

**Kompozycyjny opis semantyki
polskich fraz nominalnych w oparciu
o metody semantyki dynamicznej**

Maciej Piasecki

Wydziałowy Zakład Informatyki
Wydział Informatyki i Zarządzania
Politechnika Wrocławska

Plan

- Cel pracy
- Realizacja celu
- *Samorganizująca się Logika Struktur (SLS)* – język reprezentacji znaczenia
- Notacja graficzna *SLS*
- Model znaczenia polskiej frazy nominalnej
- Koncepcja wykorzystania *SLS* w implementacji
- Ograniczenia i dalsze prace

Cel pracy

Celem pracy było skonstruowanie formalnego języka reprezentacji znaczenie polskiej frazy nominalnej (FN) a następnie jego wykorzystanie do opracowania formalnego opisu znaczenia szerokiego podzbioru polskich FN.

Fraza nominalna — definicja

(Saloni i Świdziński, *Składnia współczesnego języka polskiego*, PWN, 1998)

- fraza wymagana frazy finitywnej (*czasownikowej*),
- realizowana jako grupa: rzeczownikowa, przymiotnikowa (*eliptyczne imięsłowy*), liczebnikowa (*liczebnik+rzeczownik*), zdaniowa (*zdania względne*).

Realizacja celu

- Etap 1
 - konstrukcja *języka reprezentacji znaczenia* (JRZ) - języka logicznego wyższego rzędu z typami - nazwanego *Samoorganizująca się Logika Struktur* (*SLS*)
- Etap 2
 - opracowanie formalnego opisu znaczenia szerokiego podzbioru *polskich FN*, opartego na *niedospecyfikowanych wyrażeniach SLS*, jako reprezentacji znaczenia poszczególnych jednostek leksykalnych

SLS – zakres opisu

- *Samorganizująca się Logika Struktur*
- Reprezentacja znaczenia **frazy nominalnej** w **dyskursie** (wielozdaniowej wypowiedzi)
- Kompozycyjny i łączny opis wielu aspektów znaczenia FN:
 - **deskryptywnego** – informacja opisowa
 - **kwantyfikacji** – **kwantyfikator uogólniony** Mostowskiego
 - **referencji** – odniesienia wyrażenia językowego do bytów pozajęzykowych
 - **presupozycji** – warunków sensowności użycia
 - **anafory** - relacja semantyczno-składniowa pomiędzy dwoma wyrażeniami językowymi w dyskursie (np. FN, zwanej **poprzednikiem anaforycznym**, i odnoszącym się do niej zaimkiem)

SLS - podstawowe założenia

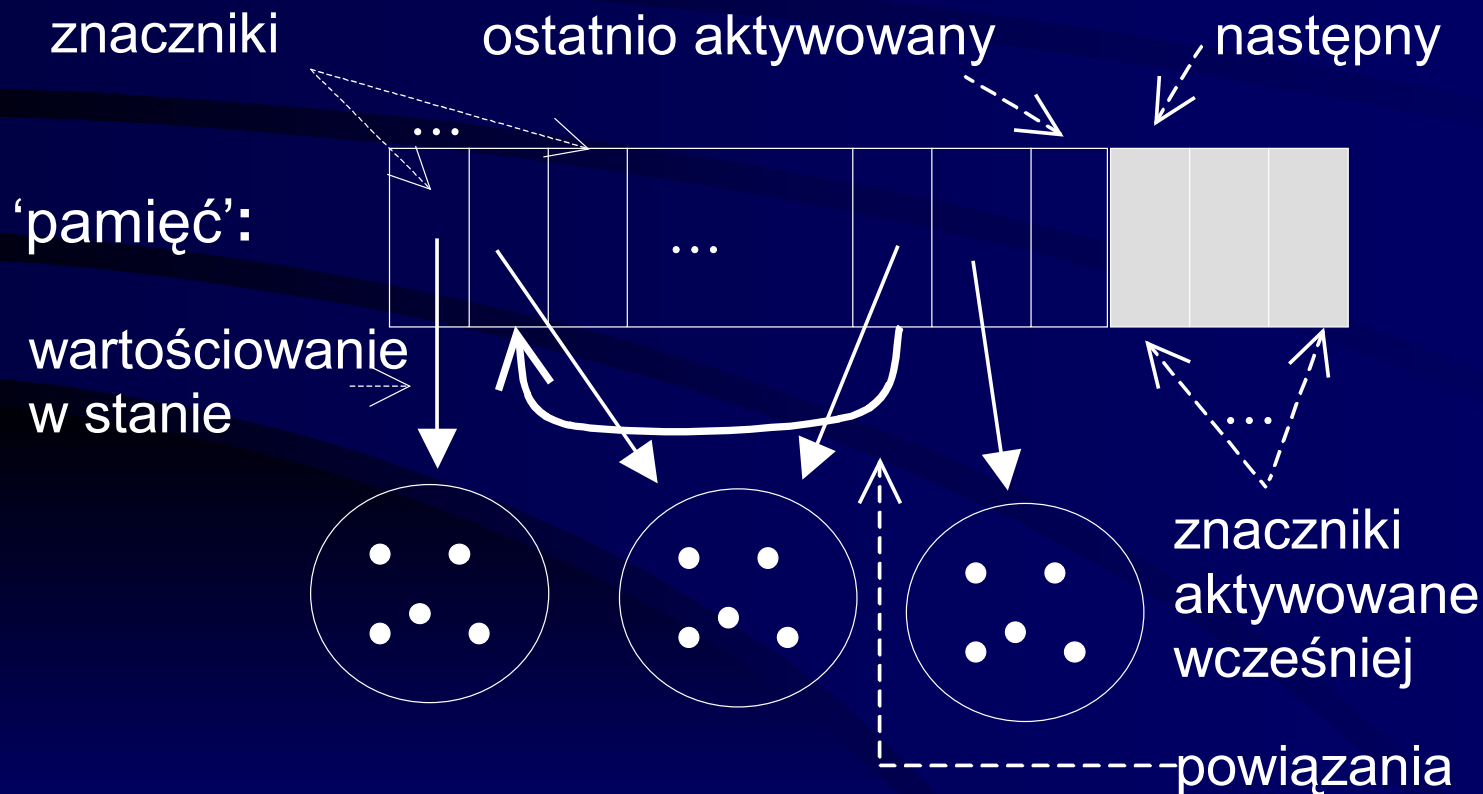
- Punkt startu to **kompozycyjna reprezentacja anafory**
 - referencyjna FN – oczekiwanie identyfikacji referenta
 - anaforyczna FN – oczekiwanie identyfikacji poprzednika anaforycznego
 - wspólny jest proces identyfikacji elementu reprezentacji kontekstu interpretacji
 - element ten nazywany będzie **znacznikiem dyskursu**
 - identyfikacja na poziomie języka przedmiotowego
- Konieczność reprezentacji
 - wykorzystywanych znaczników
 - zmiennych powiązań pomiędzy znacznikami
- **Dynamiczna koncepcja znaczenia**
 - znaczenie jako relacja na stanach kontekstu interpretacji

SLS – charakterystyka formalna

- SLS to język logiki wyższego rzędu z typami
- Typy podstawowe
 - e (*entities*) – wartości typu e reprezentują byty
 - t (*truth values*) – wartości logicznych
 - m (*'memory'*) – wartości typu m to **znaczniki dyskursu**
 - denotacja typu m to dowolny nieskończony zbiór z nałożonym porządkiem liniowym $<_M$
 - istnieje element minimalny P_0
- Konstruktory typów złożonych
 - typów funkcyjnych
 - typów, których denotacja to produkt kartezjański

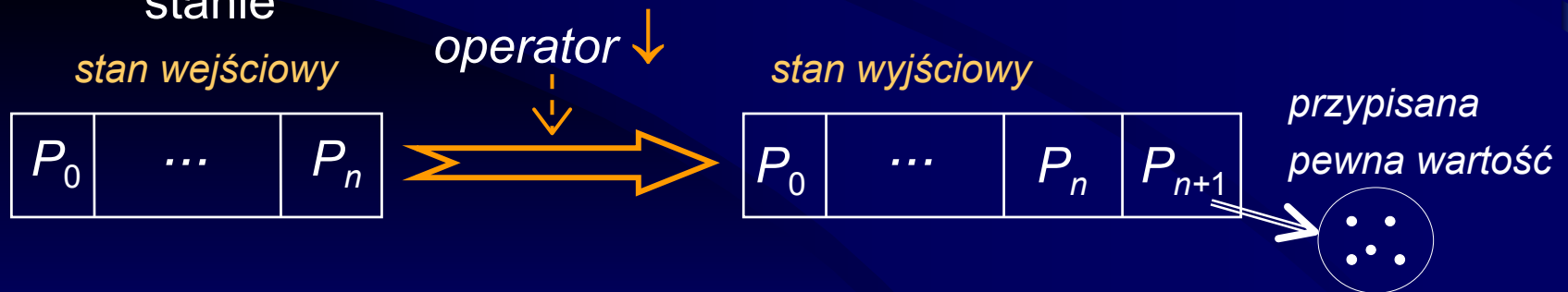
SLS – stany i dynamika znaczenia

- **Stan** to element denotacji złożonego typu
- *Stan początkowy* $S_0 = \langle P_0, \emptyset, \{\langle P_0, \emptyset \rangle\} \rangle$



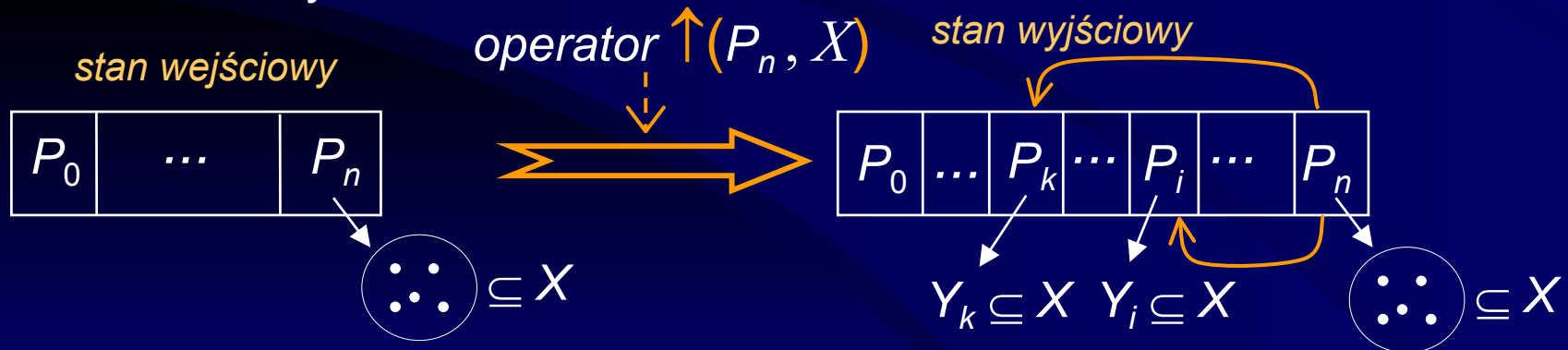
SLS - komponent kontekstu (1)

- **Formuły dynamiczne** - termy interpretowane jako relacje na stanach
- Aktywacja znacznika:
 - \downarrow – operator – relacja na stanach
 - zwiększa zbiór aktywnych znaczników o następnik ostatnio aktywowanego znacznika
 - rozszerza niedeterministycznie funkcję wartościującą w stanie



SLS - komponent kontekstu (2)

- Operator referencji \uparrow
 - dla podanego znacznika i zbioru obiektów generuje relację na stanach
 - odnajduje spośród znaczników dostępnych (warunek strukturalny) te, które są dopasowane pod względem właściwości (warunek semantyczny),
 - zawieranie się pojęciowe wartości w zbiorze podanym jako argument
 - dodaje dla odnalezionych znaczników pary do listy powiązań w danym stanie

