

# Reconnaissance d'écriture manuscrite et Interaction homme-document

Eric Anquetil

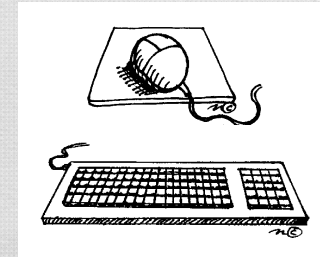
Équipe IMADOC  
IRISA/INSA

Soutenance d'HDR — 27 octobre 2008

## Préambule

### Écriture manuscrite

- *Mode de communication fondamentale chez l'homme*
  - Epanouissement d'une personne (graphothérapie)
  - Maîtrise d'une langue maternelle
  - Formation intellectuelle
- *Constat :*  
*l'informatisation réduit le rôle de l'écriture manuscrite*
  - Suprématie du clavier et de la souris — clavier virtuel
- *Défi : repositionner l'écriture au centre de la communication*
  - Nouvelles technologies orientées « stylo »
  - Nouveaux Usages (Couplage : modalité écrite — gestuelle)





# Préambule

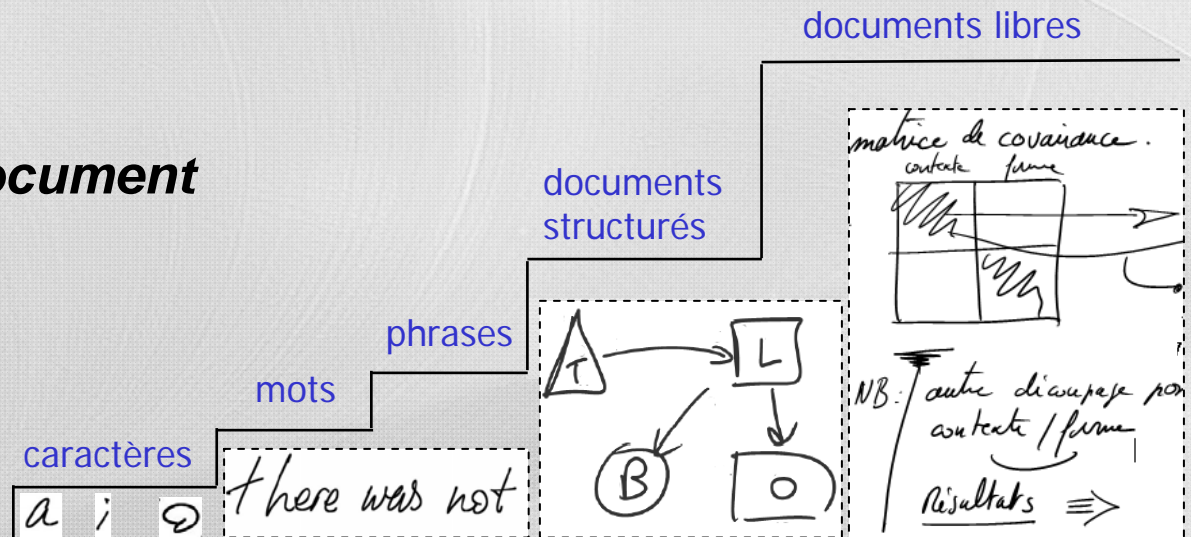
## Interaction homme-document

### ■ Interaction **homme**-document

- Préserver le processus créatif
- La machine s'efface dans cette interaction

### ■ Interaction **homme**-document

- Reconnaissance
- Interprétation



### ■ Interaction **homme**-document

- Orientée « stylo »
- **Utilisateur** rentre dans la **boucle** d'interprétation
  - Il devient un **acteur** du processus d'interprétation

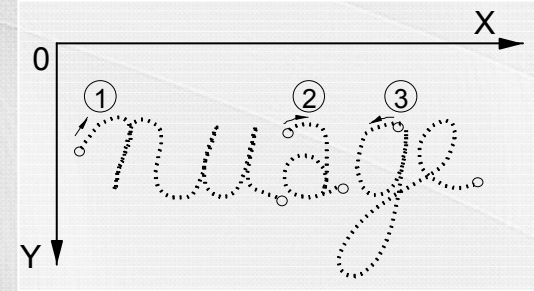


## Préambule

### Contexte de ma recherche — Reconnaissance d'écriture et IHD

#### ■ *Signal : en-ligne*

- Suivi de la trajectoire de la pointe du stylo
- ➔ restitué sous forme d'encre électronique



#### ■ *Applications*

- Méthodes de saisie, éditeurs de schémas, de documents structurés
  - « à main levée » ↔ tracés manuscrits
- Ordinateurs stylo : téléphonie, tabletPC, TBI, ...
  - Contraintes temps de calcul, encombrement mémoire, ergonomie

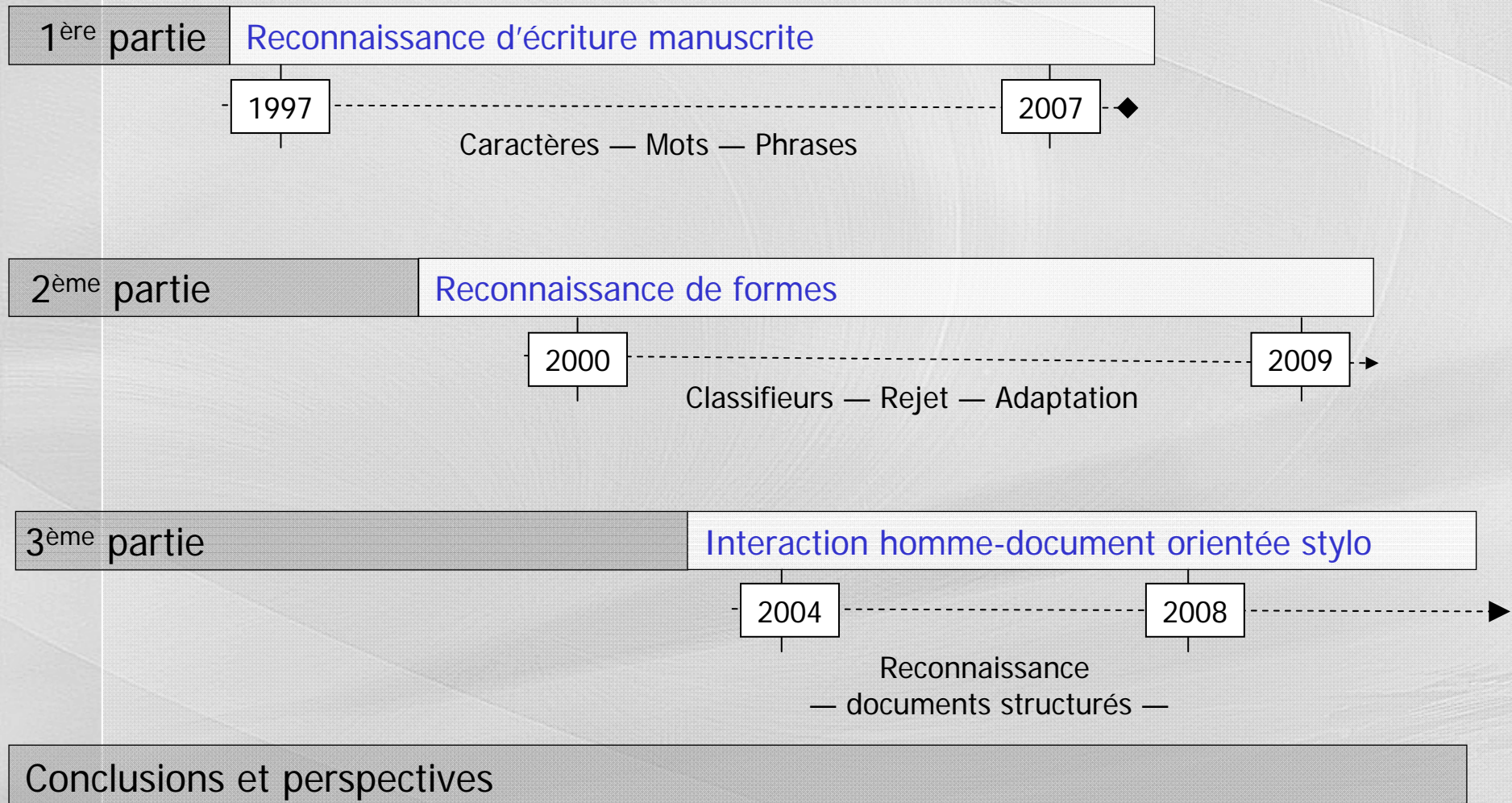
#### ■ *Thématiques scientifiques*

- Reconnaissance de Formes (écrits et documents manuscrits)
  - *sélection de caractéristiques, segmentation, modélisation, classification, apprentissage, combinaison et fusion d'informations, gestion du contexte, intégration de connaissances linguistiques...*
- Interaction Homme-Machine
  - Ergonomie, utilisabilité, acceptabilité des interfaces stylo



# Fil conducteur

## Trois axes de recherche



1<sup>ère</sup> partie  
Reconnaissance d'écriture  
manuscrite

1997

2007

Caractères — mots — phrases

a ; o

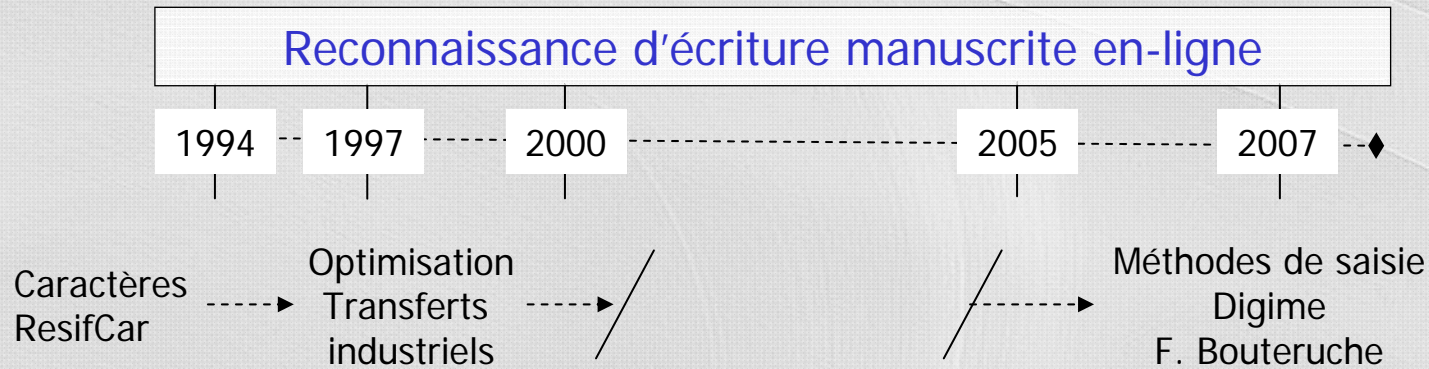
was

there was not a bit of trouble



# 1<sup>ère</sup> partie | Reconnaissance d'écriture manuscrite

## Reconnaissance de caractères isolés



### Objectifs

- Reconnaissance de caractères manuscrits
- Ergonomie des méthodes de saisie (*Thèse de F. Bouteruche*)

### Approches

- Reconnaissance hybride
  - Représentation hiérarchique structurale des caractères
  - Modélisation par Système d'Inférence Floue (SIF)

- ➔ **ResifCar** : Moteur de reconnaissance de caractères isolés
- ➔ **Digime** : Méthodes de saisie d'écriture



# 1<sup>ère</sup> partie | Reconnaissance de caractères

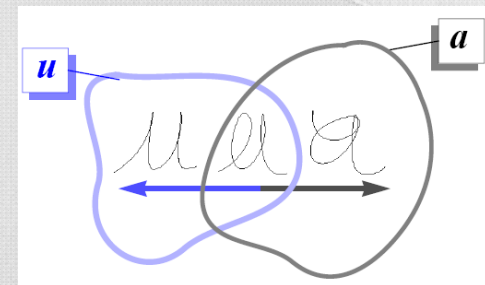
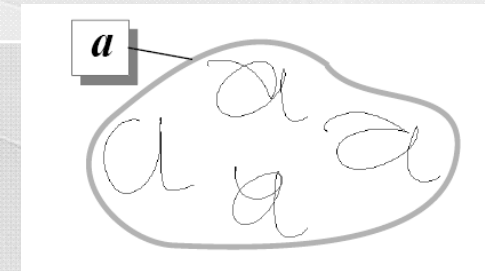
## Positionnement de l'approche isolés

### ■ *Approches*

- Statistiques : point de vue global
  - Vecteur de caractéristiques
  - Classifieurs connus : Kppv, MLP, RBFN, SVM, ...
- Structurelles : analyse fine
  - Décomposition en primitives élémentaires
  - Règles d'agencement des primitives
- Bilan
  - Performances : Statistique > Structurelle
  - Interprétabilité : Structurelle > Statistique

### ■ *Objectif en 1994*

- Existant : système Newton d'Apple (93) / Graffiti (96)
  - ➔ le grand gagnant : le clavier virtuel [MacKenzie'02]
- Défi pour la conception de méthodes de saisie
  - Système de reconnaissance
    - Équilibre entre : (performances — contraintes)
  - IHM
    - Nouveaux Usages — Interaction avec l'utilisateur





### ■ *ResifCar : une approche hybride*

- Représentation hiérarchique structurelle des caractères
- Modélisation par Système d'Inférence Floue (SIF)

### ■ *Pourquoi ?*

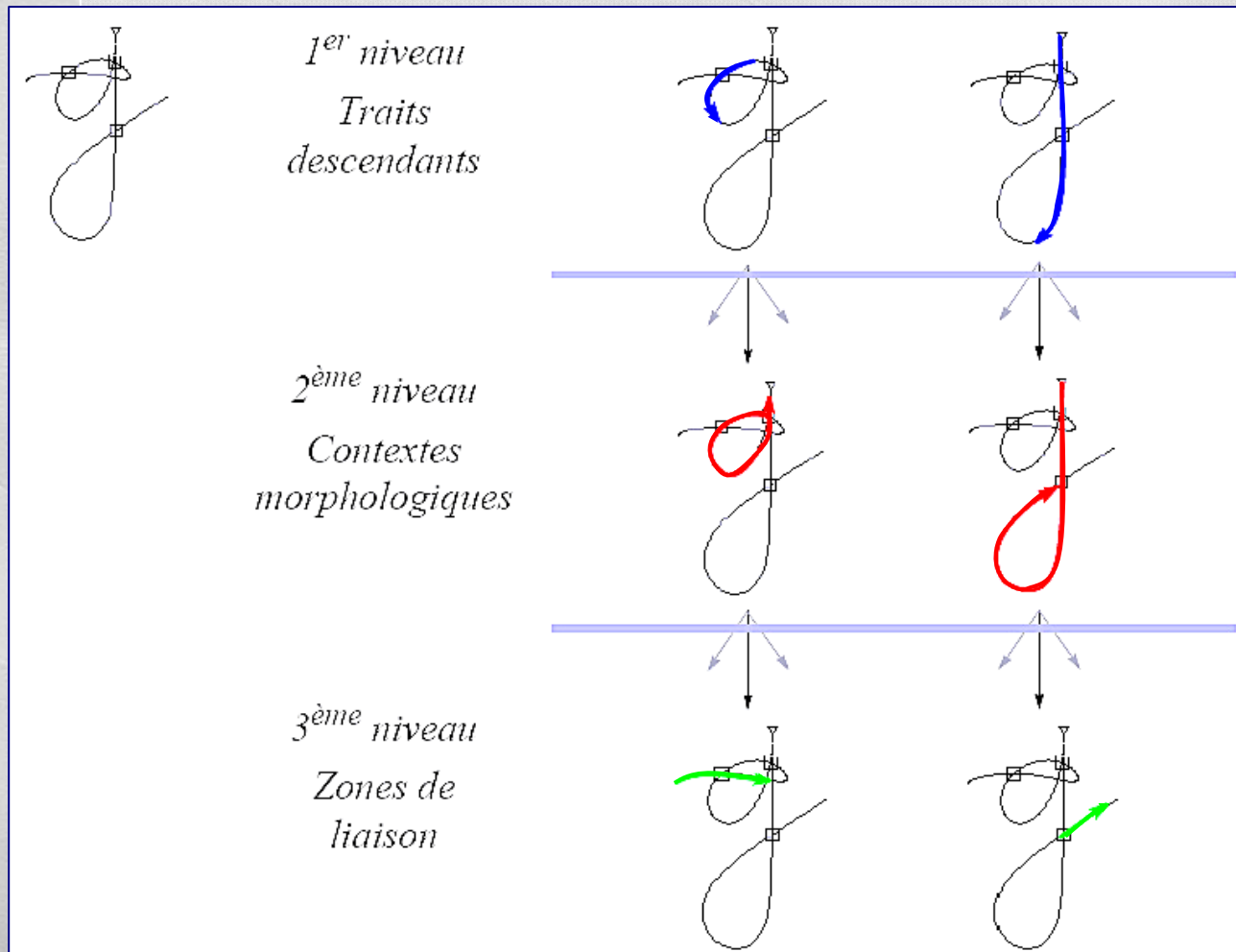
- Profiter des connaissances (expertise) a priori sur l'écriture
  - Ossature — zones fondamentales des caractères latins
  - Points d'ancrage visuel
- Garder la maîtrise des connaissances modélisées  
Modélisation explicite — système « transparent »
  - Comprendre et qualifier les erreurs
  - Faire évoluer le système (optimisation, adaptation, ...)
  - Garantir l'interopérabilité — fusion d'informations (caractères → mots → phrases)
- Contraintes d'embarquement (temps de calcul et taille mémoire)
  - Une force des Systèmes d'Inférence Floue (SIF)



# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Hiéarchisation des connaissances

### ■ *Trois niveaux de représentation*



De bons points d'ancrage pour construire une analyse structurale

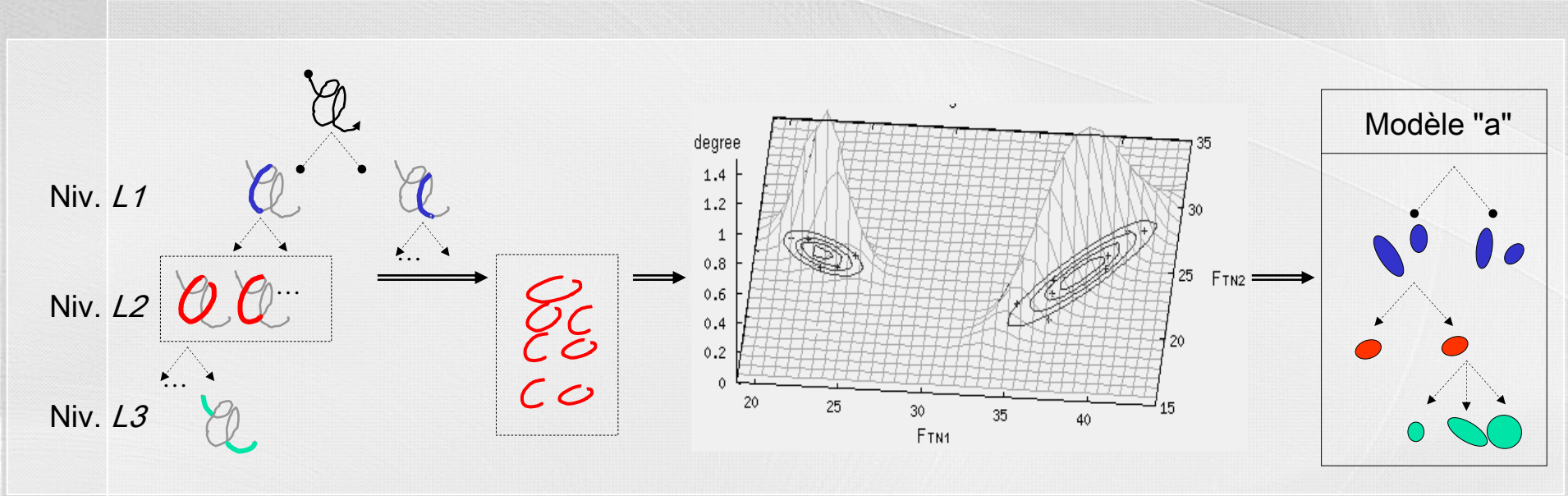
Modélisation de la morphologie du caractère

Focalisation sur des éléments plus précis qui peuvent s'avérer discriminants



# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Modélisation logique — Apprentissage



(a)  
Segmentation physique

(b)  
Vecteur d'attributs  
— L1: 4 dimensions  
— L2: 10 dimensions  
— L3: 5 dimensions

(c)  
Classification non supervisée **possibiliste**  
[Krishnapuram'93]

(d)  
Modélisation logique hiérarchique des **prototypes**  
(**Centre+Matrice de covariance**)



# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Modélisation logique — formalisation par des SIF

### Principe

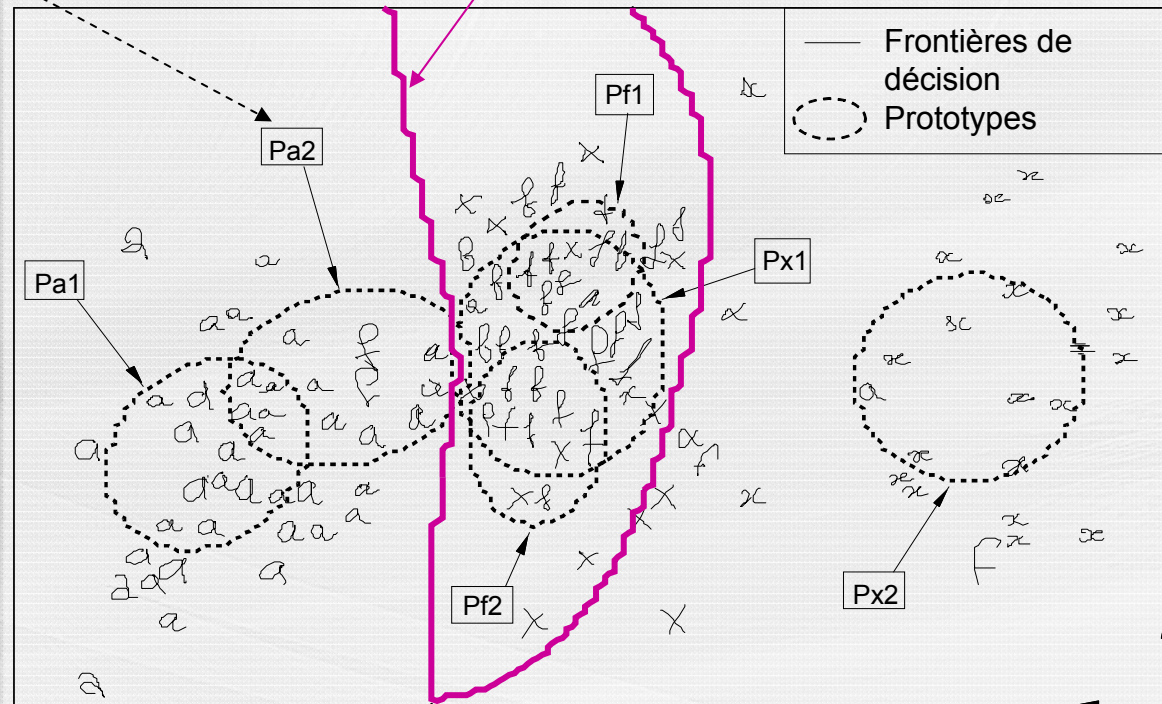
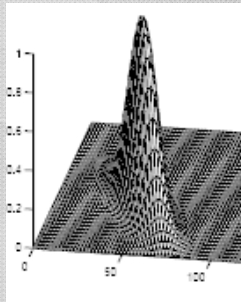
- SIF = {règles floues}
- Chaque règle est de type Takagi-Sugeno d'ordre 0 [Takagi'85]

$$r^k : \left( X \text{ est } P^r \right) \quad \text{Alors} \quad s_1^r = a_1^r \quad \dots \quad \text{et} \quad s_C^r = a_C^r$$

$$m^r(x') = \frac{1}{1 + d_{Q^r}(x', h^r)}$$

fonction de Cauchy  
avec la distance de  
Mahalanobis

$$d_{Q^r}(x', h^r)$$



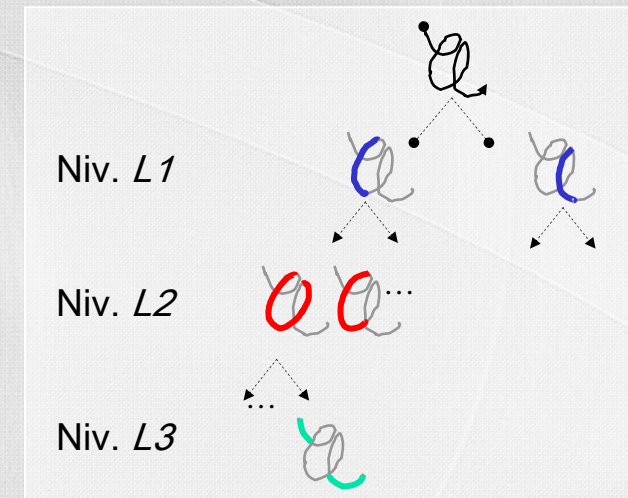


# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Modélisation logique — formalisation par des SIF

### ■ Règle $r^k$ de la classe $k$

- Prémisse (modélisation intrinsèque)
  - conjonction des prototypes pour les 3 niveaux  $L1, L2, L3$
- Conclusion (pondération discriminante)
  - impact des prototypes sur l'évaluation de l'appartenance aux  $C$  classes.



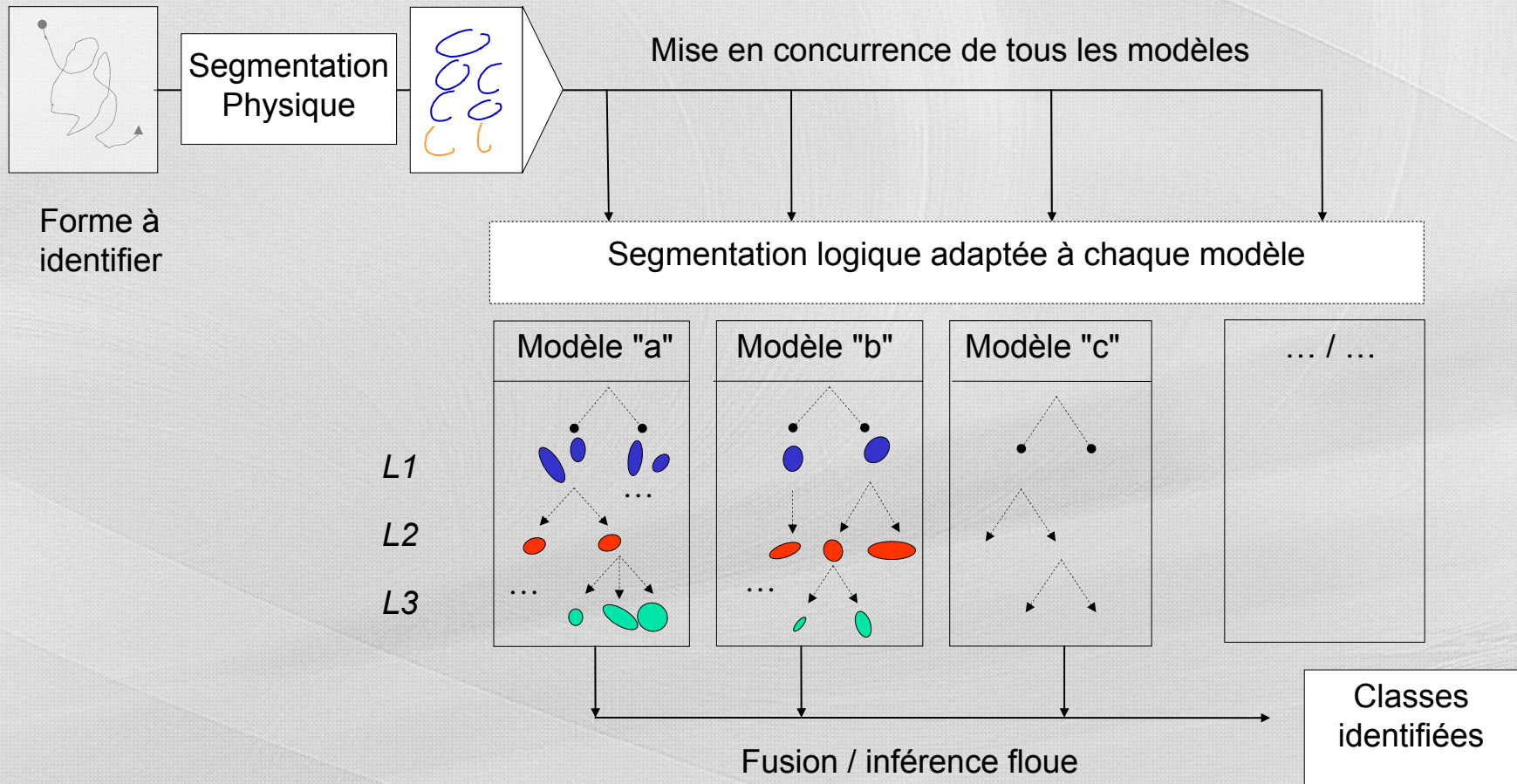
$$\begin{array}{l}
 r^k : \left( X^{L1} \text{ est } P_1^{r,L1} \text{ ou } P_2^{r,L1} \text{ ou } \dots \right) \\
 \text{et} \left( X^{L2} \text{ est } P_1^{r,L2} \text{ ou } P_2^{r,L2} \text{ ou } \dots \right) \\
 \text{et} \left( X^{L3} \text{ est } P_1^{r,L3} \text{ ou } P_2^{r,L3} \text{ ou } \dots \right)
 \end{array}
 \quad \text{Alors} \quad s_1^r = a_1^r \quad \dots \quad \text{et} \quad s_C^r = a_C^r$$



# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Phase de reconnaissance

### ■ *Vision d'ensemble*

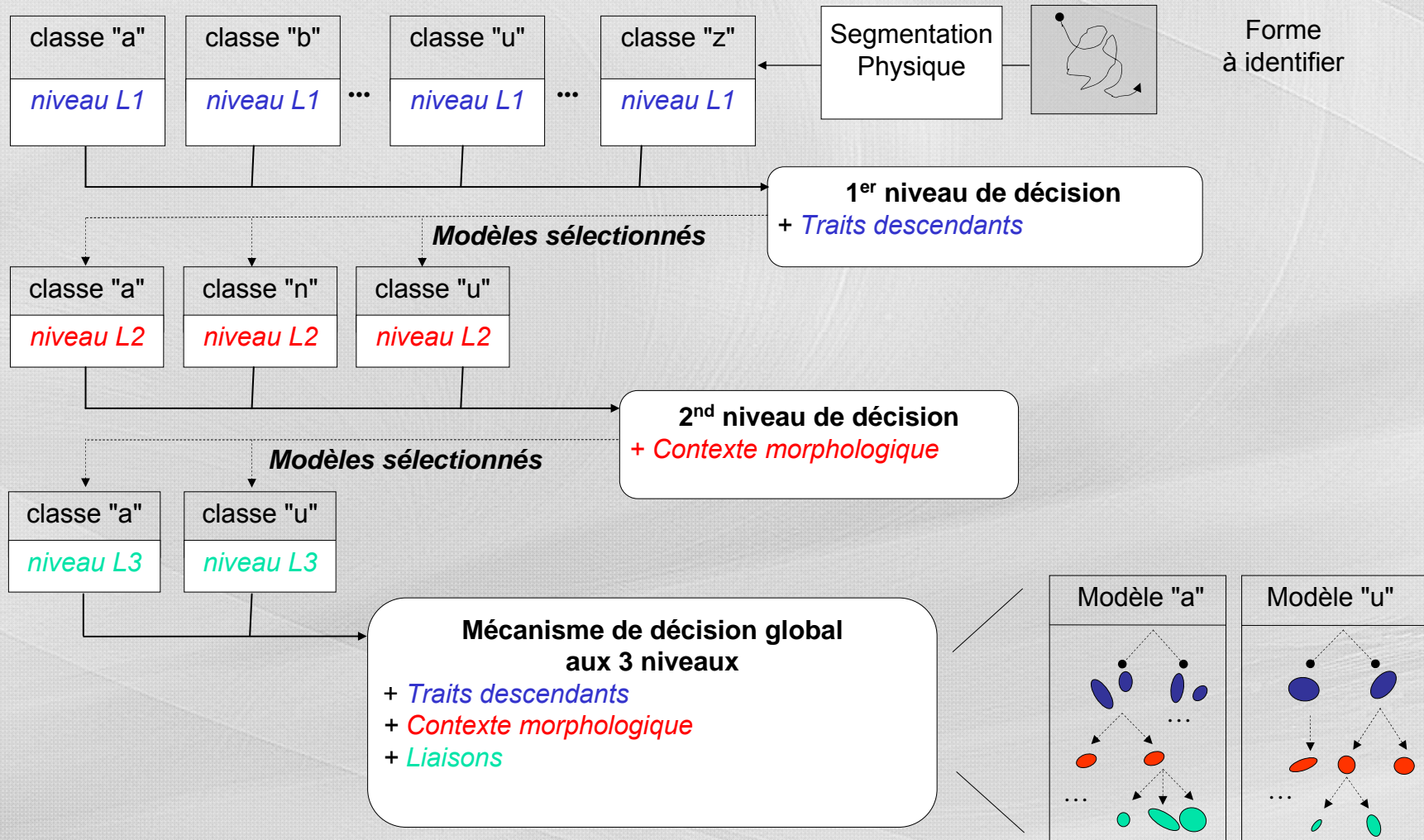




# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Optimisation du processus de décision

### ■ Sélection progressive des modèles (70% d'élagage)





# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Résultats — Transfert industriel

2000-2001 : Téléphonie – Reconnaissance de caractères

### ■ *Transfert industriel*

- Sensei (V-tech) — PurpleLabs



### ■ *Défi : embarquer ResifCar (en version bridée)*

- Processeur ARM7 TDMI : 13 Mhz, 250 Ko



### ■ *Résultats (sur 300 scripteurs)*

- Temps de calcul : (11s/car) → Après optimisation (0.5s/car)
- Performances : 92% (Top1) — 97% (Top3)
- Place mémoire : 40Ko en Ram, 120Ko en Rom



### ■ *Retombées*

- Commercialisation en Europe (Italie, Autriche, Espagne)
- Publications communes (ICPR en 2002)
- Importance de l'interaction homme-machine / Méthode de saisie  
→ Naissance de nouveaux travaux de recherche



# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Résultats — Transfert industriel

### ■ 2007 : Avionique — Reconnaissance de chiffres

- Transfert industriel
  - Thales
- Défi : Optimiser les « procédures avion » — acquisition de fréquences radio
  - Performances maximales et « cohérence » des erreurs : (ResifCar non bridée)
  - Traitement d'écritures « atypiques » (scripteurs gauchers)
- Résultats
  - Très bons retours des pilotes
  - Nouveau contrat en construction

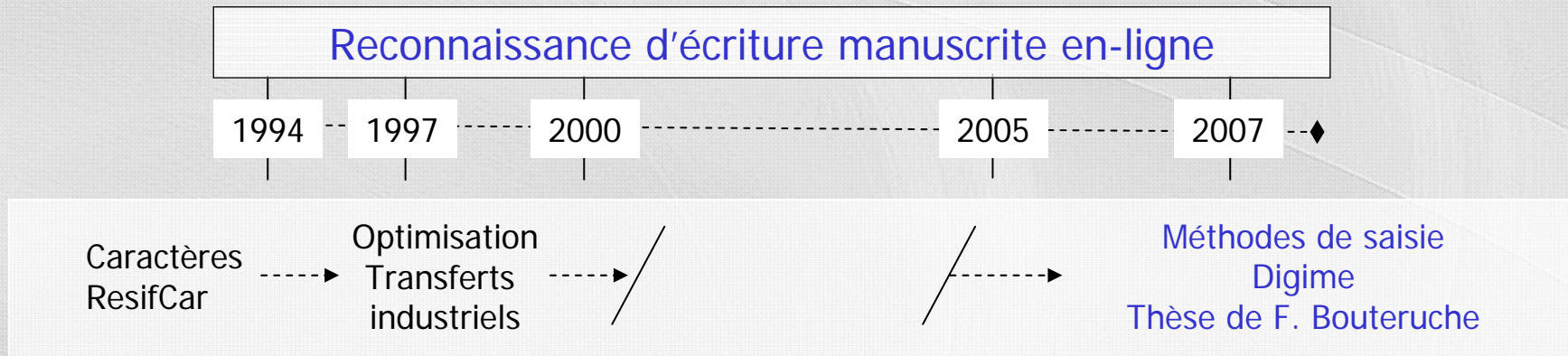


| Bases de test                            | Nombre de tracés | Taux de reconnaissance RESIFCar V5.0 |   | Taux de reconnaissance RESIFCar V5.1b | Diminution du taux d'erreurs |
|--|------------------|--------------------------------------|---|---------------------------------------|------------------------------|
| B1<br>(Unipen, Ironoff)                  | 20665            | 95,80%                               | ➡ | <b>96,20%</b>                         | <b>-9,52%</b>                |
| B2<br>(26 scripteurs gauchers atypiques) | 1150             | 87,90%                               | ➡ | <b>93,80%</b>                         | <b>-48,76%</b>               |

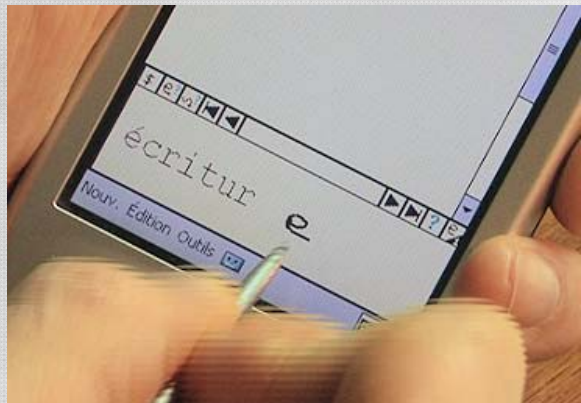


# 1<sup>ère</sup> partie | Reconnaissance de caractères isolés | ResifCar

## Méthode de saisie — Digime



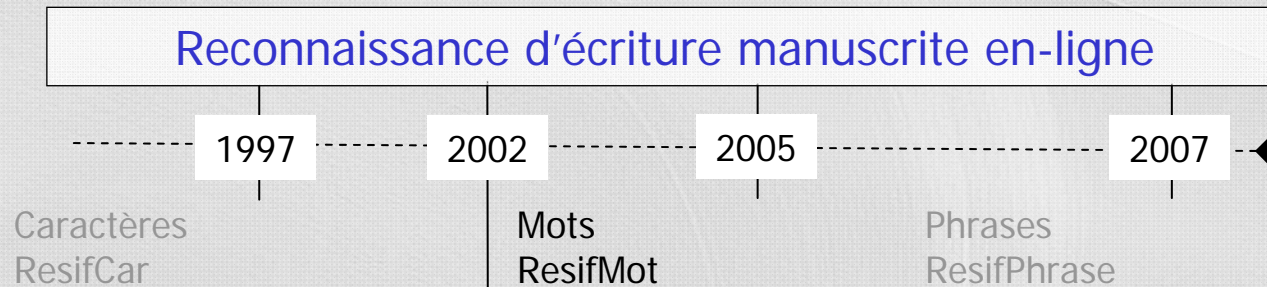
- **Micro-éditeur d'encre électronique (DIGIME)**
  - Autorise l'écriture en plusieurs tracés d'un caractère
  - Caractères en contexte de mot
    - Segmentation automatique
  - Reconnaissance des gestes graphiques
    - Fonctionnalités d'édition (suppression, insertion, ...)
    - Signes d'accentuation, ponctuation, etc.





# 1<sup>ère</sup> partie | Reconnaissance d'écriture manuscrite

## Reconnaissance de Mots cursifs



### Objectifs

- Post-traitement lexical / Grands vocabulaires
- Apprentissage de distance d'édition

### Thèse

- S. Carbonnel (2005, Ing. de Rech. à l'école des Mines d'Albi)



# 1<sup>ère</sup> partie | Reconnaissance de mots cursifs | ResifMot

## Positionnement de l'approche

### ■ *Problématique*

- Couplage : segmentation — reconnaissance
- Intégration de connaissances linguistiques/lexicales

### ➔ Trois grandes approches

#### ■ *Approches globales*

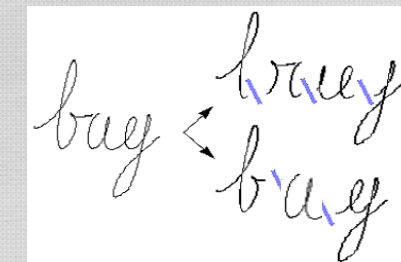
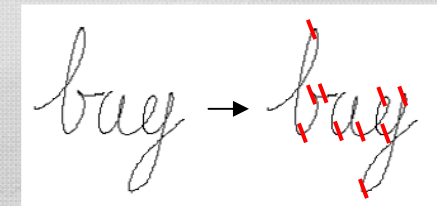
- 1 modèle / mot  
➔ petits vocabulaires

#### ■ *Approches analytiques — segmentation implicite*

- 1 modèle / classe de caractères  
➔ lexiques de grande taille

#### ■ *Approches analytiques — segmentation explicite*

- 1 modèle / classe de caractères  
➔ lexiques de grande taille
- ⊕ Modélisation plus explicite
- ⊖ Dépendant de la qualité des hypothèses de segmentation  
Limiter la combinatoire ↔ sans oublier d'hypothèses





# 1<sup>ère</sup> partie | Reconnaissance de mots cursifs | ResifMot

## Analyse physique — Segmentation explicite

*simple*

Analyse  
physique

### Objectif

- Limiter la combinatoire engendrée par le treillis de segmentation

### Principe

- Segmentation ancrée sur les traits descendants

*simple*

*simple*

traits descendants  
fondamentaux

niveau 1  
(1 trait)

*simple*

niveau 2  
(2 traits)

*simple*

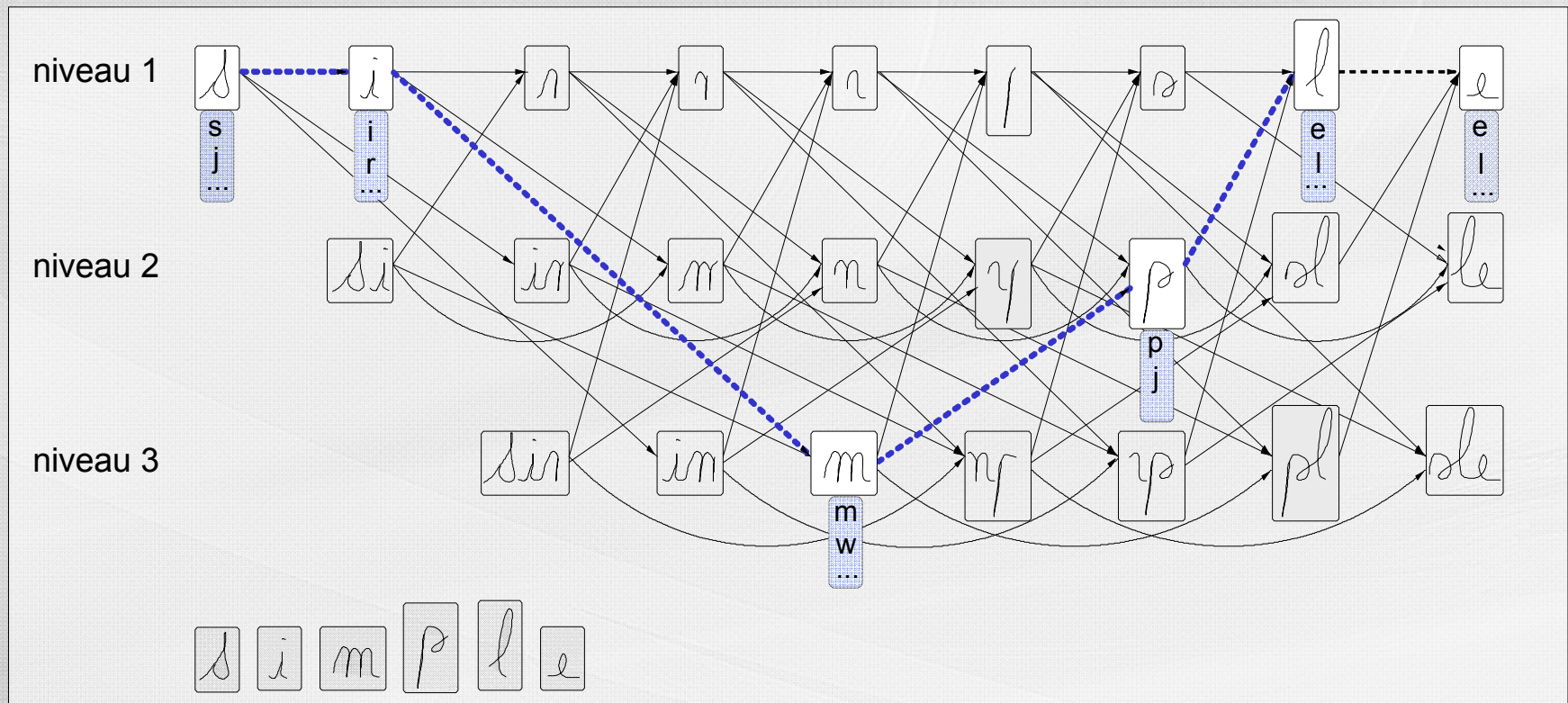
niveau 3  
(3 traits)

*simple*



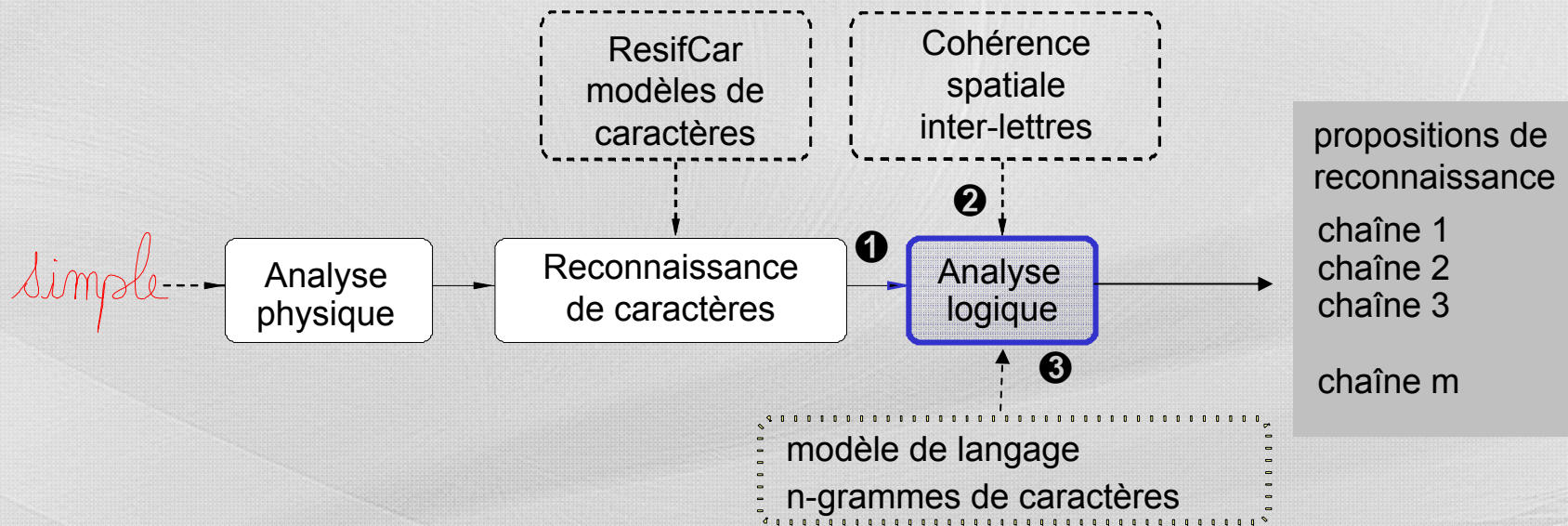
■ *Treillis de segmentation*

- Cohérence des chemins
  - Normalisés / l'ossature formée par les traits descendants
- Reconnaissance des caractères —ResifCar



