Intelligent Systems (AI-2)

Computer Science cpsc422, Lecture 13

Feb, 2, 2015

ILE: Challenges

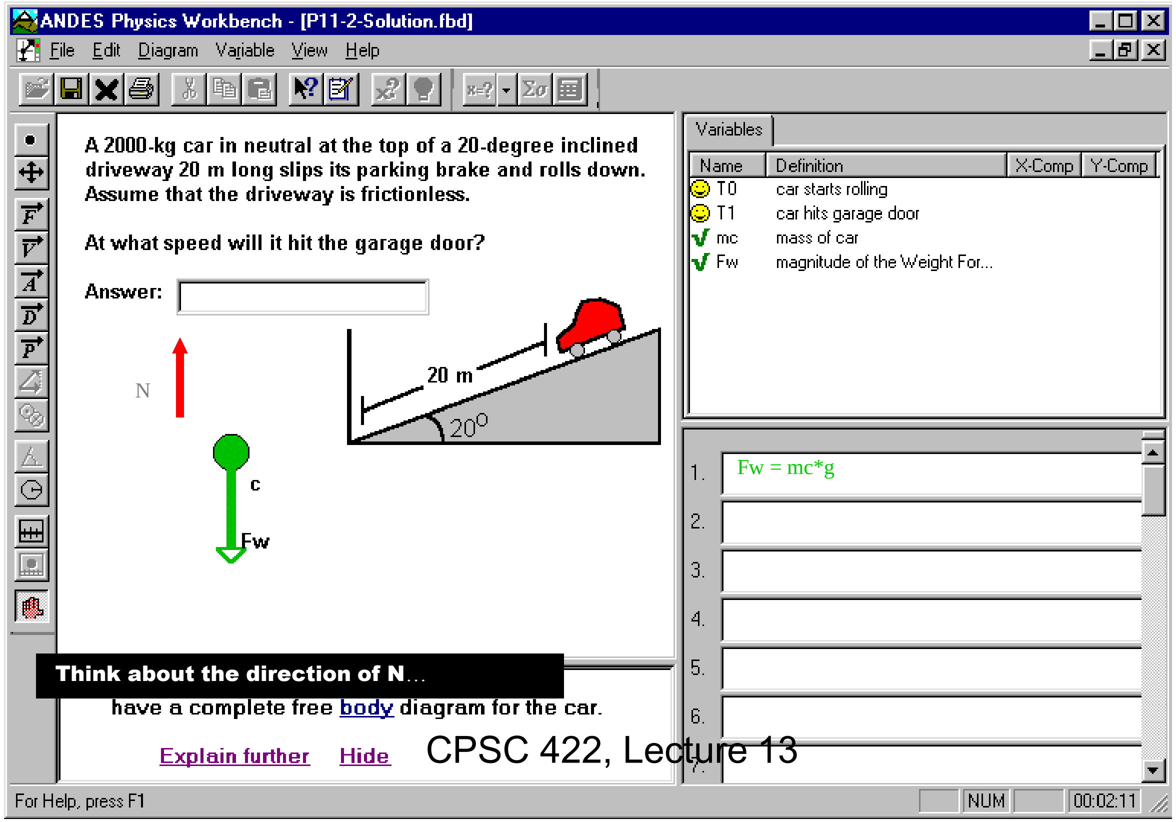
Representing the instructional domain (expert model)

Understanding the student (student model)

Providing adequate help and instruction (tutoring model)

ANDES: an ITS for Coached problem solving

- The tutor monitors the student’s solution and intervenes when the student needs help.
  - Gives feedback on correctness of student solution entries
  - Provides hints when student is stuck



Student Model for Coached Problem Solving

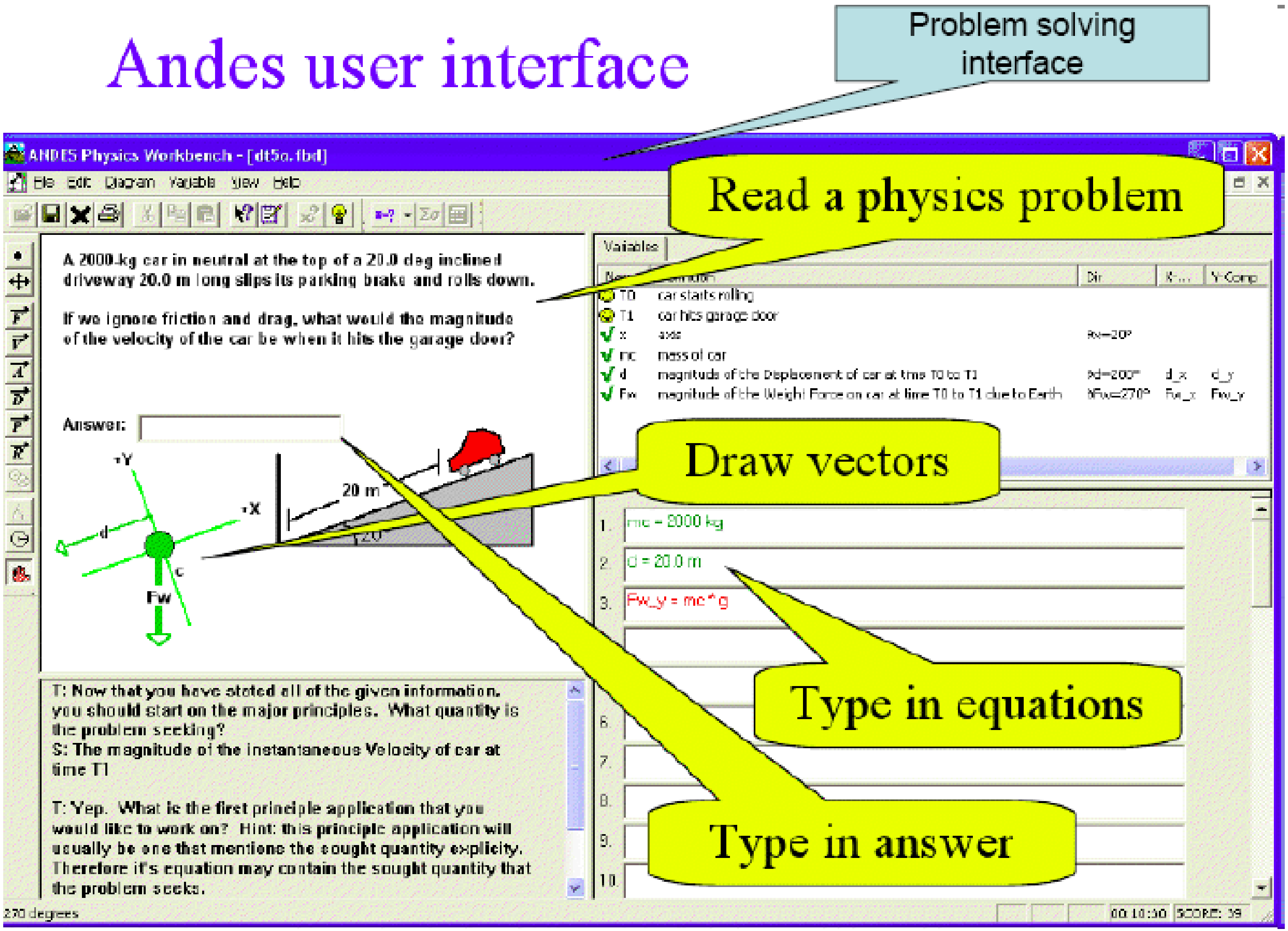
Three main functions

- **Assess from the student's actions her domain knowledge**, to decide which concepts the student needs help on (*knowledge tracing*)

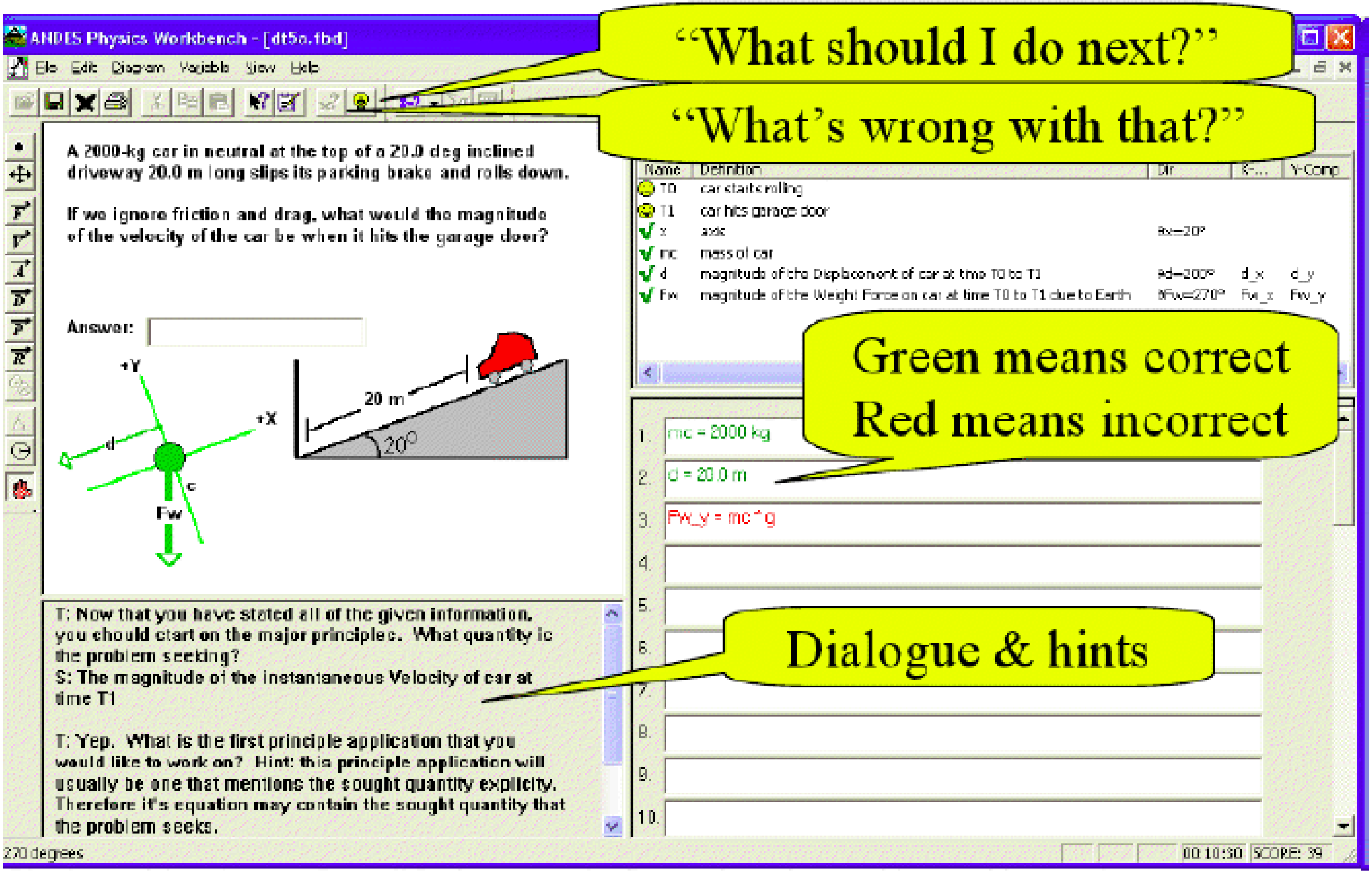
- **Infer from student's actions the solution being followed**, to understand what the student is trying to do (*plan recognition*)

- **Predict what further actions should be suggested to the student**, to provide meaningful suggestions (*adaptive procedural help*)

Andes user interface



Andes feedback and hints

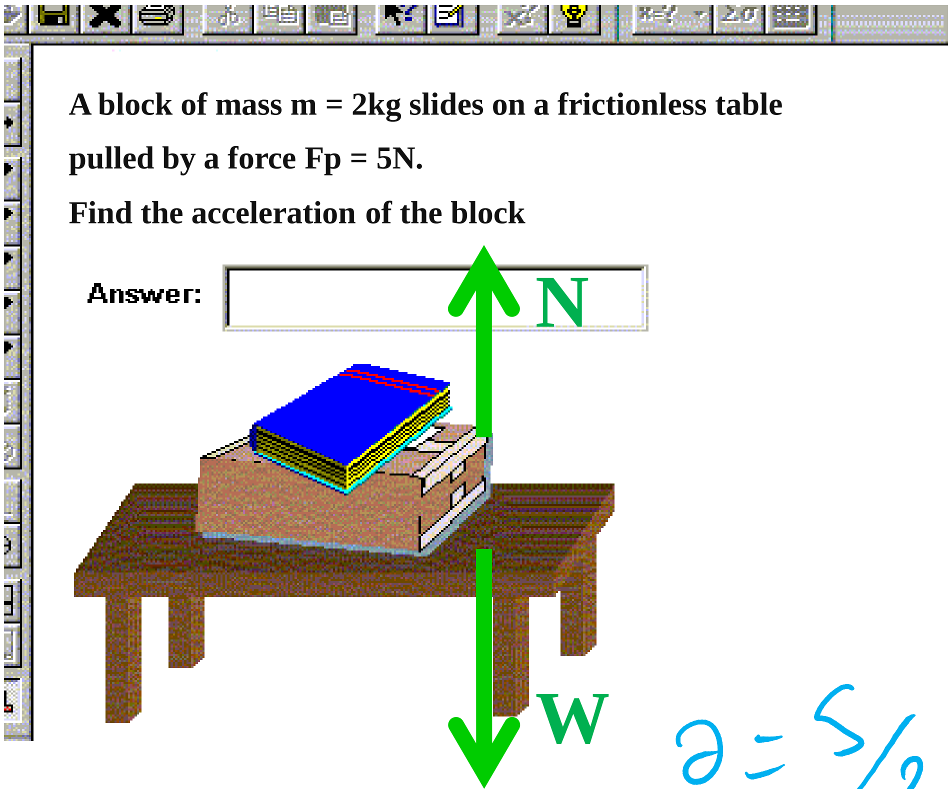


Several sources of uncertainty

- Same action can belong to different solutions
- Often much of the reasoning behind the student's actions is hidden from the tutor
- Correct answers can be achieved through guessing
- Errors can be due to slips
- System's help affects learning
- In many domains, there is flexible solution step order

Andes deals with this uncertainty by using Bayesian Networks

Example 1



a = 5/2

Correct solution

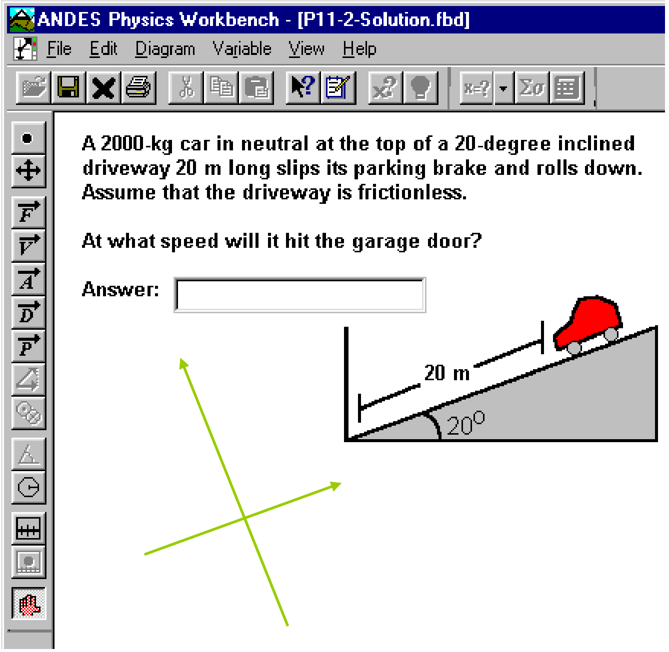
- $\Sigma \mathbf{Ft} = \mathbf{W} + \mathbf{N} + \mathbf{Fp} = m*\mathbf{a}$
- $\Sigma Ft_x = Fp = m*a_x$
- a = 5/2

Incorrect solution

- $\mathbf{Fp} = m*\mathbf{a}$
- a = 5/2

If the student only types a = 5/2 m/sec, what line of reasoning did she follow?

