

Tecnologia Adaptativa na Linguística Computacional

João José Neto

Escola Politécnica da USP

Laboratório de Linguagens e Técnicas Adaptativas

INTRODUÇÃO

Apresentação

- O LTA (**Lab. de Linguagens e Técnicas Adaptativas da EPUSP**) tem desenvolvido pesquisa teórica e aplicada na área da **Adaptatividade** e assuntos correlatos desde a sua criação, em 1985.
- Um de seus principais focos de interesse tem sido o uso da adaptatividade no processamento computacional de linguagens formais, de linguagens de programação, e de línguas naturais

- Dentre as mais importantes aplicações dessa natureza podem ser destacadas:
 - representação e manipulação do conhecimento;
 - representação e processamento de línguas naturais;
 - tradução automática de idiomas;
 - inferência gramatical;
 - sistemas inteligentes de tomada de decisão;
 - busca semântica de informação.

HISTÓRICO

Histórico

- **Programas antigos** já alteravam seu próprio código para economizar memória
- **Processadores microprogramáveis** também permitiam modificar seu próprio repertório de instruções
- Nas décadas de 60 e 70 surgiram **linguagens de programação extensíveis** e seus formalismos correlatos
- Na década de 90 publicaram-se os primeiros trabalhos sobre **técnicas genuinamente adaptativas**
- Correntemente, têm sido desenvolvidas técnicas adaptativas de diversas naturezas, envolvendo inclusive aspectos da **adaptatividade de hardware**.

Metas gerais do LTA em relação à Tecnologia Adaptativa

- Desenvolvimento de **teorias e de técnicas de ampla aplicabilidade**, baseadas em uma **fundamentação científica sólida**
- Desenvolvimento de formulações que propiciem a obtenção de **soluções simples para problemas difíceis**
- **Unificação e extensão** de diversas notações formais, formulações teóricas e métodos

Metas específicas para a área de Inteligência Artificial

- Fundamentação de um **modelo consistente e abrangente** para uso em *Inteligência Artificial*:
 - auxílio à tomada de **decisão** e **planejamento**
 - **diagnóstico** automático
 - busca, coleta e **inferência de leis de formação**
 - identificação e classificação de **padrões**
 - **aprendizagem**, treinamento, criatividade
 - **lazer**, jogos, arte
 - sistemas de **apoio a portadores de deficiências**
 - representação e manipulação de **conhecimento**
 - processamento de **linguagem natural**

Metas específicas para Linguagens de Programação

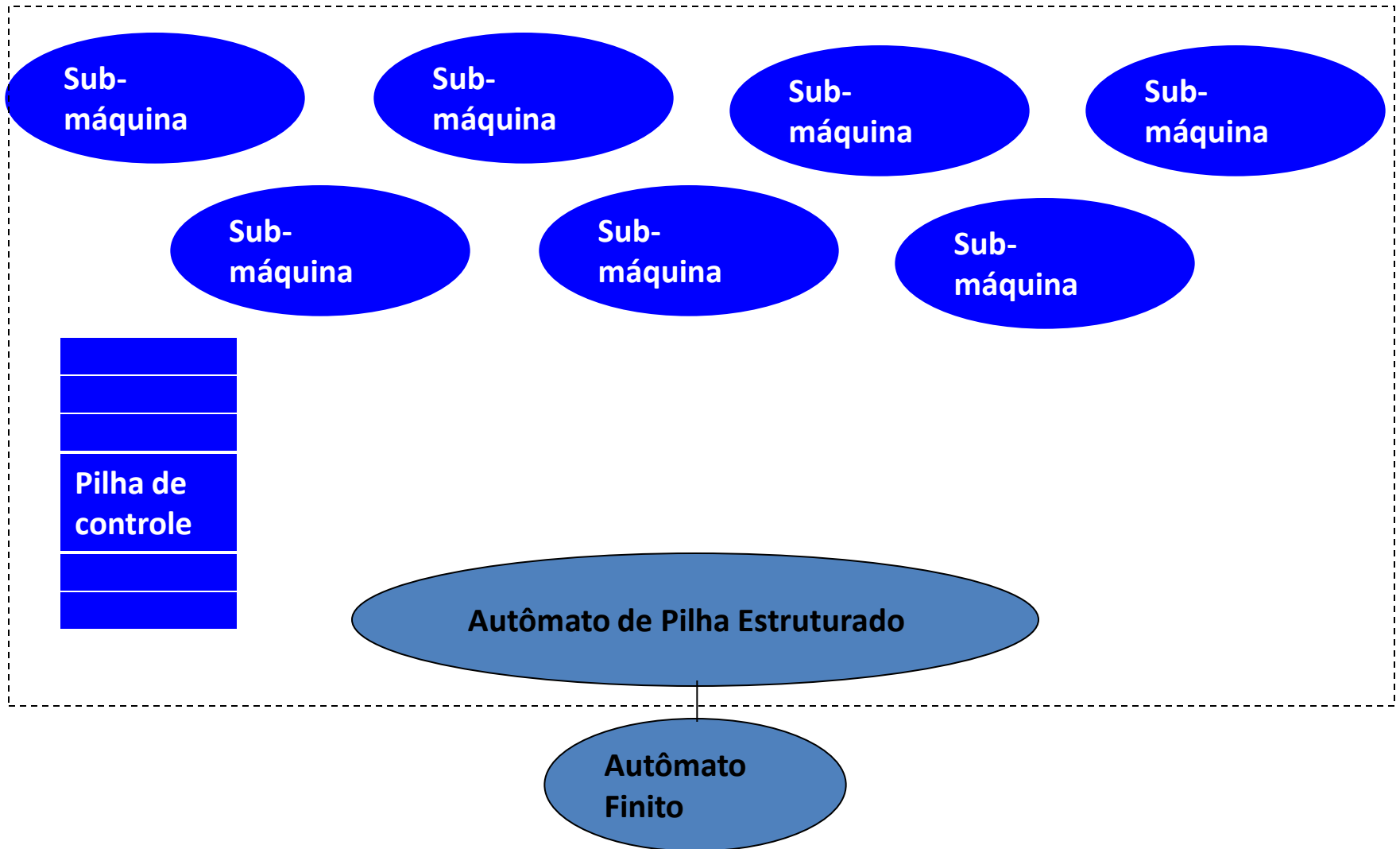
Redução à **sintaxe** do maior número possível de componentes de uma *linguagem de programação*:

- **Morfologia** (análise léxica)
- **Sintaxe** (reconhecimento e análise sintática)
- **Dependências de contexto** (tipos, escopos, macros, ...)
- **Semântica** (an.Semântica, ger.Código, Interpretação, ...)
- **Comunicação** (diálogos e protocolos com: usuário, sistema operacional, programas, computadores, ...)
- **Pragmática** (amb. de execução e de desenvolvimento; materialização do modelo de computação)

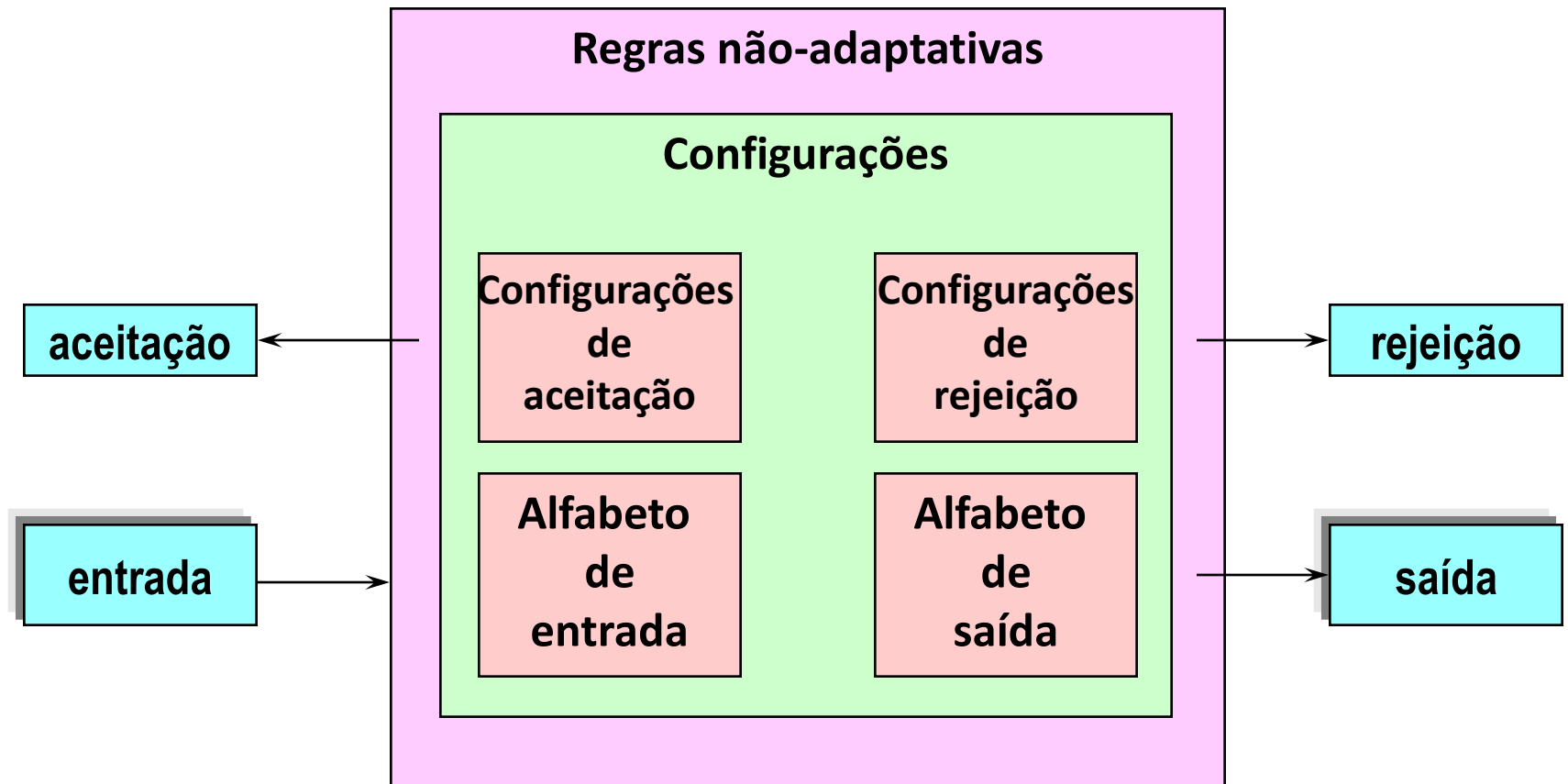
Cronologia da Evolução da Tecnologia Adaptativa no LTA

- **Fase 1** – Pré-História (antes de 1990)
(Não-adaptativa)
- **Fase 2** – Formulações Independentes (1990 a 2000) (Adaptatividade simples)
- **Fase 3** – Formulações Unificadas (desde 2001)
(Adaptatividade simples)
- **Fase 4** – Formulação Hierárquica (desde 2005)
(Adaptatividade múltipla, Modularidade)
- **Fase 5** – Formulação Paralela (em breve)
(Colaboração, Concorrência, Paralelismo, Agentes)

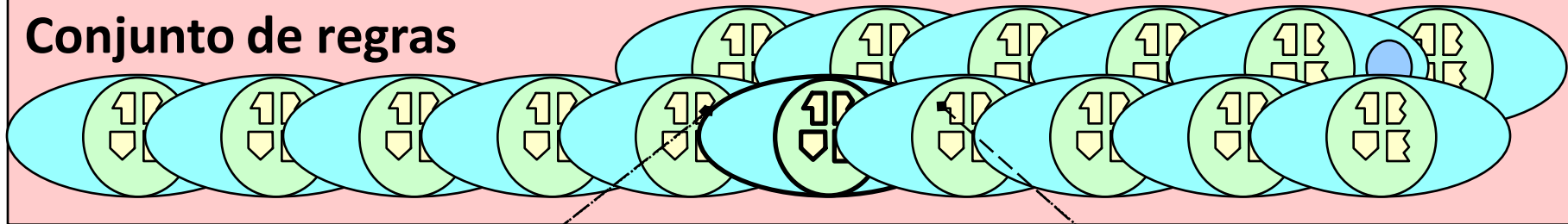
Fase 1 – Pré-História (não-adaptativa) [antes de 1990]



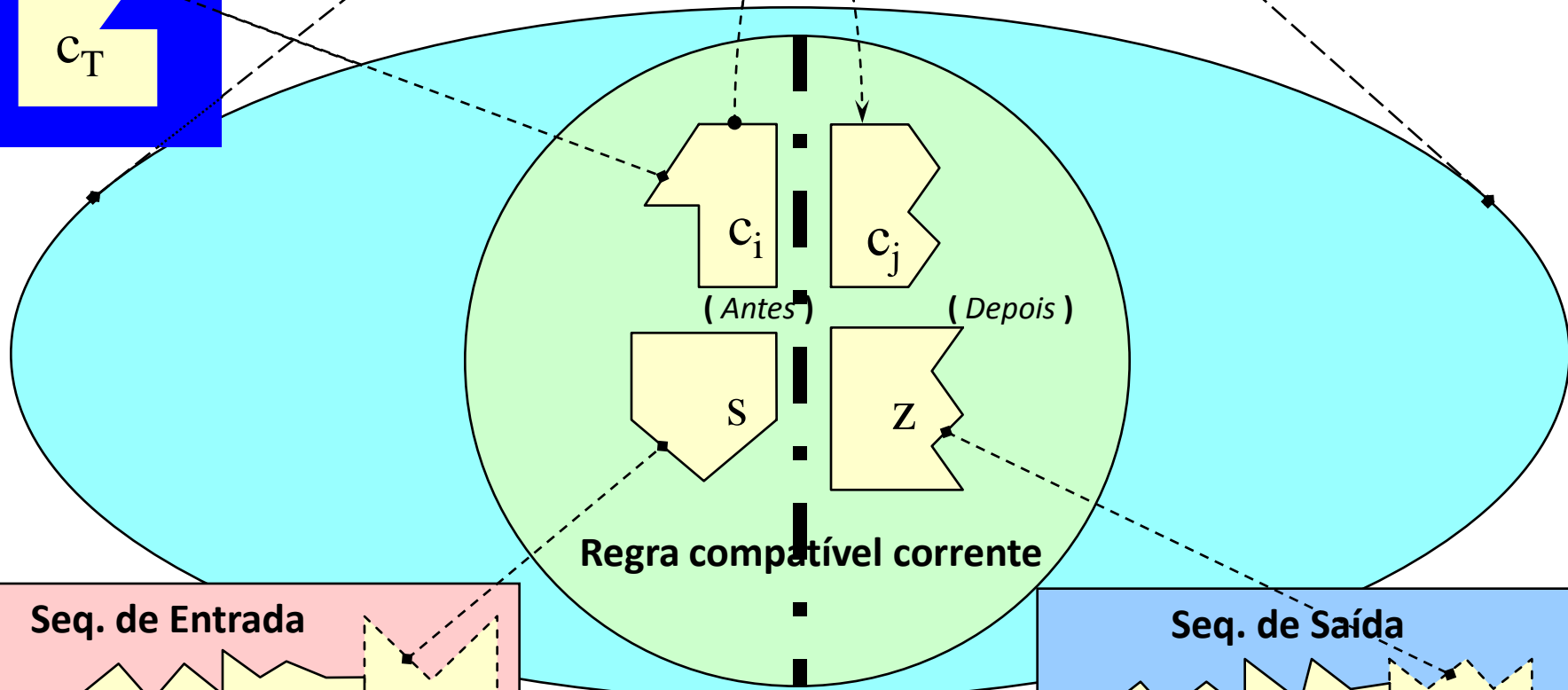
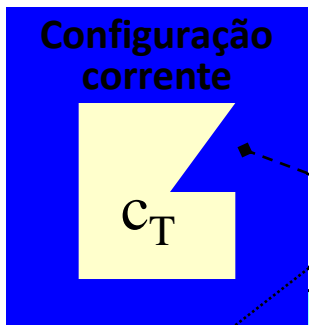
Dispositivo não-adaptativo



