitada no està disponible.

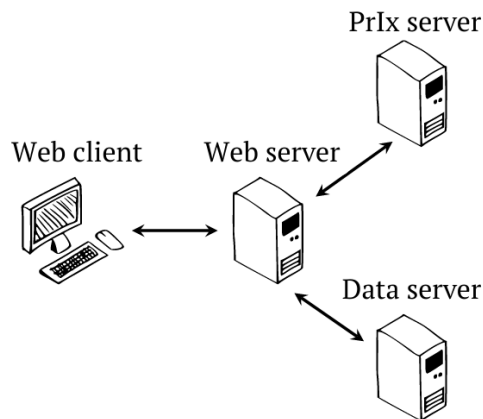


**Figura 2.2:** Esquema d'indexació probabilística en imatges de text construït a partir de KWS. Font: [16]

Aquests serveis no emmagatzemen directament les imatges i metadades (dades sobre *bounding boxes* i demès dades menors), sinó que es troben guardades en bases de dades separades que es gestionen a través del *Data server* (encara que, en la pràctica, el servidor de dades i el servidor web resideixen en la mateixa màquina) [17].

La interfície d'usuari es troba principalment en el client web. D'aquesta forma, quan l'usuari envia una *query* de PrIx a través d'ella al servidor web, aquesta és redirigida al *PrIx server* (que és realment el motor) i prepara les pàgines HTML que mostraran els resultats de la cerca a l'usuari. Les dades més importants d'indexació probabilística resideixen a aquest servidor. [17]

D'altra banda, malgrat que la figura pugui suggerir que tots els nodes es troben en computadors diferents, realment tots ells poden ubicar-se dins del mateix. En la pràctica, el més habitual és que els tres servidors residisquen (s'executen) en el mateix computador, mentre que el client ho fa en el computador de cada usuari.



**Figura 2.3:** Esquema de l'arquitectura client-servidor del demostrador. El client web (*User Interface* a la figura 2.1) és un navegador web regular, el servidor web rep les peticions de l'usuari i prepara les respostes com a fitxers HTML. El servidor PrIx (*Keyword search* a la figura 2.1) conté els índex probabilístics. El servidor de dades emmagatzema les imatges i algunes metadades dels documents a una base de dades. Font: [17]

### 2.2.3. Client web i consultes

La interfície consisteix en una barra de cerca on l'usuari pot escriure la consulta desitjada. També es pot especificar un nombre màxim de resultats que el sistema mostrarà i un interval de confiança (llindar) per filtrar els resultats que tinguen una probabilitat menor que eixe llindar.

Les consultes que pot escriure l'usuari són combinacions lògiques booleanes d'un o més termes, emprant operacions lògiques AND, OR o NOT (expressades com «&&», «| |» i «-» respectivament). Si es vol realitzar una consulta en la qual es vulga explicitar l'ordre d'aparició dels elements, és a dir es vol indicar una seqüència, s'hauran d'incloure aquests entre «[» «]».

Per exemple, si es volen buscar les melodies «L1 L2 S3» i «S1 S2 S3», la consulta s'expressarà com a «[L1 L2 S3] && [S1 S2 S3]».

#### 2.2.4. Cas d'estudi: manuscrit VORAU-253

El conjunt de dades consta d'unes 450 imatges que contenen al voltant de 1.000 tetragrames i només 100 pentagrames. A nivell pràctic, un tetragrama té la mateixa funció que un **pentagrama** amb la diferència que el tetragrama està format per quatre línies horitzontals i no cinc.

La notació dins del manuscrit consisteix en 19 símbols diferents. En notació gòtica germànica, les notes es representen com quadrats dibuixats sobre els tetragrames. D'altra banda, estan les dues claus utilitzades: c per a la clau de do i f per a la de fa.



Figura 2.4: Exemple d'una imatge de les pàgines del manuscrit VORAU-253. Font: [18]

Per entrenar el model òptic adequadament, s'ha hagut d'emprar una notació gràfica, on s'explicita la posició vertical de les notes (vore 2.5). Existeix una transcripció manual literal feta per un expert musicòleg del manuscrit [18] anomenada *ground truth* (GT). Aquesta s'analitzarà en la secció 4.5.

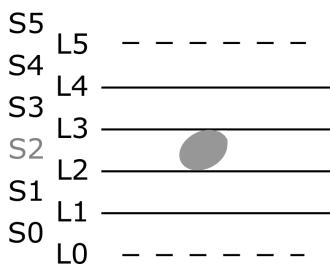


Figura 2.5: Anotació de les notes respecte de la seua posició gràfica. Font: [5]



---

---

## CAPÍTOL 3

# Tecnologies emprades

---

### 3.1 Servidor web: ús de Docker

---

Com que una part de la feina a realitzar es tracta de la modificació d'una web ja existent, és necessària una ferramenta per mantindre el servidor i que siga accessible per poder realitzar consultes i visualitzar els resultats (incloent-hi les modificacions). Per aquest propòsit, s'ha utilitzar Docker.

#### 3.1.1. Introducció

Docker és una plataforma oberta creada per desenvolupar aplicacions. Permet separar les aplicacions de la infraestructura per poder distribuir el *software* de forma eficaç, reduint així el temps entre l'escriptura de codi i l'execució en la seua producció. Docker es fonamenta en l'empaquetament d'aplicacions en un entorn aïllat anomenat contenidor. Els contenidors estan basats en imatges. D'aquesta forma, no es depèn de la instal·lació present al *host* per a la seua execució; tothom podrà rebre el mateix contenidor i funcionarà de la mateixa forma en qualsevol sistema. [19]

#### 3.1.2. Funcionament

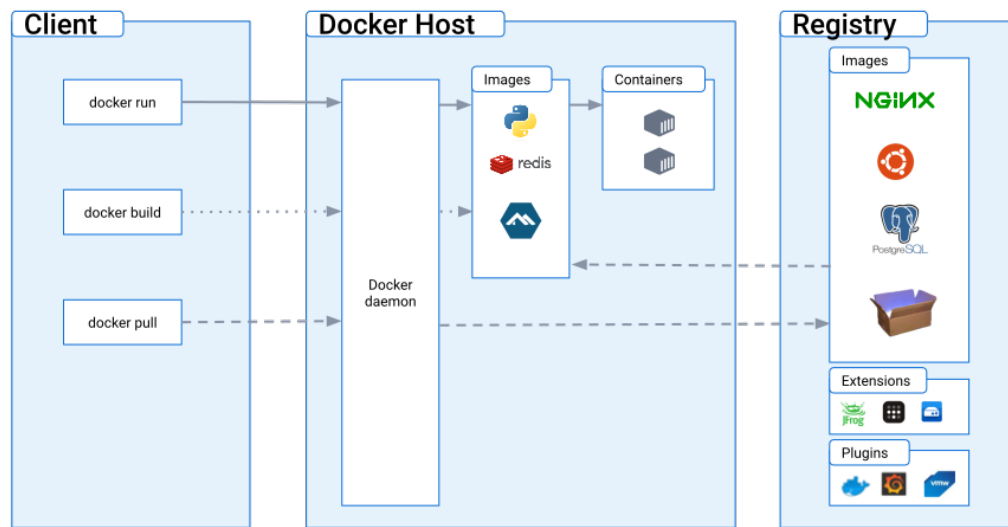
Malgrat no siga necessari conèixer en profunditat la mecànica interna de Docker per abordar aquest treball, a continuació es presenten alguns conceptes bàsics que faciliten l'ententiment de l'estructura i funcionament del projecte per si pogueren ser d'utilitat per al lector o lectora.

**Arquitectura Docker** Docker utilitza una arquitectura client-servidor. El *Docker client* es comunica amb el *Docker daemon*, qui s'encarrega de la construcció, execució i distribució dels contenidors. Ambdós es poden executar en el mateix sistema o també es pot connectar el client a un *daemon* remot. Client i servidor es poden comunicar a través d'una API (interfície de programació d'aplicacions) REST<sup>1</sup>, sòcols UNIX o una interfície de xarxa. Altre client possible és *Docker Compose*, que permet treballar amb aplicacions formades per un conjunt de contenidors.

L'esquema que segueix Docker és el següent:

---

<sup>1</sup>Respresentational State Transfer. És una arquitectura de programari pensada per a sistemes distribuïts basats en hipermèdia com ara el web [20]



**Figura 3.1:** Esquema de l'arquitectura de Docker. Font: [19]

**Docker daemon** Escolta les demandes API de Docker i gestiona objectes Docker com imatges, contenidors, xarxes i volums. A més, també pot comunicar-se amb altres *daemon* per administrar serveis.

**Imatges** Una plantilla de només lectura amb les instruccions necessàries per crear un contenidor Docker. Usualment, una imatge es basa en una altra imatge a la qual se li afegeixen modificacions. Es pot, per exemple, crear una imatge nova a partir d'una imatge Ubuntu.

**Contenidors** Una instància executable d'una imatge. És possible parar-los, moure'ls o esborrar-los. Es pot connectar un contenidor a una o més xarxes, assignar-li emmagatzemament o fins i tot crear noves imatges a partir de l'estat actual del contenidor.

## 3.2 ferramentes de programació: llibreries i APIs

El codi resultant del treball ha sigut desenvolupat dos llenguatges de programació principalment: per un costat, s'ha emprat *Python* per realitzar tasques algorítmiques que requerien treballar amb diferents dades. Per altre, s'ha utilitzat *JavaScript* (juntament amb CSS i HTML) per a la gestió d'elements web.

En molts casos s'ha necessitat treballar amb elements que dificulten la programació, ja que tenen exigències per damunt de les possibilitats dels llenguatges base. És per això que s'han fet servir diferents llibreries i APIs per tal de tindre accés a noves funcionalitats que han permès obtenir un resultat amb una gestió més fàcil i eficaç. Independentment, també s'ha fet ús d'alguns comandaments útils dins de Linux. A continuació es presenten les diferents ferramentes emprades:

### 3.2.1. NumPy

*NumPy* és una llibreria creada per realitzar computacions científiques. Aquesta proveeix operacions a partir d'*arrays* multidimensionals amb els quals es pot operar de forma ràpida a més d'incloure operacions matemàtiques com transformades de Fourier, operacions d'àlgebra lineal bàsica i operacions estadístiques entre altres [21]. Per al cas concret del projecte, s'han fet servir les utilitats que permetien treballar amb dades per tal de llegir-les i escriure-les a fitxers, analitzar-les i crear estructures amb elles.

### 3.2.2. Web MIDI API

Aquesta API<sup>2</sup> sustenta el protocol MIDI de forma que admet que les aplicacions web enumerar i seleccionar *inputs* i *outputs* MIDI a més d'enviar i rebre missatges d'aquests dispositius [22]. Per tant, ha servit per poder gestionar una nova forma d'entrada al buscador web, les entrades dels dispositius MIDI (concretament els teclats).

### 3.2.3. Altres ferramentes

Per tal de realitzar operacions de gestió de fitxers de forma ràpida i còmoda, s'ha emprat el terminal de Linux. Comandaments com `grep` i `sed` han servit per reemplaçar el contingut del fitxers o per buscar elements dins d'ells. A més, el terminal ha sigut d'utilitat per desplegar els serveis de Docker que mantenien el servidor web del buscador.

## 3.3 Manteniment del treball

---

Per tal de mantindre tant el codi com la memòria d'aquest projecte actualitzats i per tal que fora accessible en totes les seues fases als tutors, s'ha creat un repositori dins de GitHub. Així, cada canvi i assoliment d'objectius ha quedat registrat cronològicament. És accessible a través de: <https://github.com/aitamb/tfg> on es trobaran diferents fitxers que guien al lector dins de l'estructura de directoris a més de contindre informació addicional útil per a l'enteniment de les tasques realitzades.

## 3.4 Protocol MIDI

---

Una part dels objectius descrits anteriorment es tracta d'implementar una entrada musical per tal de poder realitzar noves cerques de forma accessible. Per això, s'utilitza MIDI. MIDI (*Musical Instrument Digital Interface*) és un estàndard que descriu un protocol, una interfície digital i connectors que permeten la comunicació entre instruments electrònics i altres aparells digitals [23]. El sistema no tracta la generació dels sons, sinó la transmissió de la informació relativa al que es toca en un dispositiu musical electrònic. MIDI transporta missatges d'esdeveniments que especifiquen **nota**, **to** i **intensitat**. Aquests missatges poden ser enviats a través d'un cable MIDI a altres dispositius que controlen la generació de sons [24].

---

<sup>2</sup>*Application Programming Interface*. Literalment una interfície de programació d'aplicacions.



### 3.4.1. Missatges MIDI

Un missatge MIDI està format per una seqüència de huit *bits*, que és l'estat, seguida generalment per un o dos *bytes* més d'informació (dades). Existeixen una gran varietat de missatges MIDI diferents. En un alt nivell, aquests es poden classificar en:

- **Missatges de canal.** Es refereixen a un canal específic i dins d'ells hi ha una referència a aquest. Existeixen dos tipus:
  - **De veu.** Contenen dades d'interpretació musical i comprometen la majoria del tràfic en un flux de dades MIDI típic.
  - **De mode.** Afecten a la forma en què un instrument respondrà als missatges de canal de veu que rebrà.
- **Missatges del sistema.** No fan referència a cap canal en concret, per això no inclouen un número de canal en el seu estat. Es divideixen en:
  - **Comuns.** Destinats a tots els rebadors del sistema.
  - **De temps real.** S'usen per a la sincronització entre els components MIDI basats en un rellotge.
  - **Exclusius.** Inclouen un codi identificador (ID) del fabricant i són utilitzats per transferir qualsevol nombre de *bytes* de dades en el format especificat pel fabricant.

Per al cas del projecte, com es vorà més endavant, només són rellevants els missatges de canal de veu. Els diferents tipus i la seua estructura es presenten en 3.1. Com s'indica ací, existeix un valor per identificar cada to de cada una de les notes. A 3.2, es mostren totes les equivalències entre les notes de l' **escala cromàtica** i els codis MIDI segons l'octava en la que estiguen.

Missatge en binari		Descripció
Estat b7...b0 (nnnn és el canal MIDI)	Byte(s) de dades b7...b0	
1000nnnn	0kkkkkkk 0vvvvvvv	Nota <i>Off event</i> . S'envia quan una nota finalitza. <i>kkkkkkk</i> és el to de la nota. <i>vvvvvvv</i> és la intensitat.
1001nnnn	0kkkkkkk 0vvvvvvv	Nota <i>On event</i> . S'envia quan una nota comença. <i>kkkkkkk</i> és el to de la nota. <i>vvvvvvv</i> és la intensitat.
1011nnnn	0cccccc 0vvvvvvv	Canvi de control. S'envia quan el valor d'un controlador <sup>a</sup> canvia. <i>cccccc</i> és el número de controlador. <i>vvvvvvv</i> és el nou valor.
1100nnnn	0ppppppp	Canvi de programa. S'envia quan el <i>patch</i> <sup>b</sup> canvia. <i>ppppppp</i> és el número del nou programa.

**Taula 3.1:** Estructura i descripció d'alguns dels missatges de canal de veu MIDI. Font: [25]

<sup>a</sup>Els controladors poden ser pedals o palanques, entre altres. Alguns números es reserven per a propòsits específics (vore [23]).

<sup>b</sup>Es tracta d'un so preestablert dins d'un generador de so (sintetitzador).

Número d'octava	Número de nota											
	C	C $\sharp$	D	D $\sharp$	E	F	F $\sharp$	G	G $\sharp$	A	A $\sharp$	B
-1	0	1	2	3	4	5	6	7	8	9	10	11
0	12	13	14	15	16	17	18	19	20	21	22	23
1	24	25	26	27	28	29	30	31	32	33	34	35
2	36	37	38	39	40	41	42	43	44	45	46	47
3	48	49	50	51	52	53	54	55	56	57	58	59
4	60	61	62	63	64	65	66	67	68	69	70	71
5	72	73	74	75	76	77	78	79	80	81	82	83
6	84	85	86	87	88	89	90	91	92	93	94	95
7	96	97	98	99	100	101	102	103	104	105	106	107
8	108	109	110	111	112	113	114	115	116	117	118	119
9	120	121	122	123	124	125	126	127	-	-	-	-

**Taula 3.2:** Codis MIDI en notació decimal per a la representació de tots els tons possibles del protocol ordenats per octava. Font: [25]



---

---

## CAPÍTOL 4

# Anàlisi del problema

---

De forma simplificada, el problema a abordar a aquest projecte es redueix a afegir funcionalitats al demostrador (secció 2.2) esmentat per tal de fer el servei accessible per a persones familiaritzades amb el llenguatge musical. Per això, cal entendre quin és l'ús que se li donarà.

### 4.1 Ús del demostrador

---

A nivell d'usuari, aquest buscador *on line* serveix per fer consultes dins del manuscrit musical vorau-253 (esmentat en la secció 2.2.4). Per tant, els usuaris als quals va dirigida la ferramenta seran, en la immensa majoria dels casos, usuaris experts en matèria musical. Per això, cal adaptar la interfície perquè siga accessible i utilitzable de forma senzilla per aquest tipus de persones. Farà falta que la interfície compte amb ferramentes musicals, tant visuals com auditives, per transmetre i obtenir informació.

En conclusió, es necessiten una entrada i una eixida de dades en format musical; cal modificar el *input* i l'*output*. Com ja s'ha explicat en la introducció, aquest projecte només se centra en l'entrada de dades i és per això que d'ací endavant serà eixe el problema a analitzar.

### 4.2 Introducció de dades musicals

---

Amb la finalitat d'introduir dades en format musical, cal plantejar-se quin tipus de dispositius permeten maniobrar amb eixe llenguatge, ja que incloure dades musicals de manera manual suposaria un inconvenient per a l'experiència d'usuari.

Per una banda, existeixen els dispositius MIDI (secció 3.4), que permeten transformar a través d'un dispositiu informàtic amb aparença musical (un teclat de piano, una guitarra...) el que toca la persona amb l'instrument MIDI en missatges MIDI (secció 3.4.1). Aquests missatges són fàcilment gestionables en la web gràcies a APIs com la que s'ha introduït en la secció 3.2.2, la qual cosa facilita la seua incorporació al projecte.

Per altra banda, podria donar-se el cas en que l'usuari que utilitze el buscador no compte amb cap dispositiu MIDI o no li siga possible utilitzar-lo en el moment degut a problemes de compatibilitat amb el navegador o la màquina que estiga utilitzant, entre altres.

En aquests casos caldrà comptar amb una alternativa per poder introduir consultes en format musical; alguna ferramenta semblant a un piano virtual amb la qual es pugui interactuar a través del ratolí o el teclat de l'ordinador. A més convendria permetre introduir també manualment les consultes per guanyar flexibilitat.

### 4.3 Gestió de dades musicals

---

La inserció de dades de dades en format musical, posa en relleu altra qüestió important a gestionar: com fer que aquestes dades d'entrada siguin compatibles amb el buscador existent. Actualment, les dades es representen amb notació geomètrica. És a dir, una nota es representa amb la seua posició relativa al tetragrama en què es troba.

El mateix ocorre per indicar les claus, on s'escriu el nom (f per a la de fa i c per a la de do) seguit de la línia en la qual es troba. Així, per exemple, si s'han identificat dues figures, una en la segona línia i altra en el tercer espai, precedides d'una clau de fa en la tercera línia, la notació geomètrica serà: «f3 12 s3». Per representar el bemoll (b) s'escriu literalment «flat».

