User Linguistic Model Adaptivity for Prediction in AAC Message Composition

HCII 2005

Nicola Gatti, Matteo Matteucci, and Licia Sbattella

Dipartimento di Elettronica e Informazione

Politecnico di Milano

Presentation Outline

- Application scenario
- Introduction to symbolic prediction
- CABA^{2}L
- Discrete Auto-Regressive Hidden Markov Models
- Experimental evaluation of CABA^{2}L
- Conclusion and future works

Application scenario

Verbal Impairments

- Verbal impairments represent a social problem:
  - Social exclusion
  - Seclusion

Verbal Impairments

- Verbal impairments represent a social problem:
  - Social exclusion
  - Seclusion
- Millions of verbal impaired people live in the world (e.g., more than 2 millions in USA) suffering:
  - Autism
  - Dysphasia
  - Intellectual impairments
  - Motor impairments

ISAAC

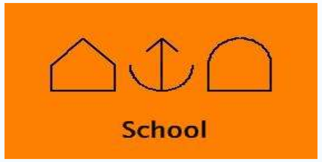
- International Society for Augmentative and Alternative Communication (1983, USA)
- Objectives:
  - Development of languages usable by verbal impaired people (AAC languages)
  - Development of communicative aids (AAC aids)

ISAAC

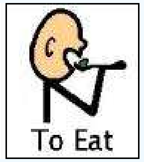
- International Society for Augmentative and Alternative Communication (1983, USA)
- Objectives:
  - Development of languages usable by verbal impaired people (AAC languages)
  - Development of communicative aids (AAC aids)
- Why *Augmentative*:
  - Non verbal communication modes do not necessarily substitute the natural language
  - They enhance residual capabilities
- Why *Alternative*:
  - Codes alternative to speech (figures, drawings, symbols, etc.)

AAC Language Examples

Bliss

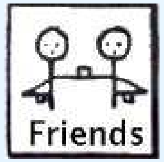


PCS



PIC

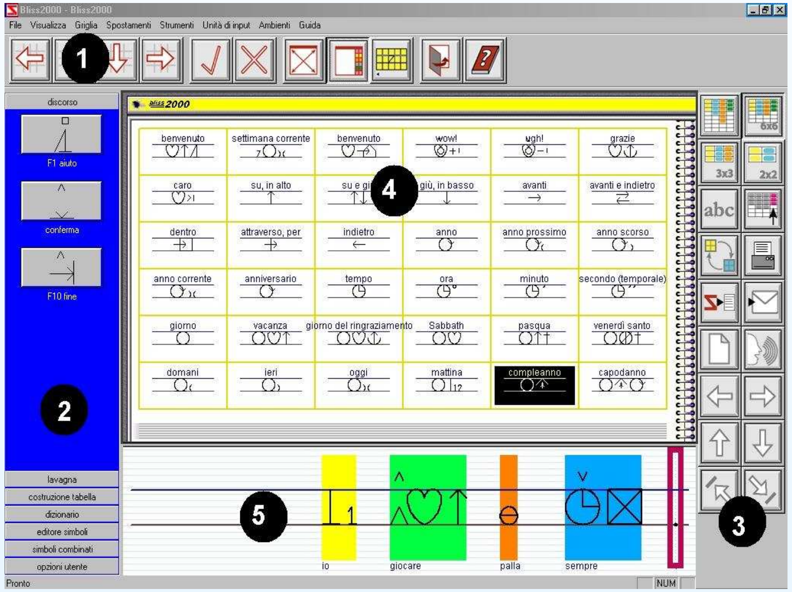
PICSYM



AAC Software Aids

Communication Software

- Provide an electronic version of AAC tables
- Provide a number of device to select symbols
- Aid the verbal impaired to compose messages
- Provide extra features:
  - Vocal synthesis
  - Message exchangement services
  - etc.