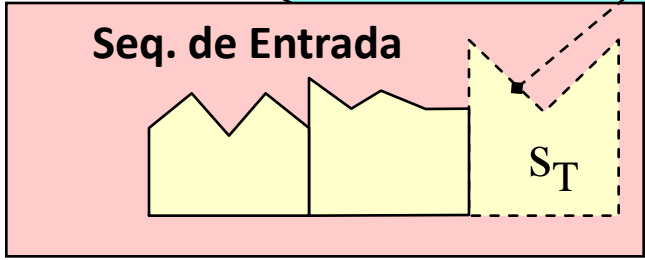
Conjunto de regras



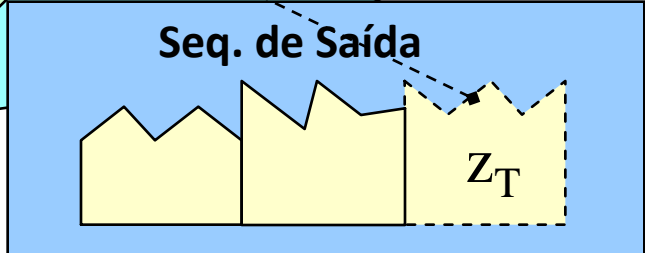
Configuração corrente



Seq. de Entrada

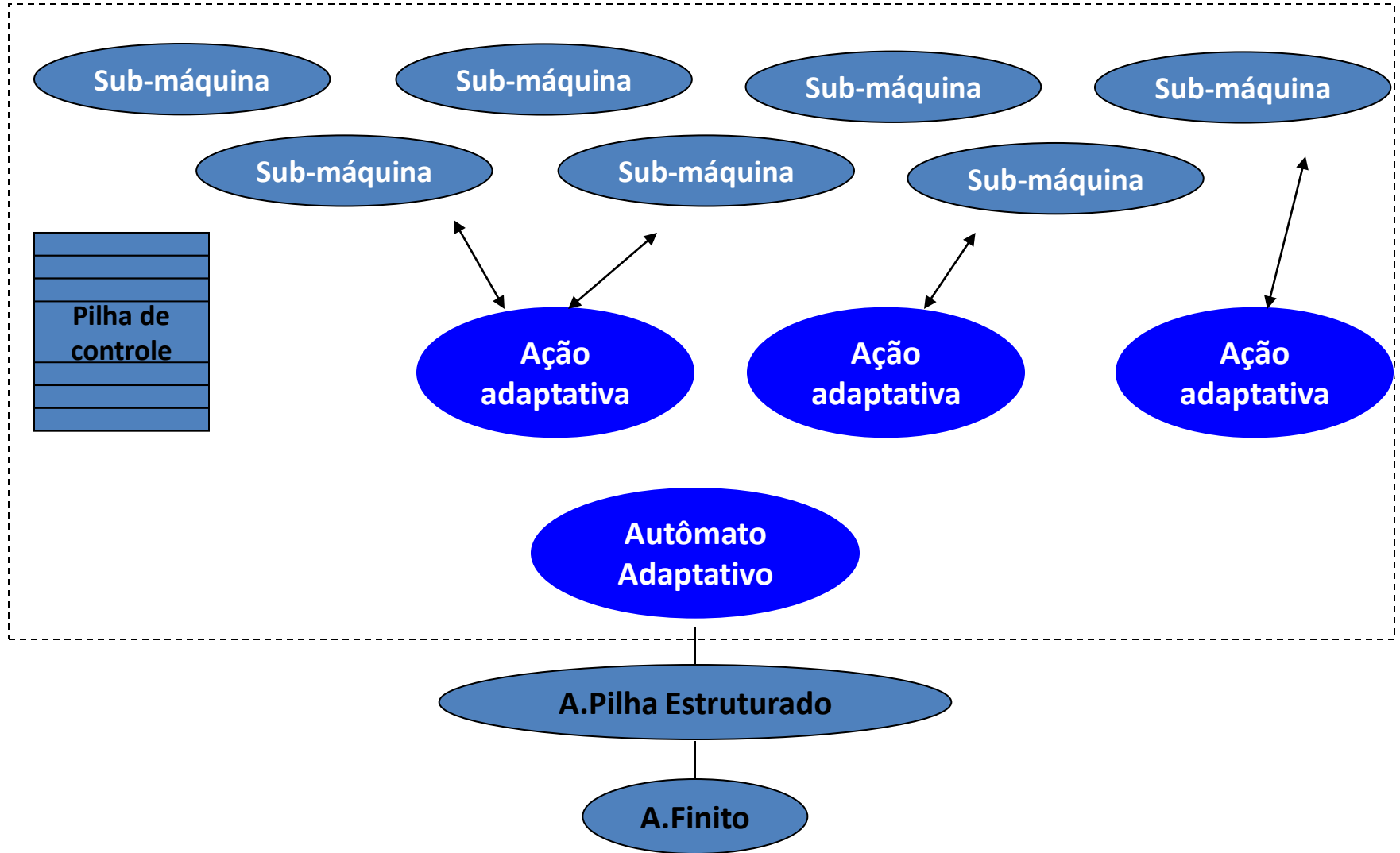


Seq. de Saída

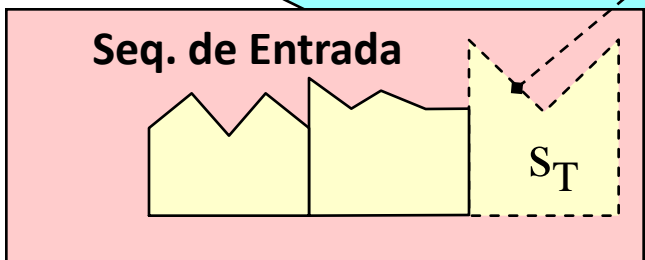
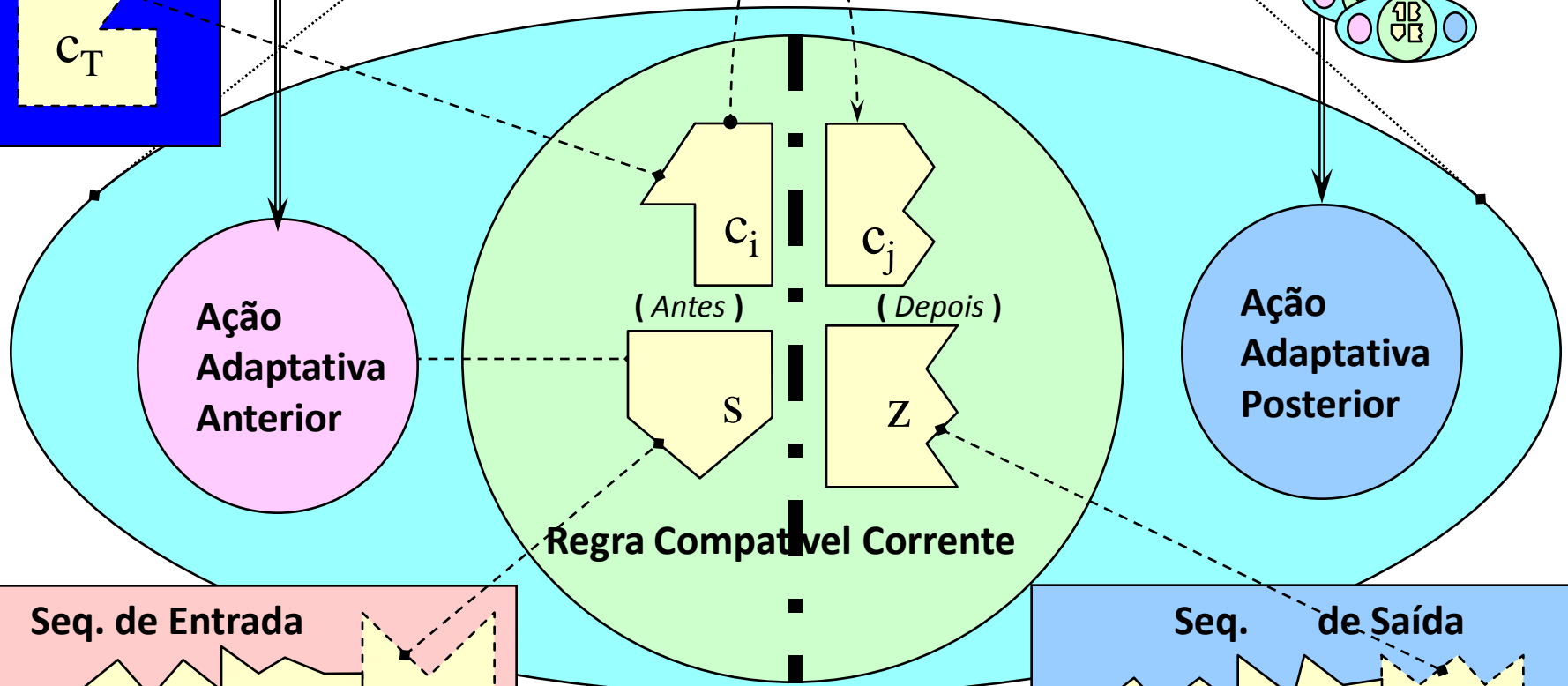
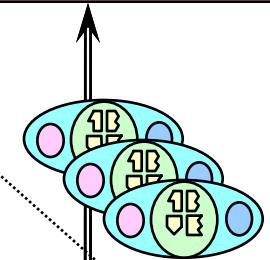
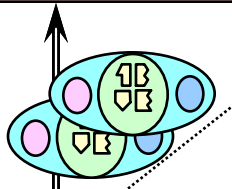
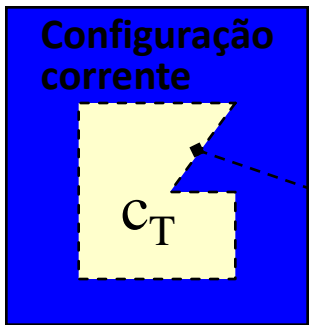
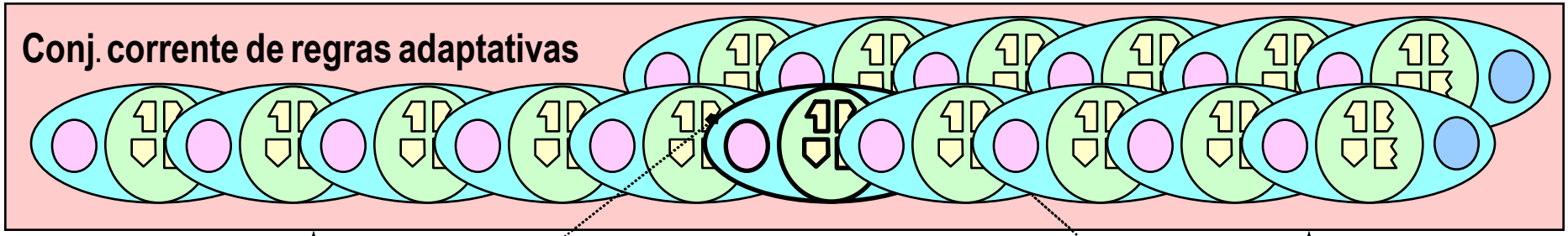


DISPOSITIVO
NÃO-ADAPTATIVO

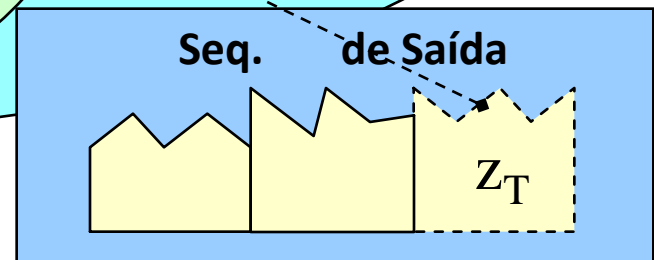
Fase 2 – Formulações Independentes (Adaptatividade simples) [1990 a 2000]