Per altre costat, amb l'entrada musical (notació musical) cada element representa una nota expressada en el sistema de notació anglosaxona juntament amb l'octava a la qual pertany eixa nota (vore apèndix B). Per tant, és necessari plantejar com es passarà d'una notació a altra, cal dissenyar una traducció de notació musical a notació geomètrica per tal d'inserir les dades.

Aquest problema és més complexe de resoldre del que sembla, ja que cada melodia musical no té correspondència directa amb la notació geomètrica. Dins d'aquesta traducció afecten les altures de les notes i les claus i, per tant, una melodia simple es transforma en una consulta *booleana* AND/OR geomètrica dins de la qual la majoria dels seus elements són al mateix temps seqüències.

### 4.4 Tipus de consultes al buscador

---

Després de plantejar la necessitat d'un esquema de traducció, cal tindre en compte en el seu disseny i implementació el tipus de consultes que es realitzen. Fins al moment, tot el que s'introduïa a l'espai de cerca del motor era buscat de forma literal. Això, però, pot limitar en alguns casos les cerques musicals establint un biaix i impedit fer cerques complexes que permetrien, a priori, arribar abans al resultat desitjat.

Així és en els casos en que la melodia que es busca no es correspon exactament amb les notes introduïdes, ja siga per error de l'usuari al transcriure l'altura a la que es troben els símbols musicals o perquè des d'un principi es volen buscar totes les melodies iguals a la introduïda però només en relació als intervals que separen les notes d'aquesta.

Per exemple, si es desitja buscar una melodia ascendent de tres notes separades entre elles per parells en un to, una proposta de cerca seria «C4 D4 E4». Encara que la cerca siguin explícitament eixes tres notes, s'hauria de donar l'opció de buscar totes les melodies que mantinguen eixa relació intervàlica per poder satisfer la consulta, ja que de forma restrictiva potser no es troben les notes tal qual, però les seqüències «G4 A4 B4» o «C3 D3 E3» sí podrien trobar-se i ambdues es corresponen la consulta inicial de les tres notes ascendents separades entre elles per un to.

A aquests tipus de consultes se'ls anomenarà amb tonalitat<sup>1</sup> (si tenen en compte el to original) o sense tonalitat (si no el tenen en compte) a partir d'ara.

Aquests tipus de consultes afectaran inevitablement a la traducció, ja que en cas que es vulguen realitzar, s'inclouran més possibilitats i caldrà traduir-les també. Complementàriament tant la traducció com les noves formes de cerca indueixen necessàriament una

---

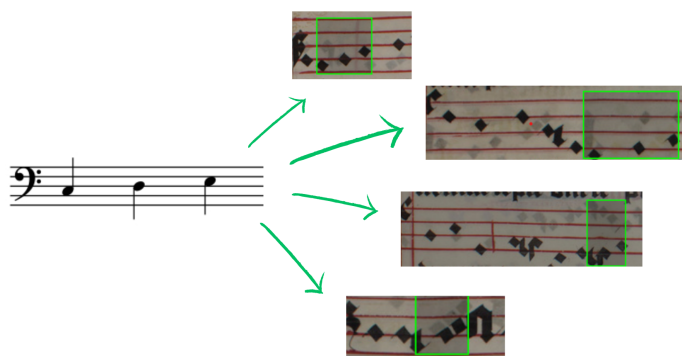
<sup>1</sup> Ací es fa un abús del llenguatge, ja que el terme tonalitat implica un context harmònic citeMichels que en aquest cas, ni es dona ni és rellevant per al desenvolupament del treball. Per fer més senzilla l'explicació, s'assumirà que la tonalitat fa referència al to i no a l'harmonia.

ampliació de la longitud de les consultes que es fan, això és degut a que més informació s'inclou per tal de satisfer les condicions musicals, com es vorà en el capítol 6.

Amb la finalitat de retallar les dades innecessàries inserides amb criteri, s'analitzarà el document del *ground truth* on es troba la transcripció real de tots els símbols musicals per a cada tetragrama del manuscrit per distingir quines consultes mai podrien obtenir resultats.

Es pot veure un exemple de cerca amb tonalitat per a la seqüència musical

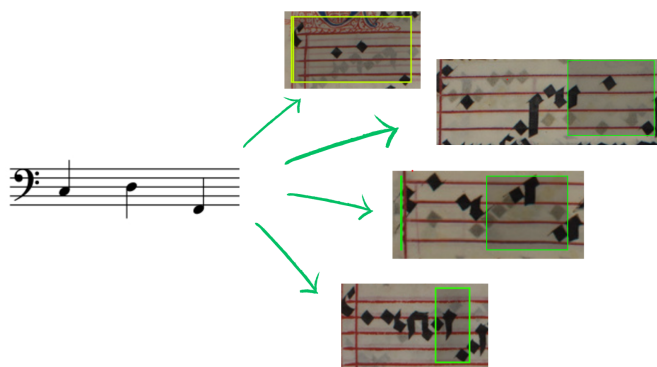
«C3 D3 E4» a 4.1 i un de cerca sense tonalitat per a la seqüència «C3 D3 F2» a 4.2.



**Figura 4.1:** Exemple de quatre possibles fragments de partitures recuperades per a la melodia «C3 D3 E3» («do3 re3 mi3») en el manuscrit musical VORAU-253, tenint en compte la tonalitat.

La consulta original (la *query* enviada) derivada de la melodia anterior és

«(c4 &20& [s0 l1 s1 ]) || (c5 &20& [s1 l2 s2 ]) || (f2 &20& [s0 l1 s1 ]) ||  
(f3 &20& [s1 l2 s2 ]) || (f4 &20& [s2 l3 s3 ])



**Figura 4.2:** Exemple de quatre possibles fragments de partitures recuperades per a la melodia «C3 D3 F2» («do3 re3 fa2») en el manuscrit musical VORAU-253, sense tindre en compte la tonalitat.

La consulta original (la *query* enviada) derivada de la melodia anterior es pot consultar en la secció 6.2.4. Els termes de l'expansió geomètrica que permeten la localització de les imatges mostrades són:

(c4 &20& [s2 l3 s0 ]) que és «sol3 la3 do3», (c4 &20& [l3 s3 l1 ]) que és «la3 si3 re3»,  
(f3 &20& [s3 l4 s1 ]) que és «sol3 la3 do3» i (c5 &20& [l4 s4 l2 ]) que és «la3 si3 re3».

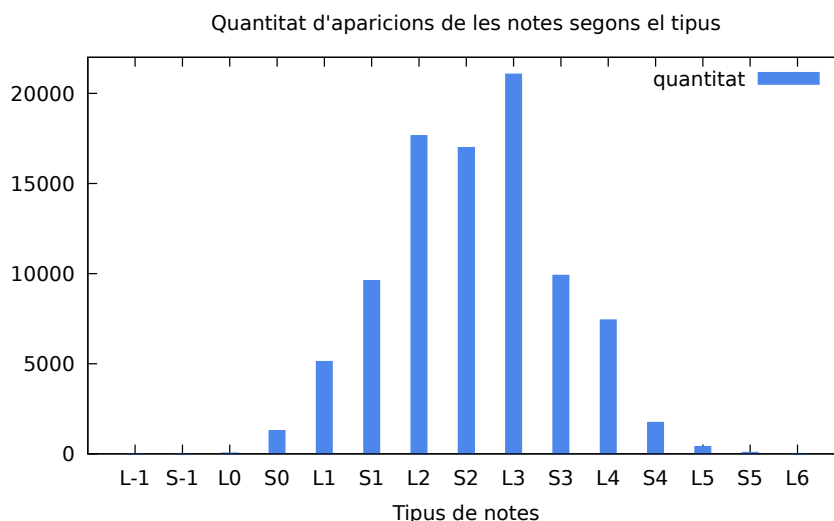
## 4.5 Anàlisi del *ground truth*

Dins del *ground truth* (GT) s'han transcrit tots els tetragrames del manuscrit, per cada tetragrama del manuscrit hi ha una línia de document amb les notes, claus i alteracions que hi ha escrites.

Els resultats de l'anàlisi del GT són els següents:

Nombre de tetragrames	4.594
Nombre d'elements	96.119
Nombre de notes	91.177
Nombre d'alteracions (b)	554

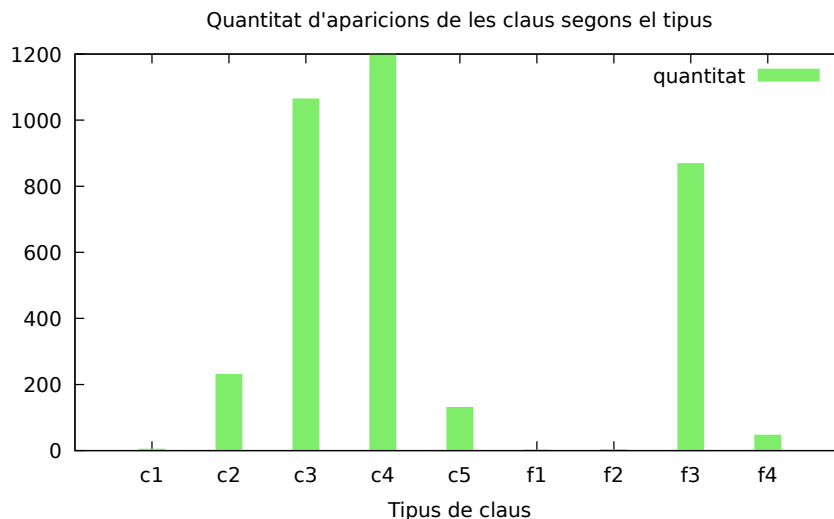
**Taula 4.1:** Resultats de l'anàlisi del document *ground truth* on es troba la transcripció del manuscrit [18].



**Figura 4.3:** Distribució del tipus de notes (segons l'espai o la línia en que es troben) que apareixen al manuscrit [18].

Per altra banda, s'ha comprovat que les alteracions trobades son bemolls (b) que només apareixen en línies i que només afecten a la nota «si» (com cabria esperar per la data en la qual es va compondre el manuscrit [26]).

Sobre les dades arreplegades en la figura 4.3 es pot apreciar que quasi no hi ha casos en que les notes sobrepassen el tetragrama per damunt o per avall. Això té sentit, ja que en música d'una època tan antiga no era freqüent l'ús de notes addicionals, sinó que quan les notes que es cantaven s'eixien del rang del tetragrama, es canviava de clau [27, p.158]. Sobre les dades a l'histograma 4.4 cal destacar que les claus que es troben amb més freqüència són les més greus. No és d'estranyar aquest fet degut a que aquest tipus d'escriptures eren cantades per homes, amb la veu més greu [26].



**Figura 4.4:** Distribució dels tipus de claus (segons la línia en que es troben) que apareixen al manuscrit [18].

## 4.6 Notes addicionals

Sobre les proves i l'avaluació de les noves funcionalitats proposades es parlarà en el capítol 7, on es realitzaran controls de funcionament a través de la interacció amb la interfície web i control de satisfacció d'obtenció dels resultats esperats. A més, es mesurarà la qualitat de l'aplicació a partir de mètriques com el *recall* o la *precision* juntament amb altres paràmetres a través de ferramentes d'avaluació existents.

Per tal de mesurar aquests serà necessari disposar d'una col·lecció relativament gran de consultes que fer al buscador. Així, s'avaluarà la qualitat dels resultats obtinguts pel demostrador tenint en compte les noves formes de consultar al motor de cerca, així com la combinació d'aquestes amb diferents formes de recuperar i buscar la informació que té el demostrador.

D'altra banda, l'hora d'incloure aquestes propostes al servidor web s'ha de tindre en compte que existeix un codi preexistent el qual s'està fent servir. Per mantindre la continuïtat i la coherència de la ferramenta, és dins d'aquest on es tractaran d'adaptar les noves propostes.

Com que l'objectiu es basa en provar d'afegir les proposicions dins del motor i no en implementar-les definitivament, no se seguirà l'estructura i gestió dels afegits a través de l'estructura explicada en 2.2, sinó que es tractaran com a una implementació superficial i momentània. Per això no s'abordarà la gestió de la bases de dades dins del treball, ja que per tal de provar el seu ús no és necessari.





---

## CAPÍTOL 5

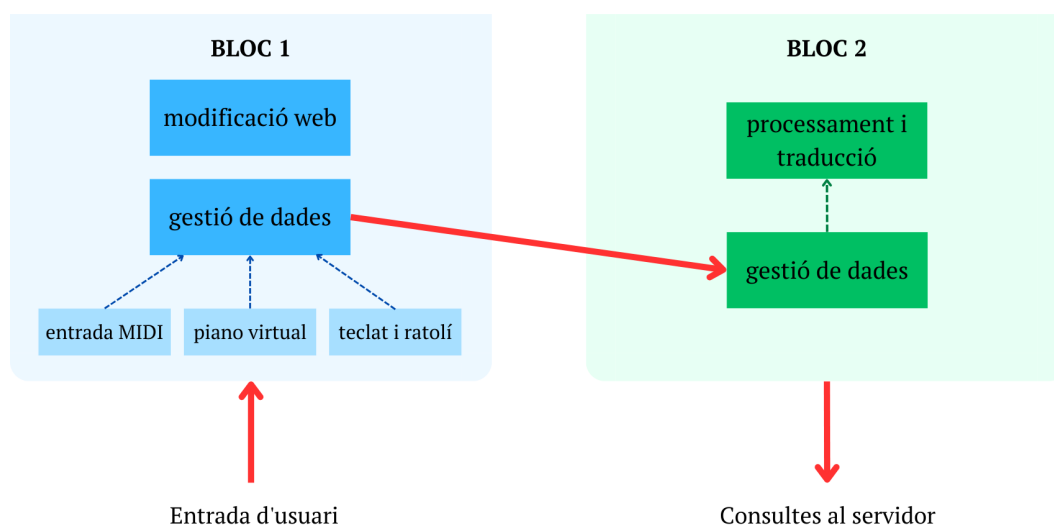
# Disseny de la solució

---

La solució proposada es pot dividir en dos grans blocs, segons s'ha analitzat. Un primer bloc centrat en el disseny web que inclou l'aplicació de les entrades musicals (el piano virtual per pantalla i per teclat, el teclat MIDI i l'entrada manual) i la gestió d'aquestes dades musicals perquè puguin ser escoltades. Dins del piano web, s'afegiran utilitats per poder escoltar les notes que s'estan tocant i que ajudaran a visualitzar de forma més clara com funciona l'entrada de dades musicals.

El segon bloc a abordar, conté el procés de creació de l'algoritme de traducció de notació musical a geomètrica. Després, les consultes musicals passaran per aquest bloc que gestionarà les dades per enviar-les en format geomètric al buscador. Aquesta traducció, a més, haurà de tindre en compte el tipus de consulta que es vol generar (tenint en compte el to original o sense tindre'l en compte), així com restriccions i possibles retalls de dades derivats de l'anàlisi del GT per tal d'evitar consultes innecessàries que sobrecarreguen el demostrador.

Un esquema simplificat del que es vol aconseguir es pot observar a la figura 5.1.



**Figura 5.1:** Esquema simplificat del projecte. L'entrada de l'usuari és la interacció amb la web per generar seqüències musicals i les consultes al servidor resulten en la recuperació dels documents (partitures) on es troben eixes les melodies consultades.



---

## CAPÍTOL 6

# Desenvolupament de la solució

---

Seguint el disseny de la solució exposat en el capítol anterior, s'abordarà el desenvolupament seguint els dos grans blocs: el bloc del desenvolupament web del piano i el bloc de traducció de notació musical a consultes lògiques geomètriques.

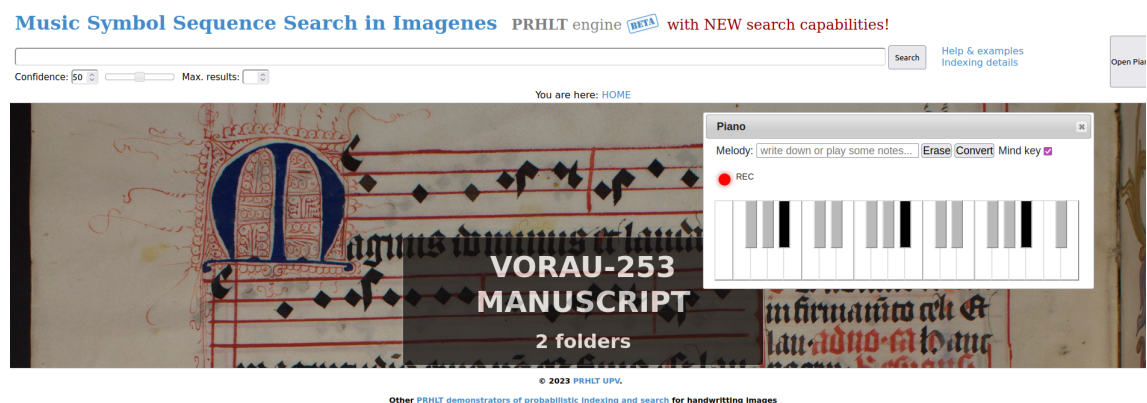
### 6.1 El piano a la web

---

L'entrada musical a la web s'ha incorporat a través de la pestanya del piano. Aquesta ha sigut implementada a partir d'un diàleg HTML dins del qual es troben diferents elements que permeten la inserció i visualització de melodies. Aquest diàleg és accessible gràcies a un botó a la part esquerra de la web que permet obrir-lo. Per comoditat, aquest element pot moure's al gust de l'usuari i tancar-se si se vol visualitzar millor la resta de la pantalla sense que açò afecte a les dades que es guarden en ell, ja que quan es torna a obrir es mantenen com estaven abans de tancar-se.

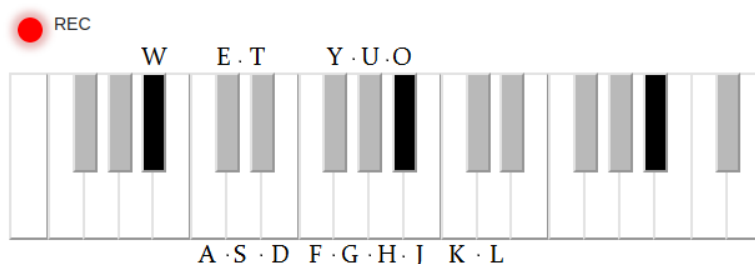
La part gràfica de l'eina ha sigut gestionada a través de fulls d'estil CSS integrats dins del codi HTML. Es poden consultar els paràmetres d'estil al codi mostrat en l'apèndix A. L'aspecte d'aquesta ferramenta es veu a la figura 6.1.

A continuació, es fa una revisió de cadascuna de les parts d'aquesta pestanya, de com s'ha construït i el seu funcionament.



**Figura 6.1:** Aspecte visual (captura de pantalla) de la ferramenta que permet la incorporació de consultes en format musical dins de la web.

### 6.1.1. Teclat



**Figura 6.2:** Teclat del piano web, simulant un teclat de piano real. S'inclou al costat de les tecles corresponent l'equivalència amb el teclat físic de l'ordinador.

Aquest teclat està format per botons HTML que simulen les tecles d'un piano real. Cadascun d'aquests botons té associat en els seus atributs el nom de la nota a la que corresponen (en notació musical, amb el nom de la nota i l'octava a la qual pertany). D'aquesta manera, cada volta que es pressiona el botó, es transmet la nota que s'ha tocat directament amb el nom.

