

オンライン手書き文字認識HMMにおける 座標情報と方向情報の利用法と効果

九州大学大学院システム情報科学府
知能システム学専攻 修士2年

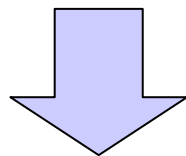
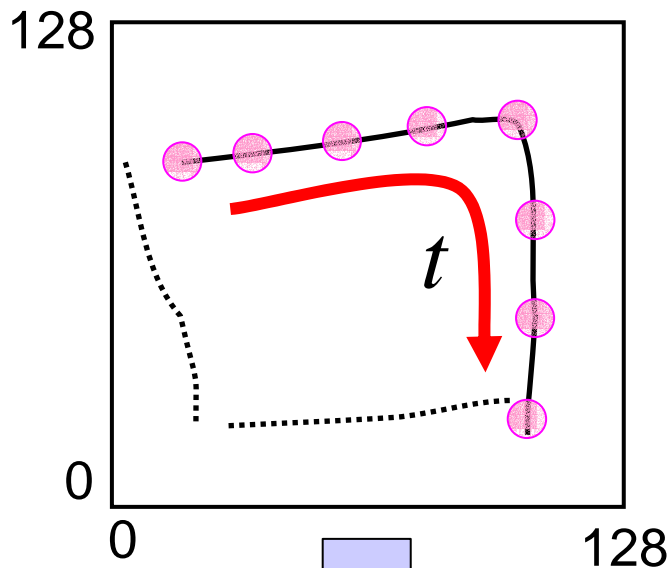
迫江・内田研究室

奥村 大樹

The slide features several decorative circles. Three light purple circles are positioned behind the text, and three more are located below the text: two solid light purple circles on the left and one hollow light purple circle on the right.

HMMによる従来の オンライン文字認識

オンライン文字



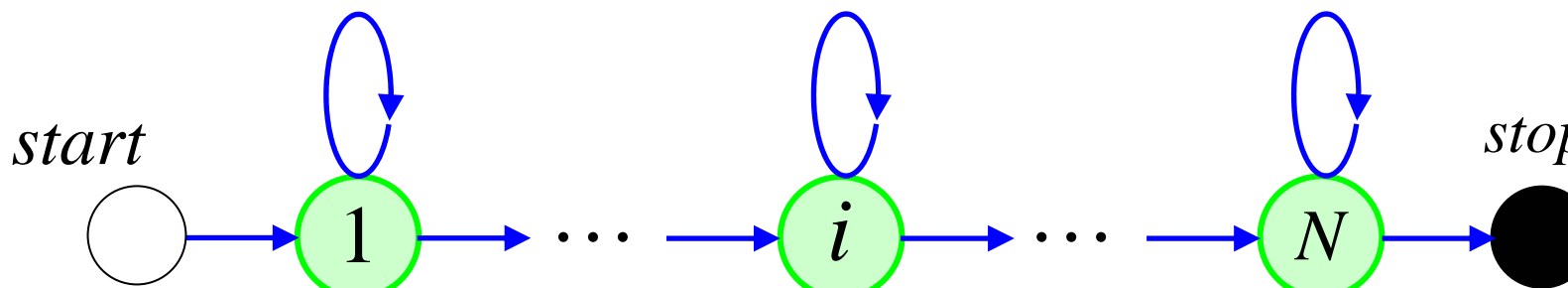
γ_i : 方向情報

x_i : 横座標情報

y_i : 縦座標情報

	→	→	→	→	↓	↓	↓	
x_i : 横座標情報	10	40	60	80	100	105	105	100
y_i : 縦座標情報	90	95	100	105	110	80	50	20
	$t \rightarrow 3$							

従来のHMM (線分方向情報出力型HMM)



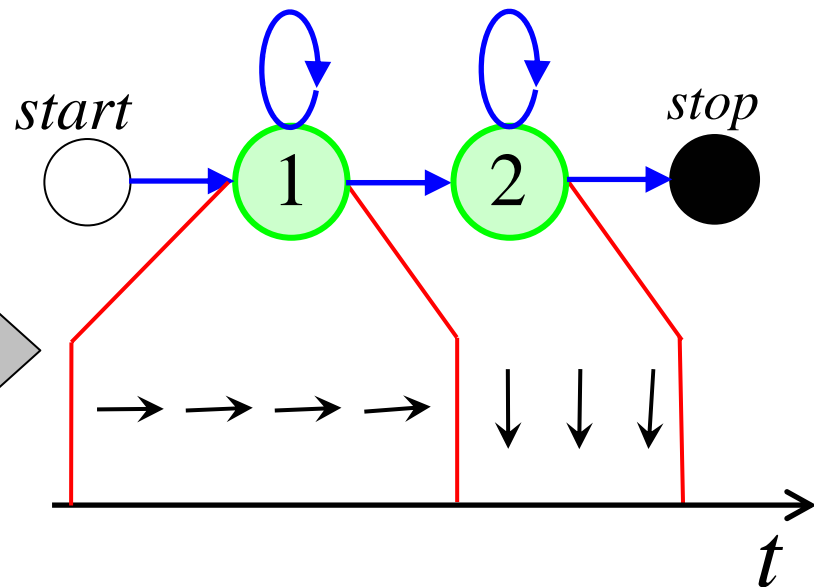
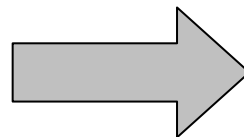
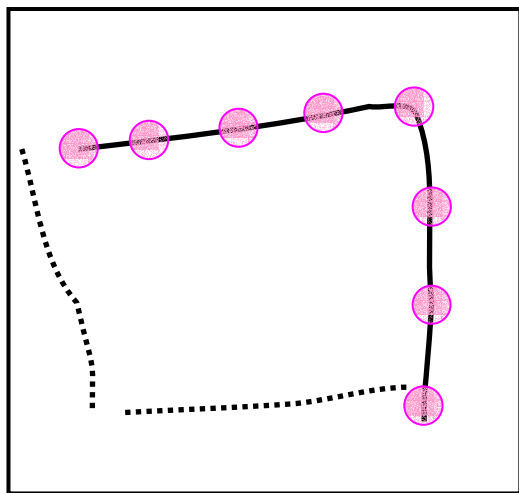
d_i : 方向情報

x_i : 横座標情報

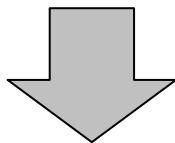
y_i : 縦座標情報

d_1	d_2	...	d_t	d_{t+1}	d_{t+2}	...	d_T
x_1	x_2	...	x_t	x_{t+1}	x_{t+2}	...	x_T
y_1	y_2		y_t	y_{t+1}	y_{t+2}		y_T

方向情報が使われる理由

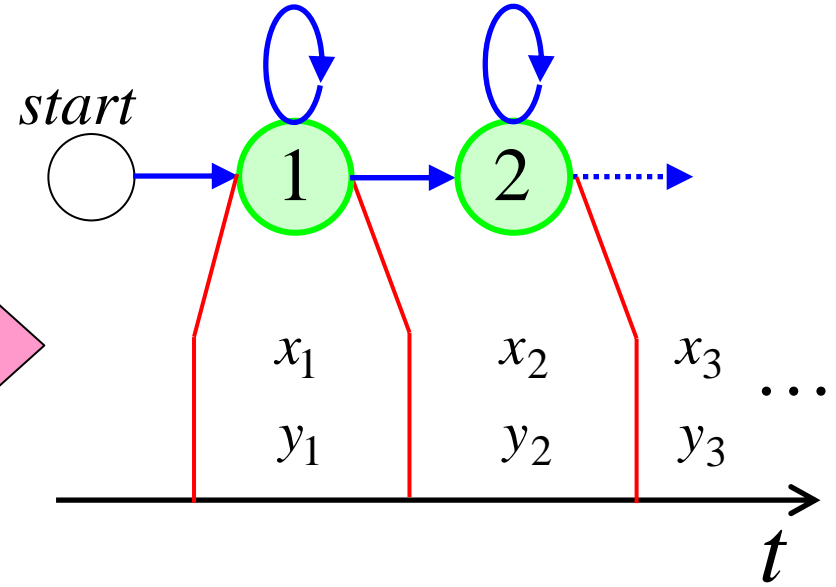
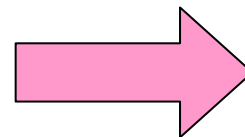
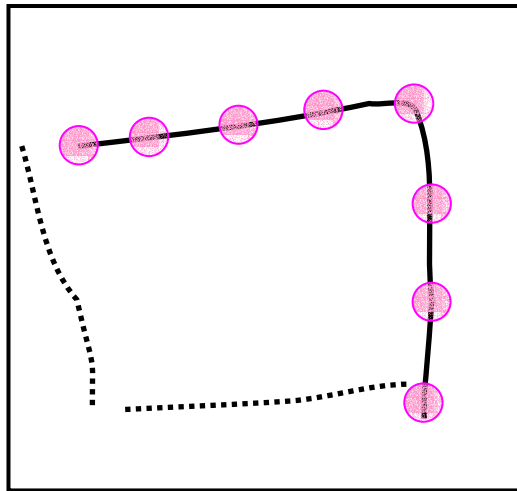