Conj. corrente de regras adaptativas



DISPOSITIVO
ADAPTATIVO



Tag→	H	+	+	+	+	-	H	-	+	1	2	3	4	5	6	7	8	9	10	
state=		p1	g1	g1	g1	g1	p1		q1	q1	S	R	R	R	R	R	R	R	R	E
input=		p2	l	d	↵		p2		↵	↵		I	I	J	J	J	J	K	L	
state:=		g1	J	J	L		J		L	K	I	J		J	J	L				
get		✓									✓	✓		✓	✓					
accept:						8						8				8	✓	✓		
X	B		✓				✓					✓								
Y					✓			A	✓							✓				
p1	P		g1	g1			p1				I									
p2	P		l	d			p2				l									
q1					g1			P	q1							J				
g1	G																			

números das regras

Tabela de decisões subjacente

chamadas de ações adaptativas

condições e ações

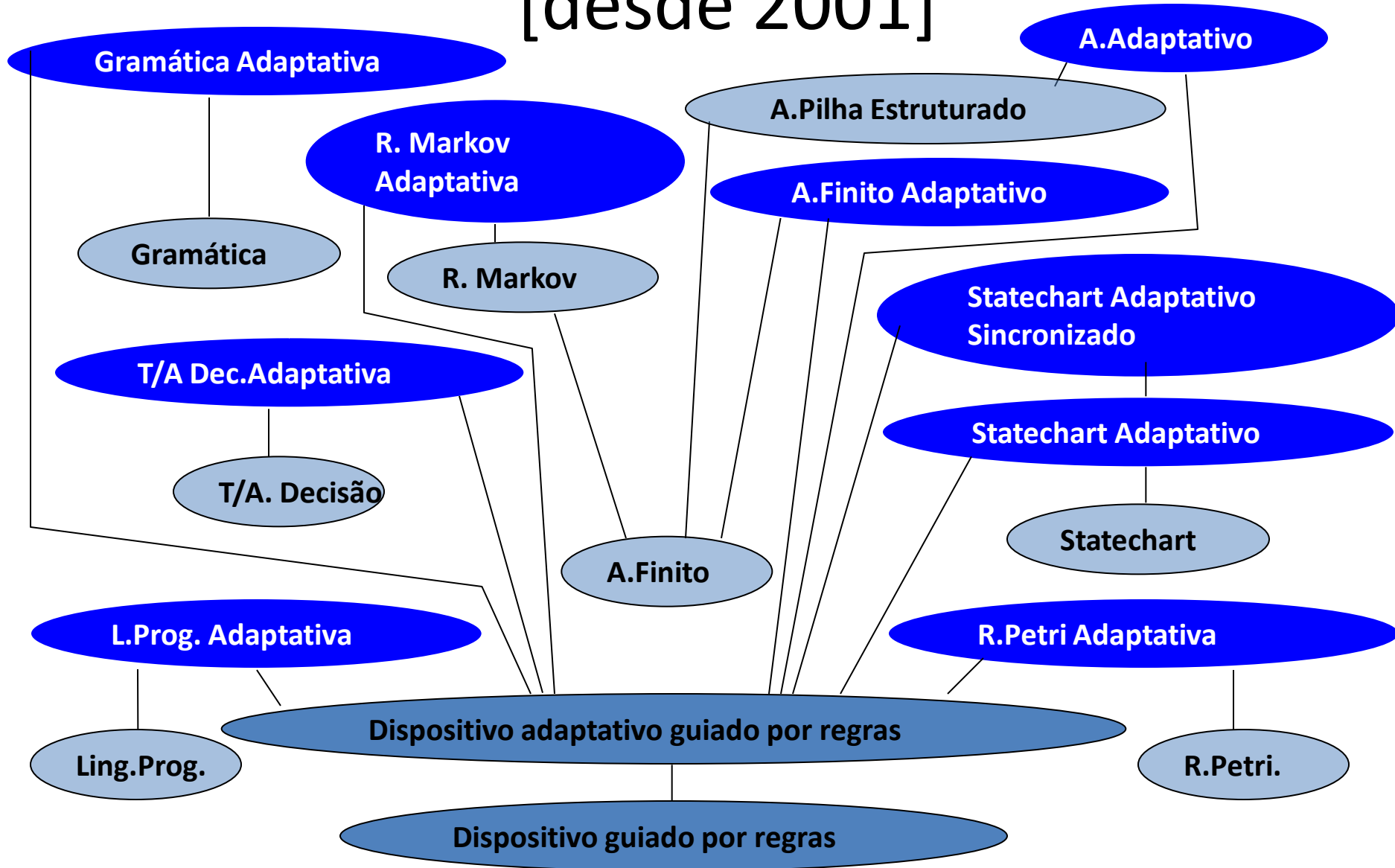
declaração da função adaptativa Y

declaração da função adaptativa X

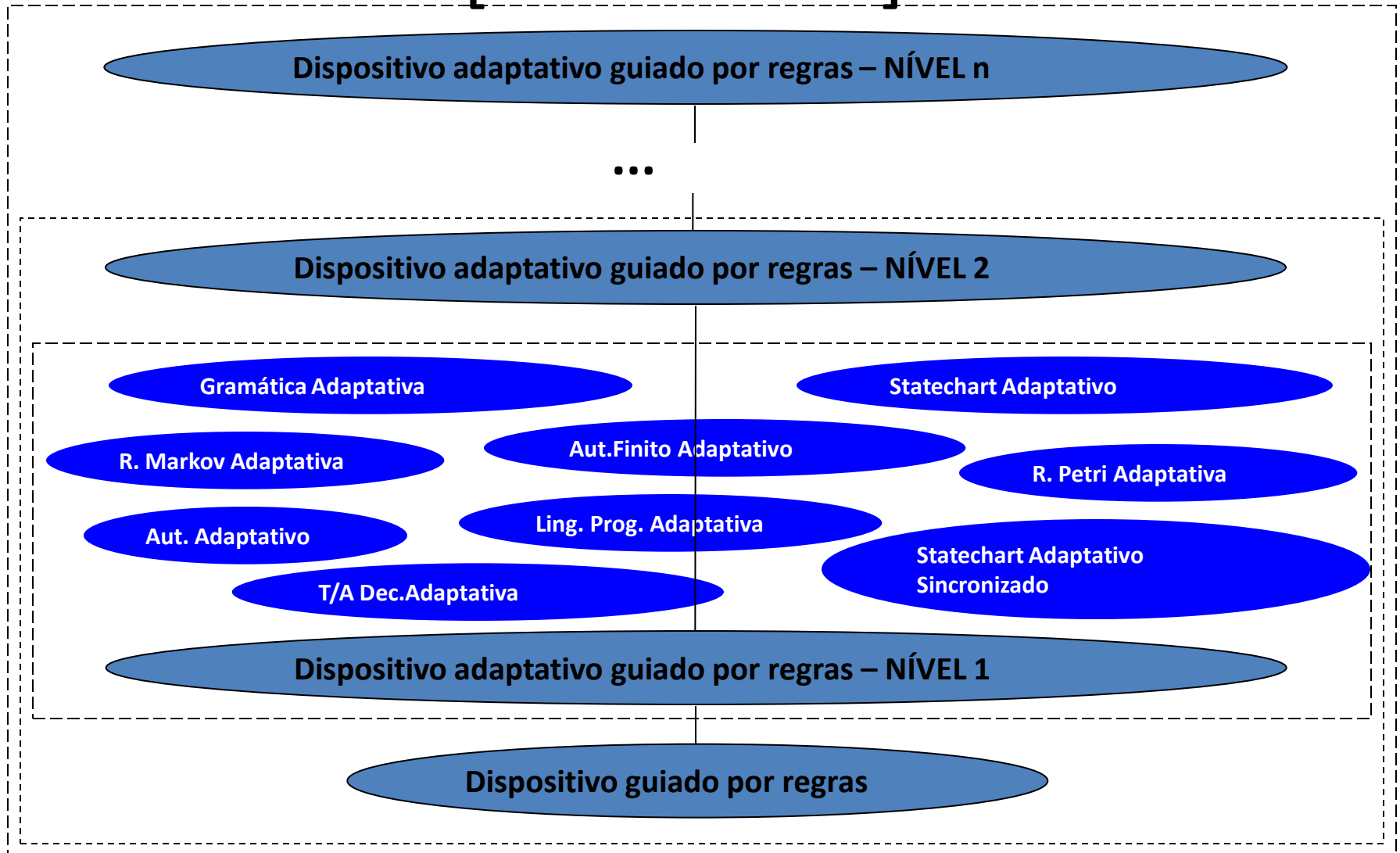
nomes de funções, parâmetros, variáveis e geradores

Tabela de decisões adaptativa

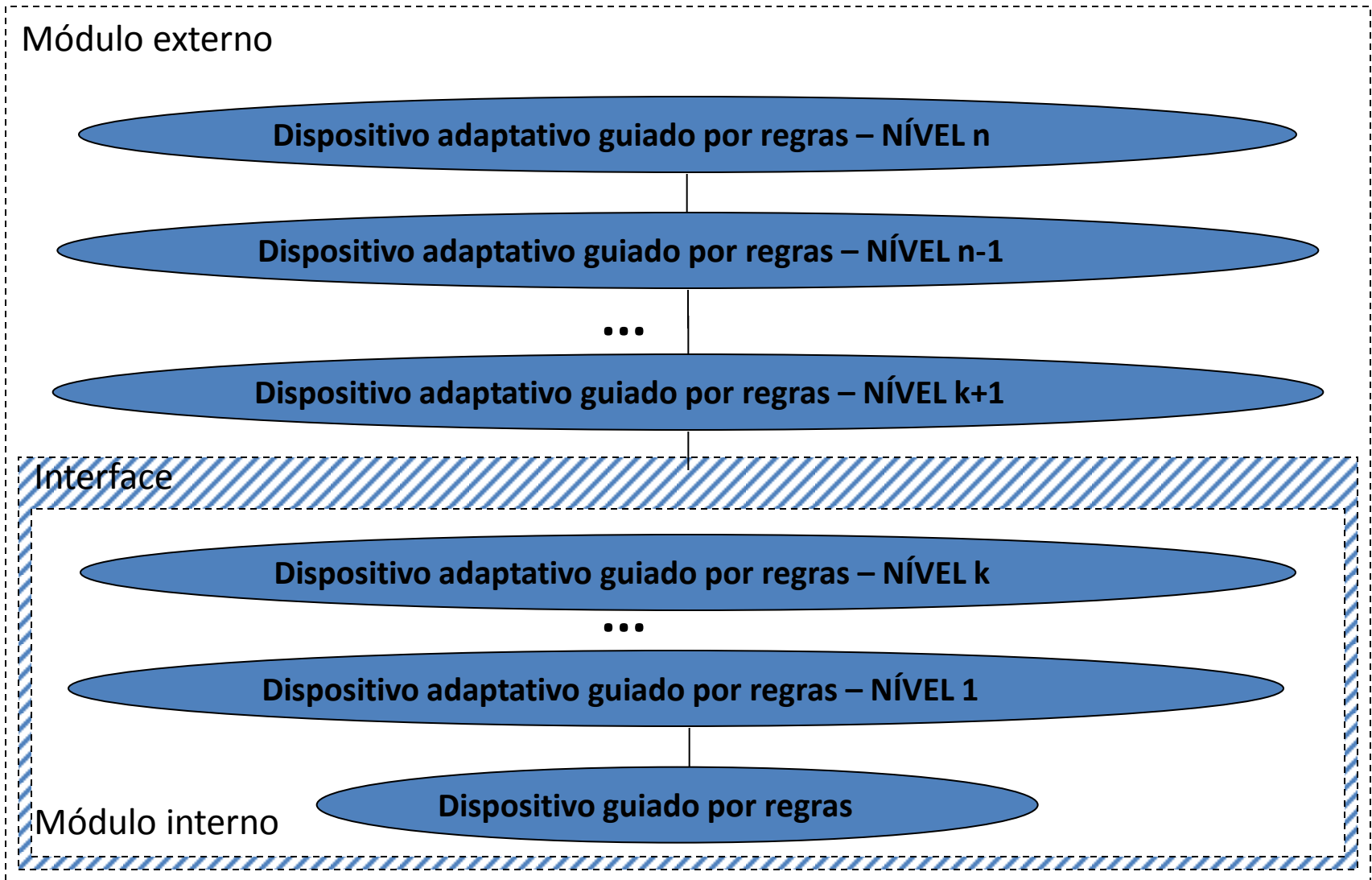
Fase 3 – Formalização Unificada [desde 2001]



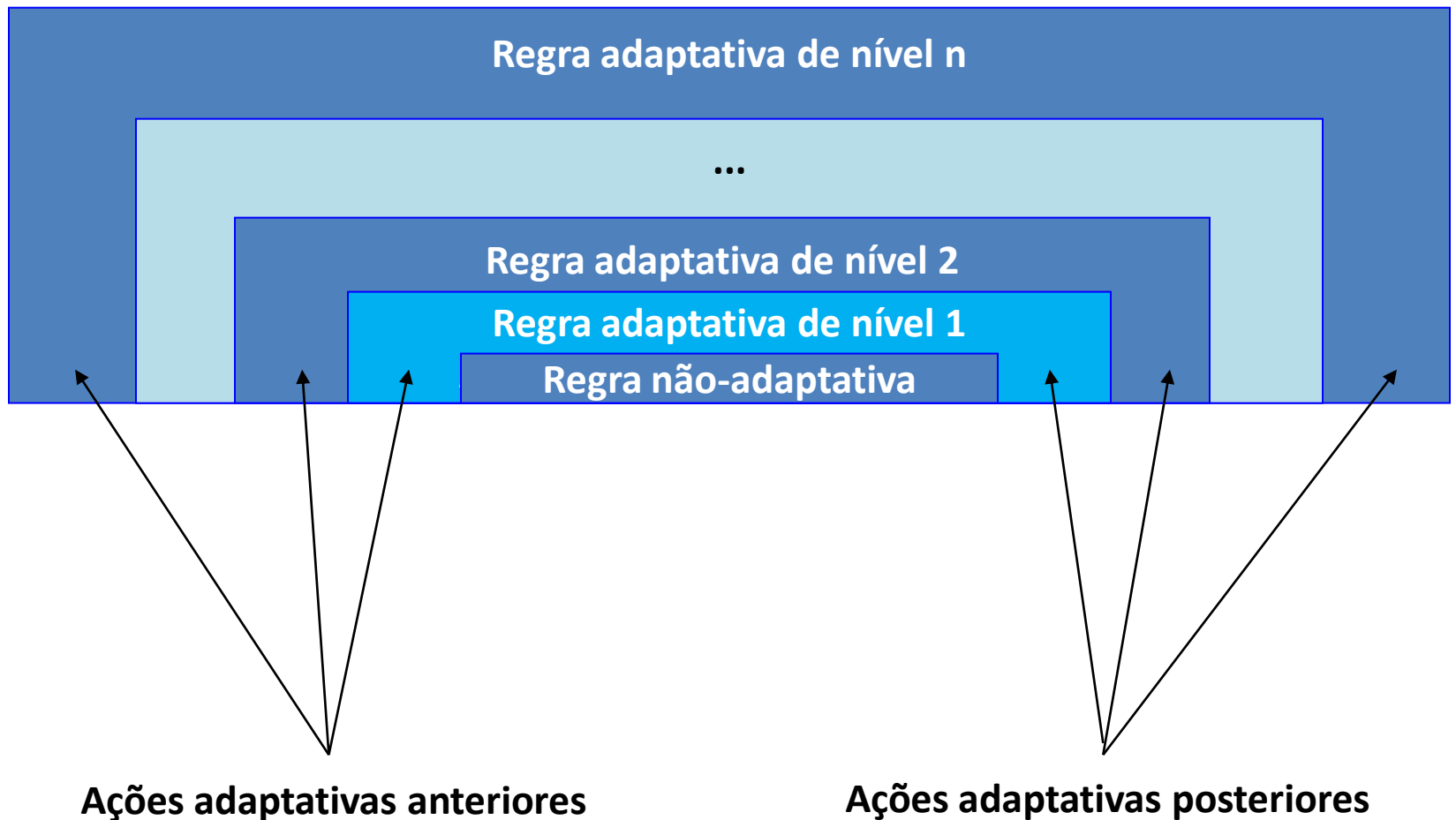
Fase 4 – Formulação Hierárquica [desde 2005]