D'altra banda, s'ha inclòs un àudio mp3 per cada nota del tecla per a que, així, gràcies a l'atribut del botó, també es pugui reproduir l'àudio corresponent a la tecla pressionada i així escoltar els sons com si es tractara d'un instrument real (A). Aquests botons també tenen associada una tecla del teclat de l'ordinador cadascú, de forma que també es poden tocar de manera física les notes. La distribució de les tecles es veu a la figura 6.2.

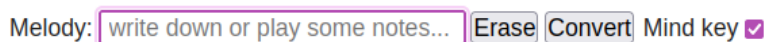
El rang de notes del teclat és de E2 a D5. S'ha pres aquesta decisió degut a l'anàlisi previ del GT on, gràcies a saber quines són la clau més aguda i la més greu del manuscrit (c1 i f5 respectivament), s'ha inferit quines eren la nota més aguda i la més greu que podia aparèixer al manuscrit. Gràcies també a aquest anàlisi, se sap que cap de les notes alterades (les no naturals, corresponents a les tecles negres) pot aparèixer en el manuscrit a menys que es tracte d'un B $\flat$ . Per això, les tecles negres referents a les notes alterades que no siguin eixe si bemoll tenen una coloració més clara.

A més, si es pressiona alguna d'aquestes tecles grises, s'escoltarà el to que els corresponga però no es transmetrà la nota que s'ha tocat, ja que si s'enviara, mai es podria recuperar en el manuscrit perquè no existeix.

Per últim, a la part superior esquerra del piano, es pot trobar un botó de gravació (REC). La finalitat d'aquest és poder tocar les tecles sense la necessitat de transmetre les notes que s'han tocat, ja que moltes voltes quan es pressionen els botons, només es volen fer proves de so o fins i tot "assajos" de la vertadera melodia que es vol consultar.

D'aquesta forma, si el botó és de color roig brillant, s'indicarà que la gravació està activa i que, per tant, es transmetran les notes que es toquen. Per altra banda, si el botó té un color apagat, no es transmetrà cap nota. Per poder activar i desactivar l'opció de gravació, només cal pulsar el botó mateix.

### 6.1.2. Barra de cerca



**Figura 6.3:** Barra de cerca del piano web, on es transmetran les notes que es toquen en el teclat en notació musical.

Aquesta barra de cerca conté un element HTML del tipus *input* on s'indiquen les consultes en notació musical. Aquesta barra s'anirà omplint de forma automàtica amb la melodia que toque l'usuari ja siga per teclat MIDI, teclat físic de l'ordinador o teclat virtual, sempre i quan l'opció de gravació estiga activa.

A la part dreta s'inclouen les opcions de cerca que es tindran en compte a l'hora de realitzar la traducció a *query* lògica geomètrica. També compra amb dos botons: un anomenat *Erase* que servix per esborrar tot el contingut de la barra de cerca, i altre anomenat *Convert* a través del qual s'inicia la traducció. Després del procés de traducció, la consulta resultant s'escriurà dins de la barra de cerca externa de la web, sense afectar al contingut de la barra de cerca del piano.

### 6.1.3. Opcions de cerca

Pel moment, l'única opció de cerca aplicable és la casella seleccionable anomenada *Mind key*. Aquesta permet realitzar la traducció tenint en compte el to de la melodia introduïda (casella activa) o sense tindre'l en compte (casella inactiva).

### 6.1.4. Entrada MIDI

Per permetre l'entrada a aquesta interfície on es troba el piano a través d'un instrument MIDI, s'han gestionat les entrades MIDI a través de la API MIDI. Aquesta ve inclosa en la majoria de navegadors, malgrat que en alguns hi ha problemes de compatibilitat i per això no sempre serà possible incloure aquesta funcionalitat. En primer lloc, es comprova si hi ha algun dispositiu connectat. De ser així, s'accedeix al dispositiu connectat i cada volta que es reba un missatge MIDI, es gestiona.

Per gestionar els missatges MIDI, gràcies a la API, només cal accedir a l'atribut *data* del missatge, a on es comprovarà que aquest missatge rebut fa referència a l'enviament d'una nota; comprovant si la seua longitud és igual a  $3^1$  i si la seua durada és major a zero (no és un silenci). Si es compleixen les condicions, es treu l'equivalència en notació musical seguint les equivalències a 3.2. A continuació, es reproduïx el so mp3 associat a eixa nota i s'incorpora (si l'opció de gravació està activa) a la consulta en notació musical.

## 6.2 La traducció

---

El problema es redueix a transformar melodies en format musical (F3 G3 A3) a una única *query* lògica geomètrica. El plantejament és senzill però se'n deriven una sèrie de problemes degut a la natura de les seqüències musicals que dificulten la transformació directa entre ambdós formats.

### 6.2.1. Plantejament del problema

El problema és que no existeix una equivalència directa entre una nota en notació musical i una nota en notació geomètrica. Com bé s'ha explicat a l'apèndix B, una figura en un pentagrama, no té un significat melòdic per si mateixa si apareix a soles. Necessita una clau de referència per poder tindre sentit. És a dir, la nota geomètrica L2, per exemple, fa referència a una figura dibuixada en la línia número 2. Però aquesta notació no transmet cap mena d'informació sobre l'altura de la nota per si mateixa, necessita eixa referència que és la clau.

---

<sup>1</sup>Açò és degut a que les notes MIDI tenen tres paràmetres diferents: altura, durada i intensitat

Si hi haguera una clau de fa en quarta (f4) davant, estaria representant a una nota diferent a la que representaria si hi haguera una clau de fa en tercera (f3). De la mateixa forma, una figura en notació musical pot ser indicada de diverses formes perquè pot ser representada en qualsevol línia o espai si la precedeix la clau indicada. Eixe és el problema principal de la traducció, no hi ha *una única* traducció directa, sinó *una sèrie de possibilitats*. De fet, hi ha una possibilitat per cadascuna de les claus.

L'esquema de traducció es basarà en traduir les seqüències representant-les en totes les claus possibles. Buscant quina és la posició que ocupa en el tetragrama cada nota de la melodia (quina línia o espai) per a una clau de referència. Així, s'obtidrien les equivalències geomètriques que després servirien per crear les *queries*. Per tant, per a una melodia, la *query* que s'enviaria al servidor serà una consulta lògica OR formada per totes les possibles representacions d'eixa melodia.

### 6.2.2. Primer pas: traducció geomètrica

El primer pas es la traducció de la melodia en totes les possibles seqüències geomètriques (una equivalència per clau). Una forma de processar una seqüència musical és traduir una a una les notes que formen la melodia. Aquestes notes, al mateix temps, s'hauran de processar en cadascuna de les claus possibles. Així, el resultat serà un conjunt de traduccions. El codi seguirà un esquema semblant a:

```
Require: estructures per guardar una consulta associada a cada clau
for cada nota en la melodia do
  for cada tipus de clau do
    traduir la nota per a la clau i afegir-la a la traducció de la clau corresponent
  end for
end for
```

Aquest primer pas tindrà com a resultat una seqüència de notes geomètriques associades a cada clau. Per realitzar la traducció, s'ha utilitzat una taula d'equivalències per claus i es pot vore a 6.1.

	E2	F2	G2	A2	B2	C3	D3	E3	F3	G3	A3	B3	C4	D4	E4	F4	G4	A4	B4	C5	B5	D5
c1	-	-	-	-	-	-	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5
c2	-	-	-	-	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-
c3	-	-	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-
c4	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-	-	-
c5	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-	-	-	-	-
f1	-	-	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-
f2	-	-	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-	-	-
f3	-	-	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-	-	-	-	-
f4	10	s0	11	s1	12	s2	13	s3	14	s4	15	s5	-	-	-	-	-	-	-	-	-	-

**Taula 6.1:** Taula d'equivalències entre notació musical i notació geomètrica segons el tipus de clau.

Després d'aquesta primera aproximació al primer pas de traducció, és observable que la quantitat de dades que se'n generaran és considerablement gran. Per això, cal plantejar-se com es poden retallar aquestes dades sense afectar a la qualitat del demostrador i, fins i tot, arribant-la a beneficiar. Una possibilitat ve donada per la taula mateixa d'equivalències (6.1). Dins d'aquesta, no s'han inclòs les equivalències de les notes que suposaven més d'una línia o espai addicional del tetragrama.

Això es fonamenta en l'anàlisi del GT, on no apareixien posicions amb aquestes característiques. D'aquesta forma, si la traducció d'alguna nota es correspon amb un espai en blanc en la taula, es dona per invàlida i s'estalvia el càlcul de la traducció de la melodia que la continga, ja que si una de les notes de la seqüència no es pot trobar en el manuscrit,

tampoc es podrà trobar la seqüència mateixa. És a dir, s'estalviaran consultes innecessàries que, amb tota seguretat, no anaven a aparèixer a les partitures.

L'esquema algorítmic es podria modificar de la següent forma:

```
Require: estructures per guardar una consulta associada a cada clau
for cada nota en la melodia do
  for cada tipus de clau vàlida do
    if la nota és vàlida per a la clau then
      traduir la nota per a la i afegir-la a la traducció de la clau corresponent
    else
      marcar la clau com a no vàlida i descartar la traducció
    end if
  end for
end for
```

Un exemple de traducció per a la melodia «C3 D3 F3» seria: «s0 11 s1» per a c4, «s1 12 s2» per a c5, «s0 11 s1» per a f2, «s1 12 s2» per a f3 i «s2 13 s3» per a f4; que, com s'observa, no inclou totes les claus possibles sinó només les vàlides segons la taula d'equivalències.

És necessària una anotació sobre la taula d'equivalències. No s'han inclòs les notes amb alteració (b). En cas que l'usuari introduísca alguna nota amb alteració, només podrà tractar-se d'un Bb. La traducció es farà sense tindre en compte aquest bemoll i s'inclourà després, ja que la posició geomètrica de la nota no es veu afectada per l'aparició d'una alteració.

### 6.2.3. Segon pas: traducció a *query* lògica

L'objectiu és enviar una consulta lògica composta per diferents seqüències en notació geomètrica. És a dir, es vol saber si existeix alguna de les melodies resultades (una per cada clau vàlida) dins de les partitures. Per tant, és necessària una consulta lògica OR on els elements intermedis són les seqüències.

Com s'explica en la secció 2.2, al demostrador una seqüència s'insereix dins de «[ «]» on els elements estan ordenats en ordre d'aparició geomètrica. Aquestes melodies sempre aniran precedides de la clau que les corresponga.

Com que les melodies no tenen per què aparèixer al principi del tetragrama (on es troba la clau), sinó que poden estar més avançades, la clau no sempre estarà immediatament abans del primer element de la seqüència. Per tant, és necessari indicar que la clau es trobarà abans de la seqüència.

Per com està implementada la cerca de seqüències, açò s'indicarà amb una distància llindar considerada del 20% de la diagonal de la partitura. S'expressarà aquesta distància entre símbols «&». La consulta final tindrà la següent forma:

```
«(clau 2 &20& [seqüència geomètrica corresponent a la clau 1]) ||
(clau 2 &20& [seqüència geomètrica corresponent a la clau 2]) ||
(clau 3 &20& [seqüència geomètrica corresponent a la clau 3]) ... »
fins arribar a l'última clau vàlida.
```

Per realitzar aquesta transformació, només serà necessari recórrer l'estructura que associa cada clau amb una melodia geomètrica i afegir els símbols lògics corresponents. Si hi haguera algun bemoll, s'escriuria flat com a element precedint el nom de la nota alterada. Si aquest bemoll precedira a alguna nota que es trobe en un espai, per l'anàlisi del GT, es descartaria la seqüència que la continguera.



En l'exemple anterior de la seqüència «C3 D3 F3» seria:

«(c4 &20& [s0 11 s1 ]) || (c5 &20& [s1 12 s2 ]) || (f2 &20& [s0 11 s1 ]) || (f3 &20& [s1 12 s2 ]) || (f4 &20& [s2 13 s3 ])».

#### 6.2.4. Traducció sense tonalitat

En el cas en que es vulga realitzar una cerca sense tindre en compte la tonalitat, és necessari realitzar una consulta més extensa encara perquè no es tracta de consultar una única melodia, sinó totes les possibles **transposicions** d'aquesta. Més concretament, al realitzar una cerca sense tonalitat, es tenen en compte no les notes, sinó només les relacions d'interval·les entre cada parella de notes.

En aquestes circumstàncies, llavors, la seqüència «C3 D3 F2» serà igual que «A3 B3 D3». La distància intervàl·lica entre una i altra és d'una sisena major; si es puja una sisena menor la primera seqüència s'obté la segona i si es baixa una sisena menor la segona, s'obté la primera.

Les transposicions són complexes i no es poden fer de forma directa a través de la notació musical. És per això que es transformaran aquestes seqüències musicals en seqüències MIDI, ja que aquesta conversió es fa de forma directa i permet així operar de forma molt més senzilla amb les notes.

La primera seqüència abans esmentada té la seua equivalent en MIDI, que és «48 50 41». Si es vol transportar la melodia en 9 semitons (sisena major), serà suficient amb sumar 9 a tots els elements de la seqüència, resultant en «57 59 50» que es correspon amb la vista anteriorment.

Ara bé, al no tindre en compte la tonalitat, es vol buscar la melodia començant a partir de qualsevol nota. Això és, a partir de totes les transposicions possibles semitò a semitò. S'haurà de transformar eixa seqüència MIDI en un conjunt de seqüències transportades, passant per tots els semitons. Per a això, només cal restar i/o sumar a tots els elements de la seqüència MIDI original cada volta una unitat.

Per tal d'evitar una sobrecàrrega d'informació innecessària, aquest procés tindrà uns límits de transposició: el rang de notes màxim del manuscrit de Vorau, de E2 (40 en MIDI) a D5 (74).

També es descarten les seqüències que continguin els números MIDI (extrets manualment) de notes que estiguen alterades però que no siguin un B $\flat$ .

Després d'haver obtingut totes les melodies possibles, l'últim pas és senzill. Només caldrà aplicar l'algoritme de traducció de notació geomètrica a *queries* lògiques (segon pas de l'algoritme de traducció, secció 6.2.3). Al final de la traducció, es tindrà una consulta lògica geomètrica com la de la traducció amb tonalitat, però considerablement més llarga. Per a l'exemple anterior de «C3 D3 F2» o «A3 B3 D3» l'equivalència seria:

```
«(f4 &20& [s2 13 s0 ]) || (c5 &20& [12 s2 10 ]) || (f3 &20& [12 s2 10 ]) ||  
(f4 &20& [13 s3 11 ]) || (c4 &20& [s2 13 s0 ]) || (c5 &20& [s3 14 s1 ]) ||  
(f2 &20& [s2 13 s0 ]) || (f3 &20& [s3 14 s1 ]) || (f4 &20& [s4 15 s2 ]) ||  
(c3 &20& [12 s2 10 ]) || (c4 &20& [13 s3 11 ]) || (c5 &20& [14 s4 12 ]) ||  
(f1 &20& [12 s2 10 ]) || (f2 &20& [13 s3 11 ]) || (f3 &20& [14 s4 12 ]) ||  
(f4 &20& [15 s5 13 ]) || (c2 &20& [12 s2 10 ]) || (c3 &20& [13 s3 11 ]) ||  
(c4 &20& [14 s4 12 ]) || (c5 &20& [15 s5 13 ]) || (f1 &20& [13 s3 11 ]) ||  
(f2 &20& [14 s4 12 ]) || (f3 &20& [15 s5 13 ]) || (c2 &20& [s2 13 s0 ]) ||  
(c3 &20& [s3 14 s1 ]) || (c4 &20& [s4 15 s2 ]) || (f1 &20& [s3 14 s1 ]) ||  
(f2 &20& [s4 15 s2 ]) || (c1 &20& [s2 13 flat s0 ]) ||  
(c2 &20& [s3 14 flat s1 ]) || (c3 &20& [s4 15 flat s2 ]) ||  
(f1 &20& [s4 15 flat s2 ]) || (c1 &20& [13 s3 11 ]) ||  
(c2 &20& [14 s4 12 ]) || (c3 &20& [15 s5 13 ]) || (f1 &20& [15 s5 13 ]) ||  
(c1 &20& [s3 14 s1 ]) || (c2 &20& [s4 15 s2 ]) || (c1 &20& [s4 s5 s2 ]) »
```

### 6.2.5. Introducció de la traducció a la web

Per poder integrar la traducció dins de la web i que siga efectiva, gràcies a l'estructura que s'ha preparat prèviament, només ha sigut necessari aplicar l'algoritme de traducció dins d'una funció JavaScript que permet el processament de forma eficaç quan s'activa el botó *Convert*.



---

---

## CAPÍTOL 7

# Proves i avaluació

---

La qüestió a abordar dins d'aquest capítol és l'avaluació dels canvis introduïts respecte a la forma de generar les consultes. És a dir, determinar si el fet de permetre consultes a partir de seqüències en notació musical, juntament amb la traducció d'aquestes consultes sense tindre en compte el to original en el qual van ser introduïdes, suposa alguna milloria en l'efectivitat del buscador; això és si els resultats de les cerques són millors.

A més, també es compararan les diferents formes de recuperació d'informació per tractar d'esbrinar quina obté millors resultats en combinació amb les *queries*.

Primer, serà necessari obtenir una quantitat raonable de consultes que s'enviaran al demostrador per poder tindre una referència més o menys representativa. Després, es podrà mesurar la qualitat dels resultats depenent de la tecnologia emprada per recuperar la informació (tres tècniques diferents) conjuntament amb el tipus de consultes que s'han realitzat (amb i sense tonalitat) per valorar el funcionament del buscador.

### 7.1 Obtenció de melodies

---

Com s'ha dit, per tal de realitzar l'avaluació, és necessari enviar una quantitat gran de consultes de melodies al buscador. Aquestes melodies poden ser de qualsevol tipus i és per això que, per tal de simular un entorn realista de proves (ja que el manuscrit data aproximadament de l'any 1450 [18]) on es mantinguen uns criteris d'estil, s'han fet servir melodies que van ser utilitzades per entrenar la ferramenta desenvolupada a [28]. Es tracta d'una aplicació d'inferència gramatical en la qual s'utilitzen gramàtiques regulars que poden ser emprades per generar melodies de forma automàtica.

Per entrenar l'aplicació s'utilitzaren melodies en diferents estils. D'aquestes, s'han agafat 100 en estil gregorià. Es troben en notació implícita, que codifica l'altura i la durada de cada una de les notes que la formen. La durada es representa amb una lletra segons la figura musical que represente (n una negra, c una corxera, s una semicorxera...). L'altura, a partir d'un número del 1 (nota més greu) al 43 (nota més aguda) [28]. Si es tracta d'un silenci, el valor serà 0.

El canvi d'una unitat a altra representa el canvi de semitò entre una nota i altra, el salt d'octava és el salt en 12 unitats; en la taula 7.1 es troben les equivalències entre notació implícita i notació musical. Un exemple de melodia i la seua equivalència en notació implícita es pot vore a la figura 7.1.

Altura implícita	Nota equivalent
1	B0 $\flat$
2	B0
3	C1
4	D1 $\flat$
5	D1
6	E1 $\flat$
7	E1
8	F1
9	G1 $\flat$
10	G1
11	A1 $\flat$
12	A1
	...
40	D5 $\flat$
41	D5
42	E5 $\flat$

**Taula 7.1:** Taula d'equivalències entre notació implícita i notació musical.



**Figura 7.1:** Exemple de l'equivalència entre les notes representades en un pentagrama i la notació implícita (amb durades i altura).