方向情報は線分内でほぼ一定

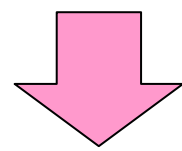


1線分を1状態で表現できる

座標情報が使われない理由



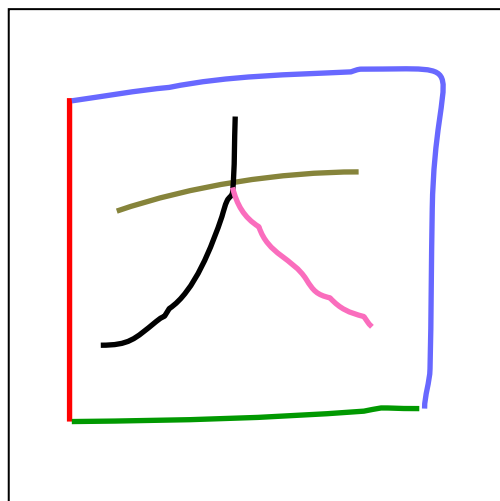
座標情報は常に変化



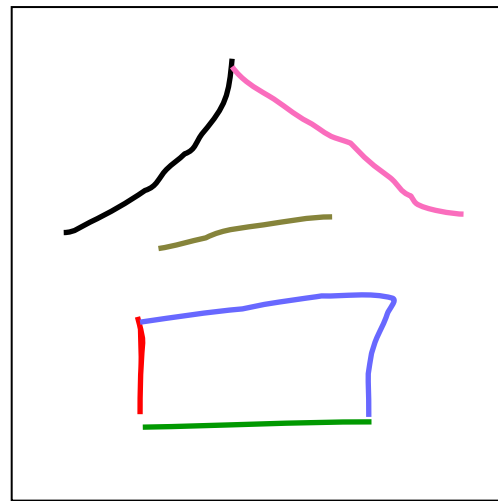
1線分を1状態で表現しにくい

座標情報の必要性

- 筆順自由化時に、方向情報のみでは特徴量として不十分
 - 以下の漢字は方向情報が同じ



「因」

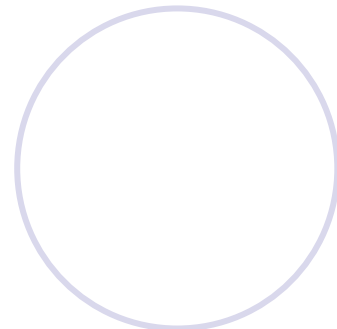
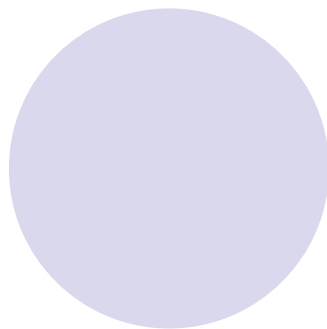
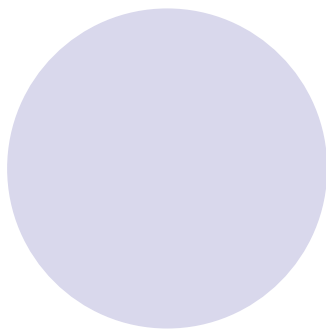


「合」

目的

- 座標情報(非定常的)と方向情報(定常的)をそれぞれの特性に合わせて使い分けるHMMを提案
 - 「線分始点座標 + 方向情報出力型HMM」
- 筆順自由認識手法への埋め込み
 - 「キューブサーチHMM」

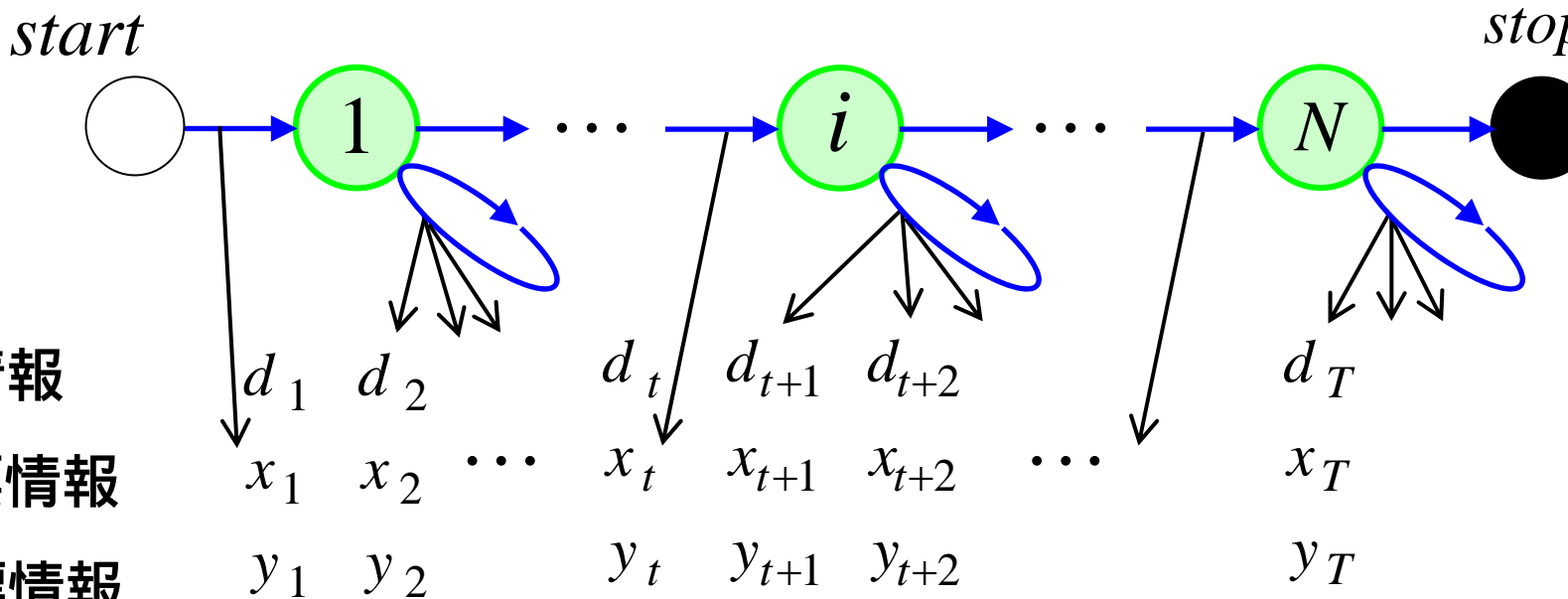
線分始点座標+方向情報出力型
HMM



基本的な考え方

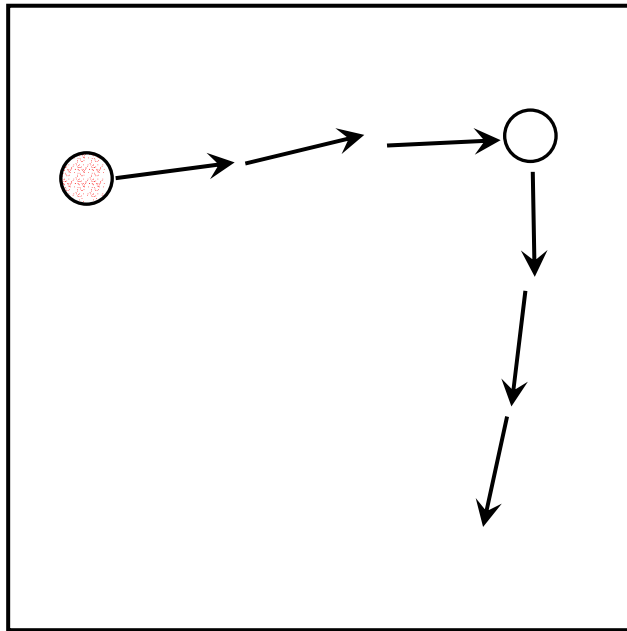
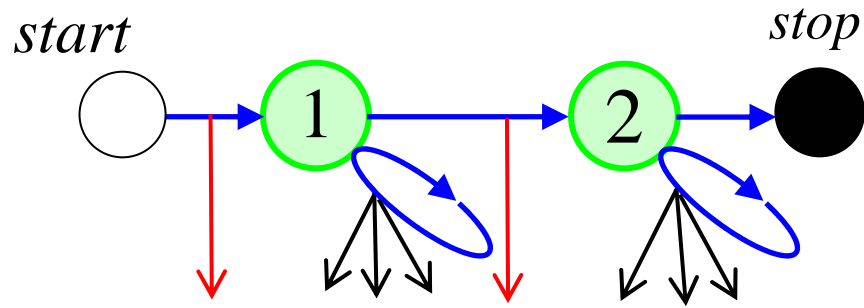
- **線分の始点座標を次状態への遷移時に出力**
 - 1線分につき1回だけ出力
- **方向情報を自己遷移時に出力**
 - 線分の長さだけ繰り返し出力
- **基本的に一つのストローク(画)を表現**

