

Tato prezentace je součástí wiki-prezentace [Metoda GUHA, LISp-Miner a typové úlohy](#)

Je dostupná z [této adresy](#)

Verze 20. 8. 2019

Typ úlohy: zvýšení spokojenosti hostů

Data: [Hotel](#)

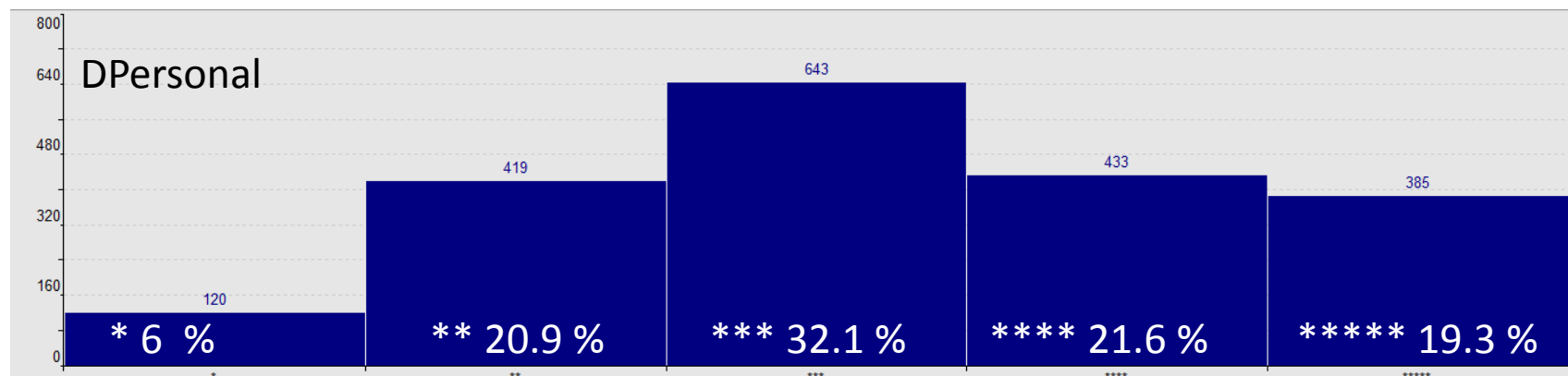
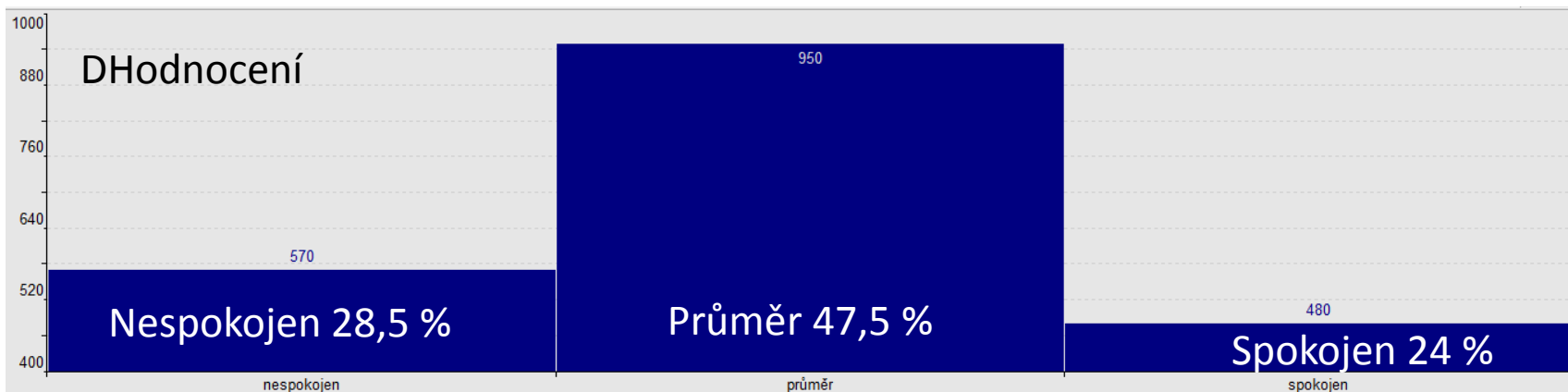
Problém: *Změna čeho a u kterých hostů zvýší spokojenost?*

