

# *Trust and Transitivity*

*Rino Falcone and Cristiano Castelfranchi*

*T<sup>3</sup> - Trust, Theory and Technology Group ([www.istc.cnr.it/T3](http://www.istc.cnr.it/T3))*

*Goal - Goal Oriented Agents Laboratory*

**Institute of Cognitive Sciences and Technologies**

**National Research Council of Italy**

**Rome- Italy**

***WOA 2010***



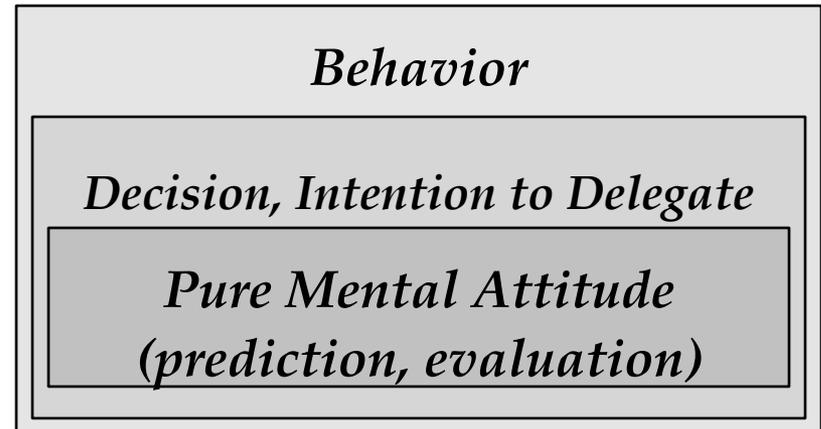
- Socio Cognitive Model of Trust
- Transitivity in Trust
- Transitivity and delegated subtasks
- Conclusions

- X trusts Y for a task  $\tau$  on the basis of a series of subjective evaluations of Y's "virtues/qualities", i.e.:
  - ◆ Abilities
  - ◆ Willingness
  - ◆ External Factors
- Not just on the basis of a statistical sampling
- Evaluations are the cognitive mediators for the decision to trust Y or not

# Socio-Cognitive Model of Trust

*X: Trustor*

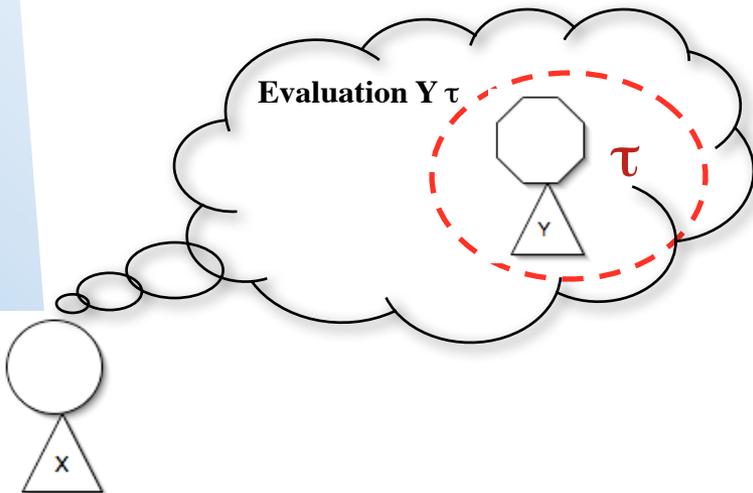
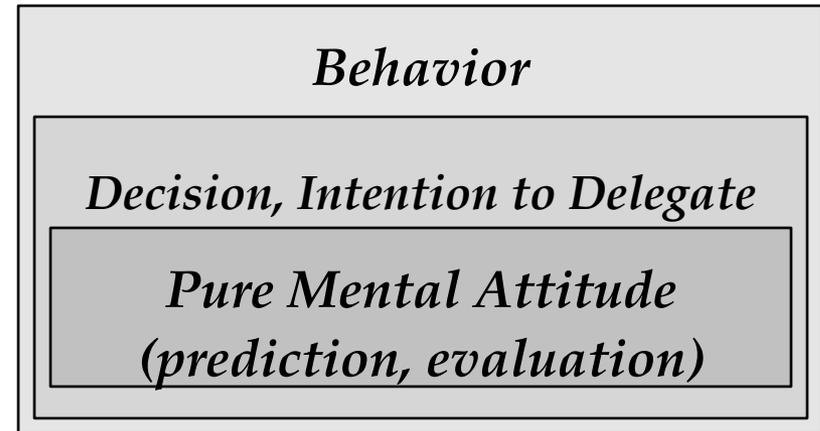
*Y: Trustee*



# Socio-Cognitive Model of Trust

*X: Trustor*

*Y: Trustee*

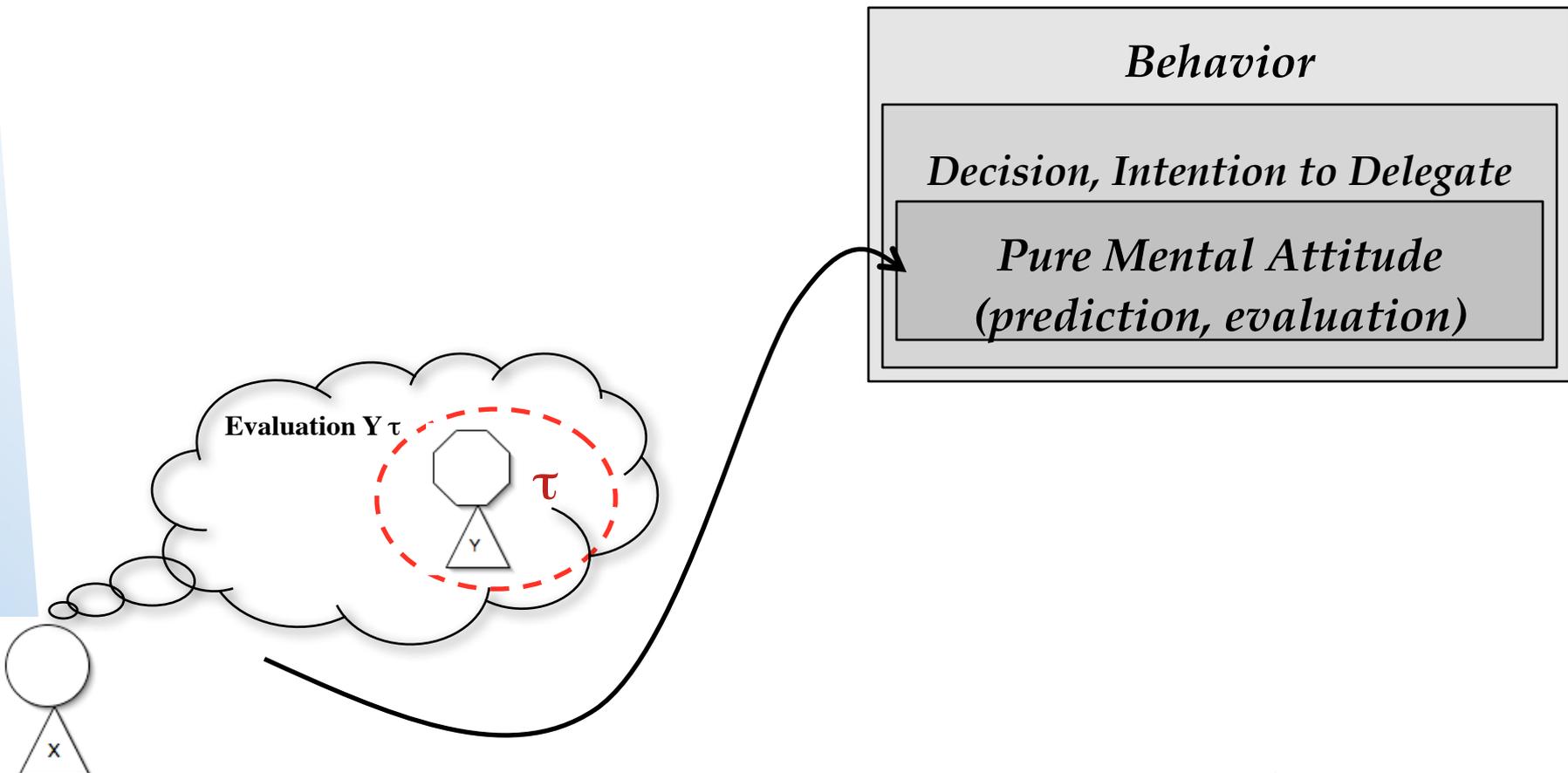


# Socio-Cognitive Model of Trust

*X: Trustor*

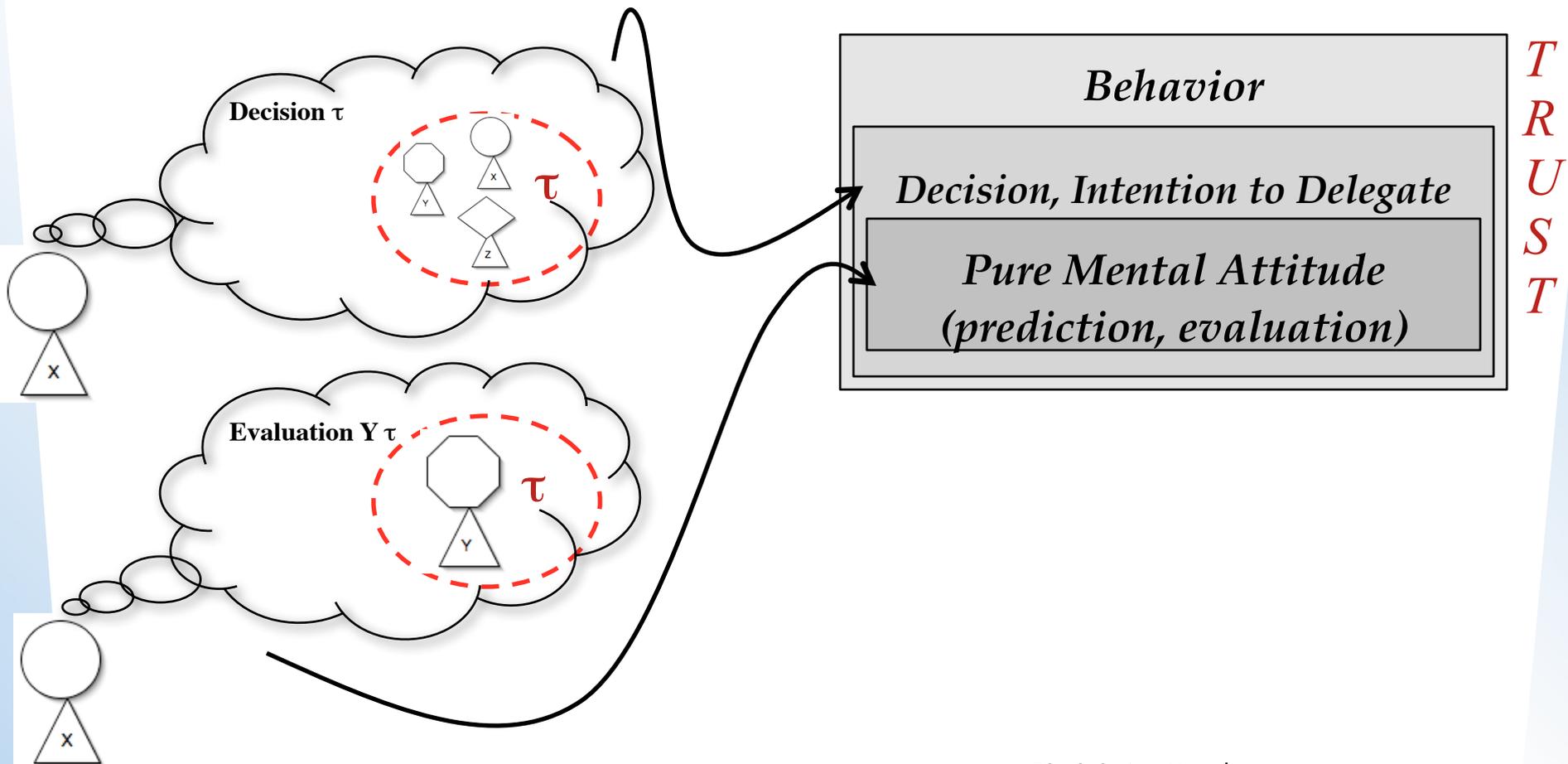
*Y: Trustee*

*T  
R  
U  
S  
T*

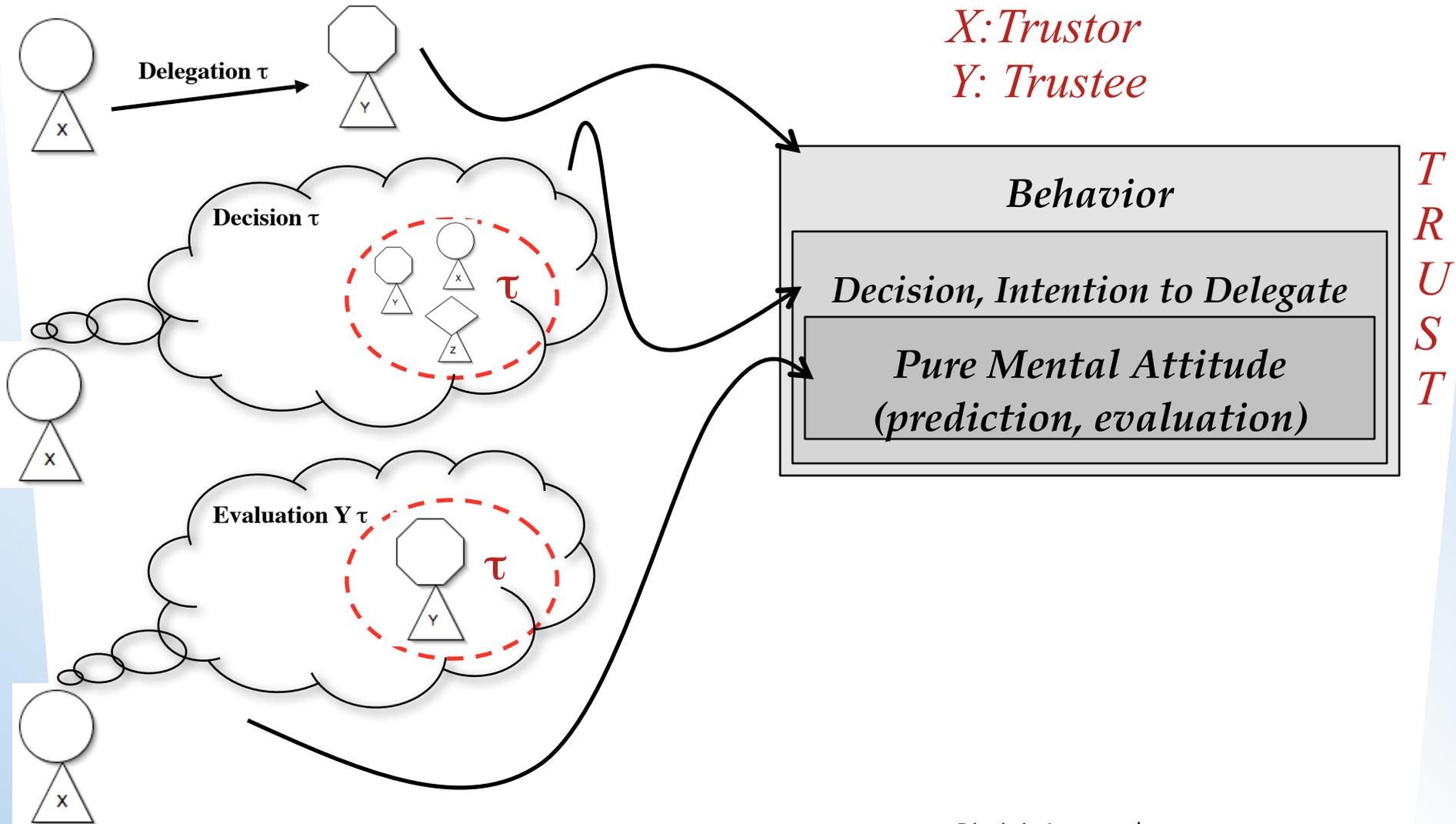


# Socio-Cognitive Model of Trust

*X: Trustor*  
*Y: Trustee*

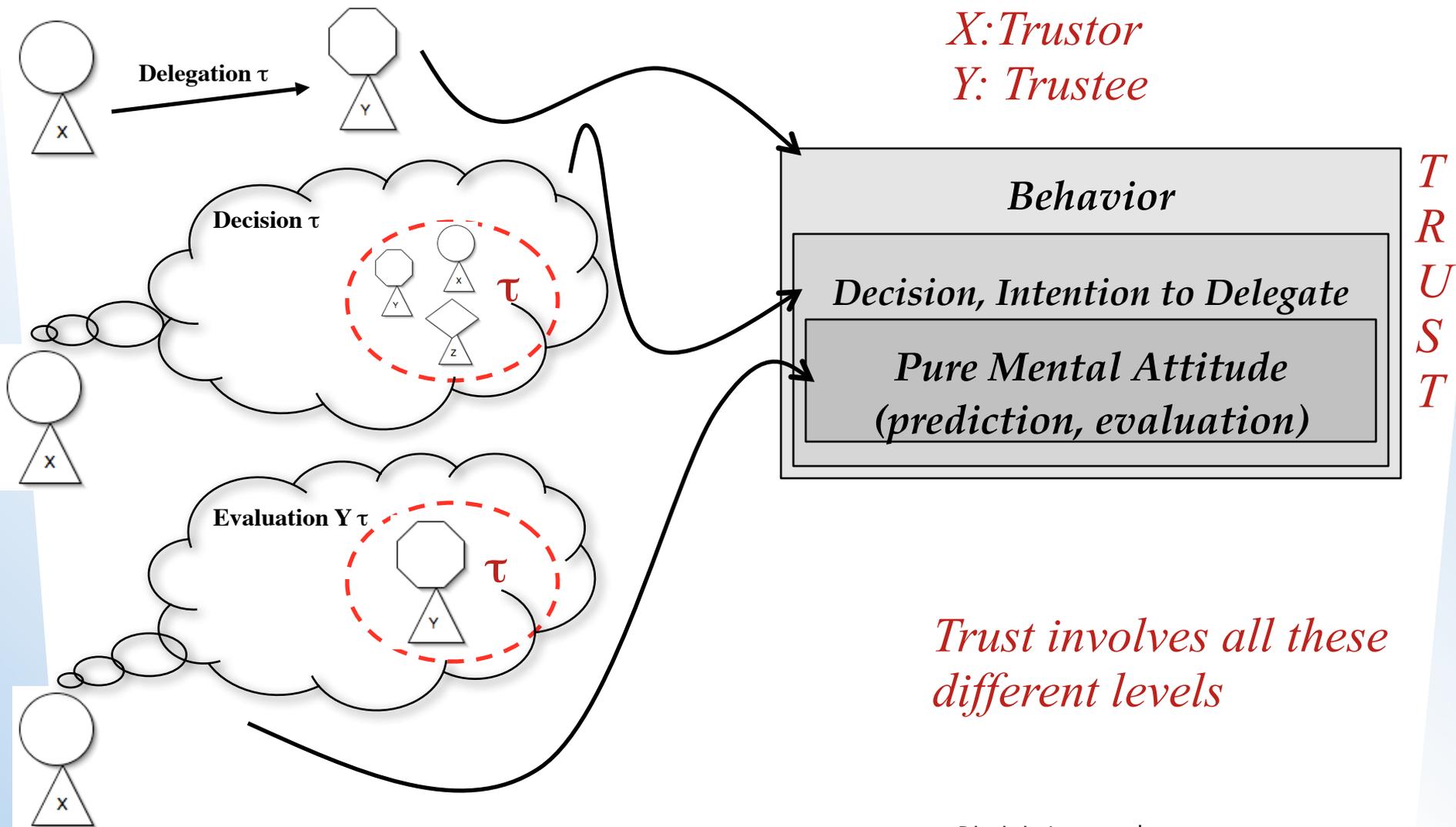


# Socio-Cognitive Model of Trust



*X: Trustor*  
*Y: Trustee*

# Socio-Cognitive Model of Trust



*Trust involves all these different levels*

- X trusts Y for the task  $\tau$  when X believes that Y is potentially :
  - ◆ able
  - ◆ willing
- to fulfill  $\tau$ , and X believes that the
  - ◆ external conditions
- are not impeding to the task realization

# Trust

*Trust* ( $X Y \tau C g_X$ )

## *Trust* (X Y $\tau$ C $g_X$ )

*X's mental ingredients of trust*

*Bel* (X *Can*<sub>Y</sub>( $\tau$ ))

*Bel* (X *Will*<sub>Y</sub>( $\tau$ ))

