

クロスワードパズル生成問題の 新しい定式化

佐藤潤一

社会情報学科, 2013160

クロスワードパズルとは

1 う	2 し	
3 さ	か	4 な
ぎ		し

ヒント

縦

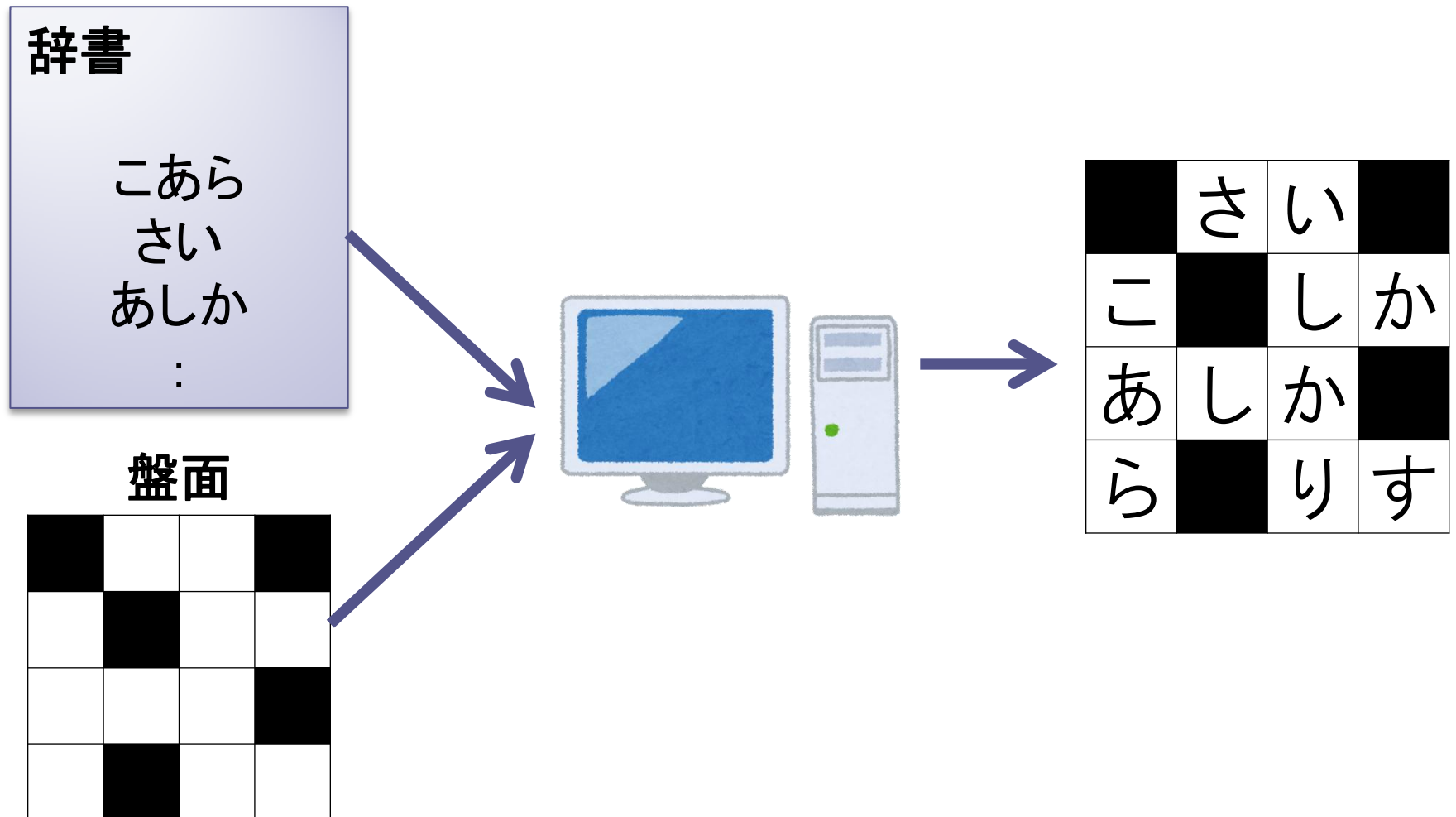
- 1 月の模様はこの動物が餅つきをしているように見える
- 2 奈良公園にたくさんいる
〇〇せんべいも有名
- 4 リンゴに似た果物
皮は黄緑色