Example 2



Solution

steps

- Find the velocity by applying the kinematics equation

$$Vt_x^2 = V0_x^2 + 2d_x*a_x$$

- Find the acceleration of the car by applying Newton's 2nd law

$$\Sigma F_x = W_x + N_x = m*a_x$$

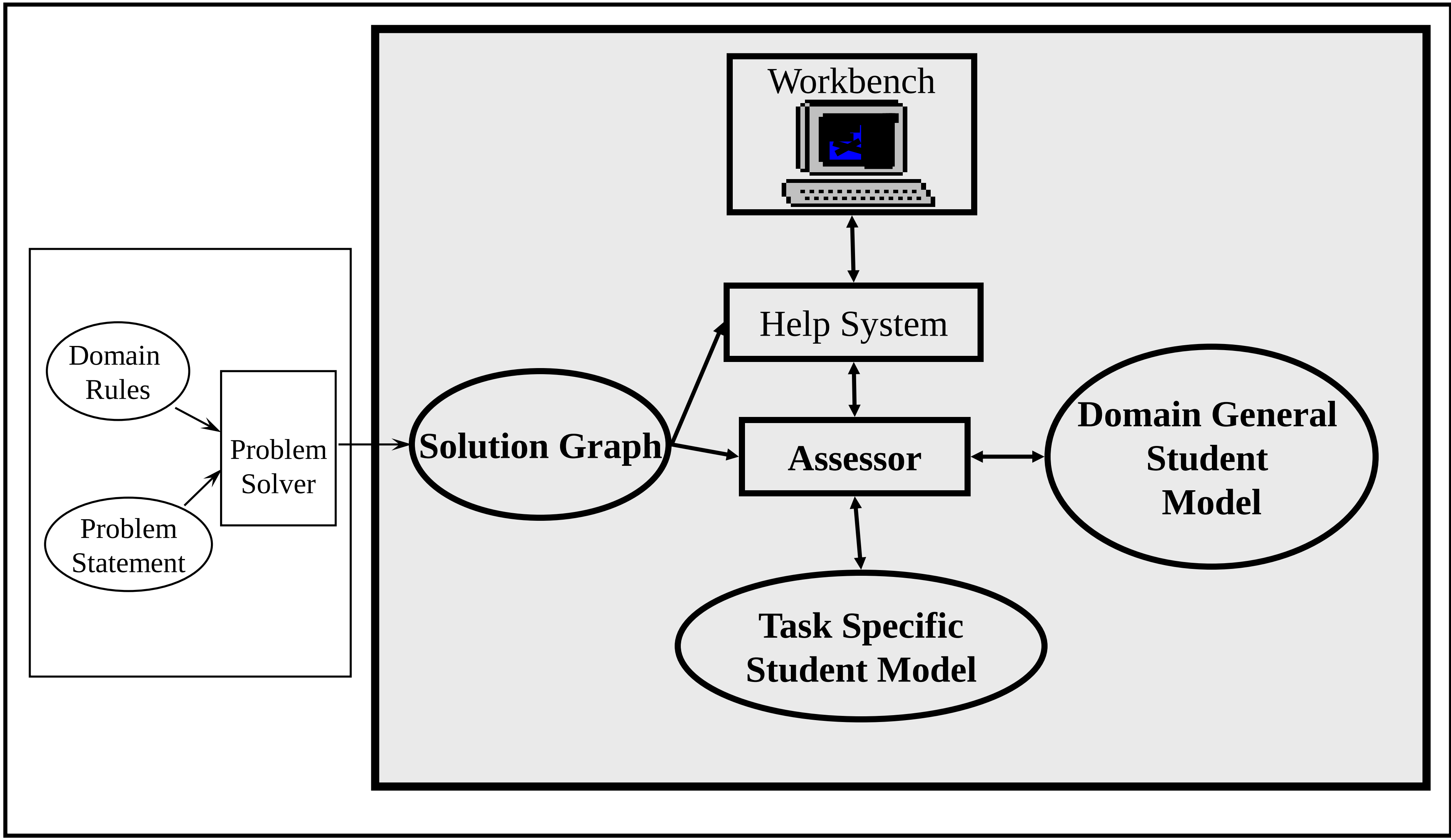
If the student draws the axes and then gets stuck, is she

- trying to write the kinematics equations to find V?
- trying to find the car acceleration by applying Newton's laws?

flexible solution step order

Architecture

Andes

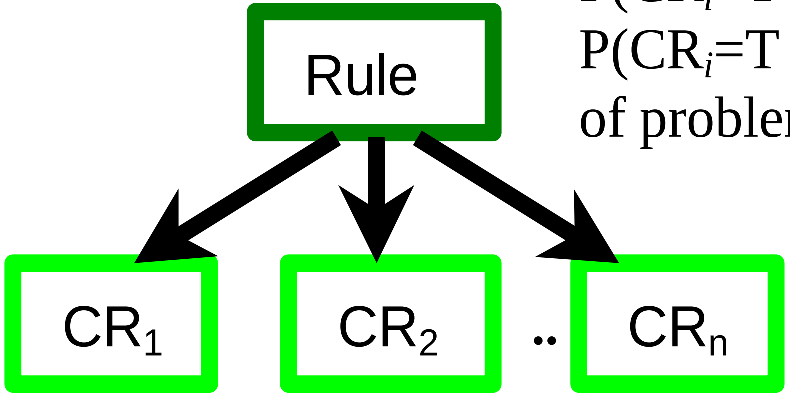


Components of Andes' Student Model

- **Domain General**
  - Reflects the content of Andes' rules
  - Defined once along with Andes' KB
  - Maintained across problems
  - Assesses the student's domain knowledge
- **Task Specific**
  - Automatically built when a new problem is opened
  - Assesses the student's task specific knowledge and problem solving behavior

Domain General Bnet

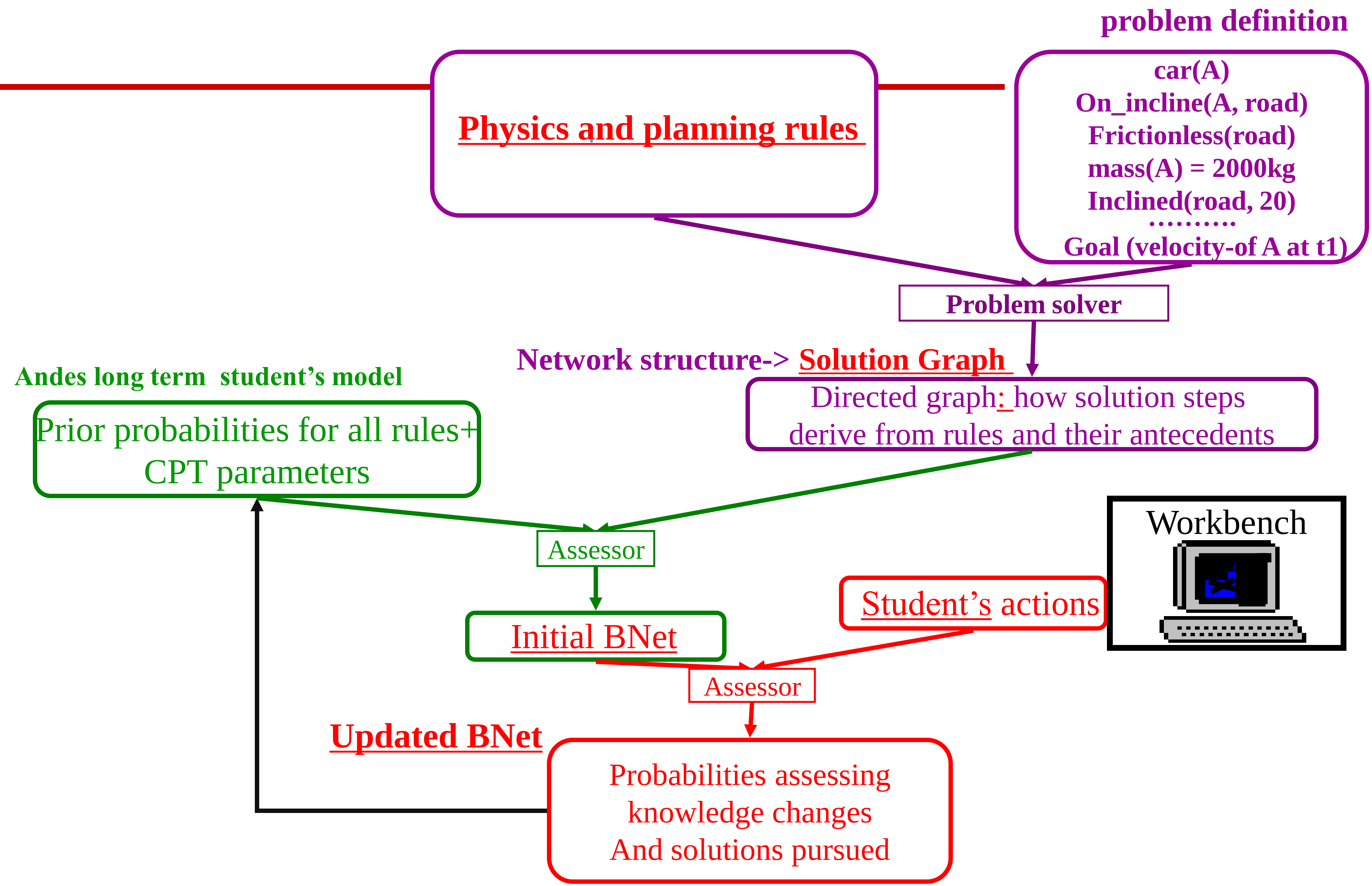
- **Rule nodes**
  - represent knowledge of generic physics and planning rules
  - P(R = T): probability that the student knows the rule (how to apply it in any context)
- **Context rule nodes**
  - Represent rules in specific problem solving contexts
  - P(CR = T): probability that the student can use the rule in the corresponding context



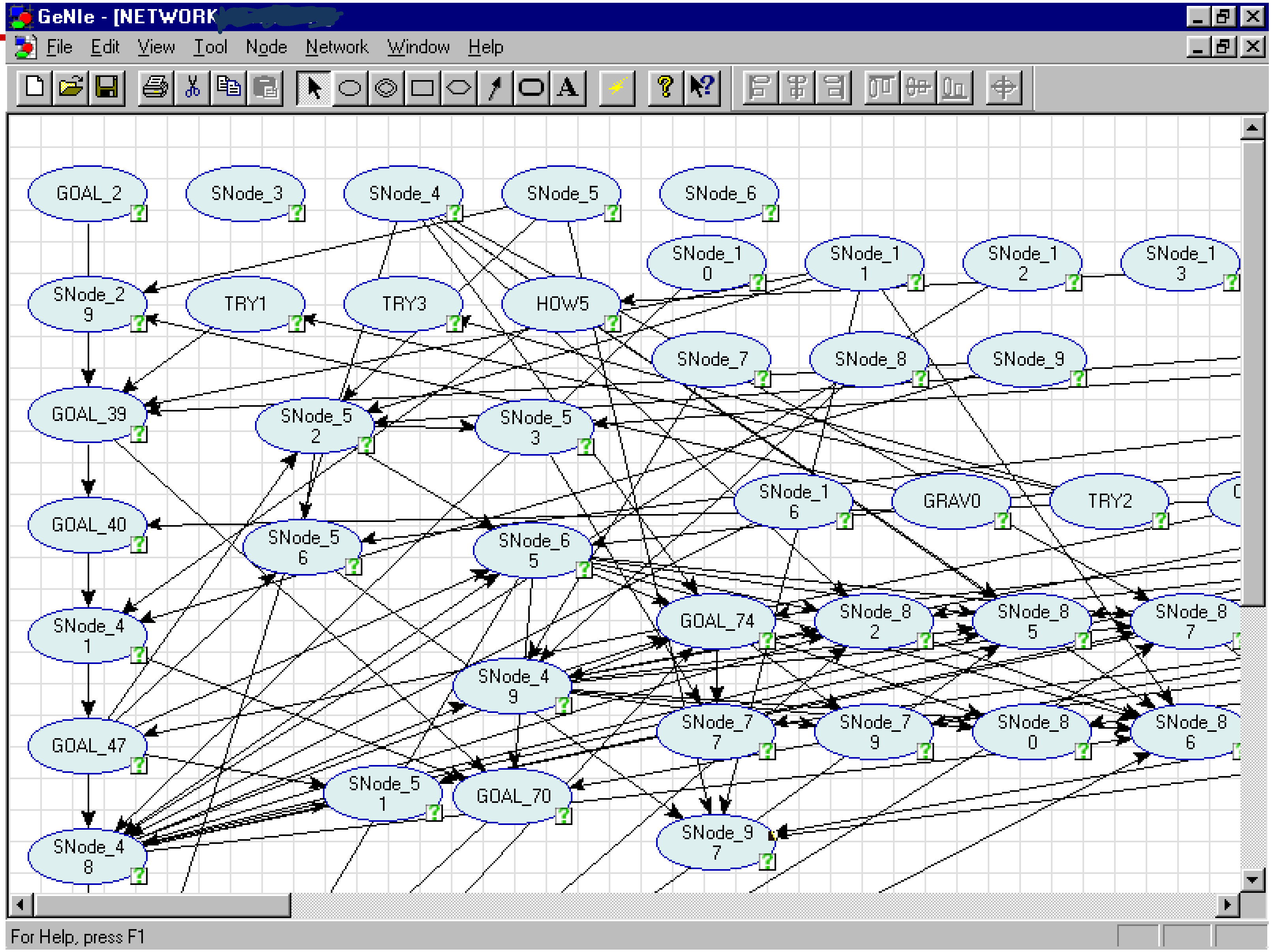
$P(CR_i = T \mid R = T) = 1$

$P(CR_i = T \mid R = F)$ measures the difficulty of problem solving context *i*

Construction of the task specific BNet



Importance of Automatic Generation



Andes rules: encode a solution approach

(a) **R-try-Newton-2law**
If the problem's goal is to find a force
then set the goal to try Newton's second Law to solve the problem

(b) **R-goal-choose-body**
If there is a goal to try Newton's second law to solve a problem
then set the goal to select a body to which to apply the law

Involving goals

(c) **R-body-by-force**
If there is a goal to select a body to apply Newton's second law
and the problem goal is to find a force on an object
then select as body the object to which the force is applied

(d) **R-normal-exists**
If there is a goal to find all forces on a body
and And the body rests on a surface
then there is a Normal Force exerted on the body by the surface

Encoding qualitative physics principles

Andes problem solver generate a solution graph

There is a block A, which has a magnitude of 50kg

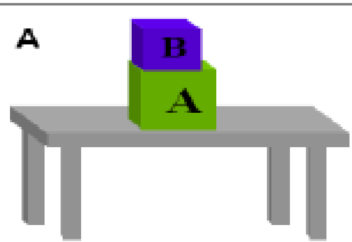
1. Encode the problem to Andes problem solver as:

```
(SCALAR (KIND MASS) (BODY BLOCK-A) (MAGNITUDE 50) (UNITS KG))
```