*Bel* (X *ExtFact*<sub>Y</sub>( $\tau$ ))  $g_X$

## *Trust* ( $X$ $Y$ $\tau$ $C$ $g_X$ )

*X*'s mental ingredients of trust

$Bel(X Can_Y(\tau))$	$Bel(X Will_Y(\tau))$	$Bel(X ExtFact_Y(\tau))$	$g_X$
----------------------	-----------------------	--------------------------	-------



trust quantification

$DoT_{XY\tau} = f( DoC_X (Ability_Y(\tau)), DoC_X (Willingness_Y(\tau)), DoC_X (Opp_Y(\tau)) )$	$Value(g_X)$
---	--------------

## *Trust* ( $X Y \tau C g_X$ )

*X's mental ingredients of trust*

$Bel (X Can_Y(\tau))$	$Bel (X Will_Y(\tau))$	$Bel (X ExtFact_Y(\tau))$	$g_X$
-----------------------	------------------------	---------------------------	-------

*trust quantification*

$DoT_{XY\tau} = f( DoC_X (Ability_Y(\tau)), DoC_X (Willingness_Y(\tau)), DoC_X (Opp_Y(\tau))$	$Value(g_X)$
---	--------------

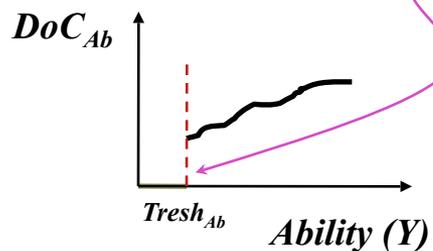
## $Trust (X Y \tau C g_X)$

*X's mental ingredients of trust*

$Bel (X Can_Y(\tau)) \quad Bel (X Will_Y(\tau)) \quad Bel (X ExtFact_Y(\tau)) \quad g_X$

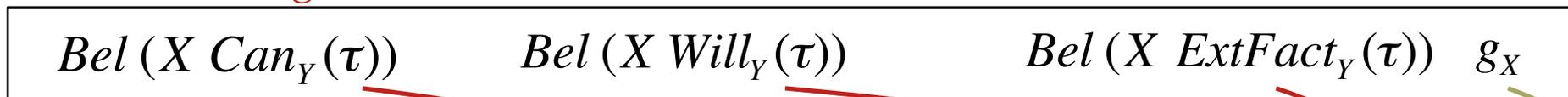
*trust quantification*

$DoT_{XY\tau} = f( DoC_X (Ability_Y(\tau)), DoC_X (Willingness_Y(\tau)), DoC_X (Opp_Y(\tau)) \quad Value(g_X)$

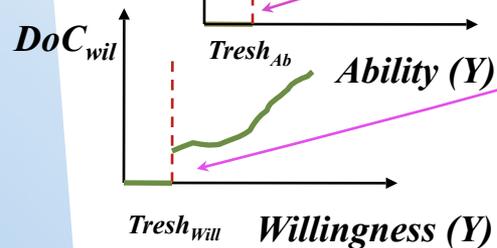
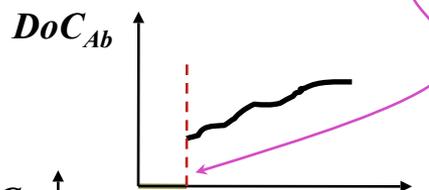


## Trust (X Y τ C g<sub>X</sub>)

X's mental ingredients of trust



trust quantification



# Trust

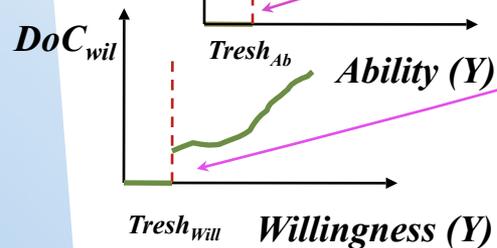
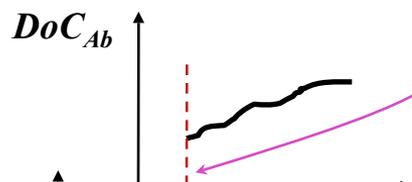
$$\text{Trust} (X Y \tau C g_X)$$

*X's mental ingredients of trust*

$$\text{Bel} (X \text{ Can}_Y(\tau)) \quad \text{Bel} (X \text{ Will}_Y(\tau)) \quad \text{Bel} (X \text{ ExtFact}_Y(\tau)) \quad g_X$$

*trust quantification*

$$\text{DoT}_{XY\tau} = f( \text{DoC}_X (\text{Ability}_Y(\tau)), \text{DoC}_X (\text{Willingness}_Y(\tau)), \text{DoC}_X (\text{Opp}_Y(\tau)) \quad \text{Value}(g_X)$$



$$\text{DoT}_{XY\tau} > \sigma$$

# Trust

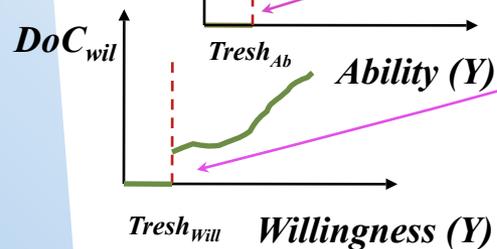
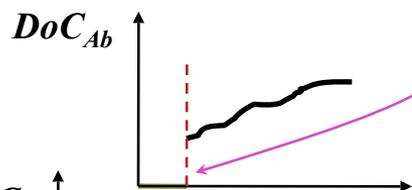
$$Trust (X Y \tau C g_X)$$

*X's mental ingredients of trust*

$$Bel (X Can_Y(\tau)) \quad Bel (X Will_Y(\tau)) \quad Bel (X ExtFact_Y(\tau)) \quad g_X$$

*trust quantification*

$$DoT_{XY\tau} = f( DoC_X (Ability_Y(\tau)), DoC_X (Willingness_Y(\tau)), DoC_X (Opp_Y(\tau)) \quad Value(g_X)$$