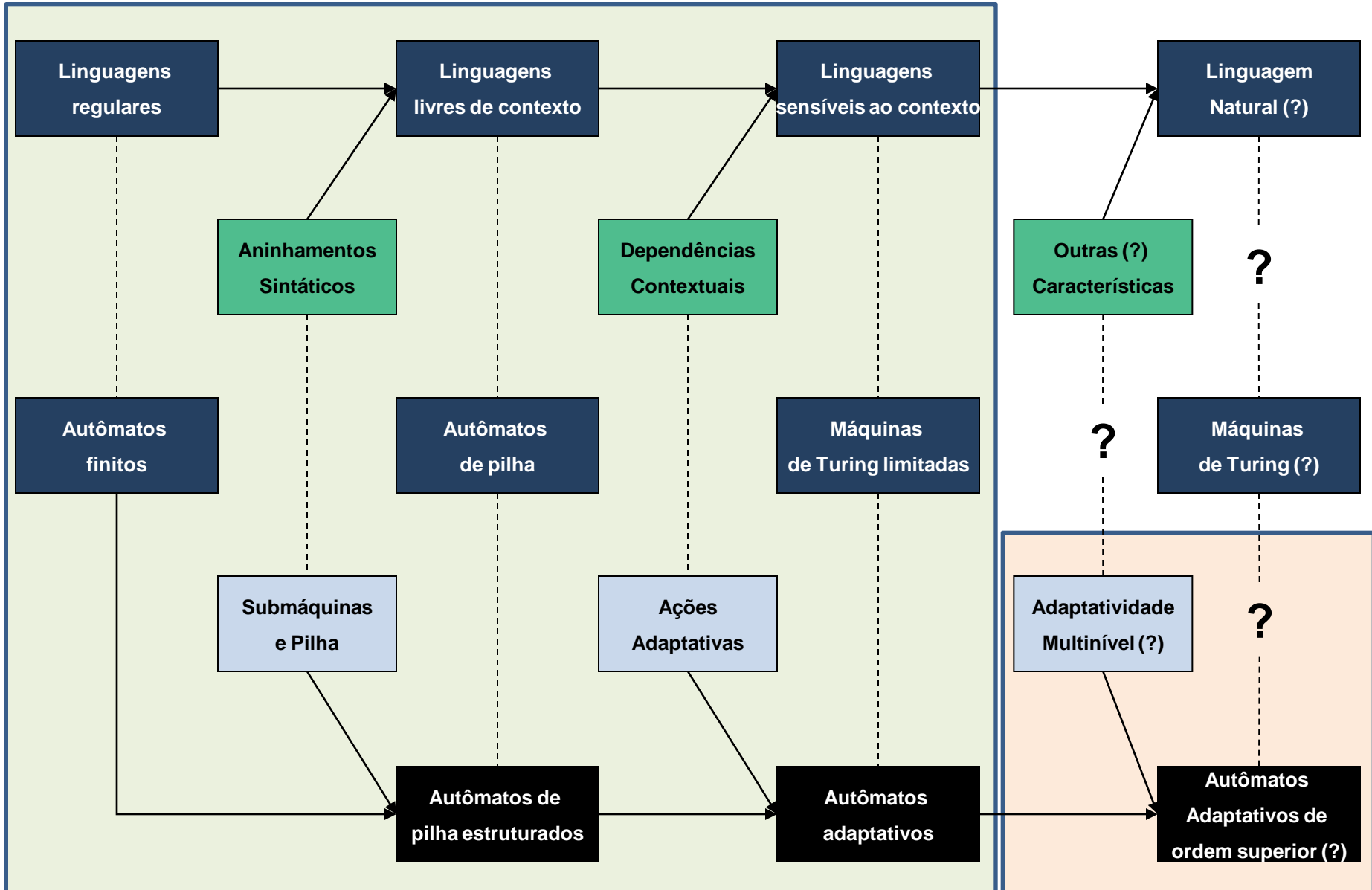
Fase 4 – Modularidade (exemplo)



Formulação Hierárquica regras adaptativas de n níveis



Resumo da Evolução da Pesquisa



A PESQUISA EM LINGUÍSTICA COMPUTACIONAL NO LTA

Estudos em Linguística no LTA

- Descrição/Identificação de Padrões Textuais
- Representação/Geração/Inferência de Linguagem
- Reconhecimento e Análise de Linguagem
- Análise/Síntese fonética/textual
- Análise Léxica/Morfológica/Sintática
- Semântica/Ontologia/Representação da Informação
- Busca e Extração de Informação textual/na Internet
- Conversão/Tradução Automática

ADAPTATIVIDADE

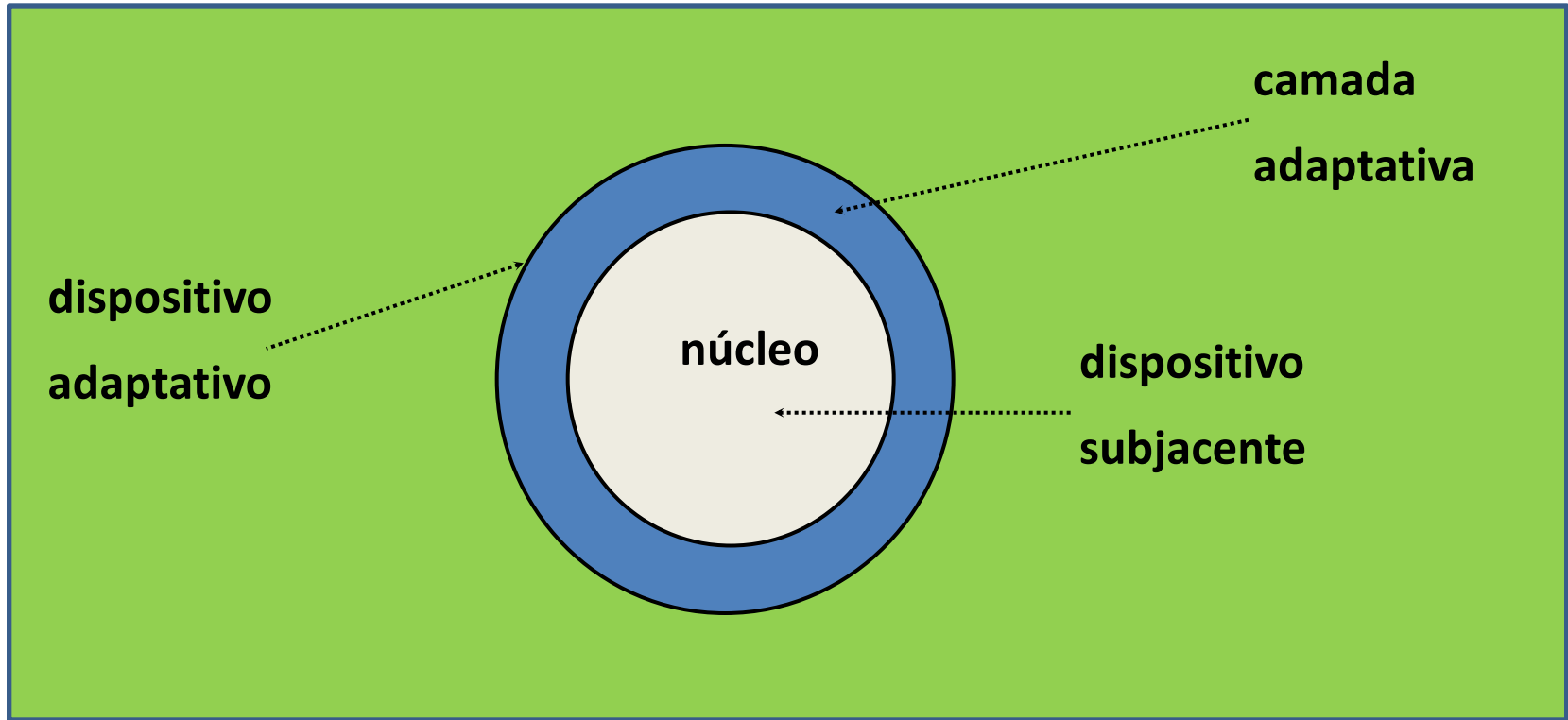
Intuição

- Um sistema se diz adaptativo quando
 - Tem comportamento/reação/operação dinâmicos
 - Essa dinâmica é autônoma e se dá em resposta à
 - Recepção de certos estímulos externos
 - Observação da ocorrência de evento raro/especial, padrão de comportamento não usual, fato inédito/inesperado, etc.
 - Constatação de desvios em relação a referenciais de mérito
- Um sistema convencional pode ser tornado adaptativo acrescentando-se-lhe uma camada responsável por sua automodificação

Dispositivo Adaptativo

- Adaptatividade: característica de um dispositivo que tem a capacidade de mudar o seu próprio comportamento de forma automática, sem interferência externa
 - Um Dispositivo Adaptativo deve se adaptar, através de mudanças comportamentais, como reação aos estímulos recebidos
 - Todo dispositivo adaptativo contém um núcleo que representa o dispositivo subjacente convencional
 - Esse núcleo pode ser dinamicamente modificado pela ação de uma camada adaptativa

Estrutura de um Dispositivo Adaptativo



- **núcleo**, composto por um dispositivo subjacente não-adaptativo
- **camada adaptativa** – mecanismo que age sobre o núcleo e define as **ações adaptativas** que alteram o comportamento, ao longo da operação do dispositivo
- **operação** do dispositivo, definida por um conjunto finito de regras

Arquiteturas Baseadas em Dispositivos Adaptativos

- Comportamento dado por **conjunto de regras**
- **Comportamento inicial** conhecido e fixo
- A **aprendizagem** altera o comportamento pela modificação do conjunto de regras
- Um mecanismo de modificação do conjunto de regras comportamentais pode permitir o **aperfeiçoamento da aprendizagem**

1. Sistemas Adaptativos com Conjunto Homogêneo de Regras

- A forma menos elaborada de organização de um sistema adaptativo consiste em considerá-lo como um **conjunto de regras** não diferenciadas
- Não há hierarquias nem agrupamentos de regras segundo critérios preestabelecidos
- Este arranjo dificulta o projeto, a compreensão, o desenvolvimento e a manutenção do sistema

2. Sistemas Adaptativos com Regras Tematicamente Agrupadas

- Nesta forma de organização, as regras de operação são **agrupadas** segundo seu papel
- Isso origina **centros de regras**, especializados e relativamente independentes
- **Metódico**: facilita o desenvolvimento e a compreensão do funcionamento do sistema
- **Poderoso**: possibilita representar de forma organizada sistemas de alta complexidade

3. Sistemas Adaptativos Apoiados na Tomada de Decisão

- Sistemas adaptativos apoiados na **tomada de decisão** guiam seus movimentos por decisões emanadas de um **conjunto dinâmico de regras**
- São muito práticas para isso as **tabelas**, ou então as **árvores de decisão** no lugar de simples regras ***se-então***
- À medida que deliberam, tais artefatos podem também se **modificar** e se **aperfeiçoar**
- Essa arquitetura propicia a criação rápida de sistemas de **fácil compreensão e alteração**
- Essa técnica é conhecida e de **fácil assimilação**
- Pode ser **facilmente combinada** com outras técnicas

4. Sistemas Adaptativos Incrementalmente Ajustáveis

- Em muitas situações de alta complexidade, pode tornar-se **abusivamente grande o número de alternativas** possíveis em cada passo de operação
- Reduz-se instantaneamente essa variedade **desconsiderando-se situações incompatíveis**
- Assim, os casos a serem analisados são reduzidos, a cada passo de operação, apenas à **parte potencialmente útil** da árvore de possibilidades
- De acordo com o progresso da operação, a cada passo, a parte da árvore mantida em operação vai sendo **incrementalmente ajustada à necessidade**

5. Sistemas Adaptativos Antecipatórios

- Aderente a arquiteturas incrementais
- Em função da trajetória já percorrida, **antecipam-se possíveis passos futuros**, incorporando-os ao conjunto de possibilidades legítimas, e **eliminando opções inviáveis**
- Muito adequado para o tratamento de **dependências de contexto**: verificação de compatibilidades, concordâncias, presença de complementos etc.

