

Katedra informatiky
Přírodovědecká fakulta
Univerzita Palackého v Olomouci

BAKALÁŘSKÁ PRÁCE

Generátor křížovek



2024

Vedoucí práce:
doc. RNDr. Miroslav Kolařík,
Ph.D.

Petr Meixner

Studijní program: Informatika,
Specializace: Programování a vývoj
software

Bibliografické údaje

Autor:	Petr Meixner
Název práce:	Generátor křížovek
Typ práce:	bakalářská práce
Pracoviště:	Katedra informatiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci
Rok obhajoby:	2024
Studijní program:	Informatika, Specializace: Programování a vývoj software
Vedoucí práce:	doc. RNDr. Miroslav Kolařík, Ph.D.
Počet stran:	27
Přílohy:	elektronická data v úložišti katedry informatiky
Jazyk práce:	český

Bibliographic info

Author:	Petr Meixner
Title:	Crossword generator
Thesis type:	bachelor thesis
Department:	Department of Computer Science, Faculty of Science, Palacký University Olomouc
Year of defense:	2024
Study program:	Computer Science, Specialization: Programming and Software Development
Supervisor:	doc. RNDr. Miroslav Kolařík, Ph.D.
Page count:	27
Supplements:	electronic data in the storage of department of computer science
Thesis language:	Czech

Anotace

Byl vyvinut generátor křížovek, který umožňuje uživatelům vytvářet křížovky na základě zadaných parametrů. Cílem práce bylo poskytnout efektivní nástroj, schopný generovat křížovky v reálném čase. Byl využitý programovací jazyk C#, zpětné procházení při generování křížovky a redukování domény proměnných odstraněním hodnot, které nesplňují omezení. Program byl úspěšně testován a demonstroval schopnost generovat různorodé křížovky.

Synopsis

A crossword puzzle generator was developed, allowing users to create crosswords based on specified parameters. The aim of the project was to provide an efficient tool capable of generating crosswords in real time. The C# programming language was used, along with backtracking for crossword generation and domain reduction by eliminating values that do not meet constraints. The program was successfully tested and demonstrated the ability to generate diverse crosswords.

Klíčová slova: generátor křížovek; backtracking; šíření omezení

Keywords: crossword generator; backtracking; constraint propagation

Poděkování za odborné vedení, rady a konzultace patří vedoucímu práce doc. RNDr. Miroslavu Kolaříkovi, Ph.D. Dále děkuji pánům Mgr. Radek Janoščík, Ph.D. a Mgr. Jiří Zaccal, Ph.D. za naučení programovacího jazyka C#. Za jazykovou korekturu děkuji Ing. Petře Meixnerové.

Odevzdáním tohoto textu jeho autor místopřísežně prohlašuje, že celou práci včetně příloh vypracoval samostatně a za použití pouze zdrojů citovaných v textu práce a uvedených v seznamu literatury.

Obsah

1	Křížovka	7
1.1	Historie křížovek	7
1.1.1	Historie křížovek v České republice	8
1.2	Struktura křížovky	9
1.2.1	Legenda	9
1.2.2	Obrazec	9
1.2.3	Pozice	10
1.2.4	Vpisované výrazy	10
1.2.5	Tajenka	10
1.2.6	Nápověda	10
1.3	Druhy křížovek	10
1.3.1	Švédská křížovka	10
1.3.2	Britská křížovka	11
1.3.2.1	Britská vs. americká křížovka	11
2	Použité technologie	12
2.1	C Sharp	12
2.1.1	Klíčové vlastnosti C#	12
2.1.2	Komunita	12
2.2	LINQ	13
2.2.1	Základní koncepty LINQ	13
3	Pomocné metody	15
3.1	Backtracking	15
3.2	Constraint propagation	15
4	Generování křížovky	17
4.1	Vytváření slovníku	17
4.2	Zpracování vstupů	17
4.3	Britská křížovka	17
4.4	Švédská křížovka	18
4.4.1	Výhody	19
4.4.2	Nevýhody	20
5	Uživatelská příručka	22
5.1	Generování křížovky	22
5.2	Luštění křížovky	22
6	Programátorská příručka	23
6.1	Požadavky	23
6.2	Spuštění a otevření projektu	23
6.3	Struktura projektu	23
6.4	Třídy	23

Závěr	24
Conclusions	25
Literatura	26

1 Křížovka

Český svaz hádankářů a křížovkářů v [1] a Miloš Zapletal v [2] se zabývají křížovkou jako hrou, popisují pravidla tvoření a tipy na hraní s dětmi.

„Křížovka je vědomostně kombinační úloha založená na úplném nebo částečném křížování výrazů vpisovaných po znacích do obrazce a obsahující tajenku“ [1].

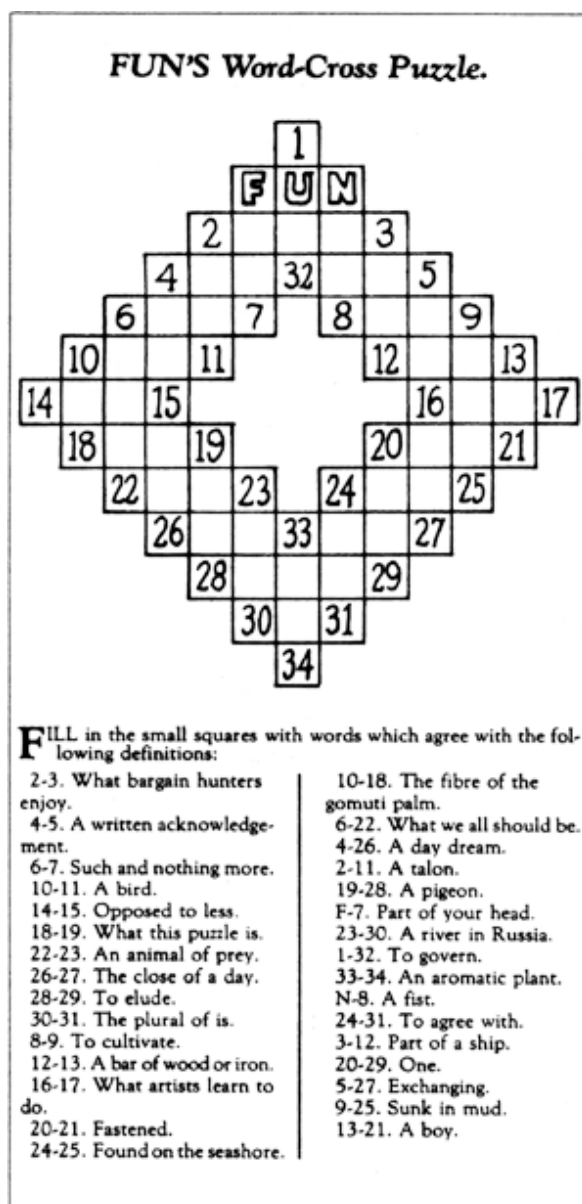
Křížovky jsou v dnešní době k nalezení snad v každém časopisu, novinách, dokonce i v samostatných knížkách na luštění. Ve světě se v novinách, časopisech či knížkách publikovalo jenom několik různých typů křížovek. Ovšem podle české klasifikace křížovek existuje přes 200 vzájemně odlišných forem a druhů křížovek. Značná část těchto variant byla vynalezena v České republice.