Jan Rauch

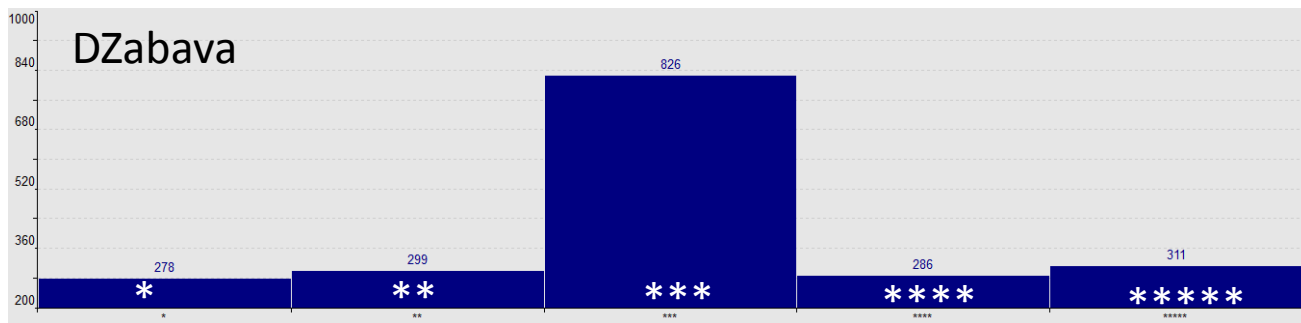
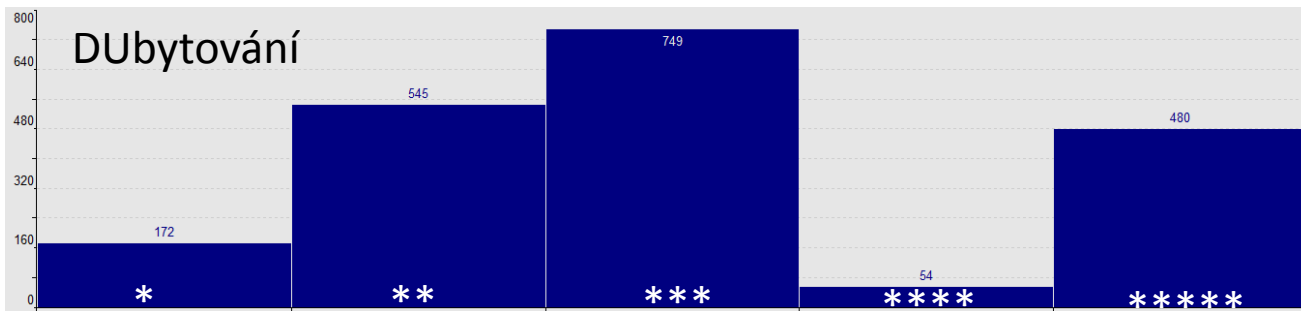
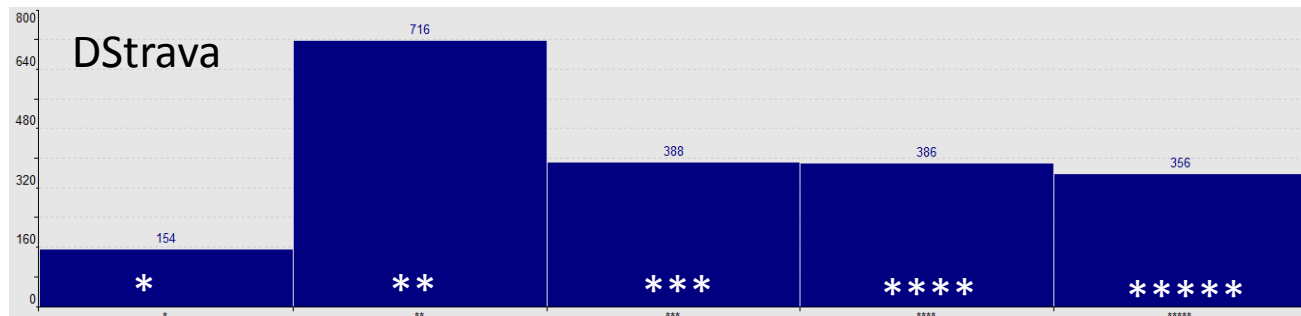
Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

Hotel – atributy „Dotazník“

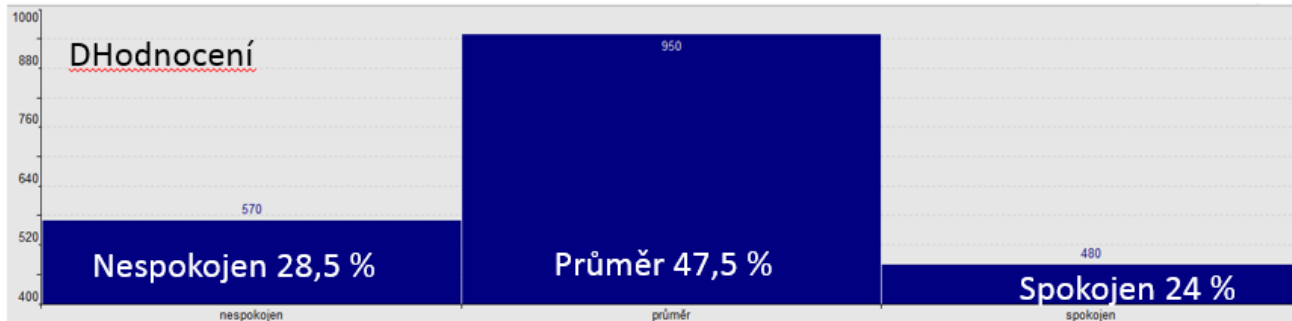


Hotel – atributy „Dotazník“



Akční pravidla

Hotel – možné využití dat „Dotazník“

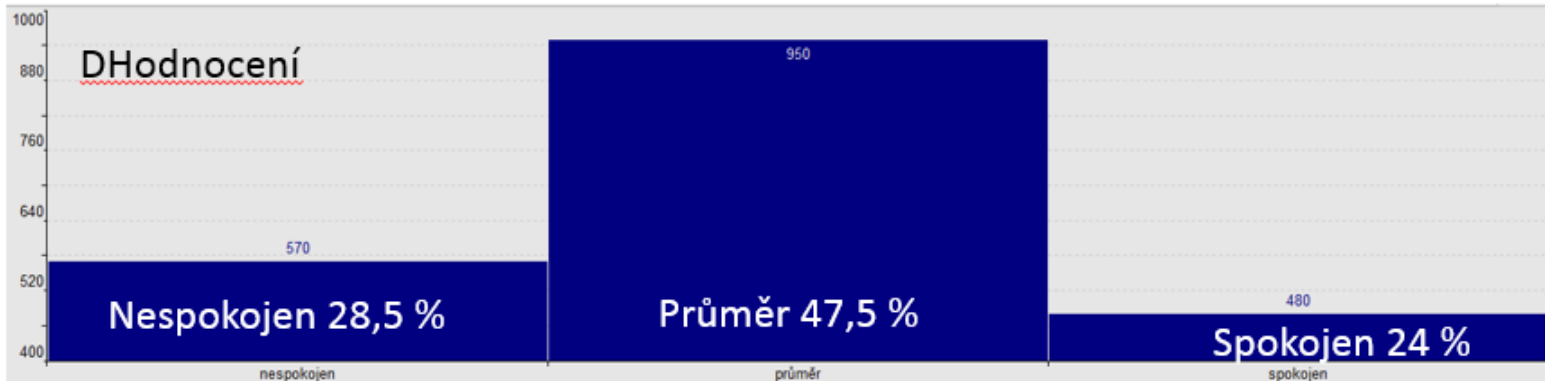


Zvýšit procento spokojených hostů prostřednictvím působením na kvalitu (počet **) u dílčích hodnocení, tedy zvýšení (změna) počtu ** u

- DZábava – např. večery s živou hubou
- DStrava – např. Michelinský kuchař
- DPersonal – např. nové uniformy
- DUbytování – např. širší postele.

Otázka: Změna čeho a u kterých hostů má žádoucí účinek?

Změna čeho a u kterých hostů zvýší spokojenost?



Příklad formalizace:

Hledáme skupiny hostů, dílčí vyhodnocení výchozí a dílčí vyhodnocení finální splňující:

- Jestliže „Skupina \wedge Dílčí vyhodnocení výchozí“, pak maximálně v 30% je DHodnocení(Spokojen)
 - tedy výchozí stav je o málo lepší než průměr
- Jestliže „Skupina \wedge Dílčí vyhodnocení finální“, pak ve výrazně více než ve 30% je DHodnocení(Spokojen)
 - výrazně více: například alespoň v 90% platí DHodnocení(Spokojen)

Rozšíření analytické otázky

Původní otázka:

Změna čeho a u kterých hostů má žádoucí účinek?

Formální zápis: $S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen})$

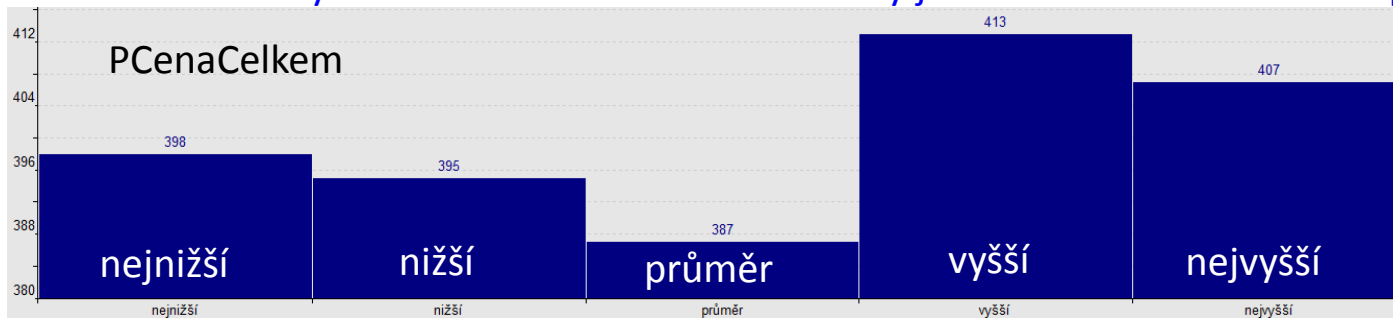
Rozšířená otázka:

Změna čeho a u kterých hostů má žádoucí účinek? Jaký další účinek ta změna má?

Příklad:

PCenaCelkem

Změna čeho a u kterých hostů má žádoucí účinek? Jaký je účinek na celkovou cenu pobytu?



Formální zápis: $S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen}) \wedge [PCenaCelkem(?) \rightarrow PCenaCelkem(?)]$

Ac4ft-kvantifikátor – jiná možnost

Čtyřpolní tabulka pro stav před změnou

	DH(Spokojen) \wedge Cena1	\neg (DH(Spokojen) \wedge Cena1)
S \wedge HI	a_I	b_I
\neg (S \wedge HI)	c_I	d_I

Čtyřpolní tabulka pro stav po změně

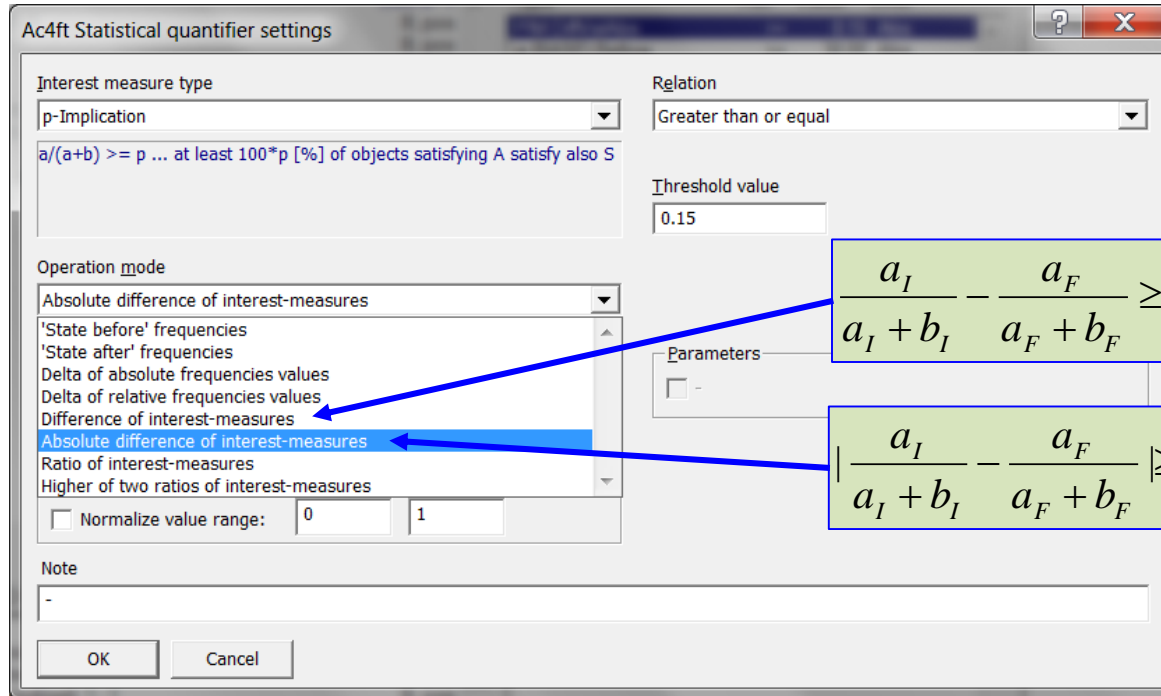
	DH(Spokojen) \wedge Cena2	\neg (DH(Spokojen) \wedge Cena2)
S \wedge HF	a_F	b_F
\neg (S \wedge HF)	c_F	d_F

Ac4ft-kvantifikátor může být dán i složitější podmínkou týkající se obou tabulek $\langle a_I, b_I, c_I, d_I \rangle$ a $\langle a_F, b_F, c_F, d_F \rangle$.

Příklad: $\Rightarrow_{0.15,30,30}^{F>I}$ je definován jako $\frac{a_F}{a_F + b_F} - \frac{a_I}{a_I + b_I} \geq 0.15 \wedge a_I \geq 30 \wedge a_F \geq 30$.

Ac4ft-kvantifikátor – jiná možnost, poznámka

Ac4ft-Miner má mimo jiné tyto volby pro Ac4ft-kvantifikátor:



Použijeme $\left| \frac{a_I}{a_I + b_I} - \frac{a_F}{a_F + b_F} \right| \ge 0.15$

a použijeme vztahy pro které je

$$\frac{a_I}{a_I + b_I} - \frac{a_F}{a_F + b_F} \leq -0.15$$

$S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen}) \wedge [PCenaCelkem(?) \rightarrow PCenaCelkem(?)]$

Zadání pro Ac4ft-Miner

$$\left| \frac{a_I}{a_I + b_I} - \frac{a_F}{a_F + b_F} \right| \geq 0.15 \wedge a_I \geq 30 \wedge a_F \geq 30$$

ANTECEDENT STABLE PART		QUANTIFIERS		SUCCEDENT STABLE PART	
Host/Bydliště	Con. 0 - 1	Type	Rel. Value Units	Default Partial Cedent	Con. 1 - 1
» HCizinec_b (subset), 1 - 1	B. pos	PIM DiffValAbs	>= 0.15 Abs	» DHodnoceni(spokojen)	B. pos
» HMesto (subset), 1 - 1	B. pos	a (BASE) Before	>= 30.00 Abs		
» HStat (subset), 1 - 1	B. pos	a (BASE) After	>= 30.00 Abs		
Meteo	Con. 0 - 2				
» MObloha (subset), 1 - 1	B. pos				
» MTeploata_exp (seq), 1 - 2	B. pos				
Pobyt	Con. 0 - 2				
» PTypPobytu (subset), 1 - 1	B. pos				
» PTurnus (subset), 1 - 1	B. pos				
		Generation information			
		Status: Not generated, 14 run(s)			
		Mode: -			
Total length:	0 - 5			Total length:	1
(1) ANTECEDENT VARIABLE PART		CONDITION		(2) SUCCEDENT VARIABLE PART	
Relevantní HI a FI	Con. 1 - 4	Default Partial Cedent	Con. 0 - 5	Default Partial Cedent	Con. 1 - 1
» DPersonal_edc5_m (subset), 1 - 1	B. pos			» PCenaCelkem_ef5 (subset), 1 - 1	B. pos
» DStrava_edc5_m (subset), 1 - 1	B. pos				
» DUbytovani_edc5_m (subset), 1 - 1	B. pos				
» DZabava_edc5_m (subset), 1 - 1	B. pos				

zadání relevantních S

DH(Spokojen)

zadání relevantních HI a FI

PCenaCelkem(?) → PCenaCelkem(?)