## 7.2 Processament de les dades

Degut a la notació implícita, cadascuna de les notes contingudes en les 100 melodies estan formades per dos camps. En aquest cas, només té rellevància el camp on s'indica l'altura de les notes, ja que al buscador s'introdueixen seqüències de notes on no s'indica cap mesura de temps, per tant caldrà extraure només la informació rellevant sobre el to de les notes.

D'altra banda, aquestes melodies són massa llargues i resulten un problema a l'hora de realitzar consultes, ja que intentar buscar seqüències extenses suposa restringir molt la cerca i afectarà de forma negativa al funcionament del demostrador; no seran uns resultats fiables. Una opció seria acurtar-les, però és difícil determinar quina llargària exacta haurien de tindre aquestes melodies per suposar una bona mostra d'avaluació ja que, si s'acurten molt, les consultes passaran a ser massa laxes, es recuperaran major quantitat d'imatges i es generaran també resultats de poca confiança.

Aleshores, l'objectiu és processar les dades per arribar a una major quantitat de seqüències (si fora possible) que siguin d'una longitud que *contamine* el mínim possible els resultats i que a més suposen respostes per part del demostrador en la majoria dels casos; per això s'han generat melodies amb un cert estil musical similar al del manuscrit i s'han descartat seqüències que o bé sobrepassaven o bé no aplegaven a un llindar de freqüència d'aparició.

### 7.2.1. Extracció de seqüències musicals

A partir de les 100 melodies originals, després d'haver eliminat les dades que indicaven la durada de les notes, s'ha realitzat un processament d'aquestes per extraure seqüències amb compromís entre la seua longitud i la probabilitat que tenen d'aparèixer al manuscrit d'aquesta forma:

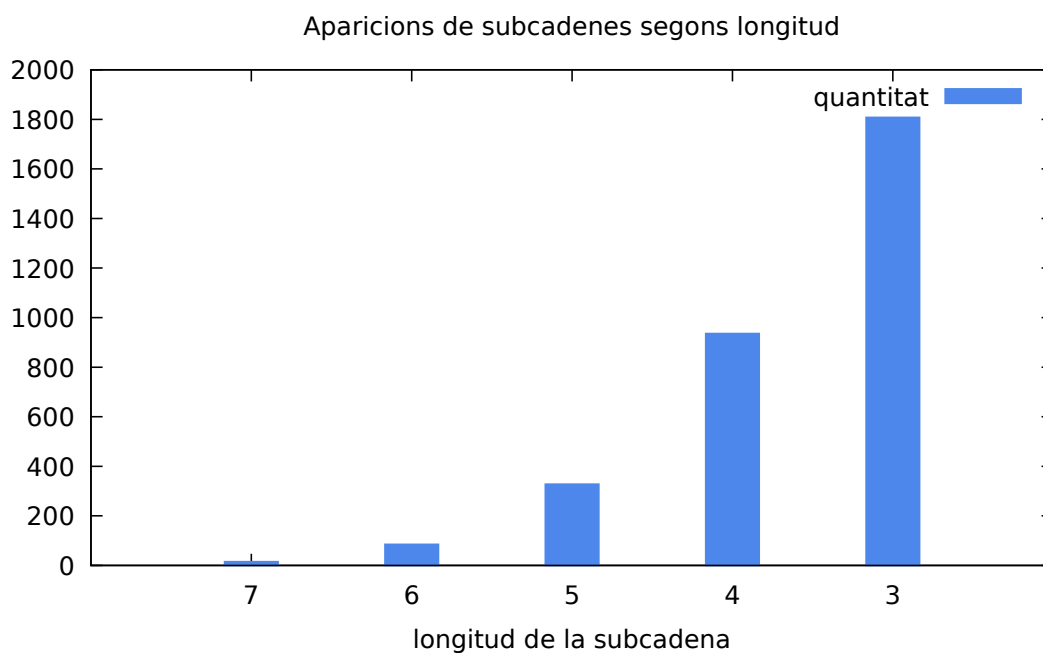
**1. Extracció de subcadenaes** Es recorre el document per extreure totes les subcadenaes de longitud  $n$ , per a  $3 \leq n \leq 9$ , que passaran a ser ara les seqüències.

**2. Eliminació de silencis** Es descarten totes les seqüències que contenen algun 0, que representa els silencis.

**3. Descart de fragments** Es calcula el nombre d'instàncies de cada seqüència per computar un paràmetre d'interès per a cadascuna. Aquest paràmetre és el producte de la longitud del fragment i la seua freqüència d'aparició. Després, s'eliminen els fragments que no tinguen un paràmetre d'interès suficientment gran (que no arribe a 30).

**4. Ordenació del resultat** Finalment, es té el document definitiu on cada fragment s'ordena per freqüència d'aparició.

El resultat d'aquest procés són 185 seqüències diferents amb longituds compreses entre 3 i 7 notes. Cada seqüència apareix com a mínim 5 i com a màxim 100 voltes en les 100 melodies originals.



**Figura 7.2:** Histograma que mostra la quantitat d'aparicions en les 100 melodies originals de les subcadenaes extretes segons la seua longitud.

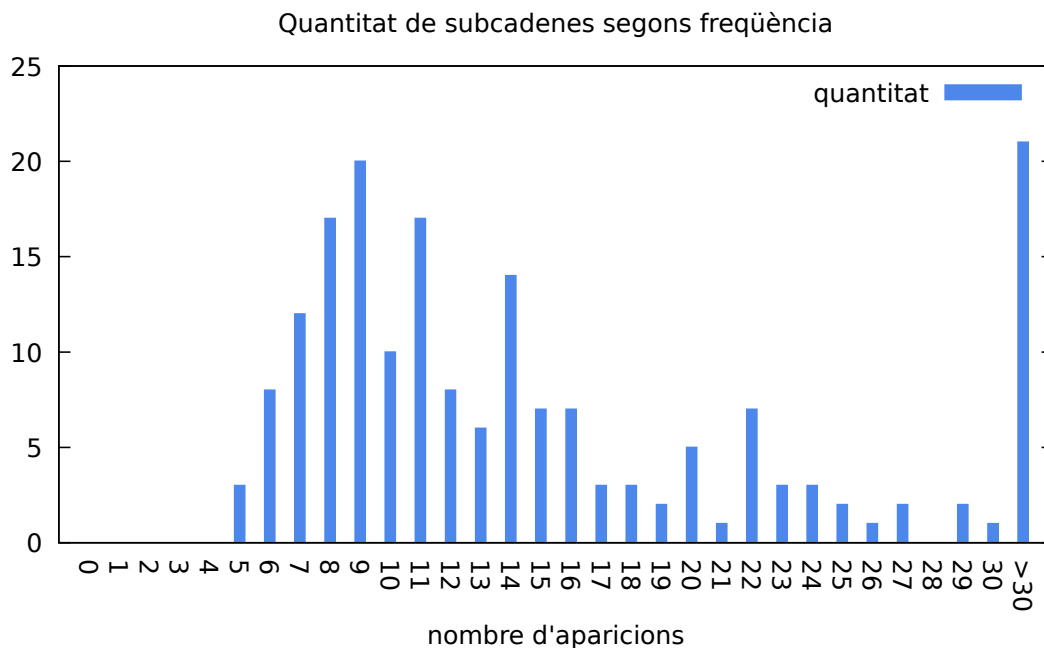


Figura 7.3: Histograma que mostra la quantitat de subcadena extreptes que apareixen en les 100 melodies originals segons la seua freqüència d'aparició.

### 7.3 Obtenció de consultes

Per tal de poder consultar al buscador, serà necessari transformar aquestes melodies, primer en notació musical i després en consultes lògiques que són les que s'enviaran. Després d'haver extret la durada de les melodies, només queden les altures, que tenen una equivalència amb la notació musical. D'aquesta forma es realitza la transformació directa a partir de la taula sense cap tipus de càlcul extra necessari (es pot consultar el codi de la transformació a l'apèndix A.4.1). Per clarificar aquest procés es pot trobar un exemple gràfic a la figura 7.4.

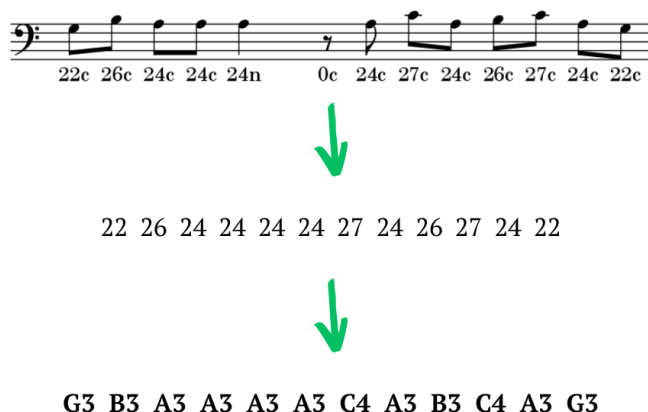


Figura 7.4: Exemple del procés de transformació d'una melodia en notació implícita a una melodia en notació musical.

Quan ja s'ha realitzat la conversió a notació musical, resta fer la traducció a consultes geomètriques. Aquest procés és el mateix que l'explicat en l'apartat de codi A.4.1. Com

que l'objectiu és valorar si les noves formes de consulta al demostrador suposen alguna diferència, serà necessari comptar amb les melodies en el format dels dos tipus de consultes diferents (amb i sense tonalitat). D'aquesta forma es podran enviar individualment al demostrador per comparar els resultats obtinguts.

Una vegada s'hagen aplicat les traduccions corresponents, es tindran dos grups de consultes: les consultes *amb tonalitat* (tenint en compte el to original en el qual van ser escrites) i les consultes *sense tonalitat* (sense tindre-ho en compte). Dins de cada grup, hi haurà una consulta per cadascuna de les 185 melodies. Aquestes consultes, com cabria esperar, seran encadenaments de seqüències lògiques AND/OR. Una vegada s'obtinguen el documents (partitures) tornats pel recuperador d'informació per a cada un dels grups de consultes, es podran avaluar i comparar els resultats.

## 7.4 Tipus de detecció de les seqüències musicals

---

Dins d'aquesta avaluació, també s'ha experimentat amb diferents aproximacions a l'hora de detectar consultes de seqüències musicals en les partitures manuscrites.

S'han provat tres alternatives de detecció diferents sciteReport: a) basada en les posicions lògiques de les deteccions indexades, b) basada en les posicions lògiques l'ordre dels quals és consistent amb la seua localització geomètrica i c) basada només en les posicions geomètriques de les deteccions indexades. Cadascuna d'aquestes aproximacions es descriu breument a continuació:

### 7.4.1. Posició lògica: LPS (*Logical-position-based Spotting*)

Aquesta aproximació usa les posicions lògiques per realitzar el procés de cerca. D'aquesta forma, donada una consulta melòdica  $q_1, q_2, q_3, \dots, q_n$ , a partir de les deteccions de PrIx de cada imatge, primer s'observa la detecció l'etiqueta de la qual és  $\sigma_k = q_1$  i després, des de la posició corresponent endavant, es busquen les deteccions amb posicions lògiques consecutives que tenen un símbol associat que es corresponga amb la seqüència consultada.

### 7.4.2. Posició lògica restringida: RLPS (*Restricted-logical-position-based Spotting*)

Les posicions lògiques de PrIx permeten que múltiples instàncies del mateix símbol estiguen a la mateixa (tetragrama) regió de la imatge. Al contrari que en els documents de text, les repeticions de símbols són extremadament comuns en partitures. Malgrat això, la posició lògica  $k$  anomenada anteriorment, resulta ser prou *fràgil* en la pràctica [29].

Aquest tipus de detecció té el mateix fonament que l'anterior, amb la diferència que, a més d'haver-se de correspondre els símbols associats a les posicions lògiques amb la consulta, les localitzacions han de ser consistents.

### 7.4.3. Posició geomètrica: GPS (*Geometrical-position-based Spotting*)

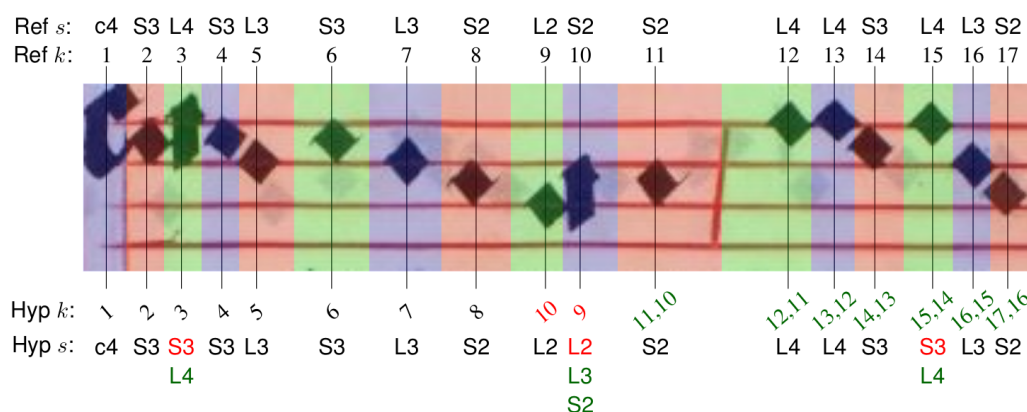
Per poder gestionar aquesta *fragilitat* de les posicions lògiques, aquesta detecció proposa una alternativa complexa i menys refinada. El mètode ignora la posició indexada  $k$  i només es basa en la informació geomètrica donada pel *Bounding Box* (BB)  $b$  de cada detecció



indexada. És a dir, es basa en l'ordenació de les *caixes* geomètriques que envolten cada nota detectada per determinar l'ordre dels elements trobats.

#### 7.4.4. Comparativa de les alternatives

A la figura 7.5 es pot veure com la tècnica LPS s'equivocaria si volguera detectar una seqüència com «[L3 S2 L2 S2 S2]» perquè no existeixen cinc posicions lògiques consecutives ( $k$ ) tals que els valors  $v$  de les deteccions es corresponguen amb eixa seqüència. GPS ignora les posicions lògiques, sinó que ordena les deteccions del pentagrama en ordre ascendent segons els marges esquerres dels seus BBs. D'aquesta forma una de les subseqüències de la llista a l'exemple es correspon amb la consulta anterior.



**Figura 7.5:** Imatge d'un tetragrama amb les etiquetes de referència i les posicions de les notes representades (dalt). També es troben les etiquetes de les deteccions indexades juntament amb les posicions lògiques (baix). Les etiquetes de les deteccions indexades i les posicions lògiques en roig denoten discrepàncies amb les referències corresponents, mentre que les negres no. En verd s'escriuen les hipòtesis alternatives de les etiquetes i les posicions lògiques. Les zones rectangulars colorades representen les posicions geomètriques (*bounding boxes*) associades a cada detecció indexada. Font: [29].

## 7.5 Mesurant els resultats

Una volta s'han executat les consultes al demostrador i s'han obtingut els resultats de les cerques, cal plantejar-se com s'avaluaran aquests resultats. L'efectivitat dels sistemes de recuperació d'informació generalment és mesurada utilitzant els estàndards de *recall* i *interpolated precision* [30].

Dins del marc de treball de la indexació probabilística, per a una consulta donada  $q$  d'un conjunt de consultes  $Q$  i un llindar de rellevància  $\tau$ , el *recall*  $\rho(q, \tau)$  mesura la proporció de regions de les imatges (tetragrames) de la col·lecció que han sigut correctament recuperades pel sistema (també anomenades *hits*) respecte del nombre total de regions rellevants de la col·lecció. El paràmetre de *Precision*  $\pi'(q, \tau)$ , d'altra banda, és la proporció de *hits* respecte del nombre total de regions recuperades (ja siguin correctes o incorrectes) [30, 14].

D'altra banda, es defineix a  $r(q)$  com el nombre de regions rellevants per a la consulta (regions a les quals sí està escrita la melodia consultada), a  $d(q, \tau)$  com el nombre de regions recuperades pel sistema amb una rellevància llindar  $\tau$  i a  $h(q, \tau)$  com el nombre de regions detectades que eren rellevants (els esmentats *hits*) [14]. Les fórmules per calcular el *recall* i la *precision* són:

$$\rho(q, \tau) = \frac{h(q, \tau)}{r(q)} \quad (7.1)$$

$$\pi'(q, \tau) = \frac{h(q, \tau)}{d(q, \tau)} \quad (7.2)$$

### 7.5.1. Corba R-P

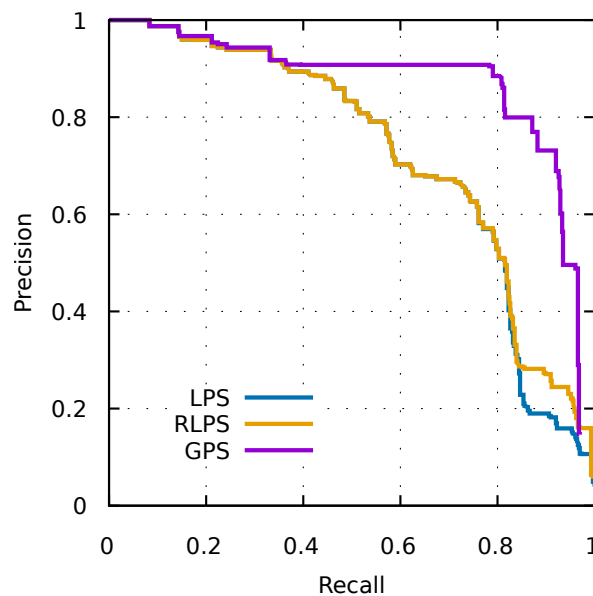
Per cada consulta, com que es tracta d'una recuperació d'informació puntuada (segons la rellevància dels documents), es pot traçar una corba de la relació *recall-precision* per al conjunt de documents obtinguts. Aquests tipus de corbes tenen una forma distintiva (i en ocasions no desitjada) de serra. La forma estàndard d'eliminar aquest efecte és utilitzant la *interpolated precision* [30], que es defineix com:

$$\pi(\rho) = \max_{\rho': \rho' \geq \rho} \pi'(\rho') \quad (7.3)$$

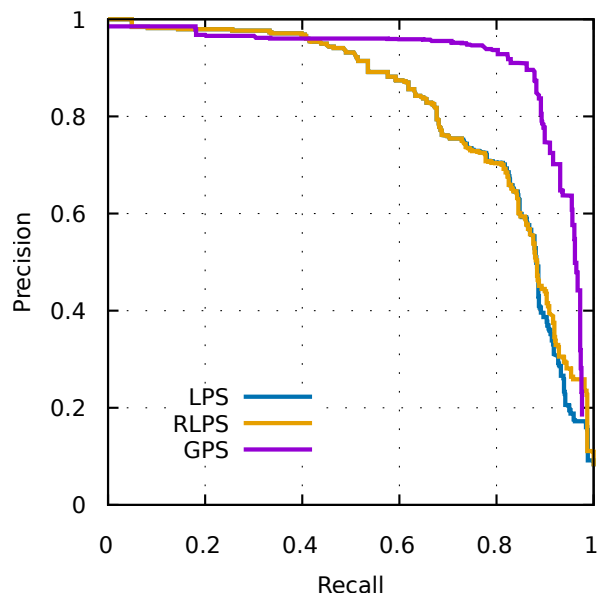
Els bons sistemes de cerca i recuperació haurien d'assolir una *recall* i una *precision* elevades per a un rang ampli de valors de  $\rho$  [4]. Una forma acceptada per mesurar açò és calcular l'àrea baix la corba R-P, referida com *average precision* (AP), quan més alta, millor serà el sistema [31]. Matemàticament es calcula:

$$\bar{\pi}_q = \int_0^1 \pi_q(\rho) d\rho \quad (7.4)$$

### 7.5.2. Resultats de les mètriques



**Figura 7.6:** Gràfica *Recall-Precision* on es mostren els resultats de R-P (una corba per cada tipus de detecció) per a les consultes **amb tonalitat**.