横

- 1 胃袋を4つもっている動物
- 3 海にたくさんいる
煮るなり焼くなり

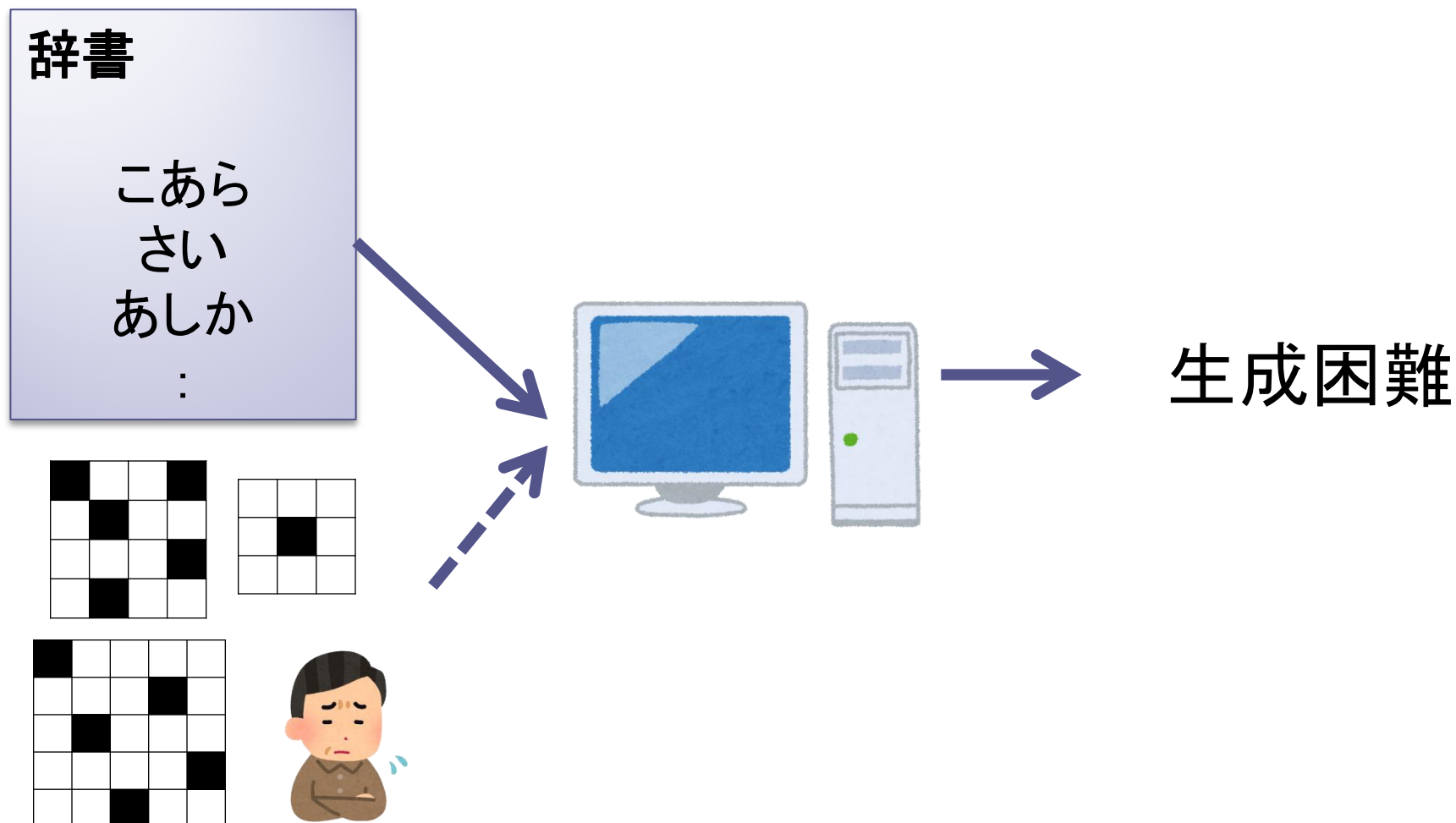
従来の自動生成

辞書と、彩色された盤面を入力として与える



従来の自動生成の問題点

盤面を決めることが難しい

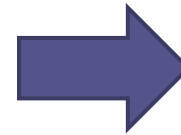
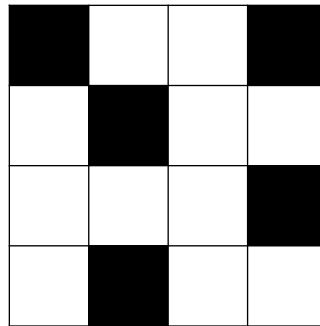