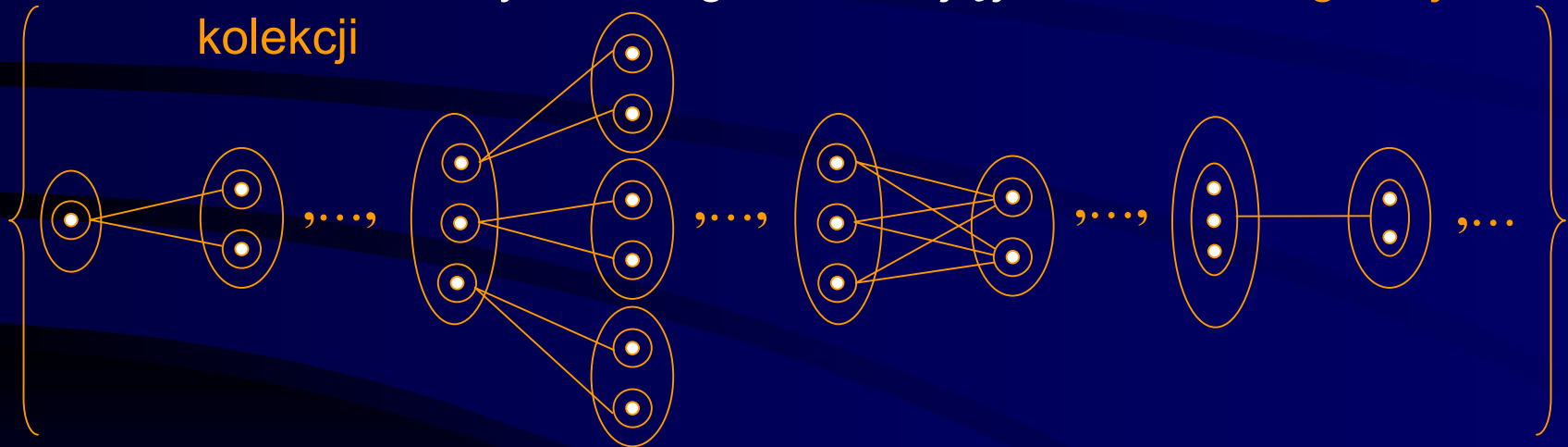


SLS - komponent kontekstu (3)

- Pseudo-operatory (skrótowy notacyjny)
 - dynamicznej: **negacji** (not), **implikacji** (\Rightarrow) i **dysjunkcji** (or)
 - łączą formuły dynamiczne
 - 'domykają' zbiór aktywnych znaczników
 - operator **sekwencji** (;)
 - odpowiada koniunkcji
 - stan wyjściowy pierwszej formuły staje się stanem wejściowym drugiej
- Presupozycja
 - operatory **presupozycji unikalnej** i **słabej**
 - warunek możliwości zmiany stanu (aktualizacji)
 - rozszerzenia listy powiązań o **dokładnie jedną** lub **przynajmniej jedną** parę znaczników dyskursu

SLS – aspekt deskryptywny

- Predykaty
 - ‘rzeczownikowy’ – którego denotacją jest zbiór obiektów
 - ‘czasownikowy’ – którego denotacją jest **zbiór konfiguracji kolekcji**

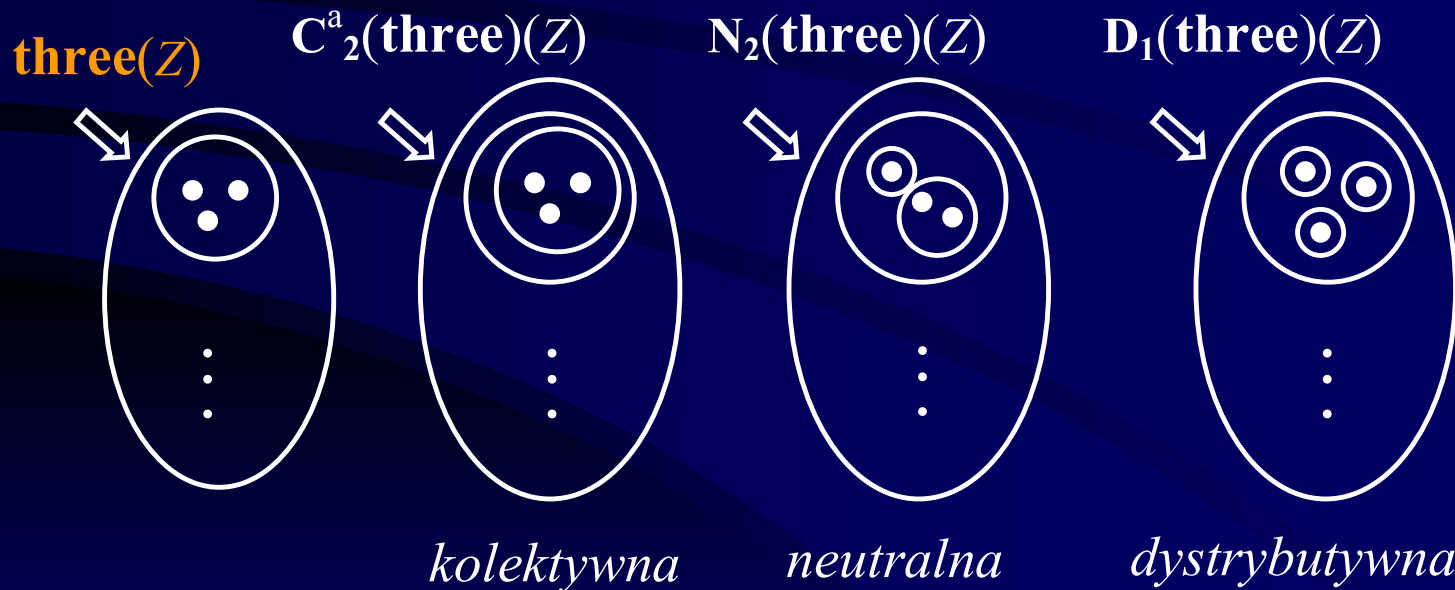


- Operatory pobrania dla podanego stanu
 - ostatnio aktywowanego znacznika
 - i wartości podanego znacznika

SLS - kwantyfikacja i mnogość

- Proto-kwantyfikatory
 - funktor generujący **kwantyfikator uogólniony**
 - wykorzystywane w reprezentacji aspektu kwantyfikacji

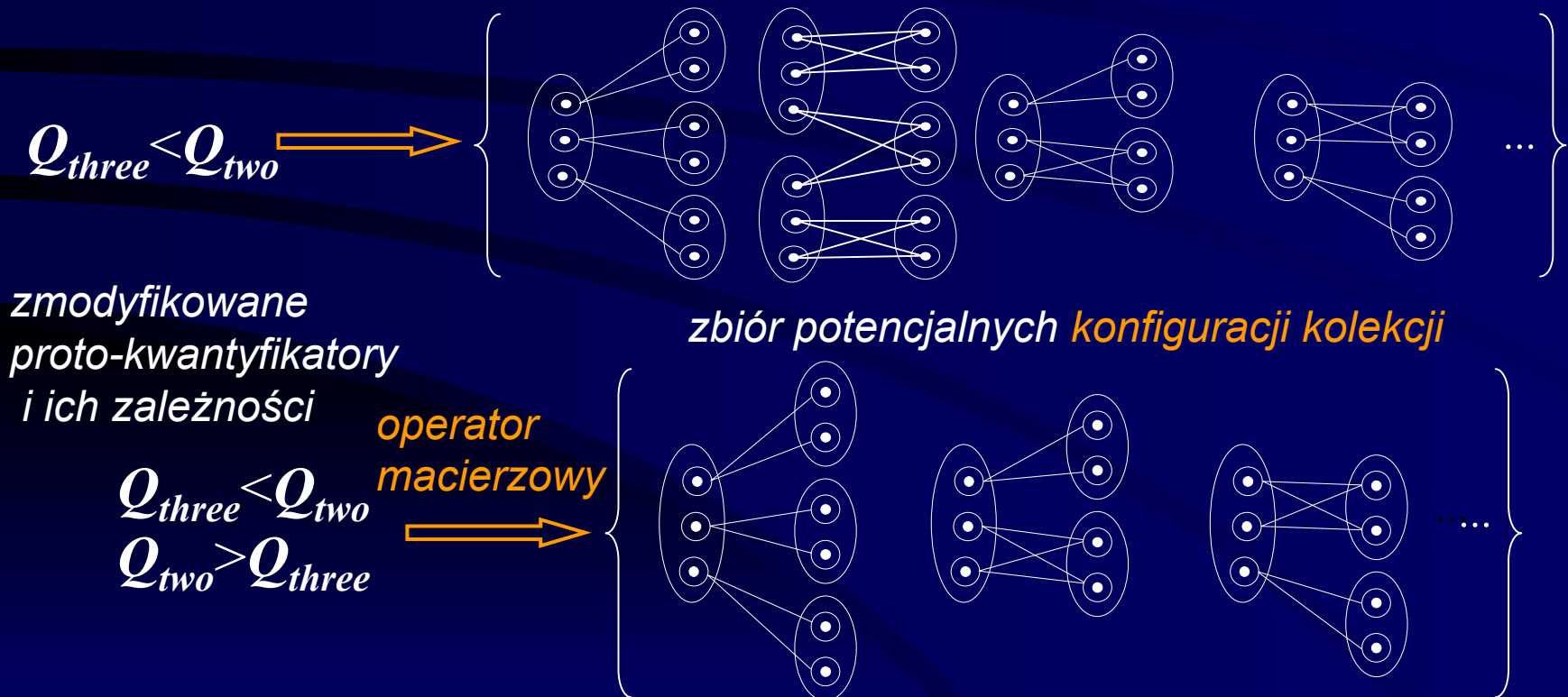
Niech Z to wartość znacznika



three := $\lambda X. \lambda Y. [|X \cap Y| = 3 \wedge Y \subseteq X]$

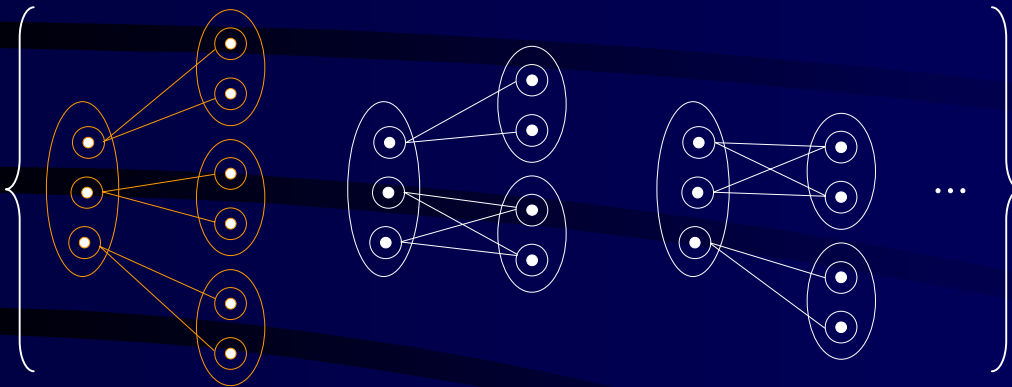
SLS – kwantyfikacja jako opis struktur relacji (1)

- Pojęcie **zależności liczebnościowej**
 - operatory niezależności i zależności liczebnościowej '<'

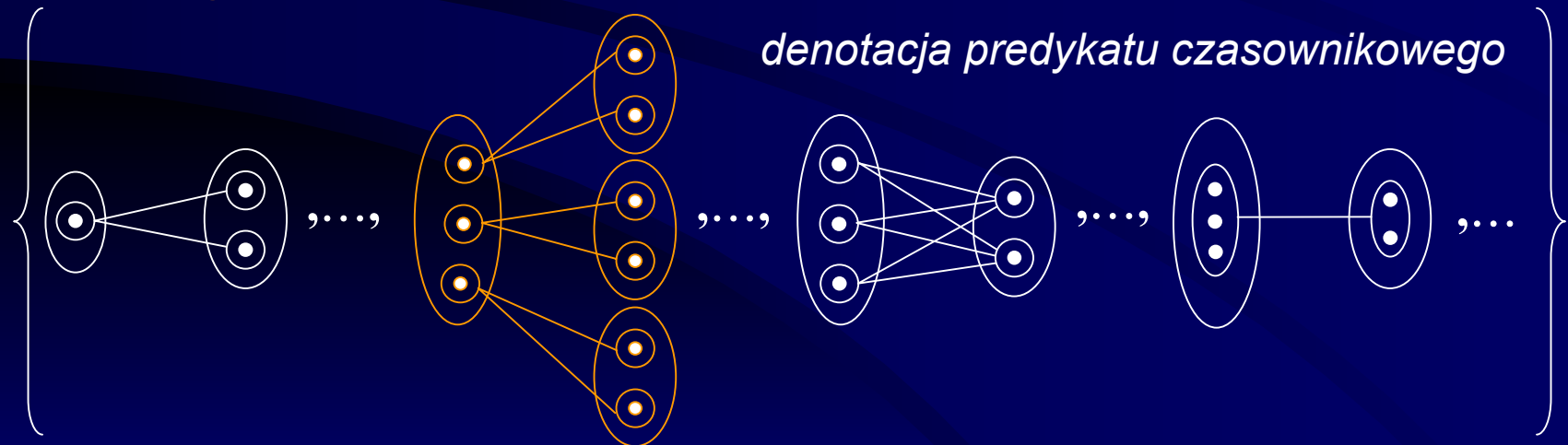


SLS – kwantyfikacja jako opis struktur relacji (2)