**Figura 7.7:** Gràfica *Recall-Precision* on es mostren els resultats de R-P (una corba per cada tipus de detecció) per a les consultes **sense tonalitat**.

	Amb tonalitat AP	Sense tonalitat AP
LPS	0,722	0,815
RLPS	0,733	0,820
GPS	0,867	0,908

**Taula 7.2:** Resultats d'*average precision* per a cadascuna de les combinacions entre els tipus de detecció i els tipus de consulta.

Després d'haver processat les melodies amb el procediment exposat anteriorment, s'han realitzat els càlculs de les mètriques. Per poder obtenir un resultat per cada tipus de detecció en combinació amb el tipus de cerca, s'han amigitjanat els resultats de cada grup de *queries*.

Així, per una banda, la taula 7.1 mostra el resultat d'*average precision* per a totes aquestes combinacions, on l'ús de GPS juntament amb les consultes sense tonalitat obtenen el millor resultat. Per altra banda, en les figures 7.6 i 7.7 es corrobora aquest fet, on s'observa que la major àrea baix les corbes R-P és la corresponent a la de la funció de GPS amb *queries* sense tonalitat.

---

---

## CAPÍTOL 8

# Conclusions

---

Arribades a aquesta part final del projecte, cal posar en discussió i analitzar el treball realitzat durant aquest temps. Primer, en relació als objectius proposats a l'inici, valorant si ha sigut factible la seua consecució i examinar en quina mesura s'han acomplert.

Després, es farà una revisió dels resultats obtinguts en les proves fetes i es detallaran les conclusions extretes dels mateixos, suposant el context teòric donat al llarg dels capítols anteriors.

Finalment, es plantejaran possibles ampliacions de les tasques ací realitzades, emmarcant-les en la implementació i creació de noves ferramentes útils per a la detecció i recuperació de símbols musicals a partitures manuscrites antigues.

### 8.1 Assoliment dels objectius

---

L'objectiu general proposat per al projecte era la implementació de noves funcionalitats per al demostrador del motor de cerca d'indexació probabilística per a textos manuscrits creat pel PRHLT *Research Center*. A continuació es desglossa l'anàlisi per subobjectius derivats d'aquest objectiu general.

En primer lloc, es volien afegir noves entrades en format musical a través de la web del demostrador. Com es pot veure en les seccions 6.1 i 6.1.4, s'han cobert les tres formes proposades d'entrada i fins i tot s'han pogut implementar elements que milloren la interacció a nivell d'usuari (com el botó de gravació o la reproducció de les notes al tocar les tecles).

Una de les parts més costoses relacionades amb les funcionalitats ha sigut la introducció dels canvis dins de la web. Com que el codi amb el qual es comptava de partida no era de creació original, a vegades ha resultat difícil entendre el funcionament dels programes i l'estructura del codi, ja que moltes vegades no hi existia una documentació associada al mateix o no ha sigut possible comunicar-se amb les persones que varen crear aquest programari per poder resoldre dubtes.

Encara així, finalment s'ha pogut unificar el codi propi amb l'anterior per, encara que fos de forma superficial, poder crear un entorn de proves a mode de pilot el qual es podria plantejar ampliar més endavant.

El demostrador provisional pot utilitzar-se a través de <http://heisenberg.prhlt.upv.es:8082/musica/> accessible amb connexió a la VPN de la UPV. Aquest encara es està en desenvolupament i en breus s'implementarà en el buscador oficial que es troba en la màquina *prhlt-carabela* (<http://prhlt-carabela.prhlt.upv.es/musica/>).

Altre dels objectius, la creació d'un algoritme de traducció de notació musical a consultes compatibles amb el motor de cerca, s'ha vist completat adequadament a la secció 6.2. A

més, gràcies a l'anàlisi previ de les dades del *ground truth* i del problema a la secció 4, en la posterior secció 6.2.4 ha sigut possible implementar un esquema de traducció a partir del primer en el qual no es tenia en compte la tonalitat a l'hora de realitzar la consulta. D'aquesta forma, s'han pogut assolir aquests objectius sense cap inconvenient.

Per últim, el treball d'implementació del codi web i creació dels algorismes ha permès la construcció d'un entorn per realitzar proves de rendiment de les quals es parlarà més endavant.

A aquesta secció, per tant, es pot concloure que tots els objectius plantejats a l'inici del projecte han estat coberts de forma satisfactòria.

## 8.2 Sobre les proves i l'avaluació

---

Degut a la comparació de les alternatives de detecció en la secció 7.4.4, no és sorprenent que la GPS siga la que millors resultats ofereix. Malgrat això, era necessari vore com responia aquesta forma de detecció de seqüències davant de les diferents alternatives de consulta (amb i sense tonalitat) abans de treure cap conclusió. Davant les dades obtingudes, comparant les consultes amb i sense tonalitat, no hi ha cap dubte que l'opció de cerca que millors resultats dona és la realitzada a través de consultes sense tonalitat.

Aquest resultat pot semblar sorprenent ja que, al ser les consultes que es realitzen més llargues, cabria esperar que el sistema augmentara la quantitat d'errors i disminuïra la taxa de satisfacció.

Després d'haver processat i analitzat les 100 melodies inicials, es posa en relleu l'existència de patrons melòdics que es repeteixen amb una freqüència considerable a les partitures. D'aquesta forma, quan es consulta un d'aquests patrons freqüents sense tindre en compte la tonalitat, es detecten una gran quantitat de correspondències.

D'altra banda, aquests resultats podrien estar relacionats amb el processament dins de la traducció que es fa de les melodies. A partir de l'anàlisi del GT s'ha observat que hi havia dades innecessàries (que mai obtindrien un resultat útil al consultar-les) que es podien eliminar abans d'enviar-les al motor de detecció i recuperació. Així, s'han estalviat possibles resultats erronis que perjudicarien l'actuació del sistema.

Com a conclusió s'extreu que la detecció de seqüències (musicals) resulta ser prou important dins dels textos manuscrits que es tracten de partitures; existeixen patrons melòdics recurrents dins d'elles; es podria dir que de forma similar a les paraules dels textos manuscrits. A més, el fet de poder fer consultes a través de notació musical no només millora la interacció, sinó que pot també vincular-se a l'obtenció de bons resultats ja que ofereix una cerca amb menys biaix i més ampla, amb una probabilitat major de trobar el resultat desitjat.

Aquest concepte de *paraules musicals* podria suposar una nova línia de treball per emprar-se en la detecció de seqüències musicals, juntament amb l'ús de notació musical per realitzar les consultes.

---

## 8.3 Línies de treball futures

---

Aquest projecte ha permès ampliar les funcionalitats del demostrador web del PRHLT *Research Center* per a la detecció i recuperació de símbols i seqüències musicals en partitures manuscrites antigues, així com obrir un nou front d'investigació sobre el tipus de consultes i les formes de detecció musical que es poden realitzar en aquest mena de documents. De cara a futurs treballs plantegen les següents propostes:

Per un costat, la implementació web completa (no únicament com a entorn de proves) del demostrador. Aquesta implementació inclouria fusionar aquest TFG, que tracta l'entrada de dades, amb el de l'alumna Carmen Bueno Hurtado (*Renderización acústica de patrones melódicos localizados en grandes colecciones de imágenes de partituras manuscritas históricas*, esmentat en la introducció), que tracta l'eixida.

Per altre, la investigació sobre noves formes de detecció de seqüències musicals, aprofitant el tipus de consultes sense tonalitat i a partir de les tecnologies fins ara emprades. En el futur es podria explorar la possibilitat de generar índex probabilístics de seqüències de notes (com a *paraules musicals* esmentades abans), en compte de notes individuals. D'aquesta forma, l'esforç de càlcul que es realitza per a la cerca de seqüències es traslladaria a la fase *off line* de construcció dels PrIx. Al eliminar aquest cost computacional de la fase *on line*, es podrien fer cerques de seqüències en col·leccions molt més grans (milions d'imatges en compte de cents com en VORAU-253). Això també pot induir a la possibilitat d'implementar cerques de seqüències més llargues emprant (si fos necessari) prefixes, sufixes, comodins i cerca aproximada, entre altres.



# Bibliografia

---

- [1] E. Davies, "Chapter 1 - vision, the challenge," in *Computer Vision (Fifth Edition)*, E. Davies, Ed. Academic Press, 2018, pp. 1–15. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128092842000010>
- [2] A. K. Jain, P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, 2000.
- [3] V. Romero, A. H. Toselli, and E. Vidal, *Multimodal Interactive Handwritten Text Transcription*. Singapur, Singapur: World Scientific Publishing, 2012.
- [4] E. Vidal, A. H. Toselli, and J. Puigcerver, "Lexicon-based probabilistic indexing of handwritten text images," *Neural Comput. Appl.*, 2023.
- [5] J. Calvo-Zaragoza, A. H. Toselli, E. Vidal, and J. A. Sanchez, "Music symbol sequence indexing in medieval plainchant manuscripts," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 2019.
- [6] M. Villarreal and J. A. Sanchez, "Handwritten music recognition improvement through language model re-interpretation for mensural notation," in *2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. IEEE, 2020.
- [7] A. P. Giotis, G. Sfikas, B. Gatos, and C. Nikou, "A survey of document image word spotting techniques," *Pattern Recognit.*, vol. 68, pp. 310–332, 2017.
- [8] "PRHLT research center," <https://www.prhlt.upv.es/>, accessed: 2023-6-28.
- [9] J. Calvo-Zaragoza, A. H. Toselli, and E. Vidal, "Probabilistic music-symbol spotting in handwritten scores," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. IEEE, 2018.
- [10] R. Malik, P. P. Roy, U. Pal, and F. Kimura, "Handwritten musical document retrieval using music-score spotting," in *2013 12th International Conference on Document Analysis and Recognition*. IEEE, 2013.
- [11] L. Likforman-Sulem, A. Zahour, and B. Taconet, "Text line segmentation of historical documents: a survey," *Int. J. Doc. Anal. Recognit.*, vol. 9, no. 2-4, pp. 123–138, 2007.
- [12] G. Louloudis, B. Gatos, I. Pratikakis, and C. Halatsis, "Text line and word segmentation of handwritten documents," *Pattern Recognit.*, vol. 42, no. 12, pp. 3169–3183, 2009.
- [13] A. H. Toselli, E. Vidal, V. Romero, and V. Frinken, "Hmm word graph based keyword spotting in handwritten document images," *Information Sciences*, vol. 370-371, pp. 497–518, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0020025516305461>



- [14] A. H. Toselli, J. Puigcerver, and E. Vidal, "Probabilistic indexing for information search and retrieval in large collections of handwritten text images," Jun. 2023, to be published in Springer Verlag.
- [15] E. Noya-García, A. H. Toselli, and E. Vidal, "Simple and effective multi-word query spotting in handwritten text images," in *Pattern Recognition and Image Analysis*. Cham: Springer International Publishing, 2017, pp. 76–84.
- [16] E. Vidal, A. H. Toselli, and J. Puigcerver, "Text search and information retrieval in large historical collections of untranscribed manuscripts," <https://www.prhlt.upv.es/~evidal/tmp/icdar19keynoteEVidal2p.pdf>, accessed: 2023-7-1.
- [17] J. Puigcerver, "A probabilistic formulation of keyword spotting," *PhD thesis*, 2018.
- [18] L. Quirós, E. Vidal, J. A. Sánchez, and M. Villarreal, "Vorau abbey library cod. 253 dataset for document layout analysis," 2021.
- [19] "Docker overview," <https://docs.docker.com/get-started/overview/>, jun 2023, accessed: 2023-6-20.
- [20] R. T. Fielding, *Architectural styles and the design of network-based software architectures*. University of California, Irvine, 2000.
- [21] "NumPy documentation," <https://numpy.org/doc/stable/>, accessed: 2023-6-26.
- [22] "Web MIDI API," <https://webaudio.github.io/web-midi-api/>, accessed: 2023-6-26.
- [23] J. Rothstein, *MIDI: A Comprehensive Introduction*. Londres, Inglaterra: Oxford University Press, 1992.
- [24] D. M. Huber, *The MIDI manual the MIDI manual: A practical guide to MIDI within modern music production*, 4th ed. Londres, Inglaterra: Routledge, 2020.
- [25] D. Back, "Standard MIDI-File format spec. 1.1, updated," <http://www.music.mcgill.ca/~ich/classes/mumt306/StandardMIDIfileformat.html>, 1992, accessed: 2023-6-20.
- [26] U. Michels, *Atlas de musica 1*. Alianza, 1999.
- [27] W. Apel, Ed., *Harvard dictionary of music*, 2nd ed. Londres, Inglaterra: Harvard University Press, 1969.
- [28] P. P. Cruz-Alcazar and E. Vidal-Ruiz, "Modeling musical style using grammatical inference techniques: a tool for classifying and generating melodies," in *Proceedings Third International Conference on WEB Delivering of Music*. IEEE Comput. Soc, 2004.
- [29] A. Menárguez Box, M. Villarreal Ruiz, A. H. Toselli, and E. Vidal Ruiz, "Search and information retrieval in historical music scores," report in preparation.
- [30] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, Inglaterra: Cambridge University Press, 2012.
- [31] M. Zhu, "Recall, precision and average precision," *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo*, vol. 2, no. 30, p. 6, 2004.
- [32] E. Herrera, *Teoría musical Armonía moderna I*. Antoni Bosch Editor, 1990.
- [33] ISO Central Secretary, "Acoustics — standard tuning frequency (standard musical pitch)," International Organization for Standardization, Tech. Rep. ISO 16:1975, 1975.

- [34] A. J. Ellis, *On the Musical Scales of Various Nations*. Journal of the Society of Arts, 1885.
- [35] D. de Pedro, *Teoría completa de la música*. Real Musical, 2014.



---

---

# APÈNDIX A

## Codi

---

Dins d'aquest apèndix s'inclou tot el codi desenvolupat amb la finalitat de completar els objectius i les funcionalitats descrites al llarg del projecte.

### A.1 Codi CSS

---

Codi escrit per canviar l'estil dels elements HTML segons l'estat en que es troben i facilitar així la visualització i interacció amb ells.

#### A.1.1. Tecles del piano virtual

```
1 .white {
2   height:134px;
3   width:32px;
4   background-color: white;
5   margin-left: -3px;
6   margin-right: -3px;
7 }
8
9 .white.active {
10  background-color: #78b5f7;
11 }
12
13 .black {
14  height:80px;
15  width:20px;
16  background-color: black;
17  position: relative;
18  top: -27px;
19  margin-left: -14px;
20  margin-right: -14px;
21  z-index: 1;
22 }
23
24 .black.active {
25  background-color: #1982f4;
26 }
27
28 .no-valid {
29  background-color: #b9b9b9;
30 }
```

## A.1.2. Botó de gravació

```

1  .record {
2    width: 20px;
3    height: 20px;
4    font-size: 0;
5    background-color: red;
6    border: 0;
7    border-radius: 20px;
8    margin: 4px;
9    outline: none;
10 }
11
12 .record:hover{
13   background-color: blue;
14 }
15
16 .rec-off{
17   background-color: darkred;
18 }
19
20 .rec-off:hover{
21   background-color: red;
22 }
23
24 .rec-on{
25   animation-name: pulse;
26   animation-duration: 1.5s;
27   animation-iteration-count: infinite;
28   animation-timing-function: linear;
29 }
30
31 .rec-on:hover{
32   background-color: darkred;
33 }
34
35 @keyframes pulse{
36   0%{
37     box-shadow: 0px 0px 5px 0px rgba(173,0,0,.3);
38   }
39   65%{
40     box-shadow: 0px 0px 5px 5px rgba(173,0,0,.3);
41   }
42   90%{
43     box-shadow: 0px 0px 5px 5px rgba(173,0,0,0);
44   }
45 }

```

## A.2 Codi HMTL

Creat per dissenyar i posicionar els elements amb els quals interaccionarà l'usuari

### A.2.1. Diàleg del piano

Codi de les tecles (botons) del teclat de piano, la barra de cerca, el botó de gravació i les opcions de cerca.

```

1  <!-- Dialog Piano -->
2  <div id="dialog-piano" title="Piano">
3    <label for="input-notes">Melody:</label>

```

```

4 <input id="input-notes" type="text" placeholder="write down or play some
   notes ..."
5 style="width:270px;padding-left:5px;padding-right:5px"></input>
6 <button id = "button-erase">Erase</button>
7 <button id = "button-convert">Convert</button>
8 <label for="mind-key" style="margin-left:2px">Mind key</label>
9 <input type="checkbox" id="mind-key" checked>
10
11 <br><br>
12
13 <button id="button-rec" class="rec-on record"></button>
14 <label for="button-rec" style="font-size:14px">REC</label>
15
16 <br><br>
17 <button id="E2" class="key white"></button>
18 <button id="F2" class="key white"></button>
19 <button id="Gb2" class="key black no-valid"></button>
20 <button id="G2" class="key white"></button>
21 <button id="Ab2" class="key black no-valid"></button>
22 <button id="A2" class="key white"></button>
23 <button id="Bb2" class="key black"></button>
24 <button id="B2" class="key white"></button>
25 <button id="C3" class="key white"></button>
26 <button id="Db3" class="key black no-valid"></button>
27 <button id="D3" class="key white"></button>
28 <button id="Eb3" class="key black no-valid"></button>
29 <button id="E3" class="key white"></button>
30 <button id="F3" class="key white"></button>
31 <button id="Gb3" class="key black no-valid"></button>
32 <button id="G3" class="key white"></button>
33 <button id="Ab3" class="key black no-valid"></button>
34 <button id="A3" class="key white"></button>
35 <button id="Bb3" class="key black"></button>
36 <button id="B3" class="key white"></button>
37 <button id="C4" class="key white"></button>
38 <button id="Db4" class="key black no-valid"></button>
39 <button id="D4" class="key white"></button>
40 <button id="Eb4" class="key black no-valid"></button>
41 <button id="E4" class="key white"></button>
42 <button id="F4" class="key white"></button>
43 <button id="Gb4" class="key black no-valid"></button>
44 <button id="G4" class="key white"></button>
45 <button id="Ab4" class="key black no-valid"></button>
46 <button id="A4" class="key white"></button>
47 <button id="Bb4" class="key black"></button>
48 <button id="B4" class="key white"></button>
49 <button id="C5" class="key white"></button>
50 <button id="Db5" class="key black no-valid"></button>
51 <button id="D5" class="key white"></button>
52 </div>
53
54 <!-- Button open Piano -->
55 <button id = "opener-piano">Open Piano</button>

```

## A.2.2. Notes musicals

Àudios mp3 que s'associen a cadascuna de les tecles (botons) del piano.

```

1 <!-- Music notes -->
2 <audio id="E2" src="views/aitana/notes/E2.mp3"></audio>
3 <audio id="F2" src="views/aitana/notes/F2.mp3"></audio>
4 <audio id="Gb2" src="views/aitana/notes/Gb2.mp3"></audio>

```

```

5 <audio id="G2" src="views/aitana/notes/G2.mp3"></audio>
6 <audio id="Ab2" src="views/aitana/notes/Ab2.mp3"></audio>
7 <audio id="A2" src="views/aitana/notes/A2.mp3"></audio>
8 <audio id="Bb2" src="views/aitana/notes/Bb2.mp3"></audio>
9 <audio id="B2" src="views/aitana/notes/B2.mp3"></audio>
10 <audio id="C3" src="views/aitana/notes/C3.mp3"></audio>
11 <audio id="Db3" src="views/aitana/notes/Db3.mp3"></audio>
12 <audio id="D3" src="views/aitana/notes/D3.mp3"></audio>
13 <audio id="Eb3" src="views/aitana/notes/Eb3.mp3"></audio>
14 <audio id="E3" src="views/aitana/notes/E3.mp3"></audio>
15 <audio id="F3" src="views/aitana/notes/F3.mp3"></audio>
16 <audio id="Gb3" src="views/aitana/notes/Gb3.mp3"></audio>
17 <audio id="G3" src="views/aitana/notes/G3.mp3"></audio>
18 <audio id="Ab3" src="views/aitana/notes/Ab3.mp3"></audio>
19 <audio id="A3" src="views/aitana/notes/A3.mp3"></audio>
20 <audio id="Bb3" src="views/aitana/notes/Bb3.mp3"></audio>
21 <audio id="B3" src="views/aitana/notes/B3.mp3"></audio>
22 <audio id="C4" src="views/aitana/notes/C4.mp3"></audio>
23 <audio id="Db4" src="views/aitana/notes/Db4.mp3"></audio>
24 <audio id="D4" src="views/aitana/notes/D4.mp3"></audio>
25 <audio id="Eb4" src="views/aitana/notes/Eb4.mp3"></audio>
26 <audio id="E4" src="views/aitana/notes/E4.mp3"></audio>
27 <audio id="F4" src="views/aitana/notes/F4.mp3"></audio>
28 <audio id="Gb4" src="views/aitana/notes/Gb4.mp3"></audio>
29 <audio id="G4" src="views/aitana/notes/G4.mp3"></audio>
30 <audio id="Ab4" src="views/aitana/notes/Ab4.mp3"></audio>
31 <audio id="A4" src="views/aitana/notes/A4.mp3"></audio>
32 <audio id="Bb4" src="views/aitana/notes/Bb4.mp3"></audio>
33 <audio id="B4" src="views/aitana/notes/B4.mp3"></audio>
34 <audio id="C5" src="views/aitana/notes/C5.mp3"></audio>
35 <audio id="Db5" src="views/aitana/notes/Db5.mp3"></audio>
36 <audio id="D5" src="views/aitana/notes/D5.mp3"></audio>

```

## A.3 Codi JavaScript

Codi creat per gestionar la interacció entre els elements HTML i el codi de traducció i gestió d'entrades MIDI.