提案 : 新しい定式化

従来の方法

辞書

こあら
さい
:



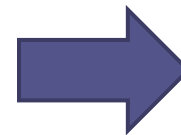
	さ	い	
こ		し	か
あ	し	か	
ら		り	す

提案手法

辞書

こあら
さい
:

盤面の1辺の長さ
 $n = 4$



	さ	い	
こ		し	か
あ	し	か	
ら		り	す

制約最適化問題

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & g_j(x) \in C_j \quad (j=1,\dots,m) \\ & x_i \in D_i \quad (i=1,\dots,n) \end{array}$$

※ n は変数の個数, m は制約の個数

- ORや人工知能における重要な基本問題
- 一般にはNP困難
- 様々な厳密/近似アルゴリズムが知られる
- 厳密アルゴリズムに基づいたソルバで解く場合, 計算時間の上限を定め, 時間内の最良解で評価するのが一般的

実装の準備

- 辞書内の全ての単語の前後に黒マスを追加

さい

しか

こあら

:

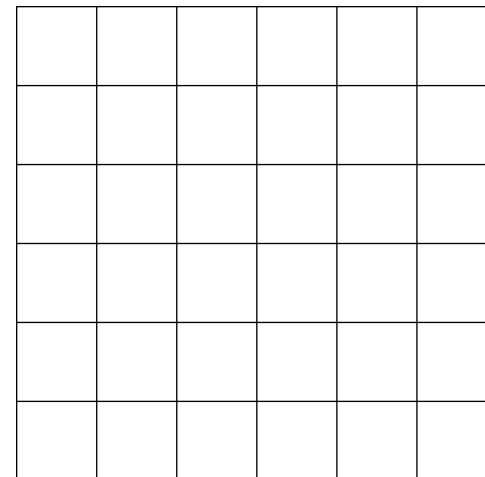
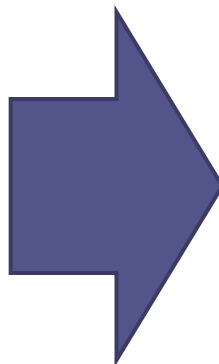
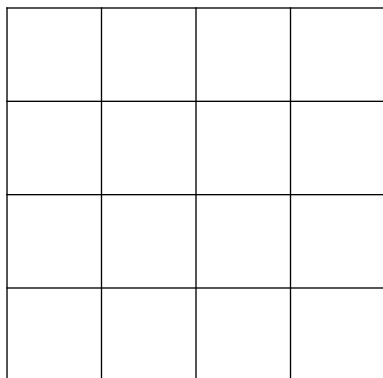
■ さい ■

■ しか ■

■ こあら ■

:

- 盤面のサイズを $(n+2) \times (n+2)$ に拡張



実装の方針

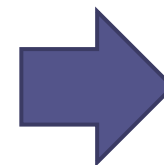
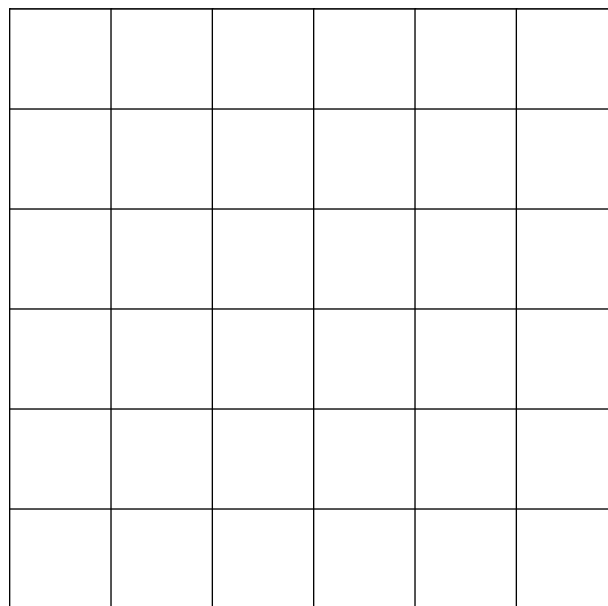
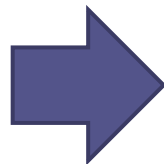
単語

■ さ い ■

■ し か ■

■ こ あ ら ■

:



■	さ	い	■
こ	■	し	か
あ	し	か	■
ら	■	り	す

maximize
subject to

(白マス数) - $w \times$ (黒マスの隣接)
隣接制約, 外周制約, 回数制約,
語交差制約, 文字一致制約,
割当制約

実験

- ・定式化した制約最適化問題を
IBM ILOG CPLEX CP Optimizer を用いて解く
- ・使用プログラム言語はC++、総行数は約800行

- ・実験内容

辞書

こあら
しか
:



盤面の1辺の長さ n



実験の結果



	あ		た		は
ひ	ら	お	か		ら
	い		は		さ
さ		い	し	か	わ
か		と		さ	
い	と	う	だ	い	き

オリンピックメダリストパズル
辞書サイズ：60
計算時間：120秒



	ほ		ひ	だ	か
さ	つ	ぽ	ろ		
	か		お	た	る
た	い	き		き	
	ど		む	か	わ
ら	う	す		わ	

北海道地名パズル
辞書サイズ：120
計算時間：120秒

実験の結果

	¹ ら		² た	ん	³ き
⁴ こ	う	か	い		よ
	ん		⁵ い	と	う
	⁶ じ	ご	く		し
⁷ ぜ			か		よ
⁸ み	す	こ	ん		く

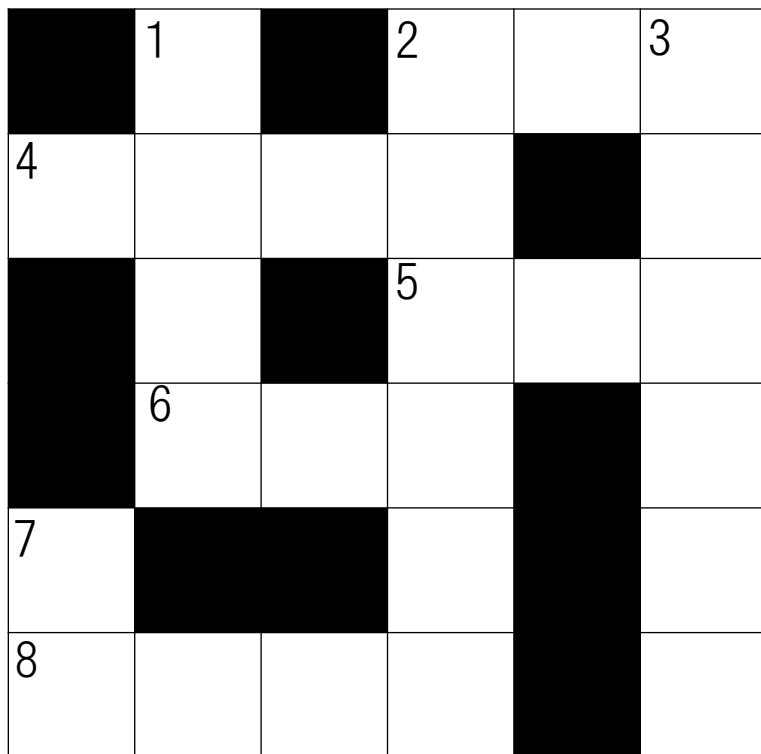
小樽商大パズル
辞書サイズ：120
計算時間：120秒



伊藤 整

画像出典：作家・伊藤整が語る小樽商大 I -
小樽商科大学同窓会 仙台緑丘会
https://sites.google.com/site/ryokyu55miyagi/umi_otaru001/ito_sei1

パズルの利用



小樽商大パズル

ヒント

縦

- 1 3号館2階、国際交流〇〇〇〇
- 2 入学式や卒業式を行う場所
- 3 先生になりたかったら
〇〇〇〇〇〇科目を取ろう
- 7 3年時から所属する

横

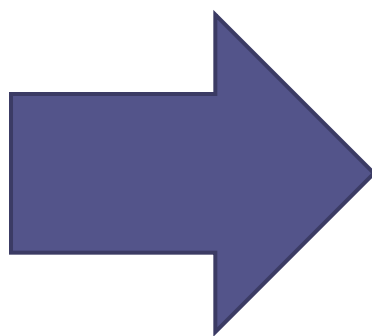
- 2 以前はあった〇〇〇大学部
- 4 一般参加可能な〇〇〇〇講座
- 5 『若い詩人の肖像』著者
〇〇〇整
- 6 〇〇〇坂を登れば商大
- 8 緑丘祭で行われるコンテスト
男性版と女性版がある