$$DoT_{XY\tau} > \sigma$$

# Socio-Cognitive model of Trust

Internal Trust

Competence

Willingness

Unharmfulness

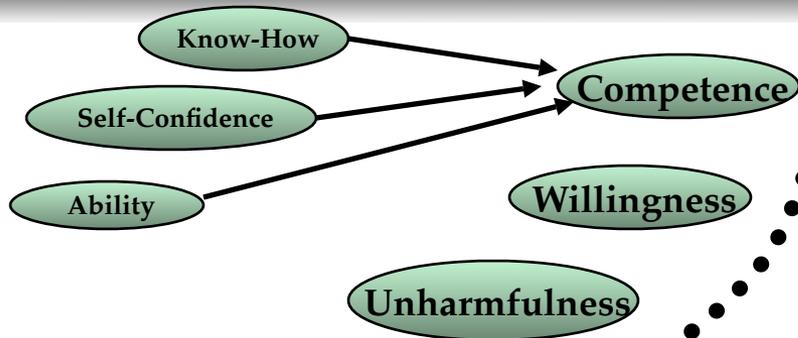
Opportunities

Obstacles

External Trust

# Socio-Cognitive model of Trust

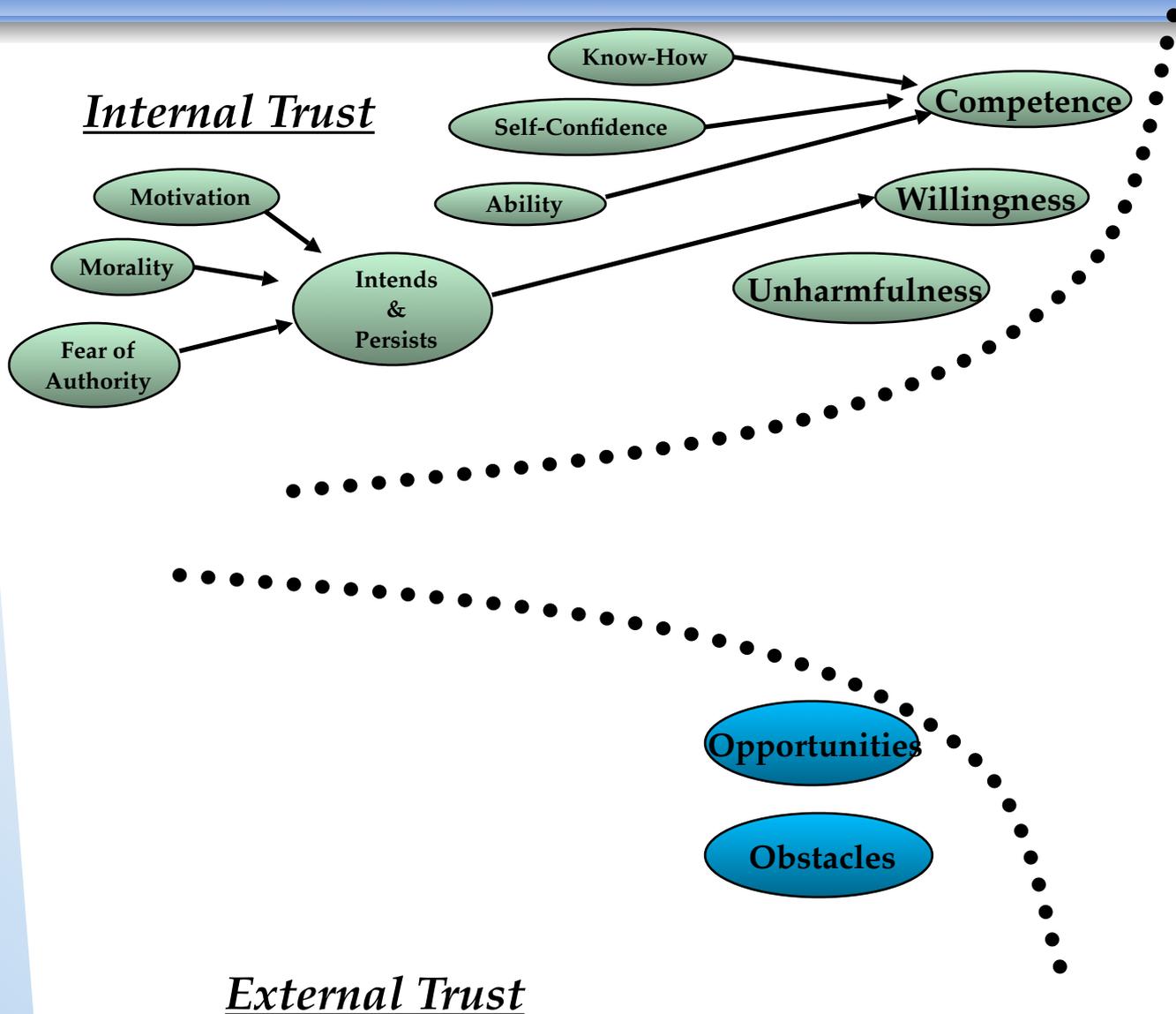
Internal Trust



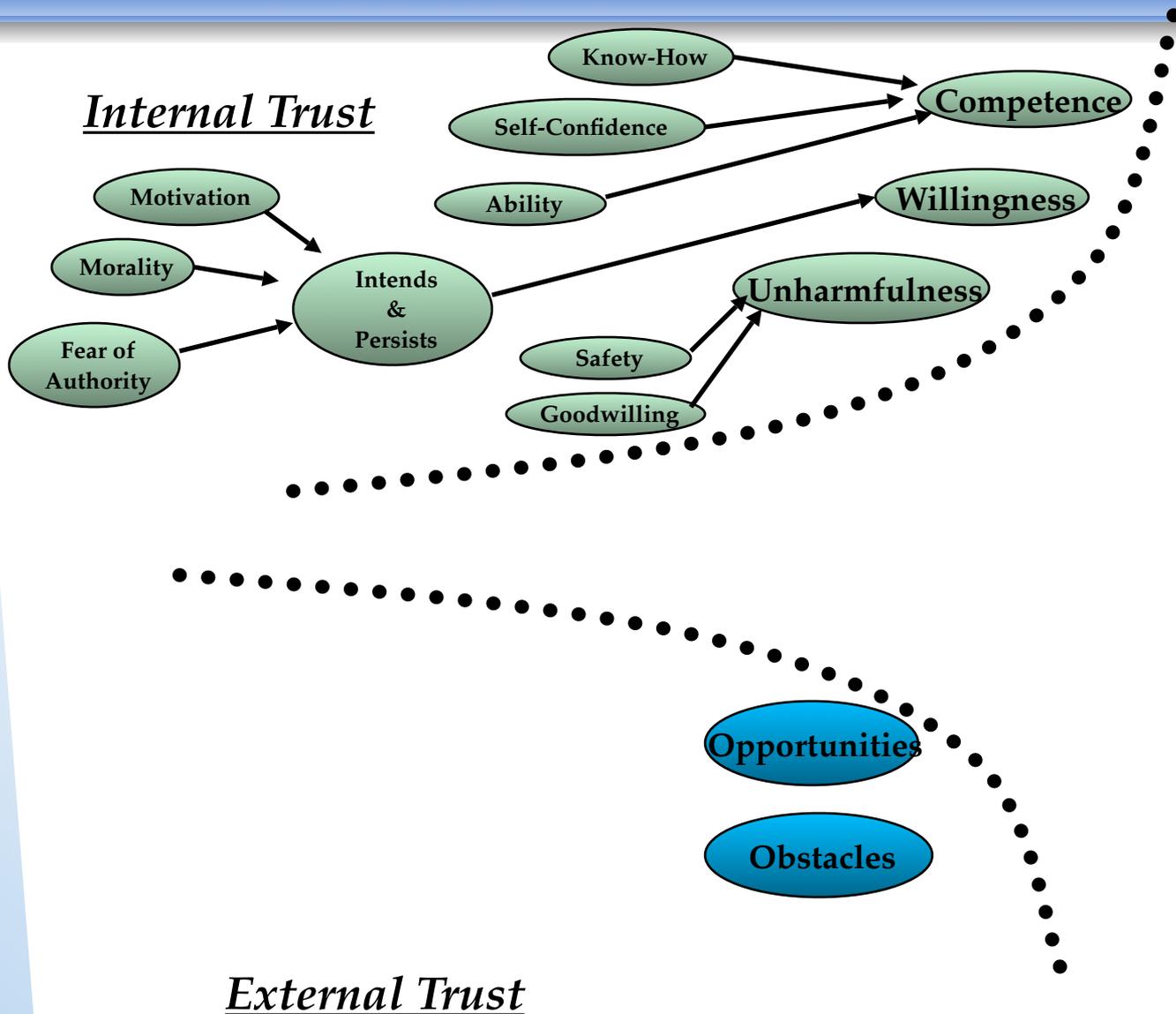
External Trust



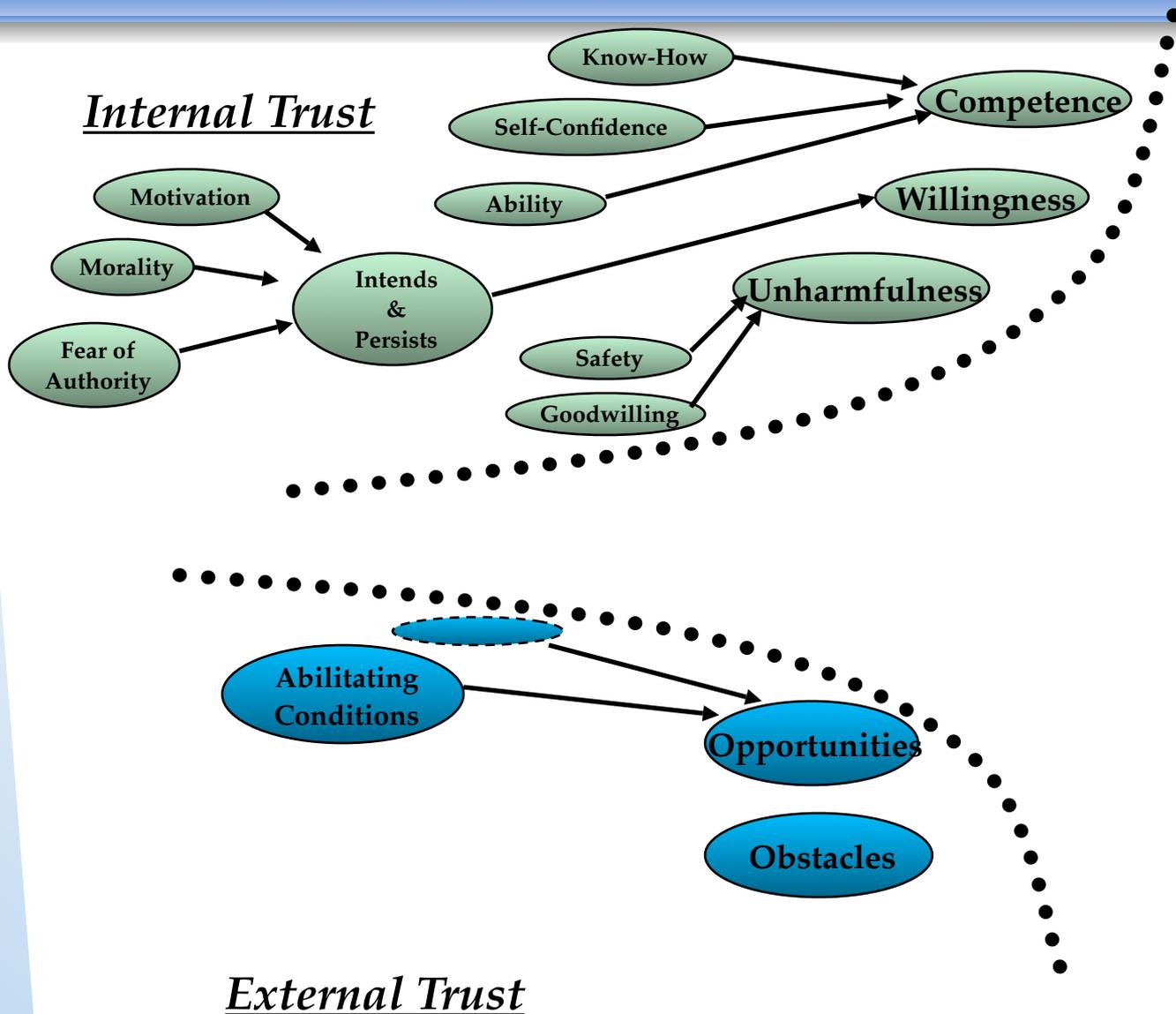
# Socio-Cognitive model of Trust



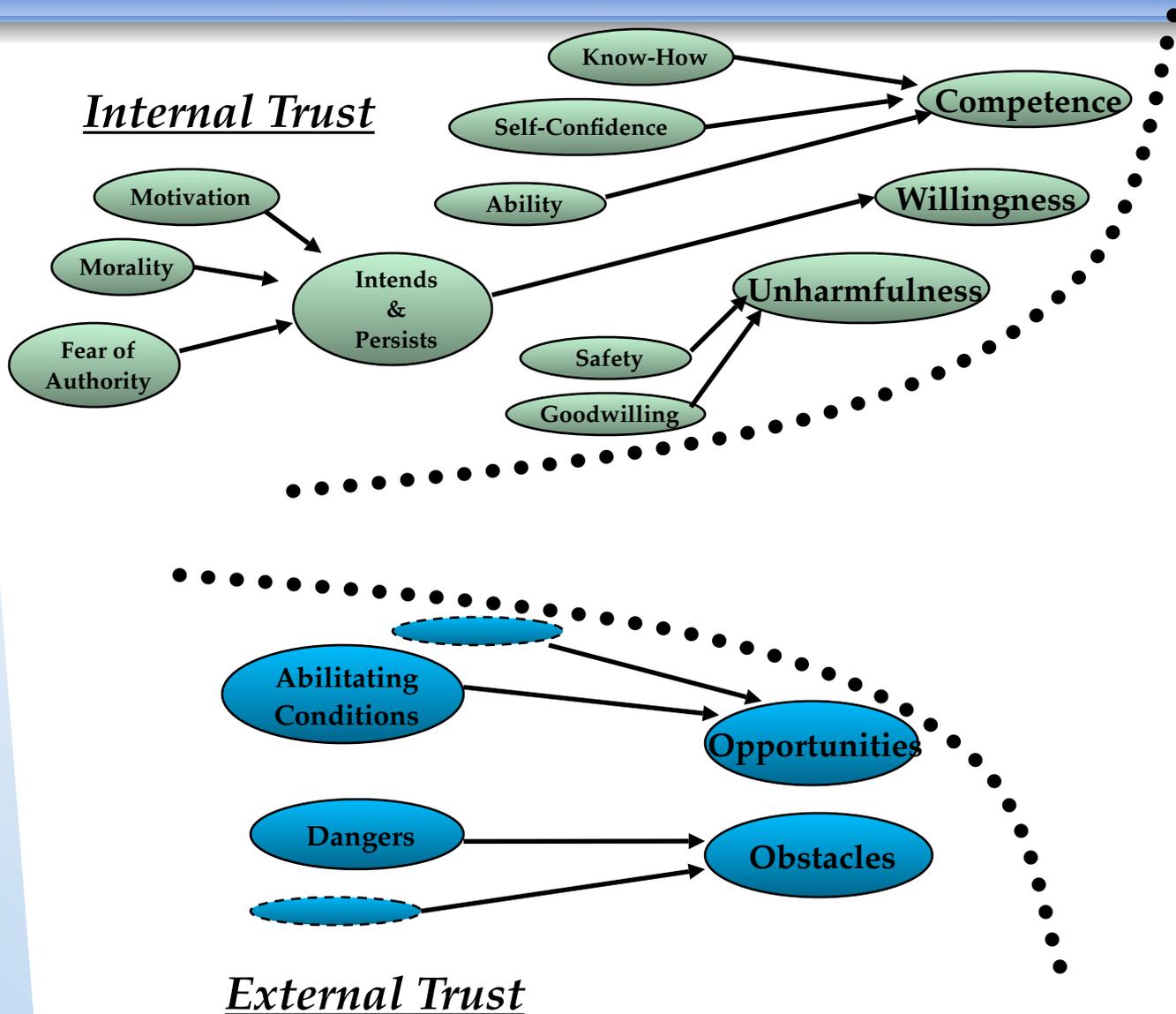
# Socio-Cognitive model of Trust



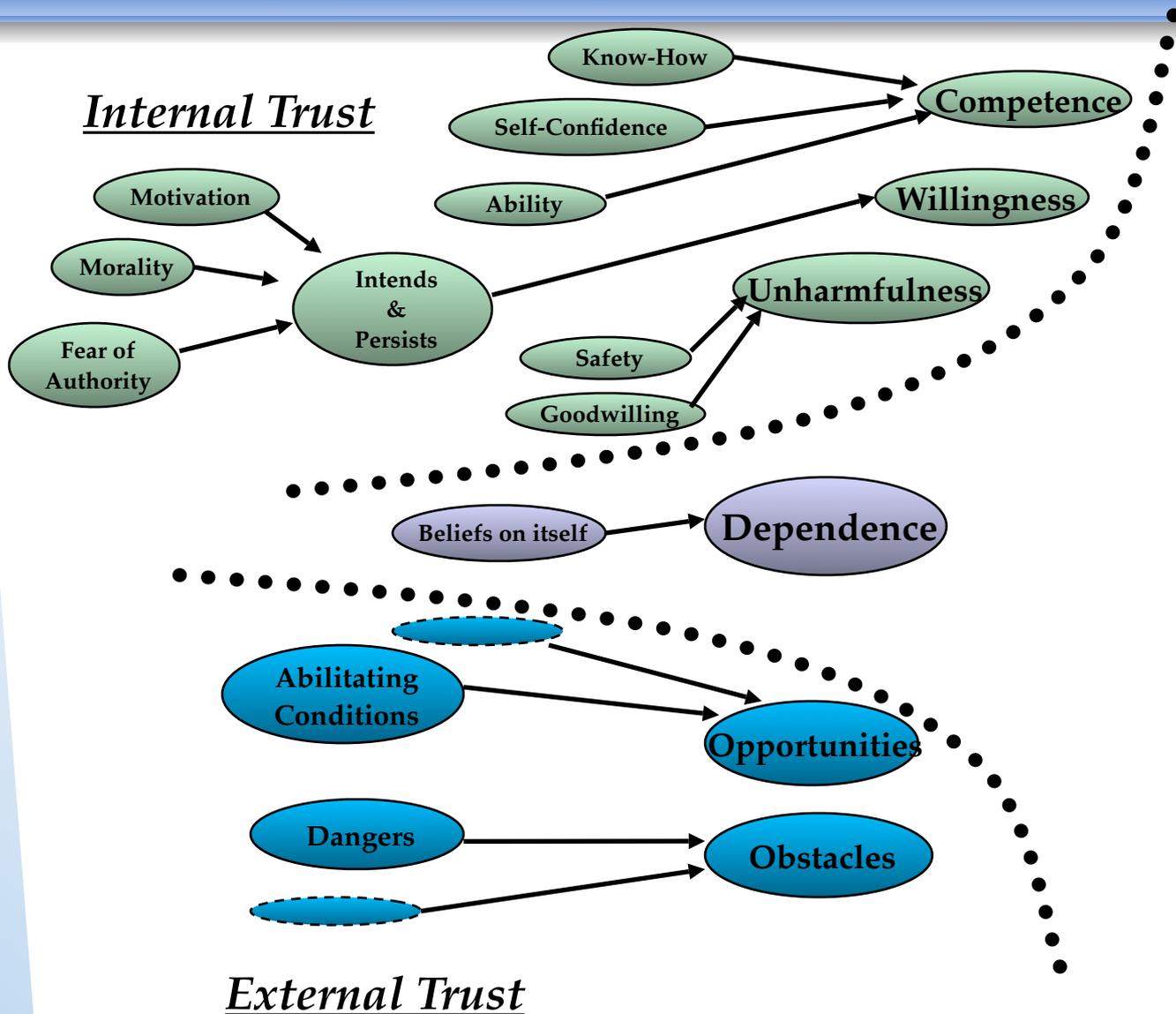
# Socio-Cognitive model of Trust



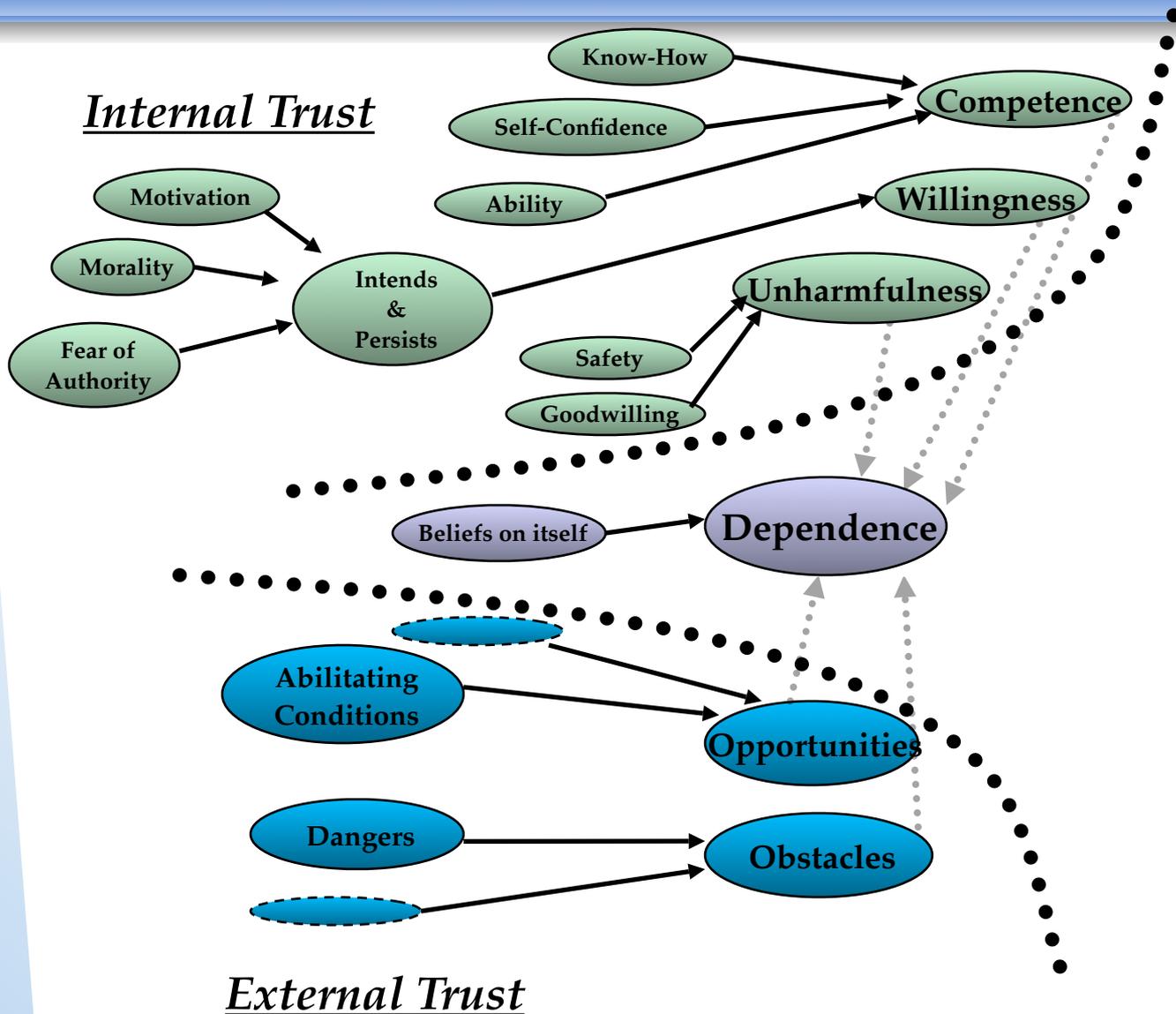
# Socio-Cognitive model of Trust



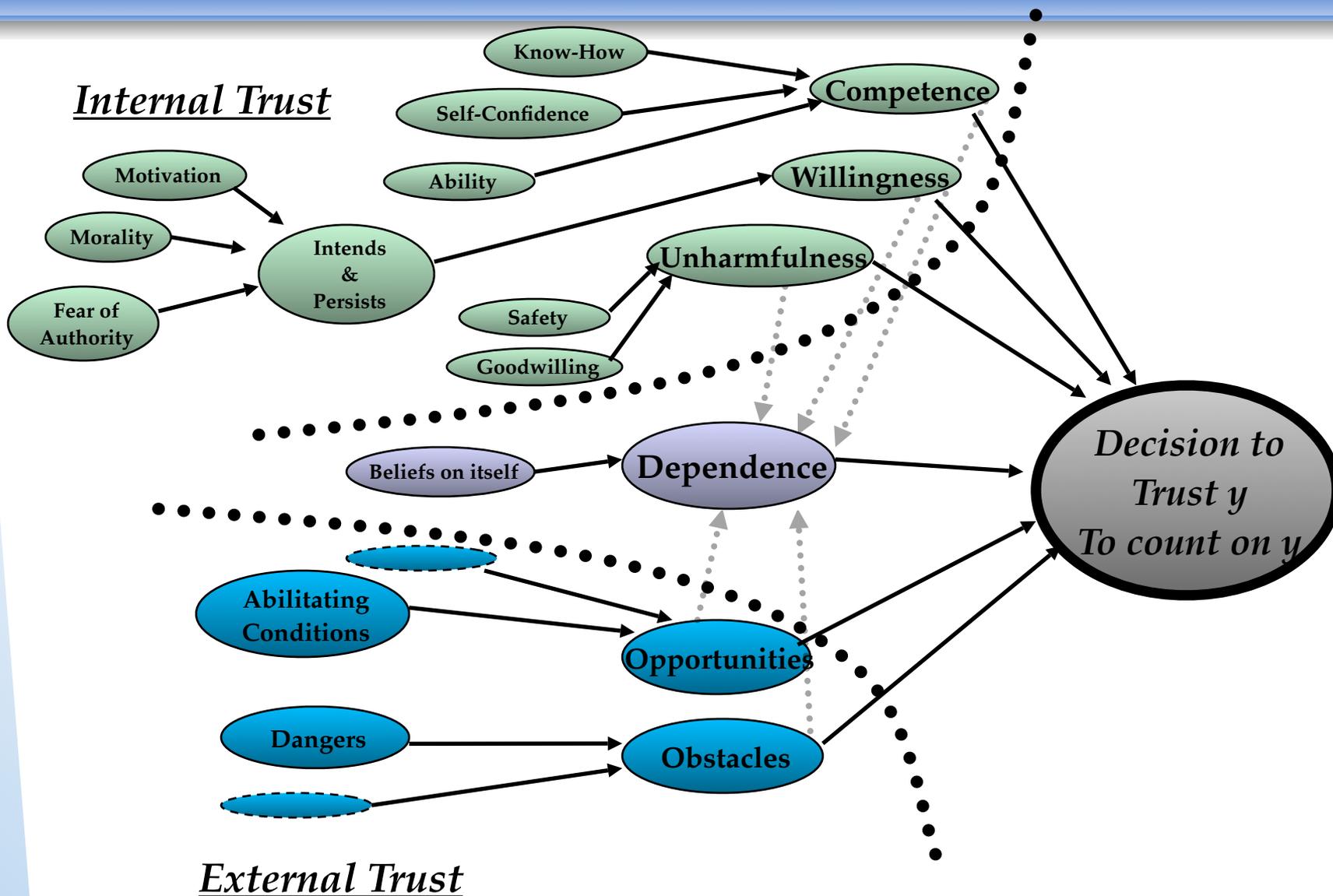
# Socio-Cognitive model of Trust



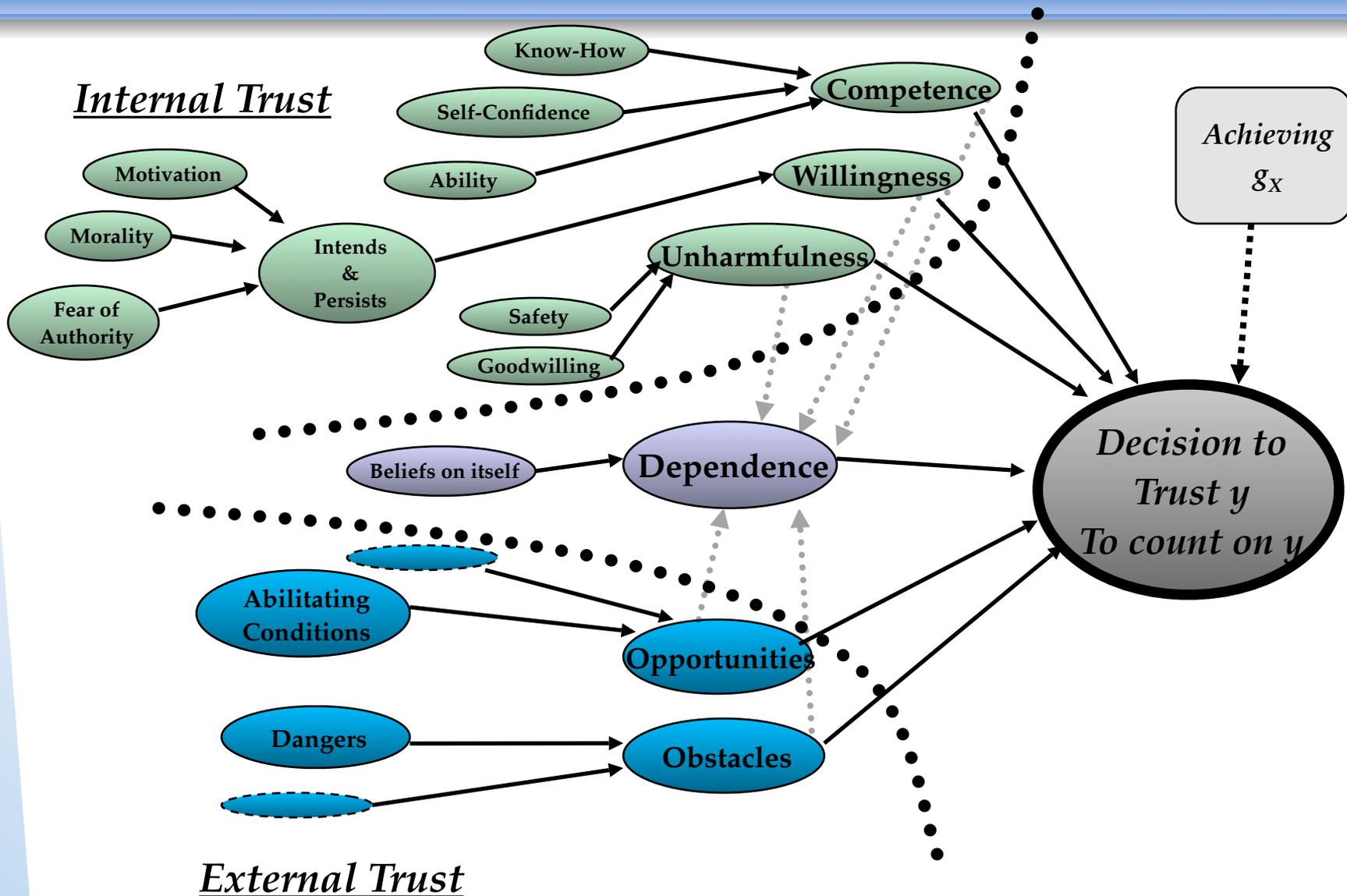
# Socio-Cognitive model of Trust



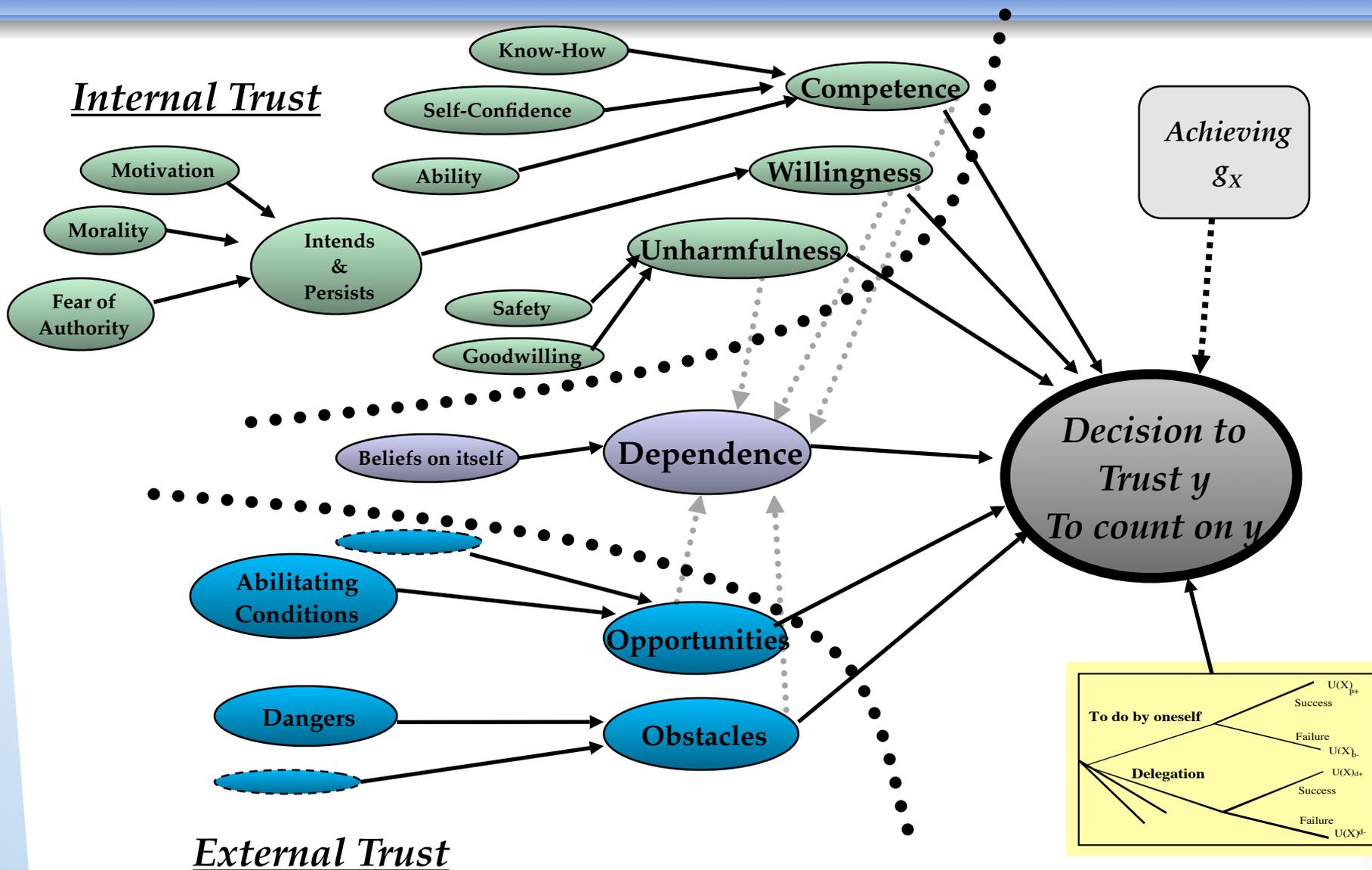
# Socio-Cognitive model of Trust



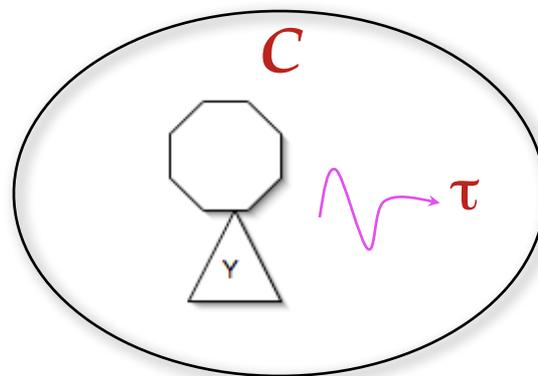
# Socio-Cognitive model of Trust



# Socio-Cognitive model of Trust

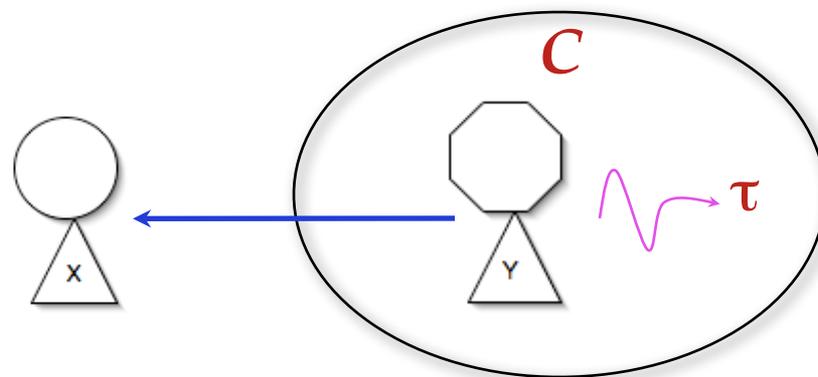


# Trustworthiness



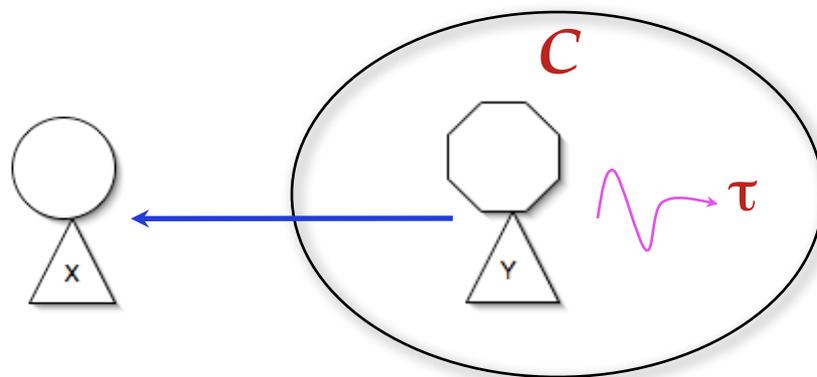
*Trustworthiness*<sub>YC</sub> ( $\tau$ )

