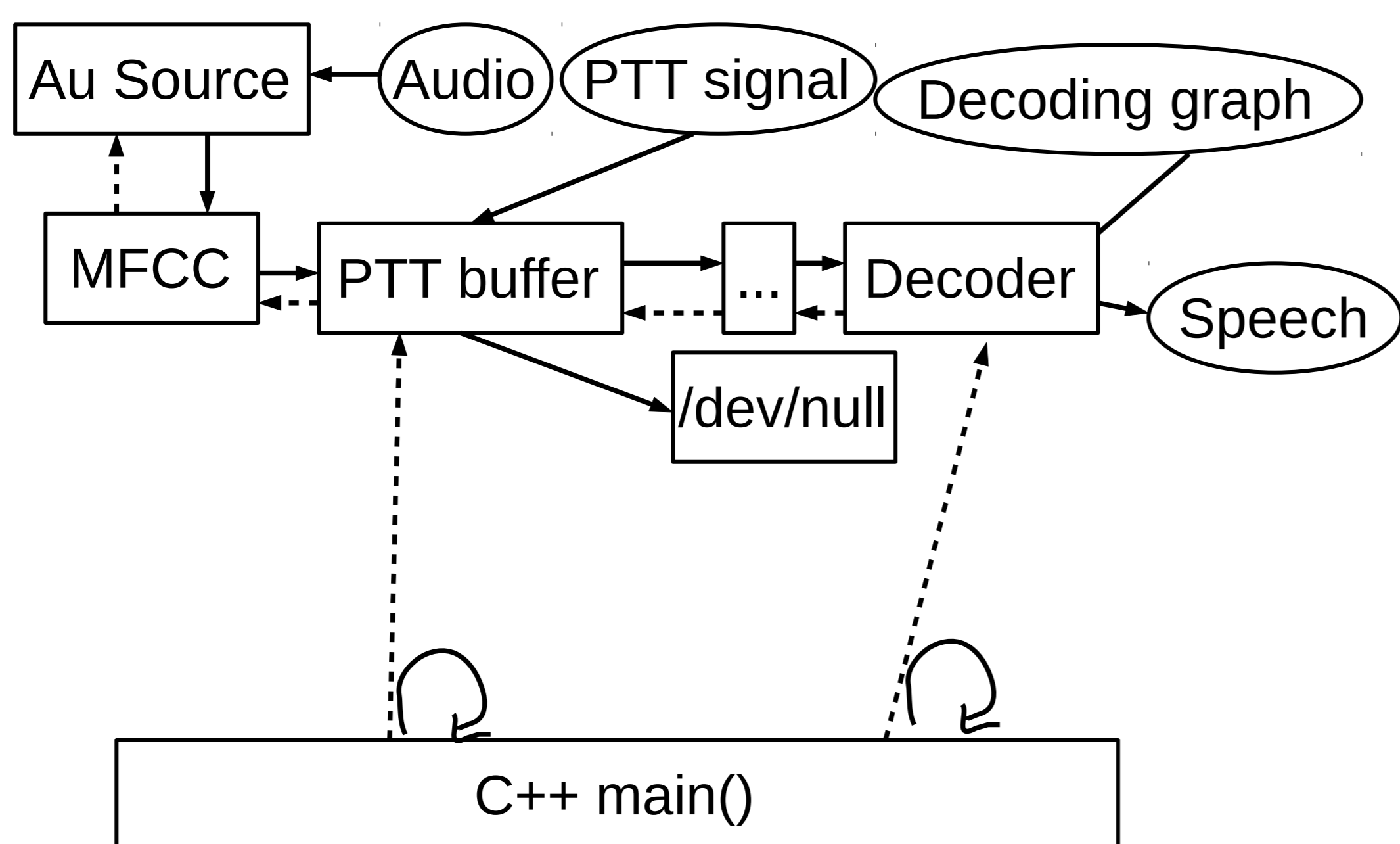
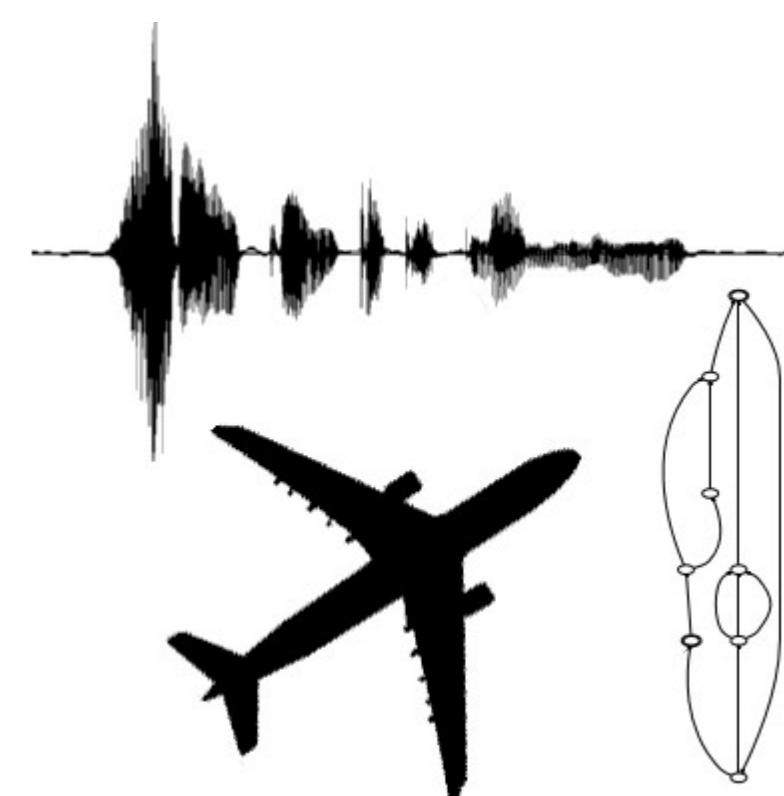