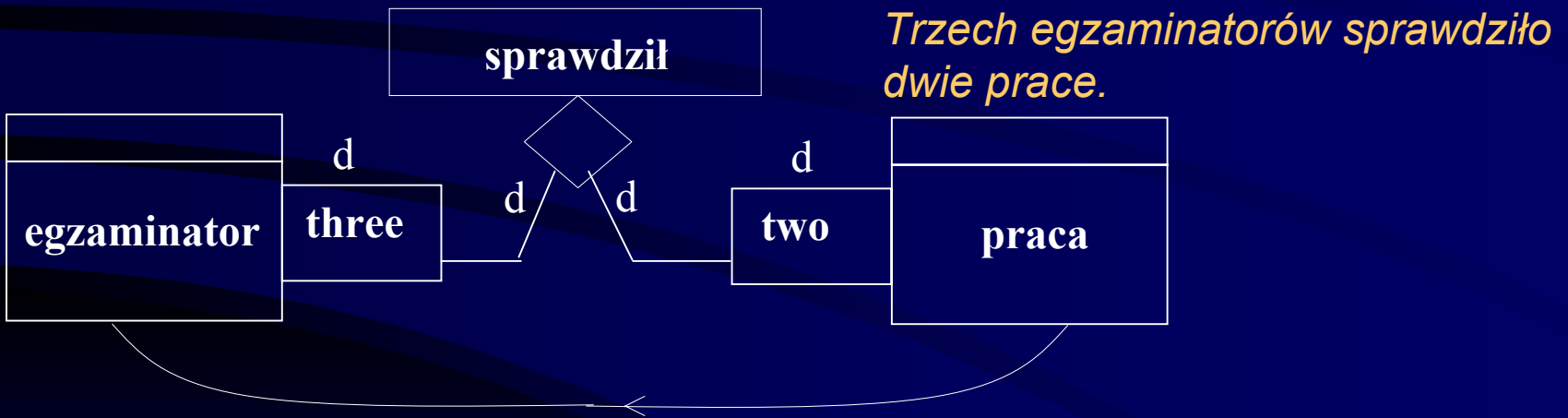
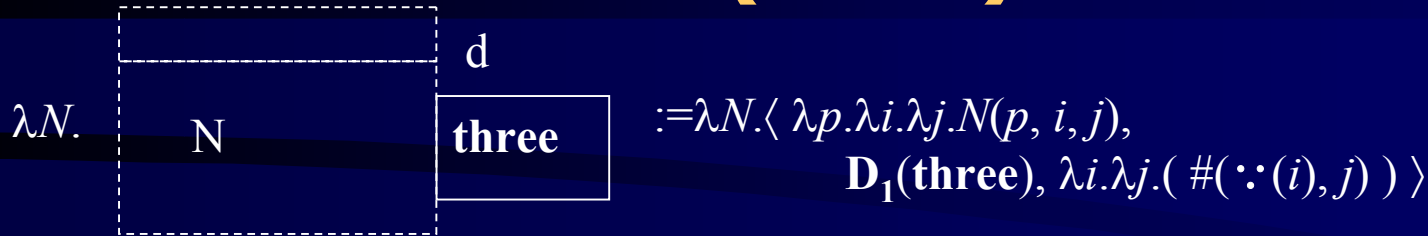
zbiór potencjalnych konfiguracji kolekcji



denotacja predykatu czasownikowego



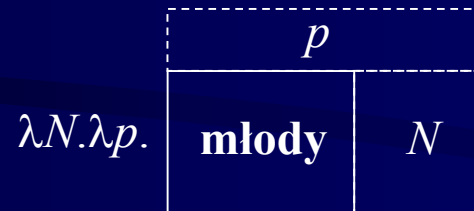
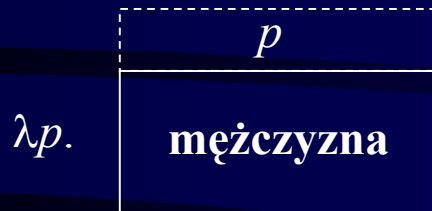
Notacja graficzna podzbioru SLS (GSLS)



$$\begin{aligned}
 := & \lambda i. \lambda j. \exists k. (\exists r. (\downarrow(i, r) \wedge ?^1(\lambda i. \lambda j. (\#(\nabla(r), i) \subseteq \text{egzaminator} \wedge i=j), \nabla(r), r, k)) \\
 & \wedge \exists l. (\downarrow(k, l) \wedge ?^1(\lambda i. \lambda j. (\#(\nabla(l), i) \subseteq \text{praca} \wedge i=j), \nabla(l), l, j)) \wedge \\
 & ?^2(\lambda i. \lambda j. \otimes^2(\Phi(\delta, \delta, \text{sprawdził}), \mathbf{M}^2(\omega(<), \omega(>), \mathbf{D}_1(\text{three})(\#(\cdot : (i), j)), \\
 & \mathbf{D}_1(\text{two})(\#(\cdot : (k), j)), \#(\cdot : (i), k), \#(\cdot : (k), j)), \cdot : (i), \cdot : (k), i, j)))
 \end{aligned}$$

Wyrażenia podstawowe GSLS (1)

- wyrażenie predykatu rzeczownikowego



- przykłady wyrażen kwantyfikatora uogólnionego



$:=\lambda N.\langle \lambda p.\lambda i.\lambda j.N(p, i, j),$
 $\mathbf{D}_1(\mathbf{three}), \lambda i.\lambda j.(\#(\cdot:(i), j)) \rangle$

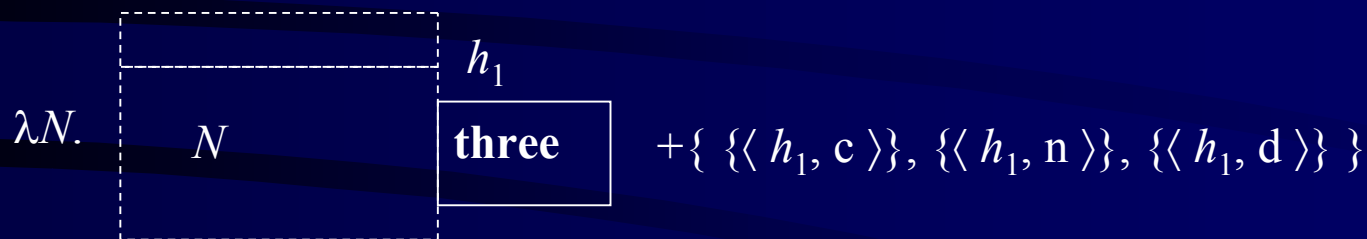
kwantyfikator absolutny



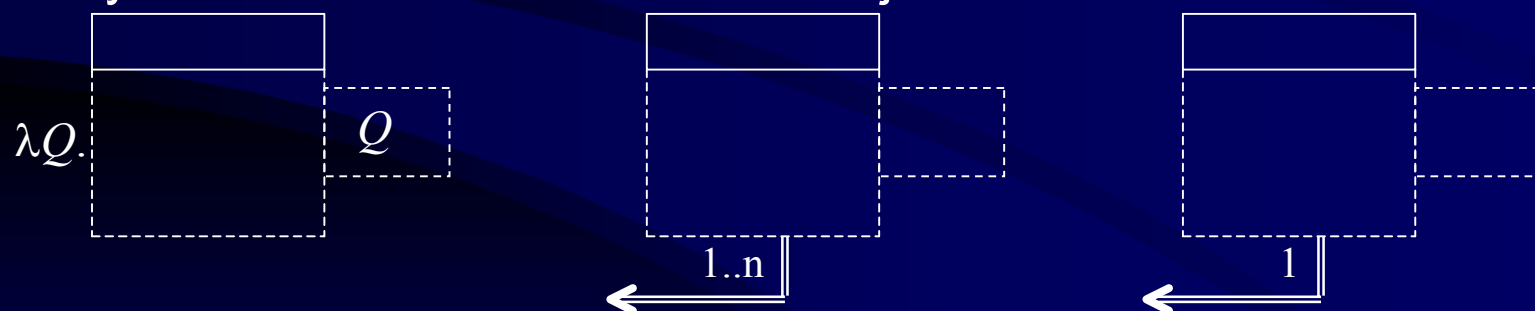
kwantyfikator relatywny

Wyrażenia podstawowe GSLS (2)

- niedospecyfikowane wyrażenie kwantyfikatora



- wyrażenie zależności kontekstowej



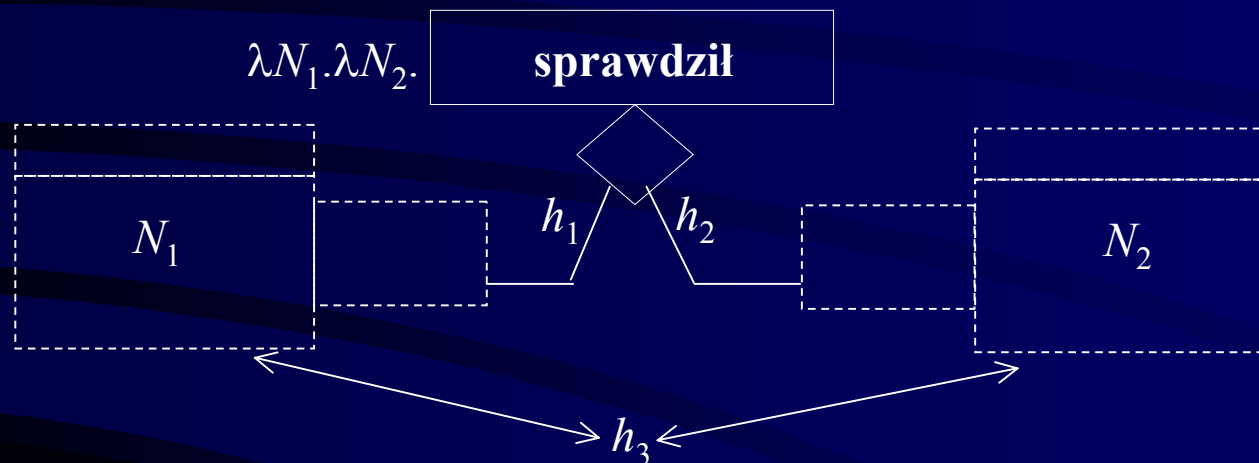
aktywacja znacznika

aktywacja znacznika +
operator referencji +
presupozycja słaba

aktywacja znacznika +
operator referencji +
presupozycja unikalna

Wyrażenia podstawowe GSLS (3)

- niedospecyfikowane wyrażenie predykatu czasownikowego



$$+ \{ \{ \langle h_1, o_1 \rangle, \langle h_2, o_2 \rangle, \langle h_3, z \rangle : o_1, o_2 \in \{c, n, d\} \wedge z \in \{ \langle ' < ', '> ' \rangle, \langle ' > ', '< ' \rangle, \langle ' : ', ': ' \rangle \} \wedge \text{cdv_match}^2(\langle o_1, o_2 \rangle, z) \}$$

Zastosowanie SLS jako JRZ

- Dostosowanie formalizmu gramatyki opartego na unifikacji wartości atrybutów
- Konstrukcja uproszczonego mechanizmu niedospecyfikowanej reprezentacji semantycznej
- Definicja leksykonu znaczeń poszczególnych leksemów
- Konstrukcja reguł gramatyki opisujących:
 - podzbiór języka angielskiego (porównanie siły ekspresji)
 - podzbiór języka polskiego (model znaczenia FN)