1.1 Historie křížovek

V [3], popř. Chromý v [4] popisují historii křížovek. O začátcích křížovek ve Velké Británii a v USA mluví na [5]. V [6] je zmínka o poplachu v britské armádě vyvolaném křížovkou na *Den D* a o prvním generátoru křížovek v [7].

Křížovky vynalezl Arthur Wyenne, původně jí nazýval „Word-Cross Puzzle“, tahle hádanka by se do češtiny přeložila jako *slovní kříž*. Poprvé ji uveřejnil 21. prosince 1913 v *New York World*. V roce 1923 už byly křížovky součástí většiny předních amerických časopisů, odkud se rozšířily i do Velké Británie. Ve Velké Británii mezi nejpopulárnější patřily křížovky z časopisu *The Sunday Times*. V USA to byly *The New York Times*, které ještě v dnešní době denně vydávají křížovky a wordle na svých webových stránkách.

Za druhé světové války v roce 1944 křížovky udělaly poplach ve Velké Británii, neboť autor křížovky pro *The Daily Telegraph*, Leonard Dawe, mezi slova zařadil tajná krycí jména („Gold“, „Sword“ a „Juno“, která označovala pláže) právě plánované operace týkající se vylodění v Normandii. Celý incident alarmoval britskou tajnou službu, ale nakonec se vše vyjasnilo jako náhoda. Od roku 1976 začal vznikat software k vytváření křížovek. Jedním z nejznámějších byl *Crossword Magic* pro Apple II. Už v 90. letech 20. století se s vytvářením křížovek spoléhali spíše na počítače. Díky tomuhle pokroku byly nově vzniklé křížovky kreativnější, zajímavější a obsahovaly více rozmanitých slov. V letech 1975–1980 a 1986–1987 byla v USA vysílaná televizní show *The Cross-Wits*, kde soutěžící spolu soupeřili v luštění křížovek. Křížovky se dodnes využívají jako zábavný způsob, jak se děti ve školách mohou učit slovní zásobu v různých jazycích a zemích, a zároveň slouží i dospělým jako příjemná volnočasová aktivita.



Obrázek 1: První křížovka od Arthura Wynnee

1.1.1 Historie křížovek v České republice

V [8] a popř. Zapletal [2] mluví podrobněji o historii křížovek v České republice. V Česku údajně byla otištěna první křížovka ve Sborníku hádanek Zlaté Prahy v roce 1899. Autorkou byla pražská učitelka Mlada Antošová z Modřan, ale i přesto je jako vynálezce křížovky v Guinnessově knize rekordů označován Arthur Wynnee. Mlada Antošová svoji první úlohu pojmenovala Čapka z písmen. V České republice byly křížovky opravdu oblíbené. K oblibě křížovek u nás přispěly velkou mírou Lidové noviny, které od roku 1925 začaly pravidelně vydávat křížovky. Této vášni propadl i spisovatel Karel Čapek. Sestavování a vyplňování křížovek

se brzy stalo koníčkem velkého počtu lidí. Občas vypadaly spíše jako malby. Luštění ale bylo trochu obtížnější. Neexistovala žádná pravidla, a proto autoři užívali slova v takovém tvaru, v jakém jej potřebovali nebo si jej vymysleli. Postupně se z těchto důvodů zpříšňovala pravidla na vytváření křížovek, na což dohlíželi všude po světě různé organizace. V tuzemsku je tou organizací Svaz českých hádankářů a křížovkářů (CSHAK), jenž vznikl v roce 1968. Podařilo se mu sjednotit pravidla, jednoznačně pojmenovat a označit všechny stávající i nové druhy hádanek, křížovek i vícesměrek a vytvořit pravidla pro tvorbu a řešení [8]. Češi a Slováci jsou nadále označováni jako národy křížovkářů. Tahle forma zábavy je u nás velmi tradiční a oblíbená. Křížovka není pouze zábavou, ale zároveň je zdrojem ponaučení o životě, nových vědomostí a také je mentálním sportem pro luštitelé [2].

1.2 Struktura křížovky

V ČSHAK [1] a [9] a popřípadě Skalický v [10] definují správnou strukturu křížovky. Každou křížovku jednoznačně určuje: název křížovky, obrazec, legenda, tajenka a vpisované výrazy. Součástí zadání jsou zpravidla název křížovky, obrazec a legenda. Tajenky a vpisované výrazy se musí vyplnit.

1.2.1 Legenda

„Legenda je souhrnem legendových výrazů“ [1]. Legendový výraz je výstižnou a aktuální charakteristikou, synonymem nebo definicí vpisovaného výrazu. Legendovým výrazem nesmí být přímo vpisovaný výraz. Obsahuje nápovědu pro hledané slovo v daném sloupci nebo řádku. Je to speciální typ buňky. Obtížnost křížovky je z části ovlivněna legendovými výrazy. Například ke slovu „míč“ můžeme napsat legendu „balón“ nebo také „habáněk“. Zatímco první výraz je jednoduché synonymum, které by napadlo téměř každého, výraz druhý pochází z nářečí a vyluštít ho vyžaduje více znalostí anebo zkušeností z luštění křížovek. Legenda by měla svou obtížností odpovídat cílové skupině řešitelů. Mezi běžné doplňky legendového výrazu patří i stylistické hodnocení vpisovaného výrazu, a to z hlediska citového zabarvení (familiárně, dětsky) a z hlediska spisovnosti (hovorově, básnicky, nářečně, slangově).

