

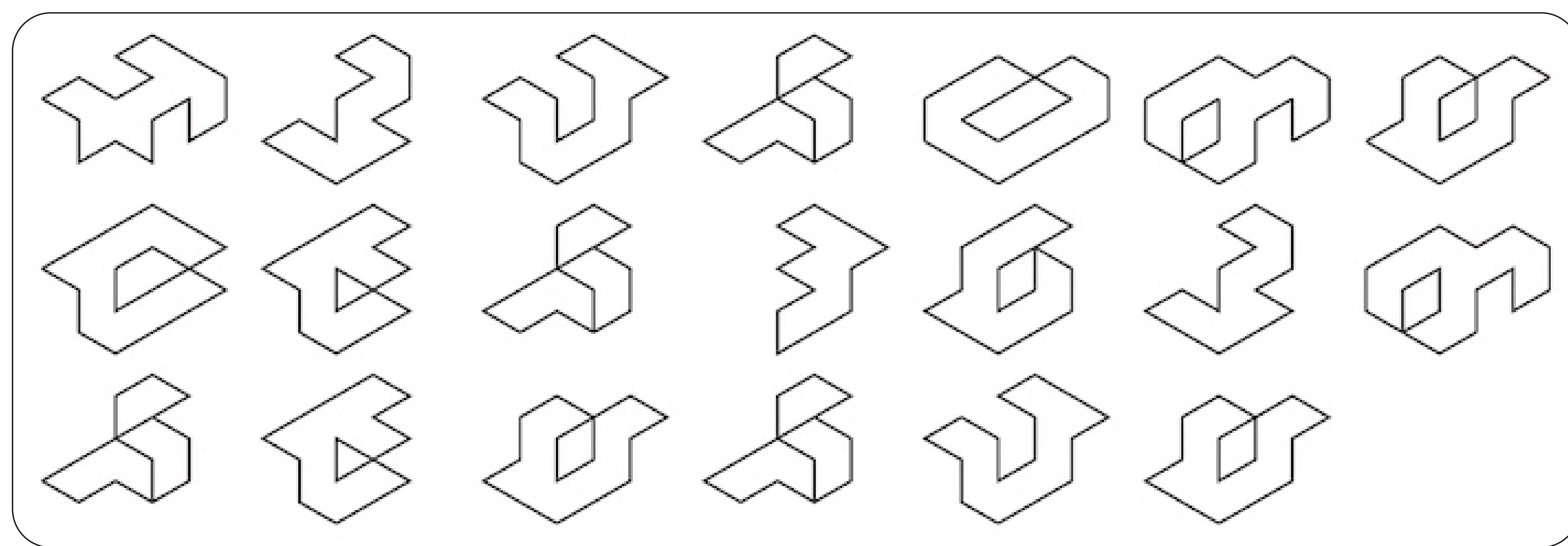
Automatický dešifrátor pro šifrovací hry

