برای فایل های موسیقی چند فرمت دیجیتال وجود دارند که در این پروژه از فرمت mid یا midi استفاده کردم

فرمت های دیگری که وجود دارد یکی musicxml و دیگری abc است که به علت نبودن سورس ها به این فرمت ها و مشکل تبدیل آن ها به یکدیگر، طی این پروژه چیزی که تولید می شود شباهت زیاری به فرمت abc دارد

در این فرمت این فرمت از دستورات 7 بایتی تشکیل شده که هر دستور شامل شروع یا پایان یافتن یک نت مخصوص روی یک کانال مخصوص می باشد. توضیحات این 7 بایت در این لینک توضیح داده شده که برای ادامه لازم نمی باشد <https://www.midi.org/specifications/item/table-1-summary-of-midi-message>

برای ترجمه ی فایل mid به متن از سایت <http://flashmusicgames.com/midi/mid2txt.php> استفاده کردم. البته قبل از شروع دستورات گفته شده دستورات توضیحاتی در این متن ها میاید که در تحلیل ما جایی ندارد ، پس قسمت های مورد نیاز را جدا کردم و در فایل های بدون پیشوند)هم نام با فایل های midi) ریختم

کد convertmidi -withlist - counterpoint – allfile یک فایل که شامل نام تمام فایل های متن هستند را دریافت می کند ( در اینجا نام این فایل 000 است) سپس به ازای هر اهنگ 3 فایل دیگر تولید می کند.

فایلی که با پیشوند 00 تولید می کند فقط عدد نُت ها به صورت خالص هستند و ترتیب آمدنشان در اهنگ. نت هایی که با علامت | کنار هم قرار گرفتند، نت هایی هستند که هم زمان نواخته می شوند.

فایلی که با پیشوند 11 تولید می کند فایل کنترپوان است یعنی هر نت نسبت به نت قبلی ای که نواخته شده چه فاصله ای داشته

فایلی که با پیشوند 22 تولید می کند فایل هارمونی است یعنی نت هایی که هم زمان تولید می شوند چه فاصله ای داشتند

توسط کد mergfiles.py این فایل ها به هم میچسبند و یکی می شوند

توسط کد converttoabc.py فایلی که شامل عدد خالص نت تمام اهنگ ها می باشد را از حالت عددی به نت های امریکایی(a,b,c) تبدیل می کند

برای این که در سایت های wordmap ساز دچار مشکل نشوم تمام # ها رو به s تبدیل کردم و |ها را به o تبدیل کردم

ورد مپ های تولید شده از فایل های 00 (نت های خالص) اهمیت دقیقی ندارند چرا که موسیقی های تولید شده توسط باخ و بتهوون بسیار ترانسفر شدندن و ممکن است از یک گام به گامی دیگر بروند ، اما به صورت اجمالی ورد مپ های 00 باخ و بهتوون اطلاعاتی که در اختیار ما می گذارند این است که بتهوون از هارمونی دو فا، ر دیز دو و فادیز لا بیشتر استفاده کرده و باخ از هارمونی های دو لا،سل می و ر لادیز.

نکته ی مهمی که از ورد مپ فایل های 11 (کنترپوان) میتوان استخراج کرد این است که باخ از فاصله های بسیار بزرگ استفاده کرده (12 22 24 26 ... ) و در فاصله های کنترپوانی بتهوون بسیار عدد 0 می بینیم یعنی یک نت را دو بار پشت هم تکرار شده است

از ورد مپ 22(هارمونی) عدد 12 که همان اکتاو است را در کار های بتهوون بسیار میبینم شاید یک دلیل آن این است که بتهوون سمفونی بیشتر دارد نسبت به باخ