

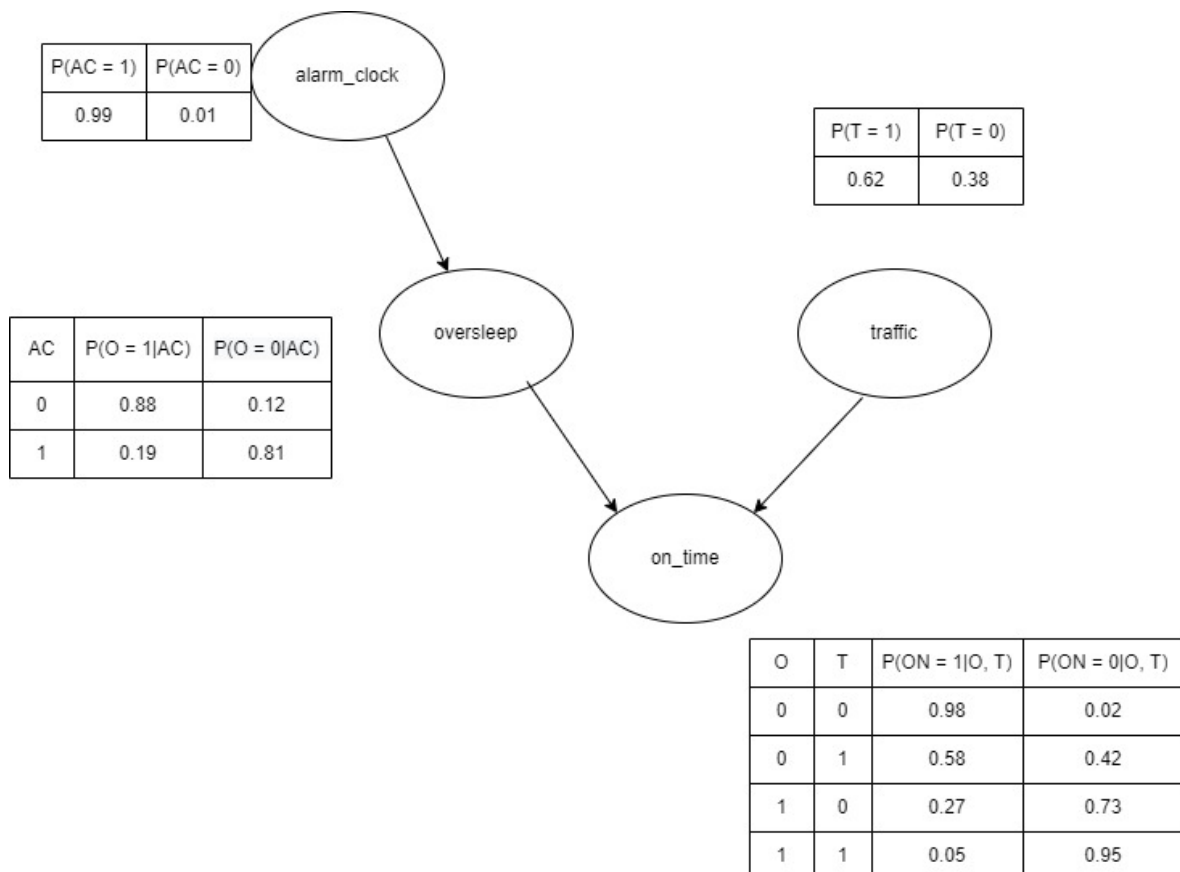
Ćwiczenie 6 – Modele bayesowskie

1) Treść zadania

Napisać losowy generator danych działający zgodnie z rozkładem reprezentowanym przez daną sieć Bayes'a. Sieć powinna opisywać zależności między zmiennymi losowymi (binarne: 0/1) Na wejściu dany jest opis grafu połączeń/zależności między zmiennymi oraz tabele prawdopodobieństw warunkowych. Format danych wejściowych jest dowolny, może być i z pliku .py, ale nie jako new MyCustomFancyGraphClassThatIWillForgetHowToUseSoon (1, a, 2, 84)... tylko raczej {'a': 'b', ...}, {'a': [0, 1], 'b': [[0,1],[1,0]]...} Wygenerować X próbek np. w liczbie [100, 1000, 10000] i sprawdzić czy są zgodne z rozkładem. Problem należy samemu zdefiniować, ale powinien zawierać kilka zmiennych i co najmniej jedną podwójną zależność

2) Przyjęte założenia

- Dane są pobierane z pliku: schema.json
- Graf sieci wraz z prawdopodobieństwami:



Gdzie:

- alarm_clock – budzik zadzwoni
- oversleep – nie obudzimy się w porę, zaśpimy
- traffic – będzie korek na drodze
- on_time – zdążymy na czas

- Pożądane wartości prawdopodobieństwa zostały obliczone za pomocą pliku on_time.py

3) Raport z przeprowadzonych eksperymentów

	prawdopodobieństwa			
	100 próbek	1000 próbek	10000 próbek	Pożądane
alarm_clock	1.0	0.992	0.9915	0.99
oversleep	0.13	0.212	0.1981	0.1969
traffic	0.51	0.615	0.619	0.62
on_time	0.68	0.602	0.6144	0.614175

4) Obserwacje i wnioski

Wraz ze wzrostem liczby próbek prawdopodobieństwa określonych zdarzeń stają się coraz bliższe pożądanym wartościom.