2. Encode the problem goal to Andes problem solver as

```
(GOAL-PROBLEM (IS FIND-NORMAL-FORCE) (APPLIED-TO BLOCK-A)
(APPLIED-BY TABLE) (TIME 1 2))
```

Find the normal force on block A applied by table

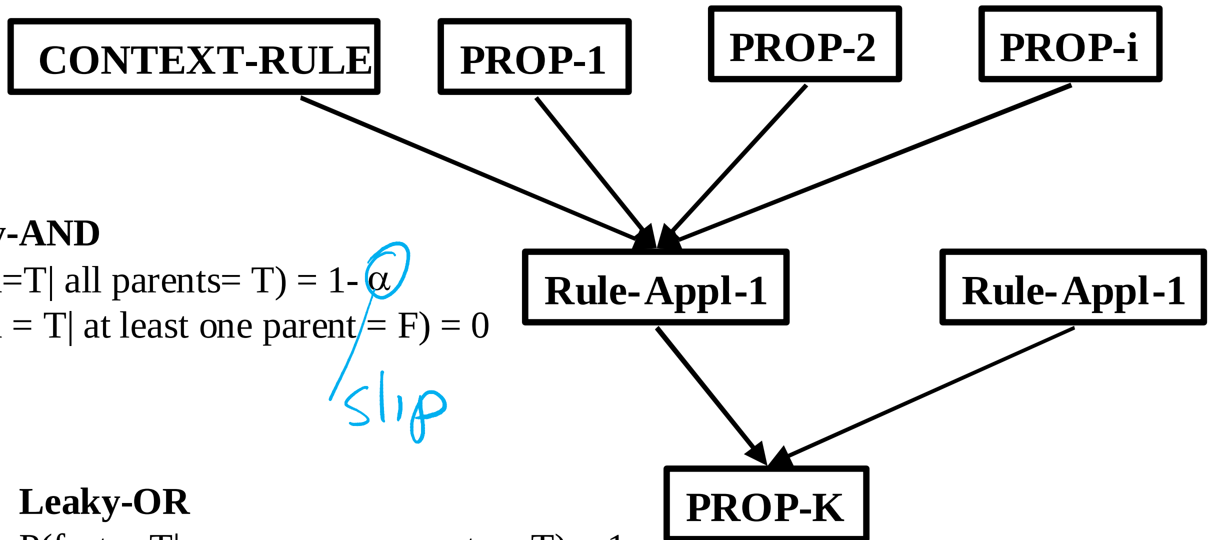


A block (A) of mass 50kg rests on top of a table. Another block (B) of mass 10kg rests on top of block A.
What is the normal force exerted by the table on block A?

3. Find the sub-goals and apply rules until to solve the sought quantity

1. choose a body/bodies to which to apply the law,
2. identify all the forces on the body,
3. write the component equations for $\Sigma F_i = m*\mathbf{a}$.

Conditional Probabilities in the Task Specific BNet



Noisy-AND

P(RA=T| all parents= T) = 1- α

P(RA = T| at least one parent = F) = 0

slip

Leaky-OR

P(fact = T| one or more parents = T) = 1

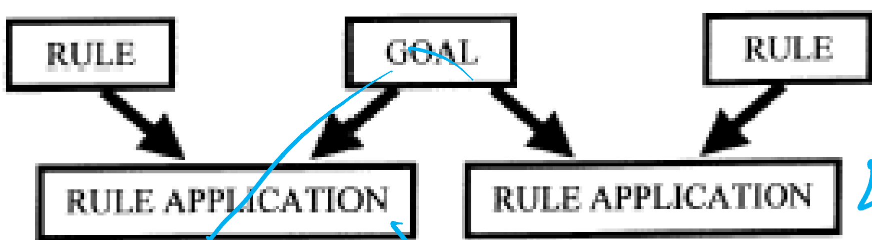
P(fact=T| all parents= F) = β

guessing

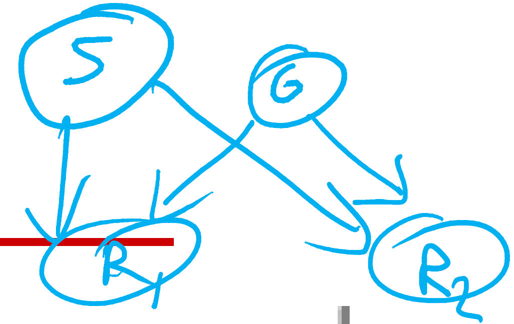
Strategy Nodes

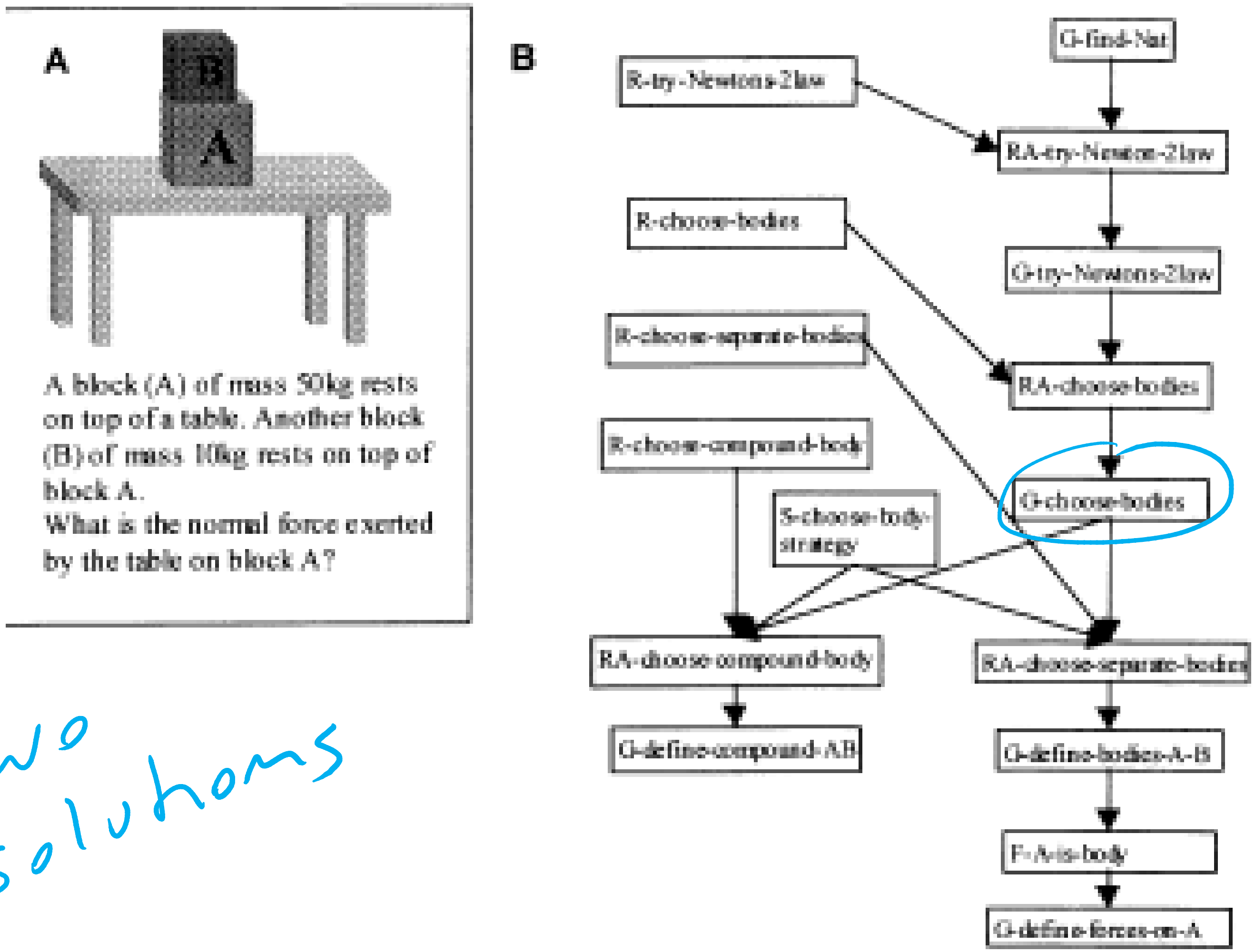
- If a given goal is involved in generating two alternative solutions, evidence that a student is following one solution should decrease the probability of the other solution
- This does not happen with the basic Andes' Bnet. Actually, evidence of a solution would increase the probability of any other alternative solution that shares a goal with it

change structure of BNet to avoid undesired evidence propagation



Example





two solutions

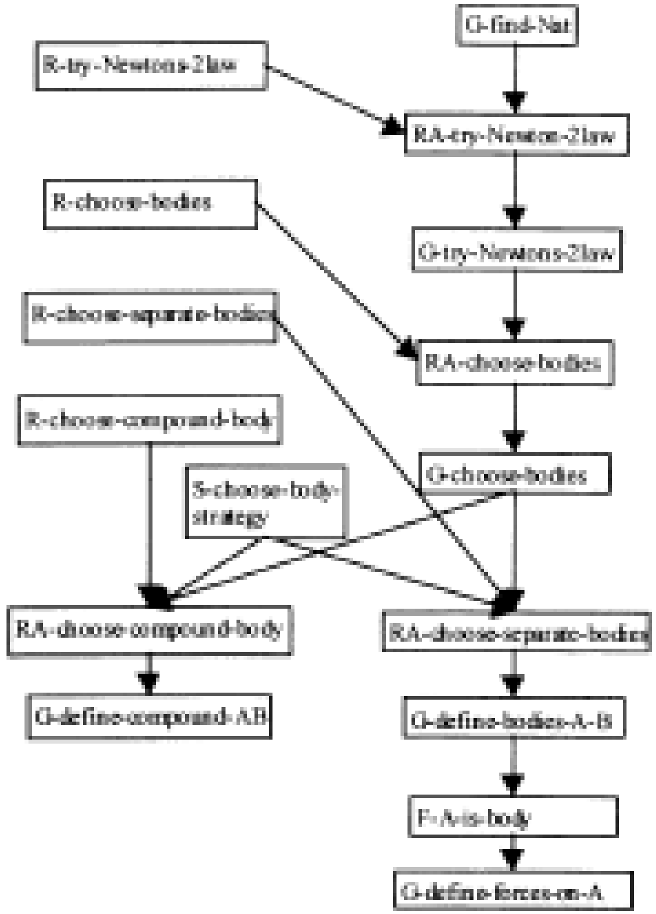
S v_1 v_2

$P(R_1 | S)$

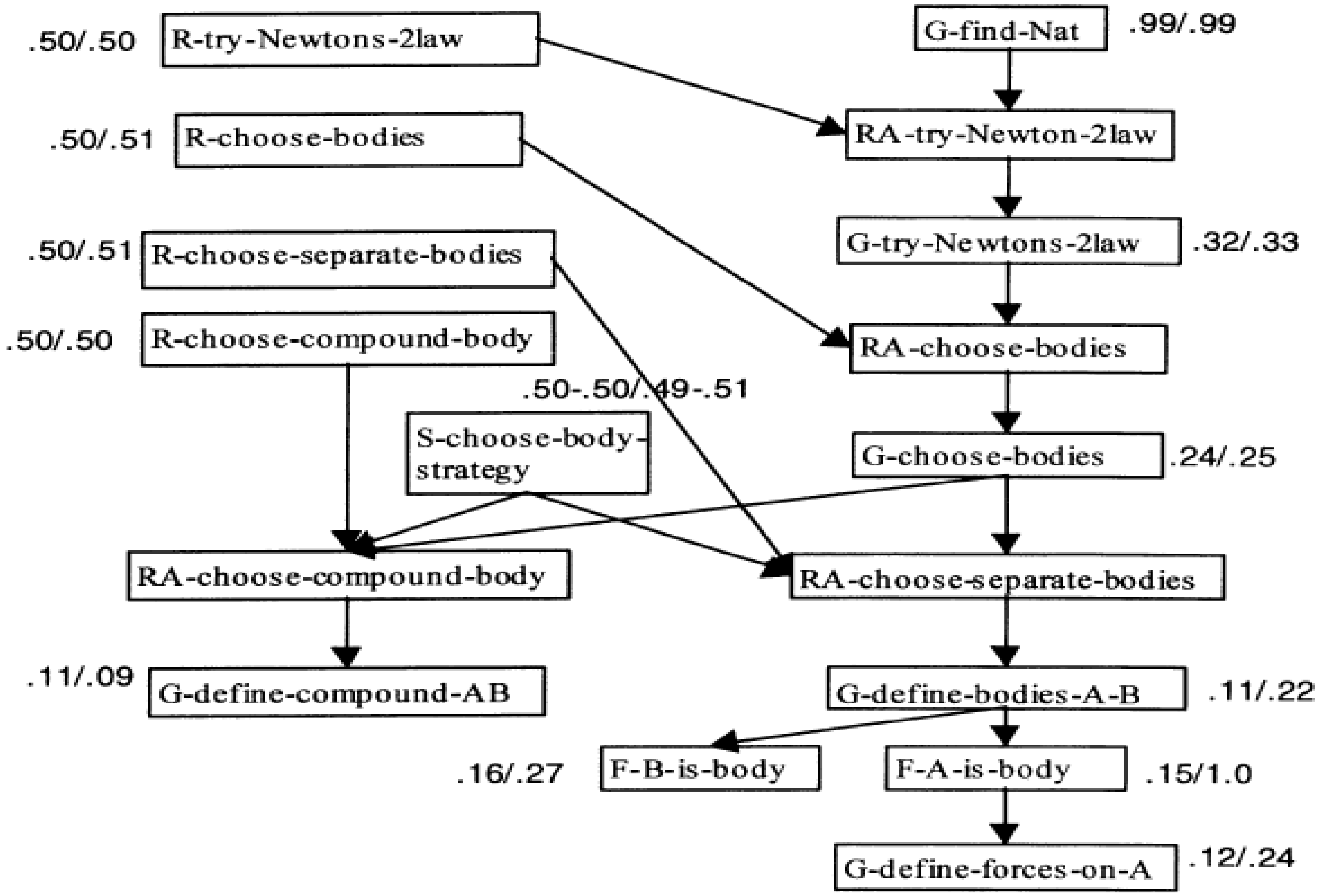
	T	F
v_1	1	0
v_2	0	1

$P(R_2 | S)$

	T	F
v_1	0	1
v_2	1	0


G-find-Nat
R-try-Newtons-2law
RA-try-Newton-2law
R-choose-bodies
G-try-Newtons-2law
R-choose-separate-bodies
RA-choose-bodies
R-choose-compound-body
G-choose-bodies
S-choose-body-strategy
RA-choose-compound-body
RA-choose-separate-bodies
G-define-compound-AB
G-define-bodies-A-B
F-A-is-body
G-define-forces-on-A

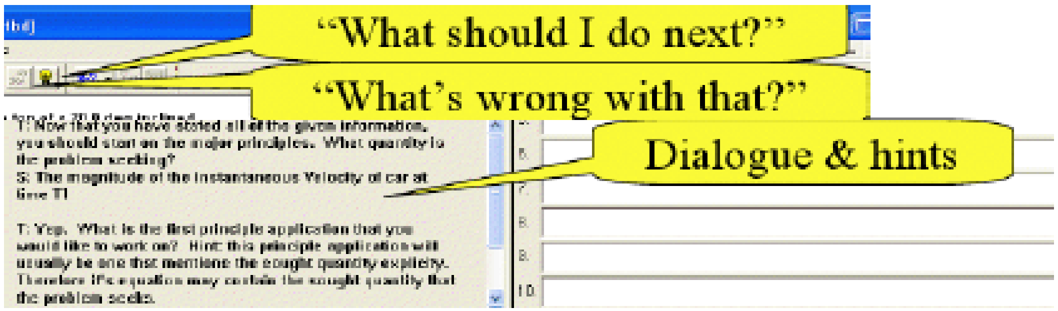
The network before and after observing F-A-is-a-body.