6. Sistemas Adaptativos com Aprendizagem

- Estrutura inicial conhecida – instinto
- Regras iniciais de aprendizagem – inteligência básica
- Centros comportamentais especializados – regiões do cérebro
- Influência mútua entre os centros
- Regras de reatividade – como responder a um estímulo
- Regras de assimilação de conhecimento – aprendizagem
- Regras de alteração da forma de aprender – evolução
- Regras de criação de novas regras – inferências, criatividade
- Regras de tomada de decisão – autonomia
- Regras de auto-avaliação e auto-crítica – planejamento
- Regras de generalização e de especialização – lógica, dedução, inferência
- Hierarquia dos centros comportamentais

ALGUNS ESTUDOS DE CASOS

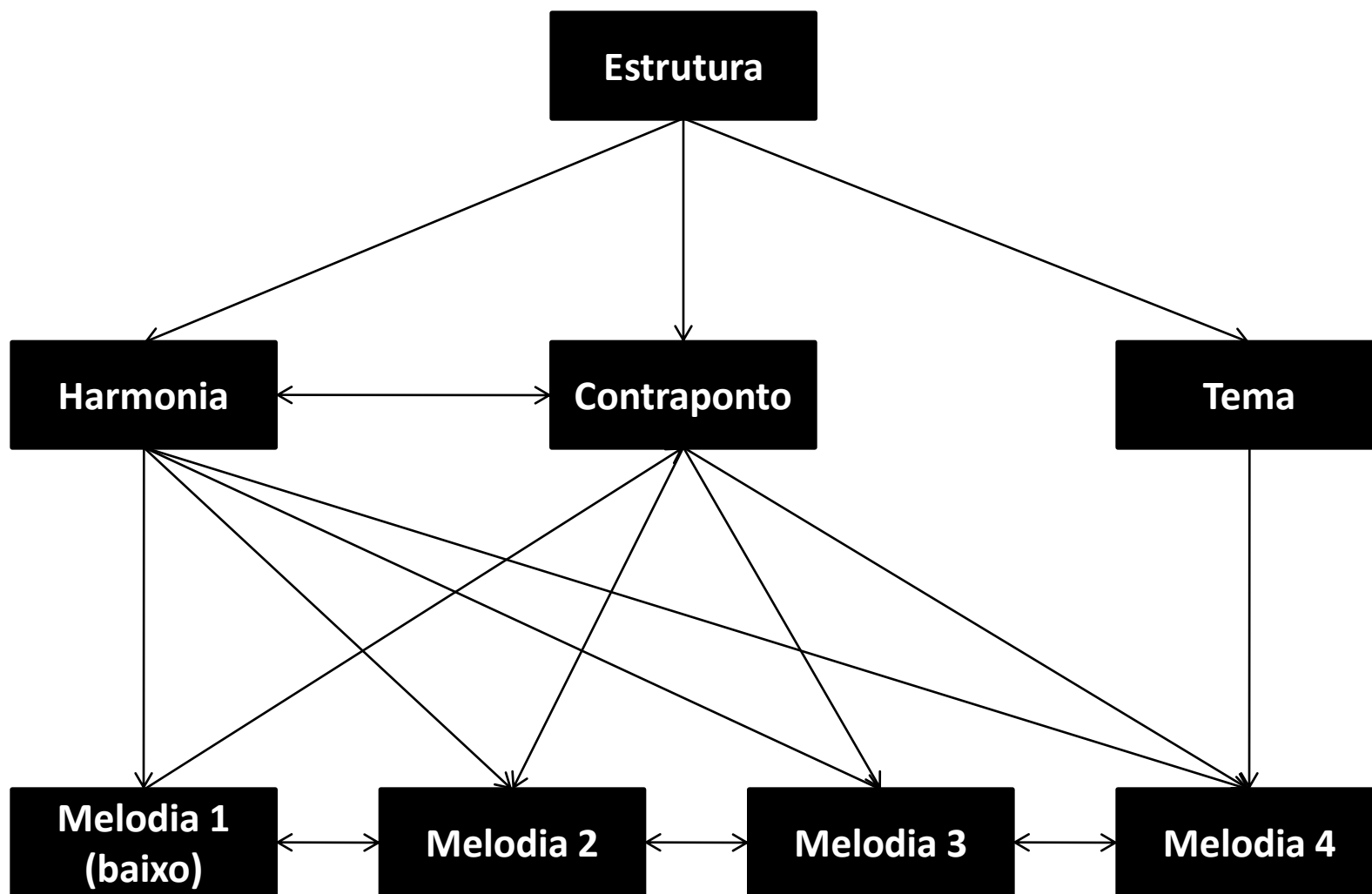
1. Composição Musical

Redes de Markov Adaptativas

Composição musical

- Música é uma forma de expressão humana , logo constitui linguagem natural
- Basseto, B. A. (Mestrado, EPUSP, 2000)
Um sistema de composição musical automatizada, baseado em gramáticas sensíveis ao contexto, implementado com formalismos adaptativos.
- Implementação organizada em centros
- Redes de Markov Mutuamente Adaptativas
 - Cada centro (rede) modela um item da composição
 - Regras de composição musical orientam decisões
 - Probabilidades variáveis calibram escolhas múltiplas

Interação entre centros (Redes de Markov Adaptativas) na composição musical



2. Tomada de Decisão

Tabelas de Decisão Adaptativas

Exemplo 2: TABELA DE DECISÃO ADAPTATIVA

	linha de cabeçalho (tag)	colunas das funções adaptativas	colunas das regras
Tabela de decisão convencional	linhas das condições	declaração das funções adaptativas	valores das condições
	linhas das ações		ações a serem aplicadas
Funções Adaptativas	linhas das funções adaptativas		ações adaptativas a serem executadas

Estrutura geral da *TDA*

ESTRUTURA DE UMA TABELA DE DECISÃO ADAPTATIVA

- a) **dispositivo subjacente (núcleo)**: tabela de decisão convencional usada para modelar problemas envolvendo condições, ações e regras:
- **condições**: valores utilizados na tomada de decisão;
 - **ações**: operações executadas em resposta às condições;
 - **regras**: associam condições às correspondentes ações;
- b) **camada adaptativa**, formada por um grupo de ações adaptativas que agem sobre o conjunto de regras modificando-o dinamicamente. A camada adaptativa associa ações adaptativas às regras, executando-as sempre que a regra for aplicada.

Ilustração de Tabela de Decisão Adaptativa

	Condições e ações							Declaração da função adaptativa Y			Declaração da função adaptativa X								
Tag→	H	+	+	+	+	+	-	H	-	+	S	R	R	R	R	R	R	R	E
state=		p1	g1	g1	g1	g1	p1		q1	q1		I	I	J	J	J	J	K	L
input=		p2	l	d	↯		p2		↯	↯		l		l	d	↯			
state:=		g1	J	J	L		J		L	K		I	J		J	J	L		
get		✓										✓	✓		✓	✓			
accept:=							✗						✗				✗	✓	✓
X	B		✓	✓			✓					✓							
Y					✓			A	✓								✓		
p1	P		g1	g1			p1					I							
p2	P		l	d			p2					l							
q1					g1			P	q1								J		
g1	G																		

Números das regras

Cabeçalhos das regras

Tabela de decisão não-adaptativa subjacente

Chamadas de funções adaptativas

Condições e ações

Declaração da função adaptativa Y

Declaração da função adaptativa X

Nomes de funções, parâmetros e geradores

Efeito da Adaptatividade em Tabelas de Decisão

Tabela de Decisão é estática:

- regras são previamente especificadas
- regras não podem ser modificadas durante a sua operação
- se novas regras surgem, a tabela deve ser refeita
- Podem resultar desperdícios de recursos computacionais

Tabela de Decisão Adaptativa

é dinâmica:

- pode dar maior flexibilidade ao processo de decisão
- permite a inclusão e a exclusão de regras durante a operação do dispositivo
- transforma a tabela de decisão numa ferramenta mais poderosa e inteligente

3. Verificação de Tipos

Sistemas antecipatórios

Aplicação em linguagens de
programação

Tratamento Sintático de Tipos

- A declaração de algum objeto em uma linguagem de programação para o seu uso posterior define atributos do objeto, que devem ser verificados quanto à coerência entre a declaração e o uso.
- A verificação é feita através de ações adaptativas que alteram a topologia do autômato, para aceitar somente as construções sintáticas em que sejam obedecidos os vínculos corretos entre os tipos declarado e esperado.
- Neste exemplo, mostra-se um caso no qual há dois tipos de variáveis, e assim, dois tipos de comandos de atribuição, envolvendo inteiros e reais.
- Os comandos de atribuição compõem-se de duas partes, uma variável destino e uma expressão aritmética (extremamente simplificada, com números e variáveis).

Funcionamento do Verificador

- Encontrado o identificador que representa a variável destino da atribuição, utiliza-se o tipo a ele associado como referência a uma ação adaptativa, que altera a topologia de forma que apenas expressões compatíveis sejam aceitas.
- Ao ser executada a função adaptativa \mathcal{F} , com parâmetro v.int ou v.real, insere-se no autômato uma transição que consome identificadores, assim classificados pelo analisador léxico, de forma que as expressões sejam compostas por elementos do mesmo tipo da variável destino.
- Ao final do reconhecimento do comando de atribuição, esta transição acrescentada é removida por outra ação adaptativa, restaurando as condições iniciais.

Verificador de Tipos (atribuição)

$(a1, \underline{v.int}) : \rightarrow a2, \mathcal{F}(\underline{v.int})$

$(a1, \underline{v.real}) : \rightarrow a2, \mathcal{F}(\underline{v.real})$

$(a2, "=") : \rightarrow a3$

$(a3, \underline{num}) : \rightarrow a4$

$(a4, "+") : \rightarrow a3$

$\{(\gamma, a4, \theta \alpha) : \rightarrow (\gamma, a5, \theta \alpha) \quad \forall \theta \in \Sigma\}$

$\{(\gamma k, a5, \alpha) : \rightarrow (\gamma, k, \underline{atrib} \alpha) \quad \forall k \in \Gamma\}$

$\mathcal{F}(\mathbf{x}) = \{ :$

$+ [(\mathbf{a3}, \mathbf{x}) : \rightarrow \mathbf{a4}]$

$\{- [(\gamma, \mathbf{a4}, \theta \alpha) : \rightarrow (\gamma, \mathbf{a5}, \theta \alpha) \quad \forall \theta \in \Sigma]\}$

$\{+ [(\gamma, \mathbf{a4}, \theta \alpha) : \mathcal{G}(\mathbf{x}), \rightarrow (\gamma, \mathbf{a5}, \theta \alpha) \quad \forall \theta \in \Sigma]\}$ }

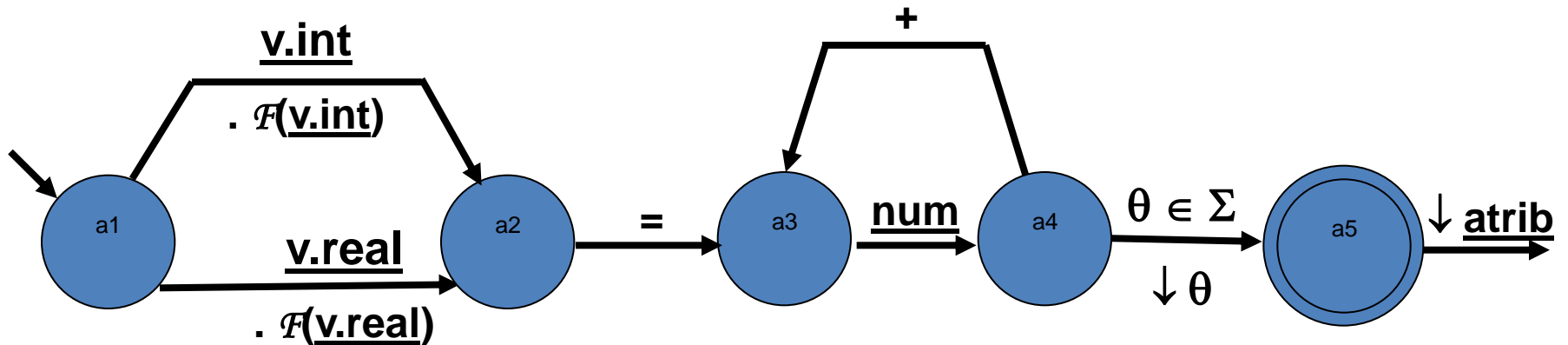
$\mathcal{G}(\mathbf{x}) = \{ j^* :$

$- [(\mathbf{a3}, \mathbf{x}) : \rightarrow \mathbf{a4}]$

$\{- [(\gamma, \mathbf{a4}, \theta \alpha) : \mathcal{G}(\mathbf{x}), \rightarrow (\gamma, \mathbf{a5}, \theta \alpha) \quad \forall \theta \in \Sigma]\}$

$\{+ [(\gamma, \mathbf{a4}, \theta \alpha) : \rightarrow (\gamma, \mathbf{a5}, \theta \alpha) \quad \forall \theta \in \Sigma]\}$ }

Verificador de Tipos (atribuição)



$\mathcal{F}(x) = \{ :$

$+ [(a3, x) : \rightarrow a4]$

$\{- [(\gamma, a4, \theta \alpha) : \rightarrow (\gamma, a5, \theta \alpha) \forall \theta \in \Sigma]\}$

$\{+ [(\gamma, a4, \theta \alpha) : \mathcal{G}(x), \rightarrow (\gamma, a5, \theta \alpha) \forall \theta \in \Sigma]\}$ }

$\mathcal{G}(x) = \{ j^* :$

$- [(a3, x) : \rightarrow a4]$

$\{- [(\gamma, a4, \theta \alpha) : \mathcal{G}(x), \rightarrow (\gamma, a5, \theta \alpha) \forall \theta \in \Sigma]\}$

$\{+ [(\gamma, a4, \theta \alpha) : \rightarrow (\gamma, a5, \theta \alpha) \forall \theta \in \Sigma]\}$ }

4. Tratamento Adaptativo de Aspectos da Linguagem Natural

Sistemas antecipatórios

Aplicação da adaptatividade no processamento de linguagem natural

Processamento de LN usando AA (1)

- É possível usar técnicas puramente sintáticas para o tratamento de dependências de contexto em linguagens naturais, a exemplo do que se faz com linguagens de programação
- Pressupõe-se o uso de um analisador léxico, que pode ser construído usando técnicas adaptativas, como foi feito no etiquetador morfológico apresentado por Menezes.
- Para dar partida ao processo, usa-se uma aproximação livre de contexto da linguagem natural, formalizada através de uma gramática livre de contexto, na Notação de Wirth
- Nessa gramática, definem-se todos os elementos essenciais das sentenças, em sua sintaxe completa, em ordem direta, sem omitir e sem repetir componentes

Processamento de LN usando AA (2)

- Num primeiro momento, não-terminais que denotam sintagmas são tratados como se fossem terminais
- Aplica-se a essa gramática o método usual de geração de reconhecedor sintático para linguagens livres de contexto, obtendo-se um autômato de pilha estruturado equivalente, que servirá como ponto de partida para a obtenção do AA que reconhece a linguagem natural
- Transforma-se esse autômato, para incluir tratamentos adaptativos específicos para uma série de fenômenos lingüísticos não-livres-de-contexto

Processamento de LN usando AA (3)

- Considerar ao menos os seguintes fenômenos lingüísticos:
 1. Não-determinismos
 2. Ambigüidades morfológicas e sintáticas
 3. Dependências contextuais próximas (p.ex. concordâncias)
 4. Omissão de componente (p. ex. elipse)
 5. Unicidade (p.ex. componente obrigatório único)
 6. Multiplicidade (p. ex. coordenações, polissíndetos)
 7. Flexões e concordâncias em gênero, número, grau, tempo, modo, pessoa
 8. Regência (p.ex. preposições em objeto indireto e complemento nominal)
 9. Dependências contextuais distantes (p.ex. sujeito oculto e anáforas)
 10. Expletivos e apostos
 11. Inversões
 12. Substituição de componentes básicos por sintagmas complexos
 13. Aninhamento de frases relativas

Processamento de LN usando AA (4)

- Para cada uma das dependências contextuais apresentadas, localizar na gramática o ponto de representação mais apropriado, e gerar alterações convenientes no autômato, de forma que seja tratada a correspondente dependência de contexto
- Considerar finalmente a presença de sintagmas complexos nos locais em que havia sido referenciado, na gramática simplificada, apenas um componente simples
- Isso completa a definição formal da linguagem natural, quanto aos aspectos formalizados pelas alterações que assim forem sendo incrementalmente introduzidas.
- Dessa forma, com um conjunto apropriado de alterações bem escolhidas, vai sendo possível criar aproximações dependentes de contexto cada vez mais satisfatórias da linguagem natural desejada

Tratamento adaptativo dos fenômenos lingüísticos citados

- Para cada um dos fenômenos dependentes de contexto mencionados anteriormente, é necessário que as necessárias verificações de contexto sejam devidamente efetuadas pelo autômato adaptativo, e para isso, propõem-se a seguir alguns padrões adaptativos correspondentes, para serem inseridos no autômato existente.

