Anaphora Resolution

Topic-wise Notes: Definition | Characteristics | Architecture | Step-by-Step Example | Applications

Topic	Title
1	Definition of Anaphora
2	Types of Anaphora
3	Constraints in Anaphora Resolution
4	Agreement Constraints
5	Syntactic and Semantic Constraints
6	Summary and exam questions

Background: In *'Anita lost her keys. She searched for them everywhere.'* the words **She**, **her** and **them** cannot be understood alone. They point back to **Anita** and **keys**. Finding what such words point to is **anaphora resolution**, one of the central tasks of discourse processing and a key component of coreference resolution.

Key terms: Anaphor = the dependent expression (she, it, the girl). **Antecedent** = the earlier expression it depends on. **Referent** = the real or discourse entity both denote. **Candidate** = any earlier expression that could be the antecedent. **Constraint** = a rule that rules candidates out. **Preference** = a heuristic that ranks the remaining candidates.

Topic 1: Definition of Anaphora

1.1 Definition

Anaphora is the linguistic relation in which the interpretation of an expression, the **anaphor**, depends on another expression in the context, its **antecedent** (Greek *anaphora*: 'carrying back'). **Anaphora resolution** is the task of automatically identifying, for each anaphor, the antecedent it refers back to.

Example: *'The **teacher** praised Anita because **she** had worked hard.'* The anaphor 'she' must be linked to 'Anita' (not 'the teacher', who is also female in some contexts, which is why resolution is non-trivial).

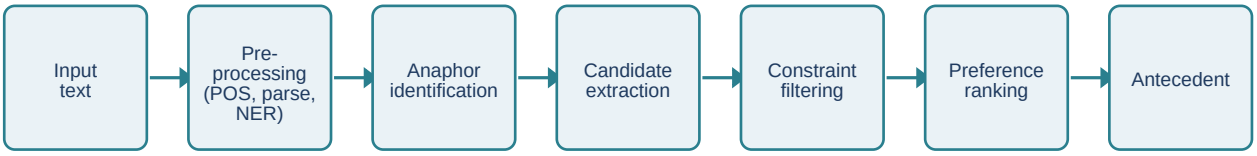
Term	Meaning	Example
Anaphora	Backward-pointing dependence on an earlier expression	'Ravi came. **He** sat.'
Cataphora	Forward-pointing dependence on a later expression	'Before **he** left, **Ravi** smiled.'
Coreference	Two expressions denote the same entity (symmetric)	'Barack Obama' and 'the former President'
Bridging	Inferable link without identity	'a car ... the engine'
Exophora	Referent outside the text (situation)	'Look at **that**!'

1.2 Characteristics

- **Context dependence:** an anaphor has little or no meaning of its own; the same word ('it') refers to different things in different places.
- **Resolved within the text** (endophoric), mostly backward, sometimes forward (cataphora).
- **Intra- or inter-sentential:** antecedent may be in the same sentence or an earlier one. Most pronoun antecedents lie in the current or previous sentence.
- **Ambiguity:** several candidates may satisfy the surface constraints ('Ravi told Mohan that he won').
- **Multiple knowledge sources:** morphological (agreement), syntactic (binding, grammatical role), semantic (selectional restrictions), discourse (salience, recency, focus) and world knowledge.
- **Not every pronoun is anaphoric:** pleonastic 'it' ('It is raining'), generic 'you', and exophoric uses must be excluded.
- **Forms chains:** successive anaphors with one referent form an anaphoric chain (Anita - she - her).
- **Language dependent:** English relies on pronouns; Hindi, Telugu, Japanese and Spanish also use zero pronouns and rely on verb agreement.

1.3 Architecture of an anaphora resolution system

Generic anaphora resolution architecture



Stage	Description
Pre-processing	Tokenise, POS-tag, parse (constituency or dependency), NER, morphological analysis, sentence splitting, optional chunking
Anaphor identification	Pick pronouns and other anaphoric expressions; remove pleonastic 'it' (patterns such as 'It is clear that...', weather verbs), generic and exophoric uses
Candidate extraction	Collect NPs from the current sentence and a window of previous sentences (typically 2-5); more for demonstratives and definite NPs
Constraint filtering	Discard candidates that violate agreement, syntactic (binding) and semantic constraints (Topics 3-5)
Preference ranking	Score the survivors using recency, grammatical role (subject > object), parallelism, repetition, centering, frequency, or a learned model
Output	Anaphor-antecedent links, merged into anaphoric chains

Approach family	Idea	Examples
Knowledge-rich / rule-based	Hand-crafted syntactic, semantic and discourse rules	Hobbs 1978, Lappin and Leass RAP 1994, centering algorithms
Knowledge-poor	Light preprocessing plus constraints and indicators, no deep parsing	Mitkov's MARS / robust approach
Statistical / machine learning	Learn from annotated corpora with feature vectors	Mention-pair, mention-ranking, entity-mention models
Neural / end-to-end	Learn span embeddings; antecedent scoring jointly with mention detection	Lee et al. 2017, BERT/SpanBERT-based coreference

1.4 Example, step by step

```
[1] The teacher praised Anita because she had worked hard. [2] The award surprised her.
```

Step	Action	Result
1. Pre-process	Tokenise, POS-tag, parse, NER	NPs: The teacher, Anita, she, The award, her. Anita = PERSON
2. Identify anaphors	Find pronouns, exclude pleonastic	she (S1), her (S2) are referential anaphors
3. Extract candidates	NPs before each anaphor	For 'she': The teacher, Anita. For 'her': The teacher, Anita, she, The award
4. Apply constraints	Agreement (3rd, sg, fem, animate), semantics	'The teacher' has unknown gender (compatible); Anita = fem (compatible); 'The award' = inanimate -> removed for 'her'

Step	Action	Result
5. Apply preferences	Causal reading of 'because ... had worked hard' (reason for being praised) plus world knowledge	'she' = Anita (the one praised and working hard); 'her' = Anita (salience: antecedent of 'she', object of praise)
6. Output	Anaphoric chain	Anita - she - her; The award separate entity

1.5 Applications

- **Machine translation:** pronouns must agree in gender and number in the target language.
- **Information extraction and question answering:** link facts to the right entity across sentences.
- **Text summarisation:** avoid dangling pronouns in extracted sentences.
- **Dialogue systems and voice assistants:** understand 'Call him' or 'Play that again'.
- **Sentiment and opinion mining:** attach opinions to the correct target.
- **Reading comprehension and commonsense benchmarks:** Winograd Schema Challenge, GAP.
- **Text simplification and accessibility:** replace ambiguous pronouns by explicit nouns.

Topic 2: Types of Anaphora

2.1 Definition

Types of anaphora are the categories of anaphoric relations distinguished by the *form of the anaphor* (pronoun, definite NP, ellipsis...), the *direction* (backward or forward), the *distance* (same or different sentence), and the *kind of antecedent* (individual, set, event, proposition). Each type needs its own resolution strategy.

2.2 Characteristics: main types

Type	Description	Example (anaphor in bold)
Pronominal anaphora	Personal, possessive pronouns point to an NP antecedent; the most studied type	'Anita lost her keys. **She** searched for them.'
Reflexive / reciprocal	Antecedent must be in the same clause (binding)	'Ravi cut **himself**.' 'They helped **each other**.'
Definite NP anaphora	Definite description repeats or re-describes the antecedent	'A girl found a wallet. **The girl** was honest.'
Demonstrative anaphora	this / that / these / those, as pronoun or determiner	'I bought a phone. **This phone** is fast.'
One-anaphora (nominal substitution)	'one(s)' replaces a noun and its modifiers	'I like the red scarf, not the blue **one**.'
Verbal / VP ellipsis anaphora	Auxiliary or 'do so' stands for an earlier verb phrase	'Meena lost her keys and Anita **did too**.'
Zero anaphora (pro-drop)	Subject or object omitted; recovered from agreement and context	Hindi/Telugu/Japanese/Spanish: 'Ravi came and [Ø] left.'
Temporal / locative anaphora	then, there, that day, that place	'We reached Goa. **There** it rained.'
Discourse deixis / abstract anaphora	this / that / it refer to a clause, event, fact or proposition	'The bank cut rates. **That** surprised everyone.'
Bridging (associative) anaphora	Inferable link, not identity	'a car ... **the engine**'
Cataphora	Anaphor precedes its antecedent	'When **she** arrived, **Anita** smiled.'
Bound-variable anaphora	Pronoun depends on a quantifier, no single referent	'**Every student** thinks **he** is clever.'

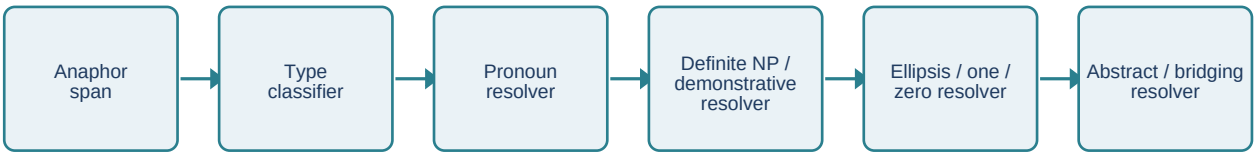
Classification along other dimensions

Dimension	Values	Example
Direction	Anaphora (backward) vs cataphora (forward)	He... Ravi (forward)
Distance	Intra-sentential vs inter-sentential	'Ravi said **he** was tired' vs 'Ravi came. **He** sat.'
Antecedent type	Individual NP; plural / split antecedent; event; proposition	'Ravi met Meena. **They** talked.' (split antecedent)
Relation	Direct (identity) vs indirect (bridging)	the girl vs the engine
Referentiality	Referential vs non-referential (pleonastic)	'**It** is raining.'

2.3 Architecture

A practical system first **classifies the anaphor type**, then routes it to a specialised resolver, because candidate sets and constraints differ by type.

Type-based routing architecture



Type	Candidate set	Key constraints / cues	Typical method
Pronoun	NPs in current + previous 2-3 sentences	Agreement, binding, salience	Hobbs, RAP, neural ranking
Reflexive	Subject / co-argument in same clause	Binding Principle A	Syntactic rule
Definite NP	All earlier mentions in document	Same head or hypernym, no contradicting modifiers	String match, WordNet, embeddings
One-anaphora	Earlier head nouns with modifiers	Compatible head noun, contrast on modifier	Rules + parse
VP ellipsis	Earlier verb phrases	Parallel structure, tense	Syntactic parallelism
Zero	Subjects / topics in discourse	Verb agreement (gender, number, person), topic continuity	Centering, neural
Abstract anaphor	Clauses and sentences	Verb selectional restriction (event / fact)	Classifier over clauses
Bridging	Entities in window	Part-of / member-of relations	Knowledge-based ranking

2.4 Example, step by step: classify every anaphor

```
[1] Anita bought a red scarf, and Meena bought a blue one. [2] They wore the scarves to
the party.
[3] Before she left, Anita smiled. [4] Meena lost her keys, and Anita did too. [5] That
upset them.
```

Anaphor	Type	Reasoning	Antecedent
one (S1)	One-anaphora	Replaces noun 'scarf', modifier 'blue' contrasts with 'red'	scarf (a blue scarf)
They (S2)	Pronominal, plural, split antecedent	Plural; no single plural NP; combine two subjects	Anita + Meena
the scarves (S2)	Definite NP (plural, set)	Sum of the two scarves just introduced	a red scarf + a blue one
she (S3)	Cataphora (intra-sentential)	Pronoun inside fronted subordinate clause precedes name	Anita (forward)
did too (S4)	VP ellipsis	Auxiliary without verb phrase	lost her keys
her (S4)	Possessive pronoun	Possessor of 'keys', agreement fem sg; subject	Meena
That (S5)	Discourse deixis (abstract)	Demonstrative pronoun with event referent; 'upset' needs an event	Both of them losing keys (S4)

Anaphor	Type	Reasoning	Antecedent
them (S5)	Pronominal, plural	Plural animate	Meena + Anita

2.5 Applications

- **Coreference systems:** routing by type improves accuracy (pronouns vs names vs definite NPs).
- **Machine translation:** pro-drop languages (Hindi, Japanese, Spanish) need zero-anaphora recovery when translating into English.
- **Dialogue and chatbots:** VP ellipsis ('so did I') and abstract anaphora ('that is why...').
- **Summarisation:** merging chains of different anaphor types for readability.
- **Linguistic analysis and cohesion measurement:** count anaphoric ties by type for text analysis and L2 writing assessment.
- **Information extraction from events:** event anaphora links reports of the same event.

Topic 3: Constraints in Anaphora Resolution

3.1 Definition

Constraints in anaphora resolution are conditions that a candidate must satisfy to be a *possible* antecedent; candidates violating them are **eliminated**. They are distinguished from **preferences** (also called factors or heuristics), which only *rank* the surviving candidates. Constraints reduce the search space; preferences choose among what is left.

3.2 Characteristics

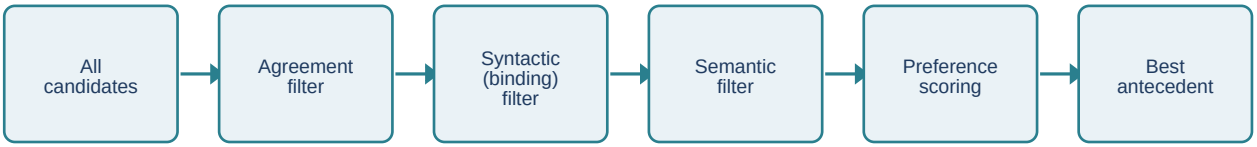
	Constraints	Preferences
Role	Eliminate impossible candidates (filter)	Rank possible candidates (score)
Strength	Strong, near-absolute (but with exceptions)	Soft, probabilistic, can be outvoted
Examples	Gender / number agreement, binding principles, selectional restrictions	Recency, subject preference, parallelism, repetition, centering, topic
Applied	First	After filtering
Cost	Cheap (lexical / structural checks)	May need learned models
Failure effect	Wrong filter removes the correct antecedent (cannot be recovered)	Wrong ranking picks wrong survivor

Constraint type	Information used	Example
Morphological / agreement	Number, gender, person, animacy	'she' cannot refer to 'Rahul'
Syntactic	Parse tree, binding theory, c-command, clause structure	'him' cannot be the subject of its own clause
Semantic	Selectional restrictions, sortal / type compatibility, plausibility	'drank it' needs a liquid
Discourse	Salience, focus, recency, centering	Prefer the entity in focus
Pragmatic / world knowledge	Commonsense, causality, scripts	'The trophy did not fit in the suitcase because it was too big' -> trophy

- **Constraints are language specific:** gender systems and binding domains differ (e.g., Hindi / Telugu verb agreement, Japanese zibun).
- **Exceptions exist:** collective nouns ('The committee ... they'), singular 'they', metonymy ('The ham sandwich wants his bill'), generic uses; so strict filters sometimes need soft handling.
- **Hard constraints for speed, soft scores for robustness:** modern neural systems often learn them as features instead of hard filters.

3.3 Architecture: constraints then preferences

Filter-then-rank design (after Mitkov)



Stage	Operation	Effect on candidate set
Candidate generation	All NPs in the search window	Large (e.g., 6-12 NPs)
Agreement filter	Match number, gender, person, animacy	Removes many (about half or more)
Syntactic filter	Apply binding principles, clause-mate and c-command conditions	Removes clause-internal impossible candidates
Semantic filter	Check selectional restrictions and type compatibility	Removes implausible referents
Preference scoring	Recency, role, parallelism, salience, centering, learned weights	Chooses one among the rest
Fallbacks	If all candidates are removed: relax constraints (e.g., unknown gender, collective nouns) or consider exophoric / cataphoric / bridging reading	Prevents empty candidate sets

3.4 Example, step by step: shrinking the candidate set

```
[1] Priya gave Rahul the laptop at the office. [2] She asked him to repair it.
```

Step	Anaphor	Candidates and filter	Remaining
0	all	Priya, Rahul, the laptop, the office	4 candidates
1. Agreement	She	fem, sg, animate: Rahul (masc) out; laptop, office (inanimate) out	Priya
2. Agreement	him	masc, sg, animate: Priya (fem) out; laptop, office out	Rahul
3. Syntax	him	Binding B: 'him' cannot corefer with its clause subject 'She' = Priya (already excluded)	Rahul confirmed
4. Agreement	it	neuter, sg, inanimate: Priya, Rahul out	laptop, office
5. Semantics	it	'repair it' requires a repairable artifact: laptop is a device; office (place) is a poor object of 'repair' in this frame	laptop
6. Preferences	it	Not needed (one survivor). Otherwise: laptop is direct object of S1 and nearer than 'office'	laptop

Result: She = Priya, him = Rahul, it = the laptop. The number of candidates for each anaphor fell from 4 to 1 using only constraints.

3.5 Applications

- **Efficient, explainable resolvers:** constraint filters shrink the search space for rule-based and hybrid systems.
- **Feature engineering for ML / neural models:** agreement and binding checks as features or masks.

- **Low-resource languages:** where little annotated data exists, constraint-based resolvers are the baseline.
- **Error analysis and debugging:** a violated constraint pinpoints why a link is wrong.
- **Machine translation:** constraint checking validates pronoun choices in generated target text.
- **Educational and writing tools:** flag ambiguous or ungrammatical pronoun use.

Topic 4: Agreement Constraints

4.1 Definition

Agreement constraints require that an anaphor and its antecedent be **compatible in morphosyntactic features**: **number** (singular / plural), **gender** (masculine / feminine / neuter), **person** (1st / 2nd / 3rd) and, in practice, **animacy** (human / animate / inanimate). A candidate whose features conflict with the anaphor is removed.

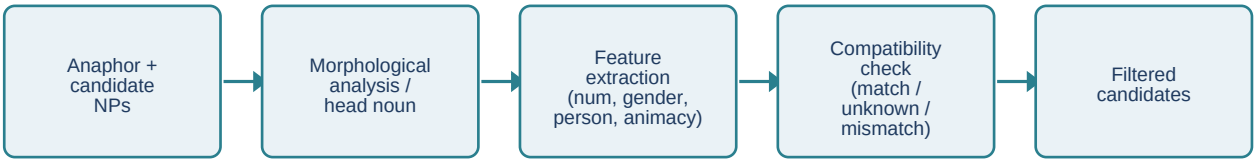
4.2 Characteristics

Feature	Rule	Example	Exceptions / difficulties
Number	Singular anaphor needs singular antecedent; plural needs plural	'The boys ... **they**' ok; 'The boy ... they' not	Collective nouns ('The team ... they'), singular 'they', 'everyone ... they', conjoined NPs ('Ravi and Meena' = plural), split antecedents
Gender	he / his = masculine, she / her = feminine, it = neuter / inanimate	'Priya ... **she**' ok; 'Priya ... he' not	Unknown or neutral names (Kiran, Alex), occupations ('the doctor'), animals ('it' vs 'he'), grammatical gender in French / German / Hindi / Telugu
Person	1st / 2nd / 3rd person pronouns refer to speaker / hearer / others	'I' = speaker, 'you' = addressee	Direct vs reported speech: 'Ravi said, "I am tired"' (I = Ravi); generic 'you'
Animacy	Human / animate pronouns need human / animate antecedent, 'it' needs inanimate (mostly)	'it' = laptop	'it' for babies, animals, robots; personification
Case / honorific	Mostly not required in English; relevant in Hindi / Telugu honorific forms	Respect forms 'aap', 'meeru' take plural verb agreement with singular referent	Language specific

- **Agreement helps zero anaphora:** in Hindi and Telugu verbs inflect for gender, number and person, so a dropped subject can be recovered. Example (Telugu, romanised): *vacchaadu* 'he came' vs *vacchindi* 'she came'; (Hindi) *gaya* (masc.) vs *gayi* (fem.) 'went'.
- **Compatibility has three outcomes:** match, mismatch, **unknown** (e.g., 'the teacher' has no gender). Unknown is treated as compatible.
- **Agreement is strong but not infallible:** use it as a hard filter in precise systems, or as a strong feature in learned systems.

4.3 Architecture of an agreement checker

Agreement constraint module



Stage	Description
Feature lexicons	Pronoun table: he (3, sg, masc, animate), she (3, sg, fem, animate), it (3, sg, neuter, inanimate), they (3, pl, any), I / you etc.
Noun features	Number from POS tag (NN vs NNS) or morphology; gender from name lists, titles (Mr, Mrs, Dr), WordNet, or web-derived gender statistics (e.g., Bergsma and Lin); animacy from NER type (PERSON vs ORG vs OBJECT)
Complex NPs	Conjoined NPs ('Anil and the nurses') become plural; collective nouns get number 'either'; split antecedents formed from two singular NPs
Compatibility function	For each feature: match (keep), unknown (keep, maybe penalise), mismatch (remove). Candidate survives only if no mismatch
Neural variants	Gender / number as embeddings or attention features; pretrained LMs implicitly encode agreement but can be biased by stereotypes

4.4 Example, step by step

```
Dr. Meera met Anil and the nurses. She told them about the schedule.
```

Step	Item	Features
1. Anaphor 1	She	3rd, singular, feminine, animate
2. Candidates	Dr. Meera; Anil; the nurses; (Anil and the nurses)	Meera: 3rd, sg, fem (name list + 'Dr. Meera' is a person), Anil: 3rd, sg, masc, nurses: 3rd, pl, unspecified gender, conjoined NP: pl
3. Compatibility for 'She'	Check each	Meera: match. Anil: gender mismatch -> removed. nurses: number mismatch -> removed. conjoined NP: number mismatch -> removed
4. Resolve 'She'		She = Dr. Meera
5. Anaphor 2	them	3rd, plural, animate (any gender)
6. Compatibility for 'them'	Check each	Meera: sg mismatch (also subject of same sentence, binding). Anil: sg mismatch. nurses: match. 'Anil and the nurses': plural match
7. Preference (agreement leaves 2)	Choose	Prefer full coordinated NP (it is the complete object of 'met', parallel role) or nurses alone: flag as ambiguous; default = Anil and the nurses

Lesson: agreement eliminates most wrong candidates but cannot always decide: plural 'them' still leaves two options. Other constraints and preferences are needed.

4.5 Applications

- **Machine translation:** choosing he / she / it, and gendered pronouns and verb forms in Hindi, Telugu, French, German.

- **Pro-drop language processing:** recovering dropped subjects from verb agreement.
- **Grammatical error correction:** detecting pronoun-antecedent disagreement in writing ('Each student must bring their...' vs 'his').
- **Bias and fairness evaluation:** testing whether models link occupations to stereotyped genders (WinoBias, WinoGender).
- **Coreference resolution preprocessing:** pruning candidate pairs to speed up training and inference.
- **Chatbots and generation:** producing consistent pronouns for users and entities.

Topic 5: Syntactic and Semantic Constraints

5.1 Definition

Syntactic constraints restrict possible antecedents using the *structure of the sentence* (parse tree, clause boundaries, c-command, binding theory).

Semantic constraints restrict possible antecedents using *meaning*: the semantic type of the candidate must fit what the anaphor's context requires (selectional restrictions, sortal compatibility, plausibility, world knowledge).

5.2 Characteristics: syntactic constraints

Binding Theory (Chomsky, 1981). C-command: node A c-commands node B if the first branching node above A also dominates B (and neither dominates the other). The **governing category** is roughly the smallest clause (or NP) containing the anaphor and a subject.

Principle	Statement	Example	Effect
A	A reflexive / reciprocal must be bound (coindexed with a c-commanding antecedent) in its governing category	'Ravi shaved **himself**.'	himself = Ravi (same clause subject)
B	A personal pronoun must be free (not bound) in its governing category	'Ravi shaved **him**.'	him != Ravi; refers to someone else
C	A name or definite description must be free everywhere	'**He** said that **Ravi** was late.'	He != Ravi

- **Cataphora restrictions:** allowed when the pronoun is in a subordinate clause that precedes the main clause ('When **he** arrived, **Ravi** smiled'), blocked when the pronoun c-commands the name ('**He** said Ravi won').
- **Clause-mate condition:** reflexives take clause-mate antecedents; pronouns usually do not.
- **Pleonastic-it patterns:** weather verbs, extraposition ('It is clear that...'), cleft sentences are non-referential.
- **Syntactic preferences (soft):** subject preference, parallelism of grammatical role, and recency are not constraints but guide ranking. The Hobbs algorithm uses a systematic tree search that builds in several binding effects.

5.3 Characteristics: semantic constraints

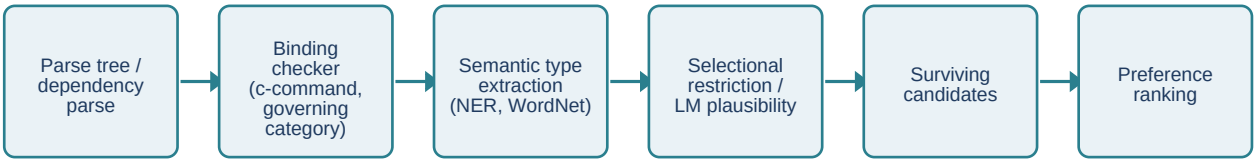
Constraint	Description	Example
Selectional restrictions	The predicate demands a semantic type for its argument	'ate it' -> food; 'drank it' -> liquid; 'read it' -> readable object
Sortal / type compatibility	Animate vs inanimate, person vs place, event vs object	'**That** surprised analysts' -> event or fact, not an object
Hypernymy / semantic relatedness	Definite NP re-describes antecedent by synonym or hypernym	'a Dell laptop ... the device'
Thematic / role fit	Agent must be animate, instrument may be object	'The key opened the door' key = instrument

Constraint	Description	Example
Plausibility / world knowledge	Choose the reading that makes the best sense	'The trophy did not fit in the suitcase because **it** was too big' -> trophy; 'too small' -> suitcase
Implicit causality	Verbs bias reference toward subject or object	'Ravi blamed Bill because **he** was careless' -> Bill

Limits: semantic constraints need large knowledge resources (WordNet, VerbNet, FrameNet, ConceptNet, language models), and metaphor and metonymy break them ('The car drank gas', 'The White House said...'). They are therefore often soft scores rather than hard filters.

5.4 Architecture

Syntactic and semantic constraint modules



Module	How it works
Parser	Constituency or dependency parse supplies clause structure, subjects, objects, governing category
Binding checker	For each (anaphor, candidate) pair compute c-command and whether they are in the same governing category; apply Principles A, B, C
Semantic type extraction	Assign semantic classes to candidates (person, organisation, artifact, liquid, event) using NER, WordNet hypernyms, embeddings
Selectional restriction check	Identify the predicate for which the anaphor is an argument and fetch its argument type from VerbNet / FrameNet / corpus statistics; keep compatible candidates
Plausibility scoring	Substitute each candidate for the anaphor and score the sentence with a language model (used in Winograd solvers); or use a neural cross-encoder over (context, pronoun, candidate)
Combination	Hard syntactic filters first, semantic constraints as hard filters for clear type clashes or as soft scores, then preference ranking

5.5 Example, step by step

Part A - Syntactic (binding).

```
[1] Ravi thinks that Mohan admires him. [2] Mohan admires himself.
```

Step	Anaphor	Analysis	Result
1	him (S1)	Governing category = embedded clause 'Mohan admires him' (subject Mohan). Candidates: Ravi, Mohan	-
2	him	Principle B: pronoun must be free in its governing category -> cannot be Mohan (clause subject, c-commands him)	Mohan removed
3	him	Ravi is in the main clause, outside the governing category -> allowed	him = Ravi
4	himself (S2)	Principle A: reflexive must be bound in its clause by a c-commanding subject	himself = Mohan
5	Check name rule	In 'He thinks that Ravi is clever', Principle C blocks He = Ravi	-

Part B - Semantic (selectional restriction and plausibility).

```
Anita poured the juice into the glass and drank it.
```

Step	Action	Result
1	Candidates for 'it'	Anita, the juice, the glass
2	Agreement / syntax	Anita removed (animate, subject of 'drank'; Principle B also blocks it); juice, glass remain

Step	Action	Result
3	Predicate and argument type	'drank' -> object type = liquid / beverage
4	Semantic types	juice = liquid (WordNet), glass = container / artifact
5	Selectional check	juice compatible; glass incompatible (unless metonymy for its contents)
6	Result	it = the juice

Part C - Commonsense (Winograd). 'The trophy does not fit in the suitcase because it is too big.' Candidates: trophy, suitcase. Agreement and syntax cannot decide (both singular, inanimate). Semantic reasoning: if something is *too big* to fit, it is the object being put in, so it = trophy. For 'too small', the container is the problem, so it = suitcase. LM plausibility scoring (illustrative): P('trophy is too big') = 0.62 vs P('suitcase is too big') = 0.11, so choose trophy.

5.6 Applications

- **Rule-based and hybrid pronoun resolvers:** binding filters (Hobbs, RAP) give precise baselines.
- **Commonsense reasoning evaluation:** Winograd Schema Challenge and WinoGrande test semantic constraints in language models.
- **Question answering and reading comprehension:** resolving pronouns in context requires semantic plausibility.
- **Machine translation:** selecting the right pronoun and agreement in the target language when the source is ambiguous.
- **Dialogue understanding:** keep references consistent with what is possible in the situation.
- **Grammar checking and teaching:** explain why a reflexive or pronoun use is ungrammatical.
- **Information extraction:** avoid attaching facts to semantically impossible entities.

Topic 6: Summary and Exam Questions

6.1 Comparison of the five topics

	Definition	Types	Constraints (overview)	Agreement	Syntactic / semantic
Core idea	Anaphor depends on antecedent	Categories by form, direction, antecedent	Filter candidates, then rank	Match features	Structure and meaning fit
Key notions	Anaphor, antecedent, chain	Pronominal, definite NP, one, ellipsis, zero, abstract, bridging, cataphora	Constraint vs preference	Number, gender, person, animacy	Binding A / B / C, c-command, selectional restrictions
Main difficulty	Ambiguity, pleonastic it	Different resolvers per type	Exceptions and over-filtering	Unknown gender, collective nouns	Needs knowledge and commonsense
Typical method	Pipeline: candidates, filter, rank	Type classifier + resolvers	Mitkov, Hobbs, RAP	Feature checker	Binding checker, LM plausibility

6.2 Likely exam questions

1. Define anaphora and anaphora resolution. Distinguish anaphora, cataphora, coreference and bridging with examples.
2. Describe the architecture of an anaphora resolution system with a neat diagram.
3. Explain the different types of anaphora with examples.
4. What are constraints and preferences in anaphora resolution? Explain with an example how the candidate set shrinks.
5. Explain agreement constraints (number, gender, person, animacy) and their exceptions. How do agreement features help in Hindi or Telugu?
6. Explain syntactic constraints in anaphora resolution: binding principles A, B, C and c-command, with examples.
7. Explain semantic constraints: selectional restrictions, plausibility and the Winograd example.
8. List applications of anaphora resolution in NLP.

Note on figures: the language-model probabilities, candidate counts and similar numbers in the examples are illustrative values chosen to explain each step, not measured results. Theory details (binding principles, constraint-preference distinction, algorithm names) follow the standard literature; check exact formulations against your textbook before quoting them in an exam.