### A.3.1. Dades

Estructures de dades amb les taules d'equivalències i els paràmetres necessaris per a la traducció.

```

1 // Mapa de conversions MIDI-nota
2 const MAX_NOTE = 74;
3 const MIN_NOTE = 40;
4
5 var midi2key = new Map([
6   [40, "E2"], [41, "F2"], [43, "G2"], [45, "A2"], [46, "Bb2"],
7   [47, "B2"], [48, "C3"], [50, "D3"], [52, "E3"], [53, "F3"],
8   [55, "G3"], [57, "Ab3"], [58, "A3"], [59, "Bb3"], [60, "B3"], [62, "C4"],
9   [64, "Db4"], [65, "E4"], [67, "F4"], [69, "G4"], [71, "Ab4"],
10  [72, "A4"], [73, "Bb4"], [74, "B4"], [75, "C5"], [77, "Db5"], [78, "E5"], [80, "F5"], [82, "G5"], [84, "Ab5"], [85, "A5"], [87, "Bb5"], [88, "B5"], [90, "C6"], [92, "Db6"], [93, "E6"], [95, "F6"], [97, "G6"], [99, "Ab6"], [100, "A6"], [102, "Bb6"], [103, "B6"], [105, "C7"], [107, "Db7"], [108, "E7"], [110, "F7"], [112, "G7"], [114, "Ab7"], [115, "A7"], [117, "Bb7"], [118, "B7"], [120, "C8"], [122, "Db8"], [123, "E8"], [125, "F8"], [127, "G8"], [129, "Ab8"], [130, "A8"], [132, "Bb8"], [133, "B8"], [135, "C9"], [137, "Db9"], [138, "E9"], [140, "F9"], [142, "G9"], [144, "Ab9"], [145, "A9"], [147, "Bb9"], [148, "B9"], [150, "C10"], [152, "Db10"], [153, "E10"], [155, "F10"], [157, "G10"], [159, "Ab10"], [160, "A10"], [162, "Bb10"], [163, "B10"], [165, "C11"], [167, "Db11"], [168, "E11"], [170, "F11"], [172, "G11"], [174, "Ab11"], [175, "A11"], [177, "Bb11"], [178, "B11"], [180, "C12"], [182, "Db12"], [183, "E12"], [185, "F12"], [187, "G12"], [189, "Ab12"], [190, "A12"], [192, "Bb12"], [193, "B12"], [195, "C13"], [197, "Db13"], [198, "E13"], [200, "F13"], [202, "G13"], [204, "Ab13"], [205, "A13"], [207, "Bb13"], [208, "B13"], [210, "C14"], [212, "Db14"], [213, "E14"], [215, "F14"], [217, "G14"], [219, "Ab14"], [220, "A14"], [222, "Bb14"], [223, "B14"], [225, "C15"], [227, "Db15"], [228, "E15"], [230, "F15"], [232, "G15"], [234, "Ab15"], [235, "A15"], [237, "Bb15"], [238, "B15"], [240, "C16"], [242, "Db16"], [243, "E16"], [245, "F16"], [247, "G16"], [249, "Ab16"], [250, "A16"], [252, "Bb16"], [253, "B16"], [255, "C17"], [257, "Db17"], [258, "E17"], [260, "F17"], [262, "G17"], [264, "Ab17"], [265, "A17"], [267, "Bb17"], [268, "B17"], [270, "C18"], [272, "Db18"], [273, "E18"], [275, "F18"], [277, "G18"], [279, "Ab18"], [280, "A18"], [282, "Bb18"], [283, "B18"], [285, "C19"], [287, "Db19"], [288, "E19"], [290, "F19"], [292, "G19"], [294, "Ab19"], [295, "A19"], [297, "Bb19"], [298, "B19"], [300, "C20"], [302, "Db20"], [303, "E20"], [305, "F20"], [307, "G20"], [309, "Ab20"], [310, "A20"], [312, "Bb20"], [313, "B20"], [315, "C21"], [317, "Db21"], [318, "E21"], [320, "F21"], [322, "G21"], [324, "Ab21"], [325, "A21"], [327, "Bb21"], [328, "B21"], [330, "C22"], [332, "Db22"], [333, "E22"], [335, "F22"], [337, "G22"], [339, "Ab22"], [340, "A22"], [342, "Bb22"], [343, "B22"], [345, "C23"], [347, "Db23"], [348, "E23"], [350, "F23"], [352, "G23"], [354, "Ab23"], [355, "A23"], [357, "Bb23"], [358, "B23"], [360, "C24"], [362, "Db24"], [363, "E24"], [365, "F24"], [367, "G24"], [369, "Ab24"], [370, "A24"], [372, "Bb24"], [373, "B24"], [375, "C25"], [377, "Db25"], [378, "E25"], [380, "F25"], [382, "G25"], [384, "Ab25"], [385, "A25"], [387, "Bb25"], [388, "B25"], [390, "C26"], [392, "Db26"], [393, "E26"], [395, "F26"], [397, "G26"], [399, "Ab26"], [400, "A26"], [402, "Bb26"], [403, "B26"], [405, "C27"], [407, "Db27"], [408, "E27"], [410, "F27"], [412, "G27"], [414, "Ab27"], [415, "A27"], [417, "Bb27"], [418, "B27"], [420, "C28"], [422, "Db28"], [423, "E28"], [425, "F28"], [427, "G28"], [429, "Ab28"], [430, "A28"], [432, "Bb28"], [433, "B28"], [435, "C29"], [437, "Db29"], [438, "E29"], [440, "F29"], [442, "G29"], [444, "Ab29"], [445, "A29"], [447, "Bb29"], [448, "B29"], [450, "C30"], [452, "Db30"], [453, "E30"], [455, "F30"], [457, "G30"], [459, "Ab30"], [460, "A30"], [462, "Bb30"], [463, "B30"], [465, "C31"], [467, "Db31"], [468, "E31"], [470, "F31"], [472, "G31"], [474, "Ab31"], [475, "A31"], [477, "Bb31"], [478, "B31"], [480, "C32"], [482, "Db32"], [483, "E32"], [485, "F32"], [487, "G32"], [489, "Ab32"], [490, "A32"], [492, "Bb32"], [493, "B32"], [495, "C33"], [497, "Db33"], [498, "E33"], [500, "F33"], [502, "G33"], [504, "Ab33"], [505, "A33"], [507, "Bb33"], [508, "B33"], [510, "C34"], [512, "Db34"], [513, "E34"], [515, "F34"], [517, "G34"], [519, "Ab34"], [520, "A34"], [522, "Bb34"], [523, "B34"], [525, "C35"], [527, "Db35"], [528, "E35"], [530, "F35"], [532, "G35"], [534, "Ab35"], [535, "A35"], [537, "Bb35"], [538, "B35"], [540, "C36"], [542, "Db36"], [543, "E36"], [545, "F36"], [547, "G36"], [549, "Ab36"], [550, "A36"], [552, "Bb36"], [553, "B36"], [555, "C37"], [557, "Db37"], [558, "E37"], [560, "F37"], [562, "G37"], [564, "Ab37"], [565, "A37"], [567, "Bb37"], [568, "B37"], [570, "C38"], [572, "Db38"], [573, "E38"], [575, "F38"], [577, "G38"], [579, "Ab38"], [580, "A38"], [582, "Bb38"], [583, "B38"], [585, "C39"], [587, "Db39"], [588, "E39"], [590, "F39"], [592, "G39"], [594, "Ab39"], [595, "A39"], [597, "Bb39"], [598, "B39"], [600, "C40"], [602, "Db40"], [603, "E40"], [605, "F40"], [607, "G40"], [609, "Ab40"], [610, "A40"], [612, "Bb40"], [613, "B40"], [615, "C41"], [617, "Db41"], [618, "E41"], [620, "F41"], [622, "G41"], [624, "Ab41"], [625, "A41"], [627, "Bb41"], [628, "B41"], [630, "C42"], [632, "Db42"], [633, "E42"], [635, "F42"], [637, "G42"], [639, "Ab42"], [640, "A42"], [642, "Bb42"], [643, "B42"], [645, "C43"], [647, "Db43"], [648, "E43"], [650, "F43"], [652, "G43"], [654, "Ab43"], [655, "A43"], [657, "Bb43"], [658, "B43"], [660, "C44"], [662, "Db44"], [663, "E44"], [665, "F44"], [667, "G44"], [669, "Ab44"], [670, "A44"], [672, "Bb44"], [673, "B44"], [675, "C45"], [677, "Db45"], [678, "E45"], [680, "F45"], [682, "G45"], [684, "Ab45"], [685, "A45"], [687, "Bb45"], [688, "B45"], [690, "C46"], [692, "Db46"], [693, "E46"], [695, "F46"], [697, "G46"], [699, "Ab46"], [700, "A46"], [702, "Bb46"], [703, "B46"], [705, "C47"], [707, "Db47"], [708, "E47"], [710, "F47"], [712, "G47"], [714, "Ab47"], [715, "A47"], [717, "Bb47"], [718, "B47"], [720, "C48"], [722, "Db48"], [723, "E48"], [725, "F48"], [727, "G48"], [729, "Ab48"], [730, "A48"], [732, "Bb48"], [733, "B48"], [735, "C49"], [737, "Db49"], [738, "E49"], [740, "F49"], [742, "G49"], [744, "Ab49"], [745, "A49"], [747, "Bb49"], [748, "B49"], [750, "C50"], [752, "Db50"], [753, "E50"], [755, "F50"], [757, "G50"], [759, "Ab50"], [760, "A50"], [762, "Bb50"], [763, "B50"], [765, "C51"], [767, "Db51"], [768, "E51"], [770, "F51"], [772, "G51"], [774, "Ab51"], [775, "A51"], [777, "Bb51"], [778, "B51"], [780, "C52"], [782, "Db52"], [783, "E52"], [785, "F52"], [787, "G52"], [789, "Ab52"], [790, "A52"], [792, "Bb52"], [793, "B52"], [795, "C53"], [797, "Db53"], [798, "E53"], [800, "F53"], [802, "G53"], [804, "Ab53"], [805, "A53"], [807, "Bb53"], [808, "B53"], [810, "C54"], [812, "Db54"], [813, "E54"], [815, "F54"], [817, "G54"], [819, "Ab54"], [820, "A54"], [822, "Bb54"], [823, "B54"], [825, "C55"], [827, "Db55"], [828, "E55"], [830, "F55"], [832, "G55"], [834, "Ab55"], [835, "A55"], [837, "Bb55"], [838, "B55"], [840, "C56"], [842, "Db56"], [843, "E56"], [845, "F56"], [847, "G56"], [849, "Ab56"], [850, "A56"], [852, "Bb56"], [853, "B56"], [855, "C57"], [857, "Db57"], [858, "E57"], [860, "F57"], [862, "G57"], [864, "Ab57"], [865, "A57"], [867, "Bb57"], [868, "B57"], [870, "C58"], [872, "Db58"], [873, "E58"], [875, "F58"], [877, "G58"], [879, "Ab58"], [880, "A58"], [882, "Bb58"], [883, "B58"], [885, "C59"], [887, "Db59"], [888, "E59"], [890, "F59"], [892, "G59"], [894, "Ab59"], [895, "A59"], [897, "Bb59"], [898, "B59"], [900, "C60"], [902, "Db60"], [903, "E60"], [905, "F60"], [907, "G60"], [909, "Ab60"], [910, "A60"], [912, "Bb60"], [913, "B60"], [915, "C61"], [917, "Db61"], [918, "E61"], [920, "F61"], [922, "G61"], [924, "Ab61"], [925, "A61"], [927, "Bb61"], [928, "B61"], [930, "C62"], [932, "Db62"], [933, "E62"], [935, "F62"], [937, "G62"], [939, "Ab62"], [940, "A62"], [942, "Bb62"], [943, "B62"], [945, "C63"], [947, "Db63"], [948, "E63"], [950, "F63"], [952, "G63"], [954, "Ab63"], [955, "A63"], [957, "Bb63"], [958, "B63"], [960, "C64"], [962, "Db64"], [963, "E64"], [965, "F64"], [967, "G64"], [969, "Ab64"], [970, "A64"], [972, "Bb64"], [973, "B64"], [975, "C65"], [977, "Db65"], [978, "E65"], [980, "F65"], [982, "G65"], [984, "Ab65"], [985, "A65"], [987, "Bb65"], [988, "B65"], [990, "C66"], [992, "Db66"], [993, "E66"], [995, "F66"], [997, "G66"], [999, "Ab66"], [1000, "A66"], [1002, "Bb66"], [1003, "B66"], [1005, "C67"], [1007, "Db67"], [1008, "E67"], [1010, "F67"], [1012, "G67"], [1014, "Ab67"], [1015, "A67"], [1017, "Bb67"], [1018, "B67"], [1020, "C68"], [1022, "Db68"], [1023, "E68"], [1025, "F68"], [1027, "G68"], [1029, "Ab68"], [1030, "A68"], [1032, "Bb68"], [1033, "B68"], [1035, "C69"], [1037, "Db69"], [1038, "E69"], [1040, "F69"], [1042, "G69"], [1044, "Ab69"], [1045, "A69"], [1047, "Bb69"], [1048, "B69"], [1050, "C70"], [1052, "Db70"], [1053, "E70"], [1055, "F70"], [1057, "G70"], [1059, "Ab70"], [1060, "A70"], [1062, "Bb70"], [1063, "B70"], [1065, "C71"], [1067, "Db71"], [1068, "E71"], [1070, "F71"], [1072, "G71"], [1074, "Ab71"], [1075, "A71"], [1077, "Bb71"], [1078, "B71"], [1080, "C72"], [1082, "Db72"], [1083, "E72"], [1085, "F72"], [1087, "G72"], [1089, "Ab72"], [1090, "A72"], [1092, "Bb72"], [1093, "B72"], [1095, "C73"], [1097, "Db73"], [1098, "E73"], [1100, "F73"], [1102, "G73"], [1104, "Ab73"], [1105, "A73"], [1107, "Bb73"], [1108, "B73"], [1110, "C74"], [1112, "Db74"], [1113, "E74"], [1115, "F74"], [1117, "G74"], [1119, "Ab74"], [1120, "A74"], [1122, "Bb74"], [1123, "B74"], [1125, "C75"], [1127, "Db75"], [1128, "E75"], [1130, "F75"], [1132, "G75"], [1134, "Ab75"], [1135, "A75"], [1137, "Bb75"], [1138, "B75"], [1140, "C76"], [1142, "Db76"], [1143, "E76"], [1145, "F76"], [1147, "G76"], [1149, "Ab76"], [1150, "A76"], [1152, "Bb76"], [1153, "B76"], [1155, "C77"], [1157, "Db77"], [1158, "E77"], [1160, "F77"], [1162, "G77"], [1164, "Ab77"], [1165, "A77"], [1167, "Bb77"], [1168, "B77"], [1170, "C78"], [1172, "Db78"], [1173, "E78"], [1175, "F78"], [1177, "G78"], [1179, "Ab78"], [1180, "A78"], [1182, "Bb78"], [1183, "B78"], [1185, "C79"], [1187, "Db79"], [1188, "E79"], [1190, "F79"], [1192, "G79"], [1194, "Ab79"], [1195, "A79"], [1197, "Bb79"], [1198, "B79"], [1200, "C80"], [1202, "Db80"], [1203, "E80"], [1205, "F80"], [1207, "G80"], [1209, "Ab80"], [1210, "A80"], [1212, "Bb80"], [1213, "B80"], [1215, "C81"], [1217, "Db81"], [1218, "E81"], [1220, "F81"], [1222, "G81"], [1224, "Ab81"], [1225, "A81"], [1227, "Bb81"], [1228, "B81"], [1230, "C82"], [1232, "Db82"], [1233, "E82"], [1235, "F82"], [1237, "G82"], [1239, "Ab82"], [1240, "A82"], [1242, "Bb82"], [1243, "B82"], [1245, "C83"], [1247, "Db83"], [1248, "E83"], [1250, "F83"], [1252, "G83"], [1254, "Ab83"], [1255, "A83"], [1257, "Bb83"], [1258, "B83"], [1260, "C84"], [1262, "Db84"], [1263, "E84"], [1265, "F84"], [1267, "G84"], [1269, "Ab84"], [1270, "A84"], [1272, "Bb84"], [1273, "B84"], [1275, "C85"], [1277, "Db85"], [1278, "E85"], [1280, "F85"], [1282, "G85"], [1284, "Ab85"], [1285, "A85"], [1287, "Bb85"], [1288, "B85"], [1290, "C86"], [1292, "Db86"], [1293, "E86"], [1295, "F86"], [1297, "G86"], [1299, "Ab86"], [1300, "A86"], [1302, "Bb86"], [1303, "B86"], [1305, "C87"], [1307, "Db87"], [1308, "E87"], [1310, "F87"], [1312, "G87"], [1314, "Ab87"], [1315, "A87"], [1317, "Bb87"], [1318, "B87"], [1320, "C88"], [1322, "Db88"], [1323, "E88"], [1325, "F88"], [1327, "G88"], [1329, "Ab88"], [1330, "A88"], [1332, "Bb88"], [1333, "B88"], [1335, "C89"], [1337, "Db89"], [1338, "E89"], [1340, "F89"], [1342, "G89"], [1344, "Ab89"], [1345, "A89"], [1347, "Bb89"], [1348, "B89"], [1350, "C90"], [1352, "Db90"], [1353, "E90"], [1355, "F90"], [1357, "G90"], [1359, "Ab90"], [1360, "A90"], [1362, "Bb90"], [1363, "B90"], [1365, "C91"], [1367, "Db91"], [1368, "E91"], [1370, "F91"], [1372, "G91"], [1374, "Ab91"], [1375, "A91"], [1377, "Bb91"], [1378, "B91"], [1380, "C92"], [1382, "Db92"], [1383, "E92"], [1385, "F92"], [1387, "G92"], [1389, "Ab92"], [1390, "A92"], [1392, "Bb92"], [1393, "B92"], [1395, "C93"], [1397, "Db93"], [1398, "E93"], [1400, "F93"], [1402, "G93"], [1404, "Ab93"], [1405, "A93"], [1407, "Bb93"], [1408, "B93"], [1410, "C94"], [1412, "Db94"], [1413, "E94"], [1415, "F94"], [1417, "G94"], [1419, "Ab94"], [1420, "A94"], [1422, "Bb94"], [1423, "B94"], [1425, "C95"], [1427, "Db95"], [1428, "E95"], [1430, "F95"], [1432, "G95"], [1434, "Ab95"], [1435, "A95"], [1437, "Bb95"], [1438, "B95"], [1440, "C96"], [1442, "Db96"], [1443, "E96"], [1445, "F96"], [1447, "G96"], [1449, "Ab96"], [1450, "A96"], [1452, "Bb96"], [1453, "B96"], [1455, "C97"], [1457, "Db97"], [1458, "E97"], [1460, "F97"], [1462, "G97"], [1464, "Ab97"], [1465, "A97"], [1467, "Bb97"], [1468, "B97"], [1470, "C98"], [1472, "Db98"], [1473, "E98"], [1475, "F98"], [1477, "G98"], [1479, "Ab98"], [1480, "A98"], [1482, "Bb98"], [1483, "B98"], [1485, "C99"], [1487, "Db99"], [1488, "E99"], [1490, "F99"], [1492, "G99"], [1494, "Ab99"], [1495, "A99"], [1497, "Bb99"], [1498, "B99"], [1500, "C100"], [1502, "Db100"], [1503, "E100"], [1505, "F100"], [1507, "G100"], [1509, "Ab100"], [1510, "A100"], [1512, "Bb100"], [1513, "B100"], [1515, "C101"], [1517, "Db101"], [1518, "E101"], [1520, "F101"], [1522, "G101"], [1524, "Ab101"], [1525, "A101"], [1527, "Bb101"], [1528, "B101"], [1530, "C102"], [1532, "Db102"], [1533, "E102"], [1535, "F102"], [1537, "G102"], [1539, "Ab102"], [1540, "A102"], [1542, "Bb102"], [1543, "B102"], [1545, "C103"], [1547, "Db103"], [1548, "E103"], [1550, "F103"], [1552, "G103"], [1554, "Ab103"], [1555, "A103"], [1557, "Bb103"], [1558, "B103"], [1560, "C104"], [1562, "Db104"], [1563, "E104"], [1565, "F104"], [1567, "G104"], [1569, "Ab104"], [1570, "A104"], [1572, "Bb104"], [1573, "B104"], [1575, "C105"], [1577, "Db105"], [1578, "E105"], [1580, "F105"], [1582, "G105"], [1584, "Ab105"], [1585, "A105"], [1587, "Bb105"], [1588, "B105"], [1590, "C106"], [1592, "Db106"], [1593, "E106"], [1595, "F106"], [1597, "G106"], [1599, "Ab106"], [1600, "A106"], [1602, "Bb106"], [1603, "B106"], [1605, "C107"], [1607, "Db107"], [1608, "E107"], [1610, "F107"], [1612, "G107"], [1614, "Ab107"], [1615, "A107"], [1617, "Bb107"], [1618, "B107"], [1620, "C108"], [1622, "Db108"], [1623, "E108"], [1625, "F108"], [1627, "G108"], [1629, "Ab108"], [1630, "A108"], [1632, "Bb108"], [1633, "B108"], [1635, "C109"], [1637, "Db109"], [1638, "E109"], [1640, "F109"], [1642, "G109"], [1644, "Ab109"], [1645, "A109"], [1647, "Bb109"], [1648, "B109"], [1650, "C110"], [1652, "Db110"], [1653, "E110"], [1655, "F110"], [1657, "G110"], [1659, "Ab110"], [1660, "A110"], [1662, "Bb110"], [1663, "B110"], [1665, "C111"], [1667, "Db111"], [1668, "E111"], [1670, "F111"], [1672, "G111"], [1674, "Ab111"], [1675, "A111"], [1677, "Bb111"], [1678, "B111"], [1680, "C112"], [1682, "Db112"], [1683, "E112"], [1685, "F112"], [1687, "G112"], [1689, "Ab112"], [1690, "A112"], [1692, "Bb112"], [1693, "B112"], [1695, "C113"], [1697, "Db113"], [1698, "E113"], [1700, "F113"], [1702, "G113"], [1704, "Ab113"], [1705, "A113"], [1707, "Bb113"], [1708, "B113"], [1710, "C114"], [1712, "Db114"], [1713, "E114"], [1715, "F114"], [1717, "G114"], [1719, "Ab114"], [1720, "A114"], [1722, "Bb114"], [1723, "B114"], [1725, "C115"], [1727, "Db115"], [1728, "E115"], [1730, "F115"], [1732, "G115"], [1734, "Ab115"], [1735, "A115"], [1737, "Bb115"], [1738, "B115"], [1740, "C116"], [1742, "Db116"], [1743, "E116"], [1745, "F116"], [1747, "G116"], [1749, "Ab116"], [1750, "A116"], [1752, "Bb116"], [1753, "B116
```

```
18 ["Bb4", 70], ["B4", 71], ["C5", 72], ["Db5", 73], ["D5", 74]]);
```

