In [ ]:
```
#Šifrování
```

In [ ]:
```
#abeceda
default_alphabet = "ABCDEFGHIJKLMNOPQRSTUVWXYZ_"
```

In [ ]:
```
#nezašifrovaný text
text = "_VOZEM_DO_NEHO_A_ZAS_MNE_BEZI_DO_CESTY__ZACHVELA_SE_TAK_KUDY_VPRAVO_
"_JE_KONEC_OTEVREL_DVIRKA_VYSKOCIL_Z_VOZU_A_POSTAVIL_SE_PRED_KOLA_JED_REKL_C
"T_POJD_MUSIME_DAL_DOVEZU_TE_ASPON_BLIZ_K_HRANICIM_KAM_CHCES_ZPATKY_SKRIPEL_
"ATKY_COPAK_MI_NEROZUMIS_MUSIM_TO_UDELAT_ABYS_VIDEL_ABY_BYLO_JISTO_ZE_JSEM_T
"A_TO_JSEM_MYSLELA_ZE_BYCH_SLA_S_TEBOU_DOPREDU_DOVEDLA_BYCH_TO_DOVEDLA_BYCH_
"ED_NIKDY_ME_NENAPADLO_PTAT_SE_TE_NA_TO_KDYZ_JE_CLOVEK_PRINCEZNA_MYSLI_SI_ZE
"A_OCIMA_PRECE_JEN_NEDOVEDL_ZAPRIT__TAK_VIDIS_VYDECHLA_TY_NEUMIS_ANI_LHAT_TY
```

In [ ]:
```
#náhodný klíč
random_key = random.sample(default_alphabet, 27)
```

In [ ]:
```
#výstup
NGQIWFNTQNXWEQNLNILPNFXWNRWIMNTQNCWPBUNNILCEGWYLNPWNBLJNJVTUNGHZLGQNXWRQNGYW
```

In [ ]:
```

```

In [ ]:
```
#M-H algoritmus k nalezení klíče
```

In [ ]:
```
#zašifrovaný text
input = "NGQIWFNTQNXWEQNLNILPNFXWNRWIMNTQNCWPBUNNILCEGWYLNPWNBLJNJVTUNGHZLGQ
```

In [ ]:
```
#text z knihy
book = "_VOZEM_DO_NEHO_A_ZAS_MNE_BEZI_DO_CESTY__ZACHVELA_SE_TAK_KUDY_VPRAVO_
"_JE_KONEC_OTEVREL_DVIRKA_VYSKOCIL_Z_VOZU_A_POSTAVIL_SE_PRED_KOLA_JED_REKL_C
"T_POJD_MUSIME_DAL_DOVEZU_TE_ASPON_BLIZ_K_HRANICIM_KAM_CHCES_ZPATKY_SKRIPEL_
"ATKY_COPAK_MI_NEROZUMIS_MUSIM_TO_UDELAT_ABYS_VIDEL_ABY_BYLO_JISTO_ZE_JSEM_T
"A_TO_JSEM_MYSLELA_ZE_BYCH_SLA_S_TEBOU_DOPREDU_DOVEDLA_BYCH_TO_DOVEDLA_BYCH_
"ED_NIKDY_ME_NENAPADLO_PTAT_SE_TE_NA_TO_KDYZ_JE_CLOVEK_PRINCEZNA_MYSLI_SI_ZE
"A_OCIMA_PRECE_JEN_NEDOVEDL_ZAPRIT__TAK_VIDIS_VYDECHLA_TY_NEUMIS_ANI_LHAT_TY
```

In [ ]:
```
#matice nezašifrovaného textu
matrix_ref = transition_matrix(bigrams, alphabet)
```

In [ ]:
```
#náhodný klíč
start = random.sample(alphabet, 27)
```

In [ ]:
```
#počet iterací algoritmu
iterations = 20000
```

In [ ]:
```
#výstup
...
Iteration: 19850, Plausibility: -10.049958489613243
Iteration: 19900, Plausibility: -10.049958489613243
Iteration: 19950, Plausibility: -10.049958489613243
```

```
Key: LRCTWODEMSJYFXQHKZPBVG_AUIN
Decrypted text: _VOZEM_DO_NEHO_A_ZAS_MNE_BEZI_DO_CESTY__ZACHVELA_SE_TAK_KUDY
```

In [ ]:

In [ ]:
```
#Dešifrování
```

In [ ]:
```
#abeceda
default_alphabet = "ABCDEFGHIJKLMNOPQRSTUVWXYZ_"
```

In [ ]:
```
#zašifrovaný text
cipher = "NGQIWFNTQNXWEQNLNILPNFXWNRWIMNTQNCWPBUNNILCEGWYLNPWNBLJNJVTUNGHZLG
```