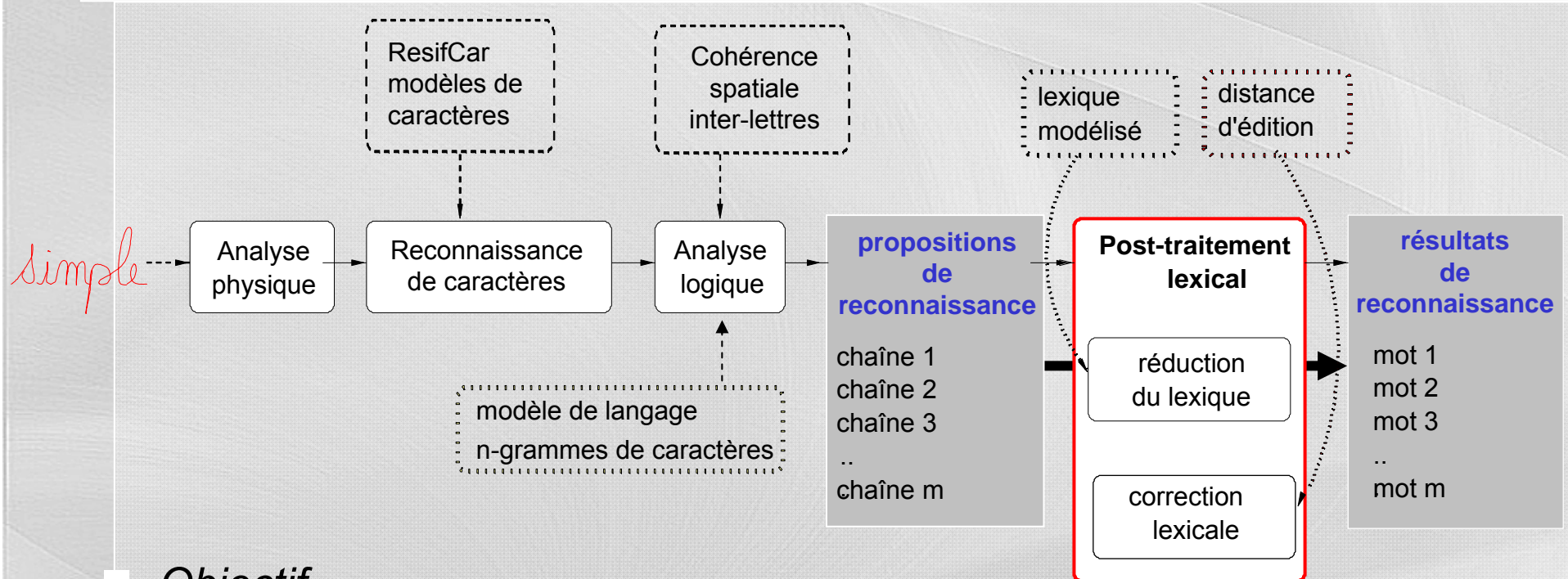


- *Evaluation de la pertinence des chemins*
  - Fusion de trois sources d'information





## Post-traitement lexical



### Objectif

- Vocabulaire de taille importante (>20 000 mots)
- Qualité de la reconnaissance — temps de réponse — encombrement mémoire

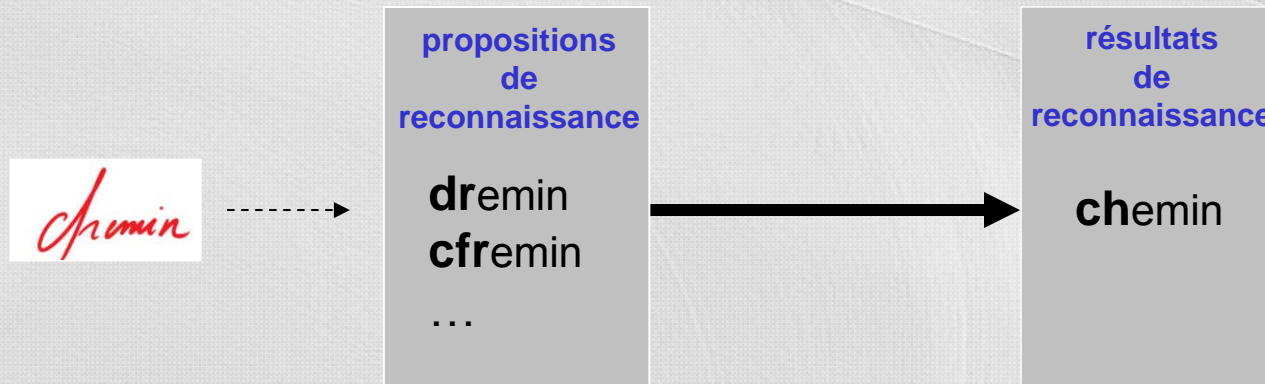
### Approche : post-traitement lexical basé sur un lexique

- Stratégie de réduction et d'organisation de dictionnaire (/20 la taille)
- Apprentissage automatique d'une distance d'édition



# 1<sup>ère</sup> partie | Reconnaissance de mots cursifs | ResifMot

## Post-traitement lexical — Distance d'édition

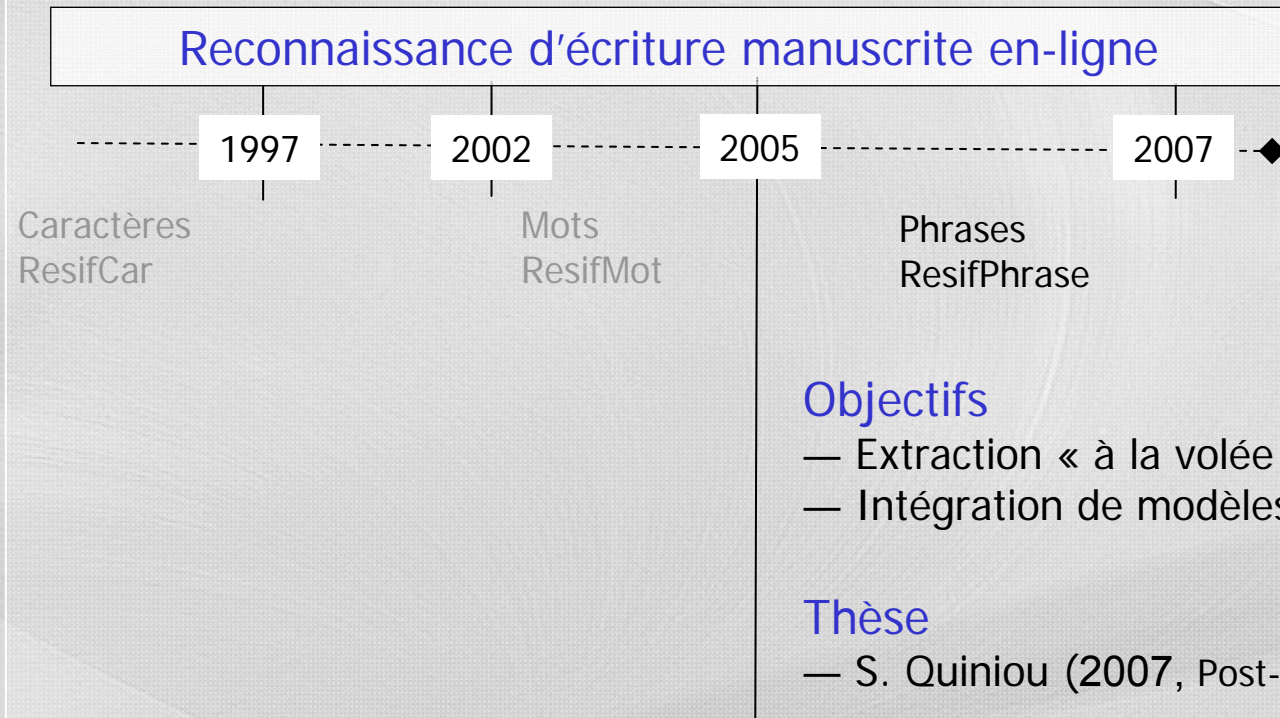


- **Existant :** *Distance empirique spécialisée pour l'écriture manuscrite* [Seni et al.'96a]
- **Contribution :** *Apprentissage de la distance d'édition* [Carbonnel'05]
  - Apprentissage des fréquences des opérations d'édition
  - Basé sur les techniques de « Boosting » [Schapire'99]
    - focus sur les erreurs du système → ajuster les opérations et les coûts d'édition
- **Résultats**
  - Amélioration des taux (lexique 25000 mots)
    - 86,03 % (Top1) — 95,38 % (Top10)
    - Réduction du taux d'erreur de 8% à 42% en top 10 / coûts empiriques
  - Système adaptable aux évolutions (version, contexte d'utilisation, ...)



# 1<sup>ère</sup> partie | Reconnaissance d'écriture manuscrite

## Reconnaissance de phrases



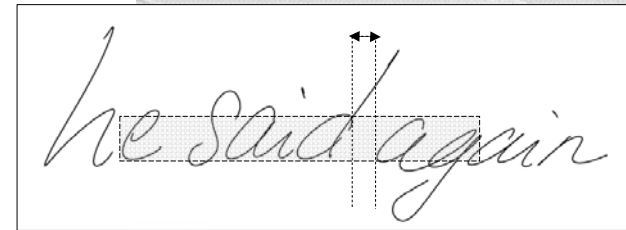


# 1<sup>ère</sup> partie | Reconnaissance de phrases | ResifPhrase

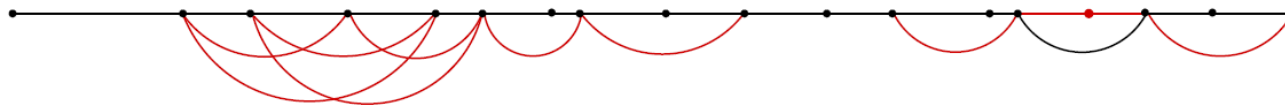
## Vision générale

generally the same level of living costs less  
in the country than in the city

Extraction des mots



generally the same level of living costs less in the country than in the city



RESIFMot

Reconnaissance de phrases

Modèles de langage

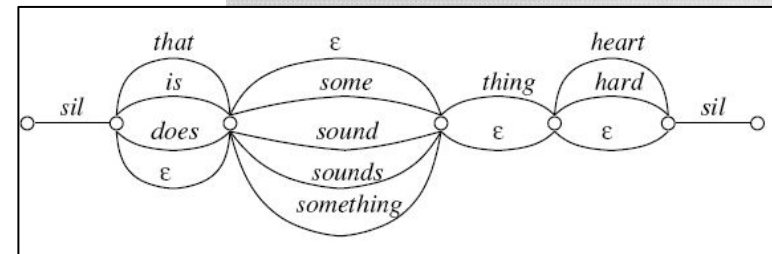
generally the same level of living costs less in ~~me~~ country ~~them~~ in the city

Détection des erreurs

Classifieurs

Rejet

Correction



generally the same level of living costs less in ~~me~~ country ~~than~~ in the city



# 1<sup>ère</sup> partie | Reconnaissance de phrases | ResifPhrase

## Modèles de langage statistiques

### ■ 1<sup>ère</sup> phase

- Impact des modèles de langage statistiques sur la reconnaissance de phrase

### ■ 2<sup>nde</sup> phase

- Combinaison de différents modèles de langage

### ■ 3<sup>ème</sup> phase : *Contributions principales*

- **Identification** : mettre en évidence les ambiguïtés sur les mots
  - réseau de **confusion** (**approche consensus** [Hillard'06])
- **Correction** des erreurs détectées
  - classifieurs dédiés (SVM — mécanisme de rejet)

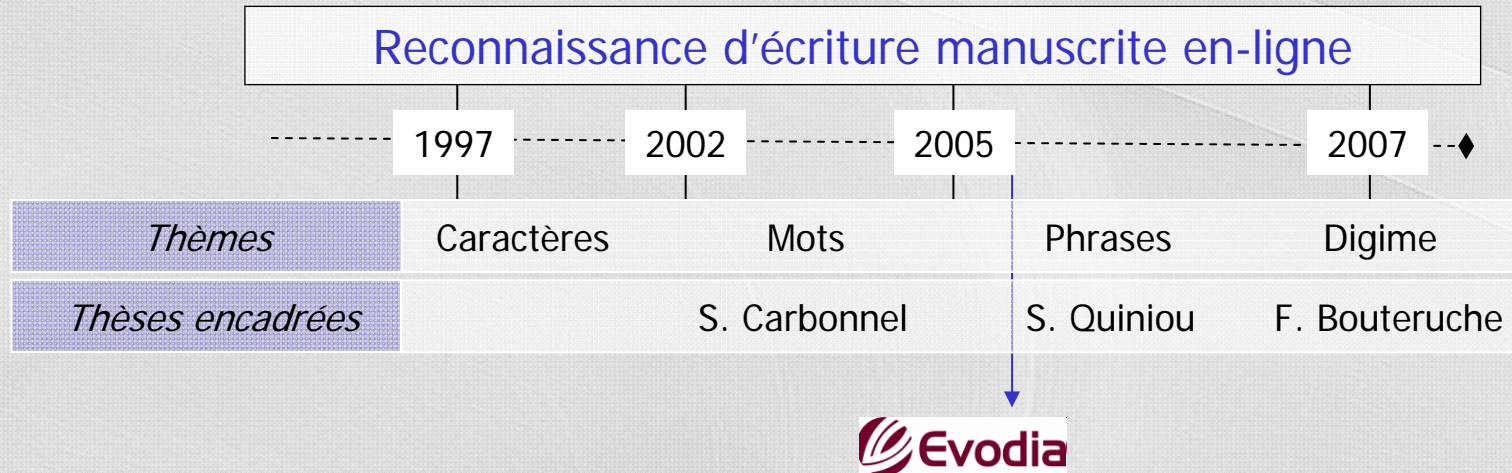
### ■ *Résultats*

- Cadre
  - **corpus de Brown** (46 836 phrases ; 900 109 mots), lexique de 13 748 mots.
  - base d'apprentissage : 517 phrases (8 047 mots) — 25 scripteurs
  - base de test : 425 phrases (6 362 mots) — 17 scripteurs.
- Taux d'erreur — Top1/mots :
  - sans modèle de langage — 28,6 % d'erreurs
  - avec modèle de langage — 13,5 % d'erreurs
  - avec modèle de langage & réseau de confusion — 11,5 % d'erreurs



