Towards A Tableaux Machine for Language Understanding

MICHAEL KOHLHASE & ALEXANDER KOLLER

Computer Science & Computational Linguistics
Universität des Saarlandes, Germany
http://www.ags.uni-sb.de/~kohlhase
http://www.coli.uni-sb.de/~koller

With anaphora, you could get pretty much anywhere, as long as someone's still attending to them.

B. Webber, 29. July 2000

Model Generation and Natural Language

- MG is becoming more and more popular in natural language semantics
- NL understanding as generating models for logical form of discourse.
  (transporting a "mental model" from the speaker's mind to the hearer's.)
  - + inference-based process that integrates world knowledge
  - – constructed model grows monotonically with the discourse
    (leads to combinatorial explosion)
- Humans cope with this by
  - concentrating on few salient discourse entities.
  - spending only limited ressources on inference in NL processing.
- Idea: Integrate this into model generation and test on NL.

Model Generation (recap)

Tableau Calculus

$$\neg(p(a) \vee p(b) \wedge \forall x.\neg p(x))^{\mathsf{f}}$$
$$p(a) \vee p(b) \wedge \forall x.\neg p(x)^{\mathsf{t}}$$
$$p(a) \vee p(b)^{\mathsf{t}}$$
$$\forall x.\neg p(x)^{\mathsf{t}}$$
$$\begin{array}{c|c} p(a)^{\mathsf{t}} & p(b)^{\mathsf{t}} \\ \neg p(y)^{\mathsf{t}} & \neg p(z)^{\mathsf{t}} \\ p(y)^{\mathsf{f}} & p(z)^{\mathsf{f}} \\ * \, [y = a] & * \, [z = b] \end{array}$$

No Model

Model generation

$$(a \wedge (b \vee \vee \neg c)) \wedge \neg b^{\mathsf{t}}$$
$$(a \wedge (b \vee \vee \neg c))^{\mathsf{t}}$$
$$\neg b^{\mathsf{t}}$$
$$b^{\mathsf{f}}$$
$$a^{\mathsf{t}}$$
$$b \vee \neg c^{\mathsf{t}}$$
$$\begin{array}{c|c} b^{\mathsf{t}} & \\ & \neg c^{\mathsf{t}} \\ \perp & \end{array}$$
$$c^{\mathsf{f}}$$

Model $\{a^{\mathsf{t}}, b^{\mathsf{f}}, c^{\mathsf{f}}\}$

Recap Tableaux (Propositional)

- formula is analyzed in a tree to determine satisfiability
- branches correspond to valuations (models)
- Connective rules of tableau calculi

$$\frac{\mathbf{A} \wedge \mathbf{B}^{\mathsf{t}}}{\begin{array}{c}\mathbf{A}^{\mathsf{t}}\\ \mathbf{B}^{\mathsf{t}}\end{array}} \qquad \frac{\mathbf{A} \wedge \mathbf{B}^{\mathsf{f}}}{\mathbf{A}^{\mathsf{f}} \mid \mathbf{B}^{\mathsf{f}}} \qquad \frac{\neg \mathbf{A}^{\mathsf{t}}}{\mathbf{A}^{\mathsf{f}}} \qquad \frac{\neg \mathbf{A}^{\mathsf{f}}}{\mathbf{A}^{\mathsf{t}}} \qquad \frac{\begin{array}{c}\mathbf{A}^{\alpha}\\ \mathbf{A}^{\beta}\end{array} \quad \alpha \neq \beta}{\perp}$$

Recap: Free Variable Tableaux

- Quantifier rules

$$\frac{\forall X.\mathbf{A}^{\mathsf{t}} \quad Y new}{[Y/X]\mathbf{A}^{\mathsf{t}}} \qquad \frac{\forall X.\mathbf{A}^{\mathsf{f}} \quad \mathbf{free}(\mathbf{A}) = \{Y, X_1, \ldots, X_n\} \quad f new}{[f(X_1, \ldots, X_n)/Y]\mathbf{A}^{\mathsf{f}}}$$

  - $\forall^{\mathsf{t}}$-rule introduces new (meta)-variable, $\forall^{\mathsf{f}}$ Skolemizes
  - Generalized cut rule: $$\frac{\begin{array}{l}\mathbf{A}^{\alpha} \\ \mathbf{B}^{\beta}\end{array} \quad \alpha \neq \beta}{\perp} \sigma = mgu(\mathbf{A}, \mathbf{B})$$

    instantiates the whole tableau by σ.