# Trustworthiness



$\text{Trustworthiness}_{YC} (\tau)$   
 $\text{Trustworthiness}_{YXC} (\tau)$

# Trustworthiness



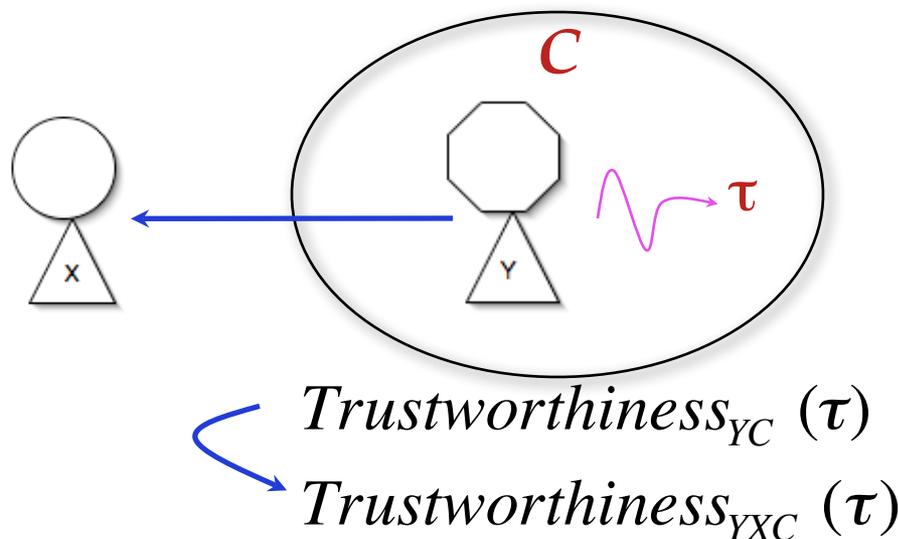
$Trustworthiness_{YC}(\tau)$

$Trustworthiness_{YXC}(\tau)$

*In general we can say:*

1)  $DoT_{XY\tau}$  corresponds with  $Bel(X Trustworthiness_{YX}(\tau))$

# Trustworthiness

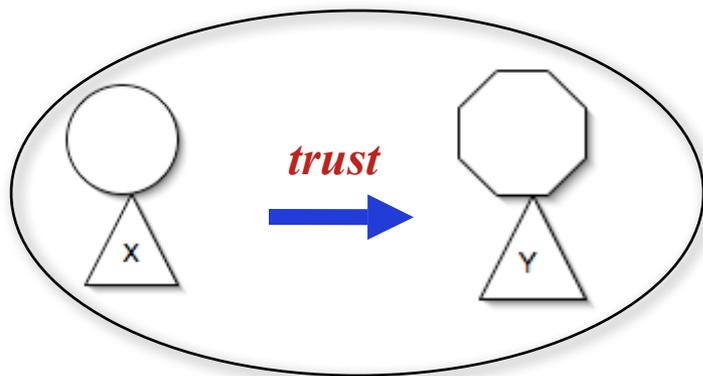


*In general we can say:*

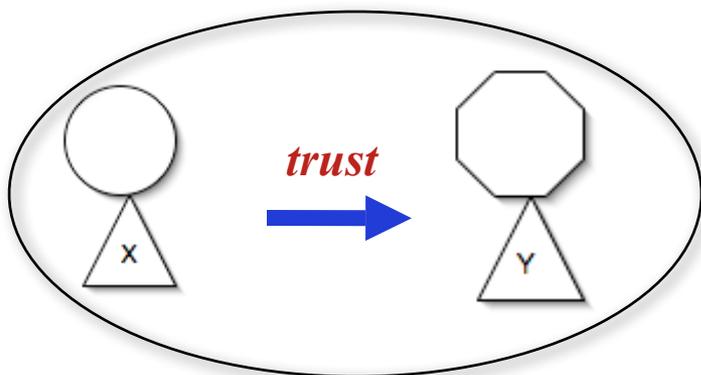
1)  $DoT_{XY\tau}$  corresponds with  $Bel(X \text{ Trustworthiness}_{YX}(\tau))$

2)  $DoT_{XY\tau} > \sigma \rightarrow Bel(X \text{ Trustworthiness}_{YX}(\tau) > \sigma)$