1.2.2 Obrazec

„Obrazec u každé křížovky je složený obrysem, sítkou a oddělovacími značkami“ [1]. V případě švédské i britské křížovky musí být sítko pravoúhlé (musí tvořit mřížku). Plochy, které vznikají mezi těmito sítkami, se označují jako políčka. Oddělovací značky, které vymezují pozice pro vpisování slov, jsou ve švédských křížovkách tvořeny celým políčkem, do kterého je vepsán legendový výraz. V těchto políčkách musí být uveden legendový výraz pro jeden nebo oba směry doplňování.

1.2.3 Pozice

Pozice je tvořena souvislou řadou políček ve vodorovném nebo svislém směru. Na každou pozici je možné zapsat pouze jedno slovo, protože se ale musí slova křížit, kříží se přes každé políčko dvě pozice – jedna svislá a jedna vodorovná.

1.2.4 Vpisované výrazy

Vpisované výrazy jsou slovní výrazy nebo slova, která luštitel zapisuje do prázdných pozic v souladu s obdržnými informacemi z legendového výrazu. Výrazy by měly pocházet primárně z běžné slovní zásoby. Kladně se hodnotí vyváženost křížovky z hlediska skladby a pestrosti výrazů, například by křížovka neměla obsahovat nadměrně moc vlastních jmen a podobně. Toto pravidlo neplatí v případě, že se jedná o monotematicky zaměřenou křížovku. Na využívání vulgárních výrazů je přihlíženo velice negativně.

1.2.5 Tajenka

„Tajenkou (tajenkovým výrazem) je souvislý text, část souvislého textu nebo samostatný pojem (např. odborný výraz)“ [1]. Vyřešení tajenky odhalí luštiteli nějaké tajemství nebo citát. Pokud je tajenka nesouvislá, tak je označena čísly, které určují pořadí, v jakém má být vyřešená tajenka čtena. Tajenka většinou bývá zvýrazněná nějakou barvou, nejčastěji žlutou. Za ideální délku tajenky se považuje 12 % z celkového počtu písmen v křížovce. Tajenka by se neměla křížit sama se sebou.

1.2.6 Náповěda

Náповěda obsahuje několik vpisovaných výrazů, které může luštitel využít. Avšak u náповědy není uvedena správná pozice, ani legendový výraz k těmto vpisovaným výrazům.

1.3 Druhy křížovek

V [8] a [9] definují druhy křížovek. Existuje více než 200 typů křížovek, velká část vznikla v České republice. Různé kouty Země mají svoji oblíbenou variantu. Tato práce se bude nejvíce zabírat variantou britskou a švédskou, která je nejrozšířenější v kontinentální Evropě.

1.3.1 Švédská křížovka

Švédská křížovka je nejrozšířenějším typem křížovky v České republice. Právě tahle verze se často vyskytuje v novinách a časopisech. Od ostatních variant se odlišuje vepsáním legendy přímo do mřížky křížovky, proto se také označuje jako křížovka s vypsanou legendou. U ostatních křížovek jsou legendové výrazy

umístěny mimo mřížku křížovky, ty si poté řešitel musí vyhledat podle souřadnic políčka. Cílem křížovky je odtajnit skrytý text tajenky.

1.3.2 Britská křížovka

Britská křížovka je nejpobulárnějši v Austrálii a ve Velké Británii, kde také vznikly. V britské křížovce je spoustu černých políček. Zhruba polovina písmen v každém slově je křížena jiným slovem. Legendové výrazy jsou zapsané mimo křížovku, luštitel si musí tuhle pomůcku vyhledat pomocí čísla. Je s oblibou vyžívaná učitelkami angličtiny na základních školách k procvičování slovíček.

1.3.2.1 Britská vs. americká křížovka

Britská křížovka se často plete s americkou verzí. Americký typ má ale méně černých políček a každé písmeno každého slova je křížené s jiným slovem. Takže je možné vyplnit tuhle křížovku jenom díky vyplňování svislých (vodorovných) slov. Vizualně je americká křížovka podobnější té švédské, ale legendy nejsou uvedeny přímo v křížovce ale mimo, podobně jako u britské.

2 Použité technologie

2.1 C Sharp

V [11], Montgomery v [11], popř. Hejlsberg [13] se zabývají jazykem C#. C# (C Sharp) je moderní univerzální vysokoúrovňový objektově orientovaný programovací jazyk. Byl vyvinutý Microsoftem jako součást jeho iniciativy .NET. Byl schválený standardizačními komisemi ECMA a ISO. Od svého vydání v roce 2000 si C# získal na popularitě a stal se klíčovým hráčem na poli programovacích jazyků především v oblasti webových, desktopových a mobilních aplikací.

C# je vhodný pro psaní aplikací pro hostované i vestavěné systémy, od velmi velkých, které používají sofistikované operační systémy, až po malé jednoúčelové funkce. Přestože jsou aplikace C# zamýšleny jako ekonomické s ohledem na paměť a zpracování požadavků na napájení, jazyk nebyl zamýšlen tak, aby přímo soutěžil ve výkonu a velikosti s C nebo jazykem symbolických instrukcí.

2.1.1 Klíčové vlastnosti C#

- Objektově orientovaný – podobně jako například Java a nebo C++ je objektově orientovaný, což podporuje opětovné použití kódu a modulární design.
- Moderní syntax – C# nabízí čistou a výraznou syntaxi, která vývojářům usnadňuje snadné čtení a psaní kódu.
- Integrace s .NET – Jednou z největších předností C# je jeho integrace s .NET frameworkem, který poskytuje rozsáhlou knihovnu tříd a funkcí, což usnadňuje rychlý vývoj aplikací.
- Multiplatformní – S příchodem .NET Core a .NET 5 (a novějších verzí) lze vyvíjet aplikace C# a provozovat je na různých platformách.
- Silně typovaný – C# je silně typovaný jazyk, který může pomoci zachytit chyby během kompilace, což vede k vytvoření robustních aplikací.
- Nepřetržitý vývoj – V průběhu let prošel C# četnými aktualizacemi, které přidaly nové funkce, jako je LINQ (Language-Integrated Query), async/await na asynchronní programování.