Recap Model Generation (RM Calculus [Konrad'98])

- Ground quantifier rules

$$\frac{\forall X.\mathbf{A}^{\mathsf{t}} \quad c \in \mathcal{H}}{[c/X]\mathbf{A}^{\mathsf{t}}} \qquad \frac{\forall X.\mathbf{A}^{\mathsf{f}} \quad \mathcal{H} = \{a_1, \ldots, a_n\}}{[a_1/X]\mathbf{A}^{\mathsf{f}} \Big| \ldots \Big| [a_n/X]\mathbf{A}^{\mathsf{f}} \Big| [c_{new}/X]\mathbf{A}^{\mathsf{f}}}$$

  - $\forall^{\mathsf{f}}$-rule introduces new constant c_{new} to Herbrand base $\mathcal{H}$ of branch
  - Apply $\forall^{\mathsf{t}}$ exhaustively (for c_{new} reapply all $\forall^{\mathsf{t}}$ rules on branch!)

Application in NLU: Rabbits [Gardent & Konrad '99]

- Interpret "the" as $\lambda P, Q.\mathsf{uniq}(P) \wedge \mathsf{every}(P, Q)$

  where $\mathsf{uniq}{:} = \lambda P.\exists X.P(X) \wedge \forall Y\, P(Y) \Rightarrow X = Y$

  and $\mathsf{every}{:} = \lambda P, Q.\forall X\, P(X) \Rightarrow Q(X)$.

- "the rabbit is cute", has logical form $\mathsf{uniq}(\mathsf{rab}) \wedge \mathsf{rab} \subseteq \mathsf{cute}$.

- $\mathcal{RM}$ generates $\{\ldots, \mathsf{rab}(c), \mathsf{cute}(c)\}$ in situations with at most one rabbit.

  (special $\mathcal{RM}(\exists)$ rule yields identification and accommodation (c^{new}))

+ At last an approach that takes world knowledge into account!

– tractable only for toy discourses/ontologies

  The world cup final was watched on TV by 7 million people. A rabbit is in the garden.

  $\forall X.\mathsf{human}(x)\exists Y\,\mathsf{human}(X) \wedge \mathsf{father}(X, Y)$ $\qquad \forall X, Y.\mathsf{father}(X, Y) \Rightarrow X \neq Y$

Abductive Interpretation of NL [Stickel89,Hobbs93]

- Example: . . . The car is red. . . . (Logical form: . . . $car(x) \wedge red(x)$. . .)
- Explain the logical form (explanandum) from the common ground $\mathcal{T}$ and extend $\mathcal{T}$ by the explanations.
  - If $car(a) \in \mathcal{T}$, then $\mathcal{T} \models_{PL1} red(a) \rightsquigarrow \mathcal{F}$
  - If $red(b) \in \mathcal{T}$, then $\mathcal{T} \models_{PL1} car(b) \rightsquigarrow \mathcal{F}$
  - else $\mathcal{T} \models_{PL1} red(c), car(c) \rightsquigarrow \mathcal{F}$
- calculate most general explanation???
  - [Freud32] Sometimes a cigar is just a cigar!
- Warning: $\mathcal{T} \models_{PL1} \emptyset \rightsquigarrow \mathcal{F}$ (Nothing new has to be abduced) if $fire\_engine(x) \Rightarrow red(x) \wedge car(x) \in \mathcal{T}$ and $fire\_engine(f) \in \mathcal{T}$