線分始点座標+方向情報出力型HMM

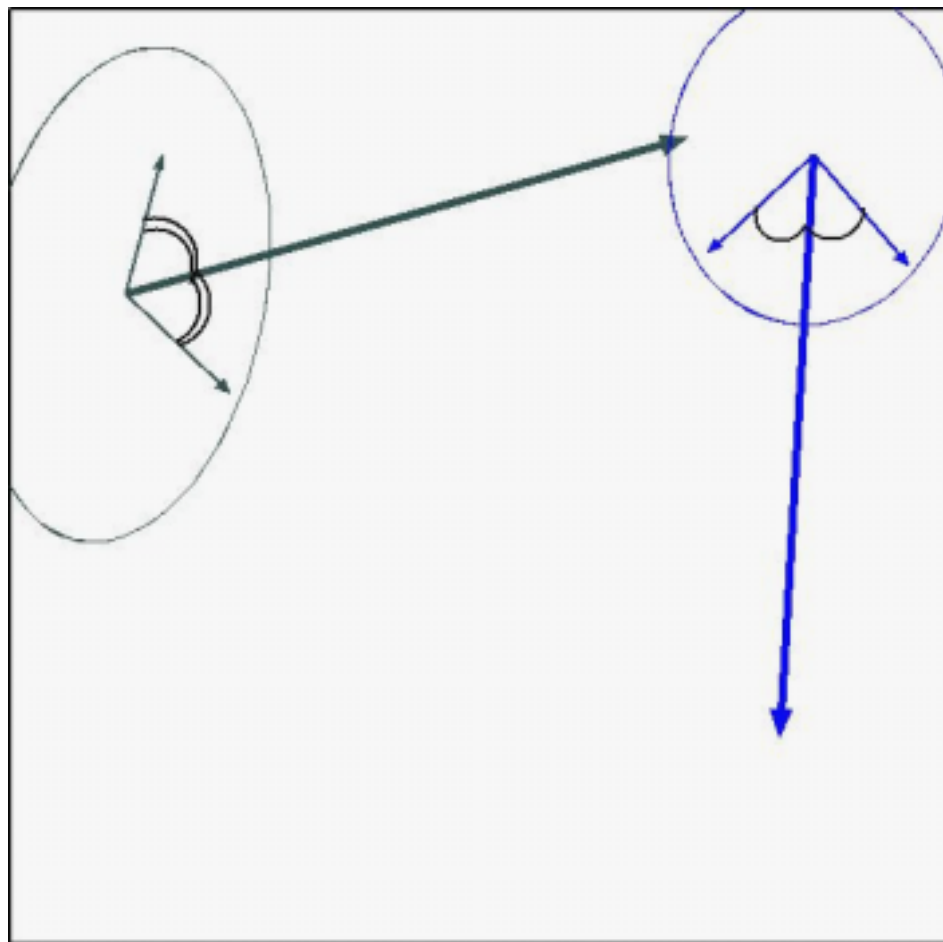


提案手法のイメージ

「口」の第2画



本手法による「口」の第2画の学習結果

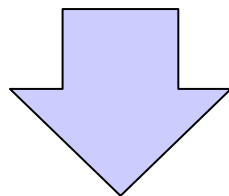




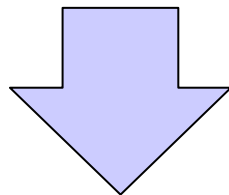
1画化した漢字を用いた 認識実験

実験の目的

- 本手法「線分始点座標 + 方向情報出力型HMM」の基本的性能の評価
- 従来法「線分方向情報出力型HMM」との比較



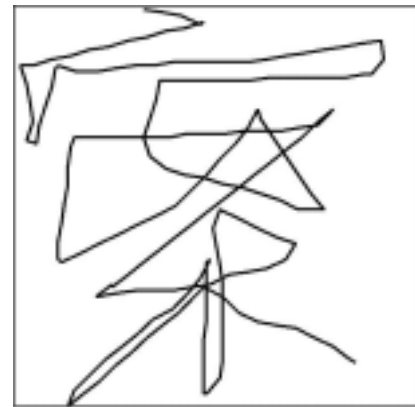
いずれも1画を表現するモデル



前処理により1画化した文字を対象

実験データ

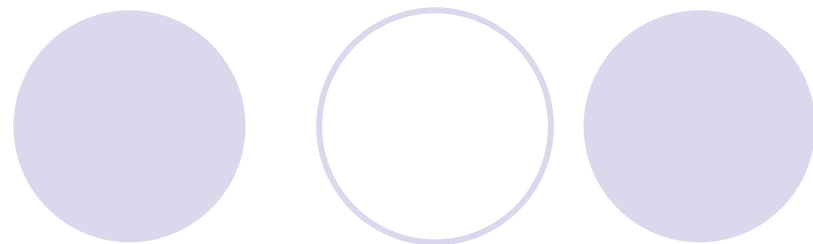
- 筆順の正しい10画漢字70文字種
- 前処理により1画化した文字を対象



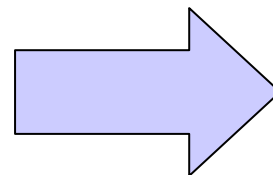
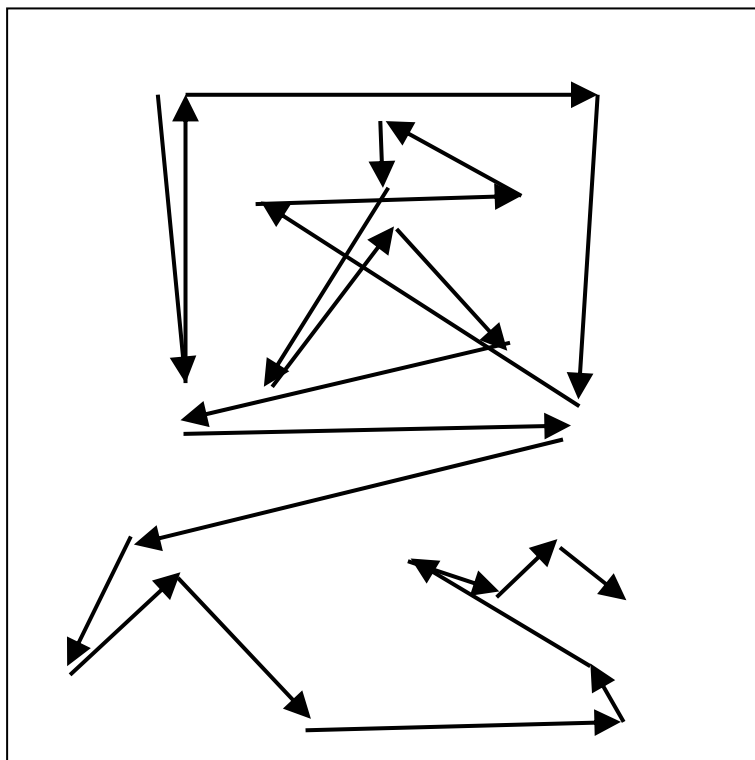
「案」の1画化文字

- データセットA
 - 当研究室の内部データ/筆記者30人分
 - 学習・認識に利用
- データセットB
 - HANDS-kuchibue_d-97-06-10/筆記者10人分
 - 認識のみに利用

状態数の決め方



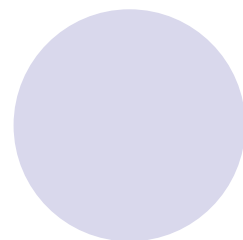
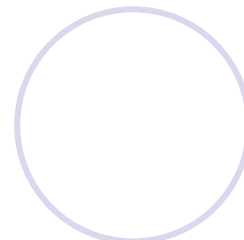
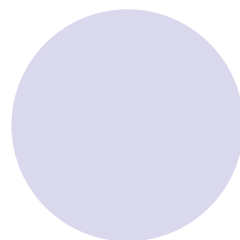
- 文字ごとに手動で決定



状態数
23

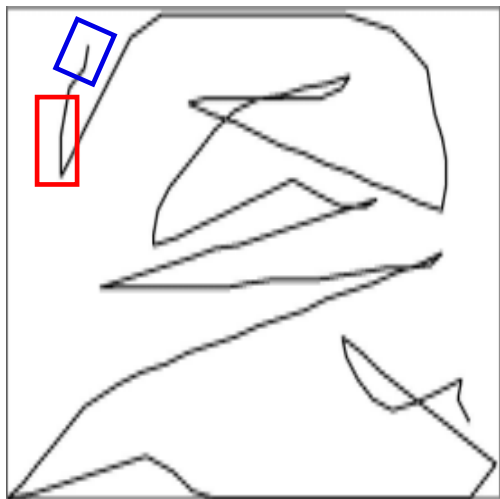
「恩」を1画化したモデル

認識率の比較

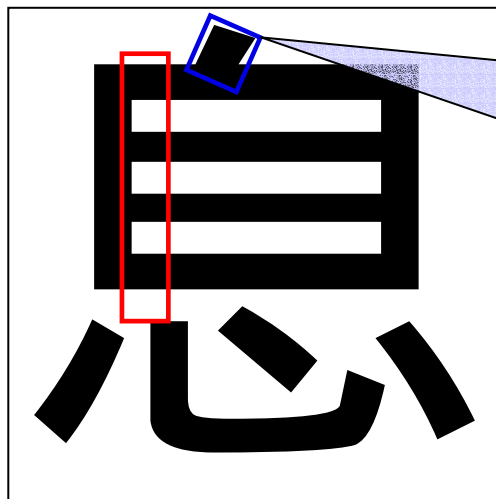


データセット	A	B
従来法	99.9%	99.3%
本手法	100.0%	99.9%

座標情報を利用による改善例



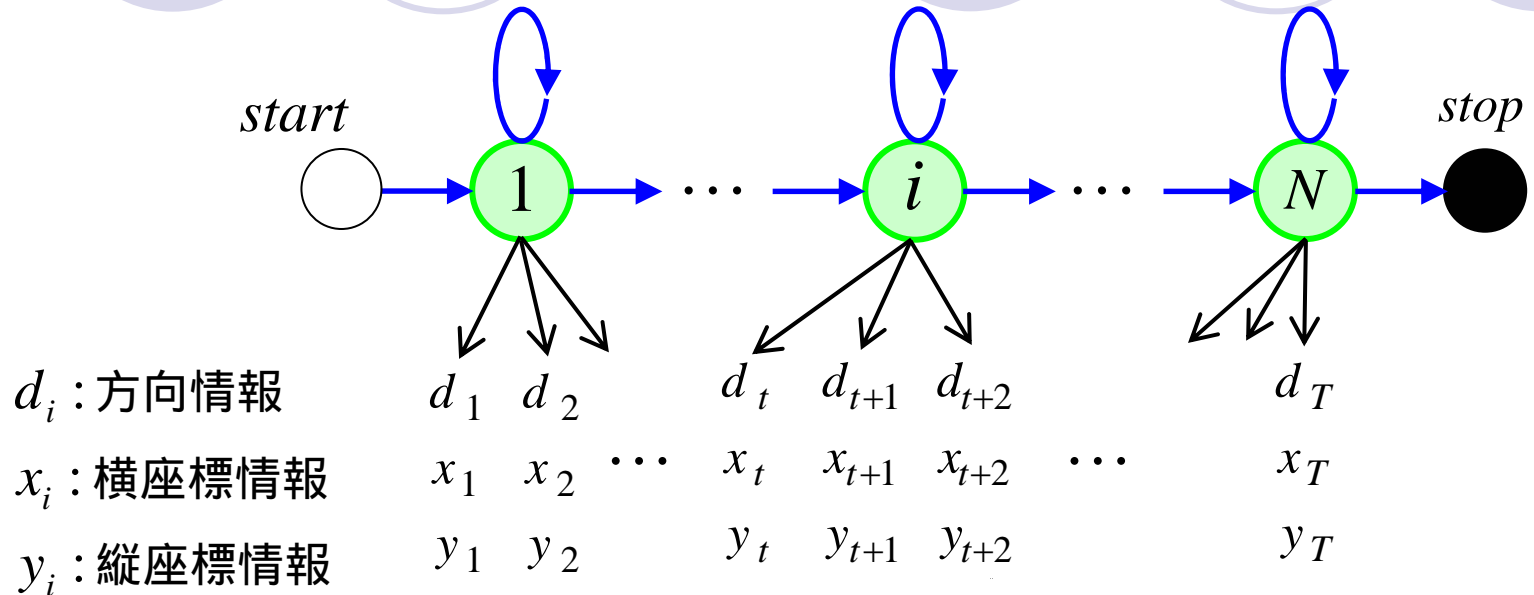
従来法で
誤認識した例
「恩」 「息」



従来法による
マッチング部分

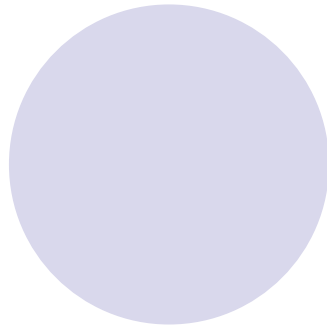
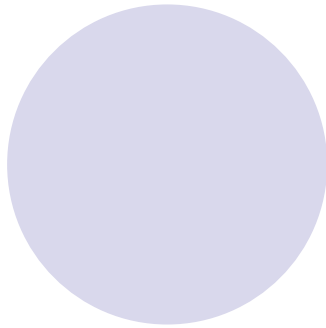
方向情報は
同じだが
位置は異なる

全情報出力型HMMとの認識率の比較



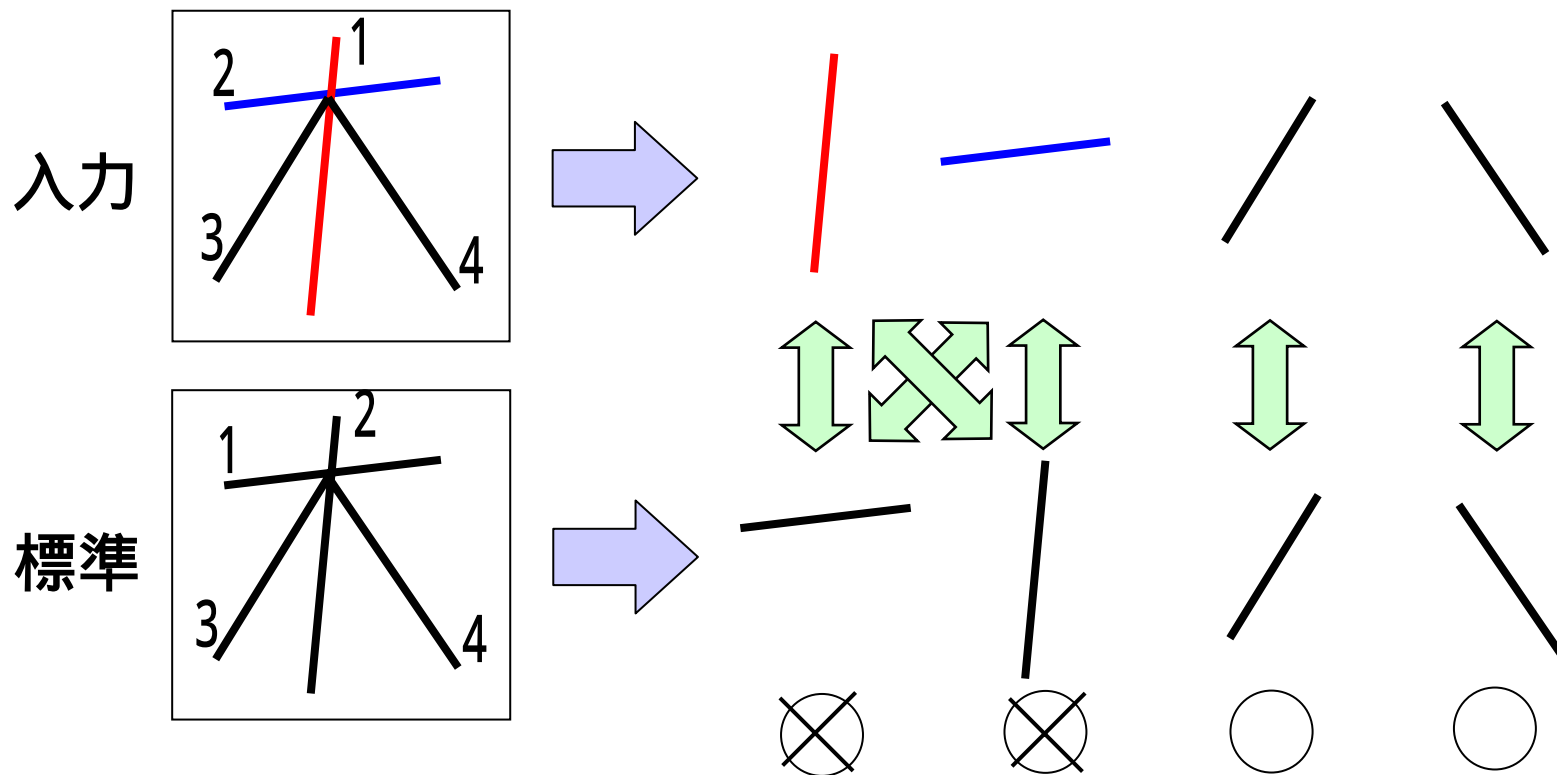
データセット	A	B
全情報出力型	89.5%	81.2%
本手法	100.0%	99.9%

筆順自由認識手法への埋め込み (キューブサーチHMM)



筆順変動

- 入力された順番の通りに比較はできない

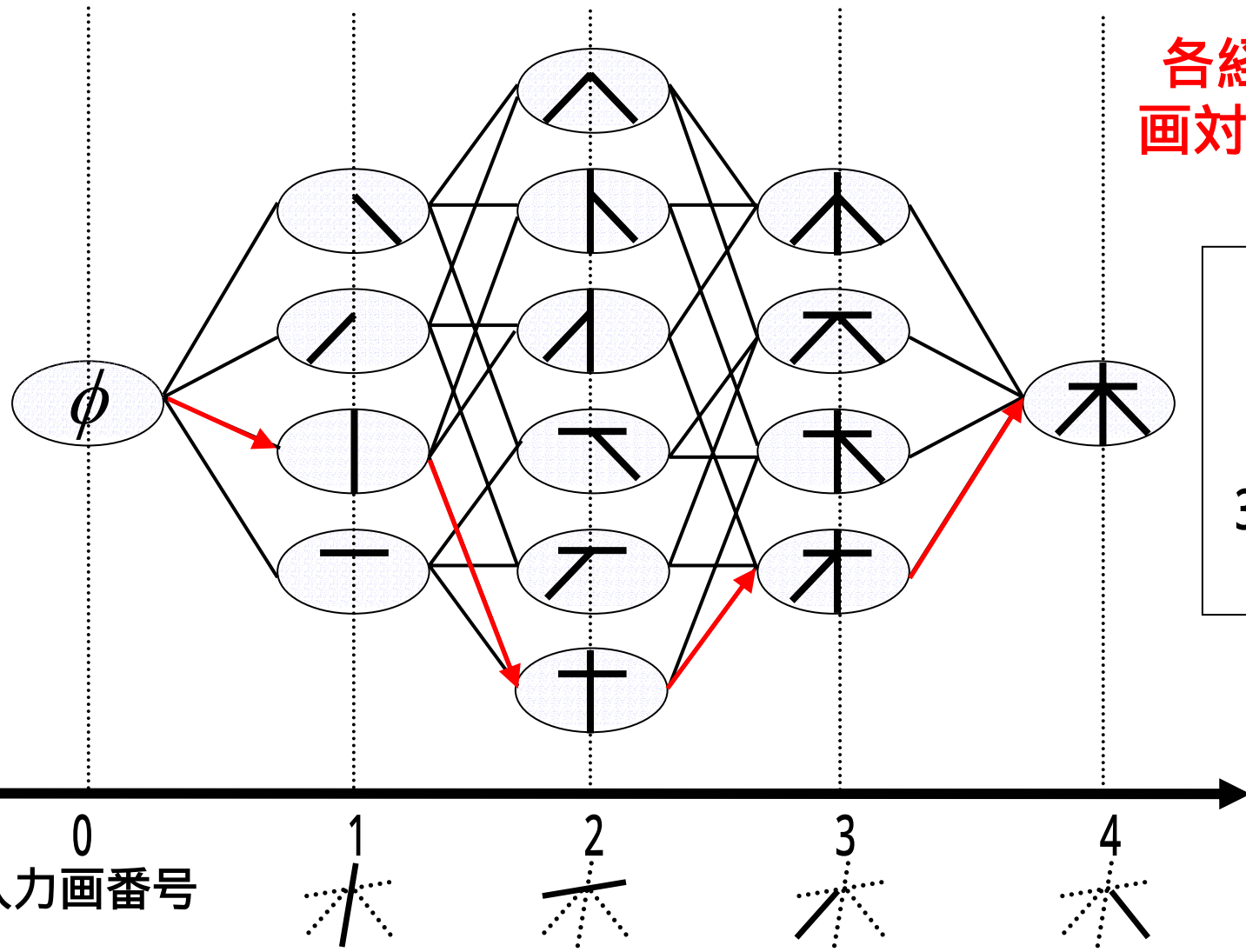


画の対応付けにより、筆順自由認識が可能に

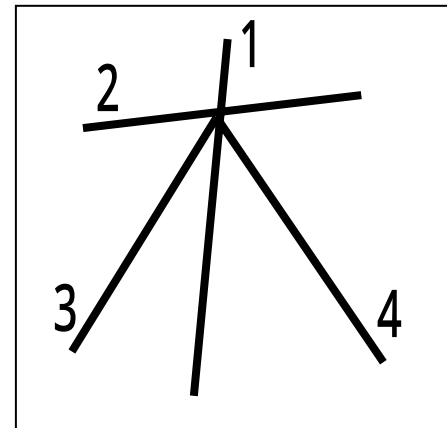
キューブサーチ法

- 最適画対応問題の解法の一つ
- 全ての画対応を表現するグラフを利用
 - キューブサーチグラフ
- そのグラフ上での最適経路として、最適画対応を出力

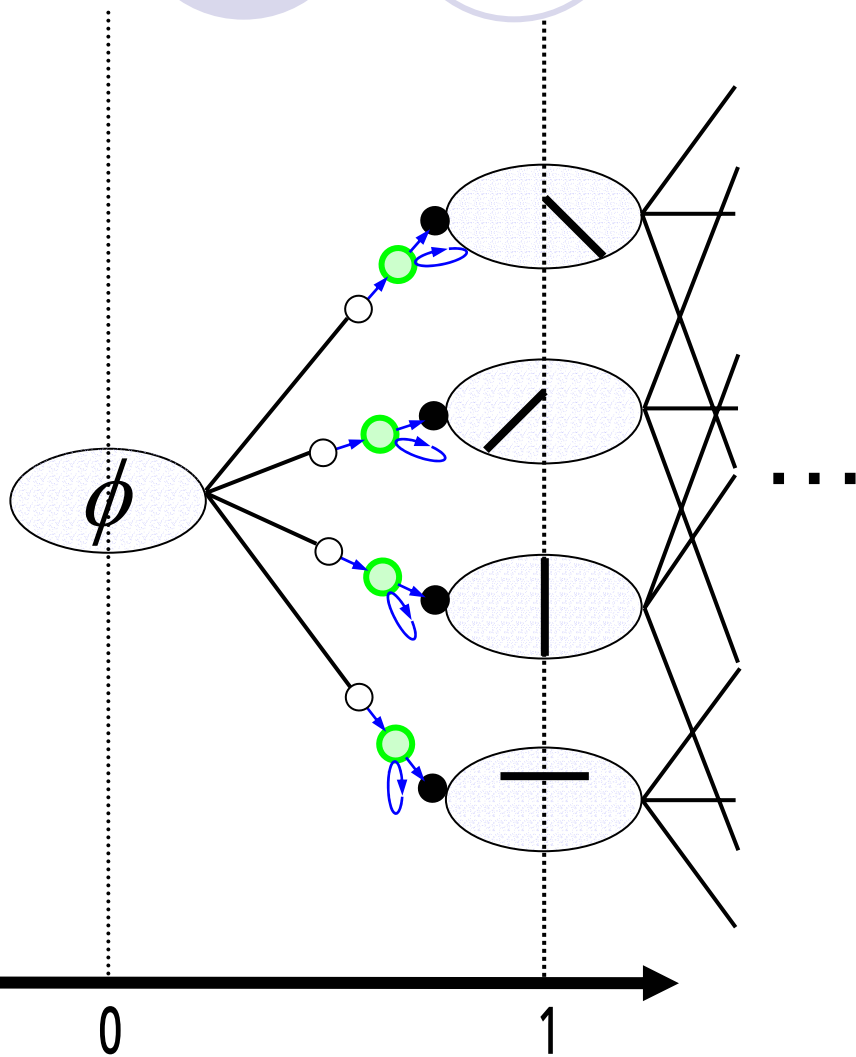
キューブサーチグラフ(「木」の場合)



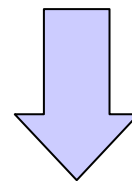
各経路が一つの
画対応付けを表現



キューブサーチへのHMMの埋め込み(1)



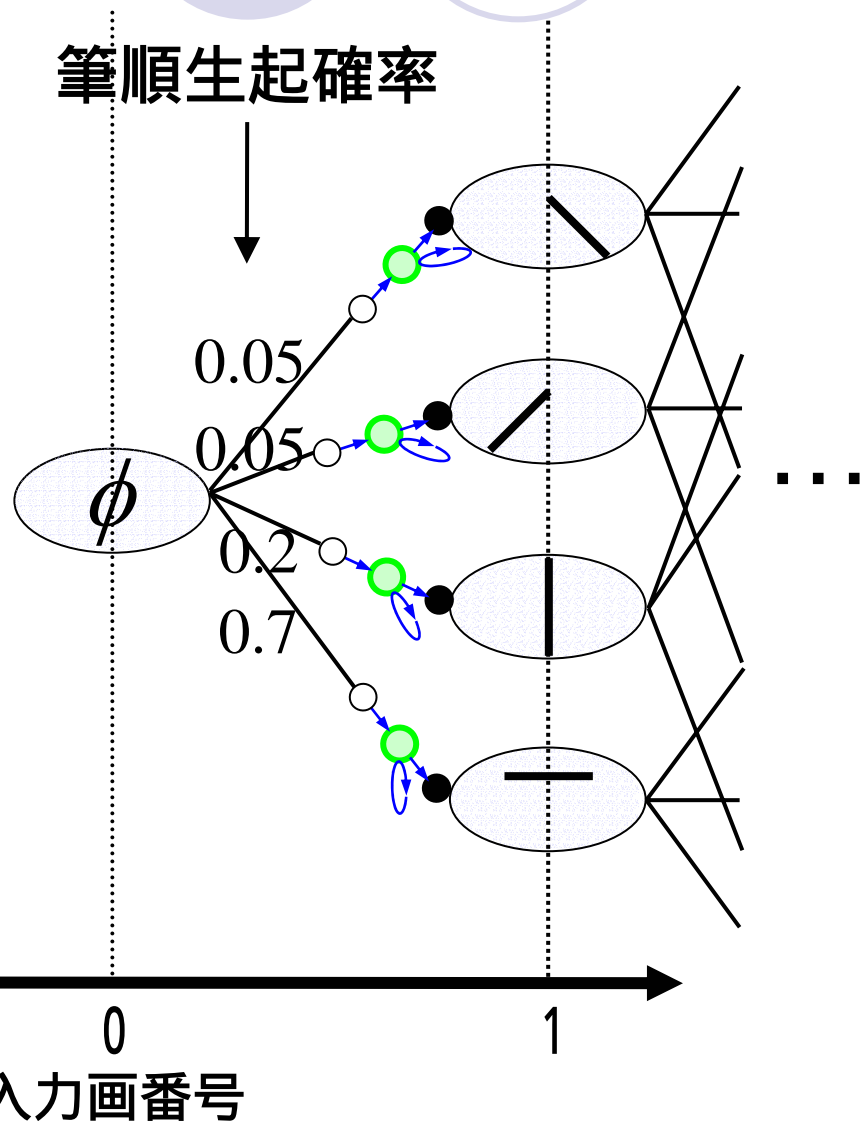
各エッジの評価値
= HMMによる
各ストロークの尤度



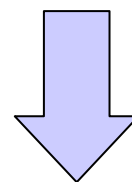
評価値を最大化する
経路を求める

キューブサーチへのHMMの埋め込み(2)