2.1.2 Komunita

Jazyk je jen tak dobrý, jak dobrý je jeho ekosystém a komunita. V průběhu let se komunita C# exponenciálně rozrostla. Platformy jako StackOverflow a GitHub se hemží C# projekty, diskuzemi a řešeními. C# je 8. nejpoužívanějším programovacím jazykem, používá jej necelých 28 % programátorů na světě [13]. Klíčové bylo to, že Microsoft nabízel nástroje jako Visual Studio, které poskytuje kvalitní vývojové prostředí pro vývojáře C#.

Vzestup open source projektů, zejména .NET Core, dále posílil popularitu tohoto jazyka. Nyní mohou vývojáři aktivně přispívat k jeho vývoji, navrhnout nové funkce a zajistit, aby C# zůstal jazykem, který je tvořen vývojáři pro vývojáře.

2.2 LINQ

V [15], popř. v [16] mluví o LINQ. Jazyk Language Integrated Query, obecně známý jako LINQ, představuje významný pokrok v možnostech jazyka C#. Poskytuje konzistentní model pro práci s daty napříč různými zdroji a formami, čímž mění způsob, jakým vývojáři pracují s daty v aplikacích. LINQ se integruje přímo do jazyka C# a umožňuje vývojářům provádět složité operace pomocí stručného a čitelného kódu.

LINQ rozšiřuje jazyk o dotazovací výrazy, které se podobají příkazům SQL a lze je použít k pohodlnému získávání a zpracování dat z polí, dokumentů XML, relačních databází a zdrojů dat třetích stran. Částečně LINQ existuje i v jazycích: PHP, JavaScript a TypeScript, ale žádná z těchto alternativ není tak zpracovaná jako právě verze pro C#. LINQ je důkazem evoluce C#, poskytuje výkonné, konzistentní a čitelné možnosti dotazování dat. Je příkladem filozofie jazyka zvýšit produktivitu vývojářů bez obětování výkonu.

V práci je LINQ využito k výběru všech slov ze slovníku, která splňují aktuální omezení (již vepsaná písmena a délku) na pozici pro legendu křížovky.

2.2.1 Základní koncepty LINQ

LINQ se zabývá kolem několika klíčových konceptů:

- Rozšiřující metody – umožňují přidávání nových metod k existujícím typům bez jejich úprav.
- Inline operace – velmi silné a flexibilní, umožňují efektivní a čitelnou práci s daty přímo v kódu bez nutnosti použití explicitních smyček nebo složitých podmínek.
- Lambda výrazy – stručný způsob definice anonymních funkcí pomocí syntaxe, která může být snadno použita pro inline operace.
- Stromy výrazů – místo okamžitého provádění operací někteří poskytovatelé LINQ vytvářejí strom výrazů, který představuje operaci a může být vykonán nebo přeložen do jiného jazyka, například SQL pro databáze.
- Konzistence – jednou z největších sil LINQ je schopnost poskytnout konzistentní způsob dotazování bez ohledu na zdroj dat.
- Čitelný kód – LINQ dotazy jsou často čitelnější než jejich protějšky, což umožňuje vývojářům vyjádřit složité operace stručným způsobem.

- Podobnost s SQL – díky podobnosti s SQL výrazy těží ze stejných výhod. Příkazy jsou srozumitelné i pro neznalého člověka, protože se dají napsat lidskou řečí.

3 Pomocné metody

3.1 Backtracking

V Bell [17] a v Korf [18] popisují využití backtrackování. Backtracking je algoritmické paradigma, charakterizované jako postupné řešení problémů pomocí zkoumání všech možných stavů prostoru řešení. Je to metoda, považovaná za brute-force techniku. Opírá se o rekursivní průchod stromem rozhodnutí, kde každý uzel představuje bod rozhodování a jeho potomci reprezentují dostupné možnosti pokračování od daného bodu. Specifickým rysem backtrackingu je jeho schopnost *prořezat se* a vrátit se (backtrackovat) k poslednímu rozhodovacímu bodu, když aktuální cesta nepřináší řešení, což umožňuje efektivní prohledávání prostoru bez nutnosti opakovaně prozkoumávat neúspěšné cesty.

Backtracking se nejčastěji používá v kontextu NP-těžkých problémů, jako jsou problémy rozhodování a optimalizace, kde je prohledávání prostoru řešení nezbytné pro nalezení optimálního nebo validního řešení. Příkladem je klasický problém osmi dam, jehož řešení demonstruje využití backtrackingu k identifikaci všech možných způsobů, jak umístit osm šachových dam na šachovnici tak, aby se žádné dvě navzájem neohrožovaly. Další aplikací je řešení logických her, jako je sudoku, kde backtracking slouží k systematickému prozkoumávání všech možných kombinací čísel, dokud není nalezeno řešení splňující všechna pravidla hry.

Přestože je backtracking výkonný v řešení určitých typů problémů, jeho efektivita může být limitována exponenciálním nárůstem prostoru stavů u složitějších problémů. Tento problém se řeší technikou prořezávání, která umožňuje algoritmu eliminovat nesplnitelné cesty předčasně, a tím značně redukovat počet stavů k prozkoumání. Další metody zahrnují heuristické přístupy, které směřují prohledávání k nejvíce slibným částem prostoru řešení, čímž se dále zvyšuje efektivita backtrackingových algoritmů.

V programové části této práce je metoda backtrackingu použita pro generování švédských křížovek, při které se procházejí všechna možná slova na konkrétním místě legendy. Pro urychlení celého procesu je rovněž aplikována metoda prořezávání, která se snaží ukončit větev prohledávání v případě, že program opakovaně zkouší slova, která již dříve neúspěšně vyzkoušel.

3.2 Constraint propagation

V Ginsberg [19] ukazuje užití constraint propagation ke generování křížovek. Constraint propagation je technika používaná v umělé inteligenci a počítačové vědě k efektivnímu řešení problémů s omezeními tím, že dynamicky redukuje domény proměnných odstraňováním hodnot, které nesplňují omezení. Tento proces nejenže zjednodušuje problémy eliminací nemožných řešení, ale také výrazně zrychluje hledání platných řešení tím, že snižuje vyhledávací prostor, což je klíčové pro aplikace vyžadující rychlé a efektivní rozhodování, jako jsou logické hry, plánování a rozvrhování, a počítačové vidění. Praktická významnost této

metody je podtržena jejím širokým využitím ve výzkumu a průmyslových aplikacích, kde přispívá k vývoji pokročilých UI systémů schopných řešit složité úlohy s vysokou účinností.