(example of communication software)

Motor Disorders and Message Composition

60% of verbal impaired people suffer motor disorders

- They have difficulties to compose messages (the composition of a single message can take several minutes)
- They need particular input devices
- They need an automatic scansion of the symbol table
- They need special human-computer interfaces

(examples ad-hoc input devices)

Automatic Symbol Scansion (1)

An highlight moves autonomously on an AAC symbol table according to a strategy

Linear sequentially (it is not useful with a high number of symbols)

Row-Column at first rows are scanned, then columns, or vice versa (it is not useful with a high number of symbols)

At subgroups in groups fewer and fewer (it is not adopted by people suffering mental impairments)

Predictive from the most probable the user will use to continue the sentence according to a linguistic model of the user (currently lacking in literature)

Automatic Symbol Scansion (2)

- Non predictive automatic scansion strategies are not effective enough (time spent to compose sentences too long)
- Intelligent scansion strategies are required:

Predictive able to predict the linguistic behavior of the user embedding the peculiar user linguistic model

Adaptive able to adapt itself on the peculiar user intellectual and motor capabilities

Introduction to symbolic prediction

Symbolic Prediction

- In literature mainly alphabetical prediction
- Alphabetical prediction techniques do not suite with symbolic prediction issue

Main Differences

Alphabetical prediction

- Strict set of items (22 signs)
- Organized in sequence of signs (words) known *a priori*

Symbolic prediction

- Variable set of items (from 50 to 300 symbols and over)
- Variable sequences of symbols

Traditional Predictive Techniques and AAC

Statistical do not take into account the peculiar AAC language

Syntactic and strong syntactic do not take into account the peculiar AAC language

Semantic do not take into account the variability related to the residual capabilities

Hybrid …

CABA^{2}L

Composition Assistant for Bliss Augmentative and Alternative Language

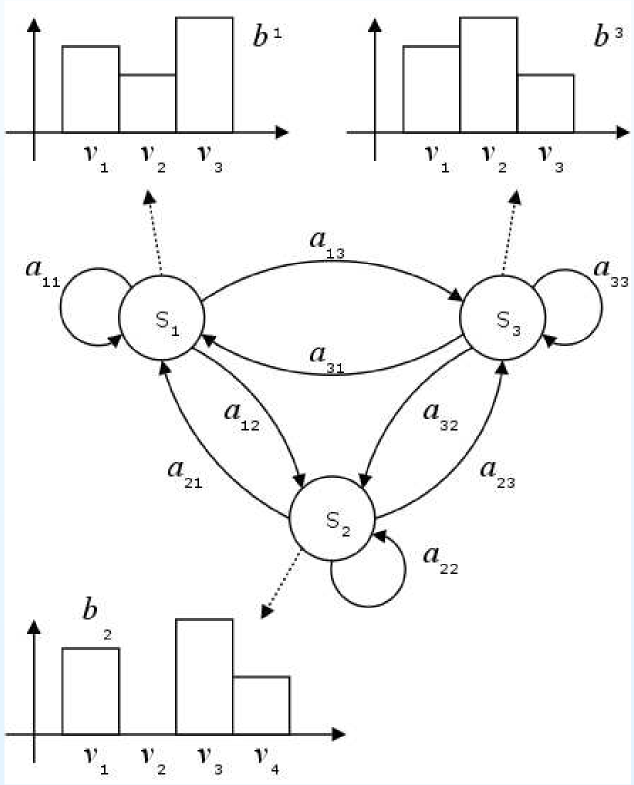
- Hybrid predictor (semantic-statistic) that takes into account:
  - Language peculiarities (linguistic and syntactic categories)
  - User linguistic behavior (probability that a symbol is selected given the last selected symbol and the related linguistic category)
- Adaptive scansion rate of the n most probable symbols

Hybrid Predictor

- Semantic network formalism to identify linguistic categories:
  - 6 syntactic categories (verbs, adverbs, adjectives, substantives, people, punctuation)
  - 30 semantic subcategories
- Identification process accomplished in collaboration with experts in verbal rehabilitation
- Note: a symbol could belong more than one category