Weighted Abduction

- There are terribly many explanations in general.
- Any definition, which one is really the best, will always be subjective.
- Idea: [Stickel89] assign costs and search for the "least expensive" one
- Example: ...The car is red. ... ($\mathcal{F} = car(x)^{\$10} \wedge red(x)^{\$1}$)
  The car is definite, therefore more expensive
  - If $car(a) \in \mathcal{T}$, then $\mathcal{T} \models_{PL1} red(a) \leadsto^{\$1} \mathcal{F}$
  - If $red(b) \in \mathcal{T}$, then $\mathcal{T} \models_{PL1} car(b) \leadsto^{\$10} \mathcal{F}$
  - otherwise $\mathcal{T} \models_{PL1} red(c), car(c) \leadsto^{\$11} \mathcal{F}$
- prefer the second explanation (cheaper)
- What do the costs actually model?

Contribution and Plot of the Paper

- Build on the intuitions and techniques from model generation and abduction approaches to NLU
- Refine them to get around the problems of
  - Model Explosion (Saliences for Individuals that decay)
  - Computational Intractability/Omniscience (Rule Costs)
- Architecture: NLU as an incremental process of
  - syntactic analysis, (interprets sentences one-by-one)
  - semantics construction, (generates logical forms passes to)
  - the tableaux machine. (MG + ATP + non-local control)

Salience: Syntactic Criteria

- Definition: Salience is a measure of the availability of an entity in the discourse universe for (anaphoric or definite) reference
- Our theory of salience (inspired by [Alshawi 87, Lappin&Leass94])
  - approximated by syntactic criteria, (e.g. NP type, role)
  - decays over time, (halved after each sentence)
  - raised whenever the entity is referred to (add saliences).
- syn/sem interface annotates salience to constants and quantifiers.
- Example: Mary chases a rabbit.
  $\exists X_{60}.\mathsf{rab}(X) \wedge \mathsf{chase}(\mathsf{mary}_{100}, X)$ (object $\prec$ subject)
- tableau branch has a salience function for each salience level.

A refined Model Generation Calculus RAMG

$$\frac{a, b \in \mathcal{H}}{a = b^{\mathsf{t}} \,\big|\, a \neq b^{\mathsf{t}}}\mathcal{R}(id)^{1} \qquad \frac{}{(\mathbf{A} \vee \neg\mathbf{A})^{\mathsf{t}}}\mathcal{R}(tnd)$$

$$\frac{(\forall X.\mathbf{A})^{\mathsf{t}} \quad a \in \mathcal{H}}{[a/X]\mathbf{A}^{\mathsf{t}}}\mathcal{R}(\forall) \qquad \frac{(\forall X_{S}.\mathbf{A})^{\mathsf{f}} \quad c^{new} \notin \mathcal{H}}{[c^{new}/X]\mathbf{A}^{\mathsf{f}}}\mathcal{R}(\exists)^{2}$$

+ standard equality rules $\mathcal{R}(=), \mathcal{R}(sym), \mathcal{R}(refl), \mathcal{R}(trans)$

[1] The saliences of a and b are replaced by their sum in the left branch.

[2] c is introduced with salience S.

- selective quantifier rules to restrict domain extensions

Chaining with World Knowledge

- The following rule is admissible in RAMG

$$\frac{\begin{array}{c}\forall X^1 \mathbf{A}^1 \Rightarrow \ldots \forall X^n \mathbf{A}^n \Rightarrow \exists Y_S.\mathbf{B}^{\mathsf{t}} \\ \mathbf{A}^1(a^1)^{\mathsf{t}} \\ \ldots \\ \mathbf{A}^n(a^1, \ldots, a^n)^{\mathsf{t}} \qquad c \notin \mathcal{H}\end{array}}{\mathbf{B}[a^i/X^i, c/Y]^{\mathsf{t}}} \mathcal{R}(\Rightarrow)^1$$

[1] c is introduced with salience $S(a^1, \ldots, a^n)\downarrow$.