Filip Hlásek, 2014

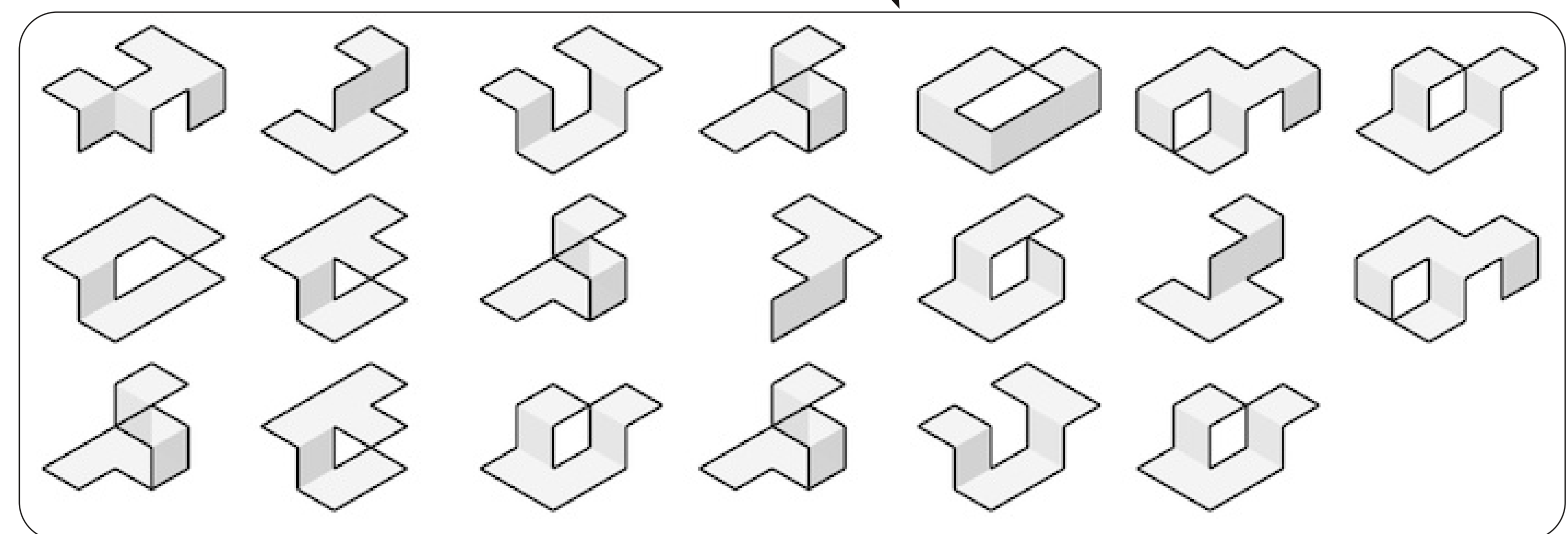


„Útvary v zadání nejsou dvourozměrné v trojúhelníkové síti, ale trojrozměrné. Když si je správně představíme, vidíme, že se všechny vejdou do kvádrů o rozměru dvě krát tři krychličky.“

Zadání šifry



Autorské řešení šifry



Všimneme si, že v zadání se některé obrázky opakují. Předpokládejme, že stejné útvary reprezentují stejné písmeno řešení. Přepíšeme si zadání tak, aby stejným obrázkům odpovídal stejný symbol, zatímco různým rozdílný.