Formalizm gramatyki i mechanizm niedospecyfikowania

- Mechanizm niedospecyfikowania (*N-SLS*)

- wprowadzenie meta-zmiennych do wyrażeń

$$\lambda N_1. \langle \lambda i. \lambda j. \exists k. (\downarrow(i, k) \wedge N_1(\nabla(k), k, j)),$$

$$\lambda i. \lambda j. h_{1_n}(\mathbf{two})(\#(\cdot : (i, j)), \omega) + \{ \langle h_1, c \rangle \}, \{ \langle h_1, n \rangle \}, \{ \langle h_1, d \rangle \} \}$$

- określenie reguł podstawiania wyrażeń za meta-zmienne
- mechanizm deterministycznego przemianowania meta-zmiennych

- Schemat reguły gramatyki (*SN-SLS*):

$$Kat_w + R(t_1, \dots, t_n) + L(F_1, \dots, F_n) \rightarrow Kat_1 + t_1 + F_1 \dots Kat_n + t_n + F_n \{Cond\}$$

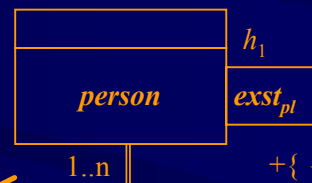
- reguły leksykalne – specyfikacja formy leksemu
- reguły konstruuujące frazę

Porównanie SN-SLS z innymi teoriami znaczenia dyskursu

- Gramatyka opisująca podzbiór języka angielskiego

- przykład reguły leksykalnej

PRO[Num=sing, Gen=male, Case=+nom/-nom, Quant=ind, Refer=+, Var= h_1 , Lex="they/them"] +



+{ {h_1, c}, {h_1, n}, {h_1, d}}

- przykład reguły konstruujące frazę

S[Num= α] + V([$\Theta(F_V, F_N)$]N)

+ { f : $\exists f_V \in F_V. \exists f_N \in F_N.$

$(f = f_V \bullet (\Theta(f_V, f_N) \circ f_N) \wedge [f]([\Theta(f_V, f_N)]u) = [f]([\Theta(f_V, f_N)]NP.Var) \}$ →

NP[Num= α , Gen= β , Case=+nom, Gap= -, Var= v] + N + F_N

VP'[Num= α , Tense= η , VarLst= uw] + V + F_V

{ $\exists f_V \in F_V. \exists f_N \in F_N. [f_V]u = [f_N]v$ }

Propozycje rozwiązań dla wybranych problematycznych konstrukcji językowych i znaczeń

Model znaczenia polskiej frazy nominalnej

- Gramatyka podzbioru języka polskiego oparta na *SN-SLS*
 - opis składniowy
 - szczegółowy w zakresie FN
 - uproszczony w zakresie opisu ‘frazy czasownikowej’ i zdania
 - kompozycyjny opis semantyczny
- Określenie możliwych znaczeń fraz FN w poszczególnych konstrukcjach ze względu na identyfikowane aspekty:
 - weryfikacja na podstawie znaczeń opisywanych w literaturze lingwistycznej
 - oraz kompetencję językową autora

Struktura składniowo- semantyczna FN – punkt wyjścia

- Analizy strukturalno-funkcjonalne
 - Topolińska (1984), Grzegorzczkowa (1996)
- Analizy ‘czysto’ semantyczne
 - Grzegorzczkowa (1996), Kosseska-Toszewa (1992), Karolak (1990), Borkowska (1984)
- Nie sformalizowane lub ‘częściowo’ sformalizowane analizy składniowe
 - Saloni i Świdziński (1998)
- Sformalizowane opisy polskiej składni
 - ‘powierzchniowe’ gramatyki DCG
 - Szpakowicz (1983), Świdziński (1992), Vetulani (w przyg., POLINT)
 - prace oparte na paradygmacie *Government and Binding Theory*
 - Węgrzynek (1995), Derwojedowa et. al. (w przyg.)
 - prace zrealizowane w formalizmie HPSG
 - Przepiórkowski (1999), Marciniak (2002), Mykowiecka (2001), Przepiórkowski et. al. (2001)
- Prace kognitywne dotyczące semantyki FN
 - Szwedek (1976)

Pojęcie determinatora

(K. Węgrzynek „Składnia przymiotnika w ujęciu generatywno-transformacyjnym” 1995)

- Składniowe, formalne kryterium przynależności do kategorii przymiotnika:
 - podstawowe pozycje w strukturze głębokiej
 - VP → V AdjP - pozycja przyłącznikowa
 - NP1 → AdjP N - pozycja w FN (przrzecznikowa),
 - NP1 → N AdjP - pozycja we FN
 - test podstawieniowy
 - 1) *Chłopak jest wysoki.*
 - 2) *błędne koło, ślepa uliczka*
 - 3) *tygrys bengalski, logarytm naturalny*
 - 4) **tygrys groźny bengalski, *bardzo biały kruk*
 - 5) **Chłopiec jest ten / ów / jakiś / pewien / żaden / każdy / niejeden.*

Podstawowa struktura składniowo-semantyczna FN (1)

- Brak całościowego formalnego opisu składni FN
- Mechanizmy obecne w większości FN:
 - aktywacja znacznika i pewna forma kwantyfikacji
- Propozycja syntetyczna

– wstępny szkielet:

FN → DetP NumNP

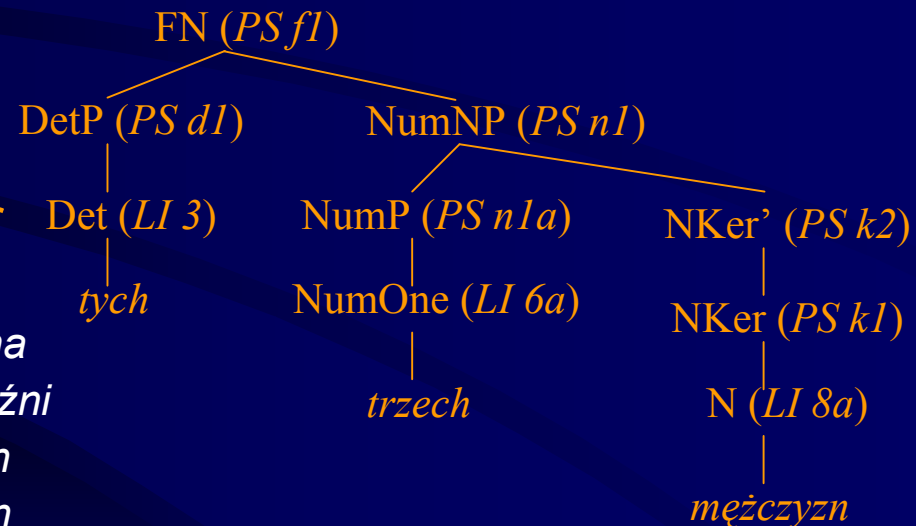
DetP → (ModP) Det

NumNP → NumP Nker

– przykłady:

- 1) *każdy ten mężczyzna
- 2) mężczyzna / mężczyźni
- 3) tych trzech mężczyzn
- 4) jakiś pięciu mężczyzn
- 5) ?każdych trzech mężczyzn

- obserwacja: maksimum jedna DetP ~ maksimum jeden znacznik



Podstawowa struktura składniowo-semantyczna FN (2)

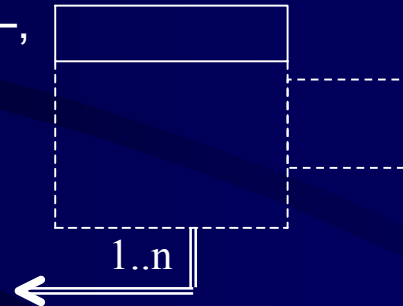
- Det jako źródło aktywacji (c.d.)
 - problem: liczne FN pozbawione DetP
 - za: FN w konstrukcji z atrybutywnym *być* nie mogą mieć DetP i nie mogą być anaforyzowane
 - 1) *Jan jest **dobrym lekarzem**¹. **Ten lekarz**₁ ...*
 - 2) *Jan jest ?**tym** / ?**jakimś** / ***pewnym lekarzem**.*
- Lokalizacja mechanizmu kwantyfikacji
 - potencjalne ‘źródła’:
 - reprezentacja semantyczna NumP
 - morfologiczna liczba mnoga
 - problem: aspekt kwantyfikacji w wielu Det np. *jakiś*
 - 1) *tych trzech mężczyzn*
 - 2) *?każdych trzech mężczyzn*
 - 3) *?wszyscy (**ci**) trzej mężczyźni*
 - 4) *każdy z tych mężczyzn*

Reprezentacja niekwantyfikujących Det

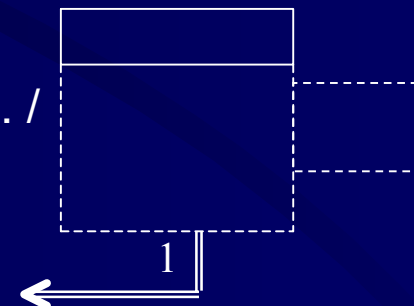
LI1) Det[Case=mian/dop/.../woł, GNum=mos-poj / mos-mno / ... / pt-mno, Neg= -/+ , Quant= -, Lex= *jakiś* / *jacyś* / ...]+



LI2) Det[Case=mian/dop/.../woł, GNum=mos-poj / mos-mno / ... / pt-mno, Neg= -/+ , Quant= -, Lex= *pewien* / *pewni* / ...]+



LI3) Det[Case=mian/dop/.../woł, GNum=mos-poj / ... / pt-mno, Neg= -/+ , Quant= -, Lex= *ten* / *ci* / ...]+



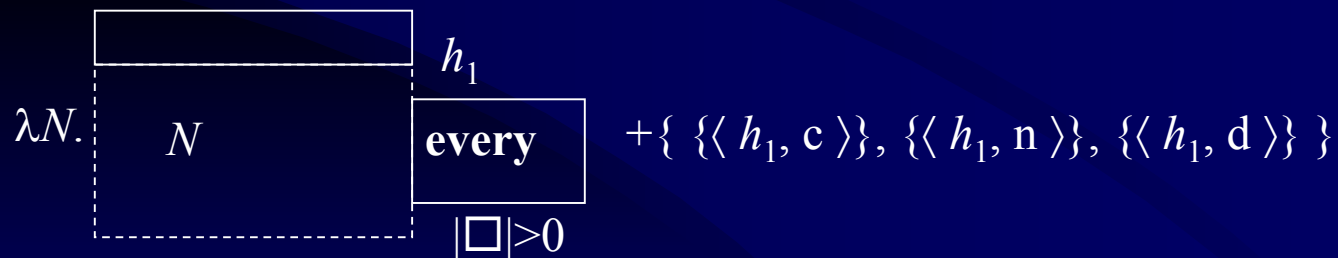
Kwantyfikujące Det

L14a) Det[Case=mian/dop/.../woł, GNum=mos-poj / mzw-poj / mnż-poj / ... ,
 Neg= -/+ , Quant= + , Var=d,
 Lex= *każdy / każdy / każde / ...*]+



$$\|\text{every}\| := \lambda X. \lambda Y. [X=Y]$$

L14b) Det[Case=mian/dop/.../woł, GNum=mos-mno / mzw-mno / mnż-mno / ..., Neg= -/+ , Quant= + , Var= h_1 , Lex= *wszyscy / wszystkie / ...*] +

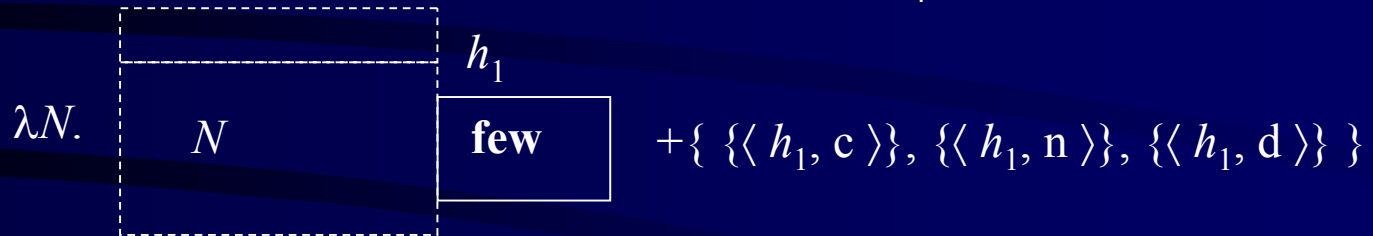


Kwantyfikujące Det - przykłady

- Odmiana *każdy* i *wszyscy*
 - 1) **Każdy* mieszkaniec zgromadził się na placu.
 - 2) Jan zapakował *wszystkie* książki, część razem, część osobno do pudełek.
 - 3) *Wszyscy* zawodnicy zaliczyli po jednej wysokości.
- Relatywny charakter
 - 4) **Istnieje(a) każdy* człowiek / *wszyscy* ludzie, który / którzy...
 - 5) **Istnieje ten* człowiek, który ...
 - 6) *Istnieje jakiś* / *pewien* uczeń, który ...