Využití constraint propagation se rozprostírá přes širokou škálu oblastí, od optimalizace procesů v průmyslové výrobě až po vývoj inteligentních algoritmů pro řešení populárních logických her. Tato metoda poskytuje robustní rámec pro řešení problémů, kde je potřeba navigovat v komplexním prostoru omezení a najít řešení splňující všechna kritéria, což dokazuje její nezastupitelnou roli v pokroku umělé inteligence a počítačové vědy.

4 Generování křížovky

4.1 Vytváření slovníku

Všechna slova jsou uložena v databázi. Každé slovo musí mít k sobě odpovídající legendu, obtížnost a jazyk. Celkem je v databázi uloženo asi 70 000 slov z různých odvětví. Slova jsou různých délek, ale nejkratší možná délka slova je 2. Slova délky 1 nemají moc velký přínos k zábavnosti křížovky. Jednoznačně je určuje slovo, kterým je jednopísmenné slovo kříženo a například v angličtině je 1 : 26 pravděpodobnost uhodnutí i bez legendy.

Při vytváření slovníku je pořadí slov zamíchané, aby byla zaručena různorodost výsledných křížovek. Všechna písmena ve slovníku jsou velká a bez čárek a bez kroužku, například *Á* se změní na *A*. To kvůli tomu, aby se snížil počet možných písmen a tím značně zrychlil proces generování křížovky a zároveň se význam slov zůstává ve většině případů nezměněný. Uživatel může psát malá písmena i klidně s čárkami, bude to přepsáno za něj.

4.2 Zpracování vstupů

Po spuštění webové aplikace si uživatel může vybrat mezi jazykem českým nebo anglickým, typ křížovky, obtížnost a také její velikost vyjádřenou v počtu políček na řádku a ve sloupci. Ajax požadavkem se pošlou tato data do backendové části v C#, kde se vytvoří slovník slov ve správném jazyce a obtížnosti. Poté se vytvoří mřížka křížovky o uživatelem zadaných rozměrech a začne se generovat vybraný typ křížovky.

4.3 Britská křížovka

Po obdržení požadavku na generování křížovky algoritmus vytvoří dvojrozměrné pole řetězců s rozměry danými uživatelem. Toto pole reprezentuje prázdnou křížovku, do které bude algoritmus postupně vkládat slova.

Algoritmus začne procházet každé políčko křížovky, zkoumá nejdříve možnosti v horizontálním směru a následně ve směru vertikálním. Pro každé políčko se pokouší najít vhodné slovo, které by bylo možné na danou pozici vložit. K tomu využívá LINQ dotazu, který vyhledává slova na základě písmen, která by nové slovo mělo obsahovat. Před vložením slova do křížovky je potřeba zkontrolovat, zda vhodné slovo splňuje daná kritéria. Musí se ujistit, že vložením slova nevytváří nesmyslná nebo nechtěná slova v křížovce a že nezmění nebo nerozloží slova, která už byla dříve vložena. Pokud je vše v pořádku, slovo se vloží do křížovky.

Po vyplnění křížovky algoritmus prověří celou plochu a zkontroluje, jestli v ní neexistují osamocená slova, která nejsou křížena s jinými. Pokud taková slova objeví, odstraní je. Na závěr zhodnotí, zda je křížovka dostatečně zaplněná. Pokud ano, uživateli se zobrazí hotová křížovka. V opačném případě algoritmus začne znovu od prázdné mřížky. Celkem tento proces projde maximálně desetkrát a uloží si nejlepší výsledek, který nakonec vrátí jako výslednou křížovku.

Při implementování algoritmu na generování britské křížovky jsem uvažoval nad používáním backtrackingu s constraint propagation, ale dle mého názoru jsou tyto metody zbytečně výpočetně pomalejší. Navíc kvůli constraint propagation by se snížil počet možných výsledků vygenerované křížovky. To je velká výhoda pro generování švédského typu, ale britský typ není tak výpočetně náročný, a proto je větší rozmanitost ku prospěchu.

^{1,1} U	P	² I	C	³ K	E		^{2,4} Č	A	⁵ D
Č		V		R			M		O
³ E	S	O		⁴ Č	A	P	E		D
B		⁵ N	A	M			L		Ě
⁶ N	A	I		⁷ A	R	T	I	K	L
I		N		Ř			C		A
⁸ C	L	O			^{9,6} X	V	I		N
E			^{10,7} B	Y	V				Ě
			A		I				
¹¹ K	R	A	V		¹² I	N	F	I	X

Legendy vodorovné:

- 1 VYRÁBĚNĚ V ÚPICI
- 2 AFRICKÝ STÁT
- 3 DOKONALÉ PODÁNÍ (TENIS)
- 4 POPADÁVÁ HOVOR.
- 5 MY 3. PÁD
- 6 VZOREC JODIDU SODNÉHO
- 7 NABÍZENÉ ZBOŽÍ
- 8 DOVOZNÍ DAŇ
- 9 ŘÍMSKÁ ČÍSLICE 16
- 10 PŘECHODNÍK SLOV. BÝT
- 11 NEMOC ŠÍLENÝCH ???
- 12 VPONA

Legendy svislé:

- 1 ŽÁKOVA KNIHA
- 2 OVLÁDÁNO IVONOU
- 3 HOSPODSKÝ
- 4 PARAZITI NA SLEPICÍCH
- 5 DOKONČENĚ
- 6 ŘÍMSKÁ ČÍSLICE 17
- 7 PRACUJI (K KOMIKOVI)

Obrázek 2: Vygenerovaná britská křížovka

4.4 Švédská křížovka

Proces generování křížovky algoritmicky začíná inicializací pracovního prostoru (mřížky), který má rozměry dle parametrů zadaných uživatelem. V této počáteční fázi algoritmus stanoví rámec pro umísťování slov, a to definicí omezení ve formě předpřipravených míst pro legendy včetně orientace (horizontální či vertikální), kudy budou slova vedena. Tento krok je zásadní pro strukturální uspořádání

křížovky. K vytváření omezení je využita metoda constraint propagation, viz 3.2.

Do křížovky se nejdříve vloží tajenka, ta se nesmí křížit, a proto jsou k sobě části tajenky vždy rovnoběžné. Náhodně se vloží horizontálně nebo vertikálně. Následně se do křížovky zapíše místa pro legendu do krajů, tedy do levé a horní strany mřížky. Protože se nepoužívají slova délky 1 a rozměry křížovky mohou být větší než délka slova, tak se do 2. a 3. sloupce i řádku musí vepsat další legendy. Ze stejného důvodu je potřeba vložit legendy i do posledního a předposledního řádku i sloupce. Nakonec se zapíše místa pro legendy do vnitřní části mřížky. Zapisují se tak, aby z každé legendy mohlo vést slovo aspoň jedním směrem a nikde nevznikal prostor pro slovo délky 1.