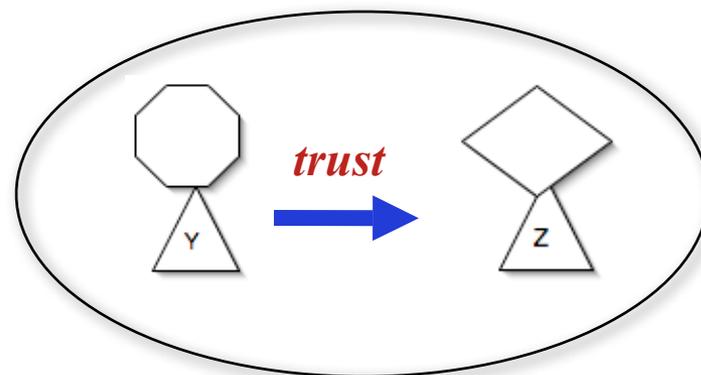
# Transitivity in Trust



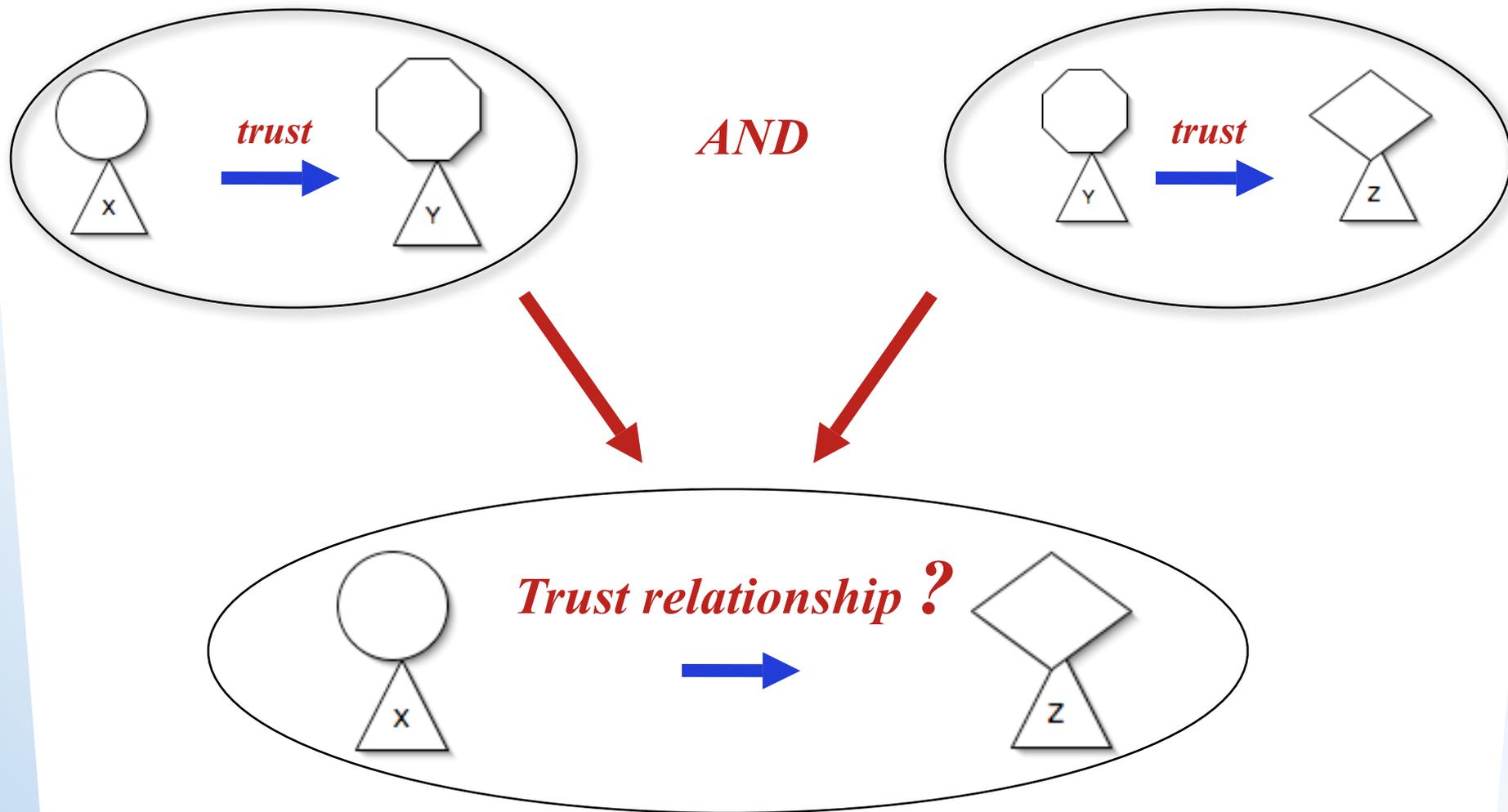
# Transitivity in Trust



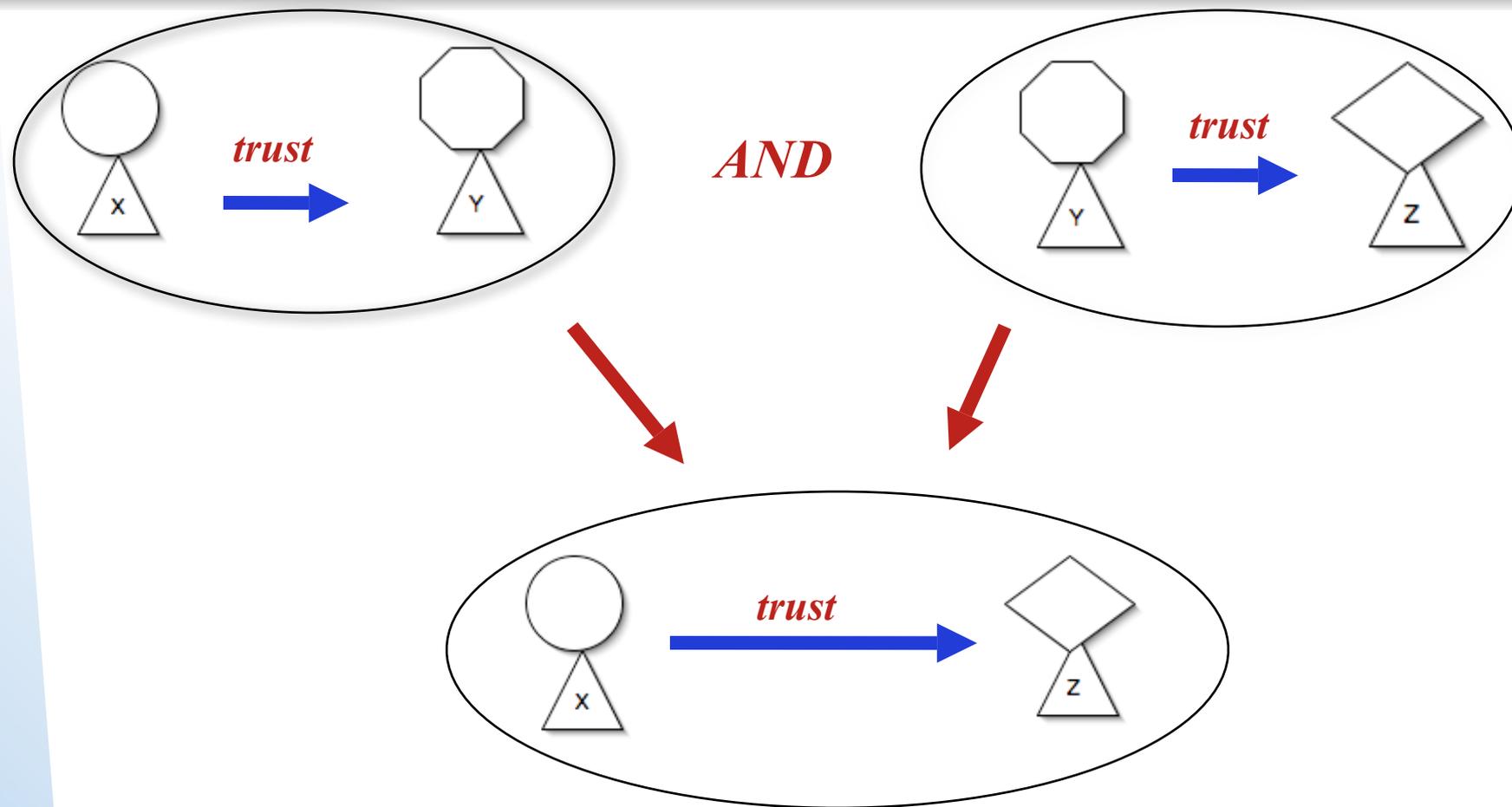
**AND**



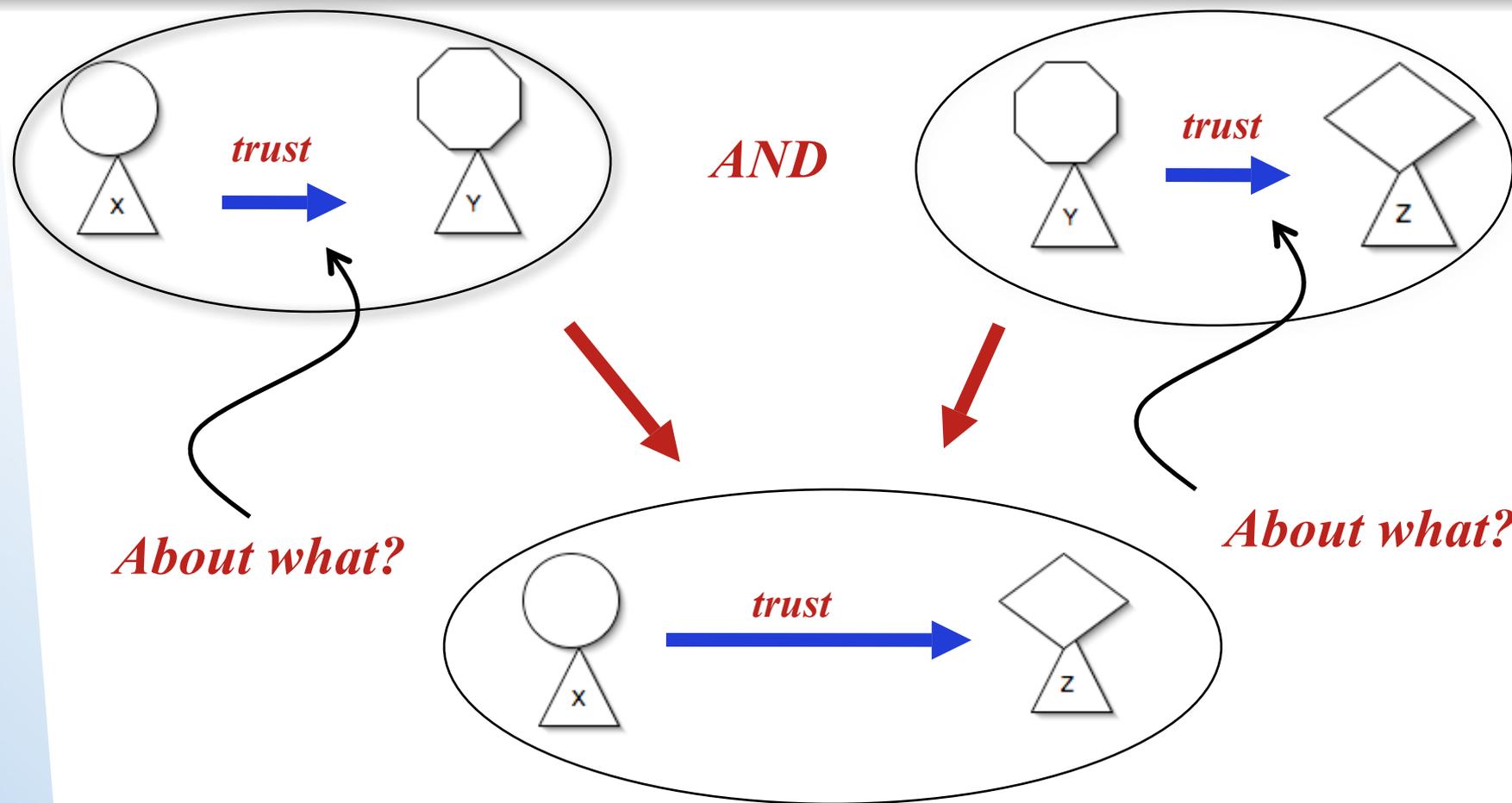
# Transitivity in Trust



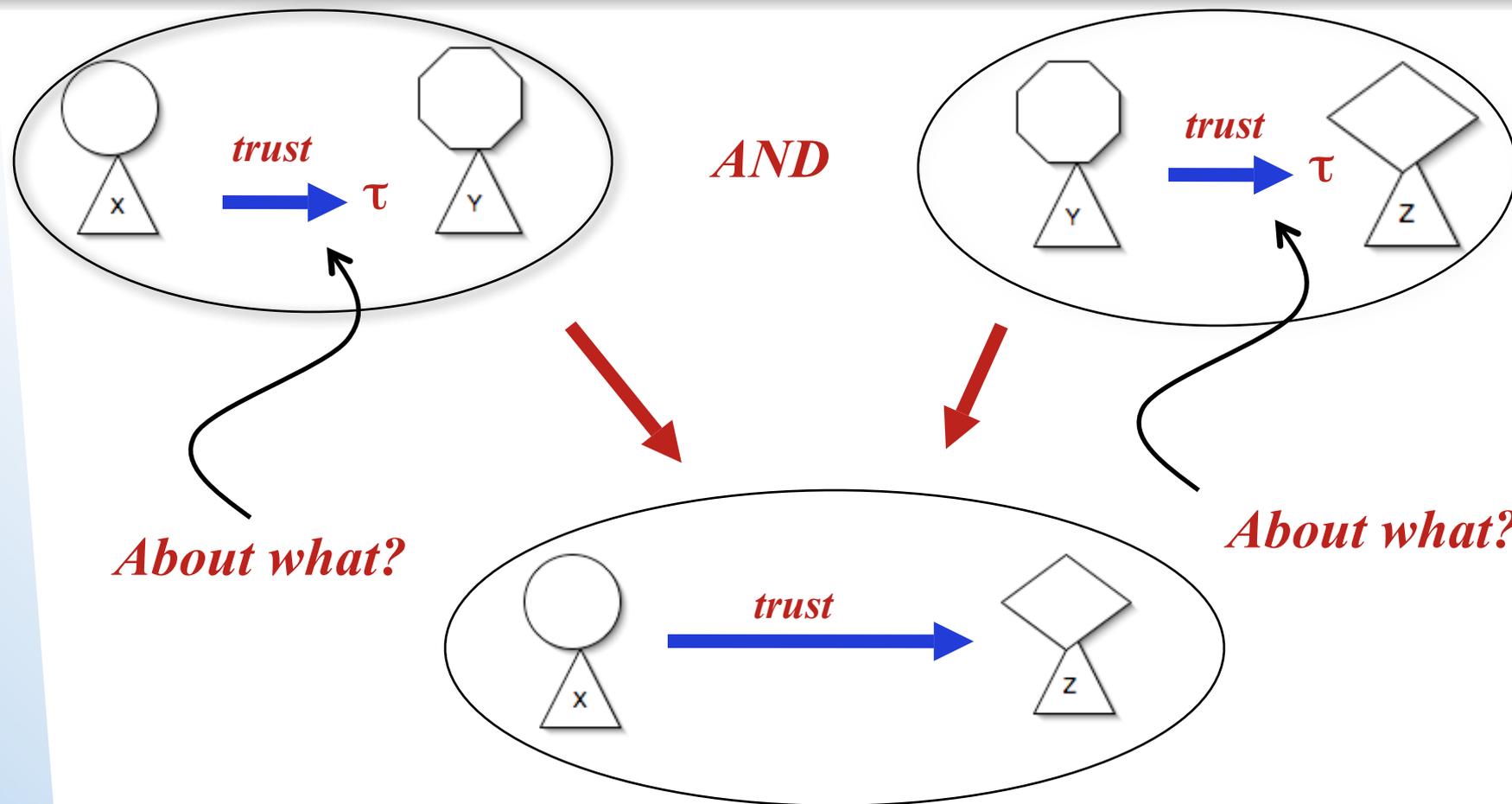
# Transitivity: the task relevance



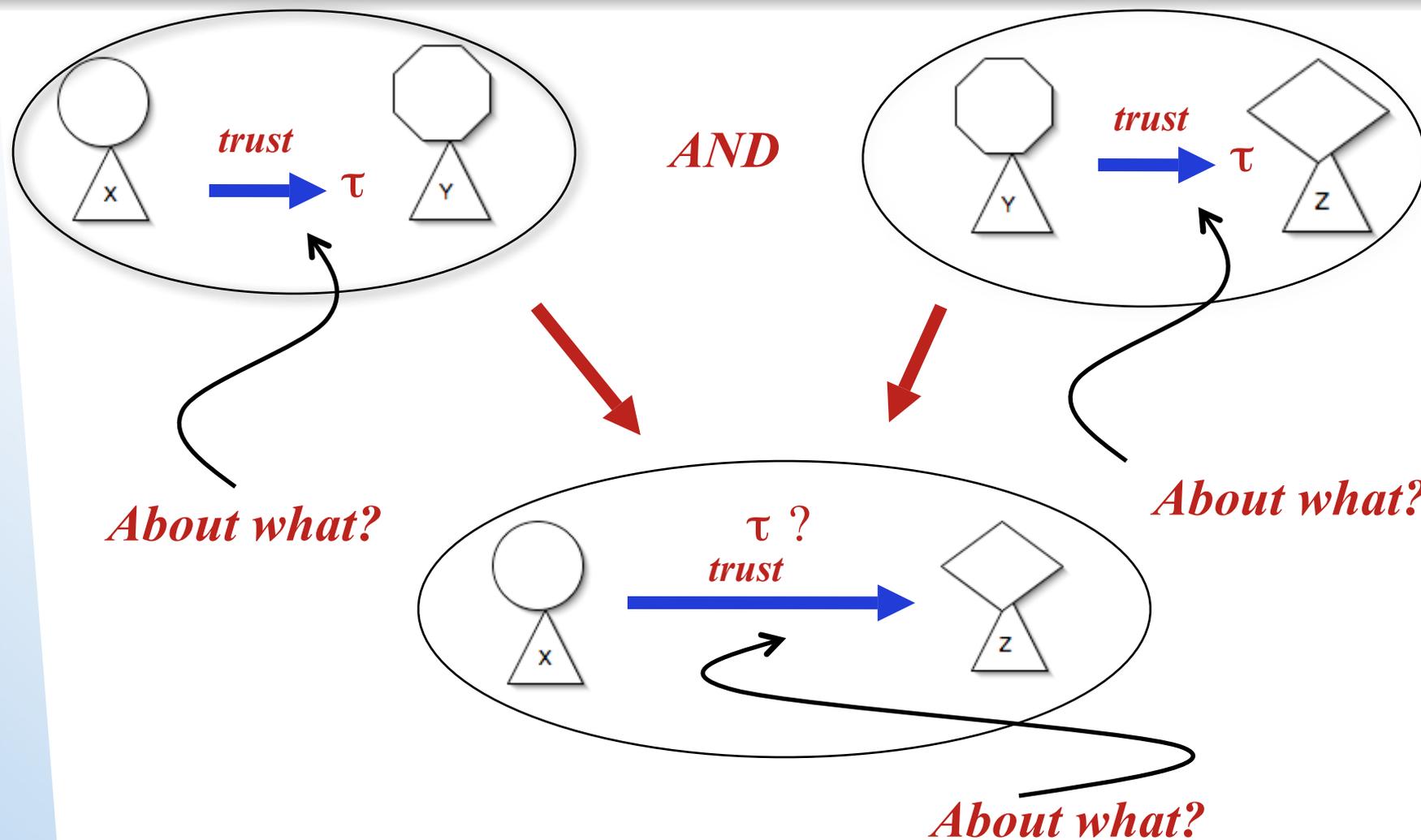
# Transitivity: the task relevance



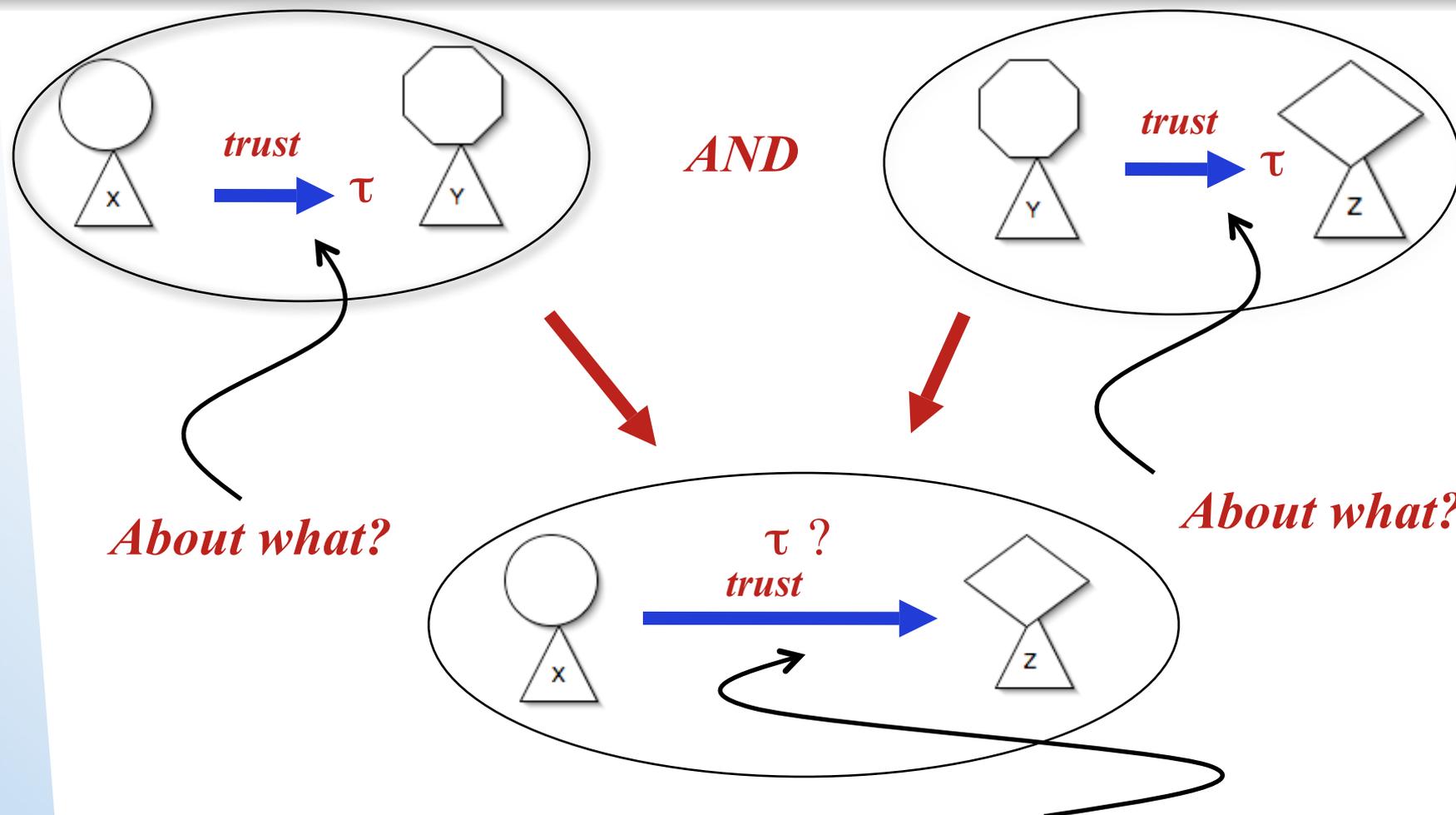
# Transitivity: the task relevance



# Transitivity: the task relevance



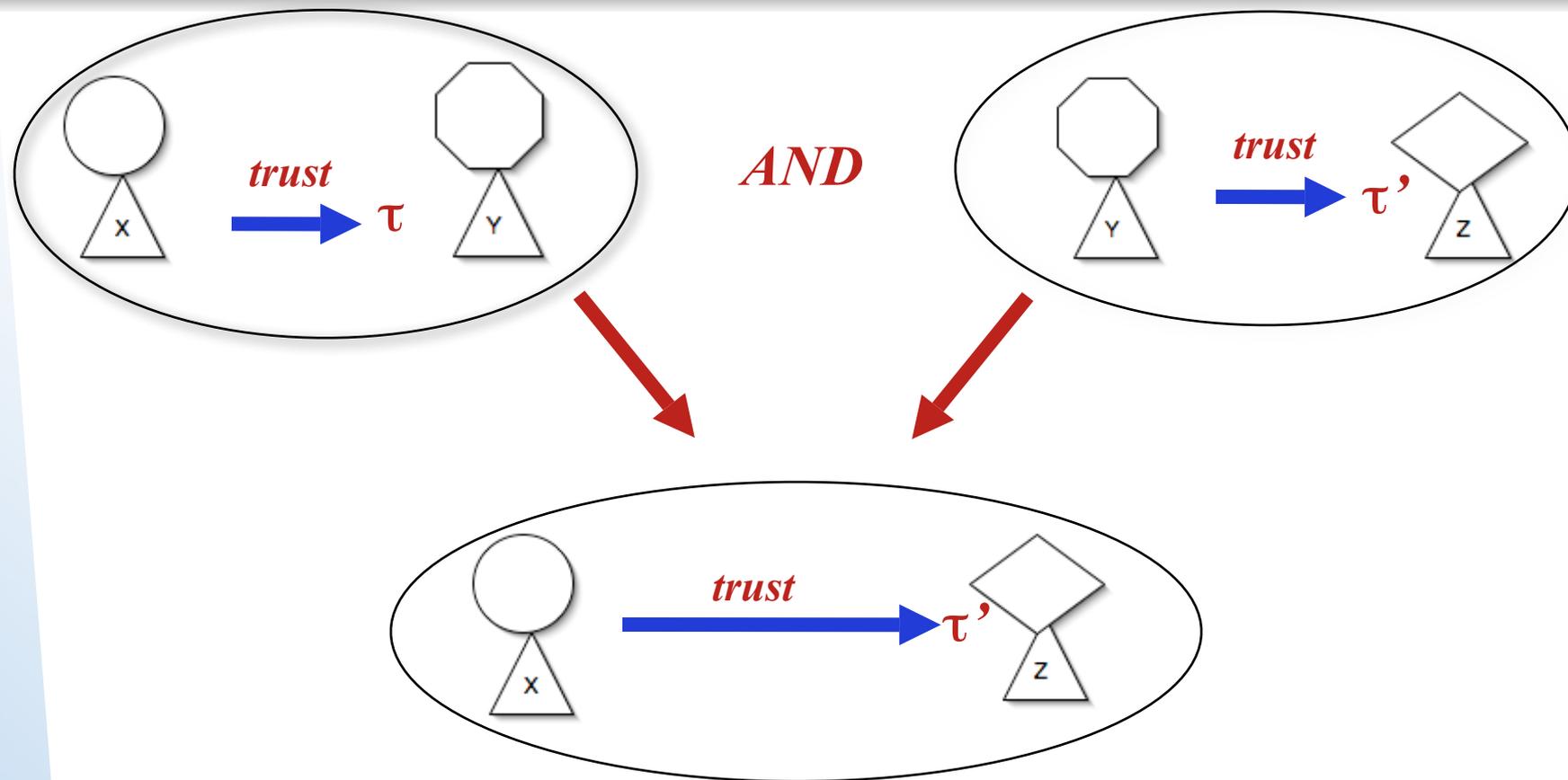
# Transitivity: the task relevance



$\tau$  = task about an health problem  
- Y is a doctor

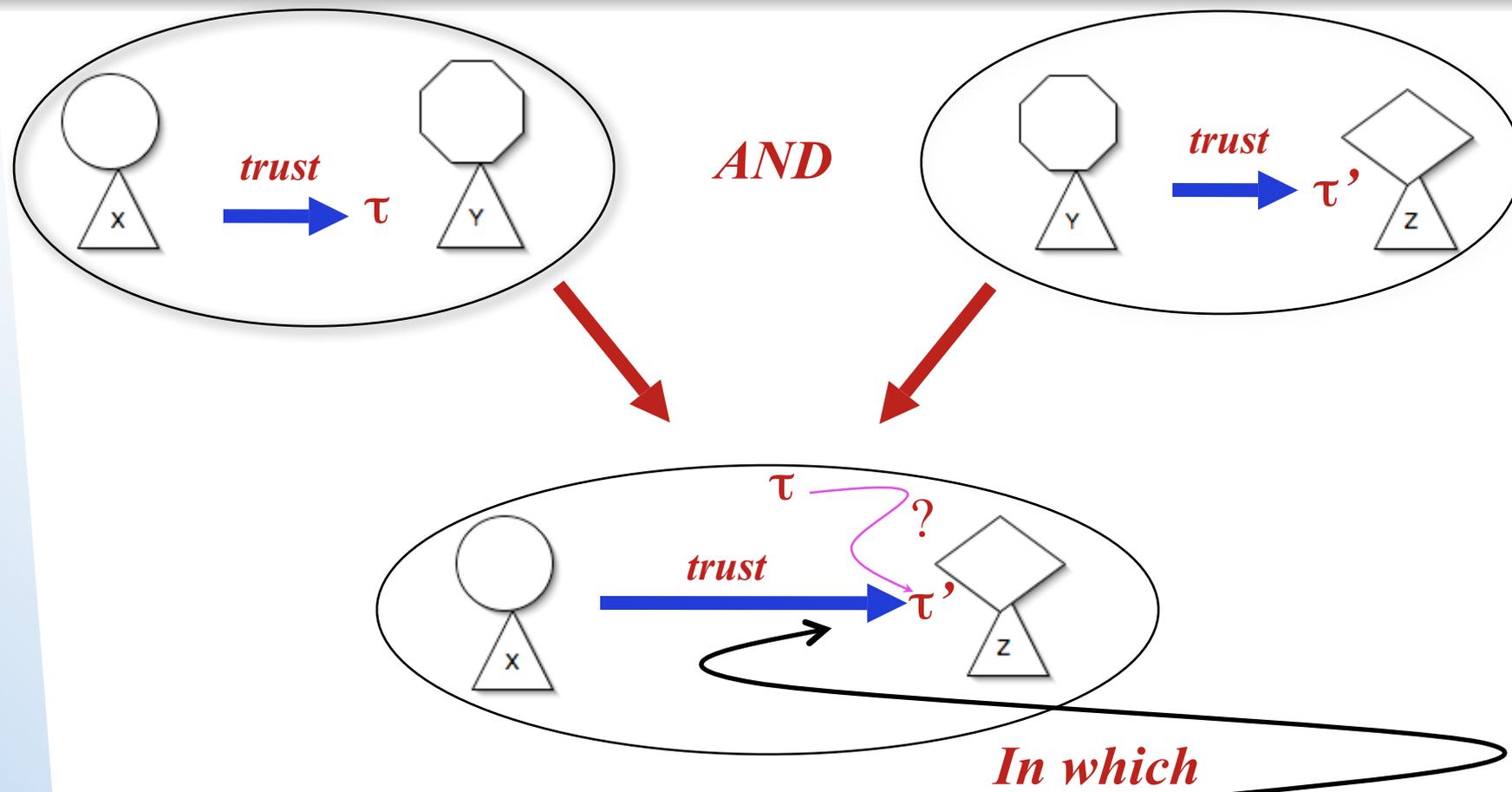
*About what?*

# Different tasks in the originating relationships



$\tau$  = about mechanic questions  
 $\tau'$  = about medical questions

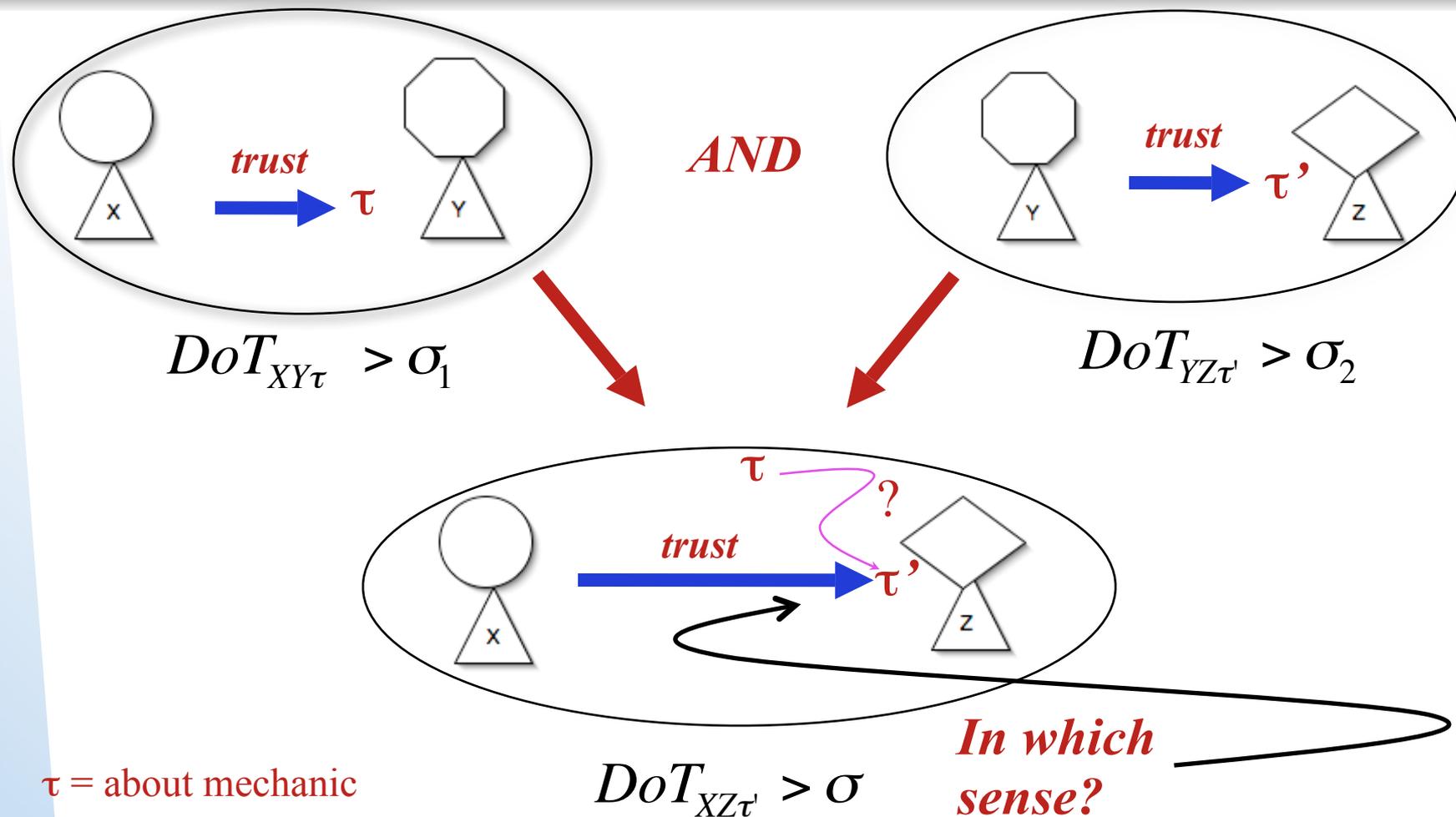
# Different tasks in the originating relationships