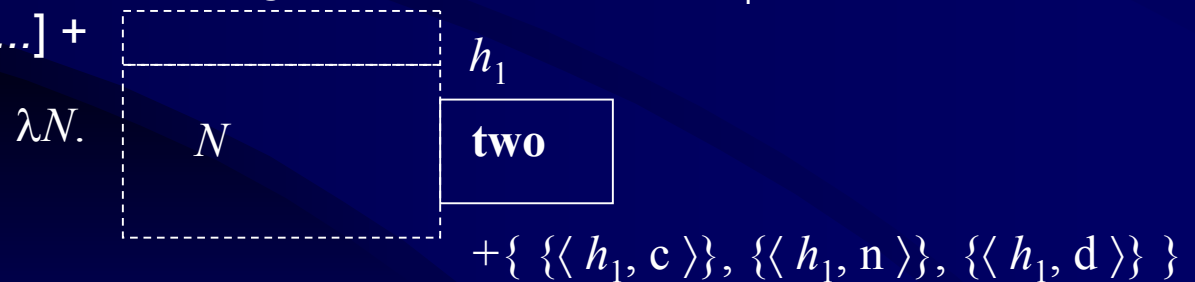
Liczebniki

LI 6) Num[Case=mian/dop/.../woł, GNum= mos-mno / mzw-mno / mnż-mno /..., NClass=C1, Neg= +, Quant= −, Var= h_1 , Lex= *kilku / kilka*] +



$\|\text{few}\| := \lambda X. \lambda Y. [|X \cap Y| > 2 \wedge |X \cap Y| < \varepsilon \wedge Y \subseteq X]$

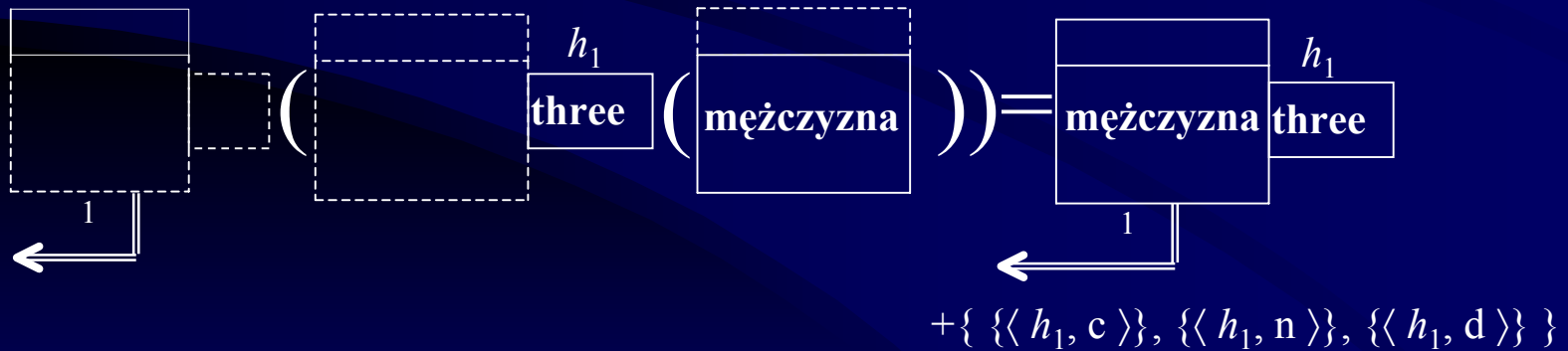
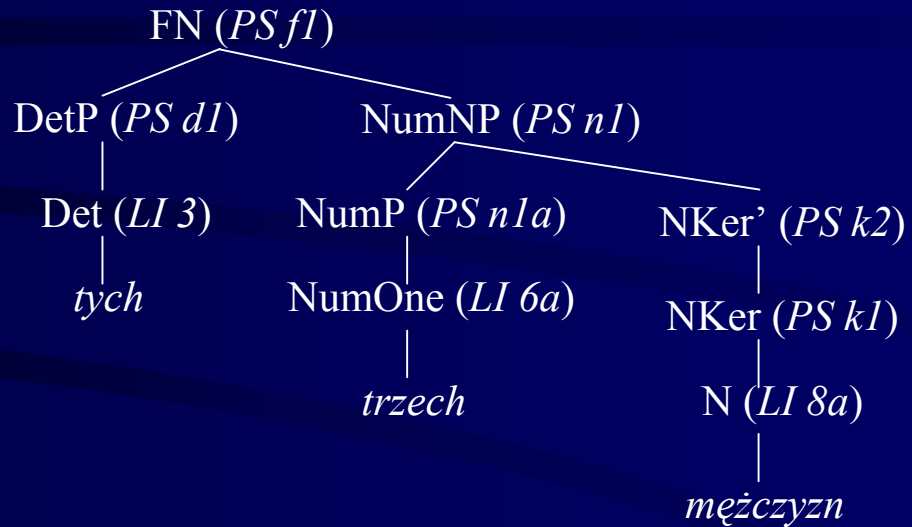
LI 6a) NumOne[Case=mian/dop/.../woł, GNum= mos-mno / mzw-mno / mnż-mno /..., NClass=C0, Neg= +, Quant= −, Var= h_1 , Lex= *dwaj / dwa...*] +



$\|\text{two}\| := \lambda X. \lambda Y. [|X \cap Y| = 2 \wedge Y \subseteq X]$

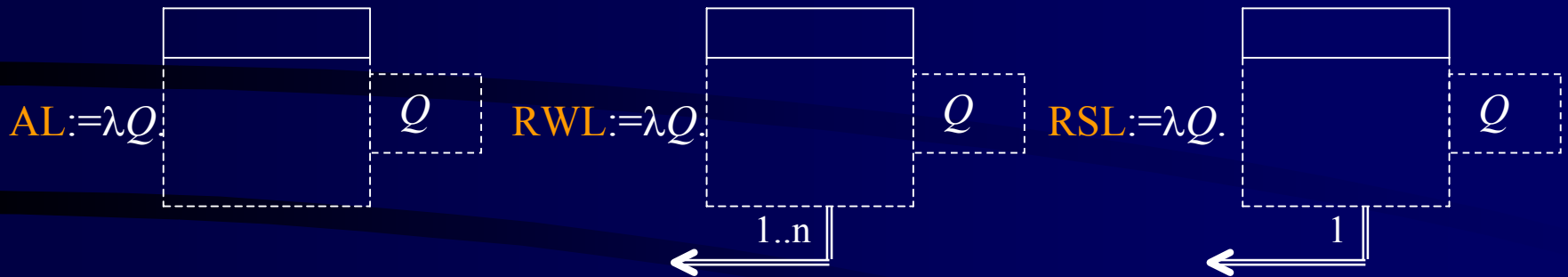
'Pełna' FN

tych trzech mężczyzn



Brak Det we FN

- *Modyfikatory referencyjności (dokonujące type lift)*



- Problem referencyjnego użycia FN z NumP

Na ulicę bawiąc się **piłką**₁ wyszła mała dziewczynka. **Piłka**₁ potoczyła się pod stopy jakiegoś przechodnia.

Dwóch / kilku profesorów₁ sprawdziło prace. **Profesorowie**₁ opuścili pokój.

***Dwóch/kilku profesorów**₁ sprawdziło prace. **Dwóch/ kilku profesorów**₁ opuściło pokój.

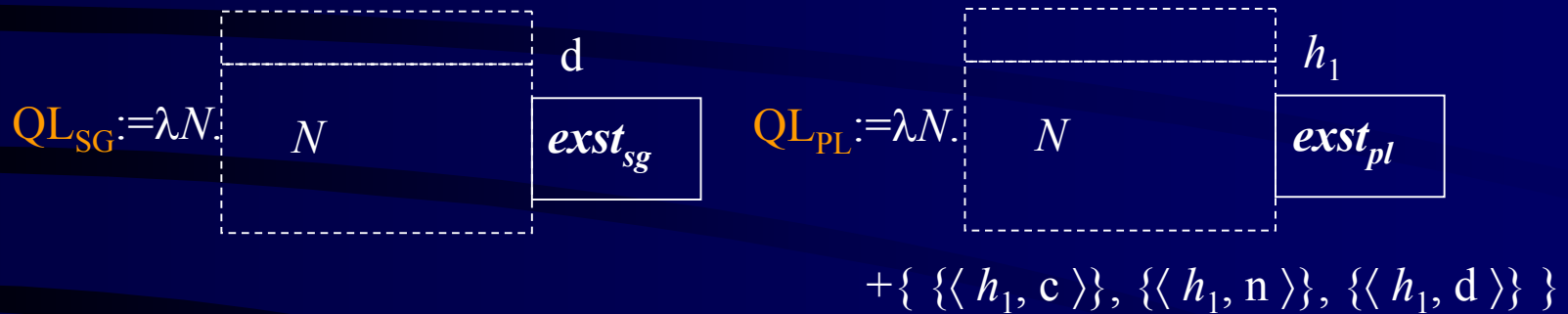
- Reguły

PS f2) $FN[... , Var=w, ...] + AL(N) + F_N \rightarrow NumNP[... , QLifted= -, ..., Var=w] + N + F_N$

PS f3) $FN[... , Var=w, ...] + h_1([\Theta(\{\{ \langle h_1, AL \rangle \}\}, F_N)]N) + \{ f: \exists f_D \in \{\{ \langle h_1, AL \rangle \}, \{ \langle h_1, RWL \rangle \}, \{ \langle h_1, RSL \rangle \} \}. \exists f_N \in F_N. (f = f_D \bullet (\Theta(f_D, f_N) \circ f_N)) \} \rightarrow NumNP[... , QLifted= +, ..., Var=w] + N + F_N$

Brak NumP we FN

- Modyfikatory kwantyfikacji
 - aktywowane morfologiczną liczbą

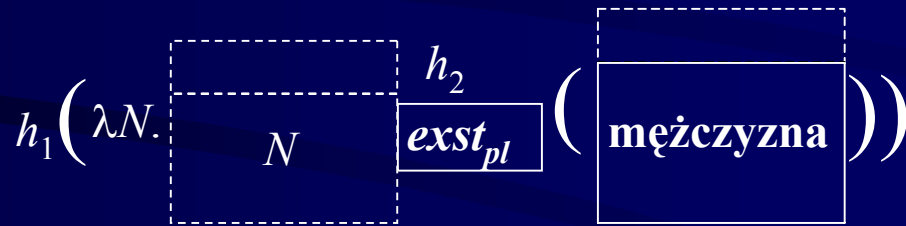
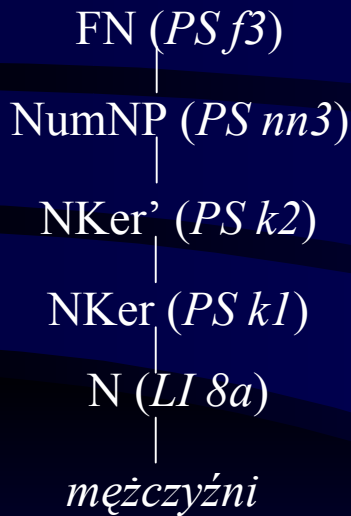


PS nn2) NumNP[\dots , $QLifted=+$, $NClass=C0$, $Var=d$] + $QL_{SG}(N)+F_N \rightarrow$
 NKer'[\dots , $GNum= mos-poj / mzw-poj / \dots$, $VarAgr=-$]+ $N+F_N$

PS nn3) NumNP[\dots , $QLifted=+$, $NClass=C0$, $Var=h_1$]
 + $QL_{PL}([\Theta(\{ \langle h_1, c \rangle \}), F_N])N$
 + $\{ f: \exists f_L \in \{ \{ \langle h_1, c \rangle \}, \{ \langle h_1, n \rangle \}, \{ \langle h_1, d \rangle \} \}. \exists f_N \in F_N. (f = f_L \bullet (\Theta(f_L, f_N) \circ f_N)) \} \rightarrow$
 NKer'[$Case=\alpha$, $GNum= mos-mno / mzw-mno / \dots$, $VarAgr=-$]+ $N+F_N$

Przykład zastosowania modyfikatorów

mężczyźni



+ { {⟨h₁,AL⟩,⟨h₂,c⟩}, {⟨h₁,RWL⟩,⟨h₂,c⟩},
 {⟨h₁,RSL⟩,⟨h₂,c⟩}, {⟨h₁,AL⟩,⟨h₂,n⟩}, ... }

$$\|exst_{sg}\| := \lambda X. \lambda Y. [(X \cap Y) \neq \emptyset \wedge Y \subseteq X]$$

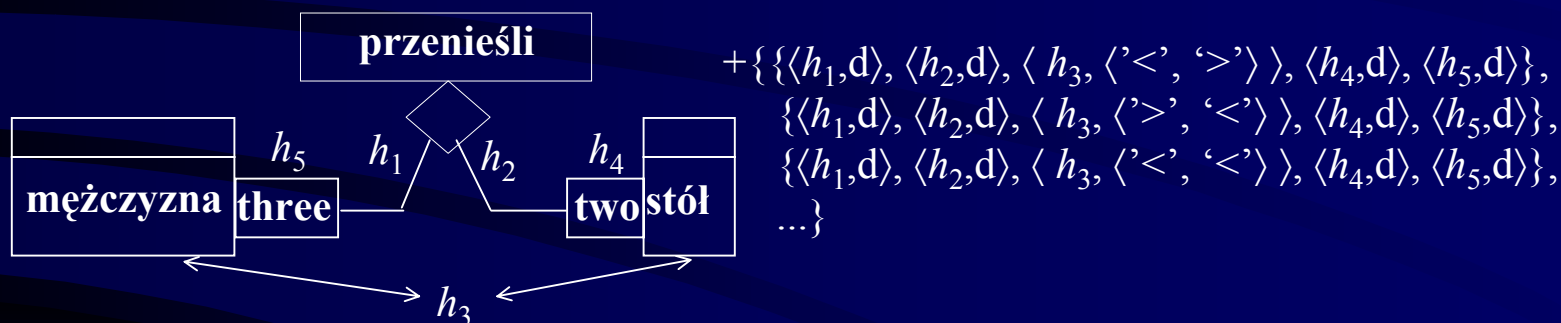
$$\|exst_{pl}\| := \lambda X. \lambda Y. [|X \cap Y| \geq 2 \wedge Y \subseteq X]$$

Zaproponowana struktura reprezentacji semantycznej FN

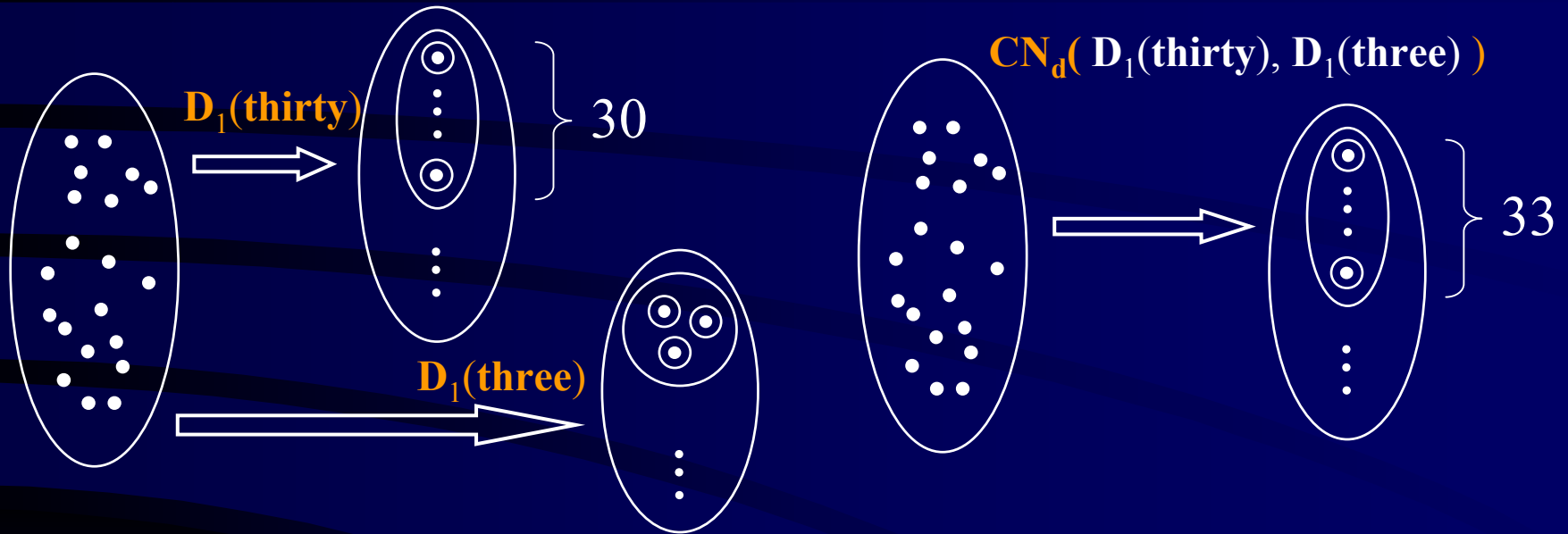
- Proponowana reprezentacja semantyczna FN to trójka:
 - aktywacja znacznika, opcjonalny operator referencji+modyfikator presupozycji, predykat reprezentujący aspekt deskryptywny
 - zmodyfikowany kwantyfikator = rodzina zbiorów kolekcji
 - reprezentacja wpływu referencyjności na relacje zależnościowe (tzw. *ogranicznik*)

Przykład niedospecyfikowanej reprezentacji znaczenia zdania

Trzech mężczyzn przeniosło dwa stoły.



Składanie (zmodyfikowanych) proto-kwantyfikatorów



$$CN_d := \lambda Q_1. \lambda Q_2. \lambda X. \lambda Z. \exists Y_1. \exists Y_2. (Q_1(X)(Y_1) \wedge Q_2(X)(Y_2) \wedge \cup Y_1 \cap \cup Y_2 = \emptyset \wedge Y_1 \cup Y_2 = Z)$$

$$CNM := \lambda L_1. \lambda L_2. \lambda N. \langle (L_2(N))^1, CN_d((L_1(N))^2, (L_2(N))^2), (L_2(N))^3 \rangle$$

$$CN_c := \lambda Q_1. \lambda Q_2. \lambda X. \lambda Z. \exists Y_1. \exists Y_2. (Q_1(X)(Y_1) \wedge Q_2(X)(Y_2) \wedge \cup Y_1 \cap \cup Y_2 = \emptyset \wedge Z(\cup(Y_1 \cup Y_2)) \wedge |Z|=1)$$

Reprezentacja liczebników złożonych

$$\begin{aligned}
 & \text{PS n3) NumTenP}[Case=\alpha, GNum=\eta, NClass=\beta, Var=w] \\
 & + \text{CNM}([\Theta(F_{\text{cnm}}, F_1)]L_1)([\Theta^3(F_{\text{cnm}}, F_1, F_2)]L_2) \\
 & + \{ f: \exists f_C \in F_{\text{cnm}}. \exists f_1 \in F_1. \exists f_2 \in F_2. (f = (f_C \bullet (\Theta(f_C, f_1) \circ f_1)) \bullet (\Theta^3(f_C, f_1, f_2) \circ f_2)) \\
 & \quad \wedge [f](\Theta(f_C, f_1)]u) = [f](\Theta^3(f_C, f_1, f_2) \circ f_2]w) \\
 & \quad \wedge (([f]h_1 = \text{CN}_c \wedge [f](\Theta(f_C, f_1)]u) = c) \\
 & \quad \vee ([f]h_1 = \text{CN}_n \wedge [f](\Theta(f_C, f_1)]u) = n) \\
 & \quad \vee ([f]h_1 = \text{CN}_d \wedge [f](\Theta(f_C, f_1)]u) = d)) \} \rightarrow
 \end{aligned}$$

$$\text{NumTen}[Case=\alpha, GNum=\eta, NClass=\beta, Var=u]+L_1+F_1$$

$$\begin{aligned}
 & \text{NumOne}[Case=\alpha, GNum=\eta, NClass=\mu, Var=w]+L_2+F_2 \\
 & \quad \{ \exists f_1 \in F_1. \exists f_2 \in F_2. ([f_1]u = [f_2]w) \}
 \end{aligned}$$

Koncepcja wykorzystania SLS w implementacji

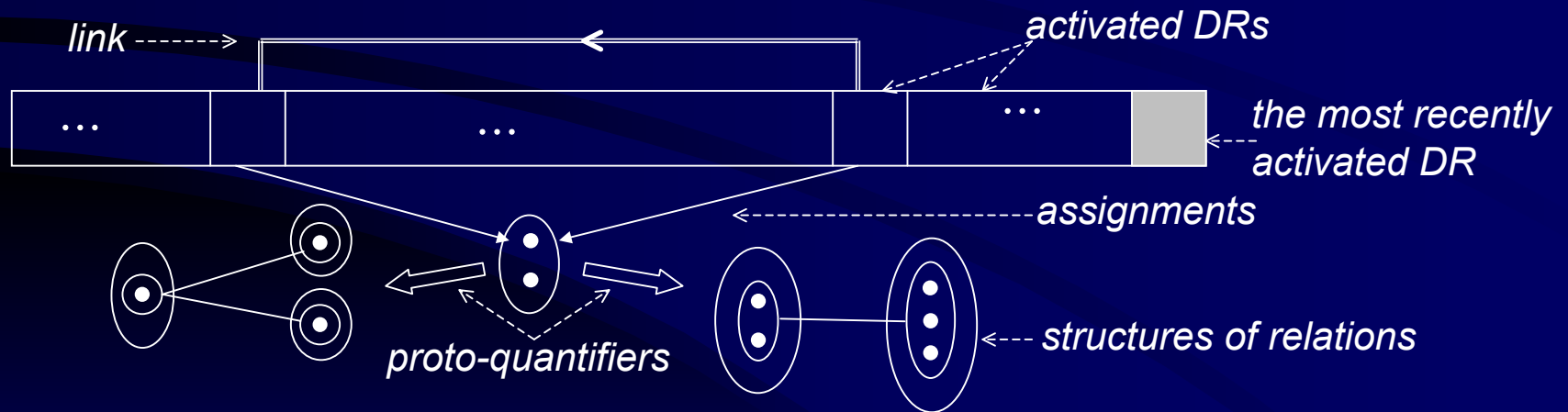
- System dialogu z użytkownikiem w języku polskim (ang. *Question and Answering*)
- Dwie drogi:
 - ‘klasyczna’ – realizacja implementacji w Prologu
 - poszukiwanie efektywnych metod przetwarzania
- Etapy przetwarzania:
 - wstępna segmentacja tekstu, analiza morfologiczna, parsing składniowy, konstrukcja reprezentacji semantycznej, interpretacja reprezentacji semantycznej (analiza pragmatyczna)
- Implementacja mechanizmu niedospecyfikowania
- Reprezentacja stanu:
 - znaczniki dyskursu jako zmienne Prologu
 - stan jako lista list (wykorzystanie meta-predykatów)
- Obsługa zapytań poprzez określenie możliwych wartości znaczników aktywowanych w formułach dynamicznych

Ograniczenia i dalsze prace

- Rozszerzenie opisu na inne rodzaje anafory np. anaforę do własności
- Rozszerzenie zakresu opisu na FN, których denotacją są masy lub substancje
- Szczegółowa analiza związku pomiędzy linearną pozycją FN w zdaniu a jej znaczeniem
 - wprowadzenie mechanizmu selektywnej eliminacji nadmiarowych wypełnień
- Bardziej szczegółowy opis frazy czasownikowej i anafory wewnątrzzdaniowej
- Integracja reprezentacji semantycznej z mechanizmami reprezentacji wiedzy
- Opracowanie systemu wnioskowania dla SLS
- Konstrukcja efektywnych algorytmów przetwarzania i wnioskowania