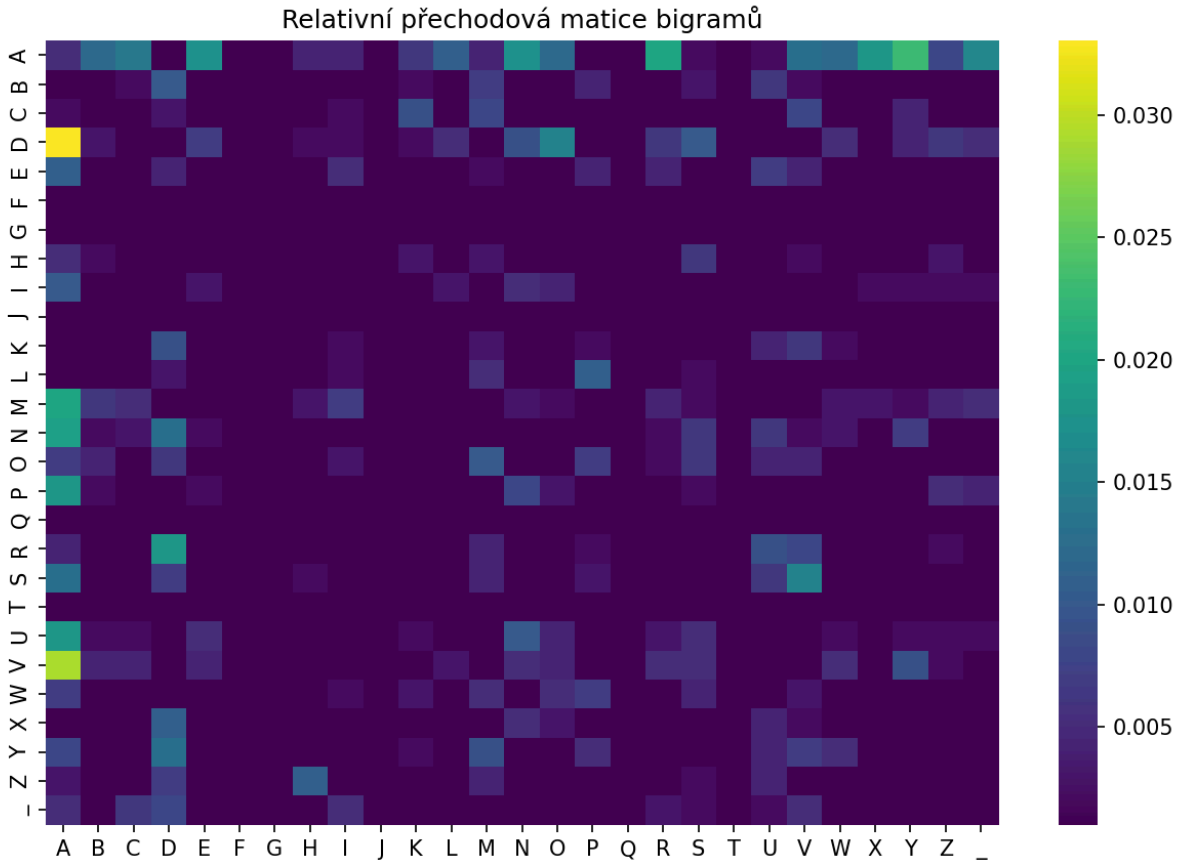
In [ ]:
```
#klíč nalezený M-H algoritmem
best_key = list("LRCTWODEMSJYFXQHKZPBVG_AUIN")
```

In [ ]:
```
#výstup
_VOZEM_DO_NEHO_A_ZAS_MNE_BEZI_DO_CESTY__ZACHVELA_SE_TAK_KUDY_VPRAVO_NEBO_VLE
```

In [ ]:

In [14]:
```
#Vizualizace relativní přechodové matice bigramů
#Z grafu můžeme vidět, že nejčastěji se v daném textu vyskytuje znak
#A po znaku D.
```

In [1]:
```
from IPython.display import Image, display
display(Image(filename='bigrams_matrix.png'))
```



In [ ]:

In [ ]:
```
#Výsledky úspěšnosti M-H algoritmu
#Na základě výsledků je zřejmé, že úspěšnost prolomení není ovlivněna
#délkou textu, ale především množstvím klíčů,
#které algoritmus přijme ve chvíli, kdy je pravděpodobnost < 1.
```

In [13]:
```
import pandas as pd
from IPython.display import HTML

data = {
    "ID textu": ["1", "2", "3", "4", "5", "6", "7", "8", "9", "10"],
    "Délka textu (znaky)": ["5229", "6246", "6688", "3903", "5322", "6027",
    "Chybné znaky": ["Ne", "Ano", "Ano", "Ne", "Ano", "Ano", "Ne", "Ne", "Ne
}

df = pd.DataFrame(data)
df
HTML(df.to_html(index=False))
```

Out[13]:

ID textu	Délka textu (znaky)	Chybné znaky
1	5229	Ne
2	6246	Ano
3	6688	Ano
4	3903	Ne
5	5322	Ano
6	6027	Ano
7	3163	Ne
8	2295	Ne
9	7241	Ne
10	10208	Ne

In [ ]: