From Text to Paragraph Representations: Working Note 15

Peter Clark
Knowledge Systems
Maths and Computing Technology
Boeing, PO Box 3707, Seattle, WA
peter.e.clark@boeing.com

April 1999

Abstract

The following informal note discusses a short (4 sentence) paragraph of text (about parachutes), its representation, and the relationship between that representation and the semantic graphs derived from automatic natural language processing (NLP) of the sentences. We discuss what common-sense knowledge is missing (assumed) in the text, what a target representation of the paragraph might look like, and the size and nature of the gap between the target and the raw semantic graphs which NLP provides. This is done with a view to migrating from hand-crafting of representations towards automatic/assisted knowledge acquisition from natural/controlled language documents.

1 Introduction

The following informal note discusses a short (4 sentence) paragraph of text (about parachutes), its representation, and the relationship between that representation and the semantic graphs derived from automatic natural language processing (NLP) of the sentences. We discuss what common-sense knowledge is missing (assumed) in the text, what a target representation of the paragraph might look like, and the size and nature of the gap between the target and the raw semantic graphs which NLP provides. This is done with a view to migrating from hand-crafting of representations towards automatic/assisted knowledge acquisition from natural/controlled language documents.

We start with the premise that to sensibly apply the knowledge in a paragraph, a system should at least have a coherent model ("p-model") of what the paragraph is saying. In a perfect world, this model could be automatically built by piecing together, jig-saw style, the (NLP-generated) semantic graphs for the paragraph's sentences + for relevant background knowledge (eg. from dictionary definitions). Some issues which arise then are:

- What might a coherent p-model look like?
- How easily do NLP-generated semantic graphs fit into this?
- What additional 'common-sense' knowledge is in the p-model but not explicitly stated?

The sample paragraph is taken from a children's book about aviation. Two versions of the semantic graphs were generated, from two different NLP systems (EGSC and NLPWIN).

2 The Paragraph

The paragraph of interest is about parachutes, and is shown below:

> **"Parachutes slow down a person falling through the air. This means that he or she can land safely when bailing (jumping) out of a plane. When open, a parachute creates lots of drag as air pushes against its underside. This slows its fall."**

Although apparently simple, there is a large amount of unspoken, background knowledge required to understand this paragraph, and to determine the unstated contexts in which the assertions are true. (This is the case even assuming the sentences are correctly parsed, word senses correctly chosen, and anaphora resolved correctly). Taken literally, and without any constraint on context, a machine may draw many incorrect conclusions from the paragraph, for example:

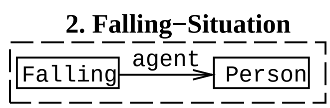
1. **"Parachutes slow down a person falling through the air."**
    - There is no need to be wearing the parachute.
    - There is no need to open the parachute.
    - There is a single person in the world who can safely jump (with or without a parachute). The fact parachutes exist will save him/her.
    - Slowing eventually results in stopping. Thus the person will eventually stop and hover in mid-air.
    - A parachute will only slow down someone if it feels so inclined (similar to "People vote in elections.")
2. **"This means that he or she can land safely when bailing (jumping) out of a plane."**
    - A pilot can jump out of a plane parked on the ground and land safely.
    - Without a parachute, the person will do something else safely instead, eg. sing safely (this conclusion follows from mistakenly identifying the import of "landing safely" to be "*landing* safely", rather than (correctly) "landing *safely*").
    - Other unstated knowledge: To understand the relationship of this to the first sentence, several additional pieces of knowledge are required:
        - Jumping (here, out of a plane) results in falling.
        - Falling is followed by a landing.
        - If an object is being slowed, its speed will be reduced.
        - Landing involves colliding with the ground, and fast collisions are dangerous.
3. **"When open, a parachute creates lots of drag as air pushes against its underside."**
    - There will be lots of drag on an open parachute lying on the ground.
    - The parachute doesn't have to be attached to anything (ie. doesn't need to have a particular shape). It could be just open and thrown out of the plane.
    - The underside is owned by the parachute (incorrect interpretation of the possessive relationship in "its underside").
    - Other unstated knowledge:

- The drag is on the parachute (unstated).

4. **"This slows its fall."**
   - Unstated knowledge:
     - It was unstated that the parachute was falling, only the person. Need the implication that if X is connected to Y, and X is falling, then Y is falling also.
     - It was unstated that if the parachute slows, then so will the person. Need the implication that if X is connected to Y, and X slows, then Y slows also.

3 The Paragraph Model (p-model), and its Components

Figure 1 shows (in an informal, graphical notation) a possible representation for the scenario the paragraph describes. The representation is hand-built and is fairly minimal, intended to be the smallest model which still covers the intended content of the paragraph. The graph depicts only ground assertions, not the rules which would derive these assertions, and thus is only a semi-formal and incomplete way of depicting content. The graphical dashed boxes denote situations, and their contents denote situation-specific (fluent) assertions. For example,



denotes the ground assertion

`agent(Falling01,Person01,Falling-Sitatuation01)`

ie. `Person01` is the agent of `Falling01` in `Falling-Situation01`. (`Falling01`, `Person01`, and `Falling-Sitatuation01` are Skolem instances denoting the individuals being talked about). Nodes with the same name are coreferential.

Figures 2, 3, 4, and 5 highlight fragments of the p-model corresponding to the four sentences in the paragraph. These diagrams clearly show that the sentences explicitly state only a small part of the scenario which they were intended to convey, and that their correct interpretation depends heavily on identifying the correct context in which to place them.

Figures 6, 7, 8, and 9 show the same framents, and compare them with the semantic graphs obtained by direct NLP of the sentences (and before intersentence pronoun resolution).