Não-determinismos e ambigüidades

- O tratamento proposto para não determinismos e ambigüidades consiste em simular o processamento paralelo dos diversos casos simultâneos, e descartando aqueles que falharem durante essa simulação. Esse procedimento está descrito em:
- Neto, Joao Jose and Moraes, Miryam de.
 - Using Adaptive Formalisms to Describe Context-Dependencies in Natural Language. Lecture Notes in Artificial Intelligence.
<http://link.springer.de/link/service/series/0558/bibs/2721/27210094.htm>
- Neto, João José e Moraes, Miryam de.
 - Formalismo adaptativo aplicado ao reconhecimento de linguagem natural.
http://www.pcs.usp.br/~lta/artigos/miryam_cisci2002.pdf

Dependências contextuais próximas (p.ex. concordâncias)

- São tratadas de forma similar à verificação de tipos em linguagens de programação:
 - os atributos (etiquetas) correspondentes a uma palavra ou sintagma são coletados na primeira ocorrência dos elementos envolvidos.
 - São então incluídos no autômato em pontos estratégicos, correspondentes ao local onde se espera encontrar os elementos que com eles guardam dependências contextuais.
 - Nesse ponto, ao ser encontrado o elemento correspondente do texto analisado, é feita a verificação pelo autômato, aceitando apenas as opções consideradas compatíveis e rejeitando as demais.

Elipses e Anáforas

- Trata-se da omissão de um componente da sentença
- Caso se trate do sujeito, pode-se tentar inferi-lo na forma de pronome através da análise da flexão verbal associada.
- Para resolver a anáfora resultante, ou alguma outra explicitamente presente na sentença, pode-se também extrair informações adicionais, e até identificar outros elementos a que se refere tal elemento oculto pela inspeção de contextos vizinhos.
- Elementos citados em frases anteriores ou seguintes são o alvo principal dessa busca. Isso não é tão simples de implementar, e nem sempre leva a resultados precisos.

Unicidade (p.ex. componente obrigatório único)

- Elementos de ocorrência única, uma vez encontrados, não mais devem ser aceitos no mesmo contexto.
- Para isso, quando possível, após sua localização no texto, convém inibir o acesso às transições que promovem sua aceitação.
- Deve-se restaurar esse acesso ao sair do ambiente em que aquele elemento foi encontrado, para que, na próxima vez que o autômato voltar a percorrer tal caminho, seja novamente possível identificar e aceitar esse elemento, agora em outro contexto.

Multiplicidade (p. ex. coordenações, polissíndetos)

- A representação da possibilidade de ocorrências múltiplas de determinadas construções sintáticas é relativamente simples, e se resume a aceitar uma nova instância dessa construção ao final de cada ocorrência encontrada.
- Na prática, o tratamento desse fenômeno pode ser tratado simplesmente retornando-se ao estado inicial do autômato que o reconhece, todas as vezes que for atingido seu estado final.
- Esgotadas as repetições, restaura-se o autômato.

Concordância em gênero, número, grau, tempo, modo, pessoa

- Esta dependência de contexto pode ser tratada de forma similar ao que ocorre nas verificações de tipos no tratamento de linguagens de programação.
- Em cada situação, determina-se, a partir das informações coletadas até o momento das palavras já processadas, os atributos que exigem concordância
- Impõe-se então, por meio de uma alteração do autômato, que nos pontos nos quais se aguarda a continuação da sentença, sejam encontrados atributos gramaticalmente concordantes.
- Restaura-se finalmente o autômato para ser reutilizado

Regência (p.ex. uso de preposições em objeto indireto e complemento nominal)

- O tratamento da regência em suas diversas variedades pode ser feito de forma similar ao da concordância.
- Inicialmente, constata-se a presença de um verbo ou substantivo que exige complemento
- Determina-se então a preposição necessária e o sintagma que deve acompanhá-la
- Finalmente, altera-se o autômato de forma que seja forçado o aparecimento da preposição e do sintagma anteriormente determinados
- Ao final do tratamento, restaura-se o autômato

Dependências contextuais distantes

- Não há como determinar antecipadamente a posição, na sentença, dos alvos de dependências contextuais a não ser caso a caso.
- Por isso, seu tratamento pode ser efetuado criando-se um repositório cujo conteúdo são as informações coletada que ainda não encontraram os elementos da sentença que com elas combinem
- As pendências nesse repositório podem ser descartadas quando completamente resolvidas.

Expletivos e apostos

- Não tendo utilidade semântica, os expletivos podem ser simplesmente descartados na análise sintática da sentença.
- Já os apostos podem contribuir com algumas informações eventualmente ausentes do texto principal, as quais podem ser extraídas e memorizadas em modificações do autômato, para uso posterior.

Inversões

- Inversões podem ser tratadas adaptativamente usando-se uma configuração do autômato que propicie a permutação dos elementos da sentença cuja ordem possa ser arbitrariamente alterada.
- A ocorrência de um elemento deve removê-lo do conjunto dos elementos aguardados.
- Quando todos os elementos obrigatórios aguardados tiverem sido identificados a sentença pode terminar corretamente se não houver mais elementos opcionais a serem tratados

Substituição de componentes simples por sintagmas complexos

- O tratamento desta operação é relativamente simples: deve-se substituir transições normais consumindo o componente simples, por transições de chamada de submáquina, para que o sintagma adequado seja reconhecido.
- Não se trata de um procedimento adaptativo, e sim manual, a ser aplicado ao autômato em evolução a cada vez que se deseje eliminar alguma simplificação anteriormente inserida.

Aninhamento de frases relativas

- O aninhamento de frases relativas não é dependente de contexto, e sim um simples fenômeno livre de contexto de aninhamento sintático, que pode ser tratado facilmente pelo próprio autômato de pilha estruturado subjacente.

5. Representação Simplificada

Representação simplificada da sintaxe da língua portuguesa para formulação adaptativa

Elaborada com base em

Celso Pedro Luft - *Moderna Gramática Brasileira*

Estrutura de referência

- Quatro casas sintáticas:

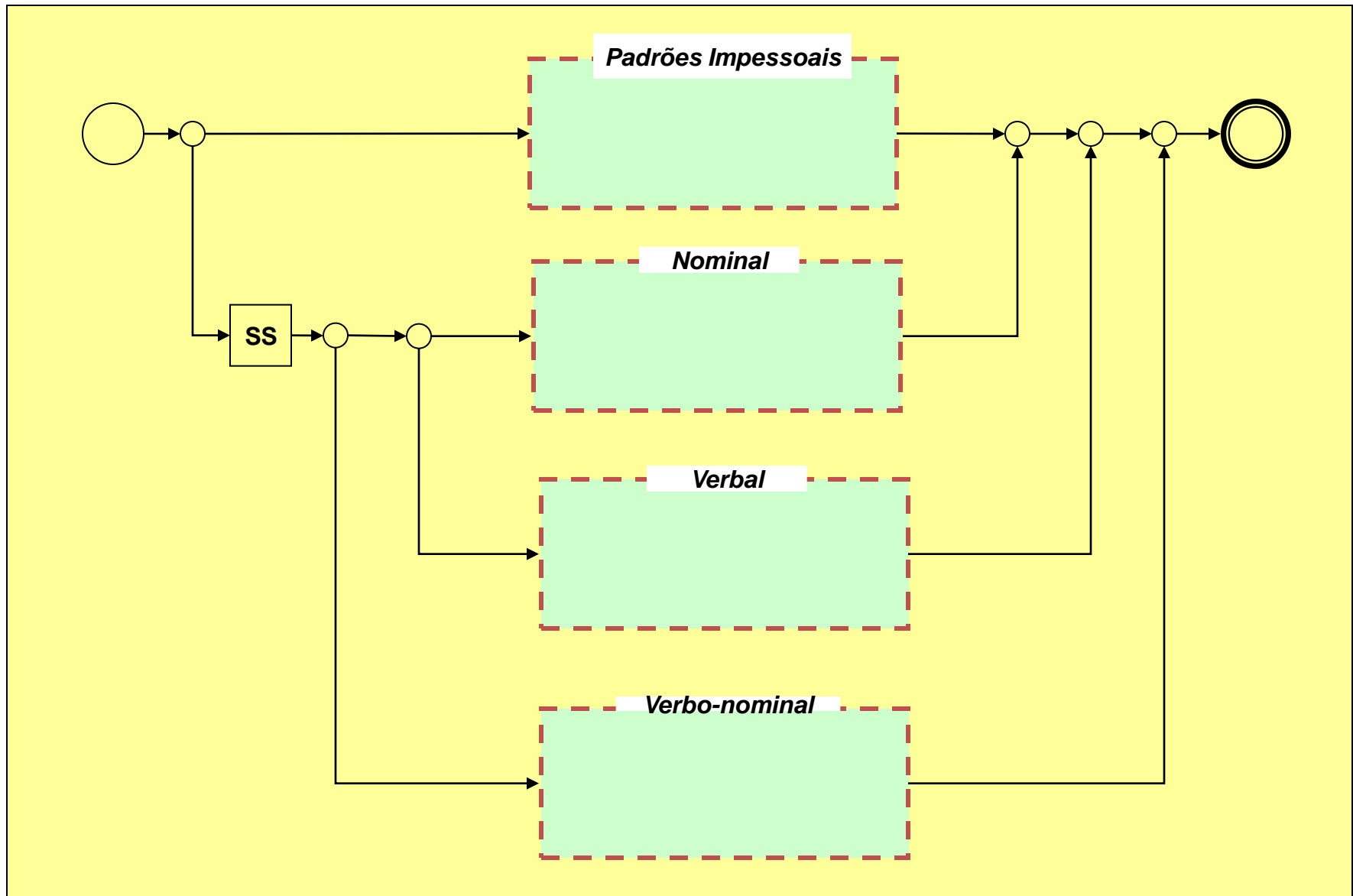
(1)	(2)	(3)	(4)
Sujeito	Verbo	Complementos	Adjuntos Adverbiais

- Casa 1: Sujeito
- Casa 2: Verbo
- Casa 3: Complementos
- Casa 4: Adjuntos Adverbiais

Abreviaturas

Abreviatura	Significado	Interpretação
SS	Sintagma Substantivo	
Sadj	Sintagma Adjetivo	
Sadv	Sintagma Adverbial	
Prep	Preposição	
SP	Sintagma Preposicional	Prep SS
Vlig	Verbo de Ligação (predicativo)	
Vi	Verbo Intransitivo	
Vtda	Verbo Transitivo Direto Apassivável	
Vtdna	Verbo Transitivo Direto Não-Apassivável	
Vtd	Verbo Transitivo Direto	Vtda Vtdna
Vti	Verbo Transitivo Indireto	
Vtdi	Verbo Transitivo Direto e Indireto	
Vt-pred	Verbo Transitivo Predicativo	
Vi-pred	Verbo Intransitivo Predicativo	