- Note: Uses salience expressions on existential variables
  - S is an arithmetic expression in the universal variables
  - compute $S\downarrow$ by replacing the constants in S by their current saliences (all variables in S have been instantiated at that time)

14 ©: Michael Kohlhase

Example (Bridging Inferences)

- Mary could not start her car. The battery was empty.
- Logical form: $\mathbf{S}^1 = \exists X_{60}.\mathsf{car}(X) \wedge \mathsf{own}(\mathsf{mary}_{100}, X) \wedge \neg\mathsf{start}(X)$
- World Knowledge: $\mathbf{R} = \forall X.\mathsf{car}(X) \Rightarrow \exists Y_{X/2}\mathsf{bat}(Y) \wedge \mathsf{part\_of}(Y, X)$

#	formula	Justification		salience	
		by	from	level	*value*
0	$\mathbf{R}^{\mathsf{t}}$	*World*			
1	$\mathbf{S}^{1\,\mathsf{t}}$	*Given*			
2	$\mathsf{car}(c)^{\mathsf{t}}$	$\mathcal{R}(\exists)$	1	1	$c : 60$
3	$\mathsf{own}(c, \mathsf{mary})^{\mathsf{t}}$	$\mathcal{R}(\exists)$	1	1	$\mathsf{mary} : 100$
4	$\mathsf{start}(c)^{\mathsf{f}}$	$\mathcal{R}(\exists)$	1	1	
5	$\mathsf{bat}(d)^{\mathsf{t}}$	$\mathcal{R}(\Rightarrow)$	0, 2	1	$d : 30$
6	$\mathsf{part\_of}(d, c)^{\mathsf{t}}$	$\mathcal{R}(\Rightarrow)$	0, 2	1	
7	$\mathbf{S}^{2\,\mathsf{t}}$	*Given*		2	$c : 30$
				2	$\mathsf{mary} : 50$
	...	...		2	$d : 15$

Salience Ordering

- The battery is resolved to d, since it is the most salient battery (had Mary been a robot, she might have had a more salient battery)
- To speak about relative saliences, assume a salience ordering
- $c \prec d$ is true iff the salience of c is much lower than that of d at the current salience level;
  (if $c : n$ and $d : m$ and $n \geq (1 + \delta)m$ for some predefined threshold $\delta \in \mathbb{R}$.)
  - relative saliences are invariant under salience decay,
    (can only be changed by inference)
- New RAMG rule:
  $\mathcal{R}(\prec)$ closes a branch ending $c \prec d^{\mathsf{t}}$ if $c \prec d$ is false, (and dual)

Tableaux Machine

- takes the logical forms (with salience expressions) as input,
- adds them to all open branches,
- performs RAMG inferences until some ressource criterion is met
- output is application dependent; some choices are
  - the preferred model given as all the (positive) literals of the preferred branch;
  - the literals augmented with all non-expanded formulae (from the discourse); (ressource-bound was reached)
  - machine answers user queries (preferred model $\models$ query?)
- model generation mode (guided by ressources and strategies)
- theorem proving mode ($\Box$ for side conditions; using tableau rules)

Model Generation Mode

- each proof rule comes with rule costs.
  - Ultimately we want bounded optimization regime [Russell'91]:
    expansion as long as expected gain in model quality outweighs proof costs
  - here: each sentence in the discourse has a fixed inference budget
    Expansion until budget used up.
- Effect: Expensive rules are rarely applied.
- Warning: Finding appropriate values for the rule costs is a major open problem of our approach.

Rule costs and Strategies

- Rule costs qualitatively:

cheap	medium	expensive
$\mathcal{T}(\wedge)$, $\mathcal{T}(\vee)$, $\mathcal{T}(\neg)$, $\mathcal{R}(\exists)$, $\mathcal{R}(\Rightarrow)$	$\mathcal{R}(\forall)$	$\mathcal{R}(tnd)$, $\mathcal{R}(id)$

- Rule costs depend on salience of constants involved,
  (low salience $\Rightarrow$ high rule cost)
- Strategies: (override other control)
  - apply $\mathcal{R}(\exists), \mathcal{R}(\Rightarrow)$ (keep discourse universe up to date)
  - error recovery for failed sub-tableaux (see accommodation)
  - folding up (add inference result high in the salience level)