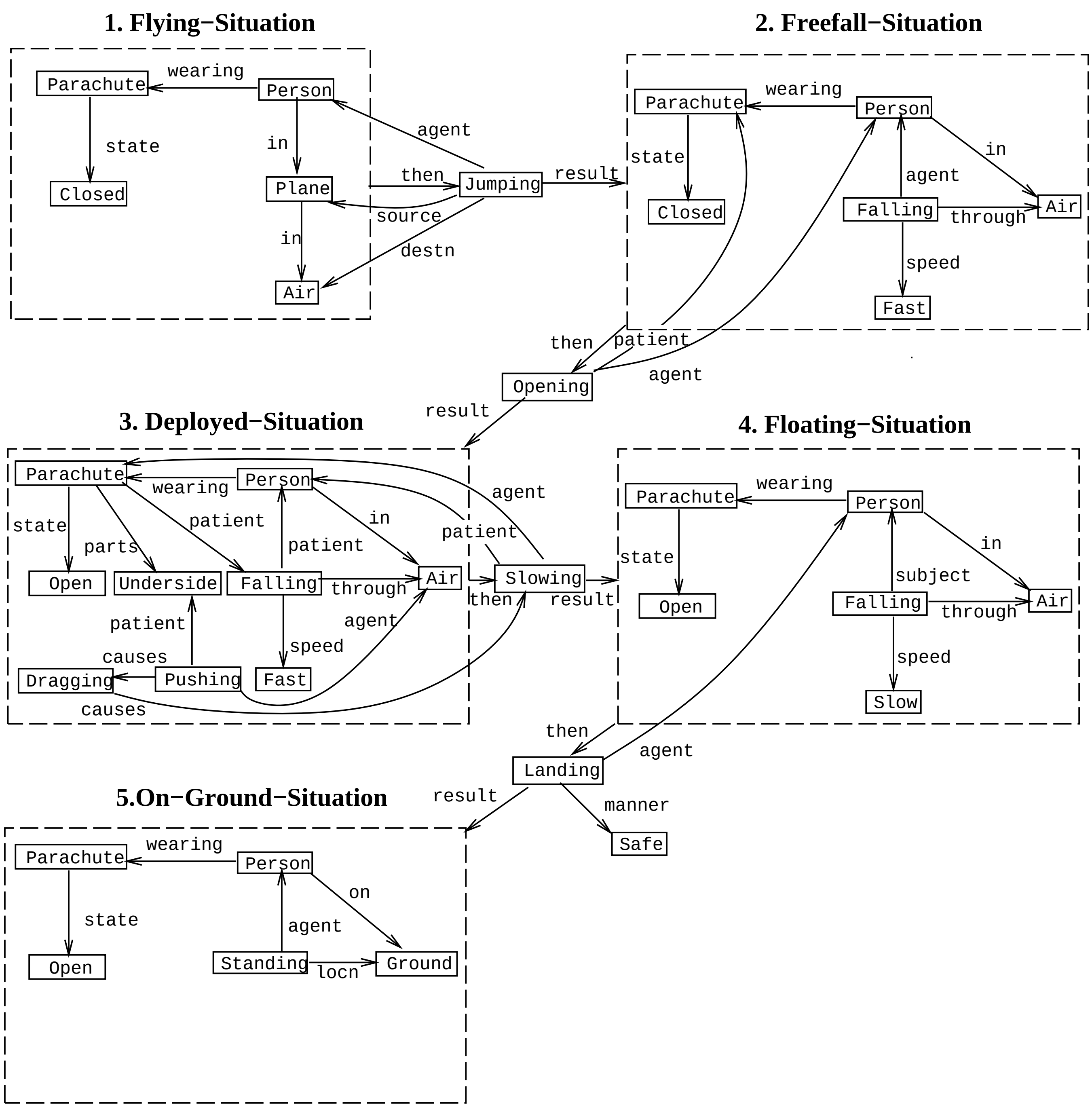


Figure 1: A possible p-model representing the parachute scenario described in the paragraph. (Dashed boxes denote situations, ie. their contained relations are fluents).

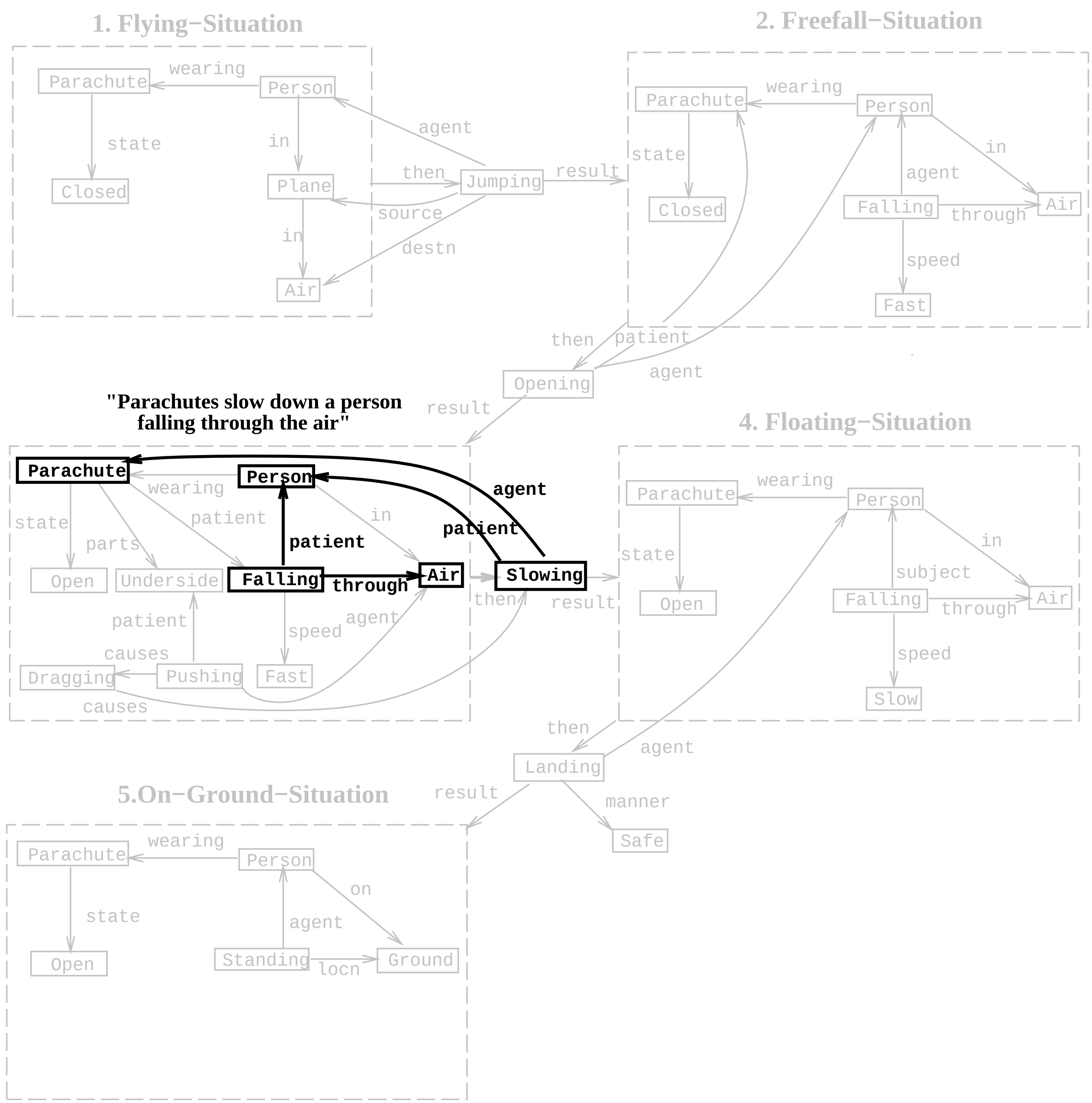


Figure 2: Subgraph of the p-model corresponding to sentence 1.

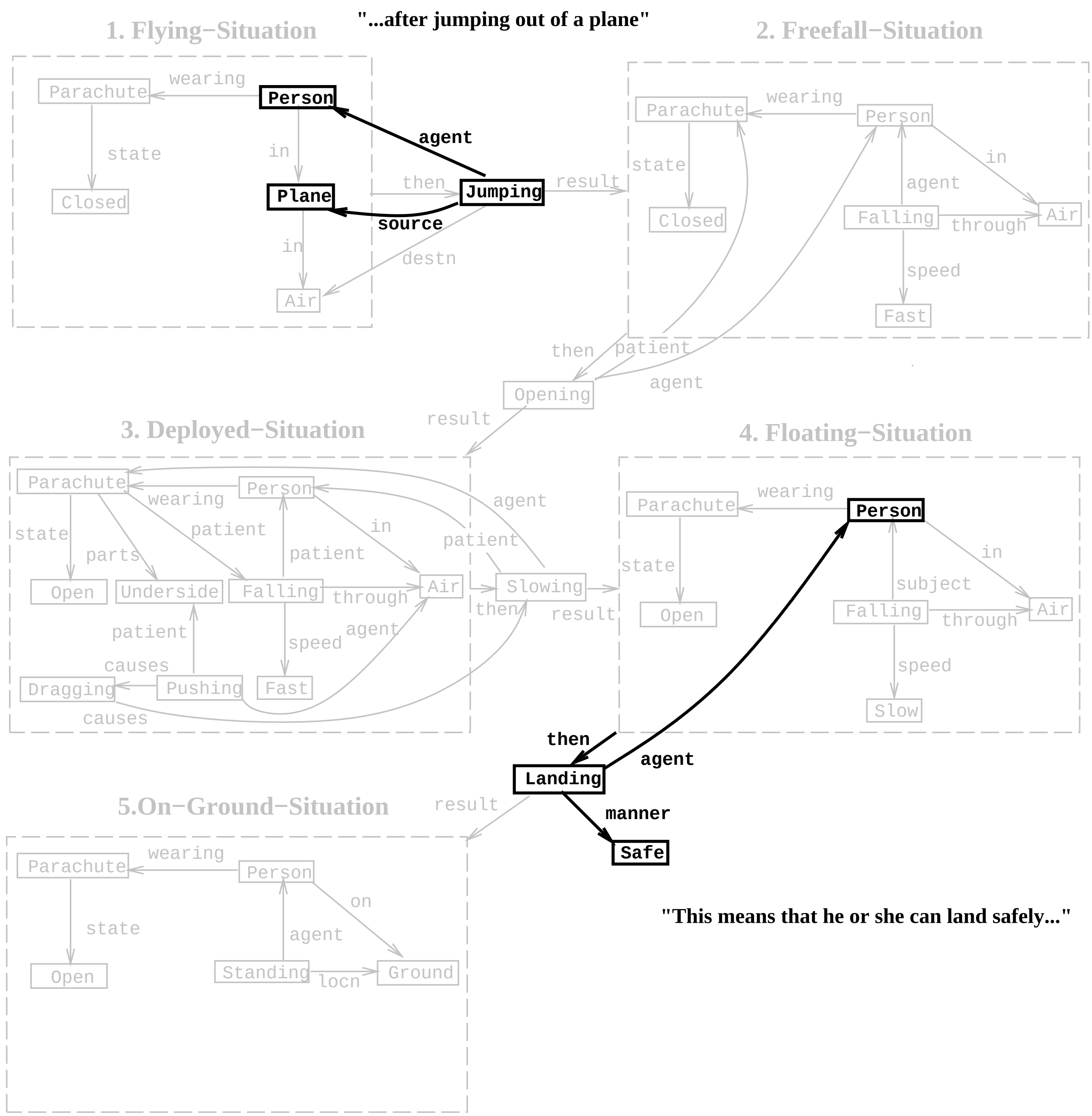


Figure 3: Subgraph of the p-model corresponding to sentence 2.