Sintaxe Geral, conforme padrões das pg. 31~36



Padrões Impessoais

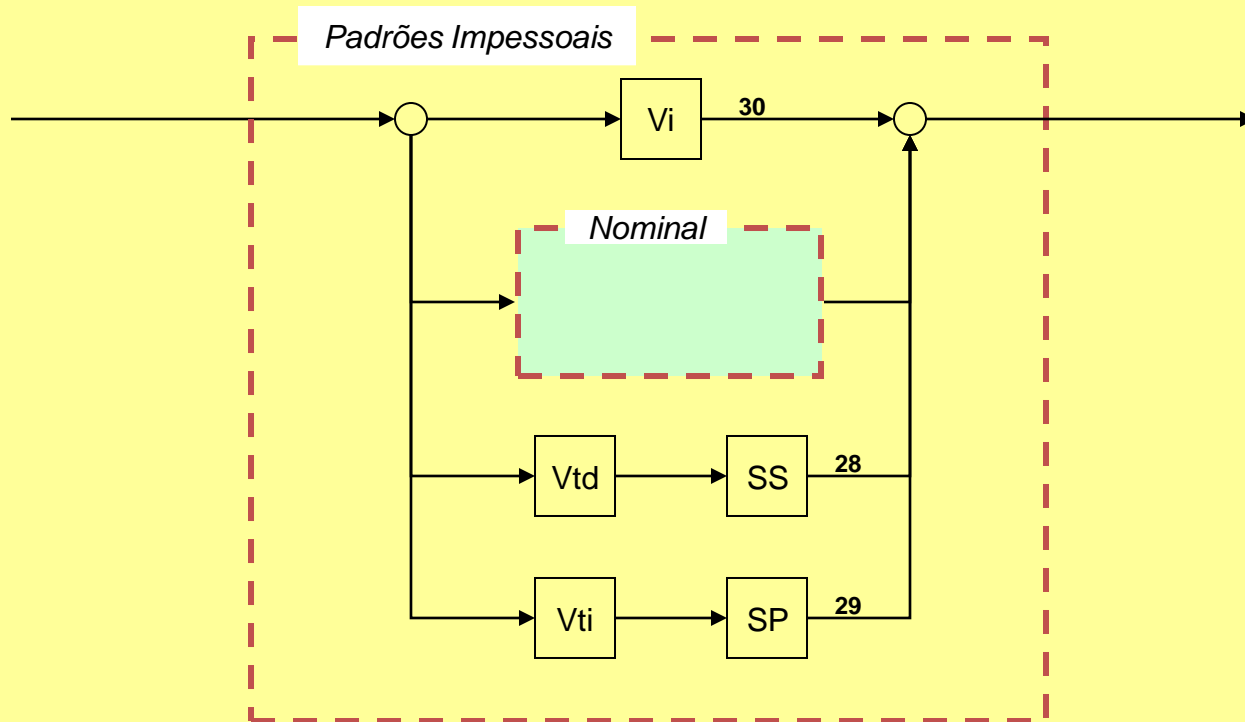
- Nominais

24	Vlig SS
25	Vlig Sadj
26	Vlig Sadv
27	Vlig SP

- Verbais

28	Vtd SS
29	Vti SP
30	Vi

Padrões Impessoais



Padrões Pessoais

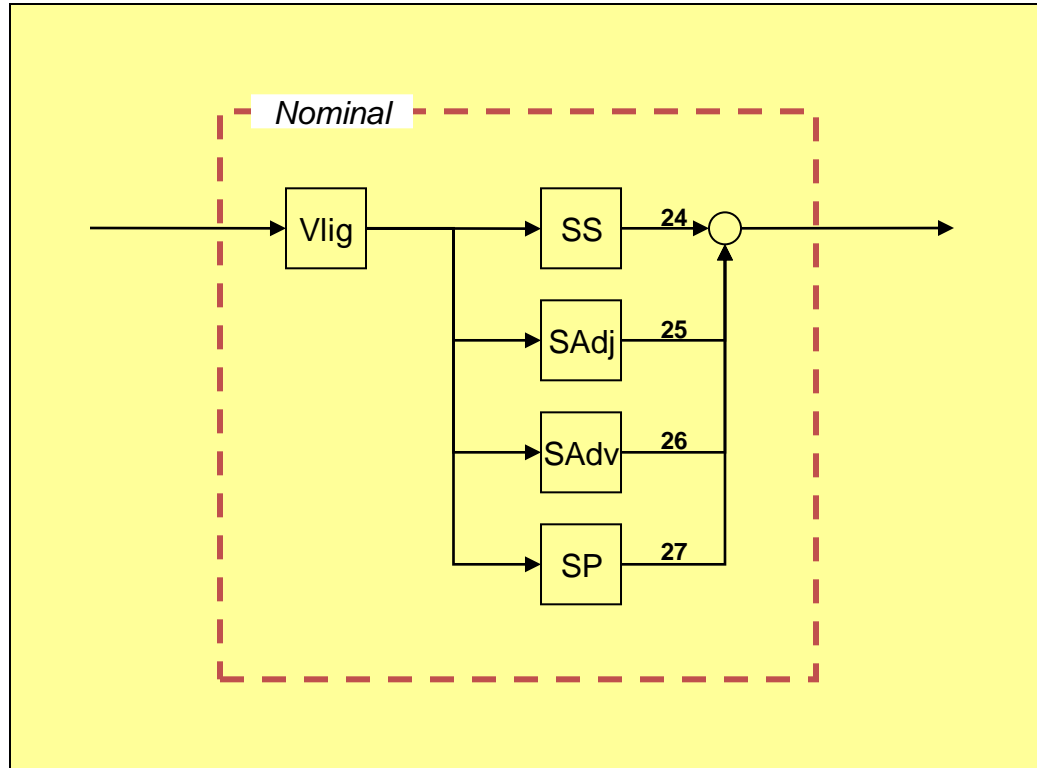
- Nominais
 - Atributivos de Predicado Nominal
- Verbais
 - Atributivos de Predicado Verbal
- Verbo-nominais
 - Estruturas Derivadas Complexas de Predicado Verbo-Nominal

Padrões Pessoais Nominais

- Atributivos de Predicado Nominal

1	SS1 Vlig SS2	
2	SS Vlig SAdj	
3	SS Vlig SAdv	SS Vlig SAdj
4	SS Vlig SP	SS Vlig SAdj
5	SS Vlig SP	SS Vlig SAdv

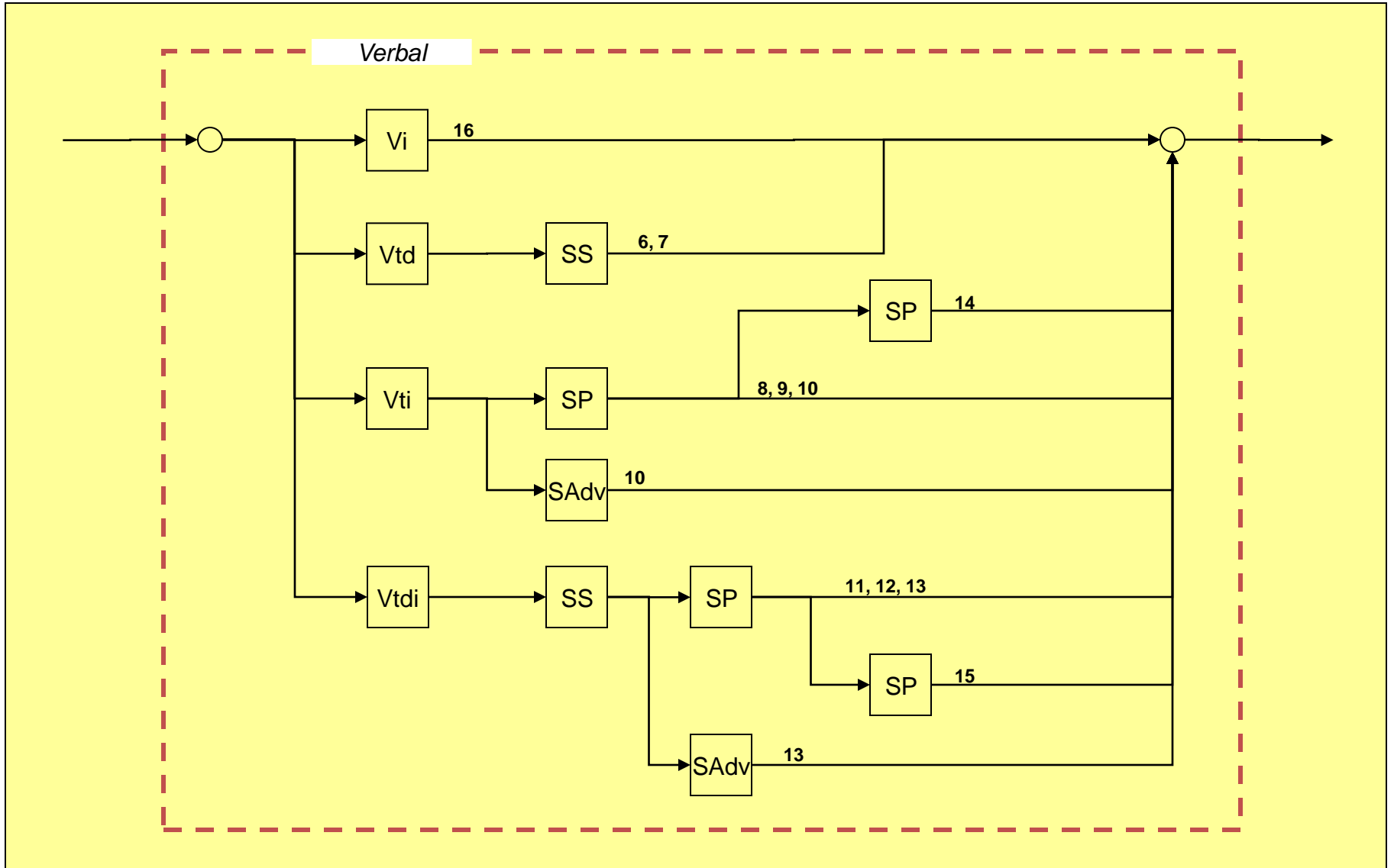
Nominal



Padrões Pessoais Verbais

6	SS1 Vtda SS2		
7	SS1 Vtdna SS2		Acusativo: Pronome Reto
8	SS Vti SP		Dativo: Pronome Oblíquo
9	SS Vti SP		Regência Verbal c/ Preposição
10	SS Vti SP	SS Vti SAdv	Locativo
11	SS1 Vtdi SS2 SP		Dativo
12	SS1 Vtdi SS2 SP		Regência Verbal c/ Preposição
13	SS Vtdi SS2 SP	SS Vtdi SS2 SAdv	Locativo
14	SS Vti SP1 SP2		
15	SS Vtdi SS2 SP1 SP2		
16	SS Vi		

Verbal

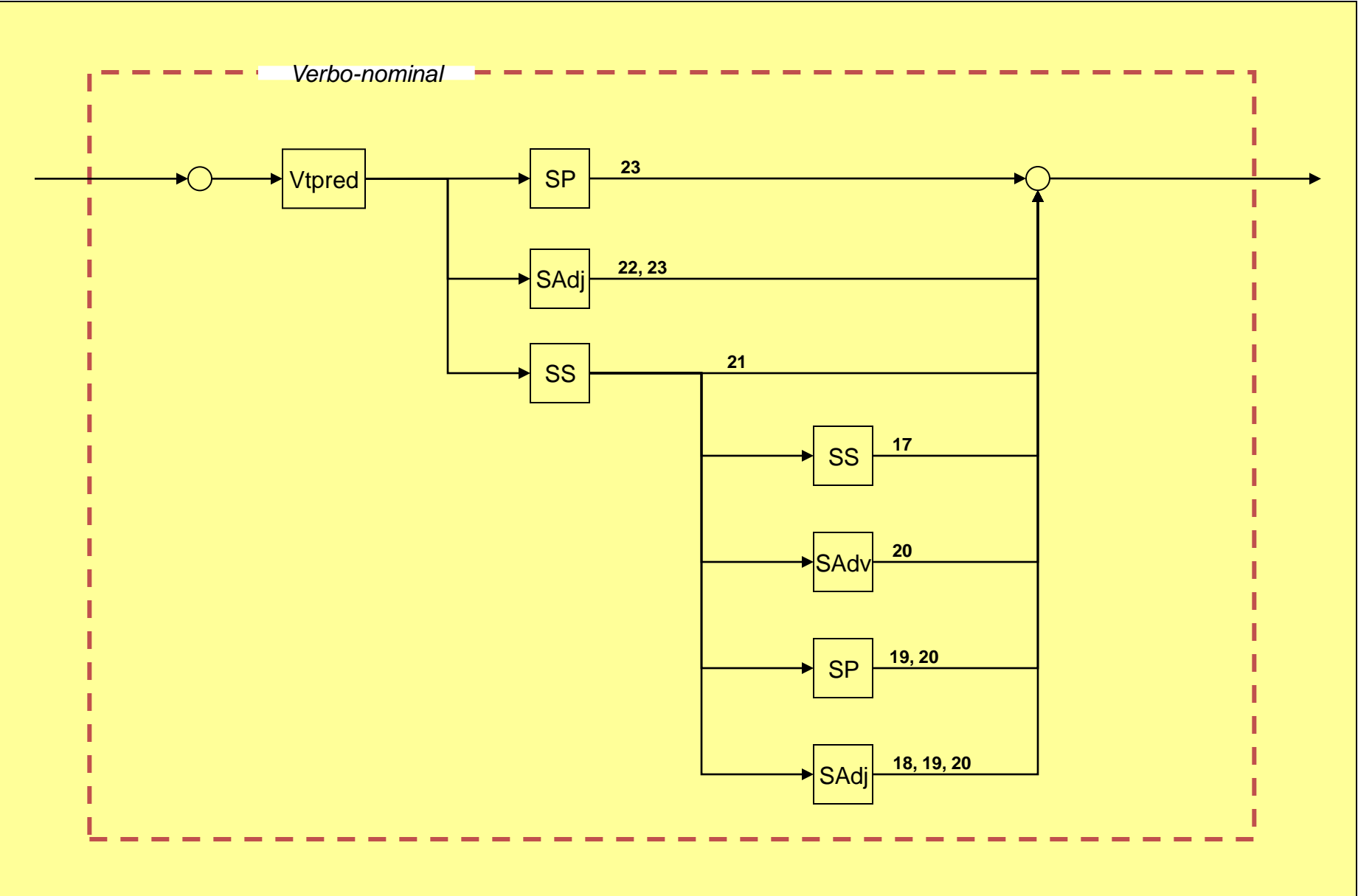


Padrões Pessoais Verbo-nominais

- Estruturas Derivadas Complexas de Predicado Verbo-Nominal (a + b)

17	SS1 Vt-pred SS2 SS3		6+1
18	SS1 Vt-pred SS2 SAdj		6+2
19	SS1 Vt-pred SS2 SP	SS1 Vt-pred SS2 SAdj	6+4
20	SS1 Vt-pred SS2 SP	SS1 Vt-pred SS2 SAdv	6+5
21	SS1 Vt-pred SS2		14+1
22	SS1 Vt-pred SAdj		
23	SS1 Vt-pred SP	SS1 Vt-pred SAdj	

Verbo-nominal



Fórmula geral (do livro)

[SS] V [SS] [SS | Sadj | Sadv | SP] [SP] [SP]

F → T O

- T: tipo
- O: oração

O → SN SV [SP]

SN → [Det] [Mod] N [Mod] | Pro | ε

Det → [Pré-Det] Det-base [Pós-Det]

- Det: determinante
- Pré-Det: quantificadores e partitivos anteriores ao determinante
- Det-base: artigo e demonstrativo (base do determinante)
- Pós-Det: numerais e possessivos posteriores ao determinante

Mod → Sadj | SP

- Mod: modificador

SP → Prep SN | Adv

Padrões cobertos pelo diagrama geral de sintaxe

Na sequência, são apresentadas formas sentenciais de diversos tipos, que podem ser identificadas e reconhecidas pelo autômato anteriormente esboçado a partir da sintaxe geral proposta.