Case Study: Definites

- Idea: Extend RAMG with a new quantifier $\tau(P, Q, S)$ (new rule $\mathcal{R}(\tau)$)
  - syn/sem interface produces this for "the Q P", with salience S
  - $\tau(P, Q, S)$ denotes one constant in the Herbrand universe which?
  - $\tau(P, Q, S)$ must be unique or most salient object in P
  - set the cost of $\mathcal{R}(\tau)$ to zero (expand τ formulas eagerly)
- Preview
  - The largest state of the US has rich oil reserves.
  - Joe Smith gave a press conference today.
    The CEO of Westinghouse announced budget cuts.
  - The president of every NY City authority is a crook

$$\frac{\tau(P,Q,S)^{\alpha} \quad \mathcal{H}=\{a^1,\ldots,a^n\}}{\begin{array}{c|c|c} Q(a^1)^{\mathsf{t}} & & Q(a^n)^{\mathsf{t}} \\ \boxed{\mathsf{uniq}_{\prec}(a^1,Q)^{\mathsf{f}}} \quad P(a^1)^{\alpha} & \ldots & \boxed{\mathsf{uniq}_{\prec}(a^n,Q)^{\mathsf{f}}} \quad P(a^n)^{\alpha} \end{array}} \mathcal{R}(\tau)^1$$

[1] For each i, the salience of a^i is raised by $S{\downarrow}$ in the i-th branch.

- Consider an arbitrary branch of the rule, say for a^i.
  - closes immediately if $\neg Q(a^i)$ in branch, (know $a^i \notin Q$)
  - otherwise try to close $\Box$ with

$$\mathsf{uniq}_{\prec}(a^i,Q) := \bigwedge_{c\in\mathcal{H}} Qc \Rightarrow a^i = c \vee c \prec a^i$$

i.e. prove for each $b \in \mathcal{H} \cap Q$ that either $b = a^i$ or b much less salient.

The largest state of the US has rich oil reserves.

$$\exists X.\mathsf{lst}(X) \wedge \forall Y.\mathsf{lst}(Y) \Rightarrow X{=}Y^{\mathsf{t}}$$
$$\mathsf{lst}(\mathsf{a})^{\mathsf{t}}$$
$$\forall Y.\mathsf{lst}(Y) \Rightarrow \mathsf{a}{=}Y^{\mathsf{t}}$$
$$\tau(\mathsf{hasror}, \mathsf{lst}, 100)^{\mathsf{t}}$$

$\boxed{1}$ $\mathsf{lst}(\mathsf{a})^{\mathsf{t}}$, $\mathsf{hasror}(\mathsf{a})^{\mathsf{t}}$	$\boxed{2}$ $\mathsf{lst}(\mathsf{b})^{\mathsf{t}}$, $\mathsf{hasror}(\mathsf{b})^{\mathsf{t}}$, $*$

$\mathsf{a}, \mathsf{b} : 100$

$\boxed{1} :=$

$$\mathsf{uniq}_{\prec}(\mathsf{a}, \mathsf{lst})^{\mathsf{f}}$$

$\mathsf{lst}(\mathsf{a}) \Rightarrow \mathsf{a}{=}\mathsf{a} \vee \mathsf{a} \prec \mathsf{a}^{\mathsf{f}}$	$\mathsf{lst}(\mathsf{b}) \Rightarrow \mathsf{a}{=}\mathsf{b} \vee \mathsf{b} \prec \mathsf{a}^{\mathsf{f}}$
$\mathsf{lst}(\mathsf{a})^{\mathsf{t}}$	$\mathsf{lst}(\mathsf{b})^{\mathsf{t}}$
$\mathsf{a}{=}\mathsf{a}^{\mathsf{f}}$	$\mathsf{a}{=}\mathsf{b}^{\mathsf{f}}$
$\mathsf{a} \prec \mathsf{a}^{\mathsf{f}}$	$\mathsf{b} \prec \mathsf{a}^{\mathsf{f}}$
$*$	$\mathsf{a}{=}\mathsf{b}^{\mathsf{t}}$
	$*$