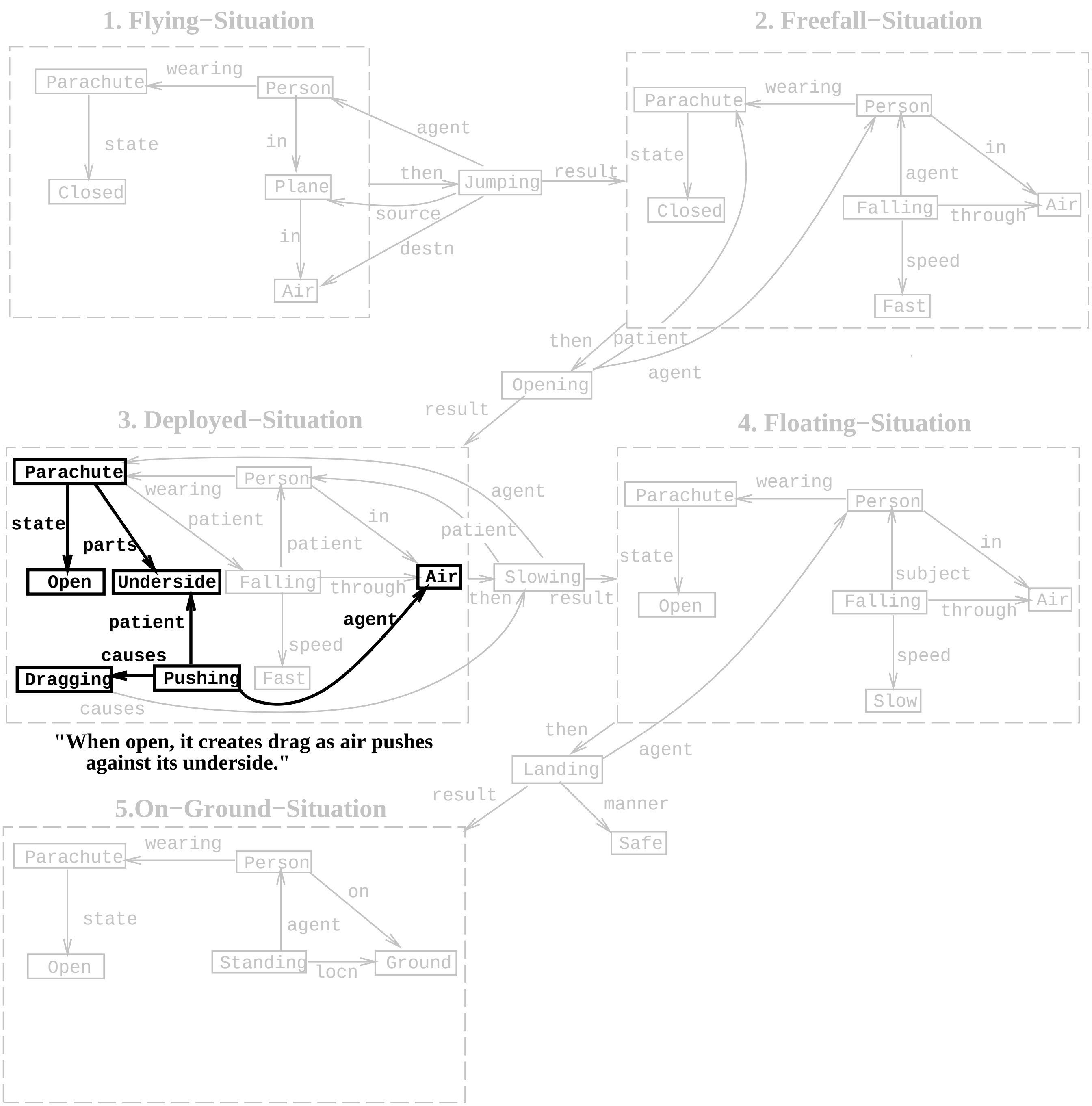


Figure 4: Subgraph of the p-model corresponding to sentence 3.