Přepis zadání

abcdefghidjkbfdigdcg

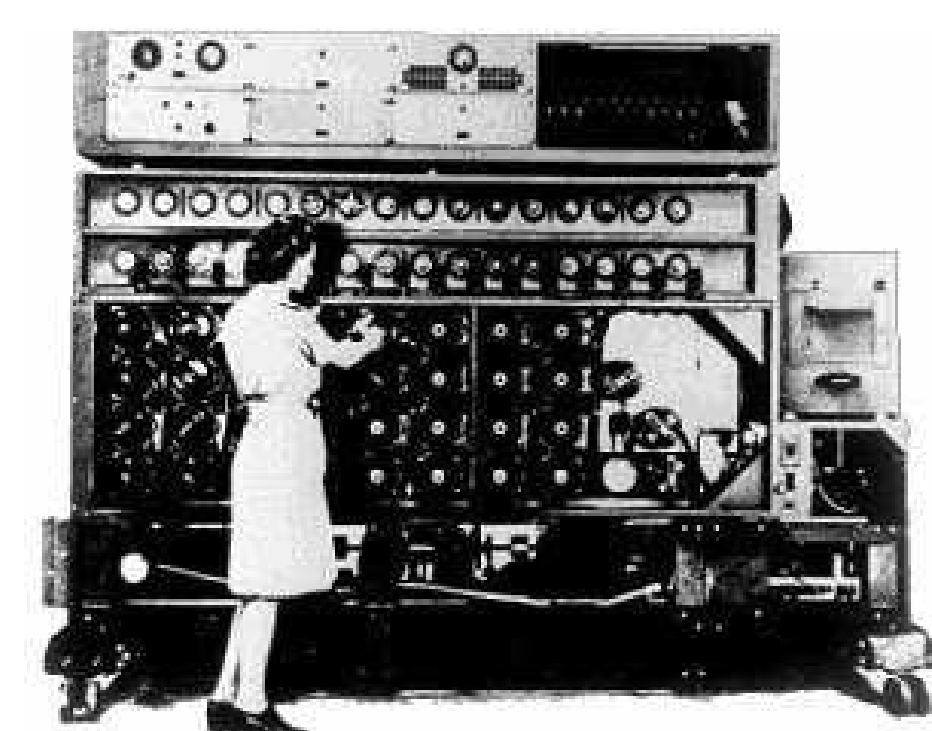
Automatický dešifrátor

je nástroj sloužící k luštění substitučních šifer. Používá pouze informaci o tom, která písmena jsou stejná, a která různá. Je určený především pro šifrovací hry a očekává tedy ve zprávě názvy zeměpisných míst, případně fyzické objekty, poblíž kterých může být umístěno následující stanoviště.

Jsou-li dešifrátoru navíc zadány zeměpisné souřadnice aktuální šifry, bude hledat následující šifru pouze v blízkém okolí. Výpočet se tím zrychlí a výsledky budou přesnější (některé z původních možností řešení mohou být příliš daleko).