$\boxed{2} :=$

$$\mathsf{uniq}_{\prec}(\mathsf{b}, \mathsf{lst})^{\mathsf{f}}$$

$\mathsf{lst}(\mathsf{a}) \Rightarrow \mathsf{b}{=}\mathsf{a} \vee \mathsf{a} \prec \mathsf{b}^{\mathsf{f}}$	$\mathsf{lst}(\mathsf{b}) \Rightarrow \mathsf{b}{=}\mathsf{b} \vee \mathsf{b} \prec \mathsf{b}^{\mathsf{f}}$
$\mathsf{lst}(\mathsf{a})^{\mathsf{t}}$	$\mathsf{lst}(\mathsf{b})^{\mathsf{t}}$
$\mathsf{b}{=}\mathsf{a}^{\mathsf{f}}$	$\mathsf{b}{=}\mathsf{b}^{\mathsf{f}}$
$\mathsf{a} \prec \mathsf{b}^{\mathsf{f}}$	$\mathsf{b} \prec \mathsf{b}^{\mathsf{f}}$
CAN'T CLOSE!	$*$

…Joe Smith gave a press conference today. The CEO of Westinghouse announced budget cuts.

$$\begin{array}{c|c}
\begin{array}{c}
\forall X.\mathsf{ceo}(X) \Rightarrow \mathsf{person}(X)^{\mathsf{t}} \\
\mathsf{person}(\mathsf{a})^{\mathsf{t}} \\
\mathsf{person}(\mathsf{j})^{\mathsf{t}} \\
\mathsf{gpc}(\mathsf{j})^{\mathsf{t}} \\
\tau(\mathsf{abc}, \mathsf{ceo}, 100)^{\mathsf{t}} \\
\hline
\begin{array}{c|c}
\begin{array}{c} \mathsf{ceo}(\mathsf{j})^{\mathsf{t}} \\ \boxed{1} \quad \mathsf{abc}(\mathsf{j})^{\mathsf{t}} \end{array} &
\begin{array}{c} \mathsf{ceo}(\mathsf{a})^{\mathsf{t}} \\ \boxed{2} \quad \mathsf{abc}(\mathsf{a})^{\mathsf{t}} \\ * \end{array}
\end{array}
\end{array}
&
\begin{array}{l}
\mathsf{j} : 100 \\
\mathsf{a} : 20
\end{array}
\end{array}$$

$\boxed{1} :=$

$$\begin{array}{c}
\mathsf{uniq}_{\prec}(\mathsf{j}, \mathsf{ceo})^{\mathsf{f}} \\
\hline
\begin{array}{c|c}
\begin{array}{c}
\mathsf{ceo}(\mathsf{j}) \Rightarrow \mathsf{j}{=}\mathsf{j} \vee \mathsf{j} \prec \mathsf{j}^{\mathsf{f}} \\
\mathsf{ceo}(\mathsf{j})^{\mathsf{t}} \\
\mathsf{j}{=}\mathsf{j}^{\mathsf{f}} \\
\mathsf{j} \prec \mathsf{j}^{\mathsf{f}} \\
*
\end{array} &
\begin{array}{c}
\mathsf{ceo}(\mathsf{a}) \Rightarrow \mathsf{j}{=}\mathsf{a} \vee \mathsf{a} \prec \mathsf{j}^{\mathsf{f}} \\
\mathsf{ceo}(\mathsf{a})^{\mathsf{t}} \\
\mathsf{j}{=}\mathsf{a}^{\mathsf{f}} \\
\mathsf{a} \prec \mathsf{j}^{\mathsf{f}} \\
*
\end{array}
\end{array}
\end{array}$$