### A.3.2. Funcions

```
1 ////////////////////////////////////////////////////////////////////
2 // CODI CONVERSIÓ PER TONALITAT ////////////////////////////////////////////////////////////////////
3 ////////////////////////////////////////////////////////////////////
4
5 function convert() {
6   // barra de cerca
7   var search_bar = document.getElementById('search-box');
8   // boolean mind key
9   var mind_key = document.getElementById('mind-key').checked;
10  // passar la seqüència a llista
11  var piano_bar = document.getElementById('input-notes');
12  var sequence = piano_bar.value.trim();
13  sequence = sequence.trim().split(" ");
14  if (sequence == "") return -1;
15  if (mind_key) { // execució normal
16    var trad = mus2que(sequence);
17  } else { // execució totes les tonalitats
18    // seqüència a número midi
19    var midiSequence = [];
20    sequence.forEach((note) => {
21      midiSequence.push(key2midi.get(note));
22    })
23    // llista amb possibles seqüències
24    seqs = [];
25    min = Math.min(...midiSequence);
26    dif_min = MIN_NOTE - min;
27    max = Math.max(...midiSequence);
28    dif_max = MAX_NOTE - max;
29
30    for (i = dif_min; i <= dif_max; i++) {
31      if (i == 0) seqs.push(sequence);
32      else {
33        new_midi = midiSequence.map(x => x + i); // midi canviat
34        new_notes = new_midi.map(x => midi2key.get(x)); // trad notes
35        if (!new_notes.includes(undefined)) { // si totes notes vàlides
36          seqs.push(new_notes); // afegir a seqüències
37        }
38      }
39    }
40    // traducció possibles seqüències
41    var trad = "";
42    seqs.forEach((seq) => {
43      trad += mus2que(seq);
44    });
45  }
46  query = trad.slice(0,-3); // eliminar últim OR
47  search_bar.value = query;
48  if (search_bar.value == "") search_bar.value = "NO RESULTS"
49 }
```

```
1 ////////////////////////////////////////////////////////////////////
2 // TRADUCCIÓ NOTACIÓ MUSICAL A QUERY ////////////////////////////////////////////////////////////////////
3 ////////////////////////////////////////////////////////////////////
4
5 function mus2que(sequence) {
6   // INICIALIZACIONS ////////////////////////////////////////////////////////////////////
7   // distància (el &n&)
8   const n = 20;
```



```

9 // claus
10 var clefs = new Map([
11   ["c1", []], ["c2", []], ["c3", []], ["c4", []], ["c5", []],
12   ["f1", []], ["f2", []], ["f3", []], ["f4", []]
13 ]);
14 // dades musicals
15 var mus;
16 // dades geometriques
17 var geo;
18 // mus: array amb les notes
19 var mus = "0,E2,F2,G2,A2,B2,C3,D3,E3,F3,G3,A3,B3,C4,D4,E4,F4,G4,A4,B4,C5,B5
20 ,D5"
21 mus = mus.split(',');
22 // geo: array amb les fileres
23 var geo = "c1,0,0,0,0,0,0,0,0,0,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5\nc2
24 ,0,0,0,0,0,0,0,0,0,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5,0,0\nc3
25 ,0,0,0,0,0,0,0,0,0,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5,0,0,0,0\nc4,0,0,0,0,10
26 ,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5,0,0,0,0,0\nc5,0,0,10,s0,l1,s1,l2,s2
27 ,l3,s3,l4,s4,l5,s5,0,0,0,0,0,0\nf1,0,0,0,0,0,0,10,s0,l1,s1,l2,s2,l3
28 ,s3,l4,s4,l5,s5,0,0,0,0,0,0\nf2,0,0,0,0,0,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5
29 ,0,0,0,0,0,0\nf3,0,0,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5
30 ,0,0,0,0,0,0\nf4,10,s0,l1,s1,l2,s2,l3,s3,l4,s4,l5,s5
31 ,0,0,0,0,0,0,0,0,0"
32 geo = geo.split('\n');
33 for (let i = 0; i < geo.length; i++) {
34   geo[i] = geo[i].split(',');
35 }
36
37 // TRADUCCIO MUS-GEO //////////////////////////////////////
38 //recorrer sequencia per traduir nota a nota
39 for (let i = 0; i < sequence.length; i++) {
40   // buscar posicio nota
41   var bemoll = false
42   if (sequence[i].includes('b')) { // si nota bemoll
43     var col = mus.indexOf(sequence[i].slice(0, -2) +
44       sequence[i].slice(-1)); // busca sense bemoll
45     bemoll = true
46   } else {
47     var col = mus.indexOf(sequence[i]); // busca la nota normal
48   }
49
50   // recorre claus per traduir nota per a cadascuna
51   let j = 0;
52   for (let entry of clefs) {
53     var eq = geo[j][col]; // eq geometrica
54     if (eq == '0') // si eq no valida, clau tampoc
55       clefs.set(entry[0], '0');
56     else if (entry[1] != '0'){ // si eq i clau valides
57       if (bemoll) eq = "flat " + eq; // si bemoll, es posa
58       entry[1].push(eq); // s'afig la nota
59     }
60     j++;
61   }
62 }
63
64 // OBTENCIO DE LES QUERIES //////////////////////////////////////
65 // recorre traduccio clau a clau
66 var complete_query = "";
67 var i = 0;
68 for (let entry of clefs) {
69   if (entry[1] != '0') { // si clau valida
70     // cada sequencia com (clau &n& [n1 n2 ...])
71     q = "(" + entry[0] + " &" + n + "& ["
72     for (let j = 0; j < entry[1].length; j++) {

```

```

64         q += entry[1][j] + " "
65     }
66     q += "]) || "
67     // afegir a la query completa RETALLAT per a no flat en linia
68     if (!(q.includes("flat 1"))) complete_query += q;
69 }
70 i++;
71 }
72 return complete_query;
73 }

```

```

1 ///////////////////////////////////////////////////////////////////
2 // FER SONAR NOTA ///////////////////////////////////////////////////////////////////
3 ///////////////////////////////////////////////////////////////////
4
5 function playNote(key) {
6     // Mostrar tecla polsada en barra de cerca si REC on
7     // nomes si la tecla es valida
8     var rec = document.getElementById("button-rec");
9     if (rec.classList.contains("rec-on") && !key.classList.contains("no-valid"))
10    {
11        var search_bar = document.getElementById('input-notes');
12        var str = search_bar.value.trimStart() + " " + key.id;
13        search_bar.value = str.trimStart(); // esborrar " " al principi
14    }
15    // Executar nota
16    const note = document.getElementById(key.id);
17    note.currentTime = 0;
18    note.play();
19    // Control tecla activa
20    key.classList.add('active');
21    note.addEventListener('ended', () => {
22        key.classList.remove('active');
23    });
24 }

```

```

1 ///////////////////////////////////////////////////////////////////
2 // CONTROL D'INPUTS MUSICALS ///////////////////////////////////////////////////////////////////
3 ///////////////////////////////////////////////////////////////////
4
5 // TECLAT VIRTUAL/////////////////////////////////////////////////////////////////
6 const keys = document.getElementsByClassName('key');
7
8 for (let i=0; i < keys.length; i++) {
9     keys[i].addEventListener('click', () => playNote(keys[i]))
10 }
11
12 // TECLAT FISIC ///////////////////////////////////////////////////////////////////
13 const WHITE_KEYS = ['a', 's', 'd', 'f', 'g', 'h', 'j', 'k', 'l'];
14 const BLACK_KEYS = ['w', 'e', 't', 'y', 'u', 'o'];
15
16 const whiteKeys = document.getElementsByClassName('white key');
17 const blackKeys = document.getElementsByClassName('black key');
18
19 document.addEventListener('keydown', e => {
20     if (e.repeat) return;
21     const key = e.key;
22     const whiteKeyIndex = WHITE_KEYS.indexOf(key);
23     const blackKeyIndex = BLACK_KEYS.indexOf(key);
24
25     if (whiteKeyIndex > -1) playNote(whiteKeys[whiteKeyIndex + 5]); // C3
26     else if (blackKeyIndex > -1) playNote(blackKeys[blackKeyIndex + 2]); // Db3
27 });
28

```

```

29 // TECLAT MIDI //////////////////////////////////////
30 // Comprovar MIDI compatible
31 if (navigator.requestMIDIAccess){
32   console.log('This browser supports WebMIDI. ');
33   navigator.requestMIDIAccess().then(success, failure);
34 }
35
36 // Accedir MIDI
37 function success(midiAccess){
38   console.log('Got midi!');
39
40   const inputs = midiAccess.inputs; // inputs MIDI
41   inputs.forEach((input) => { // afegir controlador per cada entrada MIDI
42     input.addEventListener("midimessage", handleInput);
43   });
44 }
45
46 // Gestionar MIDI input
47 function handleInput(input){
48   if (input.data.length == 3) { // input es una nota
49     const midi_note = input.data[1];
50     const vel = input.data[2];
51     // Emmagatzemar nota
52     if (vel > 0) { // note on vel > 0
53       note = midi2key.get(midi_note); // conversio MIDI a nota
54       if (note != -1) { // hi ha equivalencia
55         key = document.getElementById(note);
56         playNote(key);
57       }
58     }
59   }
60 }
61
62 // Descartar MIDI
63 function failure(){
64   console.log("Could not connect MIDI");
65 }

```

### A.3.3. Codi d'inicialització

Codi que s'executa amb la inicialització del demostrador.

```

1 $(function() {
2   // Dialog on es mostra el teclat
3   $( "#dialog-piano" ).dialog({
4     autoOpen: false,
5     modal: false,
6     show: "slide",
7     hide: "slide",
8     width: 650,
9     height: 294,
10    resizable: false,
11  });
12  // Boto per obrir piano
13  $( "#opener-piano" ).click(function() {
14    $( "#dialog-piano" ).dialog("open");
15  });
16  // Boto per esborrar sequencia de notes
17  $( "#button-erase" ).click(function() {
18    var piano_bar = document.getElementById("input-notes");
19    piano_bar.value = "";
20  });
21  // Boto per passar de notes a query

```

```

22 $( "#button-convert" ). click ( convert );
23 // Boto de gravacio
24 $( "#button-rec" ). click ( function () {
25     if ( $( "#button-rec" ). hasClass ( "rec-off" ) ) {
26         $( "#button-rec" ). removeClass ( "rec-off" );
27         $( "#button-rec" ). addClass ( "rec-on" );
28     }
29     else {
30         $( "#button-rec" ). removeClass ( "rec-on" );
31         $( "#button-rec" ). addClass ( "rec-off" );
32     }
33 });
34 });

```

## A.4 Codi Python

Algoritmes utilitzats en l'avaluació per processar dades i realitzar les consultes.

### A.4.1. Funcions de traducció

#### Codi de traducció de notació implícita (o freqüencial) a notació musical

```

1 import sys
2 import numpy as np
3
4 def fre2mus(seq):
5
6     # Entrada:
7     # - Sequencia format 22 23 20 20 (freqüencial)
8     # Eixida:
9     # - Sequencia format G4 A4b F4 F4 (musical)
10
11     # Taula equivalencies notacio freqüencial - notacio musical
12     d = np.loadtxt('low8fre-mus-equivalence.txt', dtype=str)
13
14     eq = [] # equivalencia en notes de la sequencia
15     for n in seq: # es processa frequencia a frequencia
16         if n != '-': # si es tracta d'una nota
17             row, = np.where(d[:,0] == n) # es troba ubicacio
18             nota = d[row,1] # i la nota equivalent
19             eq.append(str(nota[0])) # s'afig la nota a la sequencia
20
21     return eq
22
23 if __name__ == '__main__':
24
25     # Entrada:
26     # - Fitxer amb sequencies format "22 23 20 20" (freqüencial)
27     # - us: python fre-to-mus.py FITXER
28
29     # Eixida:
30     # Equivalencia de les sequencies en format musical
31
32     # Comprovacio correccio crida
33     if len(sys.argv) != 2:
34         print("ERROR: es necessari introduir el nom d'UN fitxer amb sequencies")
35         print("us del script: python3 fre-to-mus.py FITXER")
36         sys.exit(main())

```

```

37
38 # Parametres entrada
39 fitxer = sys.argv[-1]
40
41 # Totes les seqüències es guarden a f
42 f = np.loadtxt(fitxer, dtype=str)
43 M, N = np.shape(f) # dimensions del fitxer
44 for i in range(M): # per a cada seqüència
45     frequencial = f[i,2:] # de la tercera columna endavant esta la
        seqüència
46     musical = fre2mus(frequencial) # es busca equivalencia
47     print(musical)

```

### Codi de traducció de notació musical a consultes lògiques geomètriques

```

1 import sys
2 import numpy as np
3
4 def mus2que(entrada, n):
5
6     # Entrada:
7     # - Seqüència format G4 A4b F4 F4 (musical)
8     # Eixida:
9     # - Seqüència format L3 L0 S4 L1 (geometric)
10
11     # Taula equivalències notació musical - notació geomètrica
12     d = np.loadtxt('mus-geo-equivalence.txt', dtype=str)
13     # Dimensions taula equivalències
14     M, N = np.shape(d)
15
16     # Diccionari clau: clau - valor: llista amb les notes
17     claus = {"c1": [], "c2": [], "c3": [], "c4": [], "c5": [],
18             "f1": [], "f2": [], "f3": [], "f4": []}
19
20     # Array per guardar solució
21     sol = []
22
23     # En clau es guarden les equivalències
24     for nota in entrada:
25         bemol = False
26         if 'b' in nota: # per si la nota es bemol, es lleva
27             nota = nota[:-1]
28             bemol = True # i es marca com a bemol
29         col, = np.where(d[0,:] == nota) # col nota a la taula
30         for clau in claus:
31             if 0 not in claus[clau]: # si clau vàlida
32                 row, = np.where(d[:,0] == clau) # row clau a la taula
33                 eq = d[row, col][0] # es troba geo
34                 if eq == '0': # si equi no vàlida es posa un 0
35                     claus[clau] = [0]
36                 else:
37                     if bemol: # se anyade con flat
38                         eq = "flat " + eq
39                     claus[clau].append(eq)
40                 else: # se anyade sin flat
41                     claus[clau].append(eq)
42
43     # Una volta es tenen les equivalències de les notes
44     # es recorre el diccionari per obtenir les queries
45     for i, clau in enumerate(claus):
46         notes = claus[clau]
47         if 0 not in notes:
48             # Cada melodia com: (clau &n& [n1 n2 ...])

```

```
49     q = "(" + clau + " &" + str(n) + "& [" # (clau &n&
50     for nota in notes:
51         q += nota + " " # [n1 n2 ...
52     q += ") ||" # ] ) || ...
53     # Afegir equivalencia a solucio
54     sol.append(q)
55
56     sol[-1] = sol[-1][:-2] # per eliminar ultim OR (||)
57
58     return sol
59
60 if __name__ == '__main__':
61
62     # Entrada:
63     # Melodia format C4 D4 E4 C4 [-n valor]
64     # - primers parametres son la sequencia de notes
65     # - ultim parametre opcional per a distancia clau-sequencia
66     # - us: python mus-to-geo.py SEQUENCIA [OPCIONES]
67
68     # Eixida:
69     # Possibles queries per a la melodia
70
71     # Comprovacio correccio crida
72     if len(sys.argv) < 2:
73         print("ERROR: es necessaria al menys una nota en la sequencia d'entrada")
74         print("us del script: python3 mus-to-geo.py SEQUENCIA [-n numero]")
75         sys.exit(main())
76
77     # Parametres entrada
78     if 'n' in sys.argv[-2]:
79         if sys.argv[-2] == '-n':
80             distancia = sys.argv[-1]
81             melodia = sys.argv[1:-2]
82         else:
83             print("ERROR: el script s'ha de cridar amb python3 mus-to-geo.py")
84             print("SEQUENCIA [-n numero]")
85             sys.exit(main())
86     elif 'n' in sys.argv[-1]:
87         print("ERROR: el script s'ha de cridar amb python3 mus-to-geo.py")
88         print("SEQUENCIA [-n numero]")
89         sys.exit(main())
90     else:
91         distancia = 20
92         melodia = sys.argv[1:]
93
94     # Execucio de la traduccio
95     queries = mus2que(melodia, distancia)
96     print(*queries)
```



---

---

## APÈNDIX B

# Sobre teoria musical

---

Aquest apèndix té com a objectiu explicar de forma superficial alguns dels fonaments de la teoria musical per tal de facilitar l'enteniment del plantejament i la forma d'abordar el projecte.

### B.1 Unitats bàsiques: les notes

---

Es podria dir que la música es fonamenta en l'ordenació de les notes perseguint una finalitat artística [32]. Cada una de les notes es correspon amb un so, definit com les vibracions mecàniques i les ones d'un medi elàstic en l'àmbit de freqüències de l'audició humana (de 20 a 20.000 Hz) [26].

D'altra banda, un to és el resultat d'una vibració sinusoidal aïllada (en cas que siga *pur*) o d'una suma de tons sinusoidals (en cas que siga *natural*) [26].

Per tal de saber quin so correspon a quina nota, es fixà el to de referència en 440 Hz a 20°C [33] (que es correspon amb la nota *la*). El sistema dels Cents [34] divideix en 100 cents el pas de semitò temperat i en 1.200 el d'octava [26] (vore B.2).

L'afinació temperada es basa en divisions matemàtiques mentre que la pura segueix les proporcions intervàl·liques naturals (com en el sistema pitagòric [27, p. 734]) [26]. La temperada és la corresponent a la afinació usada en la música europea occidental moderna [32] i la que se suposarà a partir d'ara per a l'explicació.

Altre aspecte important dels sons són les seues qualitats intrínseques: altura, durada, intensitat i timbre [35]. Com que no totes elles són rellevants per aquest treball, per tal de simplificar l'explicació, a aquest apèndix només serà necessari tindre en compte el concepte d'altura, que es correspon amb la freqüència esmentada anteriorment.

### B.2 Escala cromàtica: el piano

---

Existeixen set notes d'origen: do, re, mi, fa, sol, la y si (en països llatins); en notació anglesa o alemanya (notació anglosaxona) es denominen per lletres a partir de la nota *la*: a, b, c, d, e, f, g [26]. A aquestes notes se'ls pot modificar l'altura mitjançant alteracions. Amb un sostingut ( $\sharp$ ) s'augmenta en un semitò l'altura de la nota a la que s'aplica i amb un bemoll ( $\flat$ ) es disminuirà [26].

Tenint en consideració que entre les notes mi i fa i les notes si i do només hi ha un semitò de diferència i no un to com entre la resta de parells (per exemple un *mi $\sharp$*  seria un *fa*)



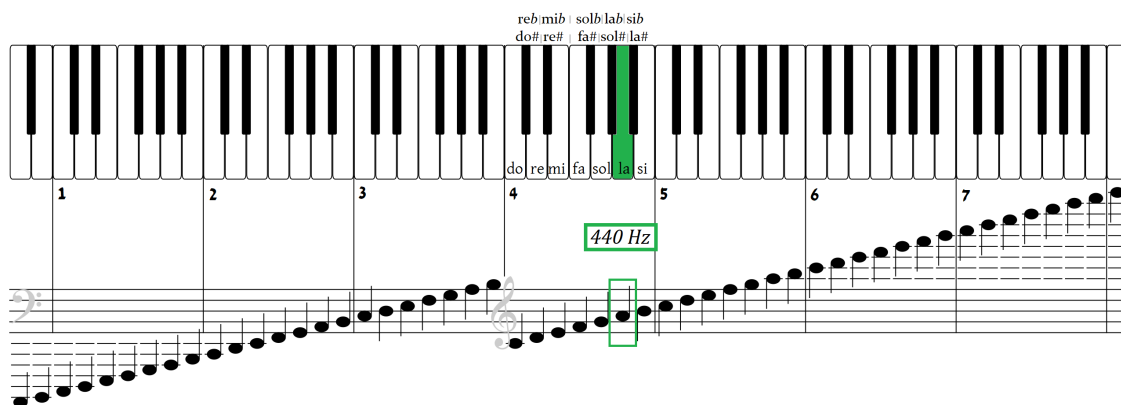
[32], el resultat és que es tenen dotze notes musicals. Aquestes notes ordenades de do a si formen una octava [26].

Les notes tenen origen en l'afinació temperada (esmentada en B.1). El principi fonamental del qual és dividir l'octava en 12 semitons iguals [26]. Com la raó de freqüència d'una octava és 2 [26], la raó  $s$  d'un semitò s'expressarà com:  $s^{12} = 2$ ;  $s = \sqrt[12]{2} = 1,05946$  [27, p. 735].

Així, cada grup de dotze notes forma una octava i una o més octaves en conjunt formen part de l'escala<sup>1</sup> cromàtica [27, p. 662]. Les notes d'aquesta escala vindran donades per les potències de l'expressió anterior. Per tant,  $re = 1,05946^2$  o  $re\sharp = mib = 1,05946^3$  [27, p. 735], per exemple.

L'escala cromàtica es correspon amb el teclat del piano. Dins, les tecles blanques són les set notes d'origen i les negres els passos de semitò entre elles.

Com es pot observar a B.1 el teclat es forma superposant octaves; les notes són cícliques. La diferència entre el primer re que apareix al piano i l'últim és la seua altura. Ambdós són un re però amb una diferència de 7.200 cents ( $1200 \times 6$ ). És per això que, per tal d'evitar confusions, juntament amb la nota s'expressa a quina octava pertany aquesta. Per exemple, el la central (a 440 Hz) és el *la* 4, ja que es troba en la quarta octava del teclat [35].



**Figura B.1:** Representació musical i identificació de les tecles del piano, corresponents a l'escala cromàtica.

### B.3 Representació gràfica de les notes: claus i pentagrames

Al igual que amb l'escriptura de textos, la representació de la música ha anat evolucionant al llarg de la història d'una tradició oral a una forma escrita [26]. Amb l'evolució de la música i les necessitats artístiques de cada època, també ha anat avançant la forma en la qual es s'escriu aquesta.

Actualment (parlant des d'un punt de vista europeu occidental) les notes es representen a través d'un pentagrama. Els pentagrames són una sèrie de cinc línies horitzontals (ampliables amb línies addicionals per a casos individuals que ho requerisquen) on s'il·lustren les notes, indicant així el seu to [27, p. 708].

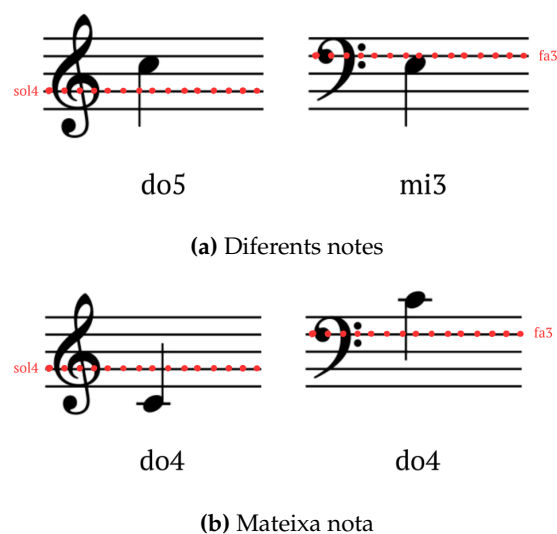
Les notes s'escriuen en correspondència amb una clau. Aquesta clau és un símbol que es troba al principi del pentagrama per indicar el to de les notes d'aquest. Hi ha tres tipus:

<sup>1</sup>Una escala és el *material tonal* obtés al ordenar els tons de forma ascendent. [27, p. 662]

sol, fa i do (vore B.3) i indiquen en quina posició es troba cada nota respectivament [27, p. 158]. En B.1 es troben les claus de *fa en quarta* i *sol en segona*. La primera indica que en la quarta línia del pentagrama es representa el *fa* 3 i la segona que en la segona línia es representa el *sol* 4 [35].

Un exemple il·lustratiu es troba a B.2. En B.2a, malgrat que s'haja dibuixat les figures en la mateixa part del pentagrama, al tercer espai, no es tracta de la mateixa nota. Això és deu a que la clau que les precedeix tampoc és la mateixa. En el primer cas, al haver una clau de sol en la segona línia, el tercer espai es correspon amb un *do*<sup>5</sup>; en el segon cas es tracta d'un *mi*<sup>3</sup>.

D'altra banda, en B.2b s'han dibuixat les figures en llocs molt diferents. En el cas de l'esquerra, en la primera línia addicional per baix del pentagrama i en el de la dreta, en la primera línia addicional per damunt del pentagrama. Encara així, en ambdós casos es tracta d'un *do*<sub>4</sub>, semblança deguda també al canvi de clau.



**Figura B.2:** Variació de la representació de les notes en el pentagrama depenent de la clau que les precedisca.



**Figura B.3:** Representació gràfica dels diferents tipus de claus musicals existents en la música occidental europea.

## B.4 Relacions entre les notes: els intervals i transposicions

Un interval és la distància d'altura entre dos sons musicals [32]. Com s'ha expressat anteriorment, en música occidental la distància mínima entre dues notes és el semitò, és a dir 100 cents. Si entre dues notes hi ha una distància de dos semitons, es diu que la distància que les separa és un to [32].

<sup>2</sup>Do de la cinquena octava del teclat del piano.

Per anomenar els intervals es mesuren el nombre de graus (notes naturals, les tecles blanques del piano) que es troben continguts dins del mateix [27]. Per exemple: l'interval entre do<sub>4</sub> i mi<sub>4</sub> és un interval de tercera<sup>3</sup> perquè, comptant els extrems do<sub>4</sub> i mi<sub>4</sub>, hi ha tres graus (do, re i mi) entre cada nota. A més, aquests intervals es poden classificar segons la seua qualitat. Aquesta dependrà de la quantitat de semitons que s'inclouen dins de l'interval [32]. En l'exemple anterior la qualitat de la tercera seria major (vore B.4), ja que es pot separar l'interval en dos tons (quatre semitons), un de do<sub>4</sub> a re<sub>4</sub> i altre de re<sub>4</sub> a mi<sub>4</sub>. Per tal de saber la qualitat dels diferents intervals vore B.1.

Cal destacar que, sobrepassada la distància d'octava entre dues notes, la qualitat dels intervals serà la mateixa que la corresponent amb la nota de la mateixa octava i a la quantitat de semitons que els separen en l'octava original només s'hauran de sumar dotze semitons per cada octava que separe les notes. Es poden vore tres exemples d'intervals a B.4, on es representen els graus entre les notes extremes i en la part inferior de cada interval es mostren la quantitat de tons (I) i semitons (-) que conté el mateix.

Semitons	Interval	Qualitat
0	uníson	just
1	segona	menor
2	segona	major
3	tercera	menor
4	tercera	major
5	quarta	justa
6	quarta / cinquena	augmentada / disminuïda
7	cinquena	justa
8	sisena	menor
9	sisena	major
10	setena	menor
11	setena	major
12	octava	justa

**Taula B.1:** Classificació dels intervals musicals i la seua qualitat segons els semitons que s'inclouen en ells. Font: [35]

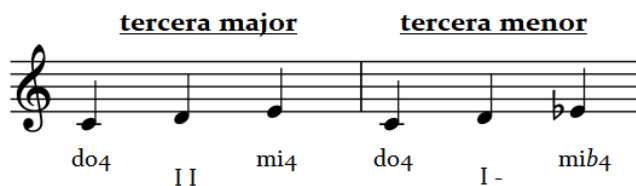
A partir de les relacions entre les notes, aplicant els conceptes dins del marc melòdic, entra en escena la transposició. La transposició no és més que la reescriptura d'una melodia o sèrie de notes en altre to diferent a l'original [27, p.757].

Segons aquesta definició, encara que una melodia es transporte, continua sent la mateixa. No sonarà igual que l'original en relació a les freqüències que l'oïda escoltarà, però les relacions intervàliques entre les notes es mantindran.

Per transportar una melodia només caldrà augmentar o disminuir totes les notes en l'interval al qual es vulga fer la transposició. És a dir, si la melodia comença en do<sub>4</sub> i es vol fer una transposició una sisena menor per damunt (començar-la en la<sub>4</sub>), caldrà augmentar totes les notes de la melodia original en una sisena menor.

És per això que en l'exemple B.5 es pot vore l'equivalència entre tres melodies que, malgrat comencen en notes diferents, mantenen els intervals entre les notes intactes; en alguns casos necessitant alteracions per fer-ho.

<sup>3</sup>L'interval s'anomena en femení ja que es fa referència a la distància [35].

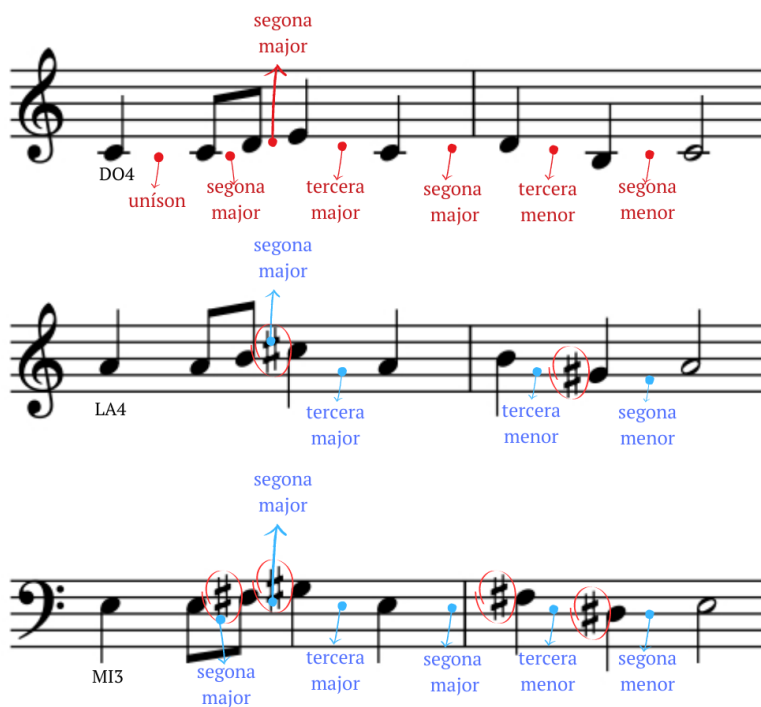


(a) Tercera major (4 semitons) i tercera menor (3 semitons)



(b) Novena menor (13 semitons)

**Figura B.4:** Exemple de dos intervals de tercera i un de novena amb els graus intermedis corresponents entre les notes extremes.



**Figura B.5:** Exemple de transposició amb els intervals entre els parells de notes consecutives. Les alteracions necessàries per respectar els intervals entre les notes es troben assenyalades.

## ANNEX

### OBJECTIUS DE DESENVOLUPAMENT SOSTENIBLE

Grau de relació del treball amb els Objectius de desenvolupament sostenible (ODS).

Objectius de desenvolupament sostenible	Alt	Mitjà	Baix	No procedeix
ODS 1. <b>Fi de la pobresa.</b>				X
ODS 2. <b>Fam zero.</b>				X
ODS 3. <b>Salut i benestar.</b>				X
ODS 4. <b>Educació de qualitat.</b>			X	
ODS 5. <b>Igualtat de gènere.</b>				X
ODS 6. <b>Aigua neta i sanejament.</b>				X
ODS 7. <b>Energia assequible i no contaminant.</b>				X
ODS 8. <b>Treball decent i creixement econòmic.</b>			X	
ODS 9. <b>Indústria, innovació i infraestructures.</b>	X			
ODS 10. <b>Reducció de les desigualtats.</b>				X
ODS 11. <b>Ciutats i comunitats sostenibles.</b>				X
ODS 12. <b>Producció i consum responsables.</b>				X
ODS 13. <b>Acció pel clima.</b>				X
ODS 14. <b>Vida submarina.</b>				X
ODS 15. <b>Vida d'ecosistemes terrestres.</b>				X
ODS 16. <b>Pau, justícia i institucions sòlides.</b>				X
ODS 17. <b>Aliances per lograr objectius.</b>				X



Aquest treball de fi de grau es pot enmarcar dins dels objectius de desenvolupament sostenible (ODS), especialment dins de l'objectiu número 9: indústria, innovació i infraestructures, encara que també podria relacionar-se en menor mesura amb el número 4 (educació de qualitat) i el número 8 (treball decent i creixement econòmic).

Per una banda, a partir d'aquest projecte s'han pogut plantejar línies de recerca i investigació futures, que impliquen una contribució al progrés tecnològic i científic. Aquests aspectes no només es relacionen amb la tecnologia sinó que podrien suposar una contribució al desenvolupament científic dins de l'àmbit de la investigació històrica, ja que el treball realitzar involucra una eina útil per a l'anàlisi musical històric.