# 1<sup>ère</sup> partie | Reconnaissance d'écriture manuscrite

## Bilan



- Approche complète pour la reconnaissance de l'écriture manuscrite en-ligne

- **Collaborations**

- France Telecom R&D
- Univ. Rennes 2 (CRPCC, psychologie expérimentale)
- équipe IRISA Texmex : traitement automatique des langues
- équipe IRISA Metiss : reconnaissance de la parole



- **Transferts industriels**

- Téléphonie
  - Sensei
  - Purplelabs
- Avionique
  - Thales (via Evodia)





# 1<sup>ère</sup> partie | Reconnaissance d'écriture manuscrite

## Perspectives

### Reconnaissance d'écriture manuscrite en-ligne

1997

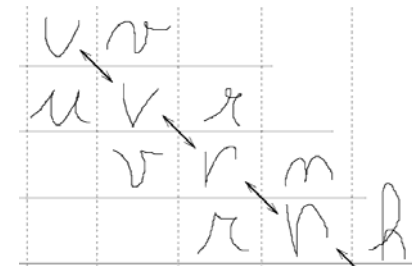
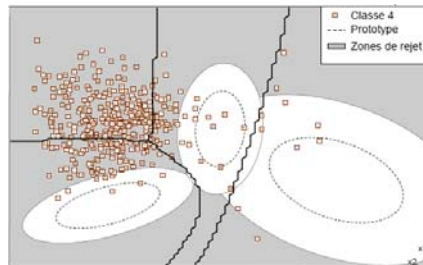
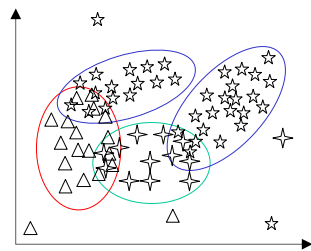
2007

#### ■ *Perspectives*

- Optimiser globalement le système
  - Interaction entre les différents étages de reconnaissance
  - Synergie entre les sources d'information : fusion de connaissances
- Passer à l'échelle au niveau des expérimentations (mots et phrases)
  - Etendre les tests à un cadre d'utilisation en vraie grandeur
- Considérer d'autres langues que l'anglais
  - Approfondir les modèles de langage



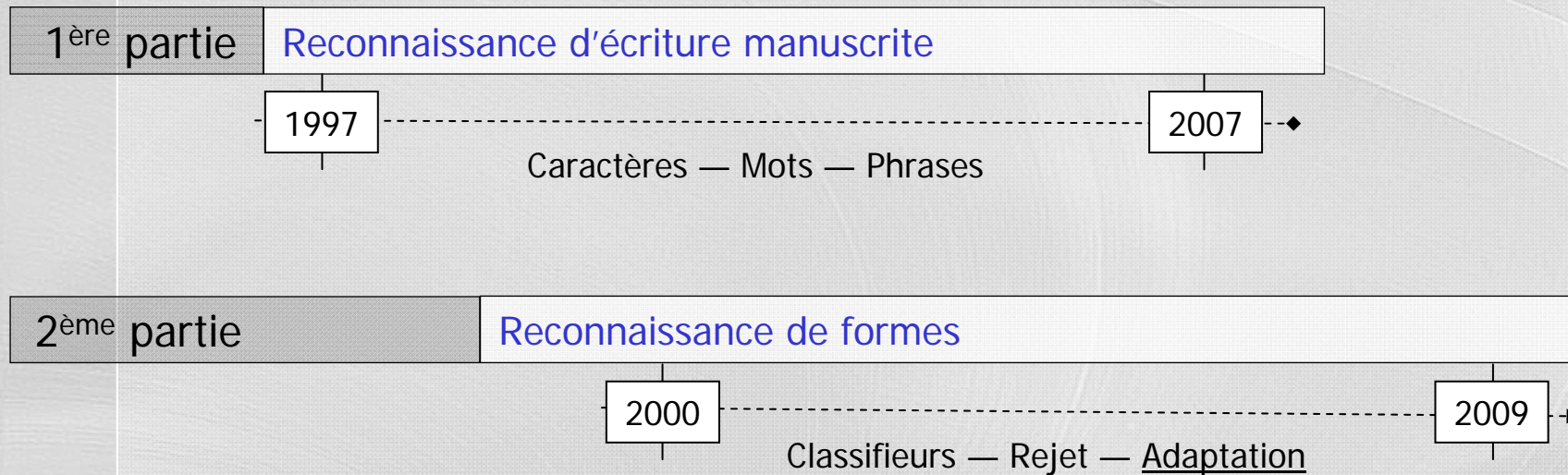
## 2<sup>ème</sup> partie Reconnaissance de formes





## 2<sup>ème</sup> partie | Reconnaissance de formes

### Classifieurs, rejet, adaptation



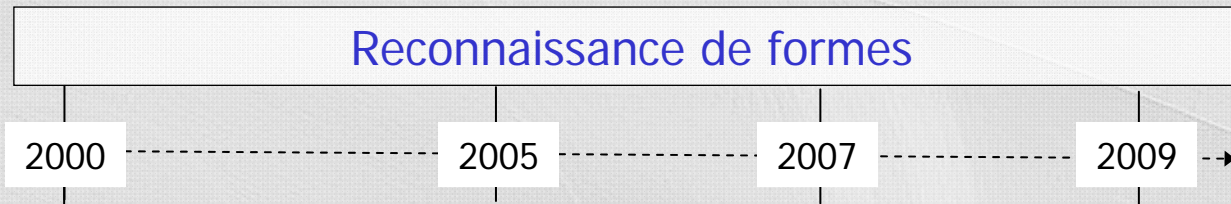
#### ■ *Recherches plus fondamentales*

- Objectifs
  - Lever certains verrous pour faire évoluer nos systèmes de reconnaissance
- Cadre
  - Formes manuscrites
  - Contraintes d'embarquement et de temps de calcul
  - Classification à base de prototypes / Systèmes d'Inférence Floue (SIF)



## 2<sup>ème</sup> partie | Reconnaissance de formes | Mélidis

### Classifieurs mixtes



#### Thèmes

Classifieurs mixtes

Rejet

Adaptation

#### Objectif

- Etendre l'approche de ResifCar
  - *Reconnaitre d'autres formes manuscrites*

#### Thèse

- N. Ragot (2003, MC à Tours)

#### Approche

- Classifieurs mixtes
  - *Modélisation générative et discriminante*
  - *Mécanisme de focalisation*

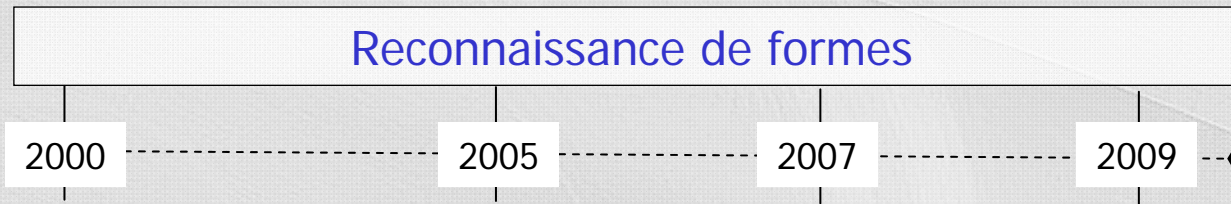
#### Résultats : le système Mélidis

- Bon compromis en termes de compacité — performance
  - *Voisins des SVM avec 10 à 30 fois moins de paramètres*



## 2<sup>ème</sup> partie | Reconnaissance de formes | AMTL

### Mécanismes de Rejet



#### Thèmes

Classifieurs mixtes

Rejet

Adaptation

#### Objectif

— Estimer la qualité des réponses du classifieur

#### Thèse

— H. Mouchère (2007, MC à l'IUT de Nantes)

#### Approche

— Mécanisme de rejet

◦ *Rejet d'ignorance*

◦ *Rejet d'ambiguïté*

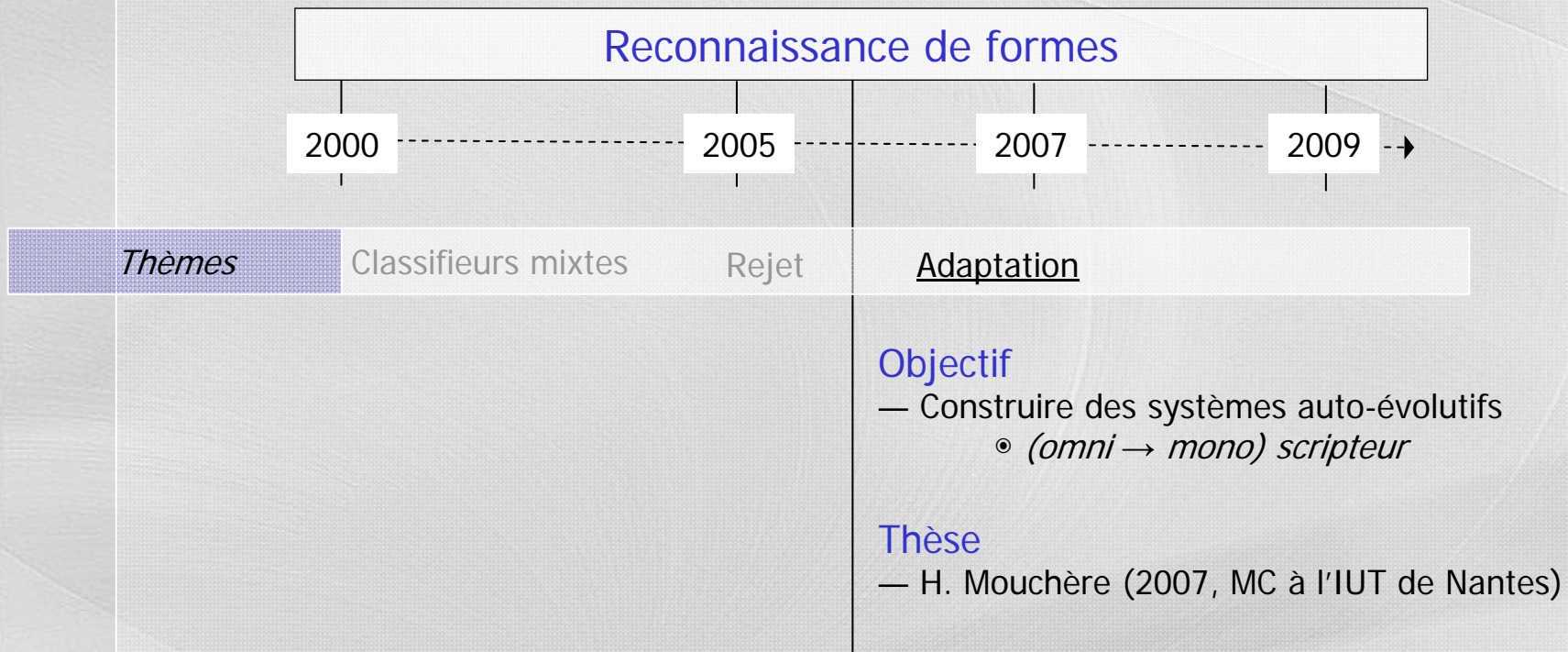
— Approche à base de fonctions de confiance

#### Résultats : Algorithme AMTL

— Optimisation efficace des seuils de rejet

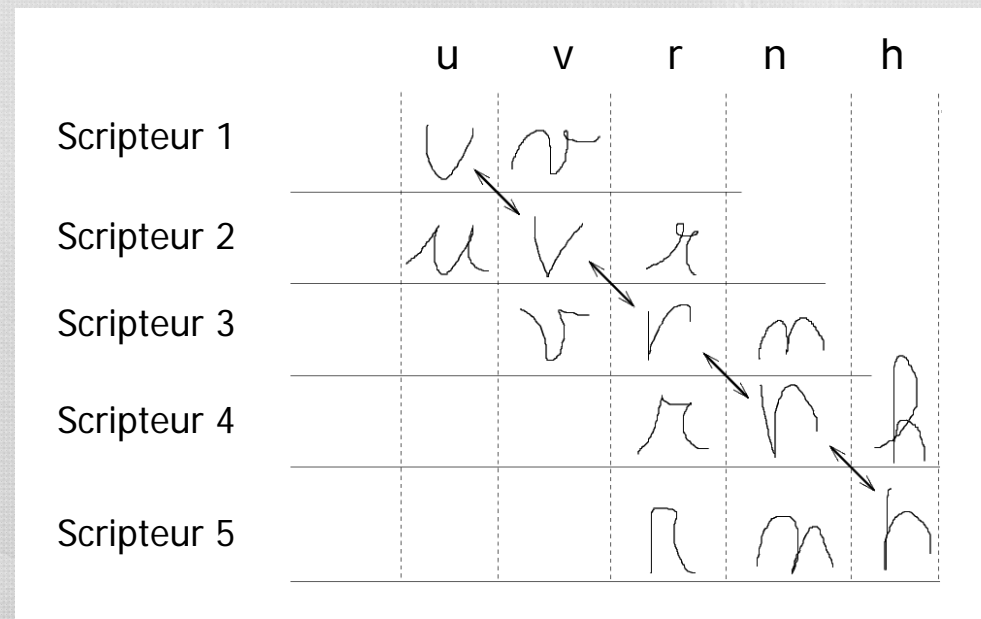


## 2<sup>ème</sup> partie | Reconnaissance de formes | Adaptation





- *Pourquoi avoir des systèmes de reconnaissance adaptés*
  - Un scripteur = Un jeu propre d'allographes
    - Personnalisation et limitation des risques de confusion
  - Meilleur compromis performance/encombrement (mono-scripteur)



- *Stratégie*
  - Système omni-scripteur → Système mono-scripteur
    - spécialisation progressive



## 2<sup>ème</sup> partie | Reconnaissance de formes | Adapt Stratégie d'adaptation — incrémentale

### ■ *Défi*

- Système d'adaptation « à la volée »
  - Très réactif
    - s'adapter à partir de très peu de caractères du scripteur
  - Stable
    - éviter les comportements chaotiques pendant l'apprentissage
  - Facile d'usage
    - transparent Mais perceptible à l'utilisateur
- Contraintes applicatives / téléphonie
  - Embarquer sur la machine
    - le *système de reconnaissance* et son *mécanisme d'adaptation*

### ■ *Stratégie d'adaptation*

- Partir avec un moteur de reconnaissance **simple**
  - Systèmes d'Inférence Floue (SIF) — Takagi-Sugeno d'ordre 0
    - SIF : Prototypes (Centre+Matrice de covariance) avec conclusions numériques
- Adaptation **incrémentale**
  - Inspirée des méthodes d'apprentissage incrémental
    - LVQ (Learning Vector Quantization) [Kohonen'90]
    - FLVQ (Fuzzy Learning Vector Quantization) [Chung et al.'94]