まとめ

今後の課題

- ・他の最適化手法を試みて、より良いパズルを生成する

	ほ		る	も	い
や	く	も		り	
	と		う		あ
そ		し	り	べ	し
う	す		ゆ		べ
や		と	う	べ	つ



	さ		え	さ	し
ほ	ろ	の	べ		べ
	ま		つ	べ	つ
て		ち		つ	
し	ん	と	つ	か	わ
お		せ		い	

- ・自然言語処理や情報検索の技術で、ヒント文や辞書を自動生成できるようにする

本研究の内容の一部は

平成28年度電気・情報関係学会北海道支部連合大会で
発表したものである

また、観光情報学会 第14回研究発表会で発表予定である

※ここから予備スライド

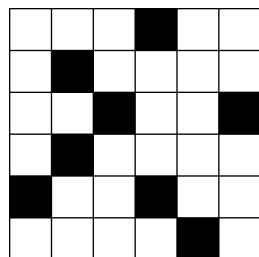
英単語パズル

A	S	P	S		O		P		R
V			I		A		I		A
I		A	F	A	R		L		P
S			T		S	H	A	R	P
	R	H	E	A			T		E
P			R		A	H	E	A	D
I		K		F				B	
N		E		A	O	R	T	A	E
E		R		M		D		S	
D	R	I	V	E	S		L	E	W

英単語パズル

クロスワードパズルを作る

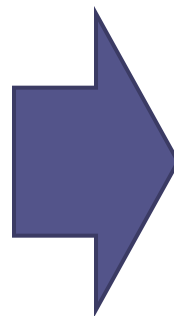
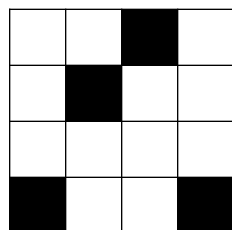
人力で作る



単語群a

?

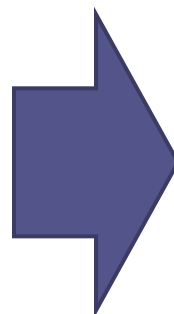
単語群b



	さ	い	
こ		し	か
あ	し	か	
ら		り	す

コンピュータを使用して自動で作る

テーマは
○○
大きさは△



	さ	い	
こ		し	か
あ	し	か	
ら		り	す

変数 : 単語の割り当て

$x_{k,ij}$ は、単語 k を (i, j) を先頭にして横方向に配置する(1)か否(0)かを示す0-1変数

辞書

1 ■ ほつかいどう ■

2 ■ さつぽろ ■

:

$x_{2,4,1} = 1$ のとき
(4,1)を先頭にして
横方向に
単語番号2の単語を
割り当てる

	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

縦方向への割り当てを表す変数 y も同様に定義される

変数 : 文字の割り当て

$z_{a,ij}$ は、文字 a を (i, j) に配置する(1)か否(0)かを示す
0-1変数

文字

0



1

あ

2

い

:

$z_{1,2,3} = 1$ のとき
(2,3)に文字「あ」を
割り当てる

	1	2	3	4	5	6	7	8
1								
2			あ					
3								
4								
5								
6								
7								
8								

目的関数

$\text{maximize}(\text{目的関数}) = (\text{白マス数}) - w \times (\text{隣り合う黒マスの対})$

※wはペナルティへの重み

日本語クロスワードパズルの慣例

- ・白マスはすべて連結している
- ・黒マスは連続しない

:

い	か	だ
わ		
し	ら	す

白マス : 7

隣り合う黒マスの対 : 1

目的関数値 : 6

い	か	
ね	つ	こ
	お	し

白マス : 7

隣り合う黒マスの対 : 0

目的関数値 : 7

制約

- 回数制約

同じ単語は多くとも1つのスロットにしか割り当てられない

か	ら	す
ら	■	み
す	い	か

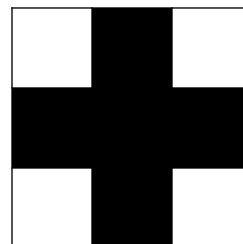
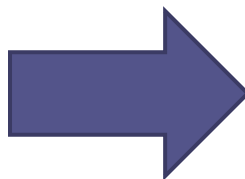
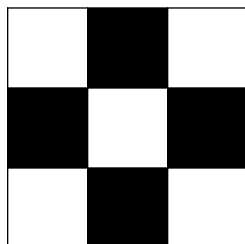


か	ら	す
え	■	き
る	び	ー



- 隣接制約

(i,j)に隣接するマスがすべて黒マスのとき、(i,j)は黒マスになる



- 割当制約

すべてのマス目に必ず文字(■を含む)を割り当てる

制約

- 文字一致制約

単語kを割り当てたとき、単語kの文字が該当するマスに入る

辞書

2 ■あしか■

:

文字

0 ■

1 あ

2 い

:

	1	2	3	4	5
1					
2					
3					
4					
5					

$$x_{2,3,1} = 1$$

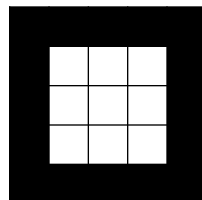


	1	2	3	4	5
1					
2					
3	■	あ	し	か	■
4					
5					

$$z_{1,3,2} = 1$$

- 外周制約

端のマスは全て黒マス

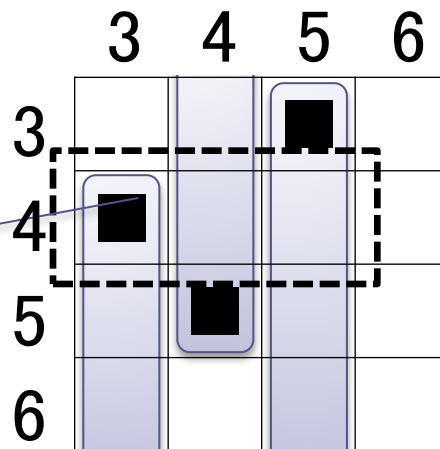


制約

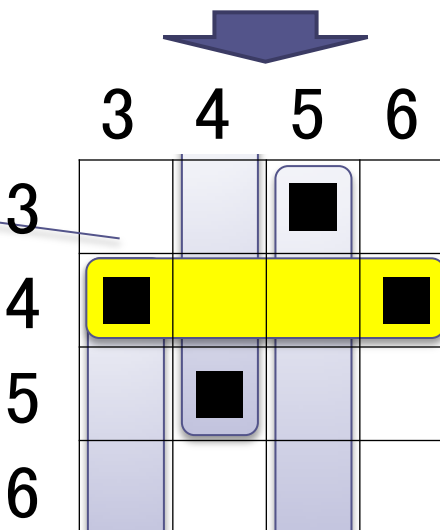
・語交差制約

黒マス-文字-文字 と連続した場合は該当箇所に単語を割り当てる

横方向に
黒一字一字
と連続



該当箇所に
単語を割り当てて
辞書と矛盾しない
ようにする

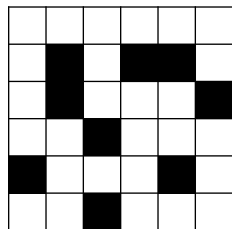


変数の個数は
 $O((m + q)n^2)$
制約の本数は
 $O(mn^3)$

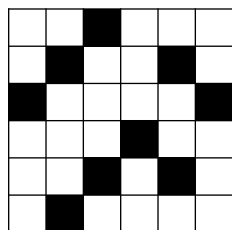
m : 単語数
q : 文字数

日本語クロスワードパズルの慣例

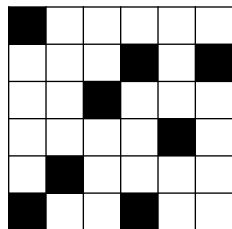
- ・盤面上で黒マスが連続してはならない



- ・盤面上で全ての白マスが連結していなければならない



- ・盤面の4隅が白マスでなければならない



実験

実験環境

生成実験の環境

CPU	Intel® Core™ i7-3520M CPU
クロック周波数	2.90GHz
メインメモリ	8.00GB
ソフトウェア	IBM ILOG CPLEX CP Optimizer (ver. 12.6.2)

生成実験

表1. 辞書サイズ=120,重み=1 のときの生成結果

盤面の次数n	6	7	8	9
計算上限時間(秒)	120	600	1800	3600
白マス数	24	34	42	54
黒マス数	12	15	22	27
隣接する黒マスの対の数	1	4	0	5
連結成分数	1	2	6	5

	ほ		ひ	だ	か
さ	つ	ぽ	ろ		
	か		お	た	る
た	い	き		き	
	ど		む	か	わ
ら	う	す		わ	

生成実験

表2. 辞書サイズ=60,重み=2 のときの生成結果

盤面の次数 n =	6	7	8	9
計算上限時間(秒)	120	600	1800	3600
白マス数	25	31	42	51
黒マス数	11	18	22	30
隣接する黒マスの対の数	3	5	7	13
連結成分数	1	3	6	4

	あ		た		は
ひ	ら	お	か		ら
	い		は		さ
さ		い	し	か	わ
か		と		さ	
い	と	う	だ	い	き