$S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen}) \wedge [PCenaCelkem(?) \rightarrow PCenaCelkem(?)]$

Výstup

Task run

Start: 25.3.2016 22:05:46 Total time: 0h 2m 44s

Number of verifications: 3204560

Number of hypotheses: 8 Mode: Standard

Add group Dej group Edit group

Actual group of hypotheses: All hypotheses

Hypotheses in group: 8 Shown hypotheses: 8 Highlighted: 0

Nr.	Id	Df:Conf	B:Conf	A:Conf	Hypothesis
1	5	0.173	0.258	0.085	(empty): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nizsi) -> PCenaCelkem(vysi)) ←
2	6	0.165	0.258	0.093	(empty): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nizsi) -> PCenaCelkem(nejvysi)) ←
3	8	0.157	0.278	0.121	MTeplota(zima, neutrální): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nizsi) -> PCenaCelkem(vysi)) ←
4	4	0.155	0.229	0.074	(empty): (DPersonal(*****) -> DPersonal(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nejvysi) -> PCenaCelkem(nizsi)) ←
5	1	-0.155	0.074	0.229	(empty): (DPersonal(*****) -> DPersonal(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nizsi) -> PCenaCelkem(nejvysi)) ←
6	7	-0.157	0.121	0.278	MTeplota(zima, neutrální): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(vysi) -> PCenaCelkem(nizsi)) ←
7	3	-0.165	0.093	0.258	(empty): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(nejvysi) -> PCenaCelkem(nizsi)) ←
8	2	-0.173	0.085	0.258	(empty): (DStrava(*****) -> DStrava(*****)) ><< DHodnoceni(spokojen): (PCenaCelkem(vysi) -> PCenaCelkem(nizsi)) ←

$$\frac{a_I}{a_I + b_I} - \frac{a_F}{a_F + b_F} \leq -0.15$$

- cena vzrostla ←
- cena klesla ←
- nutná pečlivá interpretace (generovaná data)

$S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen}) \wedge [PCenaCelkem(?) \rightarrow PCenaCelkem(?)]$
 Výstup – detail

Antecedent:	$(DPersonal(****) \rightarrow DPersonal(*****))$
Succedent:	$DHodnoceni(spokojen) : (PCenaCelkem(nižší) \rightarrow PCenaCelkem(nejvyšší))$
State before:	$DPersonal(****) > \div < DHodnoceni(spokojen) \ \&\& \ PCenaCelkem(nižší)$
State after:	$DPersonal(*****) > \div < DHodnoceni(spokojen) \ \&\& \ PCenaCelkem(nejvyšší)$

$S \wedge HI \Rightarrow_{pI, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(?)$

$DPersonal(****) \Rightarrow_{pI, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(\text{nižší})$

$S \wedge HF \Rightarrow_{pF, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(?)$

$DPersonal(*****) \Rightarrow_{pF, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(\text{nejvyšší})$

$S \wedge [HI \rightarrow HF] \approx_{I, F} DH(\text{Spokojen}) \wedge [PCenaCelkem(?) \rightarrow PCenaCelkem(?)]$
 Výstup – detail

$S \wedge HI \Rightarrow_{pI, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(?)$

$DPersonal(****) \Rightarrow_{pI, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(\text{nižší})$

$S \wedge HF \Rightarrow_{pF, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(?)$

$DPersonal(*****) \Rightarrow_{pF, 30} DH(\text{Spokojen}) \wedge PCenaCelkem(\text{nejvyšší})$

$$\frac{32}{32 + 401} = 0.07$$

	Succedent		¬Succedent	
Antecedent	32	401	88	297
¬Antecedent	79	1488	17	1598

$$\frac{88}{88 + 297} = 0.23$$