## 2<sup>ème</sup> partie | Reconnaissance de formes | Adapt Stratégie d'adaptation incrémentale

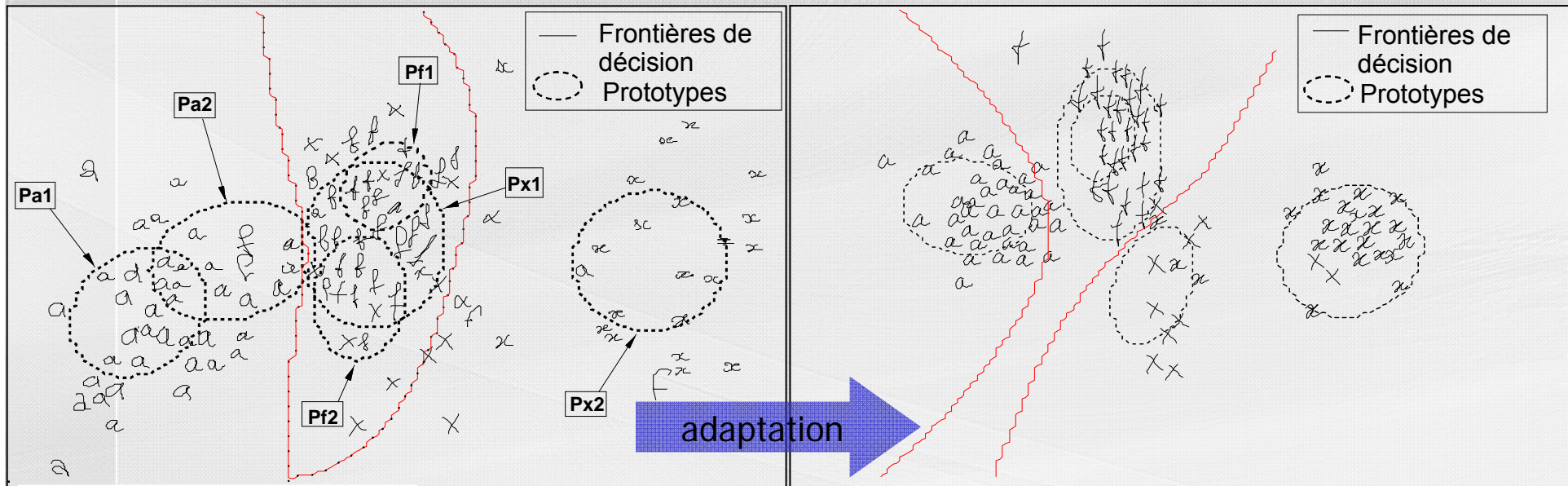
### Contributions

#### La méthode « Adapt » : ADaptation par Ajustement de ProTotypes

- Ré-estimer « à la volée » l'ensemble des paramètres
  - prendre en compte toutes les connaissances du SIF
  - garantir la stabilité et un pouvoir de généralisation du système

### Principe

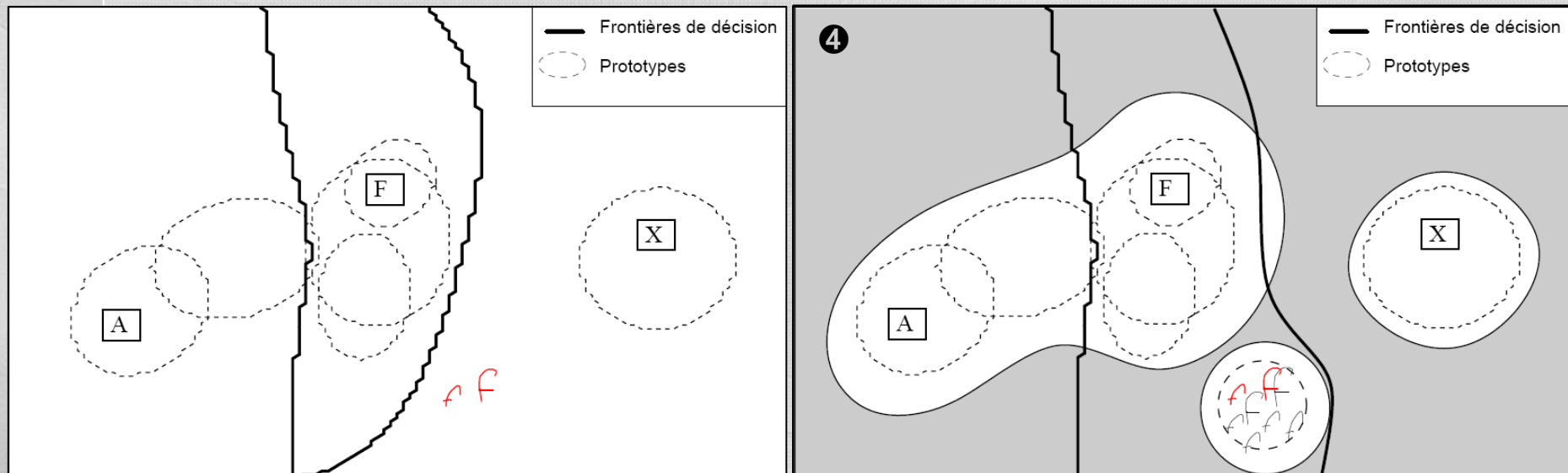
- Adaptation en optimisant conjointement
  - déplacement des centres des prototypes
  - déformation des prototypes
  - réévaluation des conclusions numériques





## 2<sup>ème</sup> partie | Reconnaissance de formes | Adapt Stratégie d'adaptation incrémentale

- **Création de nouvelles règles (nouveaux prototypes et conclusions)**
  - ❶ Identifier les nouveaux allographes (nouveaux styles d'écriture)
    - Rejet « d'ignorance »
  - ❷ Etendre la représentativité des nouveaux allographes
    - Faible nombre de données d'apprentissage
      - Création/Synthèse de caractères artificiels — respectant le style d'écriture
  - ❸ Générer le prototype
  - ❹ Mettre à jour les frontières de décision





## 2<sup>ème</sup> partie | Reconnaissance de formes | Adapt

### Quelques Résultats

#### → Apprentissage omni-scripteur sur la base IRONOFF

##### ■ *Expérimentation sur Benchmark*

- Test d'adaptation sur 12 scripteurs (Texte de 266 caractères)

doing so, the adaptation mechanisms presented here can be transposed to the original systems. in this first premise rule is composed of a fuzzy prototype which represents intrinsic properties of a class. The consequent part of rules associates a score to the prototype for each class.

- Taux de reconnaissance : 88.53%  
Après adaptation → 96.75% (réduction de 72% de l'erreur)

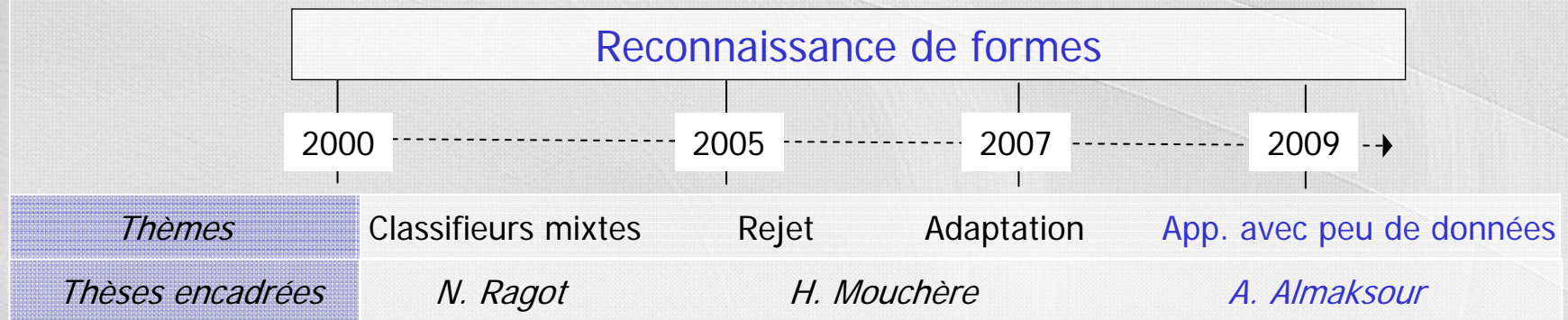
##### ■ *Cas d'utilisation réelle : assistant personnel de type PDA*



- Utilisateur impliqué dans la boucle d'adaptation
- Texte en anglais (1564 caractères)
- Contraintes d'embarquement
  - pas de déformation des prototypes, ni de synthèse de caractères
- Taux de reconnaissance : 88.6%  
Après adaptation → 93.3% (réduction de 41% de l'erreur)



### Bilan



### Collaborations

- Synthèse de caractères artificiels par analogie
  - Equipe Cordial de l'IRISA



### Perspectives

- Systèmes de reconnaissance « auto-évolutifs »
  - Apprentissage
    - « à la volée »
    - à partir de très peu de données réelles
  - Thèse d'A. Almaksour (→ fin 2009)

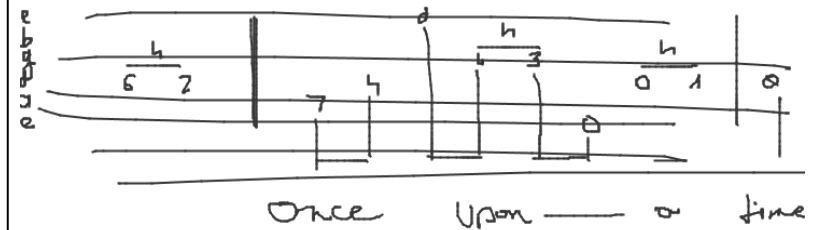
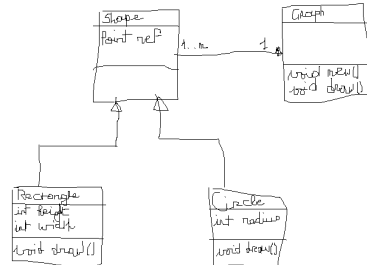
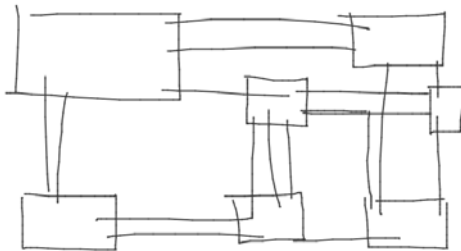


# 3<sup>ème</sup> partie Interaction homme document (IHD) orientée stylo

2004

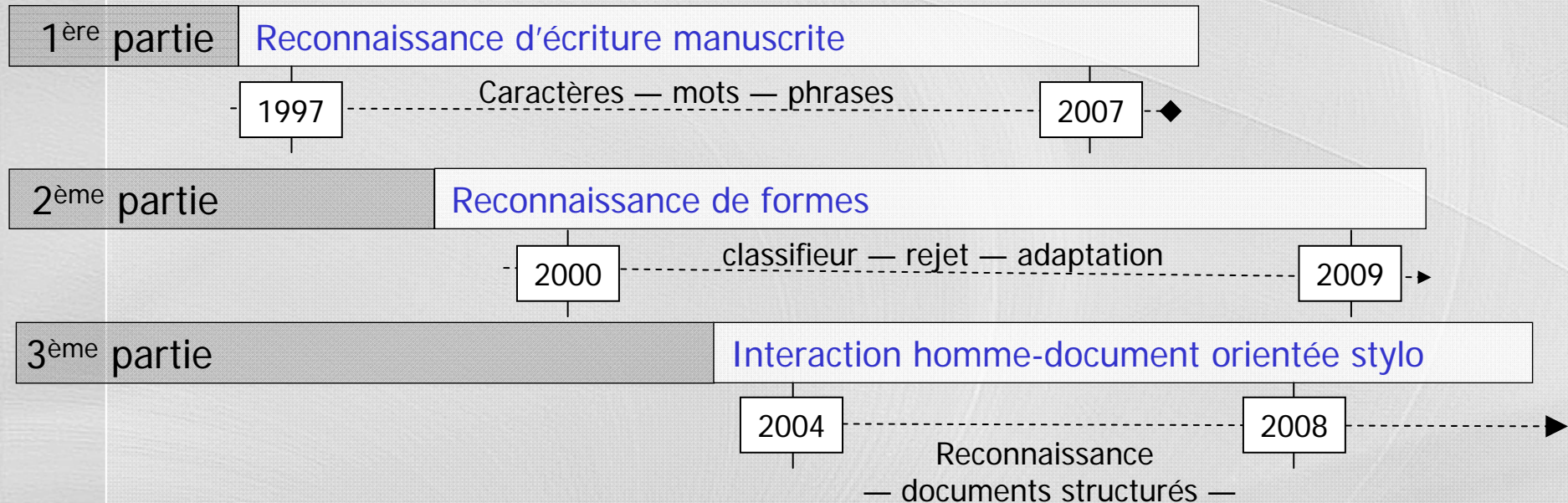
2008

Reconnaissance de documents  
structurés (en-ligne)





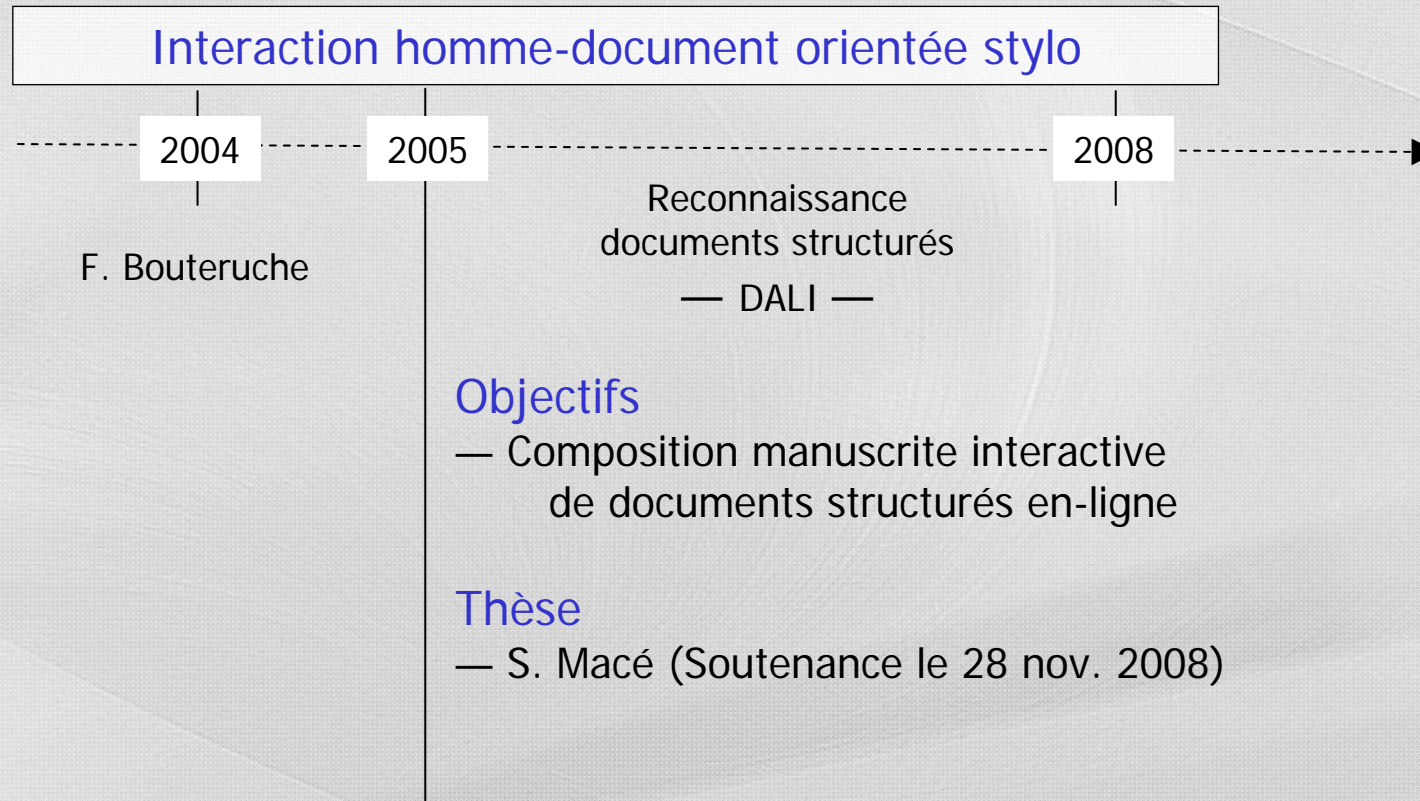
## 3<sup>ème</sup> partie | Interaction homme document (IHD) orientée stylo



- *2004 : Nouvel axe de recherche dans l'équipe IMADOC*
  - Reconnaissance de documents structurés en-ligne
  - Interaction homme-document
    - **placer l'utilisateur dans la boucle d'interprétation**



## 3<sup>ème</sup> partie | Interaction homme document (IHD) orientée stylo

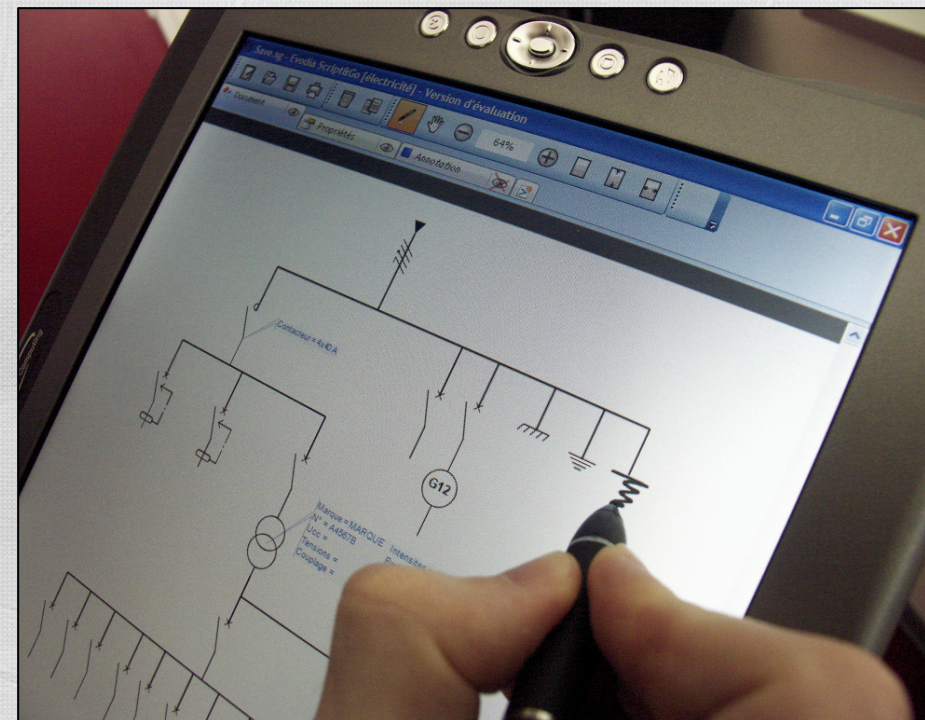
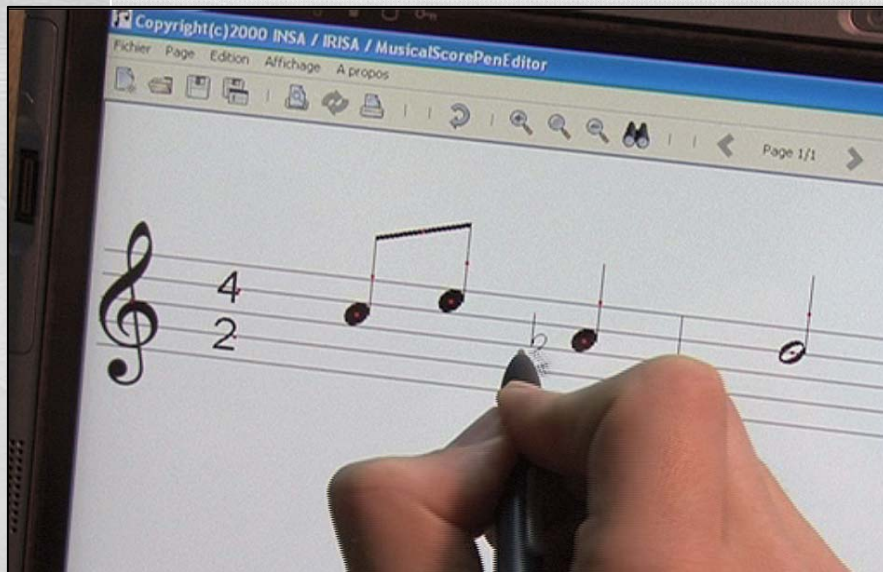




# 3<sup>ème</sup> partie | Interaction homme document (IHD) orientée stylo

## Introduction

- *Vision scientifique ambitieuse ...*
  - Reconnaissance de documents manuscrits structurés (en-ligne)
    - Approche générique
      - Reconnaissance de partitions musicales, de graphes, de schémas, ...
    - Avec de très fortes performances à moyen terme
      - Compatible avec les exigences du monde industriel



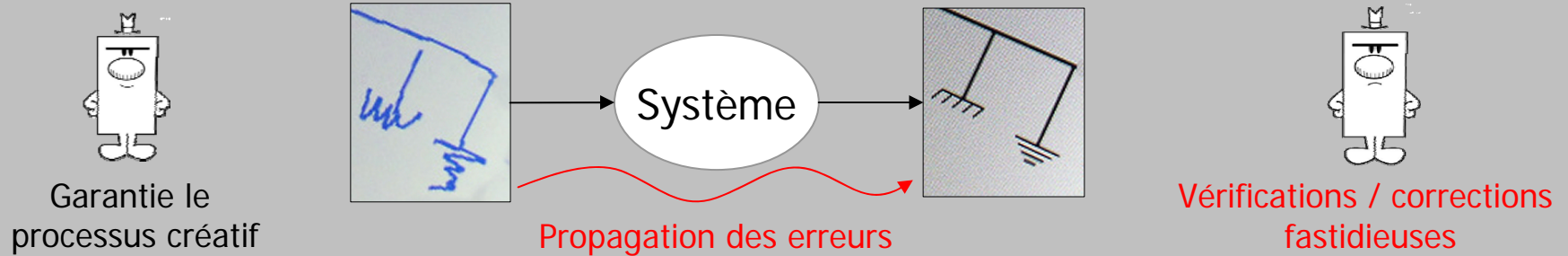


# 3<sup>ème</sup> partie | Reconnaissance de documents structurés | Dali

## Positionnement de l'approche

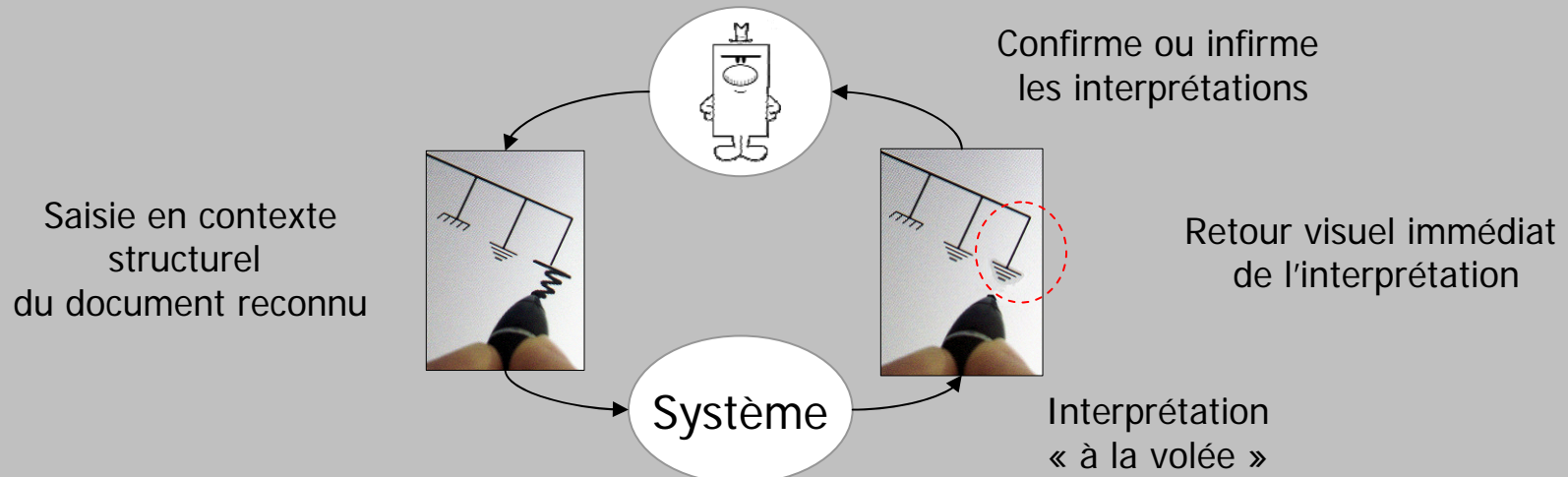
### ■ Approche « idéale » ?

- Interprétation « a posteriori » [Nakagawa'93, Gennari'05, Guihuan'06, Hall'07, ...]



### ■ Approche « pragmatique » !

- Reconnaissance « à la volée » [Alvarado'04, Hammond'03, Miyao'07]
- ➔ Originalité : L'utilisateur devient acteur du processus d'analyse





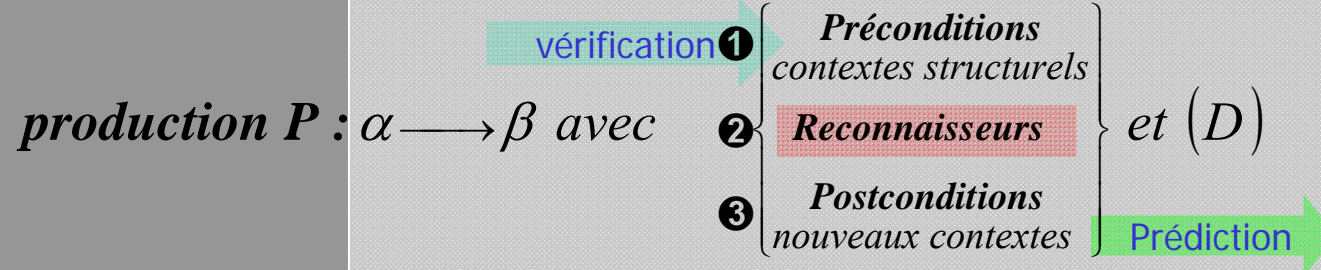
- *1<sup>er</sup> axe : Approche de modélisation générique*
  - Théorie des langages visuels
    - Extension des Grammaires de Multi-ensembles à Contraintes (GMC) [Marriott.'94]
  
- *2<sup>ème</sup> axe : Interprétation de formes manuscrites en contexte*
  - Reconnaissance hybride structurelle/statique des formes
    - Règles floues / SIF
  - Modéliser du contexte structurel
    - Relations spatiales floues [Bloch'99, Bouteruche'06]
  
- *3<sup>ème</sup> axe : Interaction homme-document*
  - Rôle de l'utilisateur (Acteur) dans le processus d'interprétation
    - Pertinence des retours visuels
    - Mise en place d'un mécanisme de rejet (ignorance et d'ambiguïté)



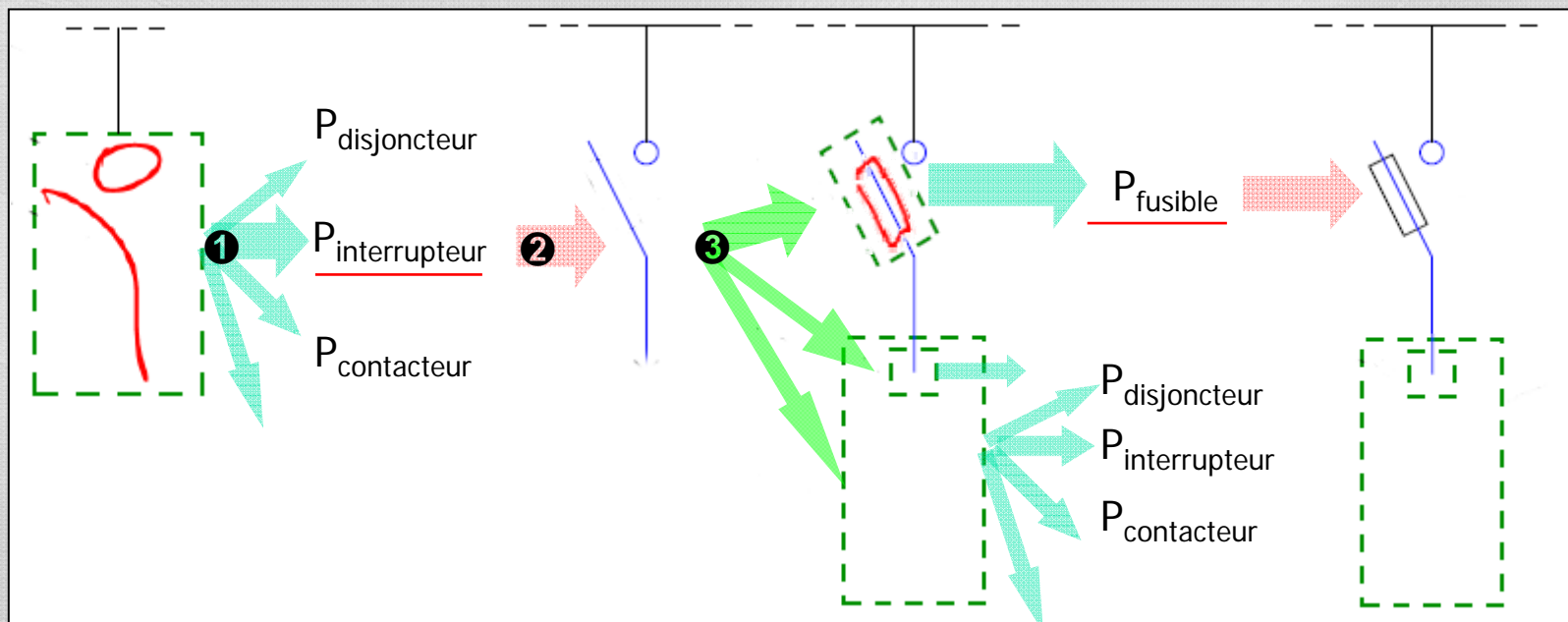
# 3<sup>ème</sup> partie | Reconnaissance de documents structurés | Dali

## Approche DALI : théorie des langages visuels

- Extension des Grammaires de Multi-ensembles à Contraintes (GMC)  
→ GMC pilotées par le contexte structurel (GMC-PC)

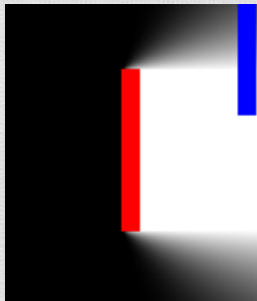


→ Déclencher uniquement les productions contextuellement cohérentes

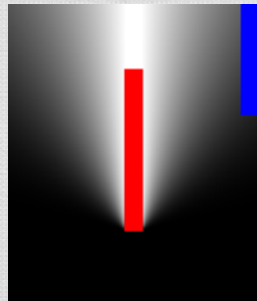




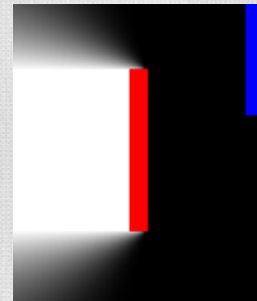
- Représentation par des régions floues [Bloch'05]