Algoritmus klade důraz na náhodnost, čímž zaručuje, že každá vygenerovaná křížovka je jedinečná a poskytuje luštiteli vždy nový a originální zážitek. Náhodnost je využita jak při promíchání slov ve slovníku, tak i při vytváření mřížky, viz 4.1.

Následně algoritmus přechází do fáze dynamického vyplňování, kde klíčovým faktorem je identifikace míst s nejvyšším počtem omezení, tedy takových, která jsou již částečně vyplněna písmeny z již vepsaných slov. Tato místa mají přednost při výběru dalšího slova z potenciálních kandidátů. Algoritmus systematicky prohledává slovník ve snaze nalézt slova, která odpovídají daným omezením. Uloží si kopii aktuální křížovky, vloží vhodné slovo na dané místo a uloží momentální stav křížovky do paměti. Vráť se k původní křížovce a vloží postupně všechna slova vhodná pro toto místo. Poté si načte nejaktuálnější stav křížovky z paměti a hledá místo s největším počtem omezení, až dokud není celá křížovka vyplněná.

V procesu vyplňování může algoritmus narazit na situaci, kdy další postup není možný kvůli nesouladu mezi již umístěnými slovy a zbývajících možnostmi slovníku. V takovém případě dochází k backtrackingu, tedy kroku zpět, který umožňuje algoritmu vrátit se na předchozí větvení procesu a zkusit jinou cestu vyplnění. Tento proces je podpořen ukládáním již prozkoumaných kombinací písmen a jejich souřadnic, což efektivně zabraňuje redundantnímu procházení stejných scénářů. Toto procházení slepých cest prochází celou křížovku, aby k tomuhle prořezání došlo co nejdříve a ušetřil se tím drahocenný čas.

Algoritmický proces generování křížovky pokračuje iterativně, až dokud nejsou všechna vyhrazená místa pro legendy úspěšně vyplněna. Tento konečný krok signalizuje úspěšné dokončení generování křížovky.

4.4.1 Výhody

- Redukce stavového prostoru – constraint propagation umožňuje výrazné snížení stavového prostoru tím, že jasně definuje, kde povedou slova a kolik jich bude.
- Prořezávání – díky prořezávání se rychle ukončuje zkoumání již nalezených nesplnitelných cest.

- Efektivita – díky constraint propagation lze dosáhnout rychlejšího nalezení řešení než při použití samotného backtrackingu.
- Různorodost křížovek – díky důrazu na náhodnost je velká různorodost ve vygenerované křížovce, po každé bude jiná
- Intuitivnost algoritmu – v základu velmi intuitivní algoritmus.

4.4.2 Neýhody

- Časová náročnost – i když má stále dobrou výpočetní rychlost (vzhledem k náročnosti problému), tak je o něco nižší, než za použití mask. Ty ale už byly použity v několika jiných pracích.
- Podobná slova – může nastat situace, že v jedné křížovce mohou být podobná slova, např. *kolo* a *kola*. Takováhle situace se hodnotí záporně [1].
- Omezení vstupu – generování švédské křížovky funguje spolehlivě pouze pro český jazyk. Pro anglický jazyk nenajde pokaždé řešení, protože pro vygenerování švédské křížovky je potřeba velký počet slov ve slovníku.
- Nespolehlivá obtížnost – kvůli malému počtu obtížností může být výsledná švédská křížovka *normální* s obtížností velmi rozdílná. Jednou může být její vylustění snadné, jindy zase složité.

Tisk		Napověda		Zkontrolovat										
	VYNES ORTEL	PÍSMENO (AZBUKA)	KŘEMÍK ZKR.		VYDÁVAJE ÚTRPNÉ ZVUKY	PORUČÍK ZKR.	AMERICIUM ZKR.		ČTENÍ PRŮJEMNÍ WILSON	ČÁST KOČÁRU	PROTŘELÝ ČLOVĚK HOVOR.		ZPRAVUJÍ O NEBEZPEČÍ	ZLATÁ MINCE
JSI SI (SPISOVNĚ)	S	I	S	UKRAJINŠTÍ TERORISTÉ	U	P	A	VYŠŠÍ ODOBORNÁ ŠKOLA	V	O	Š	VZOR ZKR.	V	Z
SNĚŽUJE ŠÍŘKU	U	Ž	I	KOPEN SLOVA NAPOKROKOVY ZPŮSOB VZPŘÍMAVÍ	P	O	M	ČTENÍ II V SLAVI ŠTÁSTNÉ ČÍSLO	I	J	I	HLINÍK ZKR. OFICIÁLNÍ AMALKA	A	L
ČTENÍ SKUPINY DE	Ď	E	VYKULENÉ JIKRY (SPAM, JINOUVÁŽNĚ)	T	Ě	R	PRACOVY FIDELITY LENEKA ZOBŘELENÍ POKROKOVNĚ	S	L	DOMÁCKÝ BARBORA	B	A	R	A
	DRŽ NĚCO A JDI	SUMERNÉ MĚSTO ZOKRAVKA (INTEKSIČNÍ VODA)	U	R	O NÁŘEČ.	PŘÍTEL ČLOVĚKA ASUBKÉ KOMIKY	P	E	S	VYVOLÁVÁJÍ SMUTNOST SLEHNY K PRŮJEMNĚ	R	M	U	Ť
TAJENKA 1	N	E	C	H	V	A	L	D	N	E	DOMÁCKÝ ALBINA PRACOVY (PEDAGOG)	A	J	A
ZKRATKA TOALETNÍ VODA	E	D	T	ONEN 7. PÁD TUDÍŽ HOVOR.	O	N	I	M	PTÁK PŘÍBUZNÝ ČEJCE	K	U	L	I	K
CHÝLIT SE	S	P	Ě	T	ŘÍMSKÁ ČÍSLOCE 4	I	V	SUMIVÉ VÍNO	SÍDLA ZRAKU PRŮJEMNOST VÍRE KVAL.	O	Č	I	KOLOIDNÍ KÁSE	SOLNIVĚZDÍ BERAN (ZKR.)
	DRUH VARIČE	ZÁPORKA DYLÁČKALY PŘEVODOVKY	N	E	ROZETRACÍ LER. PŘÍKLADNĚ	M	A	S	T	KONCOVKA 3. PÁD PL. KOST	ROZDÍL VÁLLE. JAK KVAL. ZTRÁCETI SE	E	G	A
TAJENKA 2	P	Ř	E	D	V	E	Č	E	R	E	M	PÍSMENO R	E	R
RADIUM ZKR.	R	A	TYT ANIL. TV PŮRAJ KYSLIČNÍK	A	Z	PACIFIČKÉ KOLOŠY	ROZDÍLA ISLA V. SNIVOSNÍ YARD ZKR.	K	A	M	I	PŘÍKLONKA LAKOVY NA OSLAVU	L	I
OSLOVENÍ IDY	I	D	O	OČISTIT POVRCH VODOU	O	M	Y	T	NORMA GOLFOVÉHO HRISTĚ	MUŠÍ ZVUK DROGA	Z	Z	DOMÁCKÝ ELENA	SPÍLAL
SMĚS	M	I	X	SKUPINA RODIN CÍN ZKR.	R	O	D	HLÁSKOVITÝ VESTULKA ČÍSLO CIZÍ ZKR.	P	L	E	V	E	L
DOTĚRNĚ POOLEZAVÁ	U	L	I	S	N	A	PŘÍHODILA SE. VZNÍKLA	N	A	S	T	A	L	A
AUSTRÁLSKÉ MĚSTO	S	Y	D	N	E	Y	RAĐOVÁ ČÍSLOCE	O	R	D	I	N	A	L

Obrázek 3: Vygenerovaná švédská křížovka

5 Uživatelská příručka

5.1 Generování křížovky

Po spuštění webové aplikace má uživatel možnost, buď nahrát svojí uloženou křížovku a nebo vygenerovat novou. Pro vygenerování nové křížovky je nutné, aby uživatel zadal výšku a šířku křížovky, typ, jazyk a obtížnost. Povolené rozměry křížovky jsou od 10 do 20 políček v obou směrech. Tohle omezení je jak kvůli rychlosti, tak i kvůli přehlednosti při tisknutí. Obtížnost má 3 úrovně, *jednoduchou* pro malé děti, *normální* pro libovolného člověka a *expert* pro velmi zkušené luštitelé. Pro generování švédské křížovky je nejvhodnější obtížnost *normální* a jazyk český, pro ostatní kombinace obtížnosti a jazyka je ve slovníku méně slov, a proto generování trvá déle a nebo se ani nepovede.

Poté se uživateli zobrazí ukazatel, že probíhá generování křížovky. Generování probíhá na pozadí, doba je závislá na typu a velikosti křížovky, maximálně však jednotky minut.

5.2 Luštění křížovky

Po vygenerování se křížovka zobrazí uživateli přímo na webu. Dle zvoleného typu budou legendy přímo vepsané v křížovce pro švédskou verzi a nebo pod ní pro britskou verzi.

Nejsou povoleny jiné znaky než písmena. Písmena v češtině jsou velká a bez čárek a bez kroužku nad písmenem *U*. Uživatel se tímhle zaobírat nemusí, bude to přepsáno za něj. Háčky je nutné napsat. Písmeno *CH* se píše do jednoho políčka ale také jenom pro češtinu. Kromě písmene *CH* (pouze v českém jazyce) je možné do každého políčka psát maximálně 1 znak. V křížovce je možné se pohybovat pomocí šipek. Pokud políčko obsahuje 2 odkazy na legendy, tedy začínají v něm 2 slova, tak 1. číslo odkazuje na vodorovnou legendu, 2. číslo na legendu svislou.

Uživatel má možnost si křížovku vytisknout i se svým postupem. Pro lepší tisknutí je lepší se navigovat do sekce Měřítka a tam vybrat možnost *Přizpůsobit podle tisknutelné oblasti*. Postup při luštění bude automaticky ukládán a je tedy možné se později vrátit zpět k rozluštěné křížovce. Také uživatel může využít možnosti nápovědy, kde má možnost si vybrat políčko, se kterým potřebuje poradit, a nebo i ověřit správnost svého vyplnění políčka. Po vyplnění celé křížovky si také může uživatel nechat překontrolovat celé své řešení. V případě úspěšného vyluštění bude uživateli nabídnuto vygenerování další křížovky. Tyto usnadnění se provádí pomocí tlačítek nad křížovkou viz. 3.

6 Programátorská příručka

6.1 Požadavky

Pro správnou funkčnost a vývoj generátoru křížovek je nezbytné mít k dispozici určité softwarové vybavení a hardware. Nejdříve je potřeba mít nainstalované Visual Studio 2019 nebo novější verzi a .NET Framework 4.7.2 nebo novější. Zároveň je potřeba mít k dispozici IIS Express ke spuštění lokálního serveru. Co se týče hardwarových požadavků, doporučuje se mít minimálně 8 GB RAM a 1 GB volného místa na disku pro uložení všech potřebných souborů a knihoven.

6.2 Spuštění a otevření projektu

Aby bylo možné aplikaci spustit, je nutné otevřít soubor *startApp.bat*, který automaticky nastaví potřebné prostředí a spustí aplikaci. Tento soubor je k naleznutí ve složce *Application*. Pro otevření projektu ve vývojovém prostředí je potřeba navigovat do složky *SourceCode*, poté do *CrosswordGenerator* a zde otevřít soubor *CrosswordGenerator.sln*. Tímto způsobem lze načíst celý projekt ve Visual Studiu, kde je možné jej upravovat a ladit.

6.3 Struktura projektu

Struktura projektu je rozdělena do několika složek, které obsahují různé komponenty aplikace. Složka *Controllers* obsahuje kontrolery, které zpracovávají HTTP požadavky. Složka *Models* zahrnuje modely dat používané v aplikaci. *Views* obsahuje pohledy, tedy HTML šablony, které se používají k zobrazování dat uživateli. Složka *wwwroot* pak obsahuje soubory JavaScript a CSS, které jsou nezbytné pro frontend aplikace. Nakonec, složka *IISExpress* zahrnuje soubory potřebné pro spuštění aplikace na lokálním serveru.