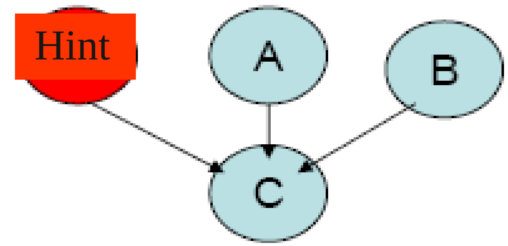


Andes problem solving Coach– handle Errors

- Two type errors:
  - Errors of omission :missing actions
  - errors of commission: disbelieve a certain correct fact or not clear what a correct action is
- Omission errors : rarely clamps nodes to F because Andes does not require explicit actions ordering
- Errors of commission:
  - Implies to disbelieve a certain correct fact, clamps nodes to F
  - otherwise not (more common)

Andes problem solving Coach– handle hint issue



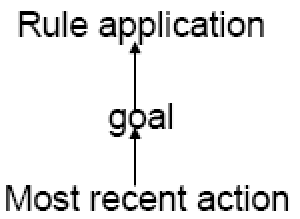


- Andes hints affect actions not domain knowledge
  - Hints just reminds knowledge not teaches it;
  - Hints increases the chance of guessing the next action.
- Hint node is added as a parent node of the proposition node

Andes problem solving Coach– generate help (plan recognition)

Use Bayesian network to figure out what is the goal trying to achieve (plan recognition) and where to get stuck (action prediction) before generate help.

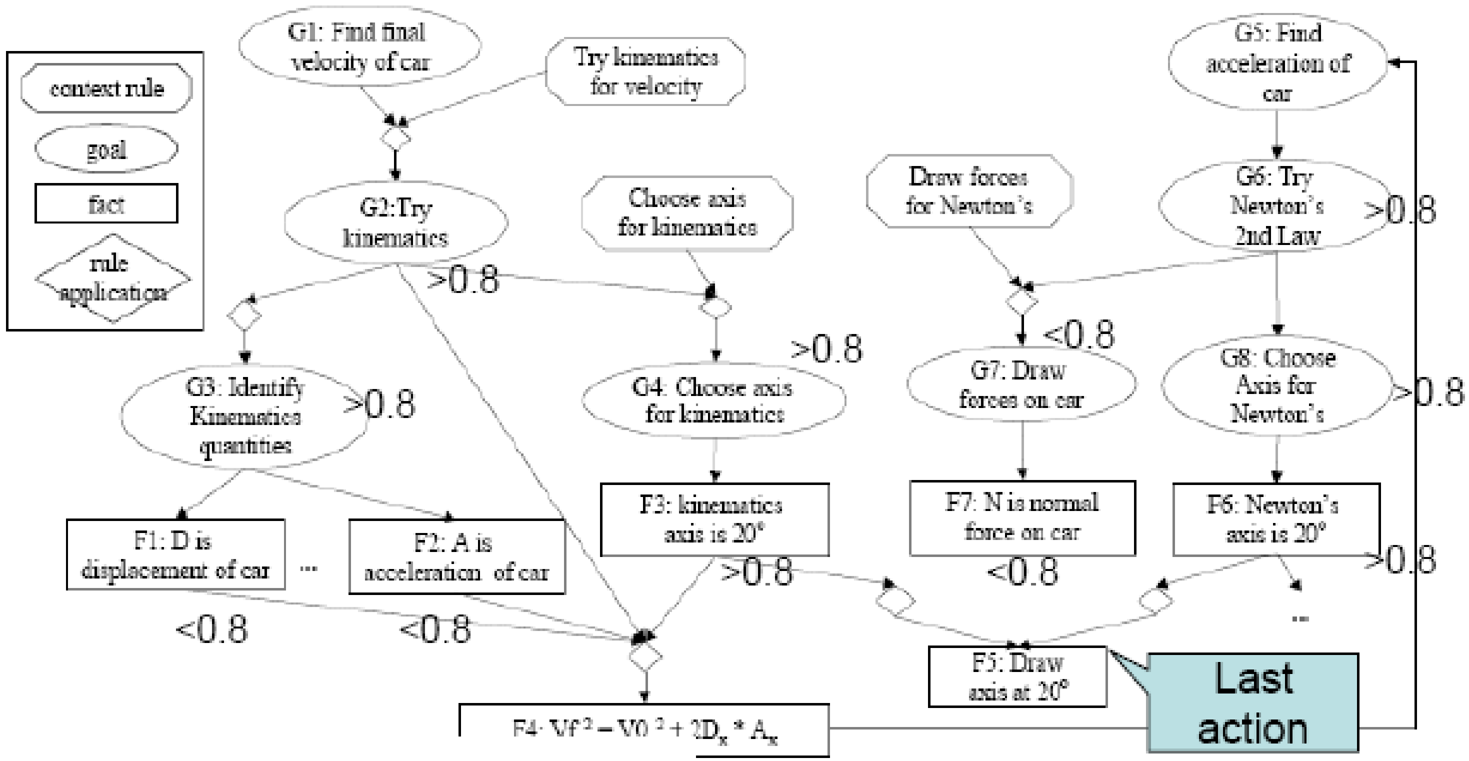
How ?

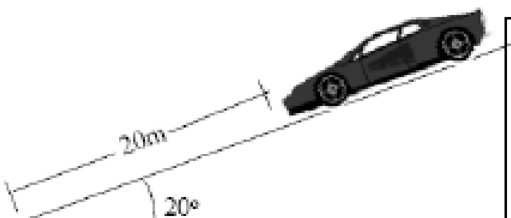


Search solution graph in depth-first way.
The depth-first traversal stops when it reaches a low probability below threshhold(0.8)

The result of this traversal is a set of paths through the solution graph beginning with the most recent action, and stopping with a node whose probability is below 0.8

Reference: procedural help in Andes: generating hints using Bayesian network student model





F5-F3-G4-G2-G3-F1

F5-F3-G4-G2-G3-F2

F5-F6-G8-G6-G7

joint probability computing for each path

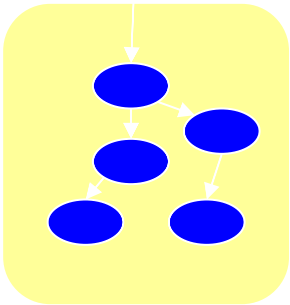
Choose the highest value path

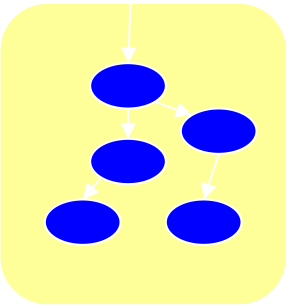
A 2000kg car at the top of a 20° inclined driveway 20m long slips its parking brake and rolls down. Assume that the driveway is frictionless. At what speed will it hit the garage door?

Reference: procedural help in Andes: generating hints using Bayesian network student model

Andes Dynamic Bayesian Network

From one problem to another





Evaluation

- Andes tutor for physics is currently in use at the US Naval Academy
- Informal studies have shown positive effect on learning
- Continuously updated through students' feedback

Outline

- ILE, background.
- Probabilistic student modeling for coached problem solving.
- Probabilistic student modeling to support learning from examples.

ILE - a step beyond

- Most ILE targets problem solving and domain specific knowledge
- Andes' SE-Coach - a framework to
  - support learning from examples
  - coach self-explanation(SE)
    - » generate explanations to oneself to clarify an example solution

Sample physics example

Problem Statement

EXAMPLE 1: Boy rescued by a helicopter

Jake, an 80Kg undergrad, is rescued from a burning building by a helicopter.

He hangs at the end of a rope dangling beneath the helicopter.

If the helicopter accelerates,

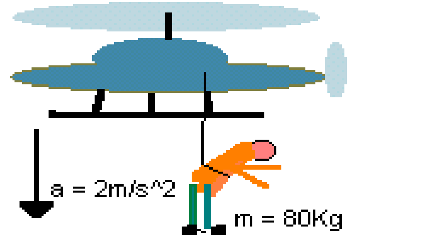
straight downward with respect to the ground,

with an acceleration a = 2m/s^2,

FIND:

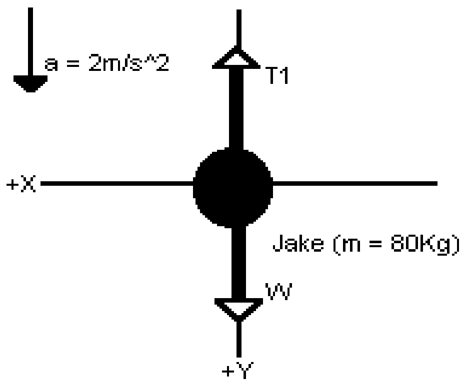
The tension T exerted by the rope.

Situation Diagram



Free Body Diagram

FREE BODY DIAGRAM:



Worked out solution

SOLUTION

Because we want to find a force, we apply Newton's 2nd law to solve this problem.

We choose Jake as the body to which to apply Newton's 2nd law.

The helicopter's rope exerts a tension force T on Jake.

The tension force T is directed upwards.

The other force acting on Jake is his weight W.

The weight W is directed downwards.

To apply Newton's 2nd law to Jake, we choose a coordinate system with the Y axis directed downward.

The Y component of Jake's weight W is

W_y = W.

The Y component of the tension T on Jake is

T_y = -T.

The net force acting on Jake along the Y axis is

Net-force_y = W_y + T_y.

Therefore, substituting

W_y = W, and T_y = -T

into the net force equation, we obtain

Net-force_y = W - T.

If we apply Newton's 2nd Law to Jake, along the Y axis, we obtain:

Net-force_y = m*a_y

The Y component of Jake's acceleration a is

a_y = a.

Therefore, if we substitute a_y and

Net_force_y = W - T

into

Net_force-y = m*a_y

we obtain:

W - T = m*a = (80*2) Newtons.

Solving the preceding equation for T gives:

Why examples and self-explanation?

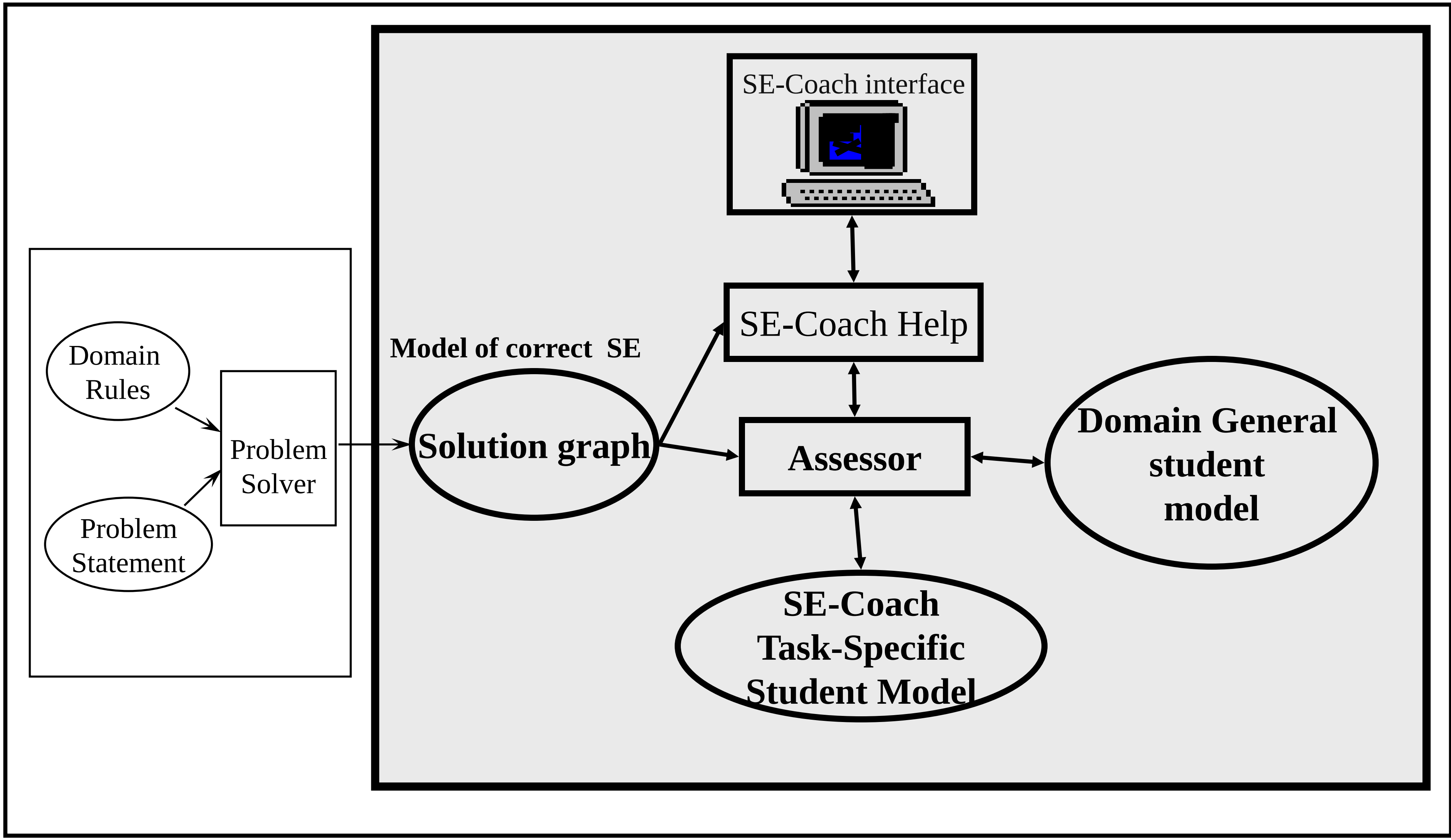
- Students who self-explain learn more
- Many students do not self-explain
  - Fail to detect their lack of understanding
  - Unable to use knowledge to self-explain
- Human tutors can guide self-explanation

SE-Coach: individualized support to SE

- Monitor students as they study examples
- Guide self-explanation to improve students' understanding
- Challenge: only prompt self-explanations that improve students' understanding