*In which sense?*

$\tau$  = about mechanic questions  
 $\tau'$  = about medical questions

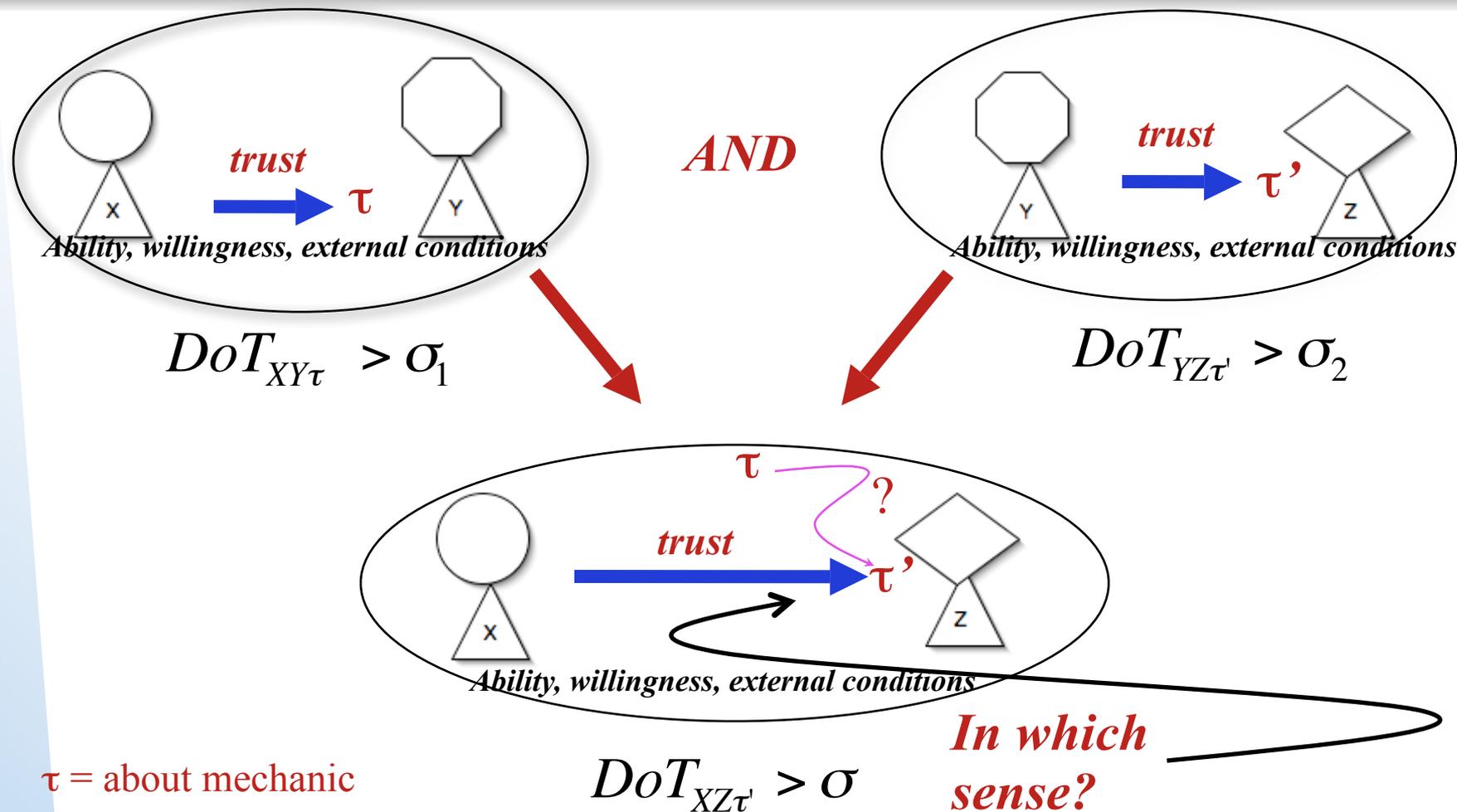
# Different tasks in the originating relationships



$\tau$  = about mechanic questions  
 $\tau'$  = about medical questions

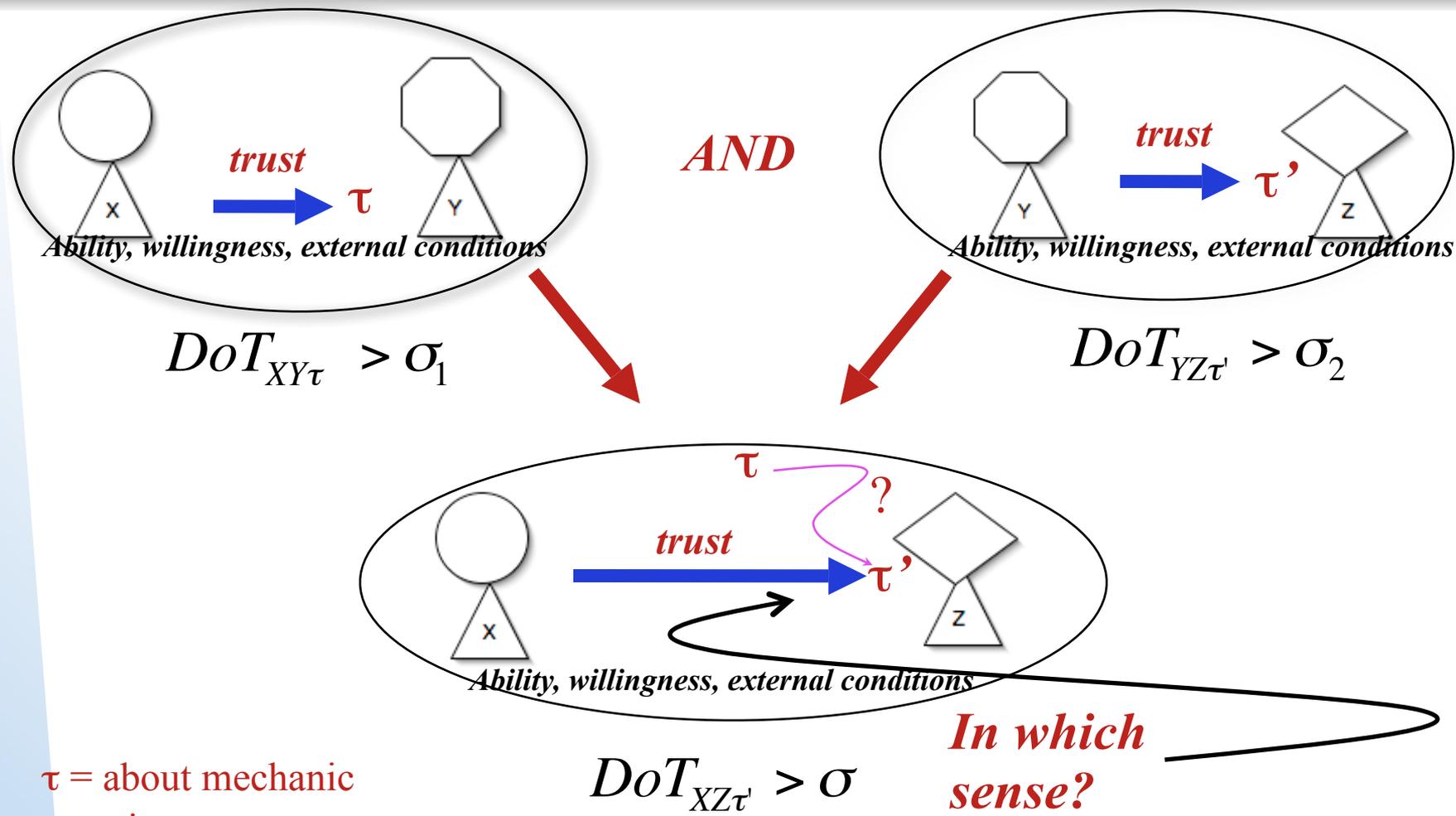
*In which sense?*

# Different tasks in the originating relationships



$\tau$  = about mechanic questions  
 $\tau'$  = about medical questions

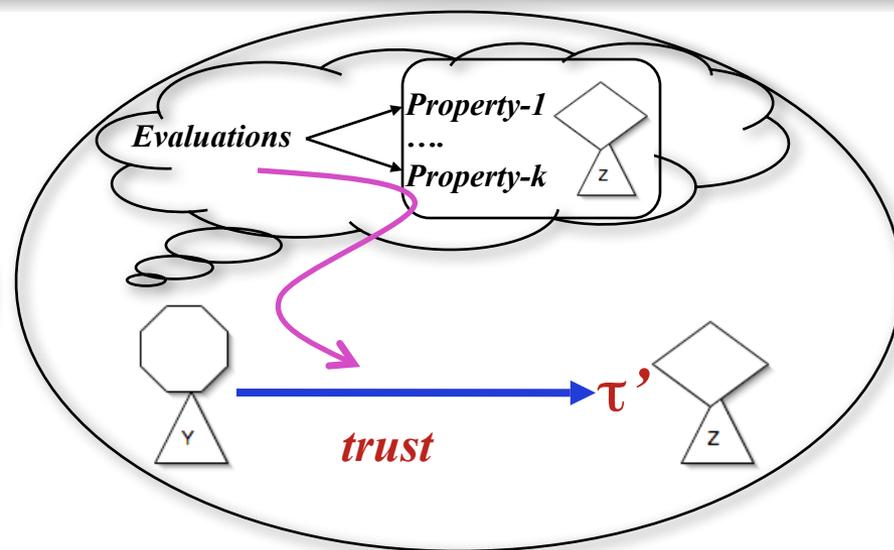
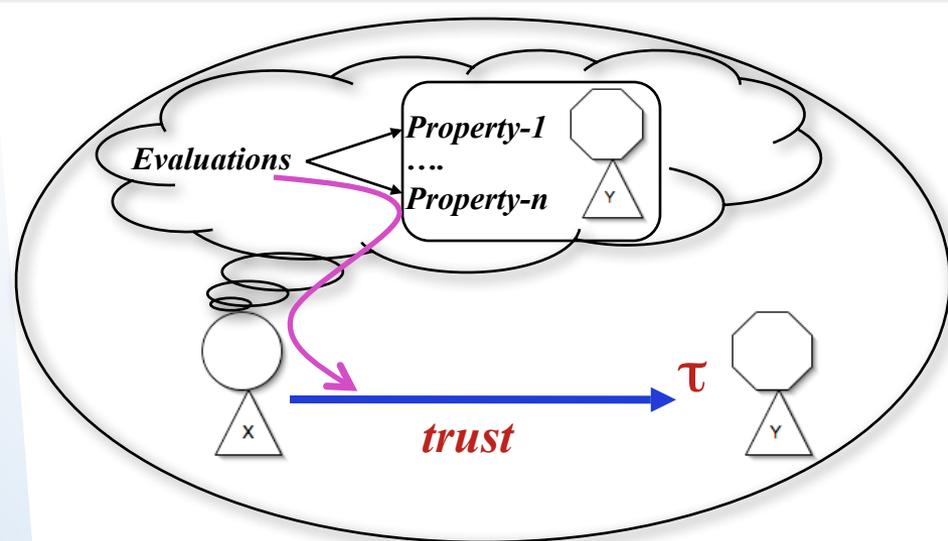
# Different tasks in the originating relationships



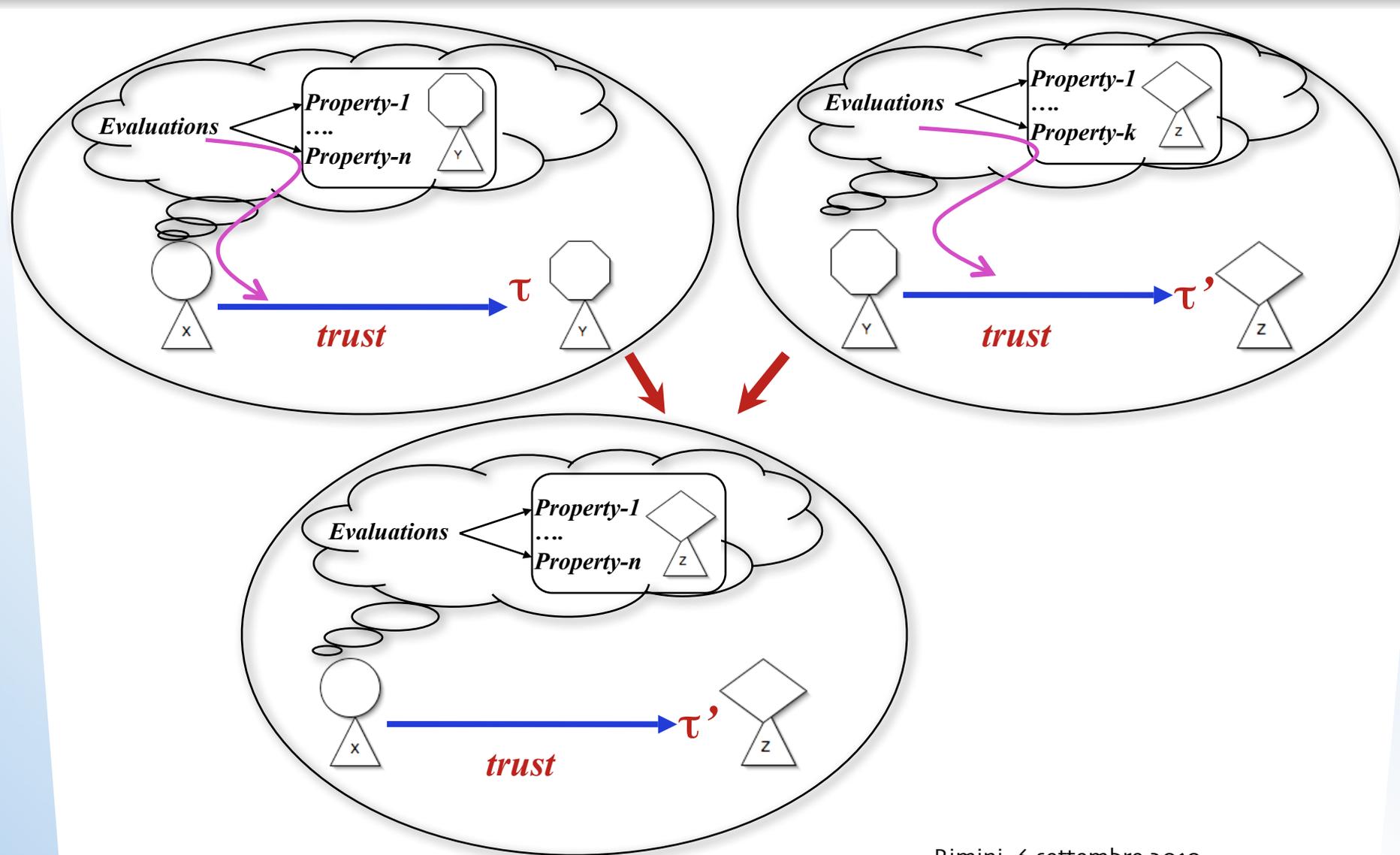
$\tau$  = about mechanic questions  
 $\tau'$  = about medical questions

**We hypothesize:**  $\sigma = \sigma_1 = \sigma_2$

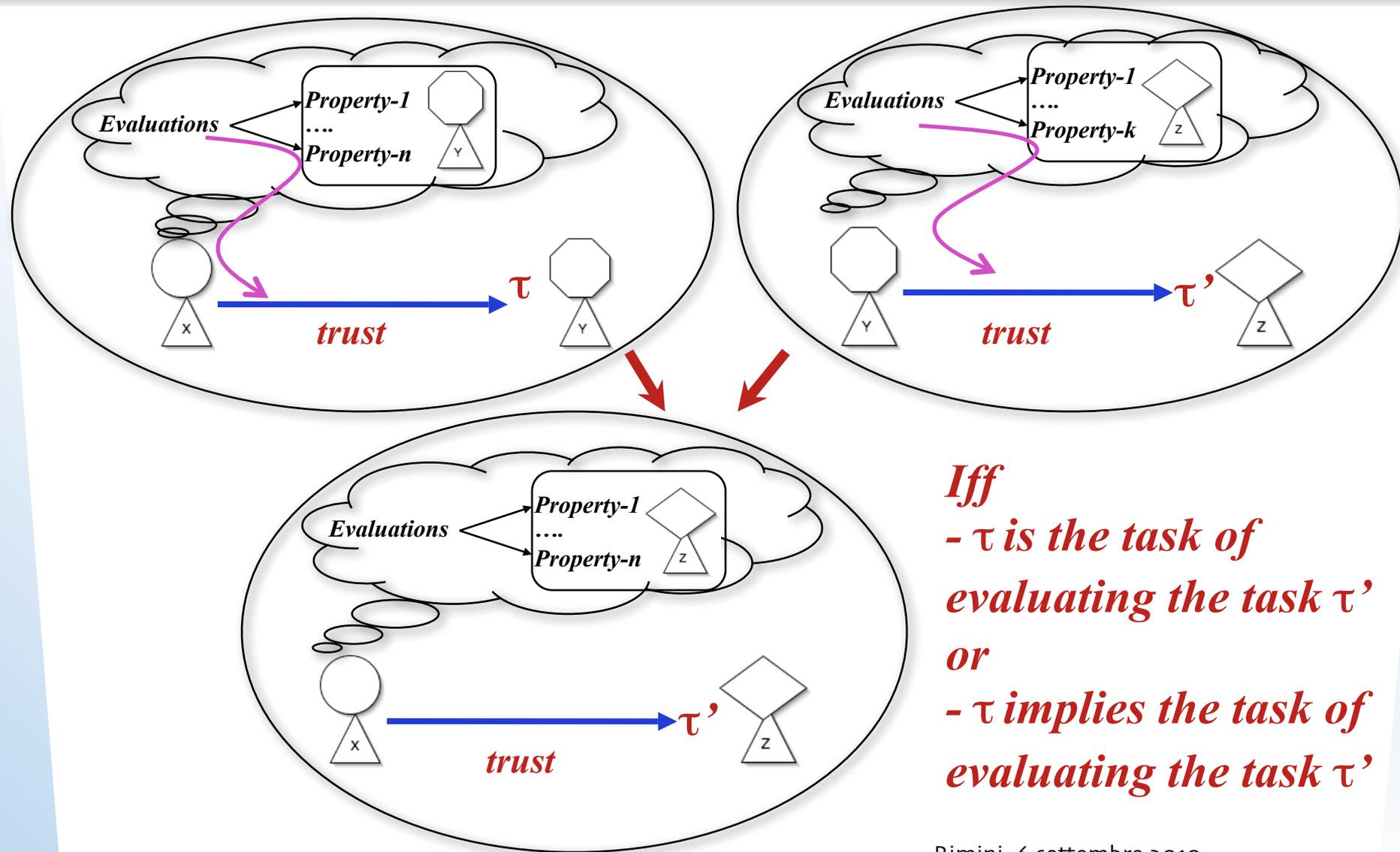
# Evaluations as mediators for trusting



# Evaluations as mediators for trusting



# Evaluations as mediators for trusting



*Iff*  
 -  $\tau$  is the task of evaluating the task  $\tau'$   
 or  
 -  $\tau$  implies the task of evaluating the task  $\tau'$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*



# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$DoT_{YZ\tau} > \sigma$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$DoT_{YZ\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau} > \sigma)$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

~~$$DoT_{YZ\tau} > \sigma$$~~

$$Bel(X DoT_{YZ\tau} > \sigma)$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

~~$$DoT_{YZ\tau} > \sigma$$~~

$$Bel(X DoT_{YZ\tau} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \longrightarrow Bel(X Trustworthiness_{YX}(eval(\tau)) > \sigma)$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

*What is the meaning?*

$$DoT_{XY\tau} > \sigma$$

~~$$DoT_{YZ\tau} > \sigma$$~~

$$Bel(X DoT_{YZ\tau} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau)) > \sigma)$$

# Resuming on trust transitivity (1)

*Same task:*  $\tau = \tau'$

*assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

*What is the meaning?*

$$DoT_{XY\tau} > \sigma$$

~~$$DoT_{YZ\tau} > \sigma$$~~

$$Bel(X DoT_{YZ\tau} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau)) > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau) \geq Trustworthiness_{ZY}(\tau))$$