Dešifrátor postupuje při lámání od začátku. Nejprve vyzkouší všechna slova, kterými může zpráva začínat, a pro každé z nich hledá následující slovo. Podobně *backtrackuje* přes *částečná řešení* šifry až do konce. Při výpočtu musí program brát ohled na to, aby skutečně stejným symbolům přiřadil stejná písmena a jiným různá. Pro každé nalezené řešení odhadne jeho pravděpodobnost a nakonec vrátí předem zvolený počet nejpravděpodobnějších řešení.

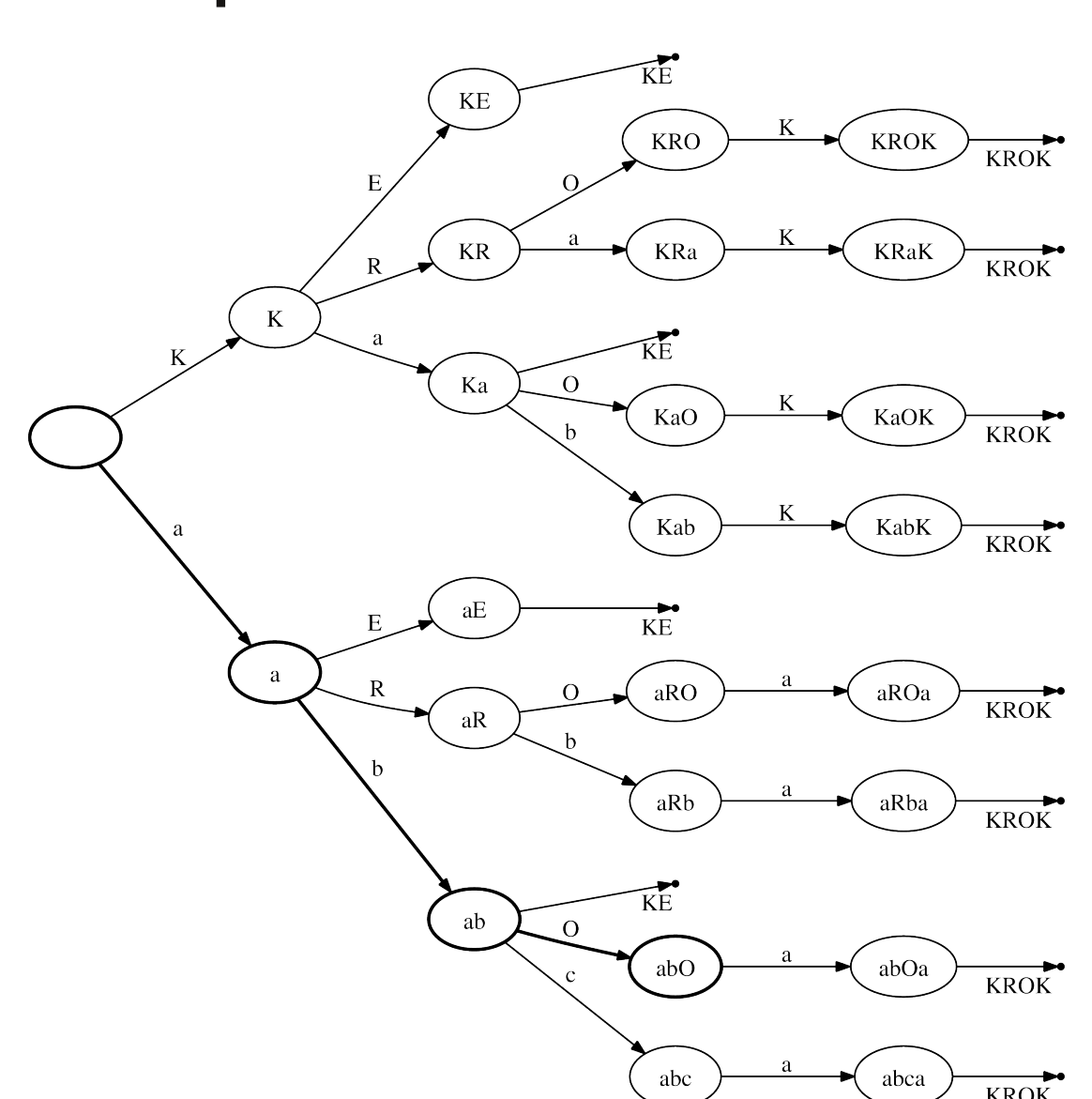
Tento zdánlivě jednoduchý problém nabízí několik zajímavých otázek.