SE-Coach Architecture

Andes

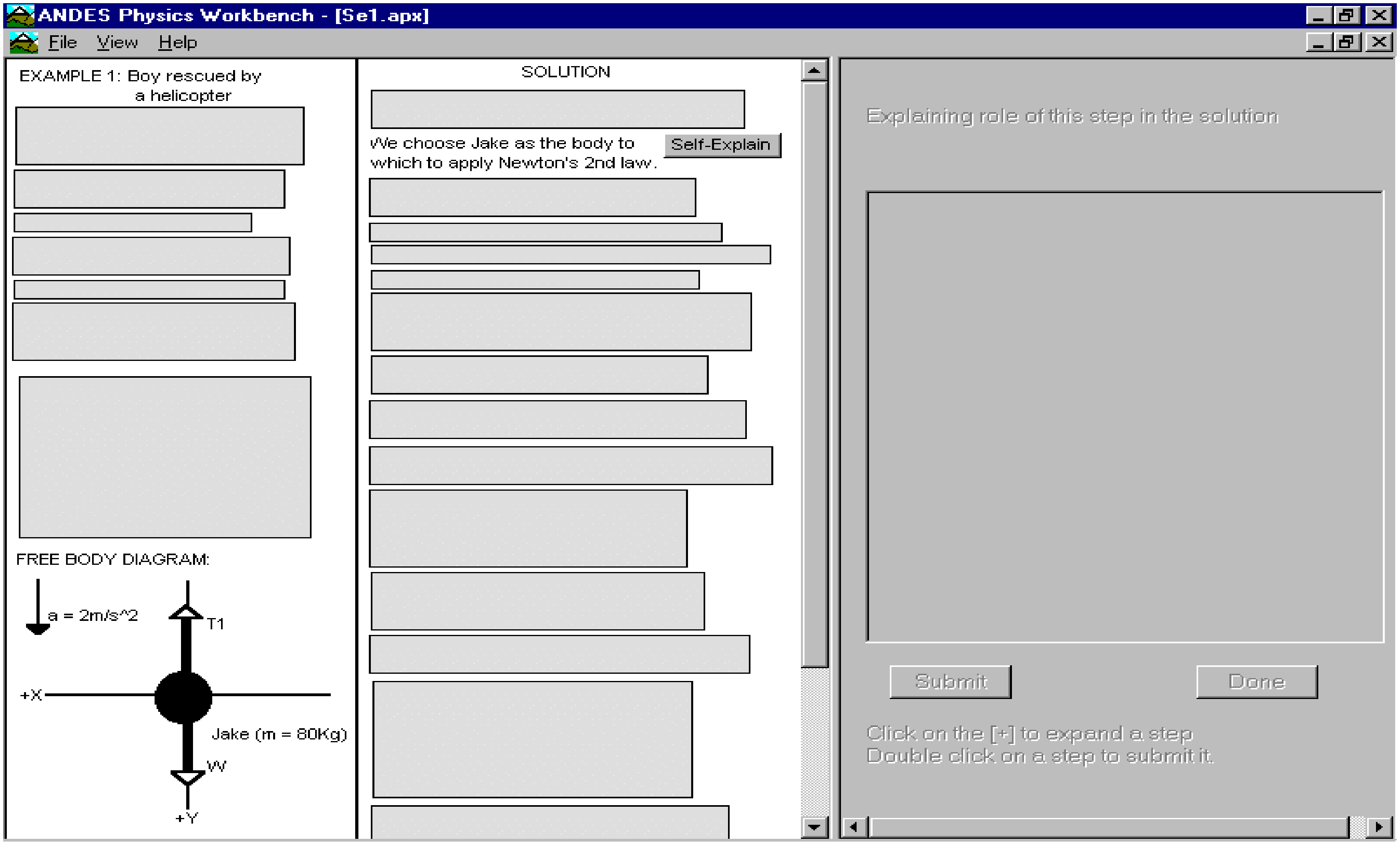


The SE-Coach Workbench

- Masking interface
  - Helps students focus attention and SE-Coach monitor it
- Prompts for relevant self-explanations
  - Justify solution steps in terms of domain principles
  - Explain role of a step in the underlying solution plan
- Menu based tools to generate self-explanations

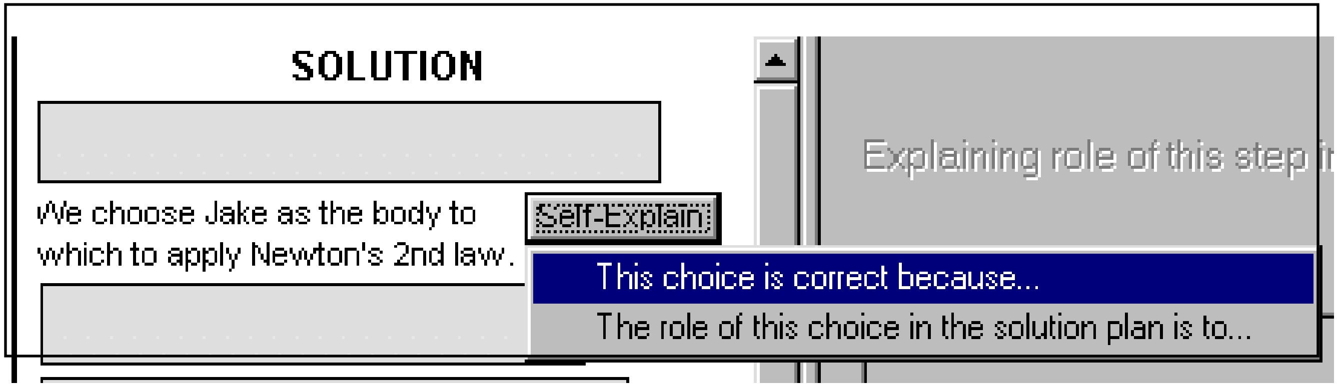
The Workbench - Masking Interface

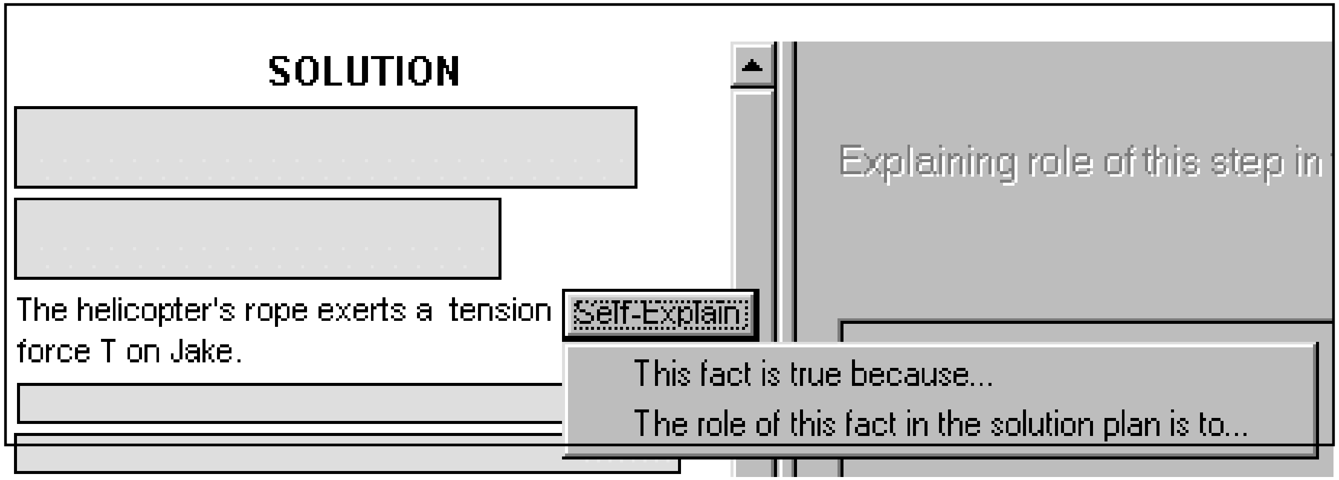
- Helps students focus attention and SE-Coach monitor it