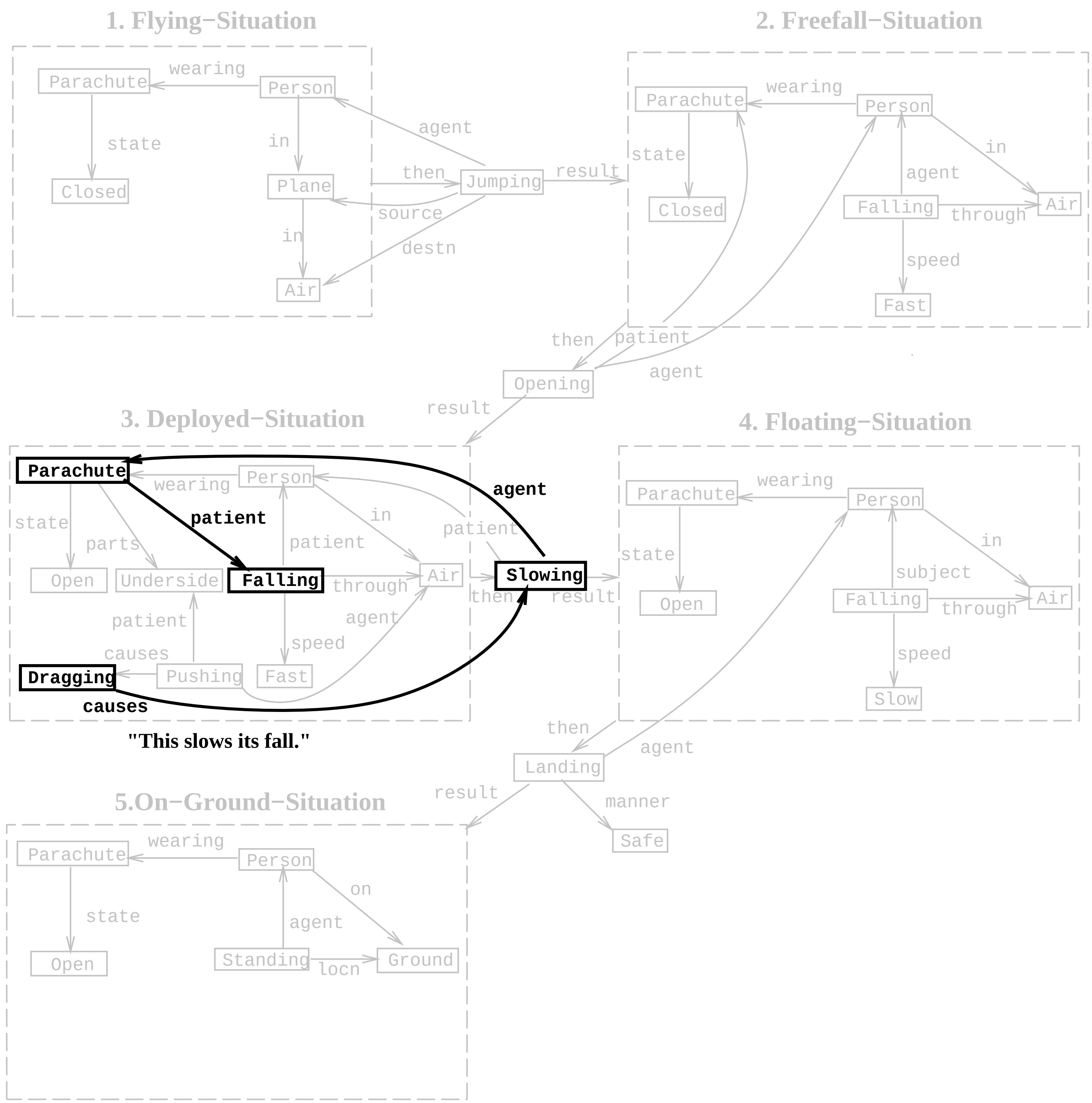


Figure 5: Subgraph of the p-model corresponding to sentence 4.

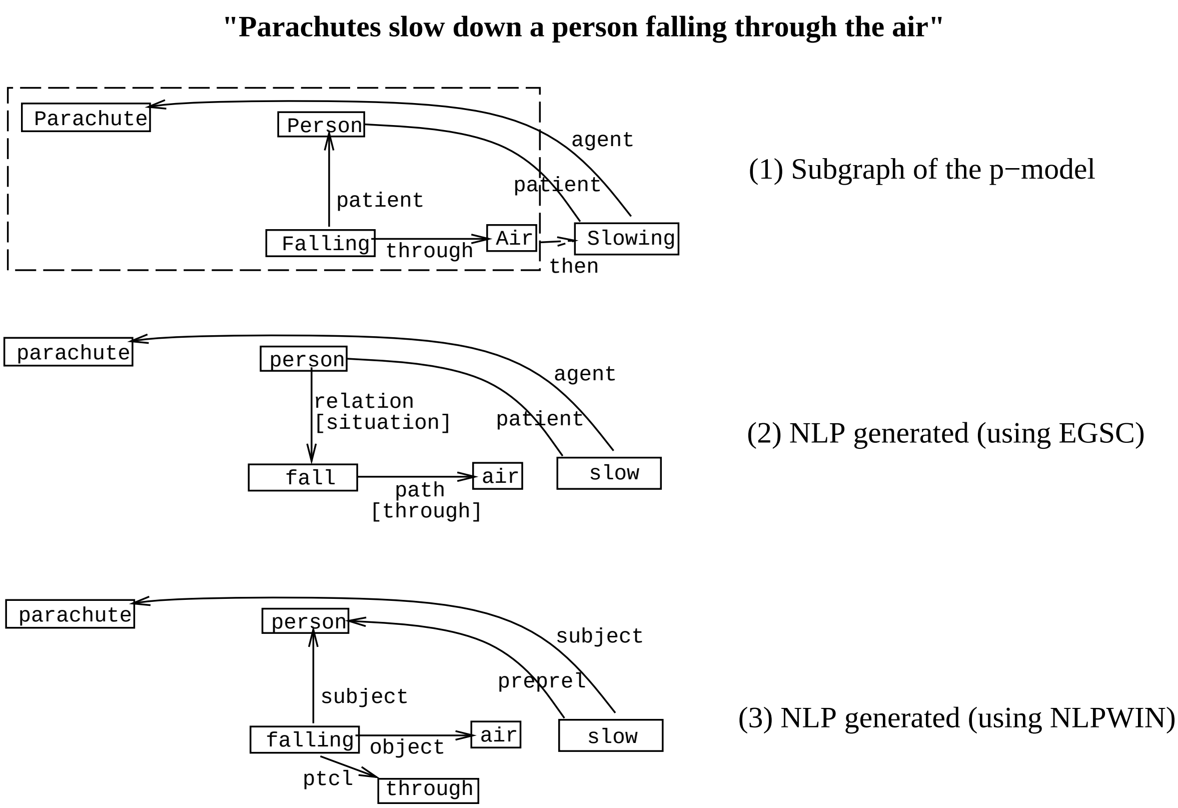


Figure 6: NLP-generated semantic graphs corresponding to sentence 1.