Sentences, States, Links and Discourse

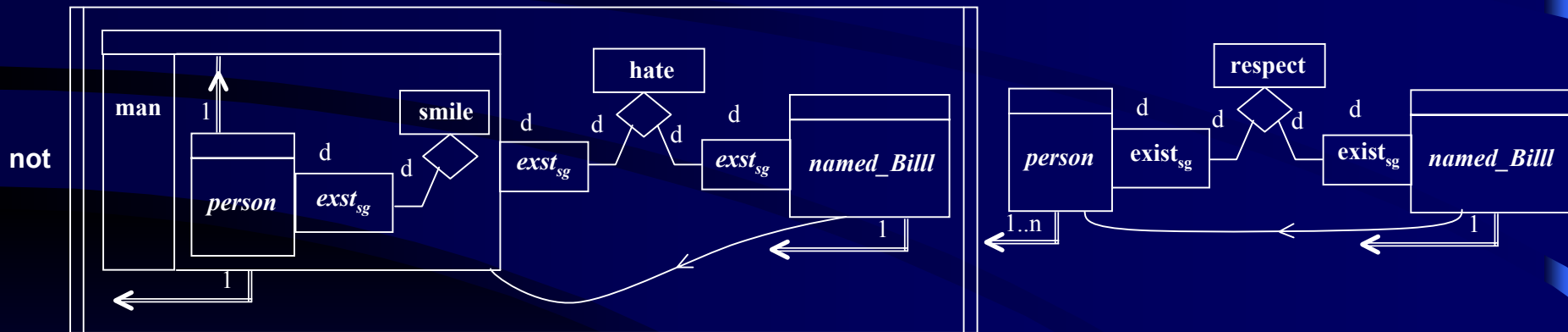
- Sequential merging:
 - the output state becomes the input state
- Dynamic formulae representing sentences are linked by self-organised links



Examples:

Reference and Accessibility (1)

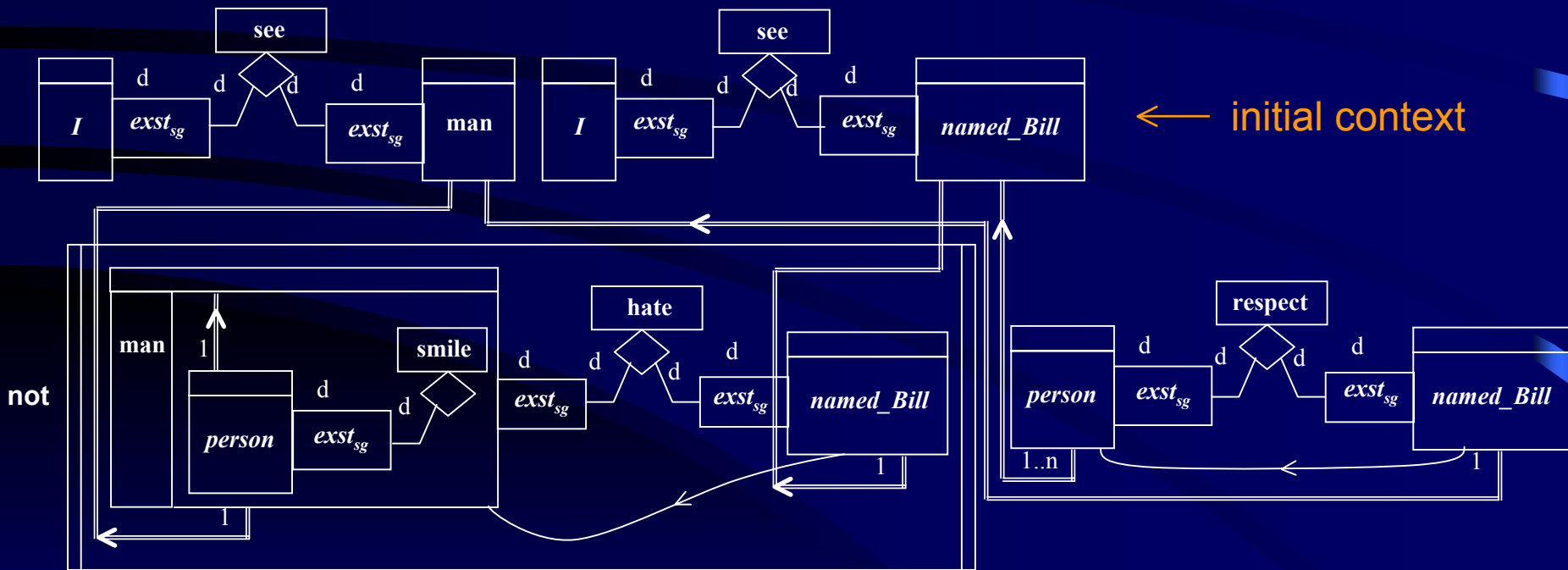
[The man who smiles]¹₂ does not hate Bill₃. He₁ respects Bill₃.



Examples:

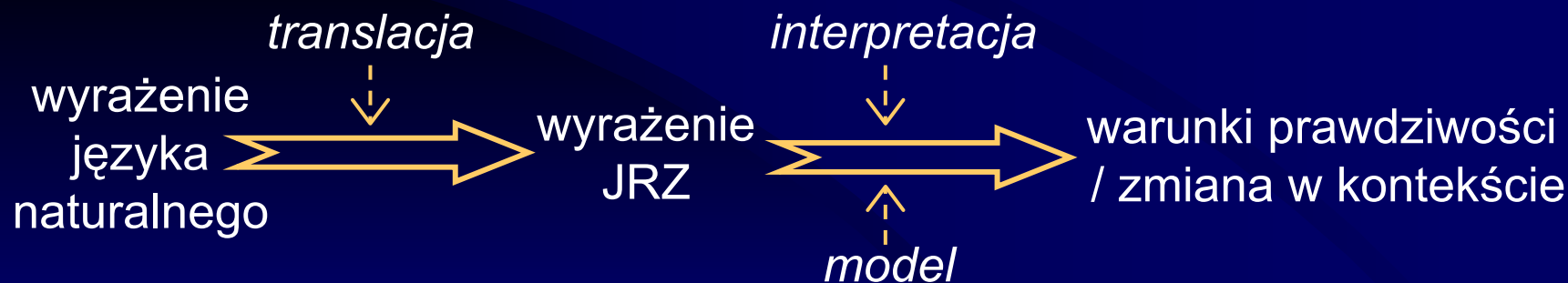
Reference and Accessibility (2)

[The man who smiles]¹₂ does not hate Bill₃. He₁ respects Bill₃.



Znaczenie i jego opis formalny

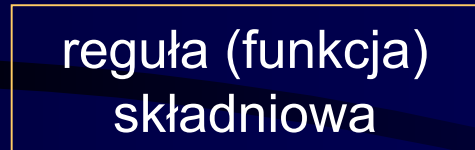
- Brak jednej, powszechnie akceptowanej koncepcji znaczenia [Encyklopedia JO]
 - znaczenie jako relacja dwuargumentowa:
 x znaczy y
- Języki reprezentacji znaczenia
 - języki formalne o precyzyjnie określonej semantyce



Kompozycyjna reprezentacja znaczenia

- Zasada kompozycyjności (Frege, Montague)

wyrażenia składowe



wyrażenie złożone

znaczenia wyrażen składowych



znaczenie wyrażenie złożone



- wprowadzona do logiki przez Fregego,
- zastosowana konsekwentnie w lingwistyce informatycznej przez Montague

Język reprezentacji znaczenia – rozwiązanie najprostsze

- Logika predykatów pierwszego rzędu (LP)
 - zalety:
 - prostota, dobrze poznane własności, łatwa implementacja
 - ograniczenia
 - kwantyfikacja (mnogość) – np. *większość mężczyzn*
 - anafora [[Encyklopedia JO](#)]
„zjawisko składniowe – relacja, która wiąże ze sobą 2 wyrażenia : anaforyzowane i anaforyczne [...] Wyrażenie anaforyczne, które nie ma własnego sensu, stanowi swoisty odsyłacz do wcześniej użytego wyrażenia anaforyzowanego, zwanego poprzednikiem (antecedensem) relacji anaforycznej, z którego odczytać można komunikowany sens.”

LP – problemy z anaforą

- Ośle zdania
 - *If Pedro¹ owns a donkey² he₁ beats it₂.*
 - $\exists x[(donkey(x) \wedge own(pedro, x)) \rightarrow like(pedro, x)]$
 - $\forall x[(donkey(x) \wedge own(pedro, x)) \rightarrow beat(pedro, x)]$
- Wpływ negacji na anaforę
 - *A man¹ is walking in the park. He₁ is whistling.*
 - **It is not the case that every man¹ is not walking in the park. He₁ is enjoying himself₁.*

Dynamiczna koncepcja znaczenia na przykładzie DRT

- *Discourse Representation Theory (DRT)*
- *Discourse Representation Structures (DRS)*

x
man(x) walk_in_park(x)

x,y
man(x) walk_in_park(x) y=x gen(x, male) whistle(y)

- **znaczniki dyskursu**
(*discourse markers/ referents*)
- **warunki** (*conditions*)
- **znaczenie = zmiana w kontekście interpretacji**

Problemy kompozycyjnych wersji DRT

- Cechy wspólne:
 - rachunek lambda z typami
 - operacja sekwencyjnego łączenia DRS'ów
- Problemy z reprezentacją anafory
 - większość rozwiązań zakłada uprzednią koindeksację powiązań anaforycznych na poziomie składniowym
 - przypisanie znaczników na podstawie indeksów
 - $\|a^1\|^{\text{CDRT}} := \lambda P' \lambda P([u_1]); P(u_1); P'(u_1)$
 - $\|man\|^{\text{CDRT}} := \lambda v [\textit{farmer}\{v\}]$
 - $\|him_1\|^{\text{CDRT}} := \lambda P(P(u_1))$
 - u_1 został zwrócony przez funkcję odczytującą drzewo struktury składniowej

SLS - typy

- Typy podstawowe

- e (entities) $\mathbf{D}_e \neq \emptyset$, t (truth values) $\mathbf{D}_t = \{0, 1\}$

- m (znaczników, 'memory') $\mathbf{D}_m = M$, gdzie:

- dowolny nieskończony zbiór z nałożonym porządkiem liniowym $<_M$

- element minimalny P_0 , $\forall p \in \mathbf{D}_m. (P_0 <_M p \vee P_0 = p)$

- Konstruktory typów złożonych

- $(a b)$, gdzie $\mathbf{D}_{(a b)} = \mathbf{D}_b^{\mathbf{D}_a}$

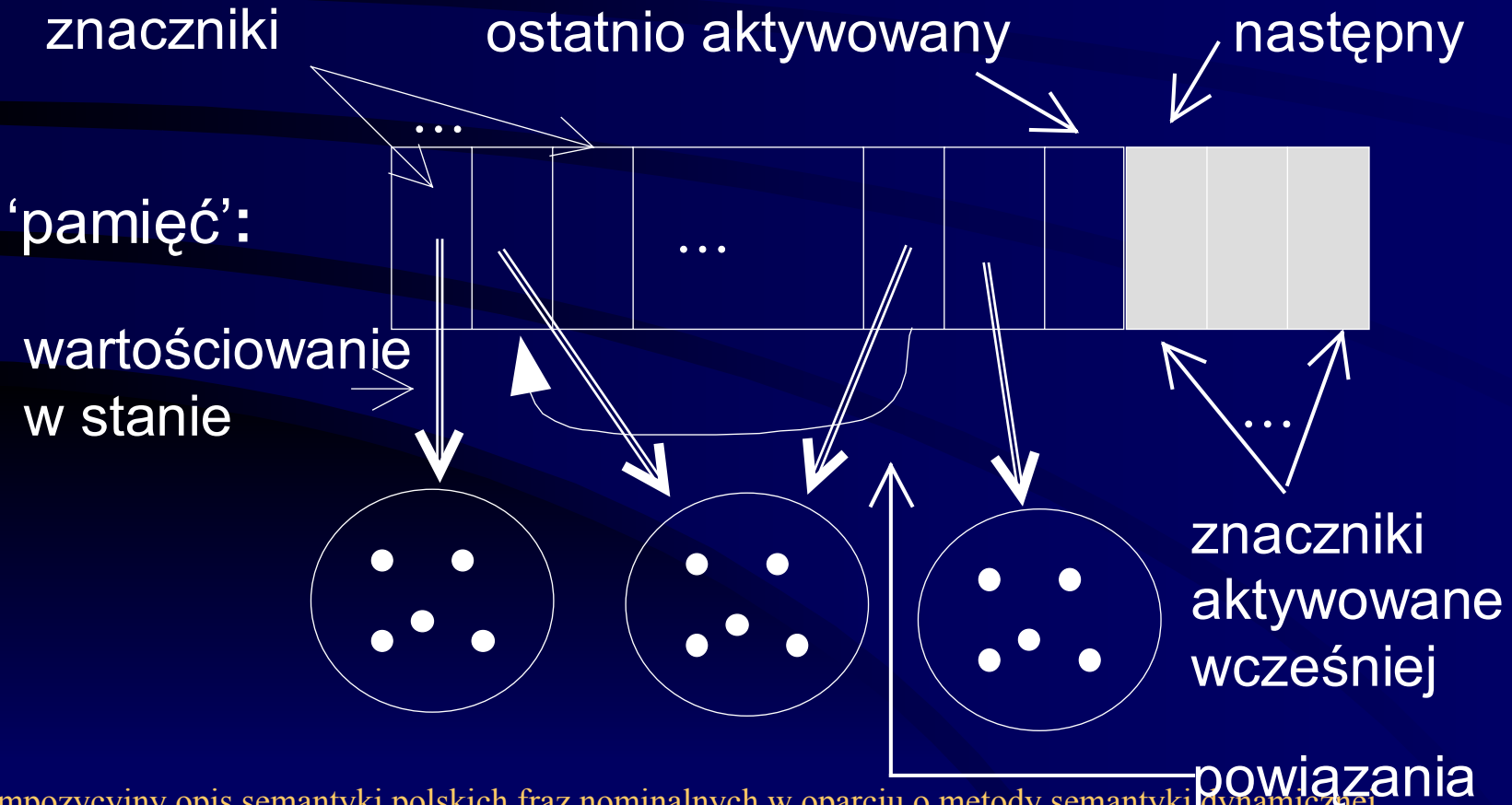
- $(a_1 \bullet a_2 \bullet \dots \bullet a_n)$, gdzie $\mathbf{D}_{(a_1 \bullet a_2 \bullet \dots \bullet a_n)} =$

$$\mathbf{D}_{(a_1 \bullet a_2 \bullet \dots \bullet a_n)}$$

$$\mathbf{D}_{a_1} \times \mathbf{D}_{a_2} \times \dots \times \mathbf{D}_{a_n}$$

SLS - typ stany

- Typ złożony $s := (m \bullet ((m \bullet m) t) \bullet (m(et)))$
 - Stan początkowy $S_0 = \langle P_0, \emptyset, \{ \langle P_0, \emptyset \rangle \} \rangle$



SLS - komponent kontekstu (1)

- *Formuły dynamiczne* - termy typu ($s(st)$)
- Aktywacja znacznika:
 - \downarrow - typu: ($s(st)$)
 - zwiększa zbiór aktywnych znaczników o następnik ostatnio aktywowanego znacznika
 - rozszerza niedeterministycznie funkcję wartościującą w stanie

$$\|\downarrow\| := \lambda i. \lambda j. [(j)^1 = succ((i)^1) \wedge (j)^2 = (i)^2 \\ \forall p. [p \neq succ((i)^1) \rightarrow [(j)^3(p) = (i)^3(p)]]].$$

SLS - komponent kontekstu (2)

- Operator referencji
 - $\hat{\uparrow}$ - typu: (m ((et) (s (st))))
 - odnajduje pośród znaczników dostępnych (warunek strukturalny) te, które są dopasowane pod względem właściwości (warunek semantyczny),
 - zawieranie się pojęciowe wartości w zbiorze podanym jako argument
 - dodaje dla odnalezionych znaczników pary do listy powiązań w danym stanie

$$\|\hat{\uparrow}\| := \lambda p_2. \lambda X_1. \lambda i. \lambda j. [(j)^1 = (i)^1 \wedge (j)^2 = (i)^2$$

$$\cup \{ \langle p_1, p_2 \rangle : P_0 <_m p_1 \wedge (i)^3(p_1) \subseteq X_1 \} \wedge (j)^3 = (i)^3 \wedge \mathbf{WNR}(j)]$$

SLS - komponent kontekstu (3)

- Pseudo-operator

- *dynamicznej: negacji, implikacji i dysjunkcji*

- ‘domykają’ zbiór aktywnych znaczników

$$\lambda T. \mathbf{not} T := \lambda i. \lambda j. (\exists k. T(i,k) \wedge (j)^1 = (i)^1 \wedge (j)^2 = (i)^2 \\ \wedge ((\neg \exists k. [T(i,k) \wedge \mathbf{WKI}(k)] \wedge (j)^3 = (i)^3) \vee \\ (\exists k. [T(i,k) \wedge \mathbf{WKI}(k)] \wedge \forall p. [(p <_m \mathit{succ}((i)^1)) \rightarrow (j)^3(p) = \emptyset]])))$$

$$\lambda T_1. \lambda T_2. (T_1(i,k) \Rightarrow T_2(k,j)) := \lambda i. \lambda j. (\mathbf{not}(T_1(i,k) ; (\mathbf{not} T_2(k,j))))$$

$$\lambda T_1. \lambda T_2. (T_1(i,k) \mathbf{or} T_2(k,j)) :=$$

$$\lambda i. \lambda j. (\mathbf{not}((\mathbf{not} T_1(i,k)) ; (\mathbf{not} T_2(k,j))))$$

- operator *sekwencji*

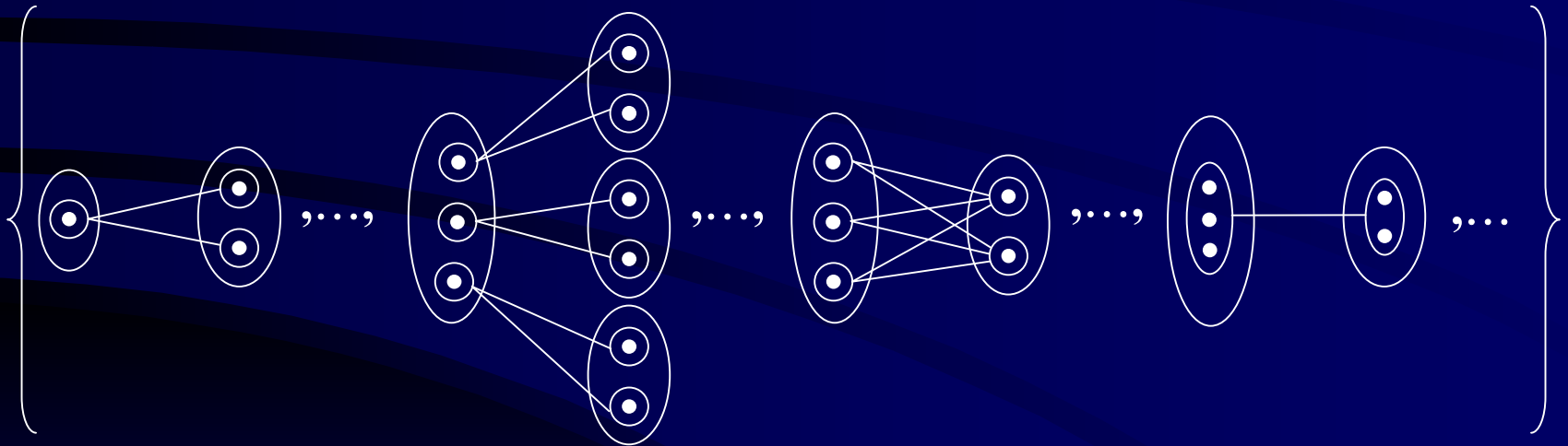
$$T_1(i,k) ; T_2(k,j) := \lambda i. \lambda j. (\exists k. [T_1(i,k) \wedge T_2(k,j)])$$

SLS - komponent kontekstu (4)

- Presupozycja
 - operatory *presupozycji unikalnej* i *słabej* (*referential indefinite*, krypto-nieokreślone)
 - ι , ε - modyfikatory operatora referencji - typu $((m ((et) (s (st)))) (m ((et) (s (st)))))$
 - warunek nałożony na możliwość zmiany - operatory wymagają aby lista powiązań:
 - ι - została rozszerzona o dokładnie jedną parę $\|\iota\| := \lambda R.\lambda p_2.\lambda X_1.\lambda i.\lambda j.[R(p, X_1, i, j) \wedge |(j)^2 \setminus (i)^2| = 1]$
 - ε - została rozszerzona o przynajmniej jedną parę,

SLS – aspekt deskryptywny

- Predykaty
 - ‘rzeczownikowy’ - typu (et)
 - ‘czasownikowy’ - typu $((et)^i t) t$, gdzie $i \in \mathbf{Nat}$



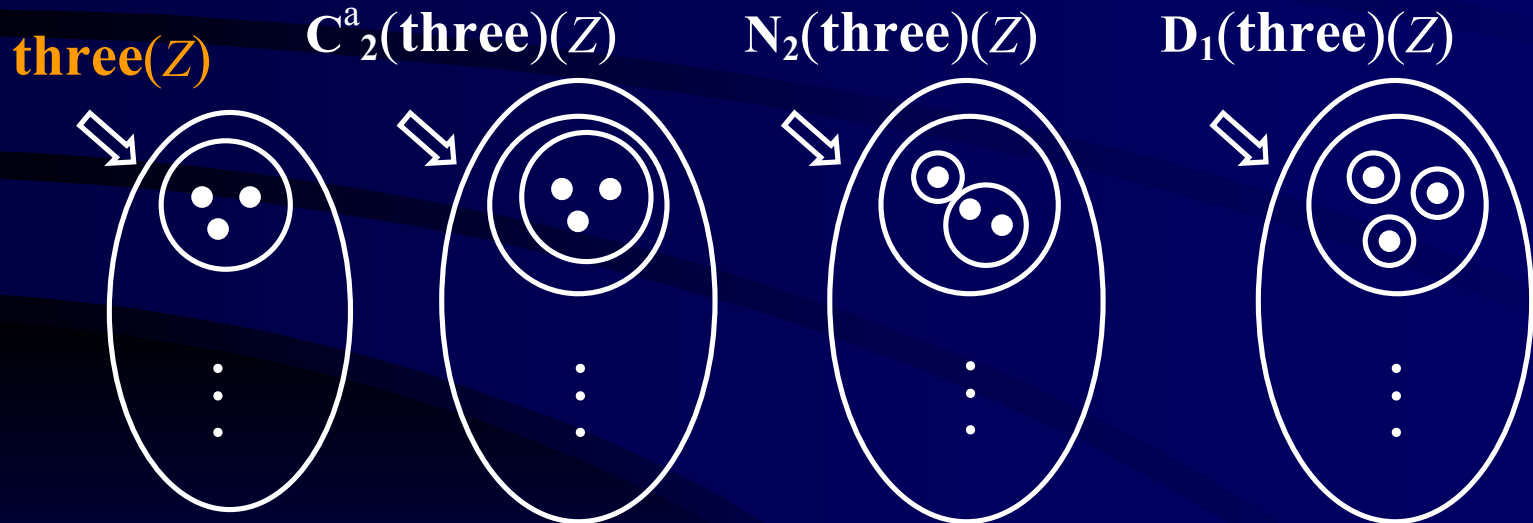
- Pobranie wartości znacznika
 - operator $\#$ - typu $(s (m (et)))$,
 - $\|\#\| := \lambda p. \lambda i. [(i)^3(p)]$ np. $\#(\nabla(j), j) \subseteq \mathbf{paper}$

SLS - kwantyfikacja i mnogość

- Determinatory jako proto-kwantyfikatory

– funktory typu $((et) ((et) t))$

$Z = \#(\nabla(j), j)$,



kolektywna

neutralna

dystrybutywna

three := $\lambda X. \lambda Y. [|X \cap Y| = 2 \wedge Y \subseteq X]$