# Resuming on trust transitivity (1)

**Same task:**  $\tau = \tau'$

**assuming:**  $\sigma = \sigma_X = \sigma_Y$

**If**

*What is the meaning?*

$$DoT_{XY\tau} > \sigma$$

~~$$DoT_{YZ\tau} > \sigma$$~~

$$Bel(X DoT_{YZ\tau} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau)) > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau) \geq Trustworthiness_{ZY}(\tau))$$

**Then**

$$DoT_{XZ\tau} > \sigma$$

# Resuming on trust transitivity (2)

*Different tasks:*  $\tau \neq \tau'$       *assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*



## Resuming on trust transitivity (2)

*Different tasks:*  $\tau \neq \tau'$       *assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

# Resuming on trust transitivity (2)

*Different tasks:*  $\tau \neq \tau'$       *assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau'} > \sigma)$$

## Resuming on trust transitivity (2)

*Different tasks:*  $\tau \neq \tau'$       *assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau'} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \longrightarrow Bel(X Trustworthiness_{YX}(eval(\tau')) > \sigma)$$

# Resuming on trust transitivity (2)

**Different tasks:**  $\tau \neq \tau'$       **assuming:**  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau'} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau')) > \sigma)$$

-  $\tau$  is the task of evaluating  $\tau'$  or  
-  $\tau$  implies the task of evaluating the task  $\tau'$

# Resuming on trust transitivity (2)

*Different tasks:*  $\tau \neq \tau'$       *assuming:*  $\sigma = \sigma_X = \sigma_Y$

*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau'} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau')) > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau') \geq Trustworthiness_{ZY}(\tau'))$$

-  $\tau$  is the task of evaluating  $\tau'$  or  
-  $\tau$  implies the task of evaluating the task  $\tau'$

# Resuming on trust transitivity (2)

**Different tasks:**  $\tau \neq \tau'$       **assuming:**  $\sigma = \sigma_X = \sigma_Y$

**If**

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau'} > \sigma)$$

$$Bel(X Trustworthiness_{YX}(\tau) > \sigma) \rightarrow Bel(X Trustworthiness_{YX}(eval(\tau')) > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau') \geq Trustworthiness_{ZY}(\tau'))$$

**Then**

$$DoT_{XZ\tau'} > \sigma$$

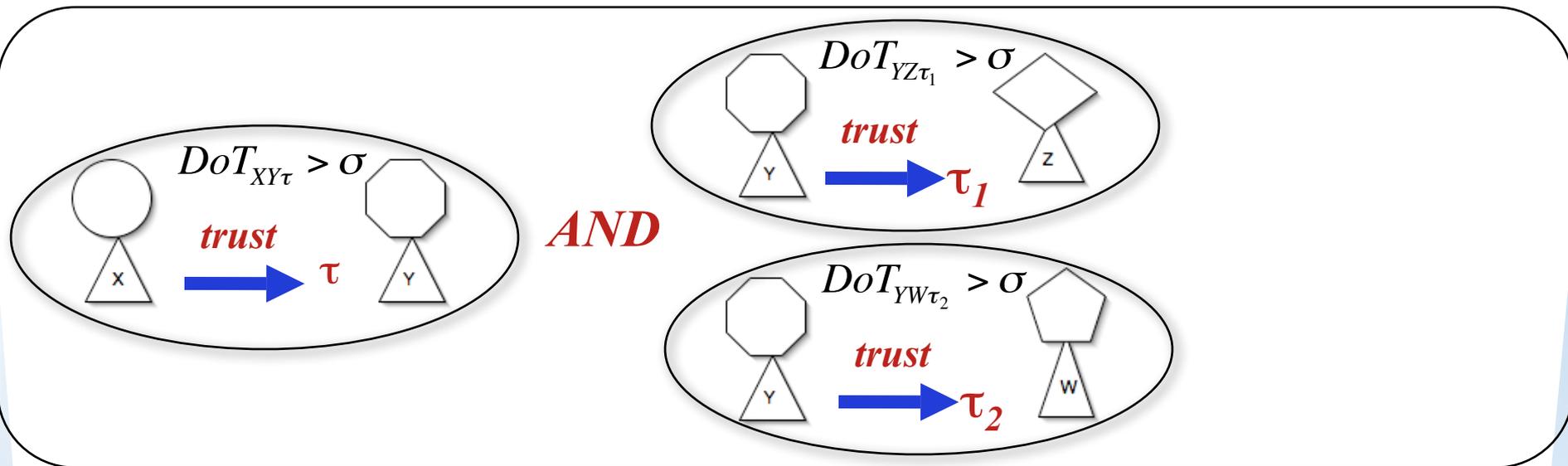
-  $\tau$  is the task of evaluating  $\tau'$  or  
-  $\tau$  implies the task of evaluating the task  $\tau'$

- X trusts Y for a task  $\tau$
- In fulfilling  $\tau$ , Y delegates  $\tau_1$  and  $\tau_2$  to Z and W
  - ◆  $\tau_1$  and  $\tau_2$  are part of  $\tau$
  - ◆ X is aware of this sub-delegation
- Which kind of trust can be established by X towards Z and W ?

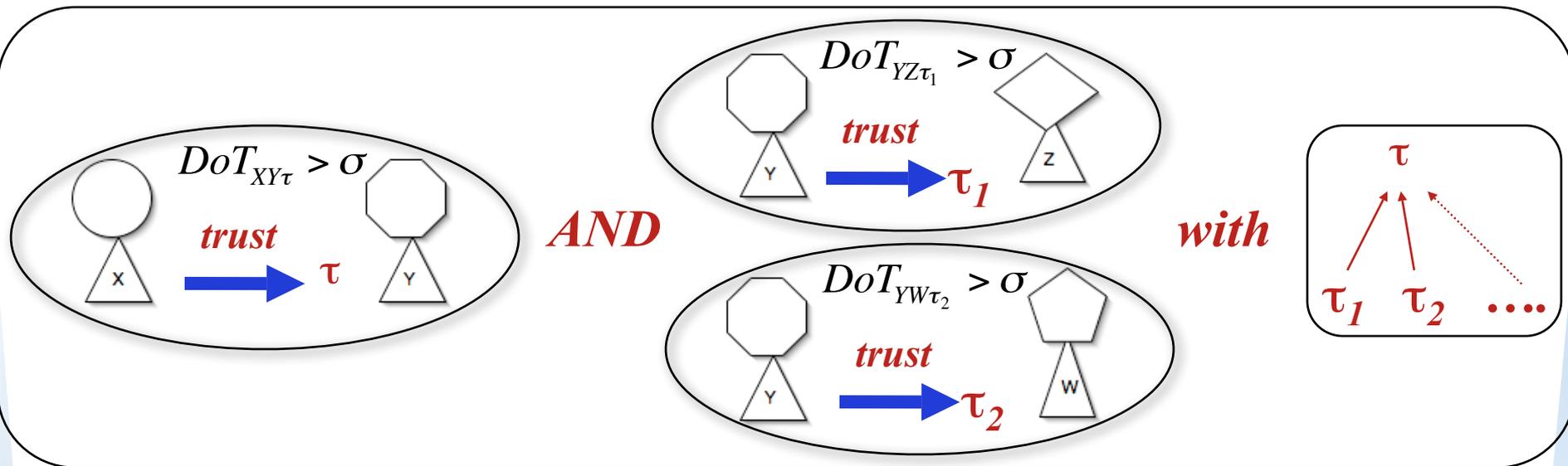
# Transitivity about delegated subtasks (1)



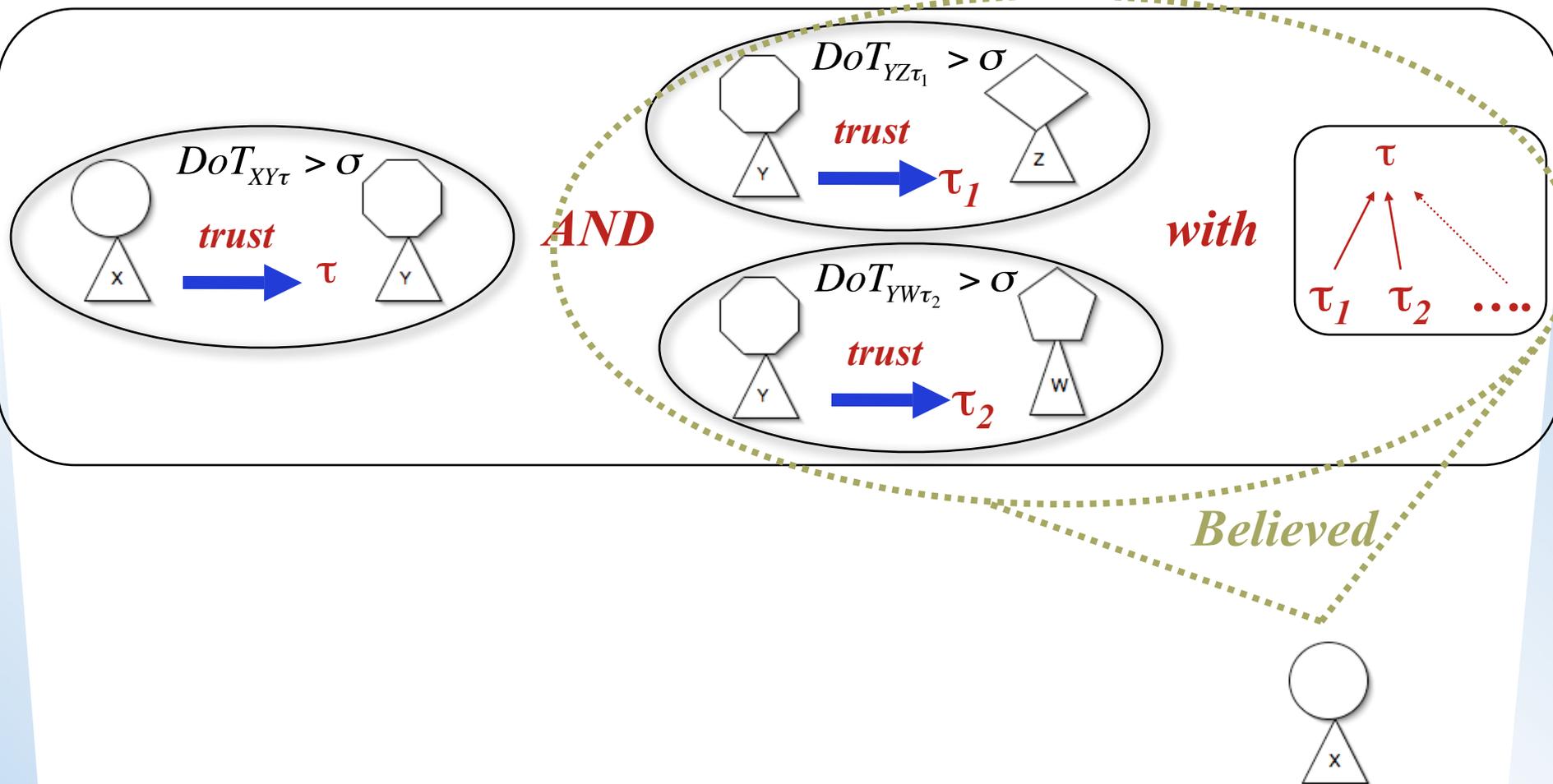
# Transitivity about delegated subtasks (1)



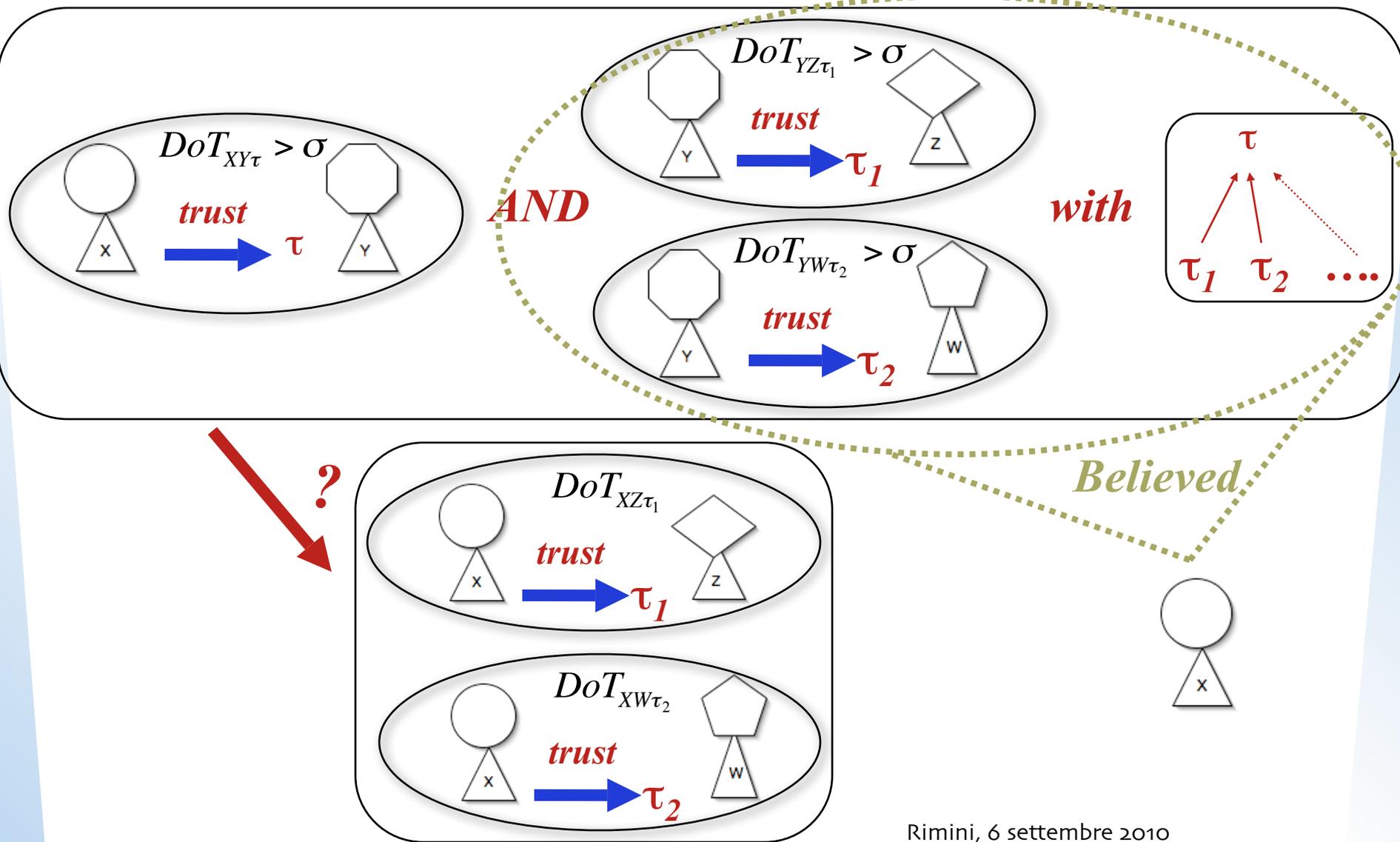
# Transitivity about delegated subtasks (1)



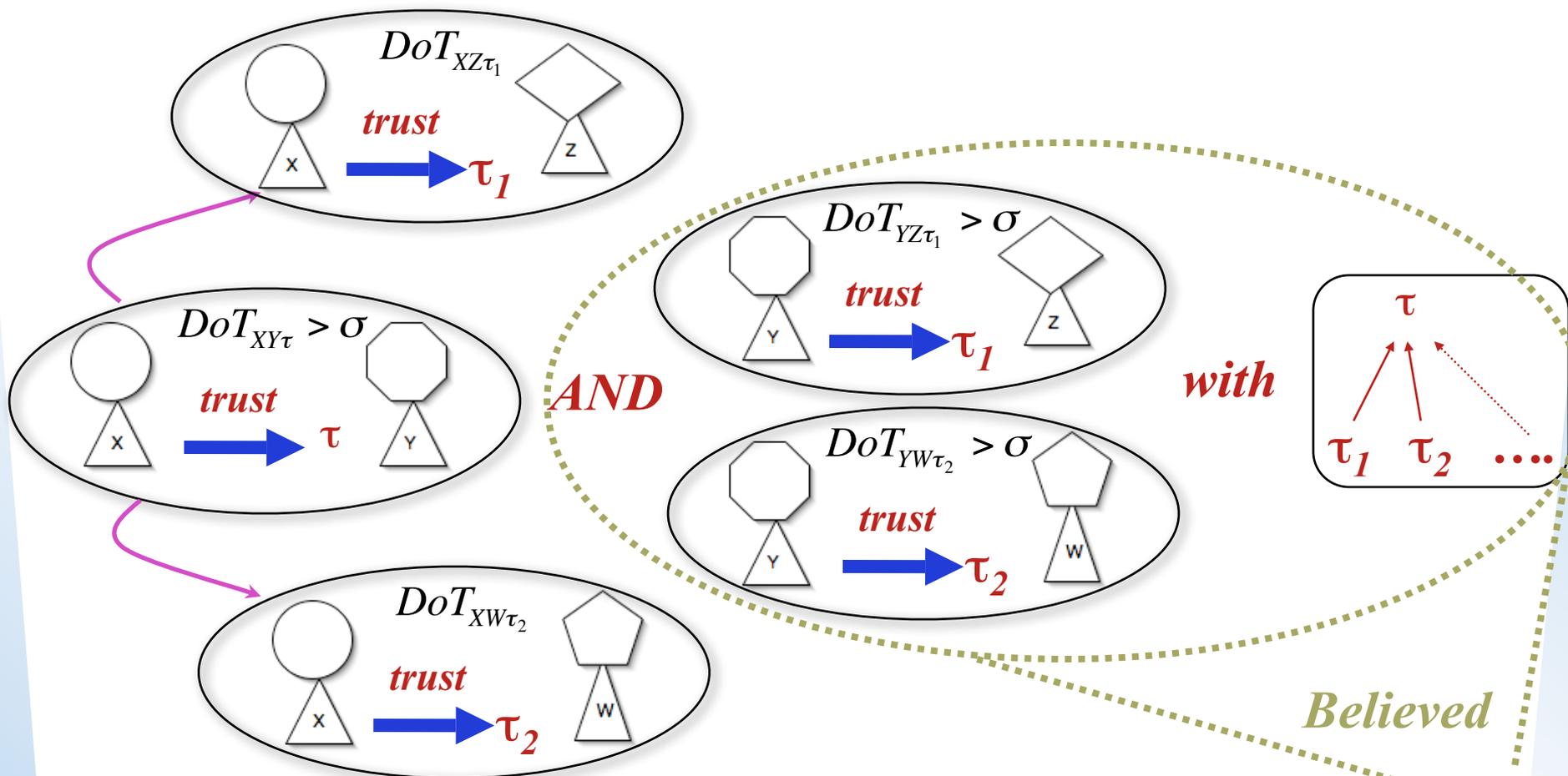
# Transitivity about delegated subtasks (1)