筆順生起確率



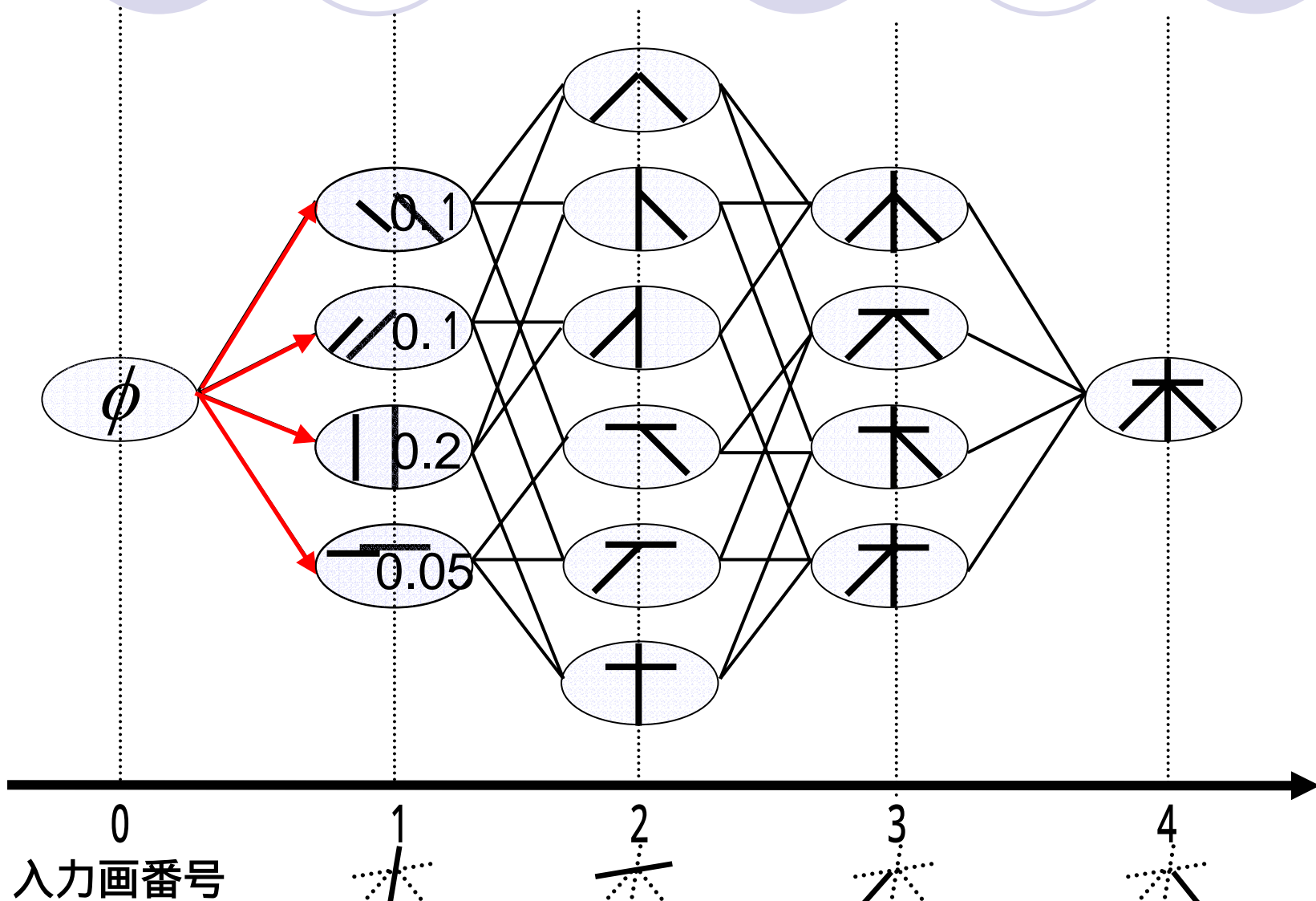
各エッジの評価値
= 筆順生起確率
× HMMによる
各ストロークの尤度



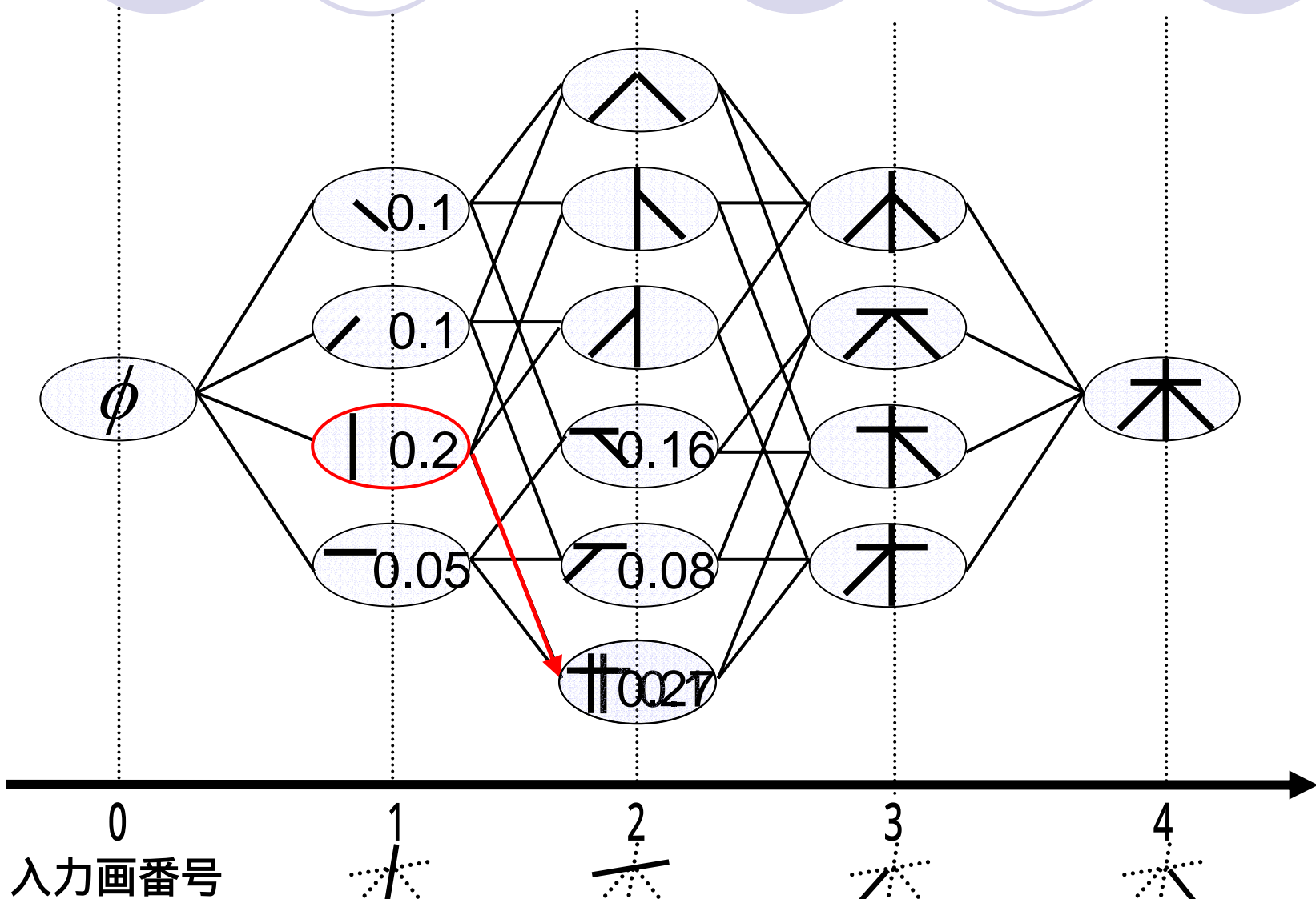
評価値を最大化する
経路を求める

確率的な枠組みで、
ストロークの尤度ならびに
筆順変動を統一的に表現

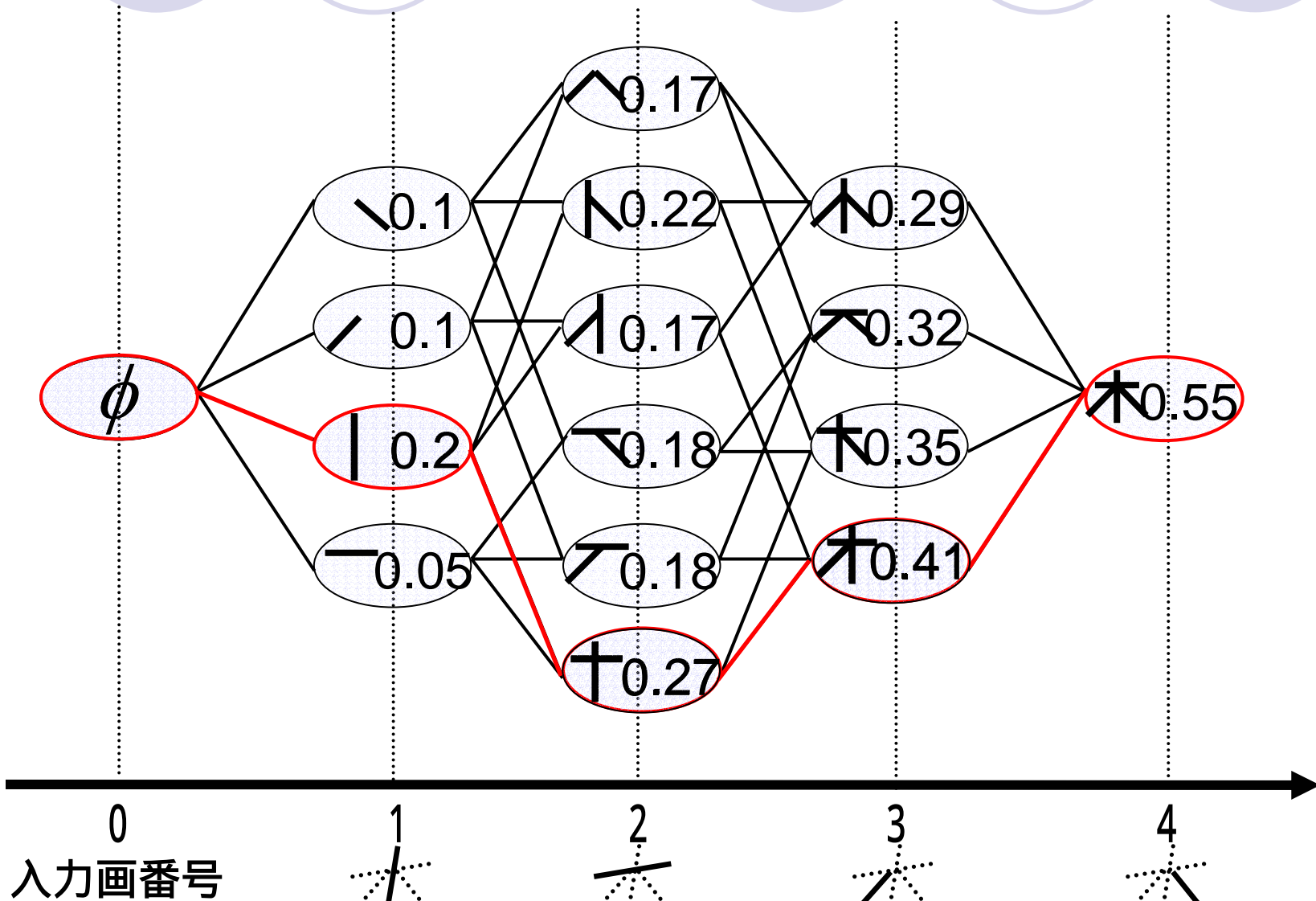
キューブサーチHMMの動作(「木」の場合)



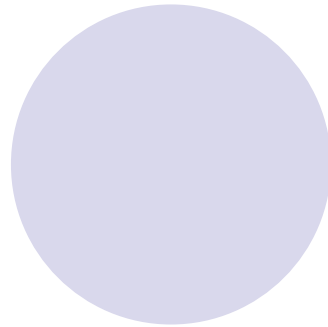
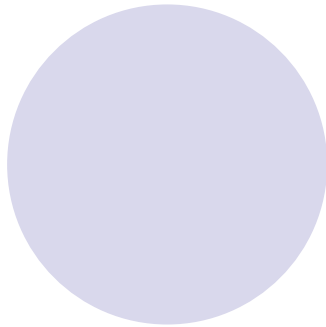
キューブサーチHMMの動作(「木」の場合)



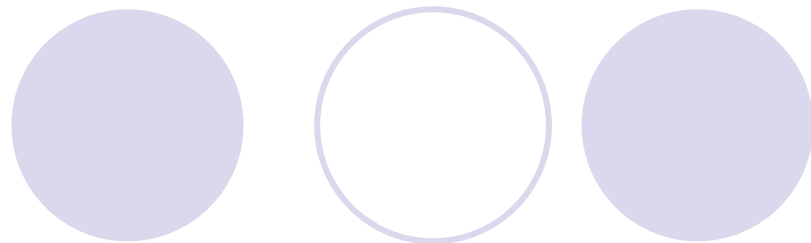
キューブサーチHMMの動作(「木」の場合)



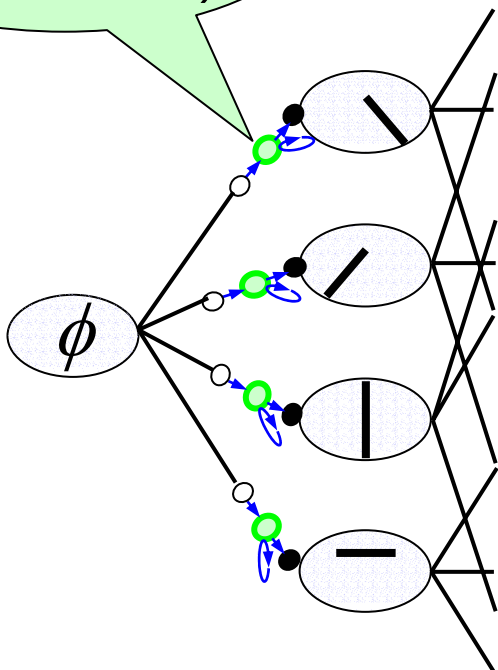
キューブサーチHMMの 評価実験



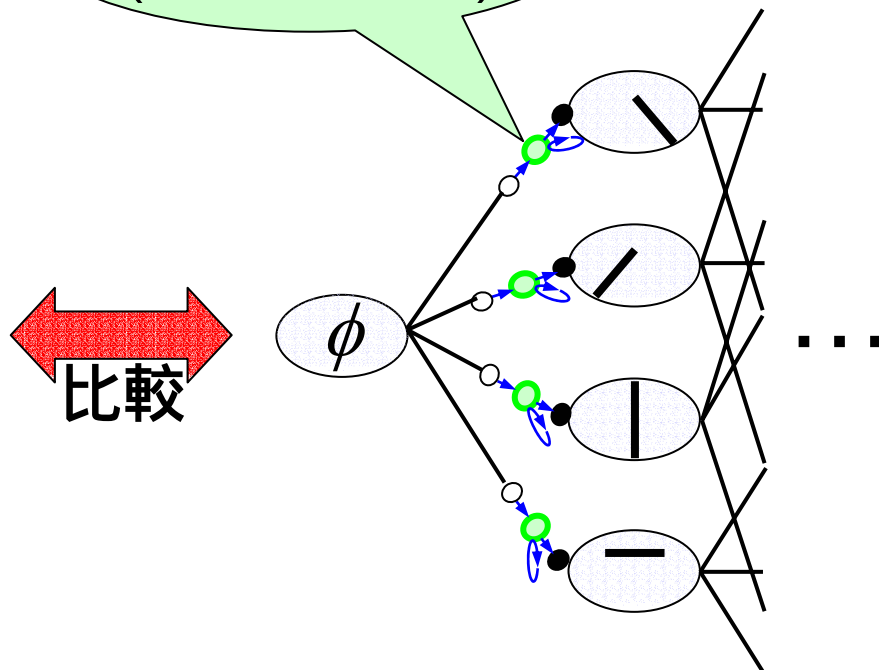
実験の目的(1)