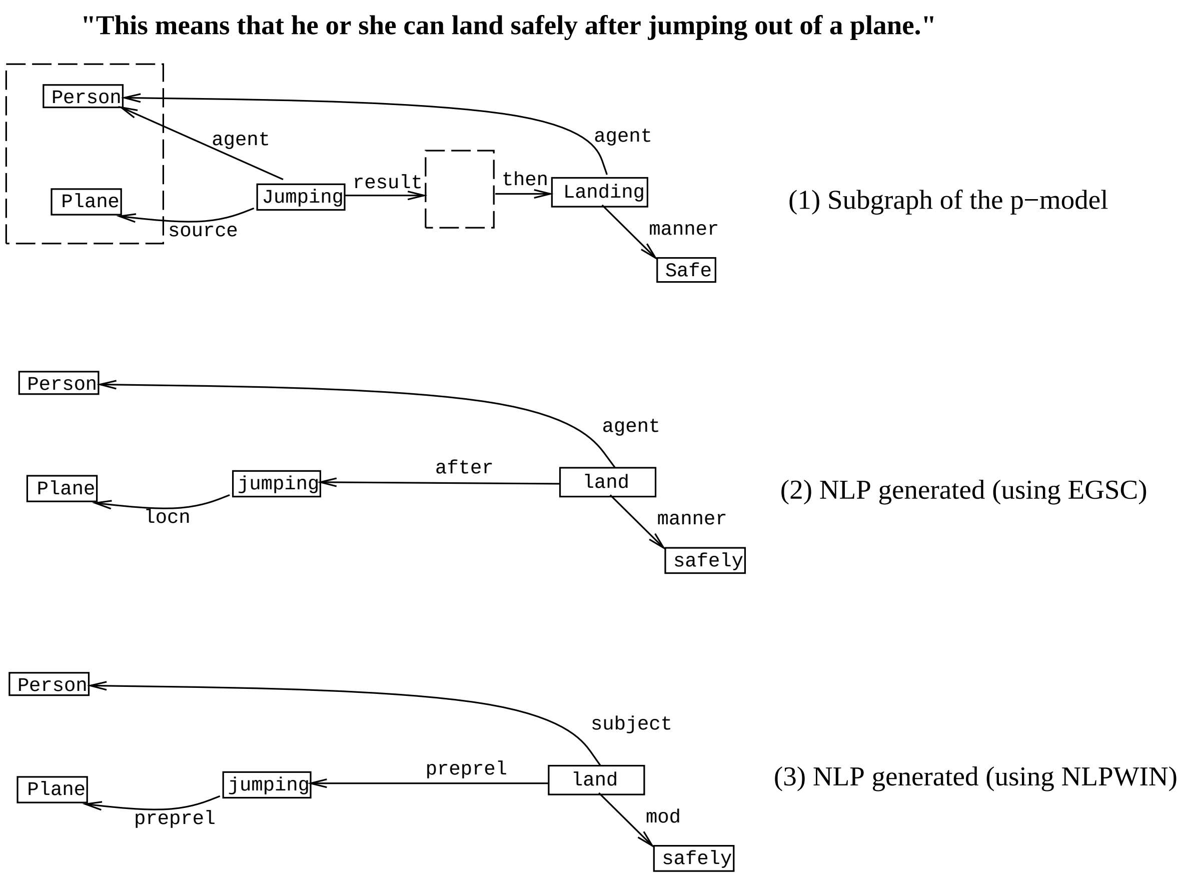


Figure 7: NLP-generated semantic graphs corresponding to sentence 2.