Hidden Markov Models

- Variation of a classical Hidden Markov Models (HMM)
- Example of classical HMMs:
  - States: S_1, S_2, S_3
  - Symbols: $\nu_1, \nu_2, \nu_3, \nu_4$
  - Transition probability: $a_{11}, a_{12}, a_{21}, \ldots$
  - Emission probability: $b_{11}, b_{12}, b_{21}, \ldots$



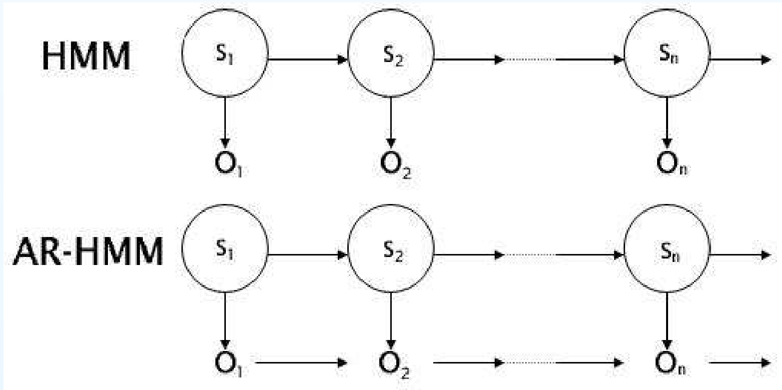
Hidden Markov Models and Symbols

Classical HMMs do not map completely symbolic prediction issue:

- Traditional HMMs allow the description of:
  - States (linguistic subcategories)
  - Symbols
  - Emission probabilities of a particular symbol in a particular state
  - Transition probabilities between states
- Traditional HMMs do not allow the description of:
  - Emission probabilities of a particular symbol conditioned to the last selected symbol

HMM vs. AR-HMM

- Regression takes into account previous emissions: Auto-Regressive Hidden Markov Models



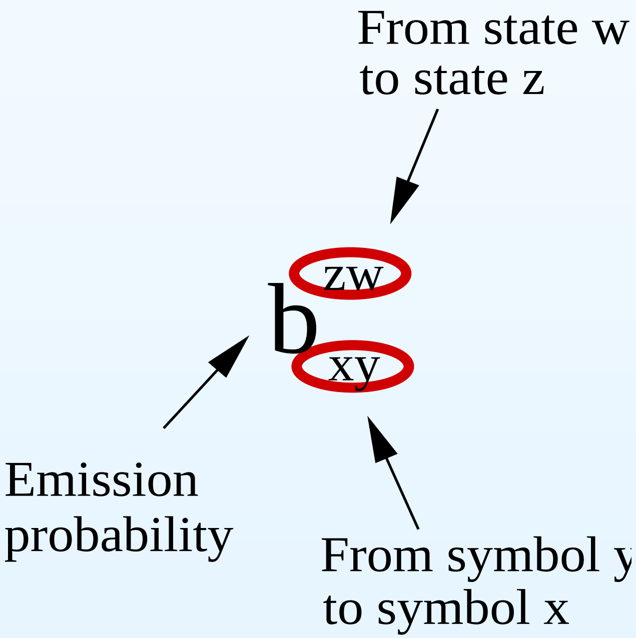
(comparison between HMM and AR-HMM)

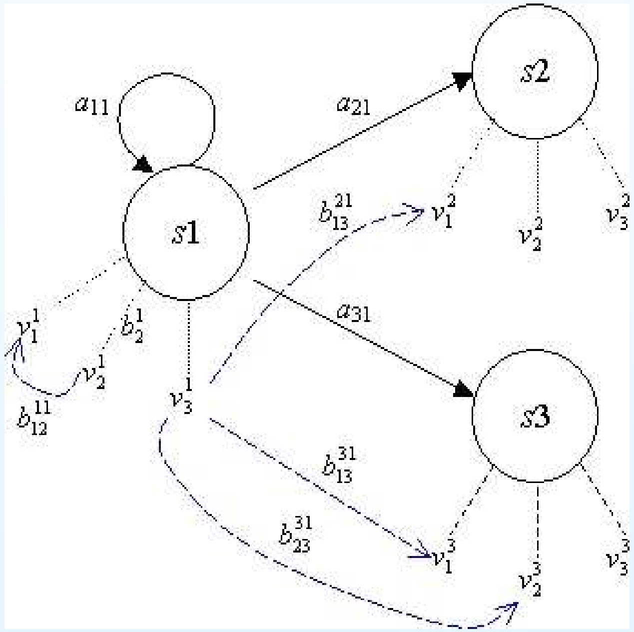
- AR-HMMs are adopted in continuous systems