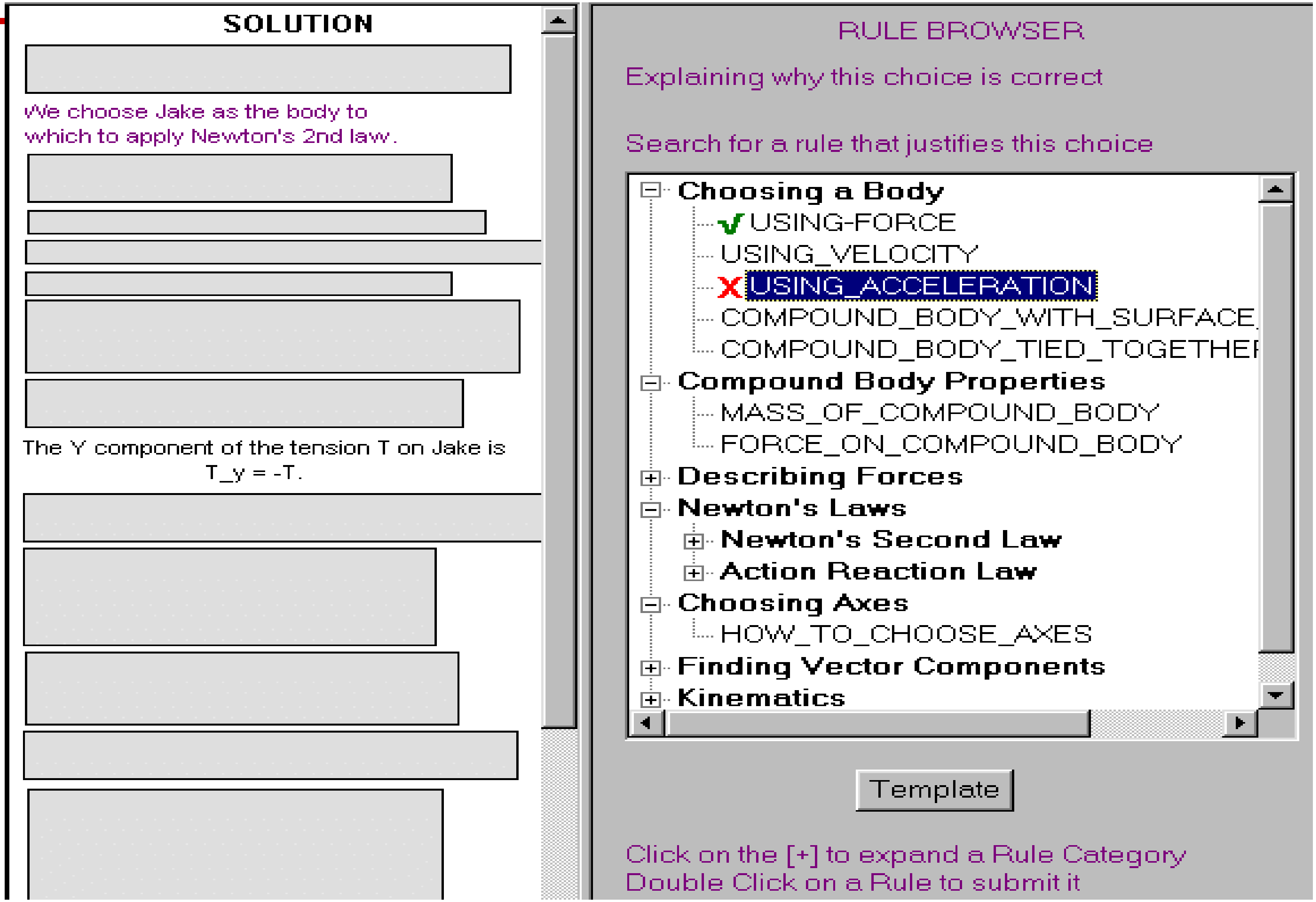
Prompts to Self-Explain

- Stimulate self-questioning on relevant explanations



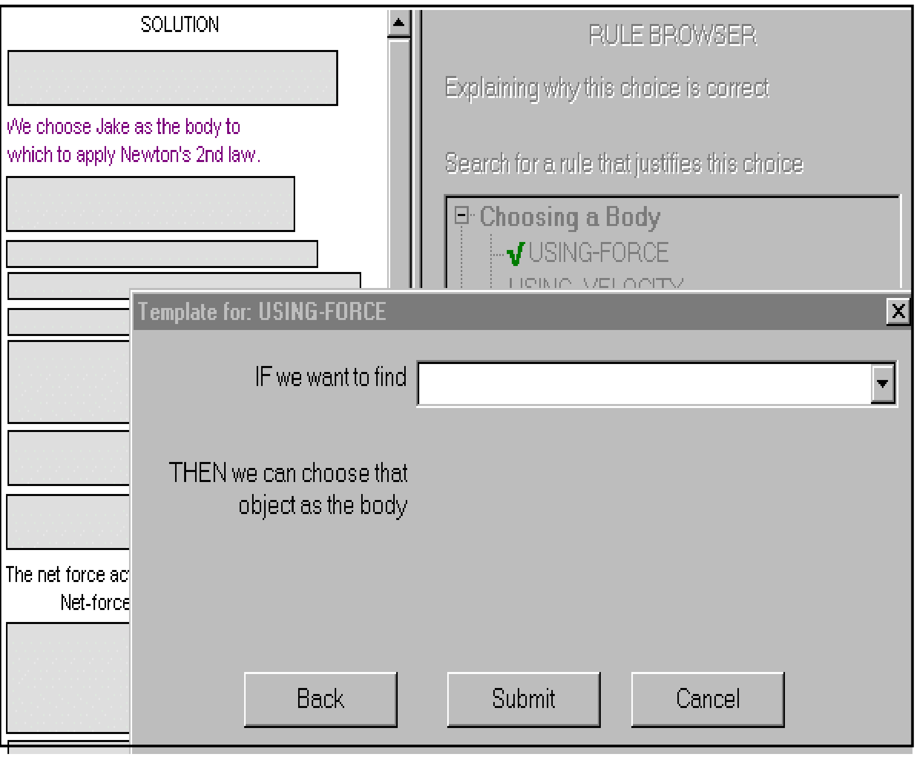


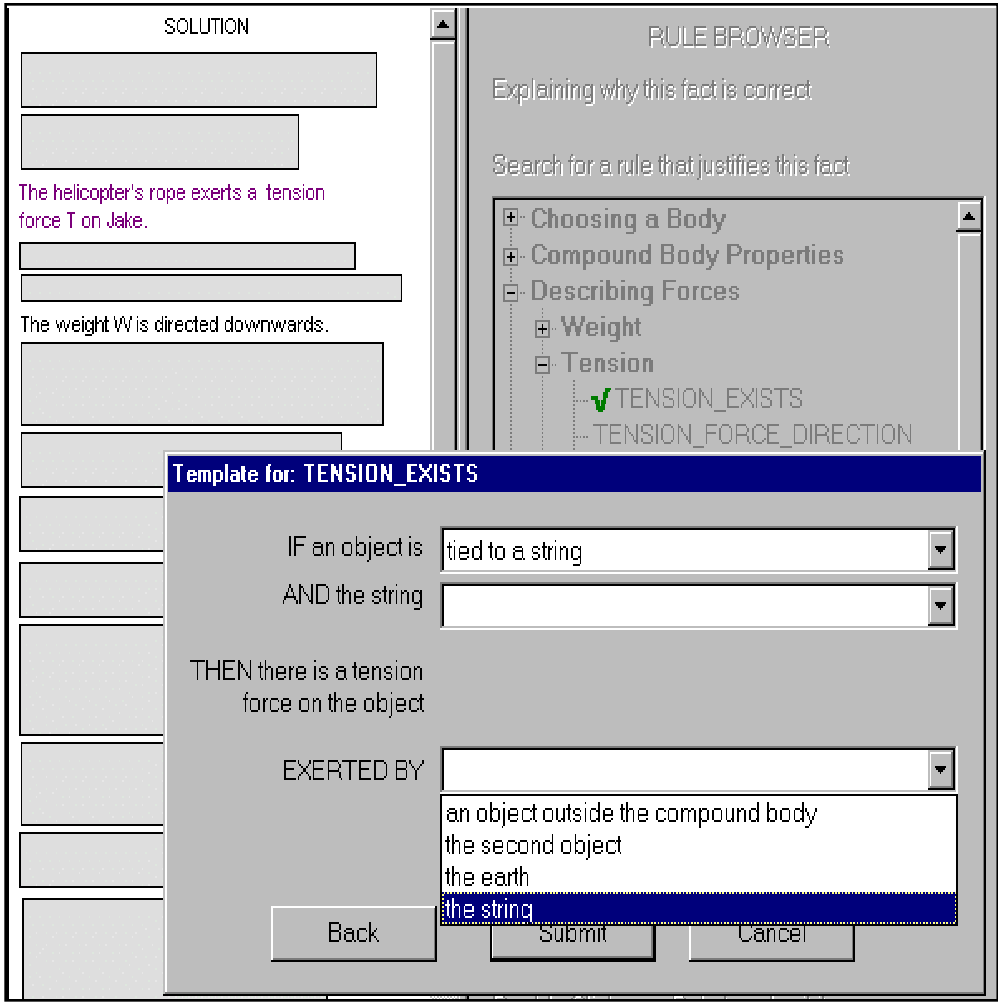
Justify Solution Steps: Rule Browser



Justify Solution Steps: Rule Templates

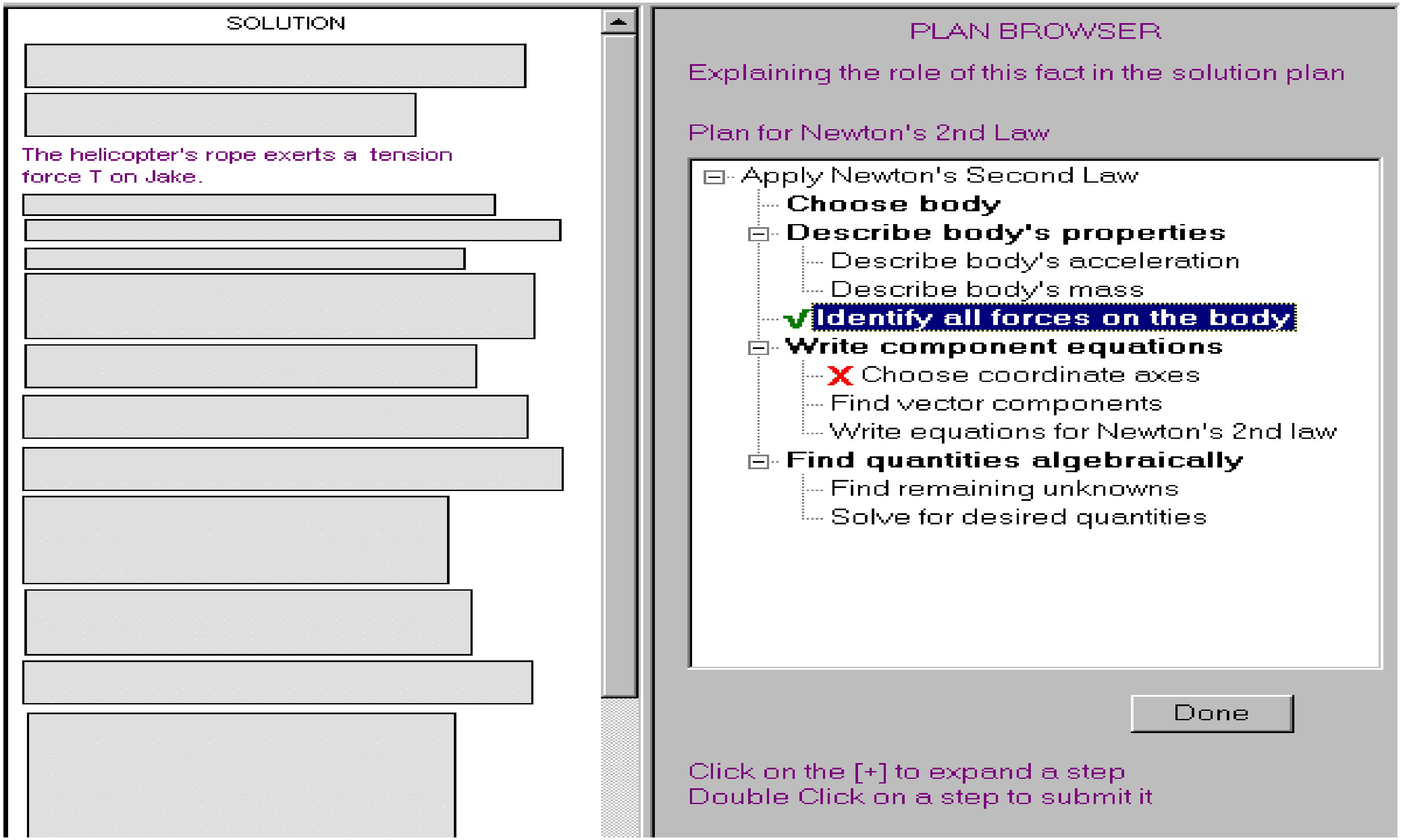
- Help students generate principle definitions



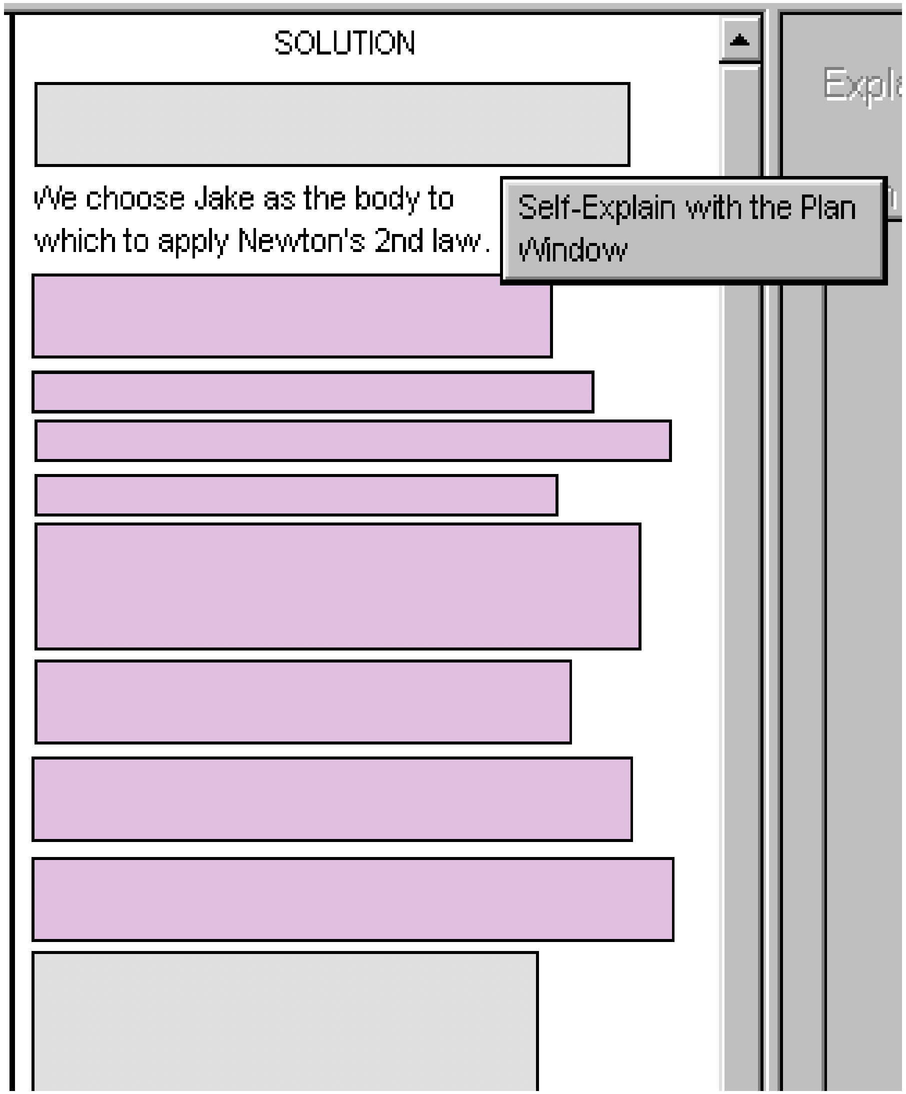


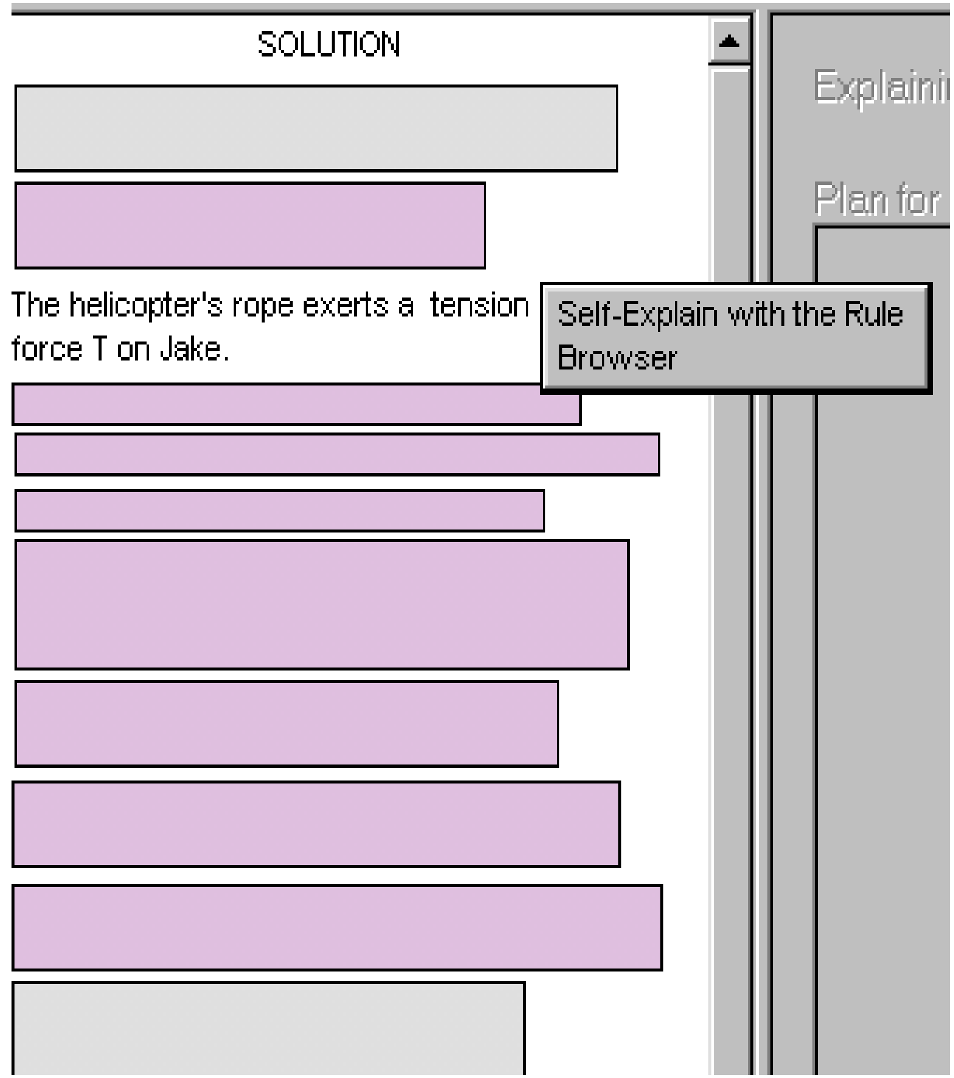
Identify Goal Structure - Plan Browser

- Encodes abstract solution plan



SE-Coach Hints



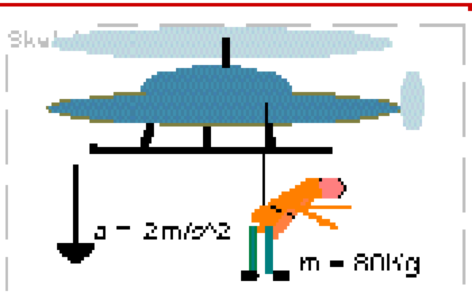


Probabilistic Student Model

Based on a Bayesian network to deal with various sources of uncertainty involved in the modeling task

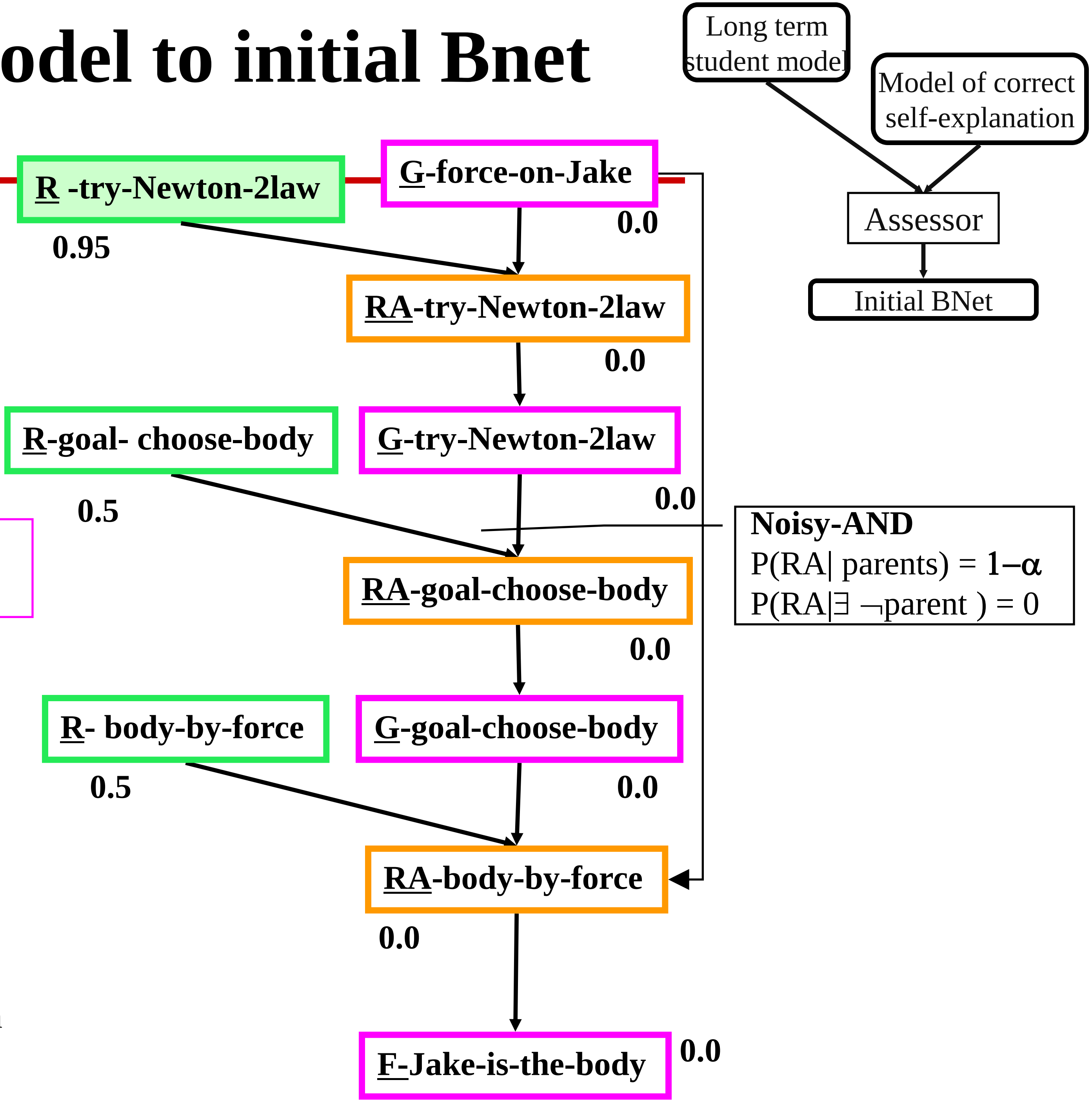
- Detecting spontaneous self-explanation from
  - Reading time
  - Student's knowledge
- Some students study examples by reasoning forward.
- Assessing learning from using the interface menu-based tools

From SE model to initial Bnet



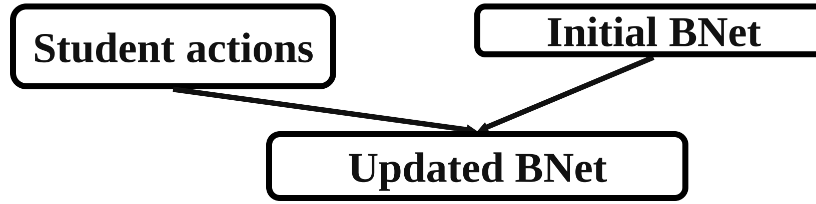
Find the force exerted on Jake by the rope.