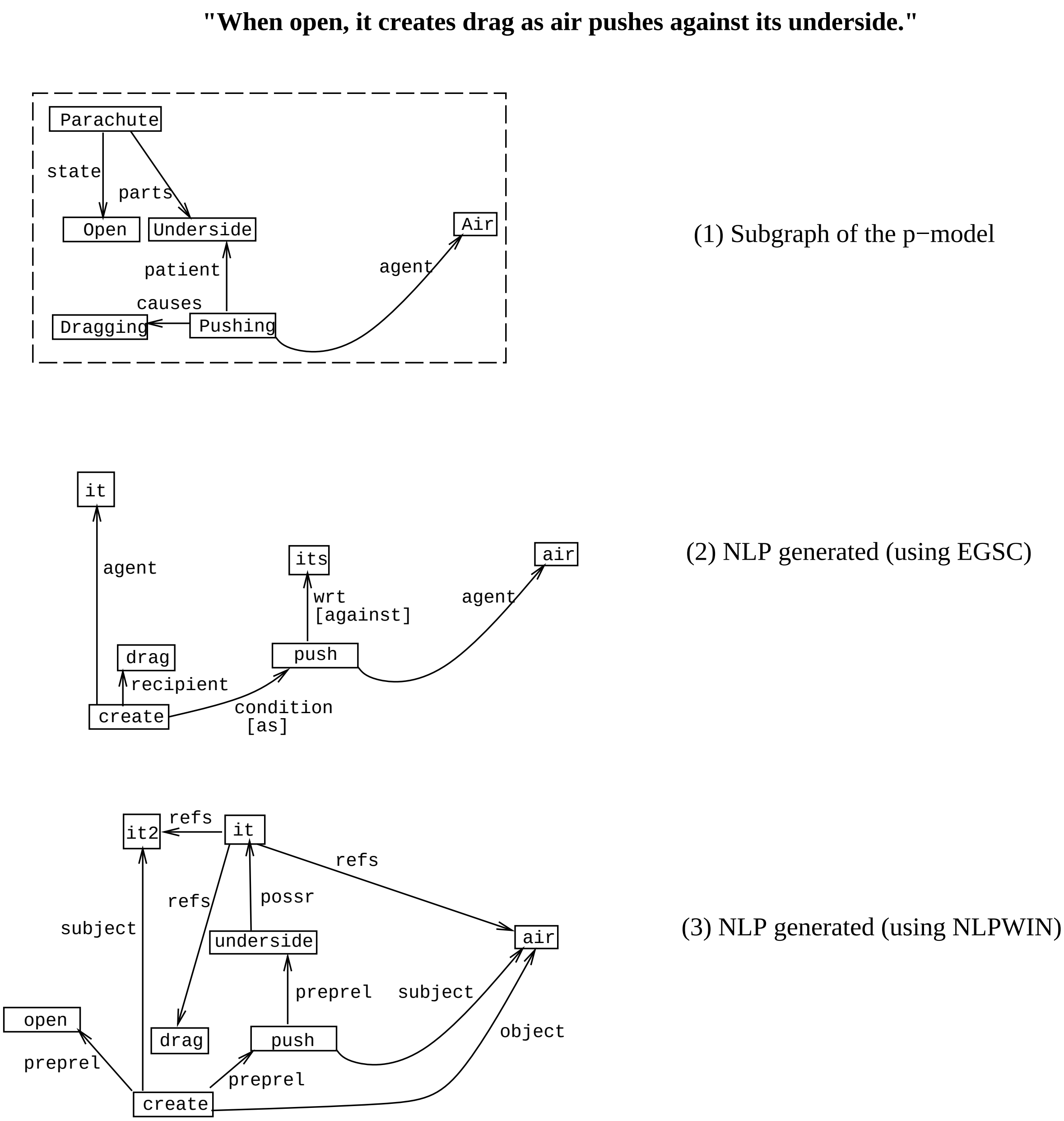


Figure 8: NLP-generated semantic graphs corresponding to sentence 3.

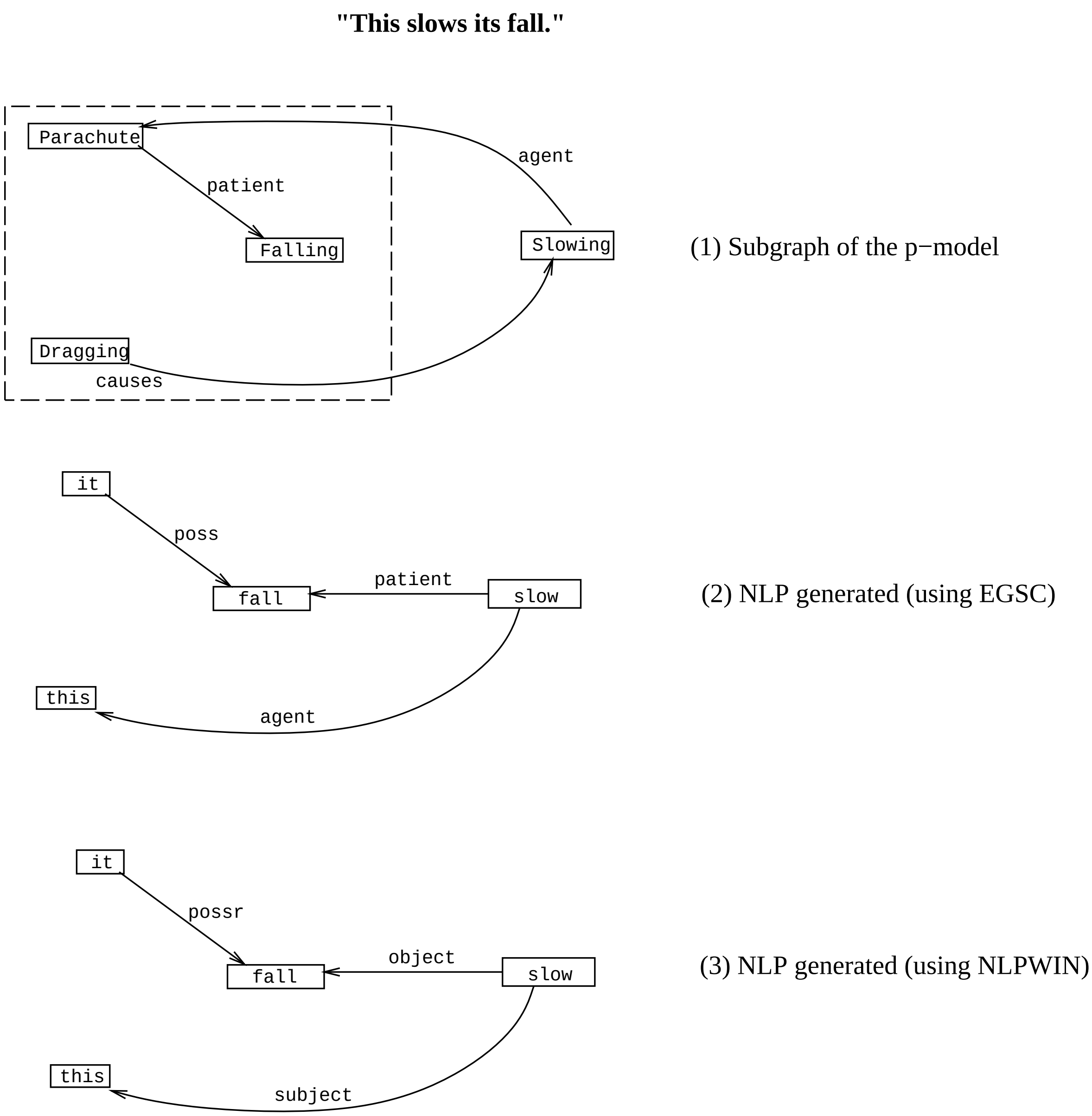


Figure 9: NLP-generated semantic graphs corresponding to sentence 4.