Adotou-se a seguinte simbologia:

X, Y, Z	=	sintagmas substantivos
A, B	=	sintagmas adjetivos
D	=	sintagma adverbial
L, M, N	=	locativos
(12)	=	número entre parênteses identifica o exemplo do livro
8-9-10	=	números indicam as decisões possíveis no caso (ambigüidade / não-determinismo)

Padrões cobertos pelo diagrama geral de sintaxe

• X é Y	(1)	1
• X é A	(2)	2-3-4
• X está A	(3)	2-3-4
• X está D	(3)	3-5
• X está com Y	(4)	4-5
• X é B	(4)	2-3-4
• X está D	(5)	3-5
• X respeita Y	(6)	6-7
• X tem Y	(7)	6-7
• X obedece a Y	(8)	8,9,10
• X depende de Y	(9)	8-9-10
• X vai D	(10)	10
• X mora em Y	(10)	8-9-10
• X entrega Y a Z	(11)	11-12-13
• X recebe Y de Z	(12)	11-12-13
• X coloca Y em Z	(13)	11-12-13
• X coloca Y contra Z	(13)	11-12-13
• X vai de Y a Z	(14)	14
• X volta de Y para Z	(14)	14
• X transporta Y de L para M	(15)	15



Padrões cobertos pelo diagrama geral de sintaxe

• X dorme	(16)	16
• X considera Y Z	(17)	17
• X acha Y A	(18)	18-19-20
• X encontra Y com Z	(19)	19-20
• X encontra Y A	(19)	18-19-20
• X julga Y L	(20)	20
• X supunha Y M	(20)	20
• X voltou Y	(21)	21
• X saiu B	(22)	22-23
• X dorme A	(22)	22-23
• X voltou com Y	(23)	23
• É X	(24)	24
• Está Y	(24)	24
• Está A	(25)	25
• É D	(26)	26
• É de X	(27)	27
• Houve X	(28)	28
• Passa de X	(29)	29
• Chove	(30)	30

Algumas obras, para consulta e referência

- Não há ainda um consenso sobre padrão de terminologia ou de formato para a representação da sintaxe da língua portuguesa.
- Uma referência interessante sobre conceitos gerais de lingüística é:
 - Edward Lopes – *Fundamentos da Lingüística Contemporânea* – 18ª. Edição, 2003
- Um manual da sintaxe da língua portuguesa se encontra em:
 - Carlos Miotto, Maria Cristina Figueiredo Silva, Ruth Elizabeth Vasconcellos Lopes – *Novo Manual de Sintaxe* – 3ª. Edição, Editora Insular, 2007.
- Há muitas referências sobre a gramática da língua portuguesa, a maioria das quais a descreve discursivamente. Só como exemplos:
 - Napoleão Mendes de Almeida – *Gramática Metódica da Língua Portuguesa* – 39ª. Edição, Editora Saraiva, 1994
 - Carlos Henrique da Rocha Lima – *Gramática Normativa da Língua Portuguesa* – 33ª. Edição, Livraria José Olympio Editora, 1996.
- As duas referências seguintes apresentam a sintaxe do português em um formato bastante compatível com as representações computacionais baseadas em formulações adaptativas:
 - Celso Pedro Luft – *Moderna Gramática Brasileira* – Editora Globo, 2002
 - M.Cecília P. de Souza e Silva; Ingedore Villaça Koch – *Lingüística Aplicada ao Português: Sintaxe* – 14ª. Edição, Cortez Editora, 2007

6. Análise Léxica

Carlos Eduardo Dantas de Menezes

**Um método para a construção de analisadores
morfológicos, aplicado à língua portuguesa,
baseado em autômatos adaptativos**

Dissertação de Mestrado, EPUSP, 2000

7. Conversão Texto-Voz

Danilo Picagli Shibata

Tradução Grafema-Fonema para a Língua Portuguesa Baseada em Autômatos Adaptativos

Dissertação de Mestrado, EPUSP, 2008.

8. Gramática Adaptativa

Margarete Keiko Iwai

**Um formalismo gramatical adaptativo para
linguagens dependentes de contexto.**

Tese de Doutorado, EPUSP, 2000

9. Comparação de Formalismos

Célia Yumi Okano Taniwaki

**Formalismos adaptativos na análise sintática de
linguagem natural**

Dissertação de Mestrado, EPUSP, 2001.

10. Inferência Gramatical

Ivone Penque Matsuno

**Um Estudo do Processo de Inferência de Gramáticas
Regulares e Livres de Contexto Baseado em Modelos
Adaptativos**

Dissertação de Mestrado, EPUSP, 2006.

11. Sintaxe dependente de contexto

Miryam de Moraes

Alguns aspectos de tratamento de dependências de contexto em linguagem natural empregando tecnologia adaptativa

Tese de Doutorado, EPUSP, 2006.

12. Semântica

Lúri Chaer

**Um estudo sobre a Teoria da Predição aplicada à
análise semântica de Linguagens Naturais**

Dissertação de Mestrado, EPUSP, 2010.

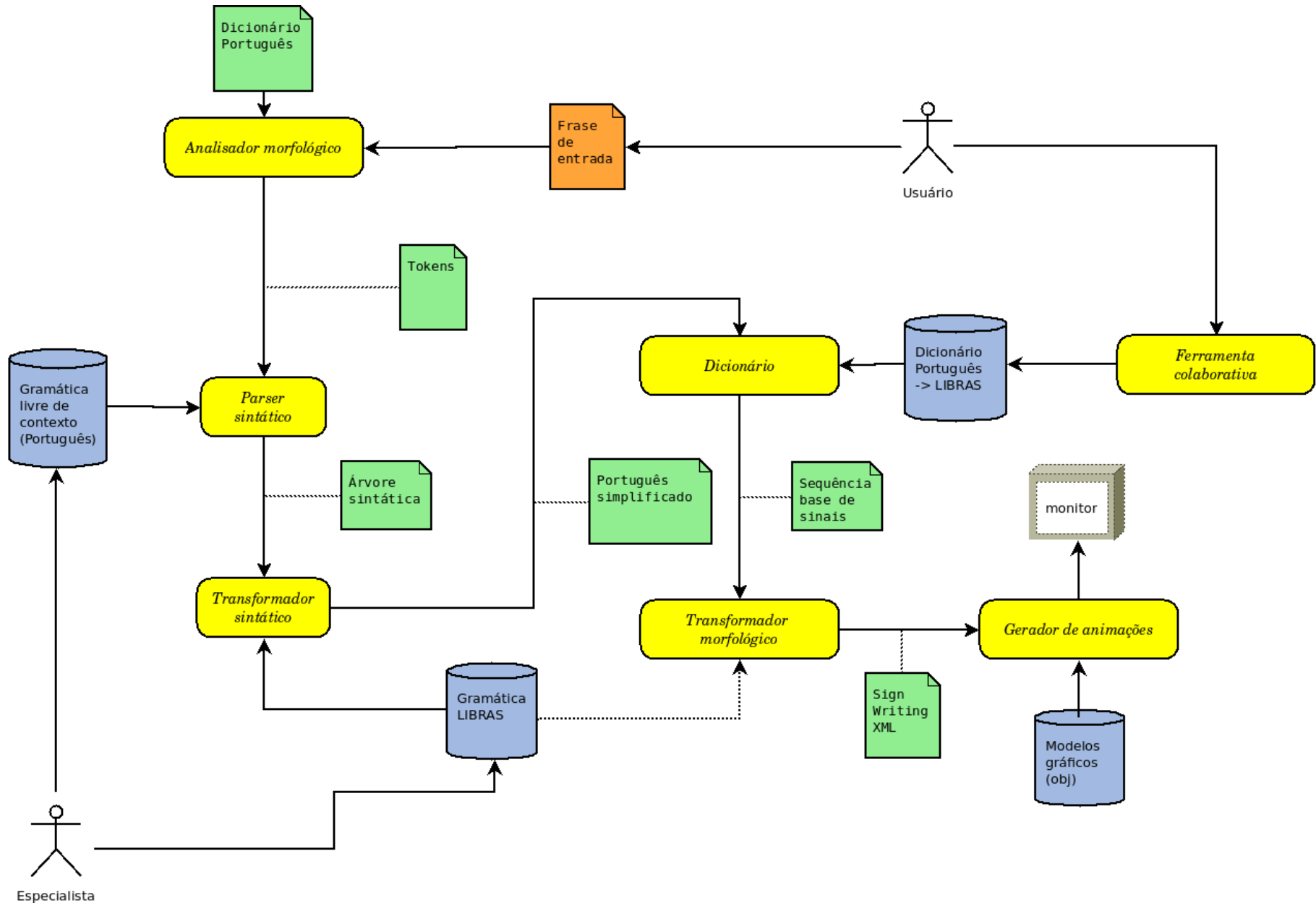
13. Tradução Português - LIBRAS

Guilherme Januário, Leonardo Leite, Marcelo Koga

Sistema Poli-Libras

Trabalho de Conclusão de Curso, EPUSP, 2010.

Sistema POLI-LIBRAS



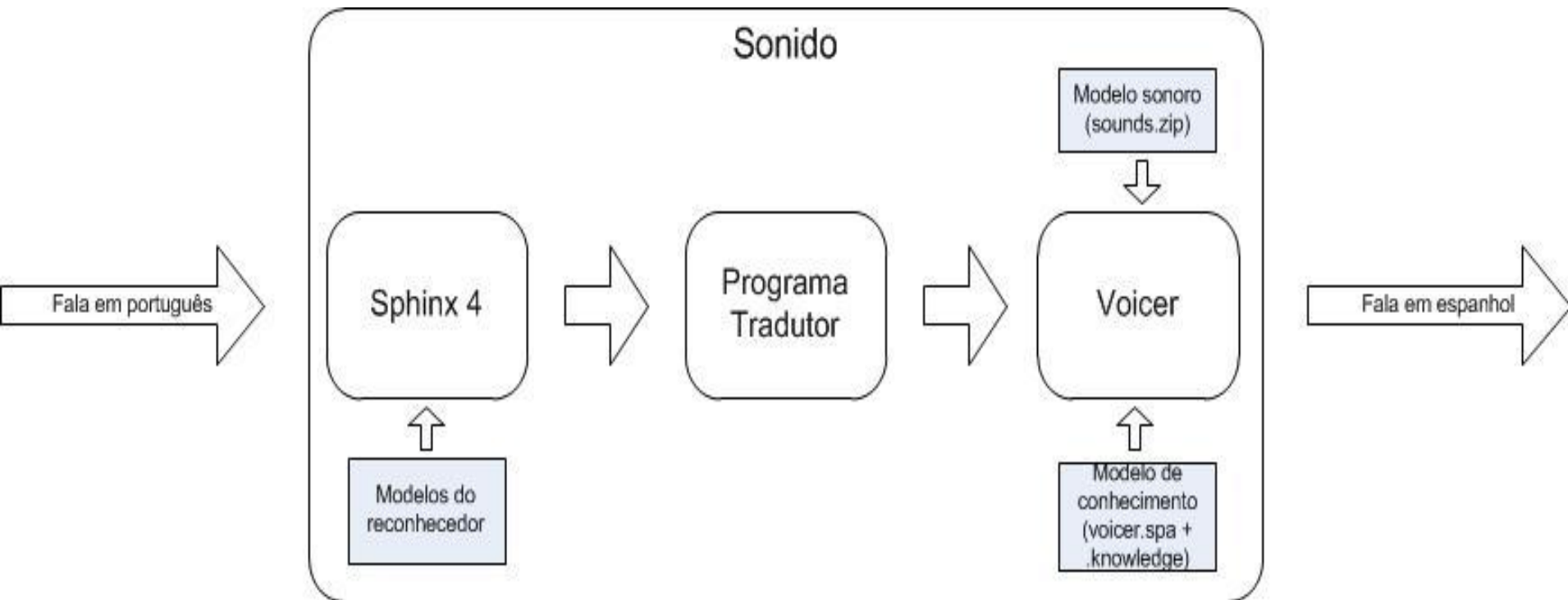
14. Tradução Português-Espanhol

Eduardo Shimizu, Pedro Massolino, Ronaldo Miura,
Wilson Martins Neto

Sistema SONIDO

Trabalho de Conclusão de Curso, EPUSP, 2010.

Arquitetura do sistema Sonido



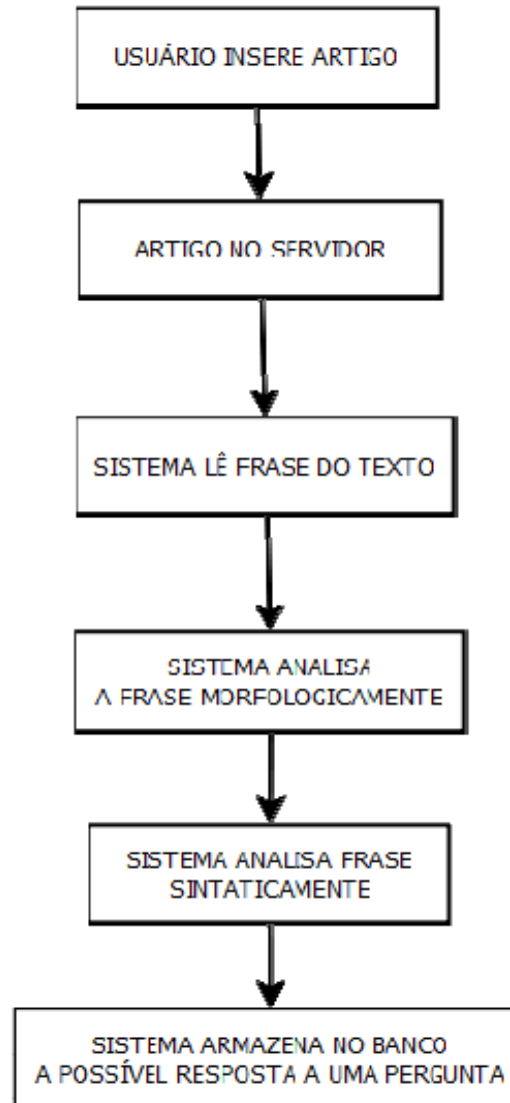
15. Busca em linguagem natural

Alencar Cerqueira, André Diniz, Carolina Dorta, Paulo Kuniyoshi

Sistema BUSCADOR

Trabalho de Conclusão de Curso, EPUSP, 2010.

Buscador



CONVITE

TODOS ESTÃO CONVIDADOS PARA O

QUINTO WORKSHOP

DE TECNOLOGIA ADAPTATIVA

– WTA-2011 –

A SER REALIZADO NOS DIAS

31 DE JANEIRO E 01 DE FEVEREIRO DE 2011,

NA EPUSP

João José Neto

joao.jose@poli.usp.br

www.pcs.usp.br/~Ita

OBRIGADO.