DAR-HMM

DAR-HMM

Symbol emission is a discrete event, a discretization of AR-HMM is required: Discrete Auto-Regressive Hidden Markov Models





DAR-HMM Parameters

- Vector of parameters: $\langle \Pi^0, A, B \rangle$
- Vector of initial subcategory probabilities: $\Pi^0 = (\pi_z(0))$
- Matrix with subcategory transition probabilities: $A[N][N] = (a_{zw})$
- Emission matrix: $B[N][M][M+1] = \left(b_{xy}^{zw}\right)$

Emission rules (1)

First observed symbol:

$$
\begin{aligned}
\hat{O}(0) &= \arg\max_{v_w^{(x)}} \left(P(O(0) = v_w^{(x)} | \lambda) \right) \\
&= \arg\max_{v_w^{(x)}} \left(P \left(O \left(0 \right) | Q \left(0 \right), \lambda \right) P(Q(0)) \right) \\
&= \arg\max_{v_w^{(x)}} \left(b_w^x \cdot \pi_x(0) \right)
\end{aligned}
$$

Emission rules (2)

Recalling that we can compute the probability of the current (hidden) state as:

$$
\begin{aligned}
P(Q(t)) &= \sum^{N} P\left(Q(t)|Q(t-1)\right) P\left(Q(t-1)\right) = \\
&= \sum_{y=1}^{N} \pi_y(t-1) a_{xy} = \pi_y(t)
\end{aligned}
$$

we obtain a recursive form for symbol prediction at time t:

$$
\hat{O}(t) = \arg\max_{v_z^{(x)}} \left(b_{zw}^{xy} \cdot \sum_{y=1}^{N} \pi_y(t-1) a_{xy} \right)
$$

Parameter estimation (1)

- Baum-Welch algorithm adapted for DAR-HMM adopting sentence dataset
- Initialization
  - Uniform distribution for Π^0
  - In A^0 the arcs between symbols and states that are not connected in semantic network have a very low probability
  - B^0 is more critical (*Segmental k-Means* and *Viterbi*)

Parameter estimation (2)