4 Discussion

We are interested in what a coherent representation of the parachute paragraph might look like (from which inference could then be performed), and its relationship to the semantic graphs derived from NLP of that paragraph – ultimately with a view to migrating from hand-crafted representations towards automatic/assisted knowledge acquisition from natural/controlled language texts. Several interesting points of reflection can be drawn.

As is well known, knowledge in surface text is highly incomplete. This fact is reflected in the preceding diagrams where the NLP-generated semantic graphs cover only a part of the complete representation. To explore this further, we distinguish and then discuss three causes of incompleteness:

Linguistic ambiguity: Inherent ambiguity in translating from the textual form to the semantic (conceptual/logical) form, eg. in word sense determination, anaphora resolution, prepositional phrase attachment, other parsing ambiguities, etc.

Missing background knowledge: Knowledge which is simply missing, both about the domain (eg. people can wear parachutes) and general common-sense knowledge (eg. jumping results in falling). The paragraph author assumes the reader will bring this knowledge to bear when reading the paragraph, to create a coherent interpretation.

Missing contextual knowledge: The paragraph sentences are only true in an implicit context – that is, they cannot be taken literally (as Section 2 makes clear). Rather, a coherent understanding also requires identifying the implicit contexts intended, and placing the assertions within them, thus (correctly) limiting their applicability to just those contexts.

Linguistic Ambiguity

Although creating semantic representations using NLP is in general an extremely hard task, both EGSC and NLPWIN fared well at least with these four sentences (though both struggled a bit with the third sentence, see Figure 8). Although neither perform intersentence anaphora resolution, it is also plausible that this task is surmountable with appropriate techniques. The Figures in this document suggest that the semantic graphs can fit together to produce part of a coherent paragraph representation.

Missing Background Knowledge

Section 2 listed some of the missing background knowledge required for a minimal understanding of the paragraph; in addition, substantially more would be needed to answer questions about the text. For example, the p-model in Figure 1 incorporates knowledge such as:

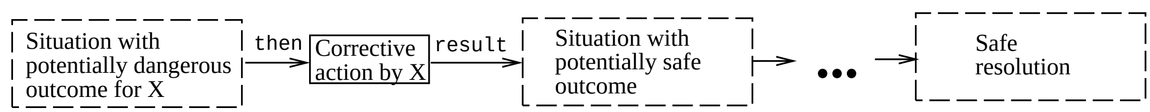
- jumping means to spring into the air,
- falling is to descend by gravity, and descend involves movement downwards,
- landing results in coming to rest at a location,
- slowing means to reduce speed,
- air is a gas,
- ground is solid,
- parachutes can be worn by people,

- two connected things will move at the same speed,
- etc.

Although none of these facts are stated in the text, it is possible that some or even much of this knowledge could be extracted, at least in some form, from processed dictionary definitions. However to do this, the same ambiguity and contextual issues above arise recursively, and definitions written from mismatching viewpoints pose an additioanl challenge. For example, "falling" can be viewed as either a state (as in this example) or an action, depending on the temporal resolution which the task at hand requires. Some mechanism for identifying and merging (or keeping distinct) definitions written from different viewpoints would be required.

Missing Contextual Knowledge

Finally, semantic graphs from both the explicit text plus background knowledge need to be placed in the correct context – sentences can rarely be taken literally, and will cause havoc for inferencing if their context of applicability is not constrained. Guidance for placing graphs in context might be provided by using scripts/schemata to specify valid p-model "skeletons", and then instantiating them with the NLP-generated semantic graphs. For example, the parachute p-model (Figure 1) can be viewed as instantiating a more general schema about recovering from danger, in which there is first some potentially dangerous situation, then some corrective action is taken resulting in a safe situation:



In the parachute scenario, the dangerous situation corresponds to some (unstated) problem during flying, the corrective action corresponds to jumping and opening the parachute, and the safe situation corresponds to the safe landing.

5 Summary

Given a paragraph, can the semantic graphs for each sentence be plugged together (jig-saw style) to create a coherent representation of it, from which reasoning/question-answering can then follow? In this exercise, the answer seems to be a qualified "sort of". On the one hand, the semantic graphs were (at least here) reasonably accurate, and do fit into a larger structure (assuming intersentence anaphora resolution can be made reasonably accurate). On the other hand, the surface text leaves a lot unsaid; some of the "missing jigsaw pieces" might be pulled from other sources, eg. processed dictionary definitions or hand-crafted; more challenging is determining the implicit and changing context/assumptions behind the text, which needs to be used to restrict the validity of the semantic graphs, otherwise an inference engine will produce garbage.

This exercise was effectively a brief excursion into looking at the full natural-language understanding task. Fortunately, there are also simpler variants of this with which some of these issues can be attacked, for example: working with a controlled (restricted) language; making the assumptions, the underlying context, and shifts in context explicit in the text; and perhaps hand-coding representations for a small number of "foundational" common-sense concepts on which target paragraphs depend. All these would be interesting to explore.