On HDTP And Human-Style Rationality

Tarek R. Besold, Martin Schmidt, Helmar Gust, Ulf Krumnack
Ahmed Abdel-Fattah and Kai-Uwe Kühnberger

University of Osnabrück, Institute of Cognitive Science

Outline

- Rationality puzzles and Analogy
- Heuristic Driven Theory Projection
  - Restricted Higher-order Anti-Unification
  - Heuristics
- Rationality Through Analogy
  - Frameworks
  - expanding HDTP
- Conclusion

Rationality Puzzles

Wason Selection Task

- Wason, P. C., & Shapiro, D. (1971)
  Natural and Contrived Experience in a Reasoning Problem.

Every card which has a D on one side has a 3 on the other side (and knowledge that each card has a letter on one side and a number on the other side), together with four cards showing respectively D, K, 3, 7, ... This problem is called "selection task" and the conditional sentence is called "the rule".

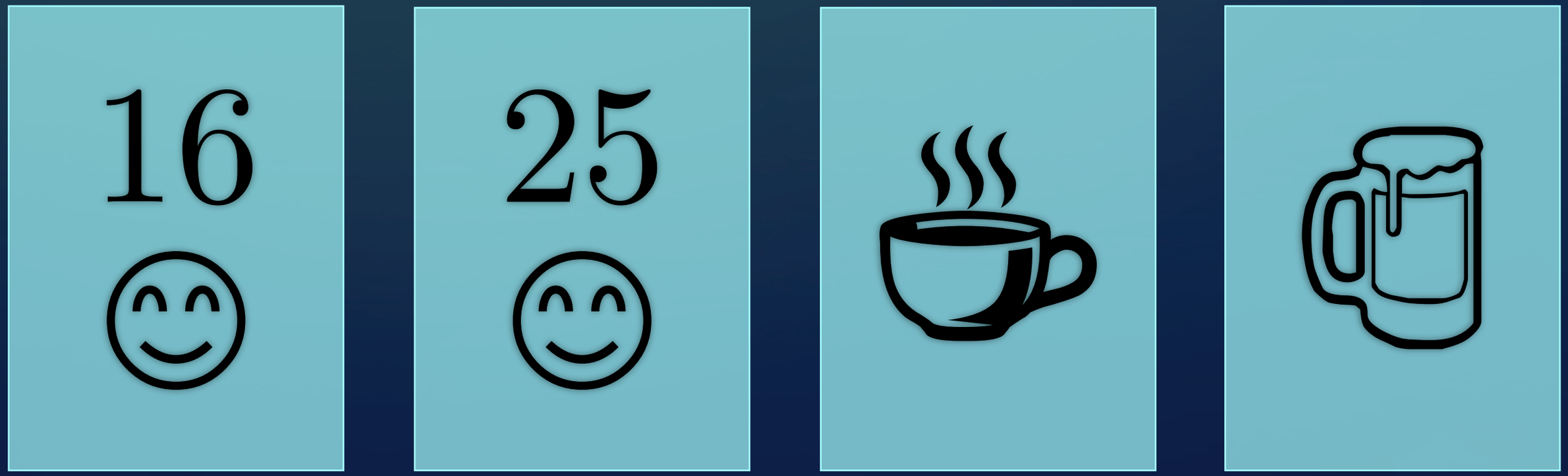
The Task

- Every card which has a D on one side has a 3 on the other side

D	K	3	7

Different Context

- Cosmides, L., & Tooby, J. (1993).
  Cognitive adaptations for social exchange
- "If you are drinking alcohol then you must be over 18"



Result

- Classic logic: turn D & 7
- Turn 7 by modus tollens:
  $((P \rightarrow Q) \wedge \neg Q) \rightarrow \neg P$
- Only 10% of subjects give the "correct" answer

Selection Task and Analogy

- Content-change makes the task easier to access for subjects
- Subjects performance is tightly connected to establishing appropriate analogies
- They fail in original task to establish a fitting analogy with an already known situation

Modeling Rationality Puzzles

- Besold, T. R., Gust, H., Krumnack, U., Abdel-Fattah, A., Schmidt, M., & Kühnberger, K. (2011, July). An Argument for an Analogical Perspective on Rationality & Decision-Making.
  - Wason, P. C., & Shapiro, D. (1971) Natural and Contrived Experience in a Reasoning Problem.
    - Wason selection task: logic
  - Byrne, R (1989): Suppressing valid inferences with conditionals.
    - reasoning with conditionals
  - Tversky, A., Kahneman, D. (1983): Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement.
    - Linda Problem: violation of the rules of probability theory

Analogy

Analogy

- Core of cognition
- Application in many domains
- No logical theory for analogical reasoning

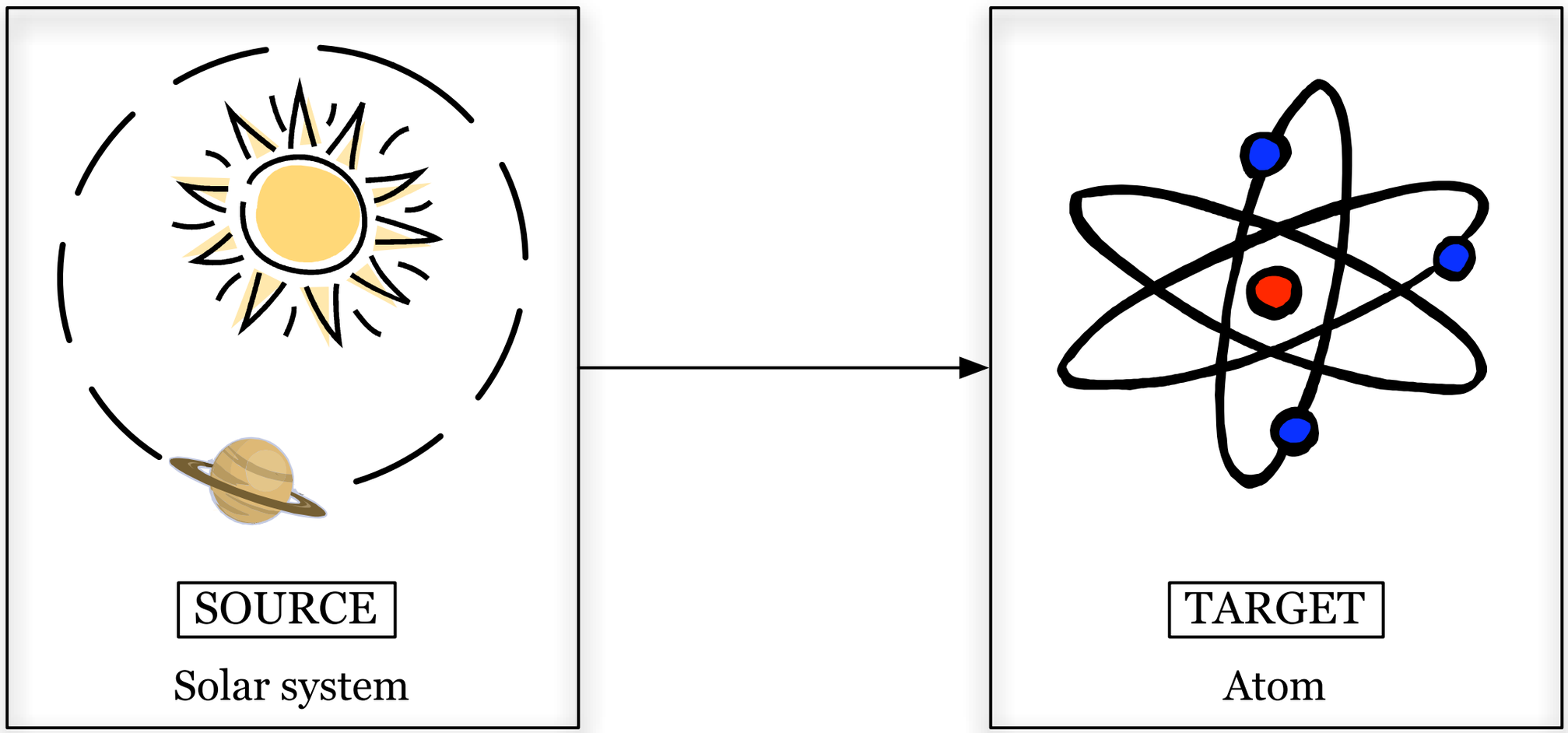
- Build a computational model

Heuristic-Driven Theory Projection (HDTP)

Properties

- Symbolic
  (First oder logic formalization of domains axioms)
- Cognitively inspired, not a cognitive architecture
  - Detects deep structural commonalities
    - Higher-order anti-unification
  - Produces explicit generalization
  - Phases: Retrieval, Mapping, Transfer

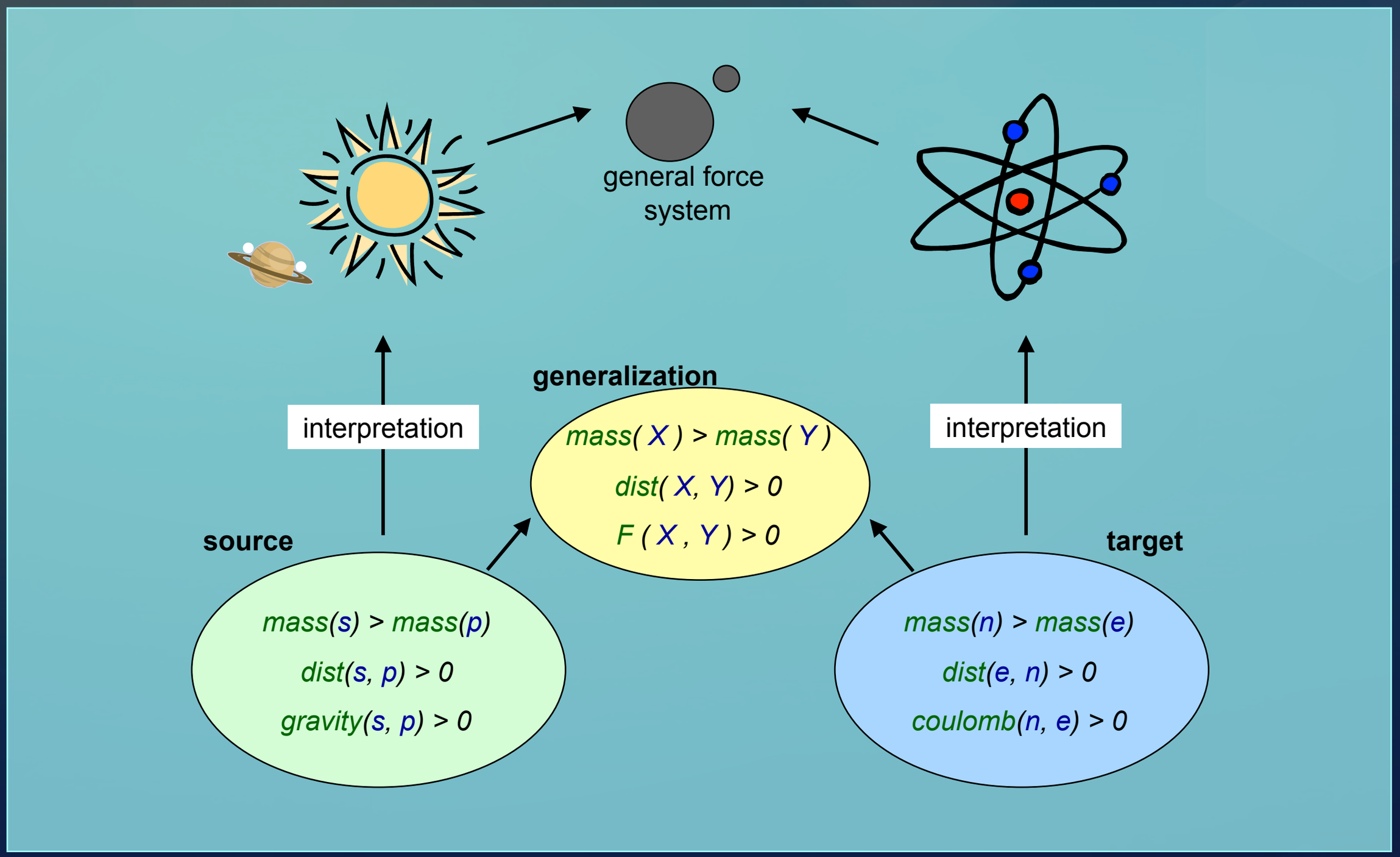
Example



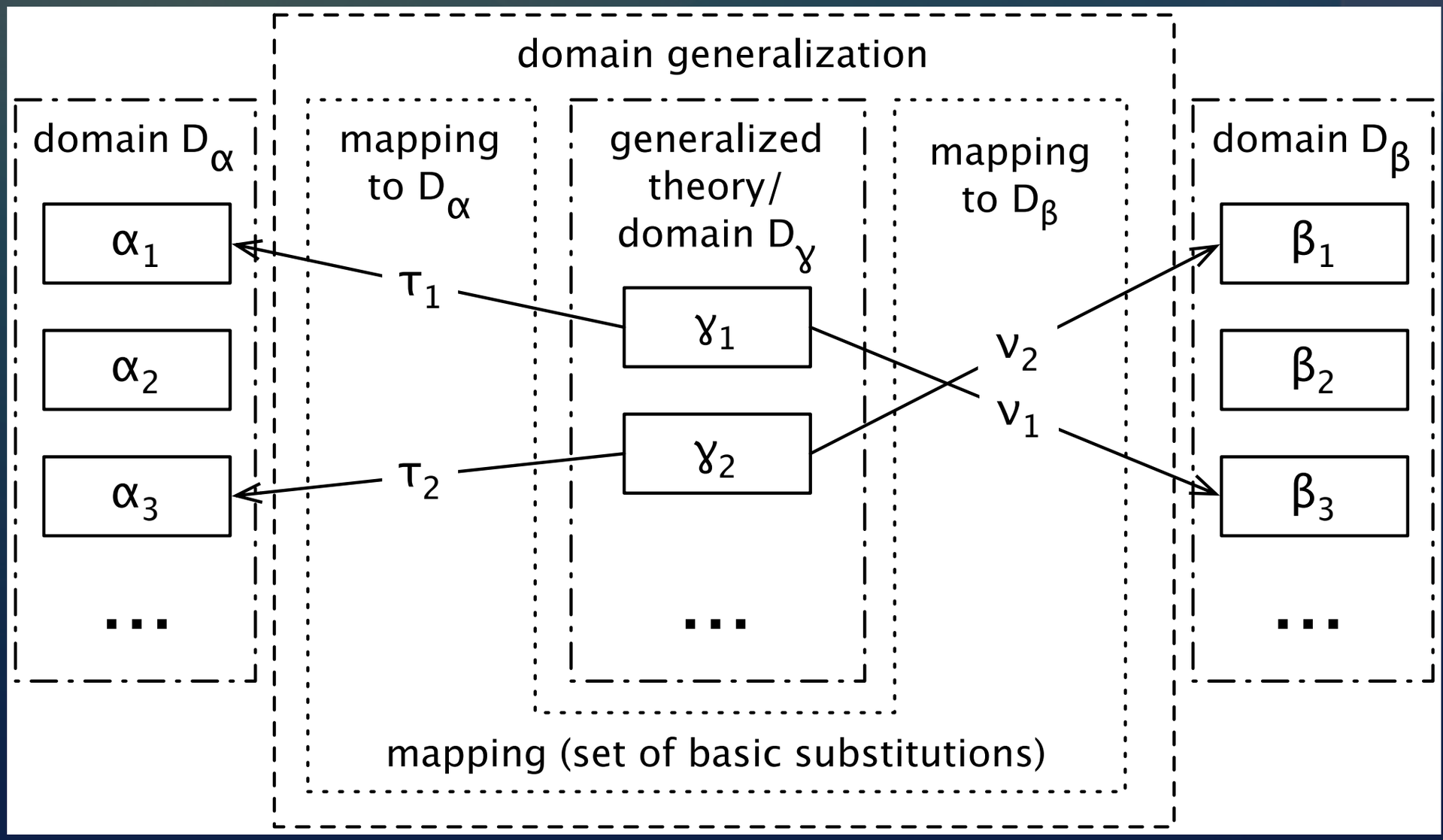
Solar System	Rutherford Atom
sorts *real*, *object*, *time*	**sorts** *real*, *object*, *time*
constants *sun* : *object*, *planet* : *object*	**constants** *nucleus* : *object*, *electron* : *object*
functions *mass* : *object* × *time* → *real* *dist* : *object* × *object* × *time* → *real* *force* : *object* × *object* × *time* → *real* *gravity* : *object* × *object* × *time* → *real* *centrifugal* : *object* × *object* × *time* → *real*	**functions** *mass* : *object* × *time* → *real* *dist* : *object* × *object* × *time* → *real* *coulomb* : *object* × *object* × *time* → *real*
predicates *revolves_around* : *object* × *object*	**facts** β_1 : $mass(nucleus) > mass(electron)$ β_2 : $mass(electron) > 0$ β_3 : $\forall(T : time) : coulomb(electron, nucleus, T) > 0$ β_4 : $\forall(T : time) : dist(electron, nucleus, T) > 0$
facts α_1 : $mass(sun) > mass(planet)$ α_2 : $mass(planet) > 0$ α_3 : $\forall(T : time) : gravity(planet, sun, T) > 0$ α_4 : $\forall(T : time) : dist(planet, sun, T) > 0$	
laws α_5 : $\forall(T : time, O_1 : object, O_2 : object)$: $dist(O_1, O_2, T) > 0 \wedge gravity(O_1, O_2, T) > 0$ $\rightarrow centrifugal(O_1, O_2, T) = -gravity(O_1, O_2, T)$ α_6 : $\forall(T : time, O_1 : object, O_2 : object)$: $0 < mass(O_1) \wedge 0 < mass(O_2) \wedge$ $dist(O_1, O_2, T) > 0 \wedge centrifugal(O_1, O_2, T) < 0$ $\rightarrow revolves\_around(O_1, O_2)$	

Prolog Input

```
ANALOGY('RUTHERFORD',
      DOMAIN('PHYSICS',[S:MASS(O:SUN) > S:MASS(O:PLANET),
                  S:DISTANCE(O:SUN,O:PLANET) > S:0,
                  YELLOW(O:SUN),
                  S:FORCE(F:GRAVITATION,O:PLANET,O:SUN) > S:0,
                  F:F_ANTI(F:GRAVITATION) = F:CENTRIFUGAL,
                  (S:DISTANCE(O:_X,O:_Y) > S:0,
                  S:FORCE(F:_F,O:_Y,O:_X) > S:0 -> S:FORCE(F:F_ANTI(F:_F),O:_X,O:_Y) < S:0 ),
                  (S:FORCE(F:CENTRIFUGAL,O:_X1,O:_Y1) < S:0, S:MASS(O:_Y1) > S:MASS(O:_X1) ->
                        REVOLVE(O:_X1,O:_Y1))
                  ]),
DOMAIN('PHYSICS',[S:FORCE(F:COULOMB,O:ELECTRON,O:NUCLEUS) > S:0,
                  S:DISTANCE(O:ELECTRON,O:NUCLEUS) > S:0,
                  S:MASS(O:NUCLEUS) > S:MASS(O:ELECTRON)
                  ])).
```


general force system
generalization
interpretation
interpretation
mass(X) > mass(Y)
dist(X, Y) > 0
F (X , Y) > 0
source
target
mass(s) > mass(p)
dist(s, p) > 0
gravity(s, p) > 0
mass(n) > mass(e)
dist(e, n) > 0
coulomb(n, e) > 0

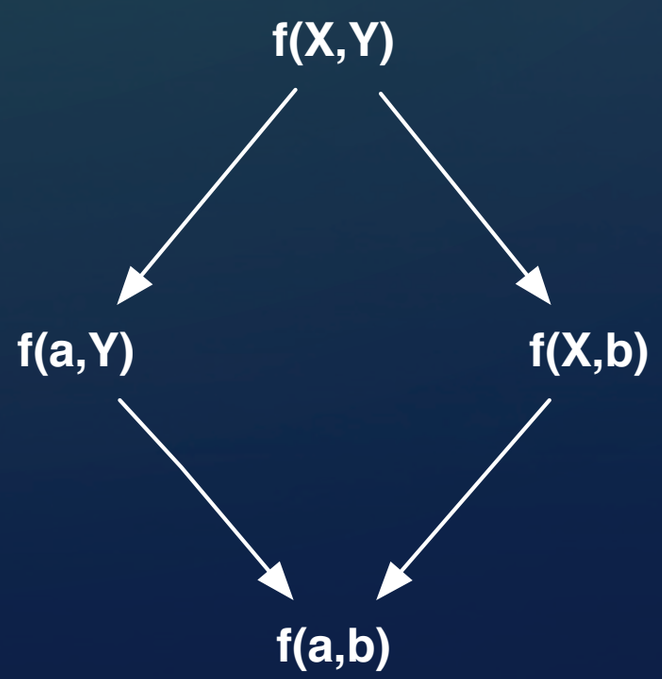
Mapping


domain generalization
domain D_α
mapping to D_α
generalized theory/ domain D_γ
mapping to D_β
domain D_β
α_1
α_2
α_3
...
τ_1
τ_2
γ_1
γ_2
...
ν_2
ν_1
β_1
β_2
β_3
...
mapping (set of basic substitutions)

Restricted Higher-Order Anti-Unification

ANTI-UNIFICATION

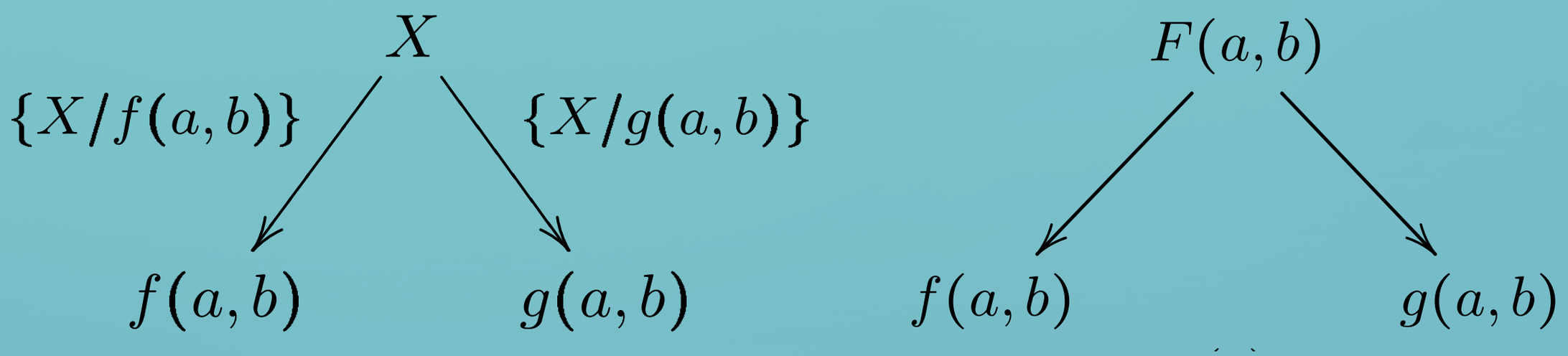
- DUAL OPERATION OF MORE WELL KNOWN UNIFICATION
- TYPE OF SUBSTITUTIONS CAN VARY
- FIND LEAST GENERAL / MOST SPECIFIC ANTI-UNIFIER



Anti-Unification Variants

- First-Order (Plotkin, 1970)
  - substitutions: replace a variable by any term
  - an anti-unifier always exists
  - finitely many most specific anti-unifier
    - unique most specific anti-unifier exists
- Second-Order
  - an anti-unifier always exists
  - captures deep structural commonalities

STRUCTURAL COMMONALITIES

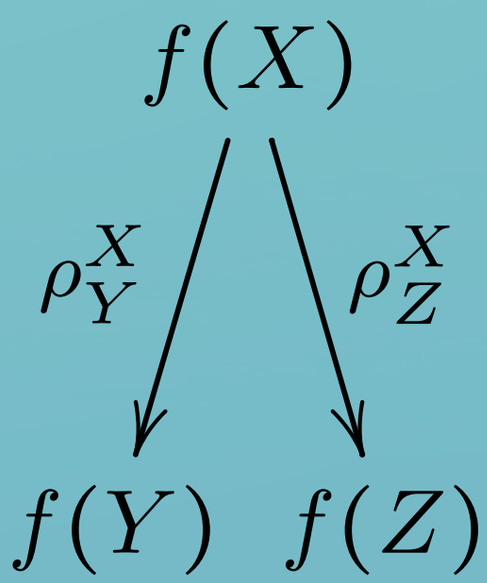


RESTRICTED HIGHER-ORDER

- GENERALIZED TERM STRUCTURALLY AS SIMPLE OR SIMPLER THAN INSTANCES
- HIGHER ORDER TERMS: VARIABLES WITH ARGUMENTS
- SUBSTITUTIONS ARE CHAINS OF BASIC SUBSTITUTIONS
  - RENAMING
  - FIXATION
  - ARGUMENT INSERTION
  - PERMUTATION

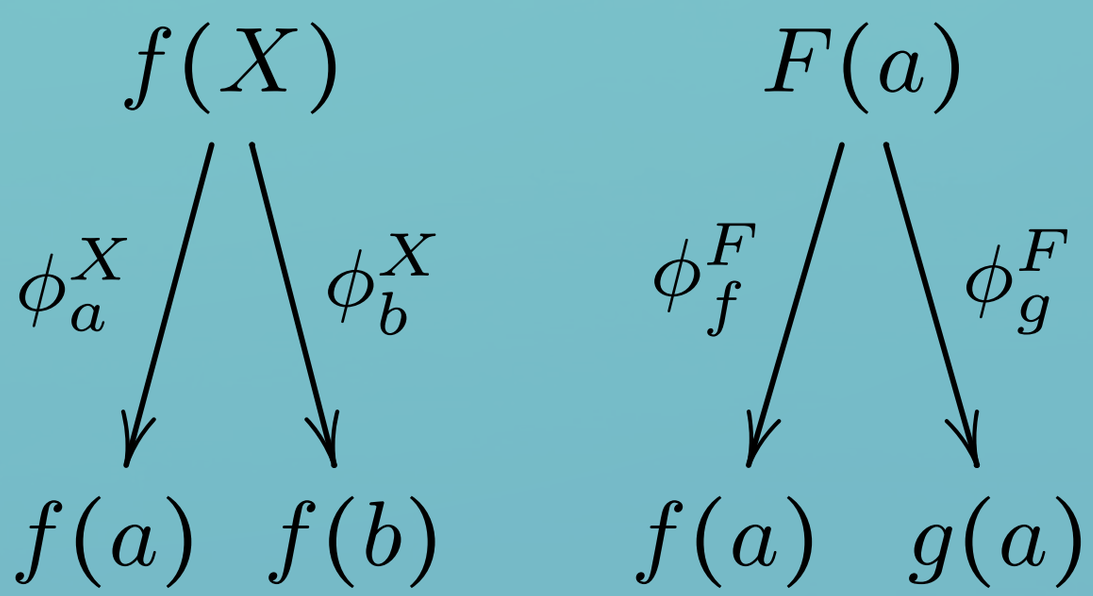
Renaming

- replace a variable by another variable



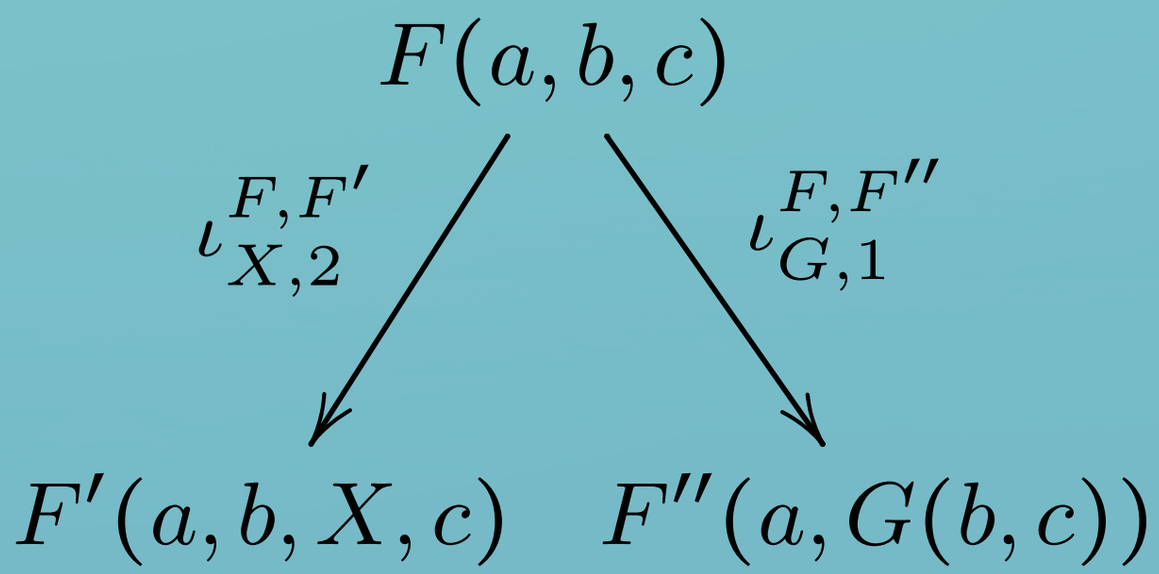
Fixation

- replace a variable by a function or predicate symbol



Argument Insertion

- simple: insert variable of arity 0
- complex: insert variable of arity > 0
  variable absorbs arguments



Permutation

- rearrange arguments (identity excluded)

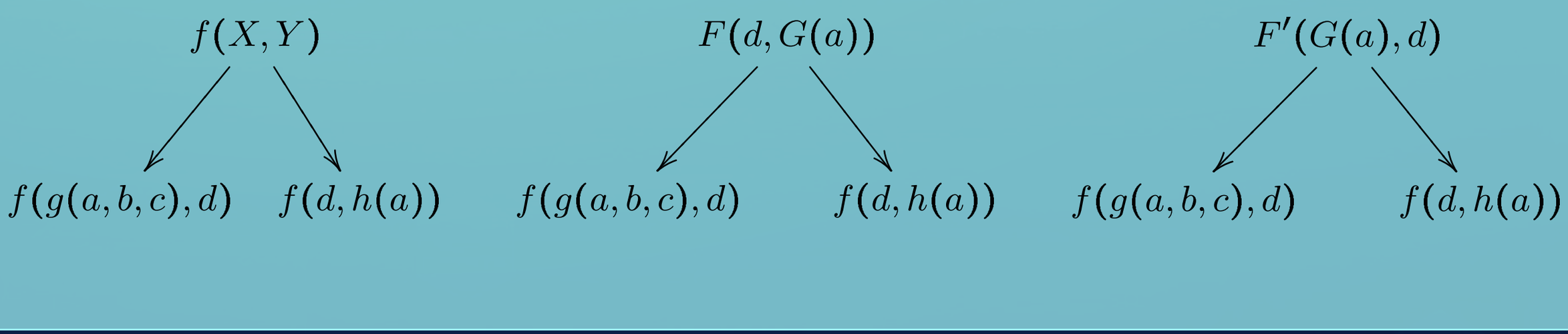
$$F(a,b,c)$$

$$\pi_{\alpha}^{F,F'} \qquad \pi_{\beta}^{F,F''}$$

$$F'(a,c,b) \qquad F''(c,b,a)$$

RESTRICTED HIGHER-ORDER ANTI-UNIFICATION

- ANTI-UNIFIER ALWAYS EXISTS
- FINITELY MANY MOST SPECIFIC ANTI-UNIFIER
- NOT ALWAYS A UNIQUE MOST SPECIFIC ANTI-UNIFIER EXISTS



COMPLEXITY OF BASIC SUBSTITUTIONS

$$\mathcal{C}(\tau) = \begin{cases} 0 & \textit{if } \tau = \rho & \textit{(renaming)} \\ 1 & \textit{if } \tau = \phi_f & \textit{(fixation)} \\ k+1 & \textit{if } \tau = \iota_{V,i} \textit{ and } V \in \mathcal{V}_k & \textit{(argument insertion)} \\ 1 & \textit{if } \tau = \pi_\alpha & \textit{(permutation)} \end{cases}$$

PREFERRED GENERALIZATION

- GENERALIZATION
  - GENERALIZED TERM TOGETHER WITH SUBSTITUTIONS
- COMPLEXITY OF GENERALIZATION
  - SUM OF BASIC SUBSTITUTION COMPLEXITIES
- PREFERRED GENERALIZATION
  - GENERALIZATIONS WITH LEAST COMPLEXITY
    - THERE EXIST ONLY FINITELY MANY

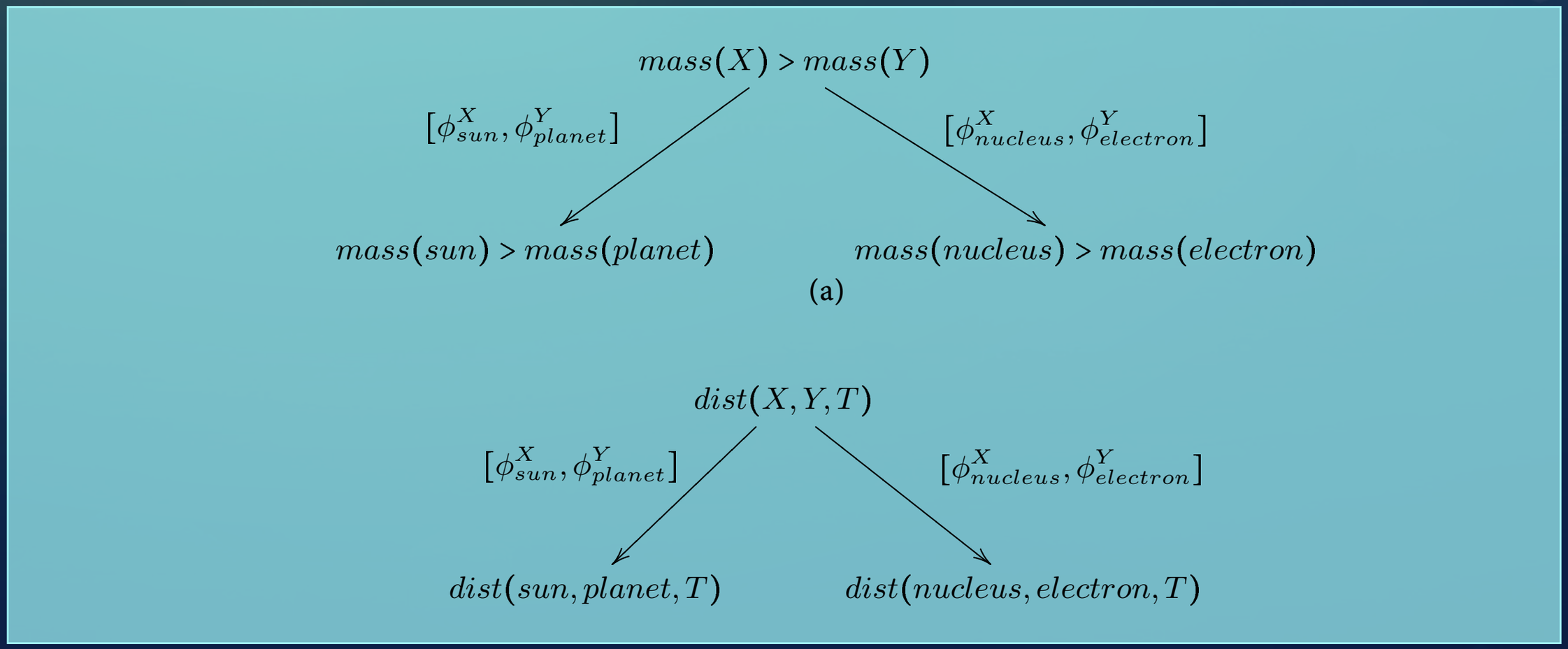
MORE CONSTRAINTS

- MANY-SORTED LOGIC WITH ONTOLOGY OF SORTS
  - SORT CONSTRICTION SUBSTITUTION
  - TIME, MASS, VOLUME, PHYSICAL OBJECT
- SYMBOL MAPPING CONSTRAINTS
  - ONE-TO-ONE, MANY-TO-ONE, MANY-TO-MANY

Mapping and Heuristics

Systematic Use Of Substitutions

- REUSE OF A BASIC SUBSTITUTION IS COMPLEXITY FREE



Mapping

- Computes an alignment between source and target domain with minimal complexity
- Local greedy search strategy
  - Repeat until all source (target) formula considered
    - Select a source formula
    - Compute complexity of anti-unification with target formulas and penalty for not matching at all
    - Select generalization with least complexity and mark involved formulas as mapped

Two Views on Heuristics

- Incorporating additional knowledge
  - Memory based reuse of generalizations from previous analogies
- Pruning the search tree
  - Intelligence in form of heuristics here is then to advert the threat of exponential search explosion

HEURISTICS

- AN ALIGNMENT WITH MINIMAL COMPLEXITY BUT MAXIMAL COVERAGE PRODUCES A GOOD ANALOGY
- SELECTING SOURCE FORMULAE
  - LEAST AMOUNT OF UNMAPPED SYMBOLS
  - MATCHING PREDICATES ON THE TOP-LEVEL
- HEURISTIC THAT APPROXIMATES ANTI-UNIFICATION COMPLEXITY
  - USE COMPLEXITY OF PAIR WITH HEURISTICALLY LEAST COMPLEXITY AS CUTOFF VALUE FOR SEARCH
- NON MATCH COMPLEXITY = (NUMBER OF SYMBOLS * 2) + 1
  - NOT ALIGN A FORMULA > FORMULA EQUAL IN ARGUMENT STRUCTURE

Rationality

FRAMEWORKS

- LOGIC BASED
  - WASON SELECTION TASK
- PROBABILITY BASED
  - LINDA PROBLEM
- GAME THEORY BASED
  - NO SINGLE OPTIMALITY CRITERION
- HEURISTIC BASED
  - OFTEN LACK FORMAL TRANSPARENCY AND EXPLANATIVE POWER
- USUALLY NORMATIVE BUT NOT PREDICTIVE FOR HUMAN BEHAVIOR
  - HUMAN RATIONALITY MODEL NEEDS TO ACCOUNT FOR "IRRATIONALITY"

Cognitive View Of Rationality

- Bounded rationality
  - H. Simon, (1959)
    Theories of decision making in economics and behavioural science
- Rationality as multilevel theory based on cognitive processes
  - B. Kokinov (2003)
    Analogy in decision-making, social interaction, and emergent rationality, Behavioral and Brain Sciences
  - Utility making is an emergent property in most cases
  - Rational rules as approximations of human behavior

Cognitive View Of Rationality

- B. Indurkhya (2007)
  Rationality and reasoning with metaphors
  - examined rationality in an interaction-based view of cognition
  - metaphors can radically alter the ontology of the target, thereby creating a new perspective
  - this role of metaphors should not be excluded from an account of rationality

- two key roles for analogy in rationality frameworks
  - restructure the problem domain (changing the view)
  - generation of new knowledge (transfer)

Existing Cognitive Frameworks

- G. Petkov and B. Kokinov (2006)
  JUDGEMAP - integration of analogy- making, judgement, and choice'
  - DUAL (cognitive architecture) + AMBR (analogy framework)
    - replicates range, frequency and sequential assimilation effects
    - reproduces contextual effects in tasks which at first sight don't seem to be related to analogy-making
- more projects that apply analogy-making to problem solving
  - M. Klenk and K. Forbus (2007)
    Cognitive modeling of analogy events in physics problem solving from examples'

HDTP as Part of An Architecture For Rationality

- Retrieval
  - compile and track limited library of previous situations
    - corresponds to human episodic memory
    - outcome does not have to be unique
- Mapping
  - multiple mappings possible
  - can be goal driven

HDTP as Part of An Architecture For Rationality

- TRANSFER
  - INTEGRATION AND CONSOLIDATION FROM MULTIPLE DOMAINS
- NEW APPLICATION PHASE
  - DOES NOT NEED TO BE A PERFECT DEDUCTIVE REASONER
  - FORMALIZATION EASY TO INTEGRATE INTO EXISTING PLANERS/REASONERS
  - PROBLEM SOLVING CAN STILL FAIL
- INTEGRATION OF UNCERTAINTY
  - OBTAINING (COMPLETE) DOMAIN THEORIES FROM (MOST LIKELY ONLY PARTIAL) OBSERVATIONS

Conclusion

- Theoretical and practical results are promising enough to justify serious and dedicated research efforts
  - Promote general application of analogy engines to rationality tasks
- Cognitive and analogy computation movement and rationality schools can profit from further intensifying their interaction and cooperation
  - Analogy-making could be expanded and enriched by some of the alternate techniques applied when modeling rationality
  - Frameworks for rationality could significantly profit from including more cognitive aspects and properties of humans into their models and theories