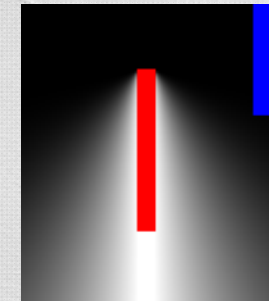
$$M_{\alpha}^{droite}(A) = 0.81$$



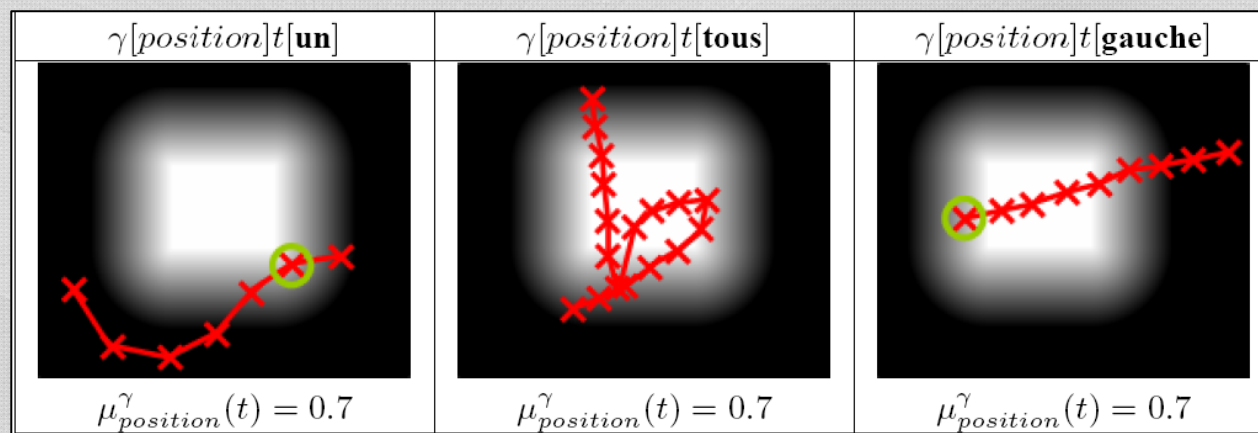
$$M_{\alpha}^{dessus}(A) = 0.66$$



$$M_{\alpha}^{gauche}(A) = 0.00$$



$$M_{\alpha}^{dessous}(A) = 0.04$$





## 3<sup>ème</sup> partie | Reconnaissance de documents structurés | Dali

### Processus de décision

#### ■ *Processus de décision : basé sur la théorie des sous-ensembles flous*

##### ■ Grammaires floues

$$\textit{production } P : \alpha \xrightarrow{\mu} \beta \textit{ avec } \left. \begin{array}{l} \textit{Préconditi ons} \\ \textit{contextes structurels} \\ \textit{Reconnaiss eurs} \\ \textit{Postconditi ons} \\ \textit{nouveaux c ontextes} \end{array} \right\} \textit{ et } (D)$$

##### ■ Évaluation d'une production

$$\mu_p = \mu_{\textit{contexte}} \bullet \mu_{\textit{reconnaisseur}}$$

##### ■ Évaluation d'une séquence de productions

$$\mu_{ps} = \left( \prod_{p_i \in ps} \mu_{p_i} \right)^{\frac{1}{|ps|}}$$

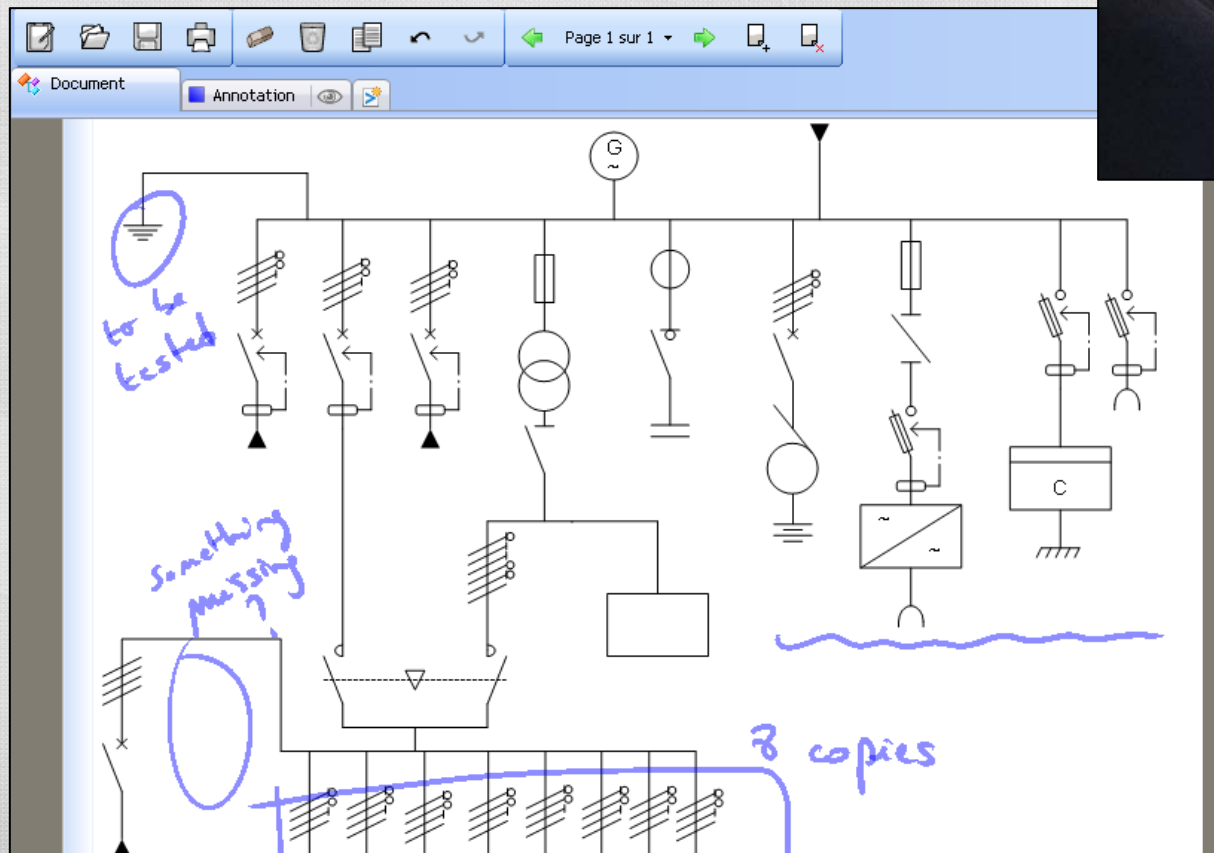


# 3<sup>ème</sup> partie | Reconnaissance de documents structurés | Dali

## Prototypage

### ■ Prototypage sur différents types de documents

- Diagrammes UML, Graphes, ...
- Partitions musicales
- Schémas électriques





■ *Taux de reconnaissance*

- 22 utilisateurs (> 8712 symboles)

➔ Taux d'erreurs moyen = 3.83 %      Taux de rejet = 3.93 %

■ *Temps de calculs*

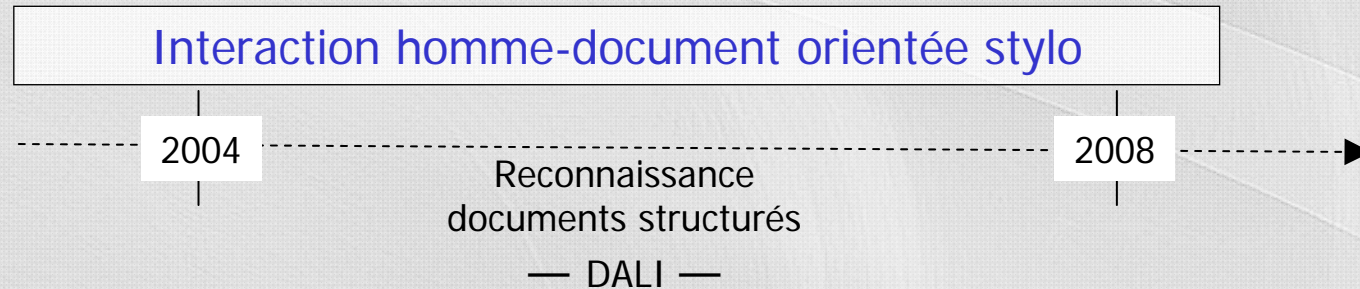
- Comportement du système avec des documents complexes (nb de symboles)
- Temps moyen d'analyse = 128 ms

■ *Ergonomie*

- Comparaison des temps de saisie entre
  - composition « orientée stylo » (DALI)
  - composition par « glisser-déposer » (Boutons / Wimp)

➔ Gain de temps moyen = 32.5 %





### Contributions scientifiques

#### *Reconnaissance « à la volée » de documents structurés*

- Extension des GMC — pilotées par le contexte
- Analyseur incrémental — utilisateur est acteur de l'analyse
- Modélisation homogène — théorie des sous-ensembles flous

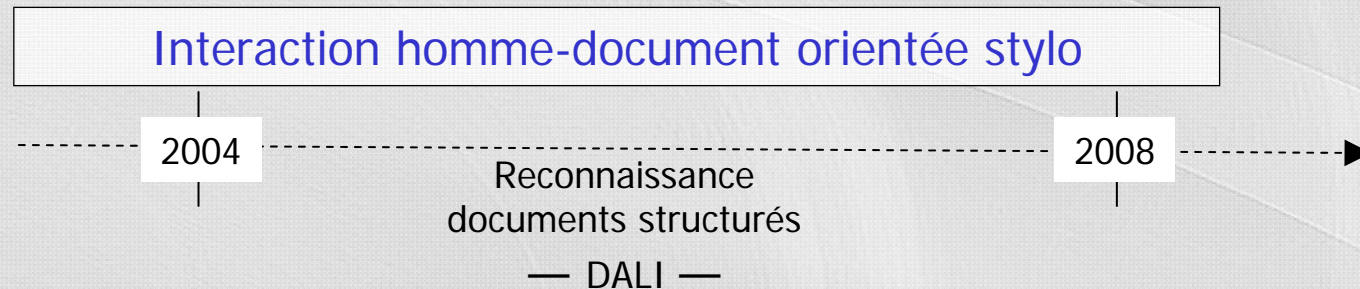


### Collaboration & Transfert industriel

- Editeur de partitions musicales
  - Laboratoire MIAC de musique de l'Univ. Rennes 2
- Editeur de schémas électriques
  - Transfert technologique avec la société Evodia — Apave
  - Commercialisation du logiciel « Script&Go schémas électriques »







■ ***Etudier d'autres natures de documents***

- Reconnaissance incrémentale de caractères chinois
  - Nouvelle thèse d'Adrien Delaye
  - Collaboration avec laboratoire Franco-Chinois de Pékin (LIAMA)
- Projet du pôle « images et réseaux » — ScriptEveryWhere (2 ans)
  - Partenaires : Evodia, APAVE Nord Ouest, l'IRISA
  - Financement : Région Bretagne - Oséo Anvar



■ ***Lever certaines contraintes qui persistent***

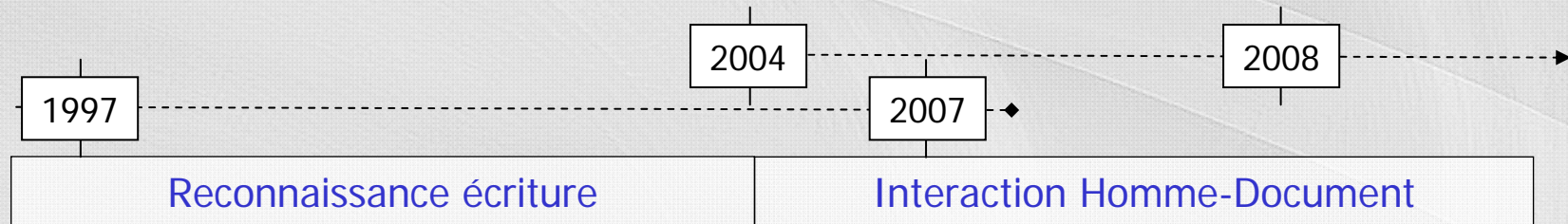
- l'utilisateur doit dessiner l'ensemble des tracés d'un symbole avant de passer au suivant ...

■ ***Apprentissage à partir d'exemples***

- Description d'un document
- Positionnement relatif flou des éléments



# Conclusions



## *Domaines*

Reconnaissance de formes — Ecrit et du document — Interaction homme-machine

## *Thèmes*

Caractères — Mots — Phrases — Gestes graphiques — Documents structurés

## *Approches*

Logique floue — formalisation explicite — approches génériques — utilisateur/acteur

## *Thèses encadrées*

6 thèses + 3 en cours

## *Collaborations*

CRPCC, MIAC — Univ. de rennes 2 — LIAMA Chine — Synchronédia Canada, ...

## *Publications*

6 revues (2 int., 4 nat.) — 5 chap. de livre — 37 conf. int.— 22 conf. nat.

## *Transferts*

France Telecom — PurpleLabs — Thales — Apave

## *Essaimage*

Création de la société Evodia en Octobre 2005

## *Enseignement*

Insa de Rennes



# Perspectives

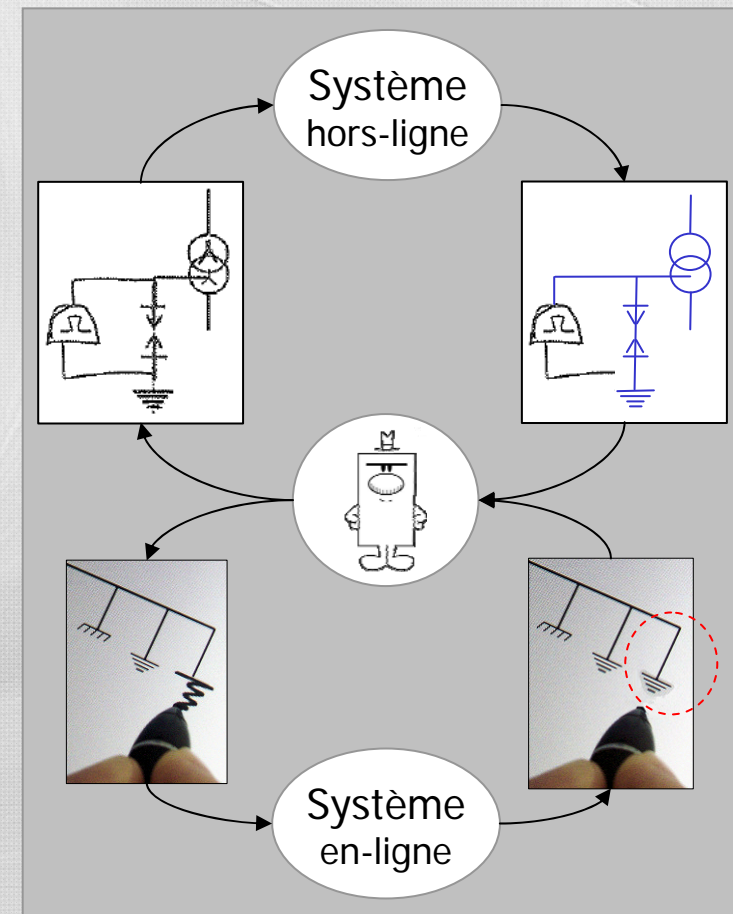
## Interaction homme-document

### A court terme

- **Composition à main levée de documents**
  - ➔ Interaction & Interprétation « à la volée » du document
- **Collaborations inter-domaines**
  - RdF — IHM — Psychologie expérimentale
  - ➔ Créer de nouveaux usages — orientés stylo
- **Retours d'expériences → défis**
  - ➔ **Fiabiliser** les réponses du système
    - Erreurs cohérentes
  - ➔ Rendre les systèmes **auto-évolutifs**
    - Apprentissage avec très peu de données.

### A Moyen Terme

- **Vision homogène en-ligne et hors-ligne**
  - Fédérer l'ensemble des activités de l'équipe Imadoc





## Remerciements

A toutes les personnes qui ont participé à ces travaux

- L. Pasquer
- N. Ragot
- S. Carbonnel
- L. Rousseau
- S. Quiniou
- H. Mouchère
- F. Bouteruche
- S. Macé
- A. Almaksour
- A. Delay
- F. Boyer
- N. Briec
- O. Combot,
- S. Decaix
- B. Dien
- N. Fortune,
- E. Garrivier
- D. Lamarche
- S. Ollivier
- B. Morin
- Y. Peng
- N. Socheleau
- B. Bossis
- H. Bouchereau,
- J. Camillerapp
- B. Couasnon,
- G. Deconde
- G. Gravier
- E. Jamet,
- E. Petit
- F. Ric
- P. Sébillot
- S. Vidal
- ...
- Guy Lorette



# Reconnaissance d'écriture manuscrite et Interaction homme-document

Eric Anquetil

Équipe IMADOC  
IRISA/INSA

Soutenance d'HDR — 27 octobre 2008