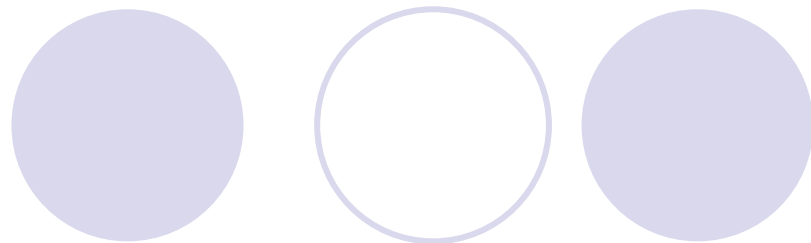
従来法
(方向のみ)



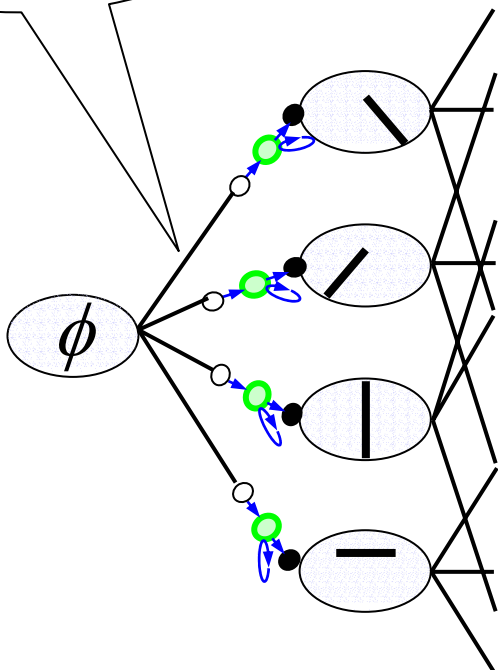
本手法
(座標 + 方向)



実験の目的(2)



一定



学習により
各々決定

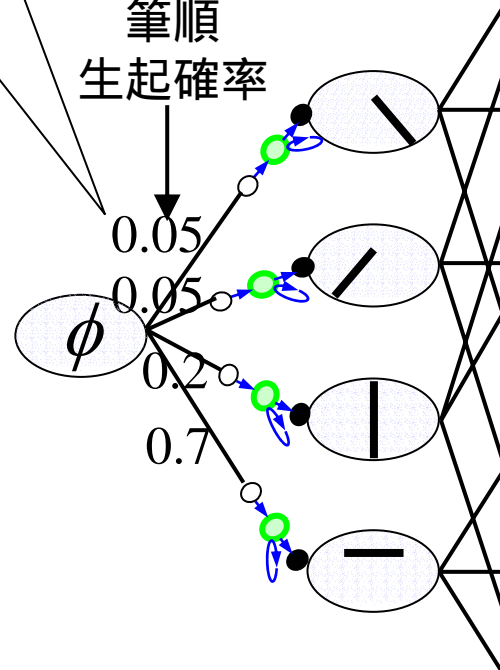
筆順
生起確率

0.05

0.05

0.2

0.7



比較

実験データ

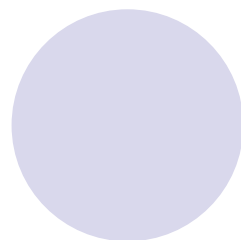
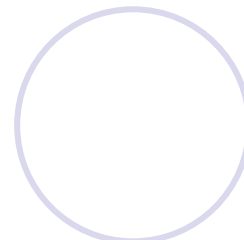
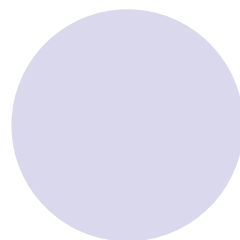
- 筆順自由な10画漢字70文字種
- 1画化の前処理は行わない



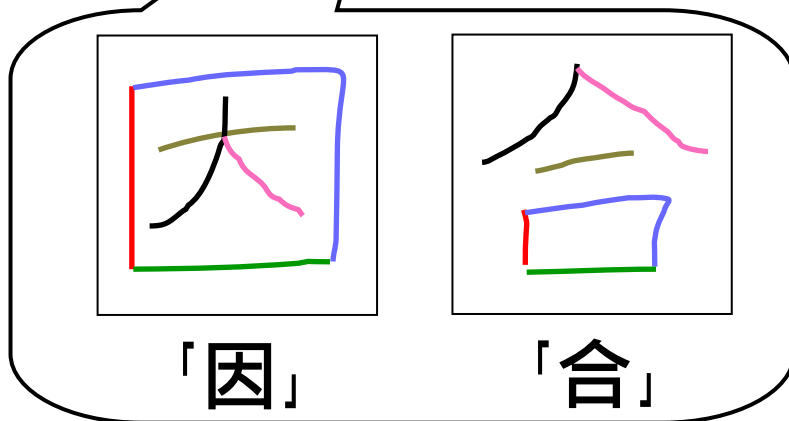
「案」

- データセットA
 - 当研究室の内部データ/筆記者30人分
 - 学習・認識に利用
- データセットB
 - HANDS-kuchibue_d-97-06-10/筆記者10人分
 - 認識のみに利用

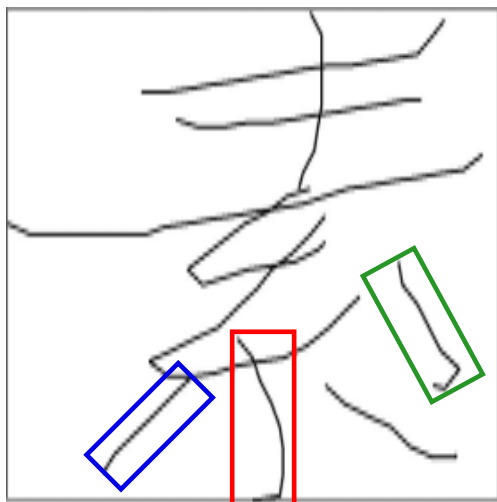
認識率の比較



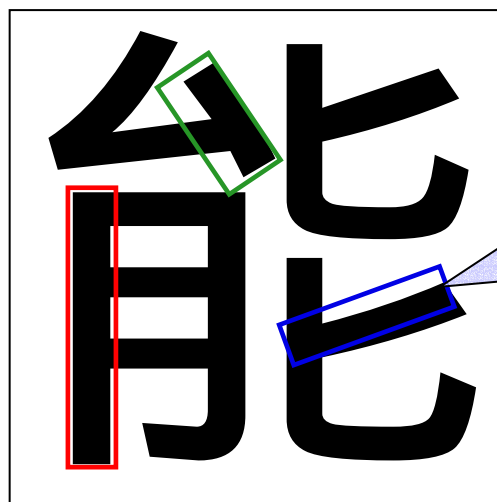
筆順生起確率	不使用		使用	
	A	B	A	B
データセット				
従来法	26.7%	26.4%	79.9%	82.0%
本手法	97.7%	93.9%	98.5%	97.4%



座標情報を利用による改善例



従来法で
誤認識した例
「素」 「能」



従来法による
画対応

方向が同じだと
位置を無視して
対応する



まとめ

- 座標情報と方向情報をHMMの定常性に矛盾なく利用する方法を提案
- 筆順生起確率を用いたキューブサーチHMMを提案
- 提案手法の有効性を確認

今後の展開

- 大規模データセットを利用した学習・認識実験
- キューブサーチHMMの学習法の検討
 - 現在は各ストロークHMMを学習し、固定した後、筆順生起確率を求める2段階の学習法
- 混合正規分布の導入
 - 現在は出力確率を正規分布で近似
- 画の終点座標の利用
- 画数自由化