6.4 Třídy

V projektu se nachází několik klíčových tříd, které se starají o specifické funkce a logiku generování křížovek. Abstraktní třída *Crossword* definuje základní metody pro generování křížovek. Z této třídy dědí třídy *CrosswordSwe* a *CrosswordBrit*, které generují švédské a britské křížovky.

Třída *Word* definuje jednotlivá slova a jejich odpovídající legendy. Třída *Dictionary* vytváří seznam instancí třídy *Word* a je zodpovědná za správu těchto slov. Pro porovnávání slov ze slovníku s vepsanými písmeny do křížovky slouží třída *WordComparator*.

Pro generování náhodných čísel podle dané váhy, což je využíváno při vytváření omezení pro mřížku švédské křížovky, je určena třída *WeightedRNG*.

Závěr

Byl vyvinut a implementován generátor křížovek, který má schopnost automaticky tvořit křížovky na základě uživatelem zvolené varianty (švédská a britská), jazyka (český a anglický) a úrovně obtížnosti (jednoduchá, normální a expert). Tento generátor je navržen tak, aby uspokojil široké spektrum uživatelů, ať už hledají jednoduché křížovky pro začátečníky nebo složitější výzvy pro zkušené luštitelé. Vygenerovanou křížovku je možné luštit přímo ve webové aplikaci, kde má uživatel k dispozici také možnost využít nápovědu, kontrolu správnosti řešení a v případě zájmu si křížovku vytisknout.

Pro generování křížovek byly použity pokročilé metody, jako je backtracking a constraint propagation. Implementace webové aplikace byla provedena v programovacím jazyce C# s využitím frameworku ASP.NET.

Vzhledem k dosaženým výsledkům, by pro další vývoj aplikace bylo vhodné zaměřit se na zlepšení a modernizaci uživatelského rozhraní. Dále by bylo přínosné rozšířit funkcionalitu aplikace o možnost třídění slov do různých kategorií. Tato funkce by uživatelům umožnila vybírat si slova pro křížovky podle jejich osobních preferencí a zájmů, což by zvýšilo uživatelský komfort a přizpůsobitelnost aplikace. Také by bylo možné umožnit generovat tvary mřížek zajímavější než jenom obdélníkové. Tento krok by zlepšil celkový zážitek z používání aplikace, čímž by se stala atraktivnější pro širší publikum.

Conclusions

A crossword puzzle generator has been developed and implemented with the ability to automatically create crosswords based on user-selected variants (Swedish and British), language (Czech and English), and difficulty levels (easy, normal, and expert). This generator is designed to satisfy a wide range of users, whether they are looking for simple puzzles for beginners or more challenging ones for experienced solvers. The generated crossword can be solved directly in the web application, where the user also has the option to use hints, check the correctness of their solution, and print the crossword if desired.

Advanced methods such as backtracking and constraint propagation were used for generating the crosswords. The web application was implemented using the C# programming language with the ASP.NET framework.

Given the achieved results, it would be beneficial for the further development of the application to focus on improving and modernizing the user interface. Additionally, expanding the application's functionality to include the ability to categorize words into different categories would be advantageous. This feature would allow users to select words for crosswords based on their personal preferences and interests, enhancing user comfort and the application's adaptability. This step would improve the overall user experience, making the application more attractive to a broader audience.

Literatura

- [1] Český svaz hádankářů a křížovkářů, z.s.. Klasifikace křížovek. <https://www.cshak.cz/klasifikace-krizovek>, [cit. 20.6.2024].
- [2] Miloš Zapletal, Velká encyklopedie her: Hry v klubovně. Olympia, 1986.
- [3] Smithsonian magazine, How the Crossword Became an American Pastime. <https://www.smithsonianmag.com/arts-culture/crossword-became-american-pastime-180973558/>, 2019-12 [cit. 20.6.2024].
- [4] Chromý Zdeněk, Zkuste si vylustit křížovku LN starou 85 let. https://www.lidovky.cz/relax/zajimavosti/zkuste-si-vylustit-krizovku-ln-starou-85-let.A100118_120036_ln-zajimavosti_mpr, 2010-1-18 [cit. 20.6.2024].
- [5] Britannica, <https://www.britannica.com/topic/The-Sunday-Times> [cit. 7.7.2024]
- [6] Telegraph, D-Day crosswords are still a few clues short of a solution. <https://www.telegraph.co.uk/news/uknews/1460892/D-Day-crosswords-are-still-a-few-clues-short-of-a-solution.html>, 2004-5-3 [cit. 20.6.2024].
- [7] Science Friday, Inside the Box: Crossword Puzzle Constructing in the Computer Age. <https://www.sciencefriday.com/articles/inside-the-box-crossword-puzzle-constructing-in-the-computer-age/>, 2014-9-12 [cit. 20.6.2024].
- [8] Český svaz hádankářů a křížovkářů, z.s. Křížovky. <https://www.cshak.cz/krizovky>, [cit. 20.6.2024].
- [9] Český svaz hádankářů a křížovkářů, z.s.. Směrnice pro tvorbu a řešení křížovek. <https://www.cshak.cz/smernice>, [cit. 20.6.2024].
- [10] Ivo Skalický, Program pro automatické generování křížovek. Diplomová práce. 2009. <https://theses.cz/id/15qnko/8909.pdf>, [cit. 20.6.2024].
- [11] C# Language Specification (6th ed.), Ecma International. [cit. 20.6.2024].
- [12] Montenegro Amanda, Montgomery Scott, Visual Studio 2019: Code faster. Work smarter. Create the future. Microsoft. 2019.
- [13] Hejlsberg Anders, The C# Programming Language. 2010.
- [14] <https://www.statista.com/statistics/793628/worldwide-developer-survey-most-used-languages/>, [cit. 20.6.2024].
- [15] Microsoft Learn. <https://learn.microsoft.com/en-us/dotnet/csharp/linq/get-started/introduction-to-linq-queries>, [cit. 20.6.2024].
- [16] Albahari Joseph, C# 7.0 in a Nutshell. O'Reilly Media, 2018, [cit. 20.6.2024].

- [17] Bell Jordan, A survey of known results and research areas for n-queens. *Discrete Mathematics*, 309(1), 1–31, 209
- [18] Korf Richard, Improved limited discrepancy search. 1996.
- [19] Ginsberg Matthew, Search Lessons Learned from Crossword Puzzles, <https://cdn.aaai.org/AAAI/1990/AAAI90-032.pdf>, [cit. 20.6.2024].