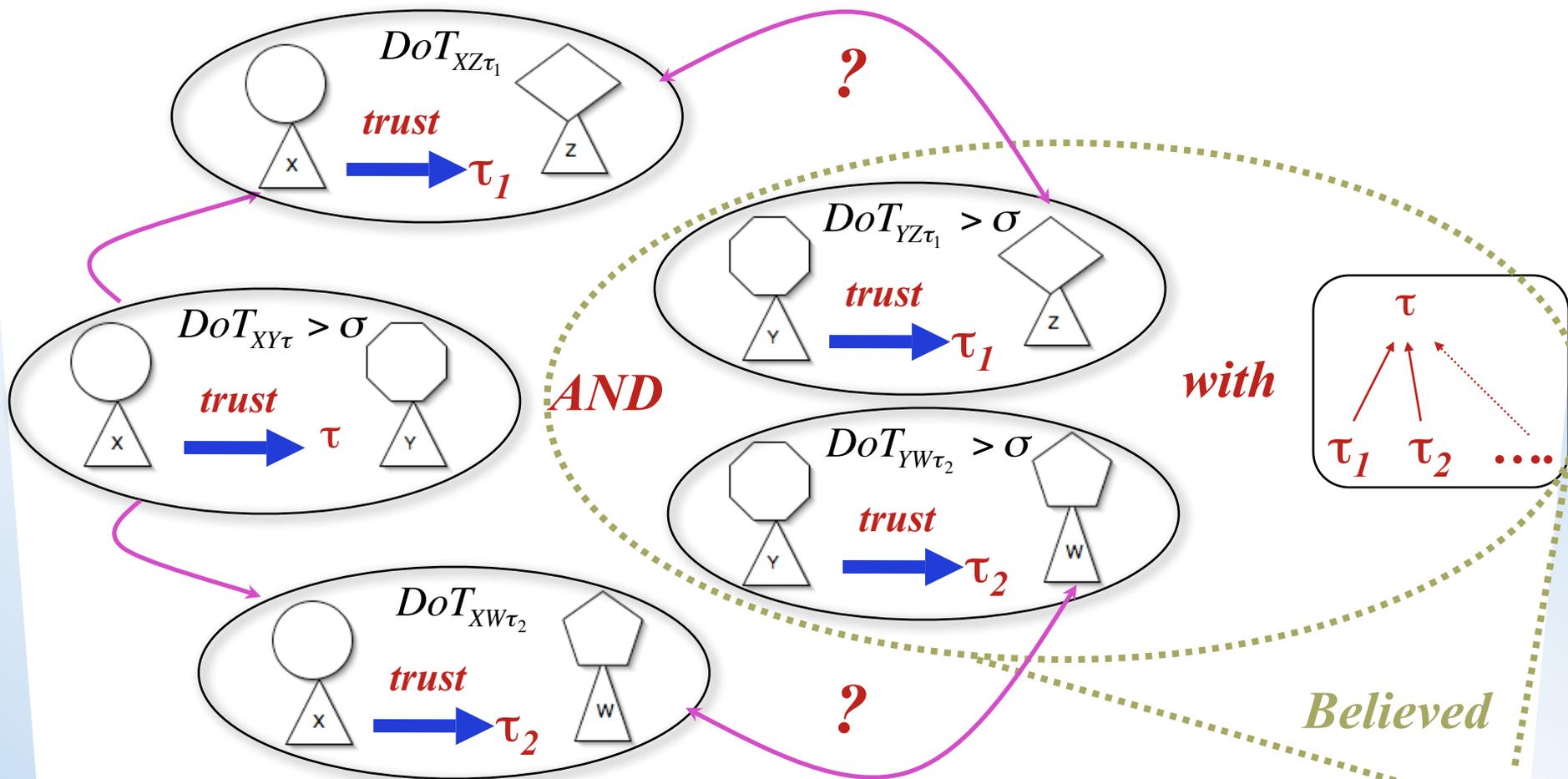
# Transitivity about delegated subtasks (1)



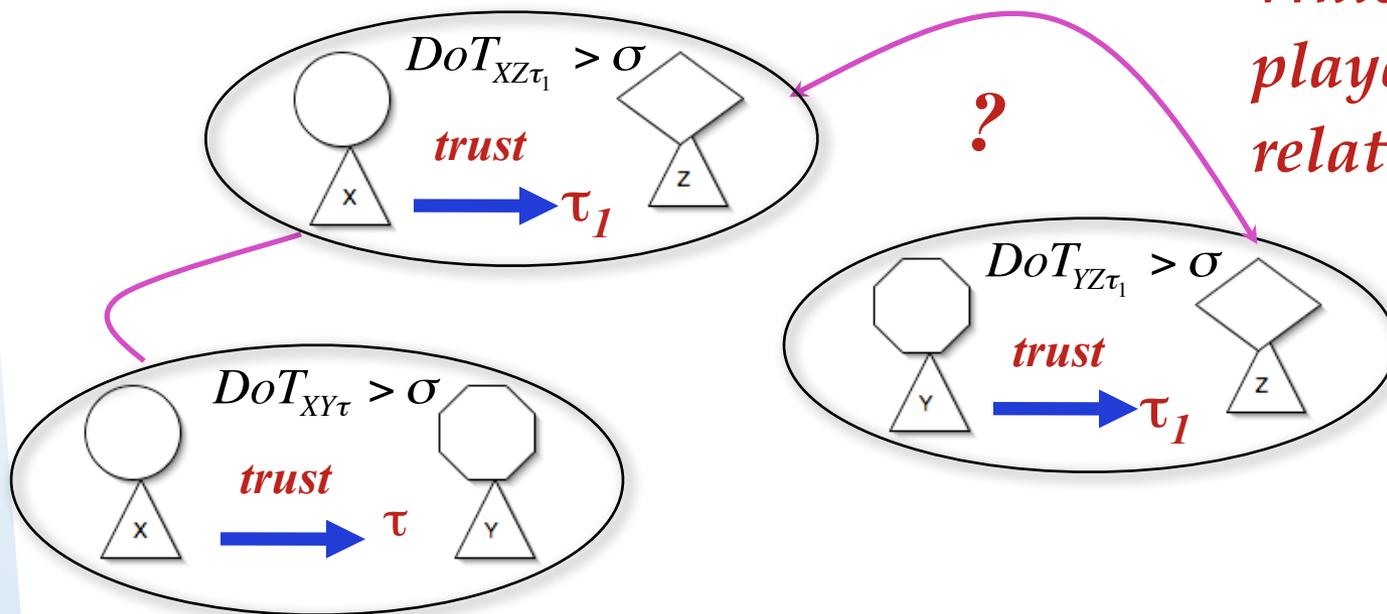
# Transitivity about delegated subtasks (2)



# Transitivity about delegated subtasks (2)



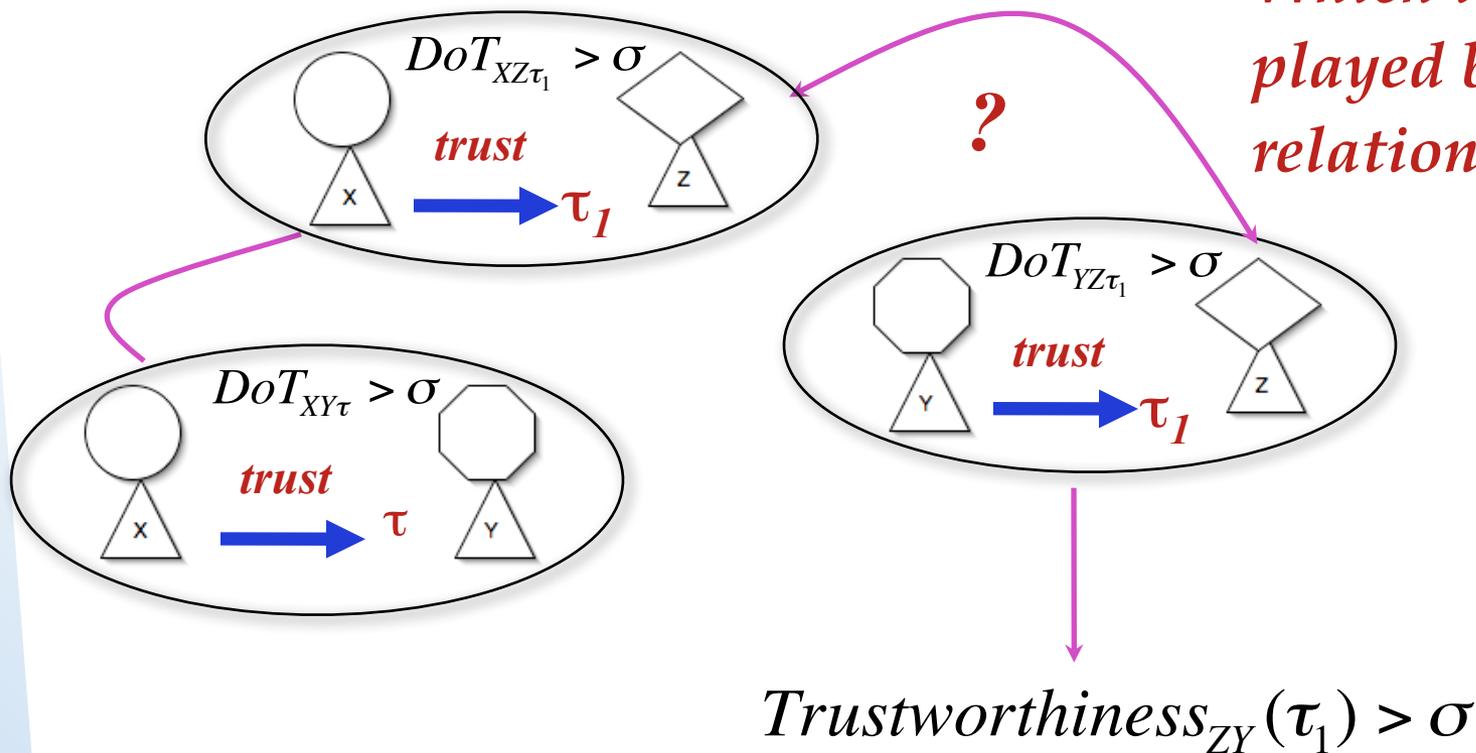
# Transitivity about delegated subtasks (3)



*Which is the role played by Y in trust relationship with Z?*

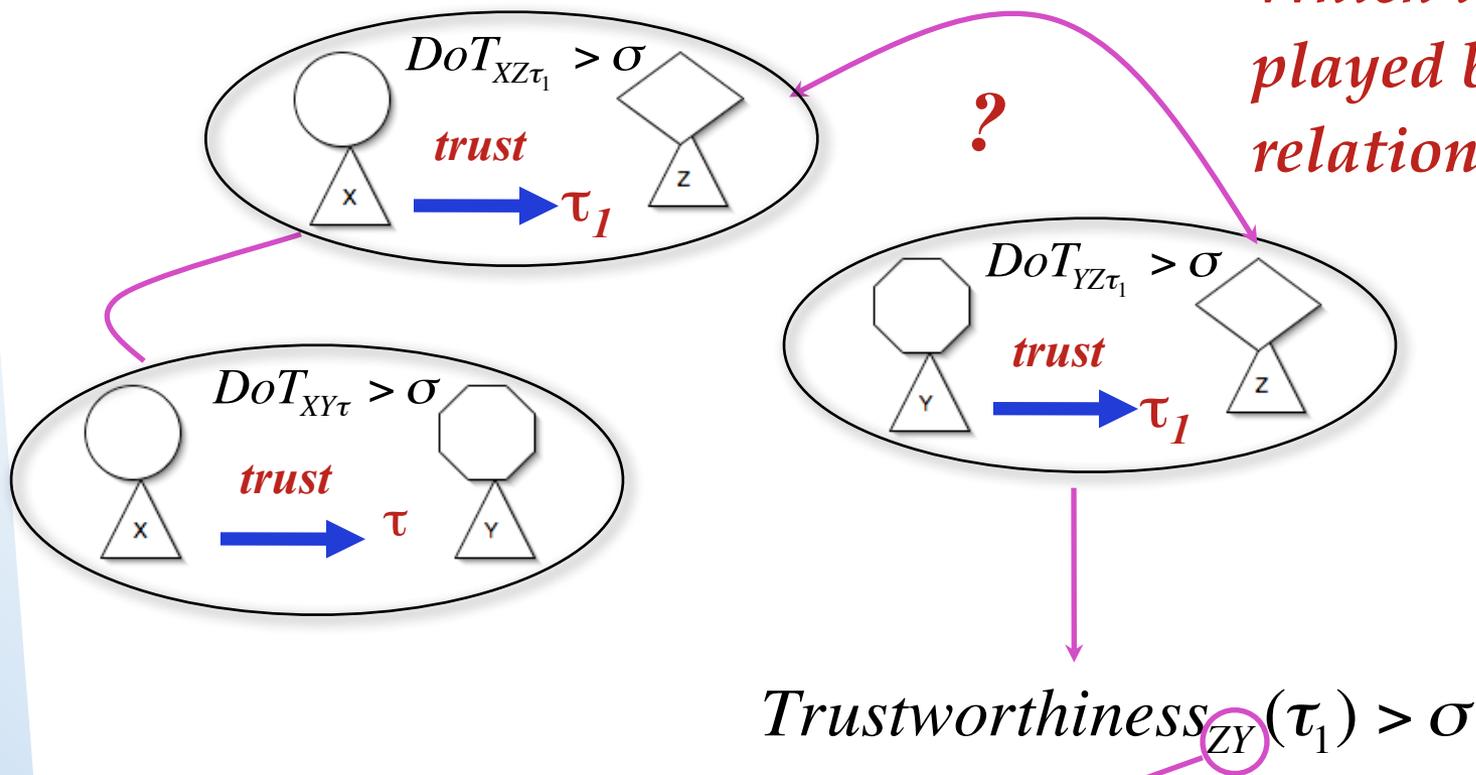
# Transitivity about delegated subtasks (3)

*Which is the role played by Y in trust relationship with Z?*



# Transitivity about delegated subtasks (3)

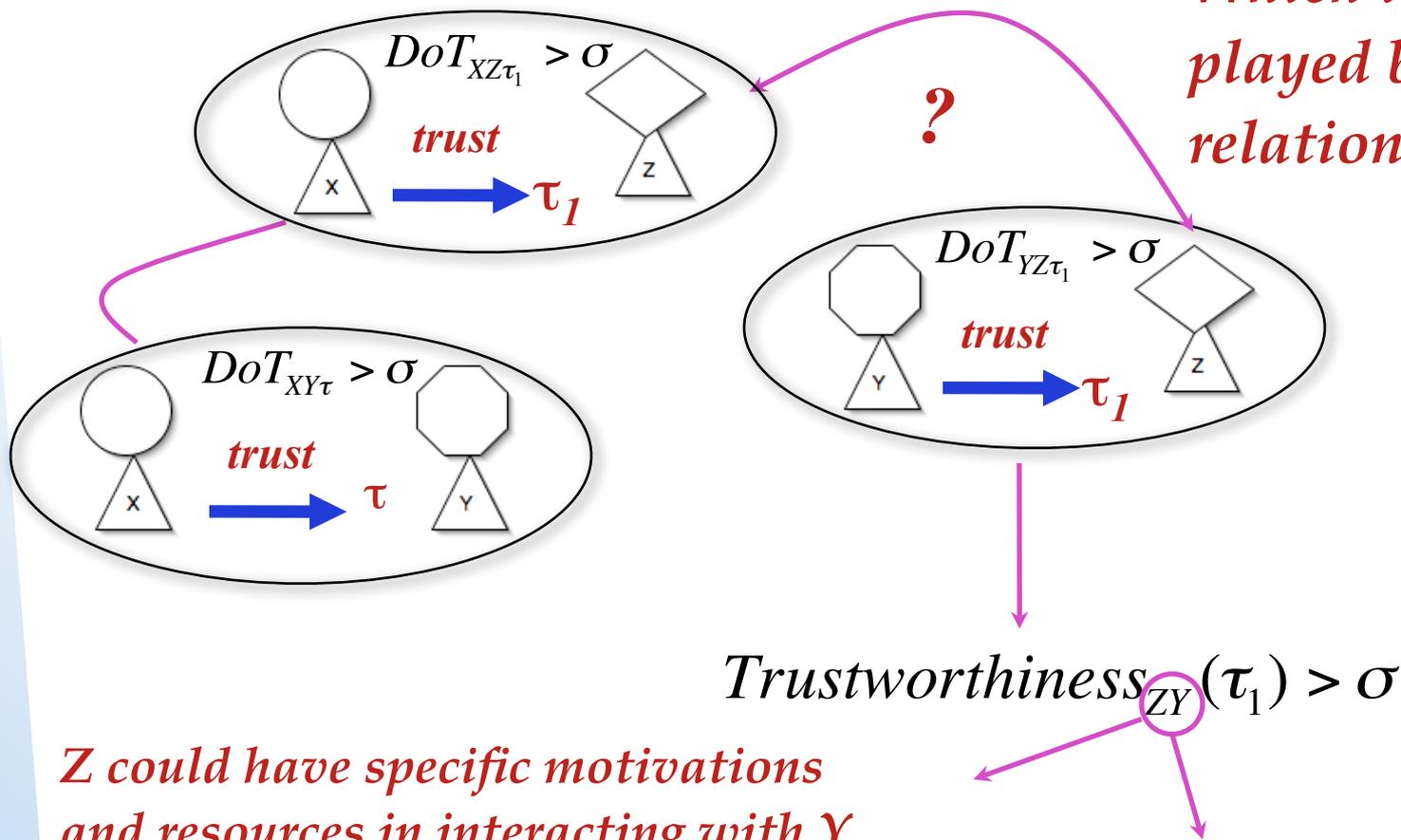
*Which is the role played by Y in trust relationship with Z?*



*Z could have specific motivations and resources in interacting with Y*

# Transitivity about delegated subtasks (3)

*Which is the role played by Y in trust relationship with Z?*

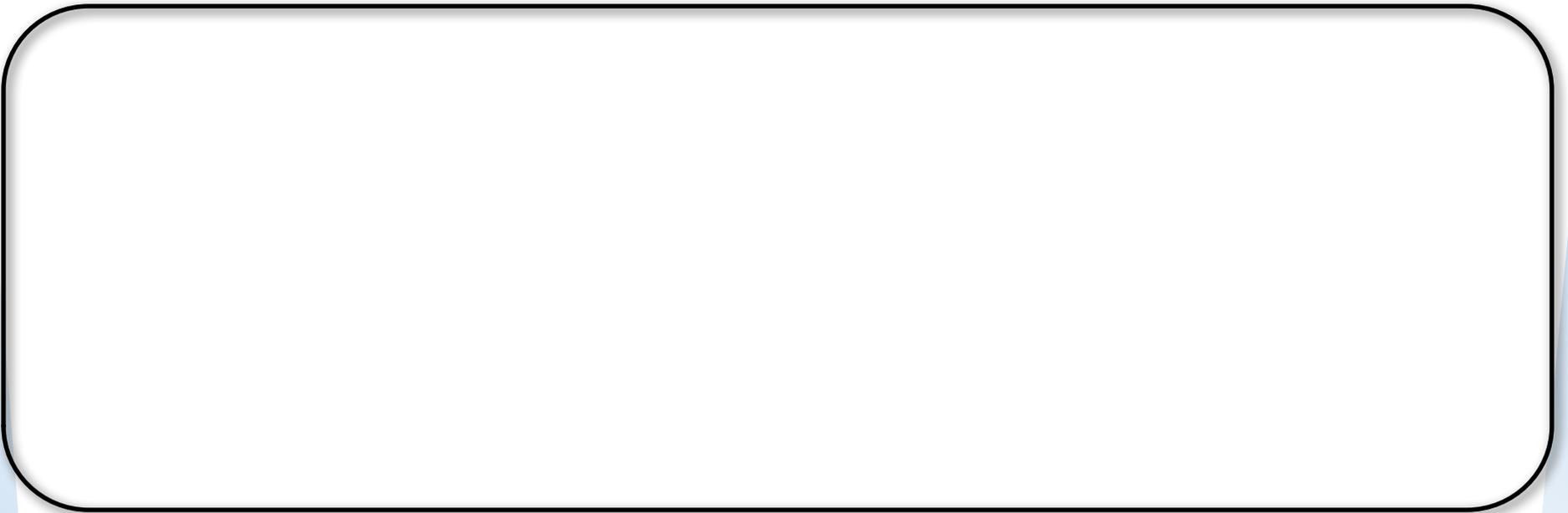


*Z could have specific motivations and resources in interacting with Y*

*Y is able to integrate and support the realization of the task*

# Transitivity about delegated subtasks (4)

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*



## Transitivity about delegated subtasks (4)

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

# Transitivity about delegated subtasks (4)

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau_1) \geq Trustworthiness_{ZY}(\tau_1))$$

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau_1) \geq Trustworthiness_{ZY}(\tau_1))$$


$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau_1) \geq Trustworthiness_{ZY}(\tau_1))$$

$$Bel(X Trustworthiness_{WX}(\tau_2) \geq Trustworthiness_{WY}(\tau_2))$$

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau_1) \geq Trustworthiness_{ZY}(\tau_1))$$

$$Bel(X Trustworthiness_{WX}(\tau_2) \geq Trustworthiness_{WY}(\tau_2))$$

# Transitivity about delegated subtasks (4)

$\tau_1, \tau_2$  *subtasks of*  $\tau$  *assuming:*  $\sigma = \sigma_X = \sigma_Y$   
*If*

$$DoT_{XY\tau} > \sigma$$

$$Bel(X DoT_{YZ\tau_1} > \sigma)$$

$$Bel(X DoT_{YW\tau_2} > \sigma)$$

$$Bel(X Trustworthiness_{ZX}(\tau_1) \geq Trustworthiness_{ZY}(\tau_1))$$

$$Bel(X Trustworthiness_{WX}(\tau_2) \geq Trustworthiness_{WY}(\tau_2))$$

*Then*

$$DoT_{XZ\tau_1} > \sigma$$

$$DoT_{XW\tau_2} > \sigma$$

# CONCLUSIONS

# CONCLUSIONS

- Without a socio-cognitive approach to transitivity of trust, a set of deep and interesting details of this property would not be clear and they would be of difficult understanding

- Without a socio-cognitive approach to transitivity of trust, a set of deep and interesting details of this property would not be clear and they would be of difficult understanding
- in Trust Transitivity we have to take in consideration two main factors:
  - the role of the agent that mediate the relationship as evaluator of the final entity (agent Z in our case) and
  - the potential differences between the trust relationships when we exchange the trustors