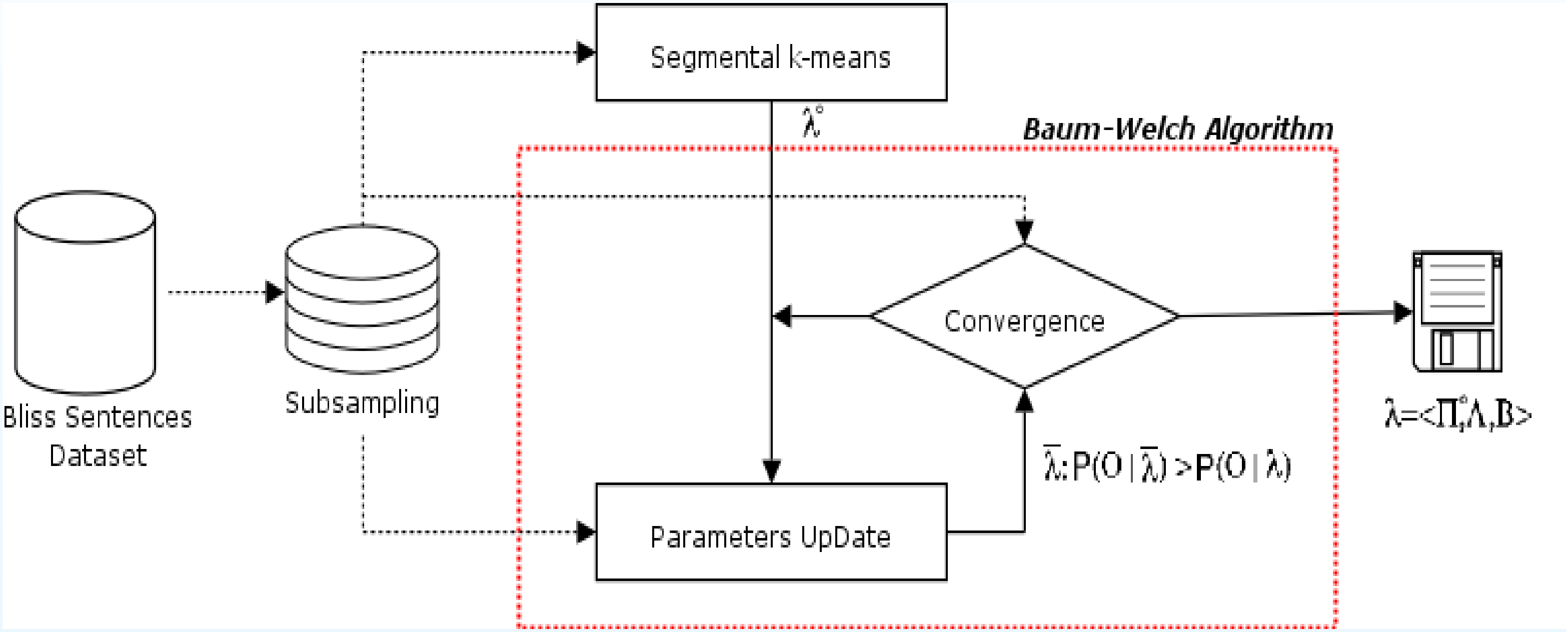
- *Baum-Welch*
- Stopping criterion based on *generalization error* adopting *K-fold cross validation* technique
  - Generalization loss:

$$Err_{Opt}(t) = \min_{t' \leq t} Err_{Val}(t')$$

$$GL(t) \triangleq 100\left(\frac{Err_{Val}(t)}{Err_{Opt}(t)} - 1\right)$$

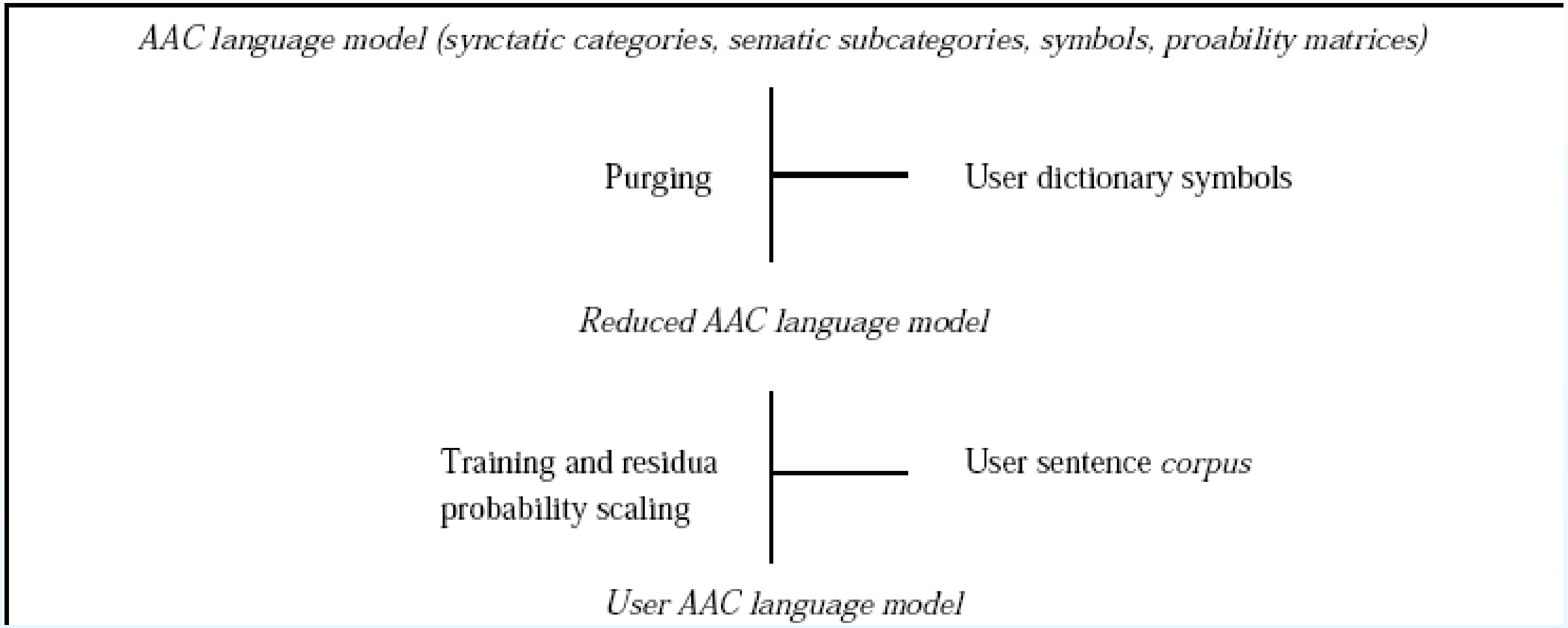
$$GL(t) > \tau$$

Parameter estimation (3)



(synthesis of the training process)

Language model production



(synthesis of the language model production procedure)

Experimental evaluation

Evaluation Indexes

Indexes:

- *Training error*
  - Express the effectiveness of the learning
  - Obtained comparing the suggestions of CABA^{2}L during the composition of sentences belonging the training dataset
- *Generalization error*
  - Express the effectiveness of the prediction system
  - Obtained comparing the suggestions of CABA^2 during the composition of sentences not presented during the training phase

Case Study

1. Dataset of 20 sentences with 4 sub-categories and 7 symbols (person unskilled in Bliss utilization)
2. Dataset of 80 sentences with 18 sub-categories and 120 symbols (person skilled in Bliss utilization without mental deficiency)

Evaluation Results (Training Error)

Probability that the requested symbol is in the first four predicted symbols according to the datasets adopted to train the DAR-HMM

Predictions	Scenario 1		Scenario 2	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
1 symbol	0.343	0.055	0.250	0.017
2 symbols	0.563	0.074	0.299	0.028
3 symbols	0.778	0.067	0.319	0.033
4 symbols	0.908	0.056	0.345	0.042
not suggested	0.092	0.056	0.655	0.042

Evaluation Results (Generalization Error)

Probability that the requested symbol is in the first four predicted symbols according to the datasets not adopted to train the DAR-HMM

Predictions	**Scenario 1**		**Scenario 2**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
1 symbol	0.202	0.082	0.185	0.089
2 symbols	0.438	0.146	0.252	0.073
3 symbols	0.666	0.181	0.304	0.070
4 symbols	0.887	0.067	0.357	0.077
not suggested	0.113	0.067	0.643	0.077

Evaluation Results (Other Issues)

- Sentence composition time reduction: 60%
- Training required time: a few minutes according to dataset size (in our case study: (1) less a minute, (2) between 10 and 15 minutes)
- Prediction required time: $<$ 1s (real time)

Conclusions

- Analysis of the AAC table symbol scansion issue
- Design of symbolic prediction model (semantic-statistical)
- Design of DAR-HMM: formalism, ad-hoc emission rules, training, etc.
- Experimental evaluation

Future works

- Refining the prediction model introducing:
  - Specific syntactic analysis of symbolic languages (AAC languages)
  - On-line adaption to the linguistic behavior of the user and the related evolution
- Analysis of the semantic/probabilistic model to study relationships between disabilities and verbal impairments