$\boxed{2} :=$

$$\begin{array}{c}
\mathsf{uniq}_{\prec}(\mathsf{a}, \mathsf{ceo})^{\mathsf{f}} \\
\hline
\begin{array}{c|c}
\begin{array}{c}
\mathsf{ceo}(\mathsf{a}) \Rightarrow \mathsf{a}{=}\mathsf{a} \vee \mathsf{a} \prec \mathsf{a}^{\mathsf{f}} \\
\mathsf{ceo}(\mathsf{a})^{\mathsf{t}} \\
\mathsf{a}{=}\mathsf{a}^{\mathsf{f}} \\
\mathsf{a} \prec \mathsf{a}^{\mathsf{f}} \\
*
\end{array} &
\begin{array}{c}
\mathsf{ceo}(\mathsf{j}) \Rightarrow \mathsf{a}{=}\mathsf{j} \vee \mathsf{j} \prec \mathsf{a}^{\mathsf{f}} \\
\mathsf{ceo}(\mathsf{j})^{\mathsf{t}} \\
\mathsf{a}{=}\mathsf{j}^{\mathsf{f}} \\
\mathsf{j} \prec \mathsf{a}^{\mathsf{f}} \\
\text{CAN'T CLOSE!}
\end{array}
\end{array}
\end{array}$$

The president of every NY City authority is a crook

- ignore the strong reading (wait for accommodation)
- $\forall X.\mathsf{nyca}(X) \Rightarrow \tau(\lambda Y.\mathsf{presof}(Y, X), \mathsf{crook}, X/2)$
- expansion suspended until tableau contains a NYC authority ($\mathcal{R}(\forall)$ expensive)
- Once a constant a: 30 is a NYCA, apply a τ-chaining rule
- this adds $\tau(\lambda Y.\mathsf{presof}(Y, \mathsf{a}), \mathsf{crook}, 15)$
- apply the $\mathcal{R}(\tau)$ rule as above.

Accommodation: Mary feeds the rabbit in her garden.

- Problem: no referent for definite NP present in the discourse universe.
- without objects which could be rabbits $\mathcal{R}(\tau)$ fails immediately
- Idea: use an “error recovery” process
  - backtrack to immediately before the application of $\mathcal{R}(\tau)$;
  - apply $\mathcal{R}(tnd)$ once with the formula $\exists X_{60}.\mathsf{rihg}(X) \vee \neg\exists X.\mathsf{rihg}(X)$
    salience($\exists$) comes from τ-formula, here 60 (direct object)
  - continue as usual: Skolemization by $\mathcal{R}(\exists)$ on the positive branch leads successful application of $\mathcal{R}(\tau)$ (new, salient potential rabbit)
  - Nothing new happens on the negative branch.
- Accommodation via $\mathcal{R}(tnd)$ similar to [Hobbs&Stickel’93]
  (introduction of “helpful” formula, at a cost)

Related Work

- Weighted Abduction [Hobbs&Stickel'93]
  - RAMG distinguishes between object-level (saliences) and literal-level (costs) influences
  - Weighted Abduction an simulate saliences by increased weights (cannot distinguish between contribution of a, b in literal $p(a, b)$)
- salience theories for pronominal anaphora [Alshawi 87, Lappin&Leass94]
  - semantic theory of salience (interaction with world knowledge)
- other model generation approaches:
  - introduces rule costs and saliences (orthogonal device)

Conclusion

- Refined model generation for natural language understanding by
  - Saliences for Individuals that decay (model hygiene)
  - Rule Costs (computational tractability/partial ignorance)
  - Refutation Theorem Proving (e.g. uniqueness presuppositions)
  - Non-local control (e.g. accommodation)
- Yet another "theory" of definites

Open Questions

- elimination of remaining handwaving, in particular
- details of rule costs and saliences (e.g. spell out interaction)
- implementation of tableaux machine is indispensable for experimentation,
  - is there a consistent choice of costs at all?
  - can the new mechanisms really limit search space sizes?
- can we use learning techniques from theories of bounded rationality?

Perspectives

- Case study on definites is not a theory (yet?)
  - combination of MG, model-entailment, and salience consideration seems to cover a lot of ground
- Relation to dynamic semantics
  - MG dynamically introduces (discourse) referents (structural similarity)
  - saliences as alternative model for accessibility?
- Relation to psycholinguistics
  - "cognitive reality" of model generation? (mental models [Johnson-Laird])