To solve this problem, we choose Jake as the body.

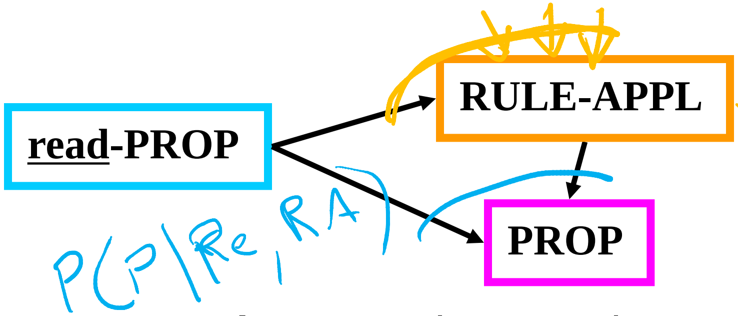


- **R** Rule
- **F/G** Fact/Goal
- **RA** Rule Application

Student Actions



- Read nodes: duration of attention to example parts



Knows RULE	Knows Goal/Fact	READ LOW	READ OK	READ LONG
T	T	$p_1 = 1- \alpha$	$p_2 = 1- \alpha$	$p_3 > \max\{p_2, 0.9\}$
otherwise		0	0	0

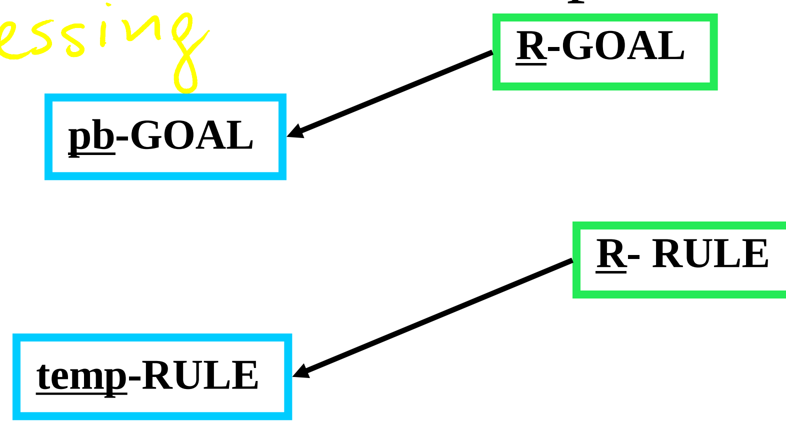
READ		LOW	OK	LONG
RULE-APPL	T	1.0	1.0	1.0
	F	$p_1 < 0.5$	$p_2 > 0.9$	$p_3 > p_2$

← it does not matter if I read or not

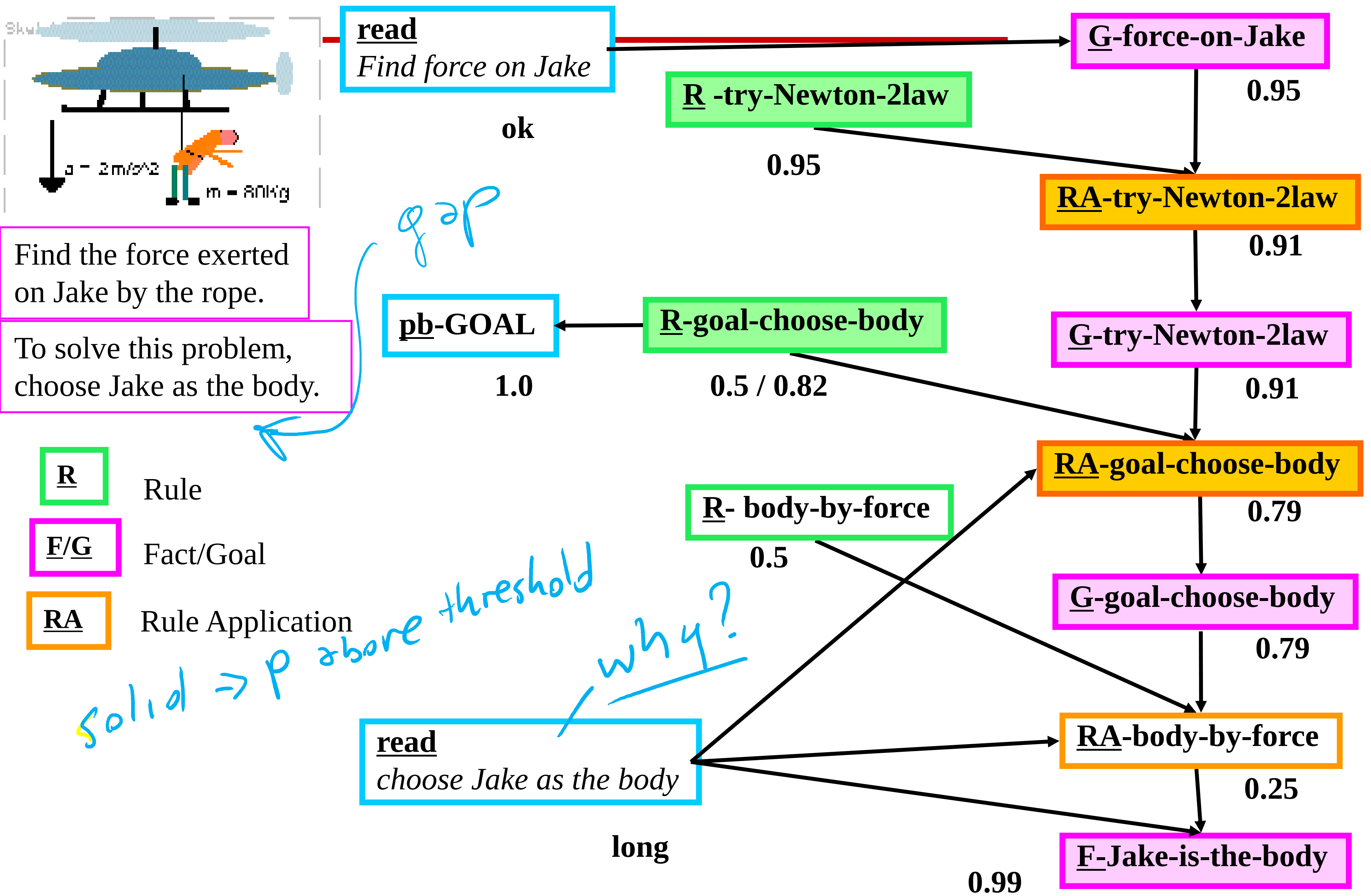
- SE nodes: actions with Plan Browser and Templates

RULE		*T*	*F*
SE Node	*T*	0.95	0.2
	F	0.05	0.8

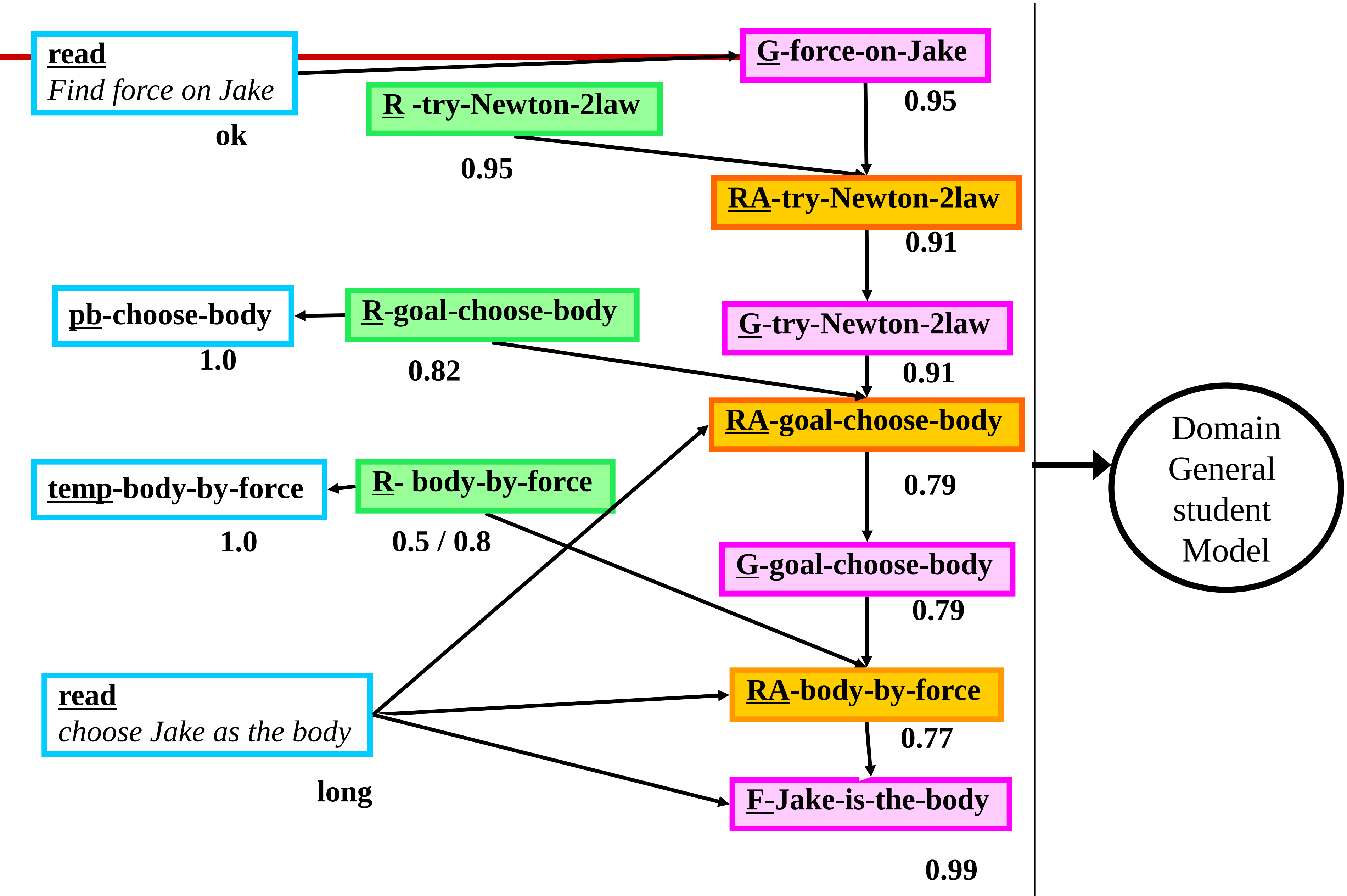
guessing



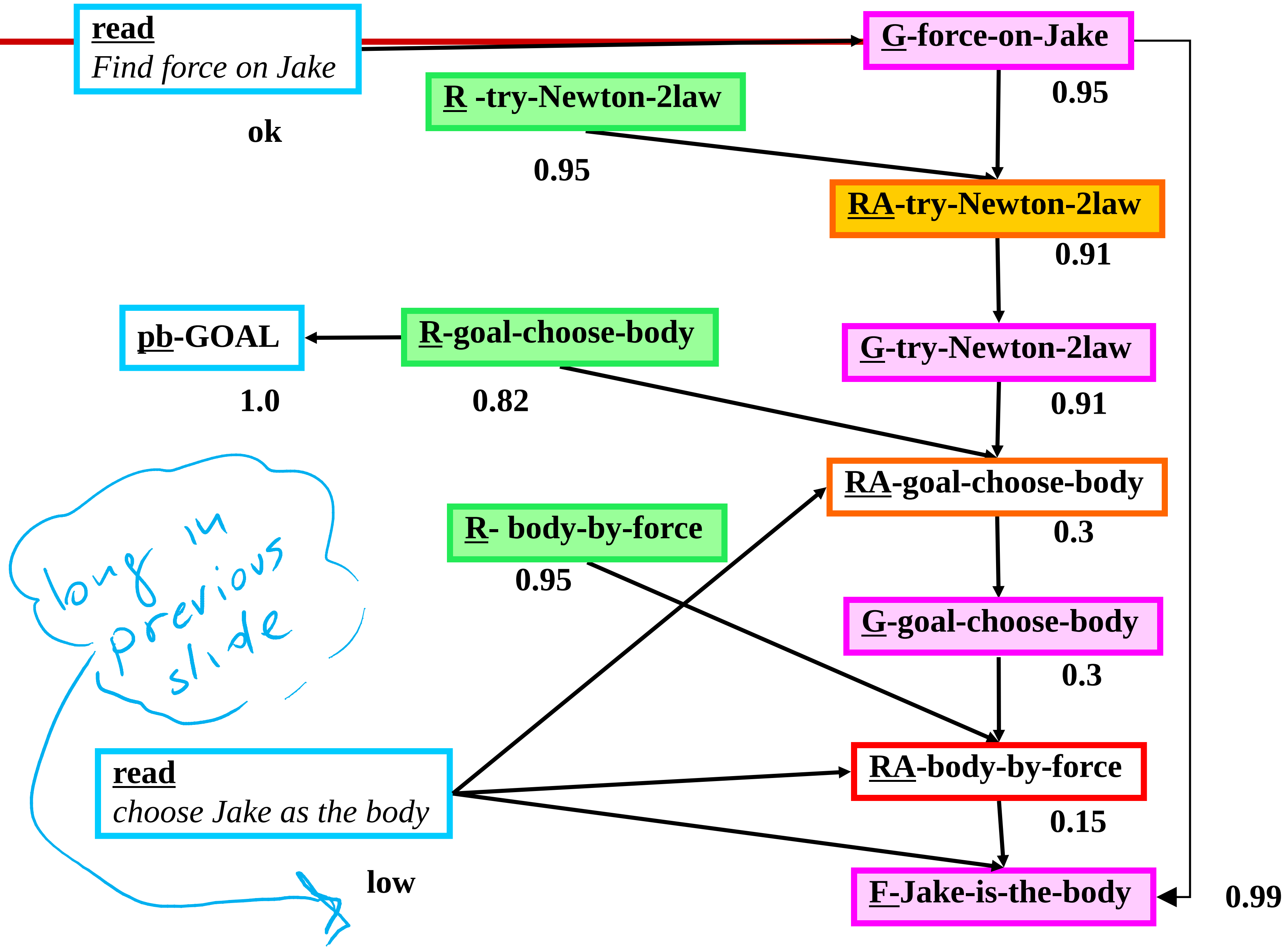
After Reading and Plan Browser Selection