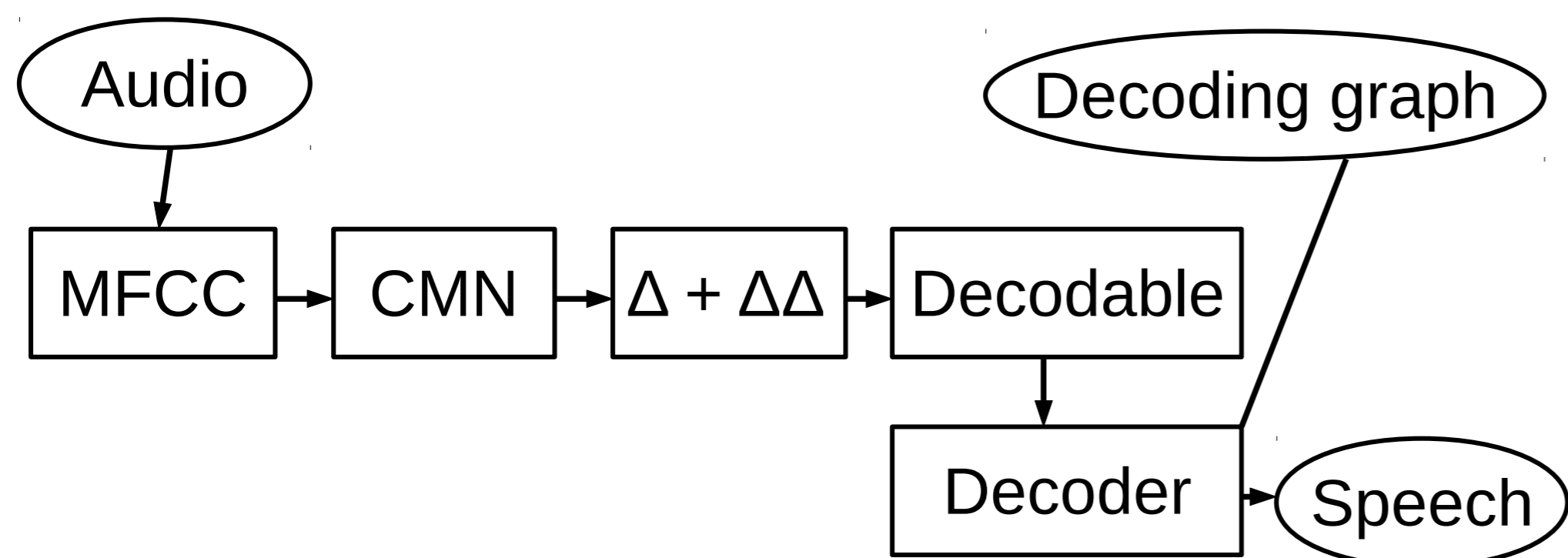
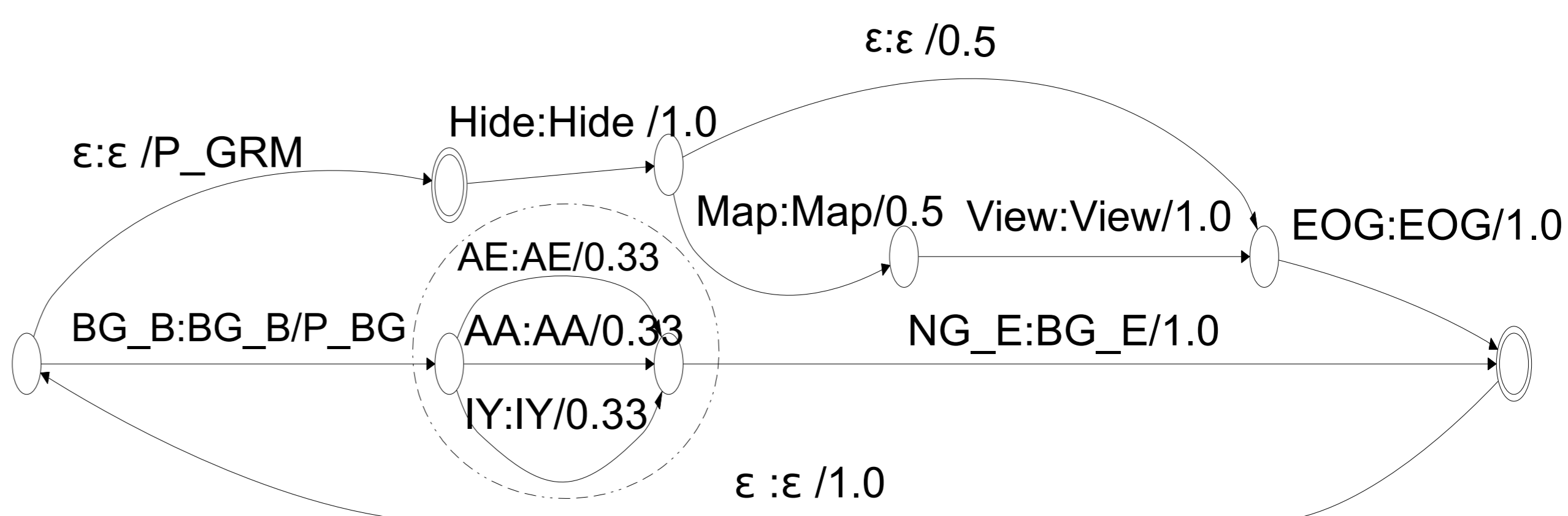
System pro rozpoznávání řeči v projektu A-PiMod

Karel Beneš, xbenes20@stud.fit.vutbr.cz



Většina zodpovědnosti nad rámec běžného rozpoznávání je na bedrech Push-to-talk bufferu. Ten shromažďuje řeč ve formě MFCC a rozhoduje, které rámce budou dekódovány. U těch pak uloží informaci o přesném časování, která je přidána k výstupu. Aby byl systém uživatelsky přívětivější, je rozpoznáváno i půl vteřiny před stiskem PTT klávesy.

Aby bylo možno detekovat věty mimo gramatiku a odhadnout přesnost rozpoznání, byl přidán model pozadí ve formě fonémové smyčky na úrovni modelu jazyka. Takto rozpoznávač úspěšně odmítá nesmyslné promluvy a je odolnější proti šumu.



Vlastní postup zpracování řeči využívá pouze základní bloky akustického modelování. Při první validaci v rámci projektu byli piloti schopni do dvou pokusů zadat 90 % testovacích příkazů. S rychlostí rozpoznávání byli velmi spokojeni (hodnocení 4.4/5).