Jak odhadnout pravděpodobnost řešení, když neznáme to správné?

Zkoumaný jazyk zachytíme Katzovým back-off modelem. K jeho trénování je potřeba co největší vzorek jazyka, ale šifer na šifrovacích hrách je bohužel dostupné pouze omezené množství. Navrhne tedy rozdělení slovníku na slovní třídy a spočítáme pouze pravděpodobnosti všech slovních tříd podmíněné předchozím textem. Na závěr pomocí nich odhadneme pravděpodobnosti jednotlivých slov

Jak lze rychle nalézt všechna slova, kterými může pokračovat rozpracované řešení?



Jaké pořadí zkoušení slov je nejvýhodnější?

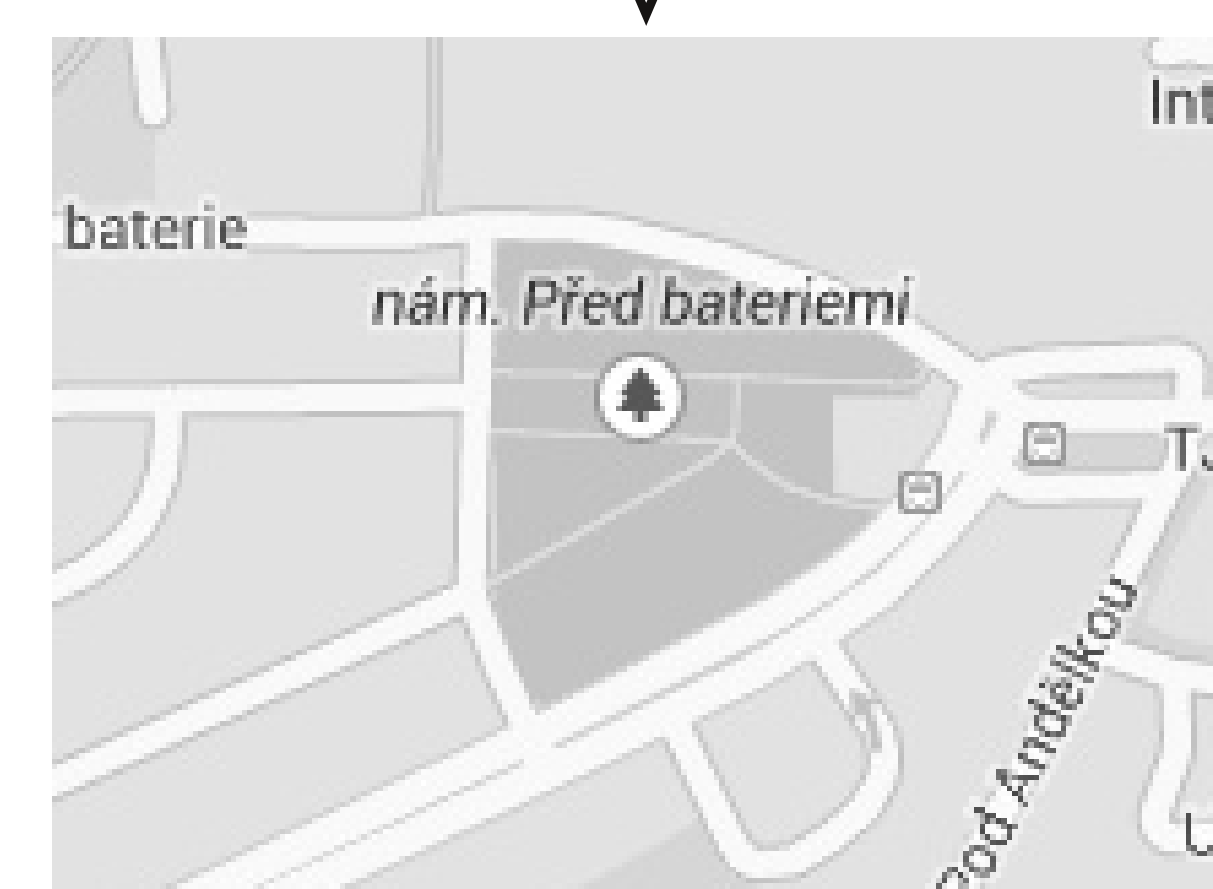
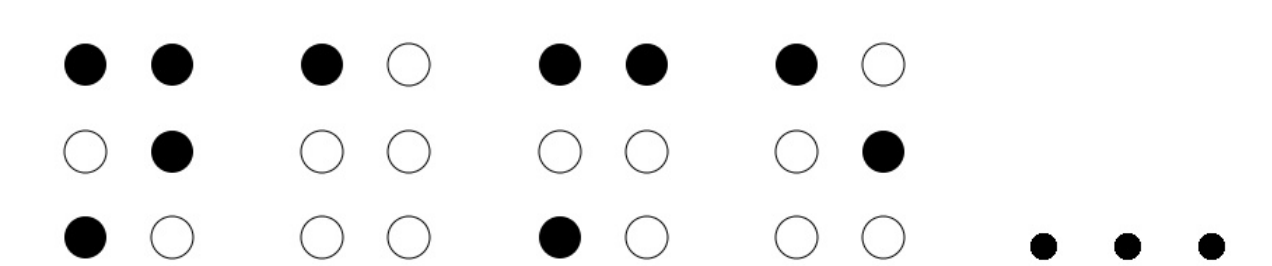
Navrhne hladovou heuristiku, která bude nejprve zkoušet slova s menším poměrem $\frac{\text{cena}(\text{slovo})}{\text{délka}(\text{slovo})}$, kde cena je vhodně zvolený odhad pravděpodobnosti, že bude následovat zkoumané slovo.

Správné řešení šifry

NAMESTIPREDBATERIEMI



„Čteme z těchto těles pouze horní stěnu a získáváme tajenku zapsanou Braillovým písmem.“



Kolik šifer dešifrátor rozluštil úplně správně? Přibližně **15 %** všech šifer. Úspěšnost roste při použití zeměpisných souřadnic a další šifry rozluštil téměř a člověk byl schopen z poskytnutých výsledků správné řešení odhadnout.

Zdroj uvedené šifry včetně řešení: <http://bedna.org/2012/sifry/13>

Dešifrátor si můžete vyzkoušet online na adrese: <http://atrey.karlin.mff.cuni.cz/~filip/decipher>

Tento projekt vznikl jako bakalářská práce pod vedením RNDr. Davida Marečka, Ph.D., Ústav formální a aplikované lingvistiky, Matematicko-fyzikální fakulta Univerzity Karlovy v Praze.