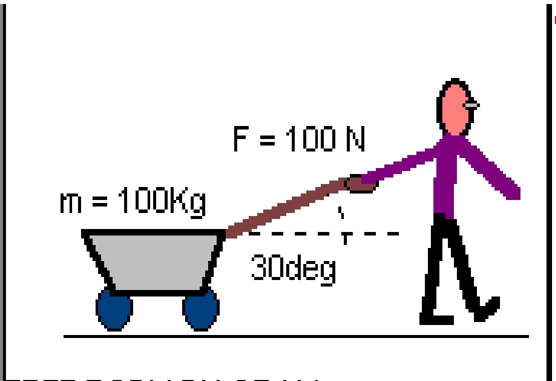
After Filling Template and Closing



After Reading and Plan Browser Selection

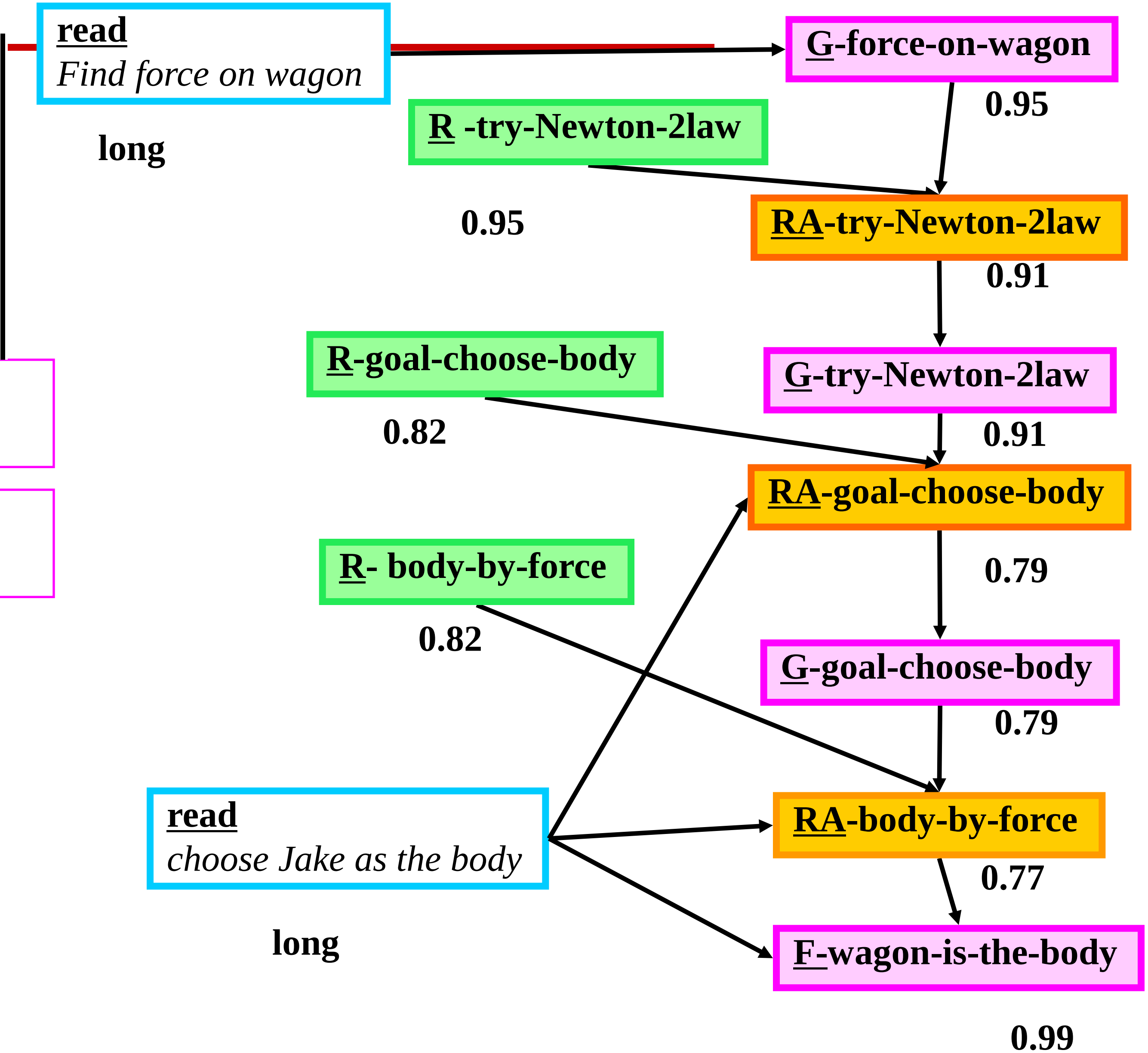


Transfer to a new example



Find the force N exerted on the wagon by the ground.

We choose the wagon as the body.



Empirical Evaluation

- Subjects - 56 students taking Introductory Physics
- Pretest - 4 problems on Newton's second law
- Treatment
  - Experimental (29): studied examples with complete SE-Coach
  - Control (27): studied examples with Masking interface and Plan Window, no feedback nor coaching
- Posttest - 4 problems analogous to pretest

Evaluation of the SE-Coach

- Interface easy to use and generally successful at stimulating SE.
- Overall effectiveness seems to depend on learning stage
  - The SE-Coach was more effective for the subjects that had just started learning the examples topic (late-start subjects).
- Student model: guides interventions that positively correlate with learning ($p < 0.05$)

Prompt Type	Max.	Generated	Followed
Use Rule Browser/Templ.	**43**	**22.6**	**38.6%**
Use Plan Browser	**34**	**22.4**	**42%**
Read More Carefully	**43**	**7**	**34%**

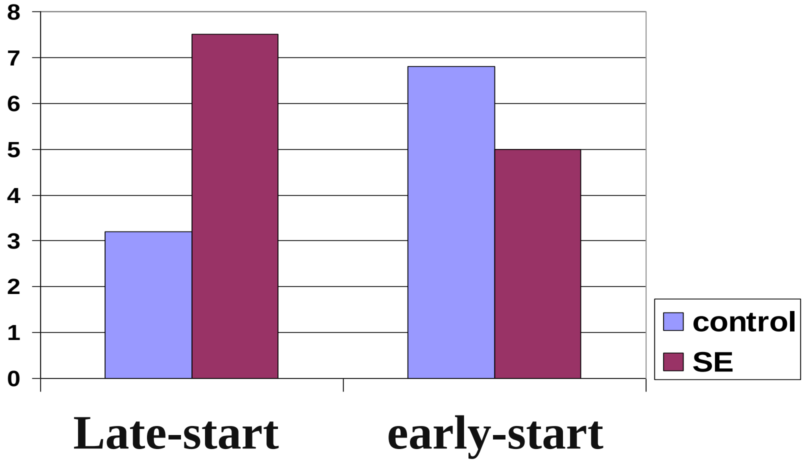
Results: Hints to self-explain

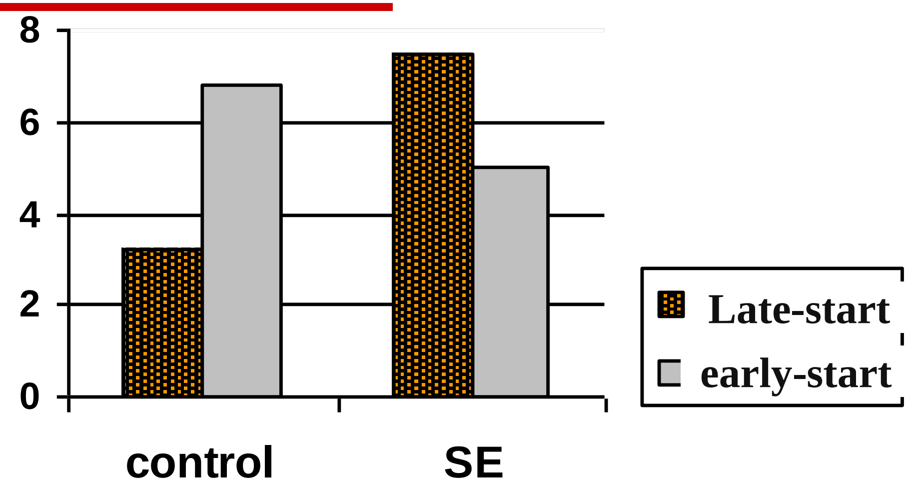
Prompt Type	Max.	Generated	Followed
Use Rule Browser/Templ.	**43**	**22.6**	**38.6%**
Use Plan Browser	**34**	**22.4**	**42%**
Read More Carefully	**43**	**7**	**34%**

- All hints positively correlated with posttest ($p < 0.0.5$)

Results: Learning

P < (0.03) P > (0.1)





- Late-start subjects in SE condition more motivated to learn from Workbench tools?
  - Significantly more (p = 0.01) attempts before abandoning template explanation
  - Larger correlation (r = 0.3 vs. r = 0.03) between learned rules and posttest
- Early-start subjects control spontaneously self-explained?
  - Mean and St.Dev. # of line accesses correlate with posttest (p < 0.08)
  - Pitt-USNA classes started semester earlier => More recall to self-explain spontaneously

Conclusions

Probabilistic student modeling for

- Coached problem solving
  - On-line knowledge tracing, plan recognition and action prediction to improve the effectiveness of the tutor's interventions
- Learning from examples
  - Assessment of the understanding of written instructional material
  - Takes into account student's attention patterns

Andes Bnet inference

- Andes' networks include anywhere between 100 and 1000 nodes
- Update needs to happen in real time
  - Starts each time a student performs a new action
  - Needs to be done when the student asks for help

- Exact when feasible
- Otherwise Approximate

Several questions about…

Why Bnets and not MDPs or POMDPs?

Actions of the agent and action of the student …..

What would be a state?.....

reason for not modelling it as a planning problem because there would be a **large number of states** due to fact that Andes stresses that the order of solving the problem is not strict

Inference

Can the model be extended to **Reinforcement learning model** based on the observations of user behaviors in the practice questions?

Is the paper using Approximate Inference at all? It may be implied somewhere but I can't find a concrete example.

Yes exact inference in some cases was taking seconds… too much for an interactive system

Selecting the problem

As a first/second year physics tutor for many years, I've seen a lot of students becoming completely lost when the problem incorporate more than 2 or more physics theories/concepts. In the paper, it says Andes will choose a problem with an appropriate complexity that involves only a few rules that the student has not yet mastered, how exactly does Andes generate such problem and how does it know what is the appropriate complexity for the student?

Student modeling…..

At what point will the Andes Student model determine that a Student has **mastered** a rule?

How does this model handle with **different difficulty levels** of questions with same rule applied to decide the **mastery** of this rule? For example, two questions might use the same rule but one of them is extremely tricky and students may fail to do this one while it cannot say they do not master the rule.

Was the approach able to predict the effect of outside knowledge affecting students answers? In the case of a student having **sufficient knowledge in calculus and linear equations the majority of Newtonian** mechanics is simplistic, but would not provide diagramming skills.

No

What happens if a student interacts with the system and the network learns about the student but then the student completely changes his behaviour in some way …..Will it take a long time for the system to re-adapt to this?

Yes… probably the same for a human ;-)

Student modeling…..

What happens if a rule of some sort is created by the teacher (e.g. these two are mutually exclusive strategies to solve a problem), realizes it's incorrect after some students …….

It seems as if the probability of knowing a rule is based on the student's reading ability. However since the AI tutor is using time as the only reference, how will it take into account if the student had opened the application and did not immediately start reading?

Self explaining

Error due to input mistake: there is a prob for that

Error due to language mistake: ESL student might …

Not covered by Andes

Problem Solving Interface

Can the students view information the system has on them, such as how likely the system thinks they are likely to self-explain, or what topics they are likely to not yet have mastered?

No but this is an interesting possibility

How do students actually use the "hint" feature? The hint is encoded so that the probability of mastery is not raised as much when a hint is given but perhaps students use the hint to confirm their solutions as opposed to solely for when they have not mastered the rule. Is there evidence that the "hint" feature is encoded in the way that is actually used by students?

?? Given that mutual exclusivity is a big issue with Newtonian physics, how does the system handle this when **presenting problems to students** and generating the probability distribution?

Bnet structure

Many of the nodes are described to have binary domains. Although the paper provides reasoning for this choice, is it common practice to do this for Bayesian networks due to the increase in complexity with having to maintain bigger probability tables if more domain values are available?

No I would say you try to model your domain as close as possible

Domain-general part

- Is there a problem with making general rule nodes observable with perhaps a simple question about a definition?

No, could be an interesting extension

- How are Context-rule nodes corresponding to a template for student's self-explanation created and how does their input get translated into Bayesian Network probabilities required for building their student model?

This is encoded in the Bayesian network by linking the SE node for a template filling action with the Context-rule node corresponding to that template's content.

- Dependencies among rules

Not captured in Andes

Task / Probabilities

- ……. It would be very useful if particular **dominant strategies** could be identified - e.g. if a problem can be solved in multiple ways, but those who solved it in one particular way were more or less likely to solve a separate problem.

Not sure this analysis was ever done. But it would be interesting and possible for similar systems

- On page 387, then definition of a **slip** is presented. Would something still be considered a slip if all preconditions were known and two rules were mastered, but one was chosen instead of the other? (i.e. is there an idea of a "best" action to take, or are all the correct actions really just as good as each other?). No
- The approach to implement Leaky-OR relationship to address the case where the student might be guessing is really interesting. What are other potential or actual usage for this in the industry? to make them seem more random and human like?

I would say yes.

- How are alpha and beta determined in 'slipping' and 'guessing' (i.e., Leaky-OR and Noisy-AND)?

Conditional/ Prior Probability Where do prior probabilities come from? Default to 0.5? Are there better starting values? Learn probabilities from more data on the student (Educational Data Mining)

Task / Probabilities

series of correct guesses? Wouldn't the model have no way to know/recover from that?

What about humans?

Reading latency

How did they end up tackling the problem of deciding what is happening during the student latency time period?

Student modeling for example studying

It is discussed that reading latency is used to evaluate the probability of self-explanation without requiring self-explanation explicitly, and an equation modeling this was provided. How is this probability value scaled compared to explicit self-explanation?

Student learning in practice

"Has this tutoring system had an impact on student learning in practice?"

Adaptation to new tasks/domains

How difficult would it be to add more physics problems to Andes system? Could this model be applied to different learning domains such as language learning, literature analysis or history?

What are the limitations of expanding this system to other problem domains?

How difficult is it to extend the system? For example, by adding new rules and problem types.

other applications exists for the Andes? Is it possible to use it in **a literature class environment**? Since there are multiple interpretation of a book/paragraph/essay/paper, how would Andes handle such high level of variety in the student's response? strictly model for student in **mathematics and theorem** related courses?

The ITS, Andes, studied in this paper is using in the subject of physics. Physics involves a lot of problem solving and formula applying, and it is a good field to apply ITS to improve self-learning. However, this tutoring model may not be good for other subject such as **Philosophy and Business** where the answer can be various due to different point of view. These could cause a even higher uncertainty and hard to make a educated guess.

What about **chemistry**, assess comprehension in the **arts**?

Future Directions

NLP: There has been huge advancement in natural language processing since the year this paper was accepted in 2002. How would expressing self-explanations work if we were to replace the interface discussed in this paper with current available natural language processing technology? Still not "easy"…. Very specific, simple proposal….

Lehman, B., Mills, C., D'Mello, S., & Graesser, A. (2012). Automatic Evaluation of Learner Self-Explanations and Erroneous Responses for Dialogue-Based ITSs. In. S. A. Cerri, & B. Clancey (Eds.), *Proceedings of 11th International Conference on Intelligent Tutoring Systems* (*ITS 2012*) (pp. 544-553). Berlin: Springer-Verlag.

NLP: What parts of the system can be changed with better natural language processing? …..

Eye tracking: With today's eye tracking technologies, can we train a separate model which is able to classify whether student is confused or satisfied on a particular problem from student's eye movement? Then, we can integrate this feature as a prior probability to student's actions in the bayesian network, as it may provide accurate information for the student model to infer student's action and generate help. Yes

Does taking into account other student's tendencies and patterns help create a better algorithm for future students?

It could

TODO for Fri

- **Reading Textbook Chp. 6.5**

- Keep working on assignment-2 (due on Fri, Oct 16)